World-Class Instructional Design and Assessment

Annual Technical Report for ACCESS for ELLs Paper English Language Proficiency Test Series 501, 2019–2020 Administration

Annual Technical Report No. 16B

Prepared by:

Center for Applied Linguistics

Language Assessment Division
Psychometrics and Quantitative Research Team

May 2021

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 Research, Measurement, and Evaluation, 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, Teacher (retired), Hanover High School, Hanover, New Hampshire
- Carol Myford, Ph.D., Professor Emerita, Educational Psychology, University of Illinois at Chicago

Executive Summary

This is the 16th annual technical report on the ACCESS for ELLs English Language Proficiency Test and the fifth report on the Paper ACCESS for ELLs assessment since the Online assessment was launched.

This technical report is produced as a service to members and potential members of the WIDA Consortium and to support states' submissions for U.S. Department of Education English language proficiency assessment peer review. 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). WIDA also produces an annual *Year in Review Report*, intended for a general audience, for readers who are interested in a nontechnical overview of the 2019–2020 ACCESS assessment.

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 Paper Series 501 was administered in school year 2019–2020 in 35 states, the Bureau of Indian Education, the Department of Defense Education Activity, the District of Columbia, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands, for a total of 40 state entities (henceforth "states").

The Series 501 Paper data set included the results of 503,365 students. The largest grade was Kindergarten, with 226,212 students, while the smallest was Grade 12, with 8,391 students. Of the participating WIDA states, the largest was Florida, with 264,969 students, while the smallest was the Commonwealth of the Northern Mariana Islands, with 50 students.

During the 2019–2020 testing year, many states suspended in-person schooling due to the COVID-19 public health emergency. Based on a comparison with prior years' numbers of participating students, WIDA believes that most students who likely would participate in ACCESS for ELLs had completed their test sessions at the time that schools closed. Further

detail on the impact of COVID-19 is contained in the ACCESS 2019–2020 *Year in Review Report*.

ACCESS for ELLs Series 501 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 16B) provides technical information pertaining to ACCESS for ELLs Series 501 Paper. A second report (WIDA ACCESS Technical Report 16A) provides technical information for the ACCESS for ELLs Series 501 Online assessment.

Part 1:
Purpose, Design, Implementation

Contents

1. Purpose and Design of ACCESS

1.1. Purpose Statement

The 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 used 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 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 Paper, described in this report, and ACCESS Online, 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.

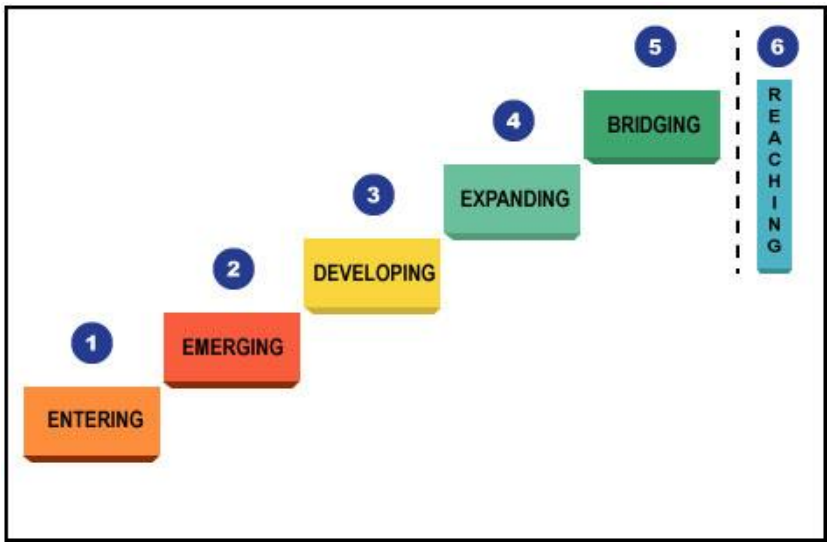


Figure 1. The language proficiency levels of the WIDA ELD Standards.

These language proficiency levels are embedded in the WIDA ELD Standards in two ways.

First, they appear in the **performance definitions**. The performance definitions describe the stages of language acquisition, providing details about the language that students can comprehend and produce at each proficiency level. The performance definitions are based on three criteria: (a) vocabulary usage at the word/phrase level; (b) language forms and conventions at the sentence level; and (c) linguistic complexity at the discourse level. 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' understanding or demonstration of oral interaction and writing of increasing quantity and variety.

Second, language proficiency levels are represented through connections to the accompanying **Model Performance Indicators** (MPIs). The MPIs provide a model of the expectations for ELL students in each of the five Standards, by grade-level cluster, across the four language domains, for each of the language proficiency levels up to level 5. The grouping of MPIs at PLs 1 through 5 for a given WIDA Standard, grade-level cluster, domain, and topic is called a strand. These MPIs together describe 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 final level, PL 6: *Reaching*, represents the end of the continuum rather than another level of language proficiency.

Each MPI has a tripartite structure, consisting of a language function, a content stem, and support. The MPIs used on ACCESS can be taken directly from the WIDA English Language Proficiency Standards (WIDA Consortium, 2007) or the amplified 2012 ELD Standards (WIDA Consortium, 2012). In addition, given that the MPIs in the WIDA Standards are truly "models"

and do not cover all possible topics within each Standard for each grade-level cluster and language domain, MPIs can be "transformed" to accommodate the needs of classroom instruction, as described in the amplified 2012 ELD Standards (WIDA Consortium, 2012, p. 11). MPIs are also transformed for the purposes of the assessment. When MPIs are transformed, one or more of the three aspects of the base MPI are changed. For example, if an MPI from the amplified 2012 ELD Standards (WIDA Consortium, 2012) has "categorize" as its language function, that could be transformed to "compare/contrast" or "infer." Likewise, if the content stem for a grades 9-10 Language of Social Studies strand of MPIs is "supply and demand," it could be transformed to "freedom and democracy." Each item specification document for a given WIDA Standard, grade-level cluster, and language domain contains an MPI for each item or task, such that the MPI is the core construct that the given item/task intends to measure. 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 grade-level cluster structure for ACCESS for ELLs Paper is as follows: K, 1, 2, 3, 4–5, 6–8, 9–12.

In the lower grades (Grades 1–5), test forms may be shared across grade-level clusters. As described in Section 2.2.1 below, the Listening and Reading tests were developed prior to the launch of the 2016 operational administration, which represented the shift to the new cluster structure of Online ACCESS. Earlier ACCESS tests had a cluster structure that differs from that of the current ACCESS items in newer development, in the lower grades. The Speaking and Writing tests were developed using the ACCESS Online cluster structure. ACCESS Paper clusters, therefore, bridge the cluster structure of the older ACCESS assessments and ACCESS Online. For example, the Cluster 2 tests in the domains of Reading and Listening are the same test forms as the Cluster 1 tests. The Cluster 2 tests in the domains of Speaking and Writing are the same test forms as the Cluster 3 tests in these domains. Table 1 details the grade-level cluster structure of ACCESS Paper and the shared forms across clusters.

Table 1
ACCESS Paper Grade-Level Clusters and Shared Forms Across Clusters

ACCESS Paper Grade-level Clusters	Shared Test Forms (Listening and Reading)	Shared Test Forms (Speaking and Writing)	Grade
K	K	K	K
1	Cluster 1 and Cluster 2	Cluster 1	1
2		Cluster 2 and Cluster 3	2
3	Cluster 3 and Cluster 4–5		3
4–5		Cluster 4–5	4
			5
6–8	Cluster 6–8	Cluster 6–8	6
			7
			8
9–12	Cluster 9–12	Cluster 9–12	9
			10
			11
			12

Note that in our analyses of student participation in the assessment (Part 2, Chapter 1), analysis is conducted by cluster (K, 1, 2, 3, 4–5, 6–8, 9–12). In our analyses of test forms (Part 2, Chapter 2), analysis is conducted at the form level (i.e., in Listening and Reading, a single analysis is conducted for the Cluster 1 and Cluster 2 form). Test form level analyses are presented for each cluster that the form appears in; if a table of results pertains to more than one cluster, it is repeated in each cluster.

1.6. Tiers

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. Tests must be at the appropriate difficulty level for each individual test-taker 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, within each grade-level cluster, students display a range of abilities. 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.

Paper ACCESS test forms are constructed at either Tier A (for students at beginning levels of English proficiency) or Tier B/C (for students at higher proficiency levels). Each Grade 1–12 test-taker takes either the Tier A form or the Tier B/C form. The Kindergarten assessment is not tiered.

In Listening and Reading, Tier A has items and tasks designed to allow students at the lowest language proficiency levels (PLs 1 and 2) to meet the WIDA ELD Standards at their language proficiency levels, and it includes some items targeted to PL 3. Tier B/C tests include items constructed to target PLs 2 (Emerging) through 5 (Bridging).

In the domain of Writing, Tier A forms include tasks written to elicit language up to PL 3, and Tier B/C forms include tasks written to elicit language up to PL 4 or PL 5. In the domain of Speaking, students at early levels of proficiency take the Tier A form, with tasks designed to elicit language at PL 1 and PL 3, and more proficient students take the Tier B/C form, with tasks designed to elicit language at PL 3 and PL 5.

2. Test Development

2.1. Test Design

This section provides information on the test design for the two forms of Paper ACCESS (Tier A and Tier B/C) and the design of each form. Note that this section applies to ACCESS Paper Grades 1–12. For detail on Kindergarten, see Section 2.4 below and the technical report on the development of the Kindergarten static form (MacGregor, Kenyon, Gibson, & Evans, 2009).

2.1.1. Listening

For the ACCESS Listening test, Table 2 shows, for each test form, the number of items, the targeted range of WIDA proficiency levels, the item types, the response format, and the scoring procedure.

Table 2
Number and Types of Items on the Listening Subtest

Grade-Level Cluster	Tier	Number of Items	Targeted PL range	Item Types	Response Formats	Scoring Procedures
1	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
1	B/C	21	PL2 - PL5			
2	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
2	B/C	21	PL2 - PL5			
3	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
3	B/C	21	PL2 - PL5			
4-5	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
4-5	B/C	21	PL2 - PL5			
6-8	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
6-8	B/C	21	PL2 - PL5			
9-12	A	18	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
9-12	B/C	21	PL2 - PL5			

Figure 2 presents the Listening test design, showing the distribution of folders by Standard for each tier. In this figure, each small gray box represents an item.

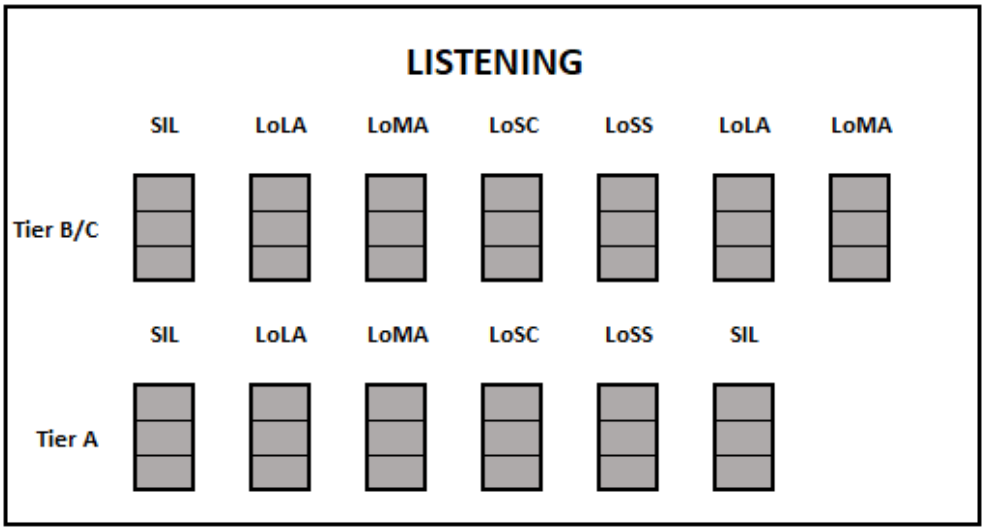


Figure 2. Distribution of items by Standard for each tier of the Listening test.

Note that the test design is slightly different between Tier A and Tier B/C. Tier B/C students, who potentially may be reclassified by the assessment, take a slightly longer test and take two folders each assessing the Language of Language Arts and the Language of Mathematics Standards. Tier A students receive a second folder assessing the Social and Instructional Language Standard, under the assumption that less proficient students will find this Standard more accessible.

Although timing guidance is provided to test administrators in the Test Administrator Manual, the Listening subtest is untimed.

2.1.2. Reading

For the ACCESS Reading test, Table 3 shows, for each test form, the number of items, the targeted range of WIDA proficiency levels, the item types, the response format, and the scoring procedure.

Table 3
Number and Types of Items on the Reading Subtest

Grade-Level Cluster	Tier	Number of Items	Targeted PL range	Item Types	Response Formats	Scoring Procedures
1	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
1	B/C	27	PL2 - PL5			
2	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
2	B/C	27	PL2 - PL5			
3	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
3	B/C	27	PL2 - PL5			
4-5	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
4-5	B/C	27	PL2 - PL5			
6-8	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
6-8	B/C	27	PL2 - PL5			
9-12	A	24	PL1 - PL4	Multiple Choice	Dichotomous Selected Response	Machine Scored
9-12	B/C	27	PL2 - PL5			

Figure 3 presents the Reading test design, showing the distribution of folders by Standard for each tier. In this figure, each small gray box represents an item.

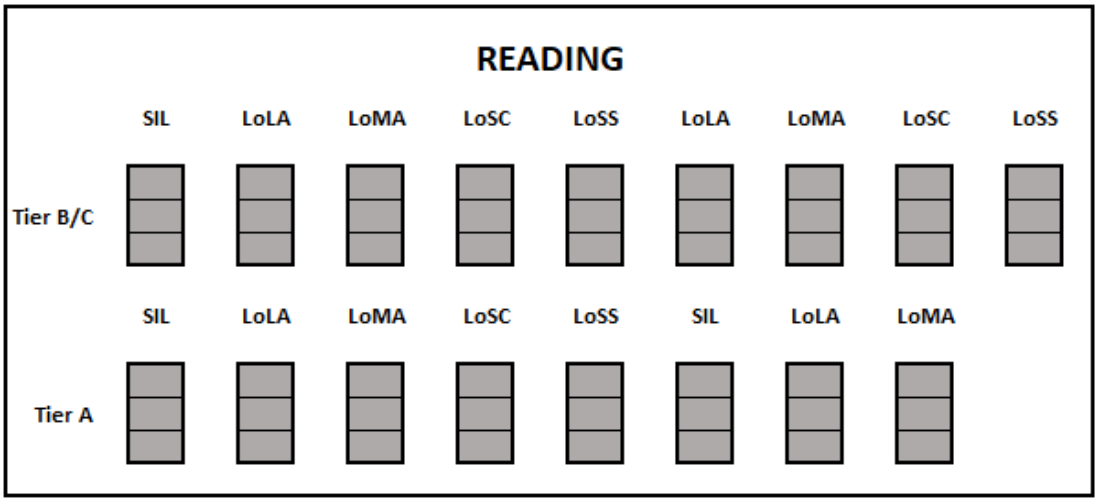


Figure 3. Distribution of items by Standard for each tier of the Reading test.

As with Listening, the Reading test is shorter and focuses on Standards deemed more accessible for lower proficiency students.

Although timing guidance is provided to test administrators in the Test Administrator Manual, the Reading subtest is untimed.

2.1.3. Writing

For the ACCESS Writing test, Table 4 shows, for each test form, the number of tasks, the targeted range of WIDA proficiency levels, the task types, the response format, and the scoring procedure.

Table 4
Number and Types of Items on the Writing Subtest

Grade-Level Cluster	Tier	Number of Tasks	Targeted PL range	Task Types	Response Formats	Scoring Procedures
1	A	4	PL1 - PL3	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
1	B/C	3	PL2 - PL5			
2	A	3	PL1 - PL3	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
2	B/C	3	PL2 - PL5			
3	A	3	PL1 - PL3	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
3	B/C	3	PL2 - PL5			
4-5	A	3	PL1 - PL3	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
4-5	B/C	3	PL2 - PL5			
6-8	A	3	PL1 - PL3	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
6-8	B/C	3	PL2 - PL5			
9-12	A	3	PL1 - PL4	Writing Constructed Response	Polytomous Constructed Response; handwritten in test booklet	Human Scored: Centrally scored by DRC
9-12	B/C	3	PL2 - PL5			

The Writing test is tiered. As Writing tasks are polytomous and elicit a range of student performances, each task is targeted to elicit language across a range of proficiency levels, rather than targeted to a single proficiency level. Tier A consists of tasks written to elicit language up to PL 3, while Tier B/C tasks are designed to elicit language up to PL 5. This is indicated by the large number in the colored rectangle in the figure. However, for both tiers of the test, all tasks

are scored using the entire breadth of the scoring scale. 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 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 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. Grade 1 Tier A has four tasks, designed specifically to allow beginning writers at this grade to demonstrate their ability in the domain of Writing. Figures 4 and 5 present the Writing test design, showing the distribution of tasks for each tier. In these figures, each colored box represents a task. The number in the box represents the targeted proficiency level of the task.

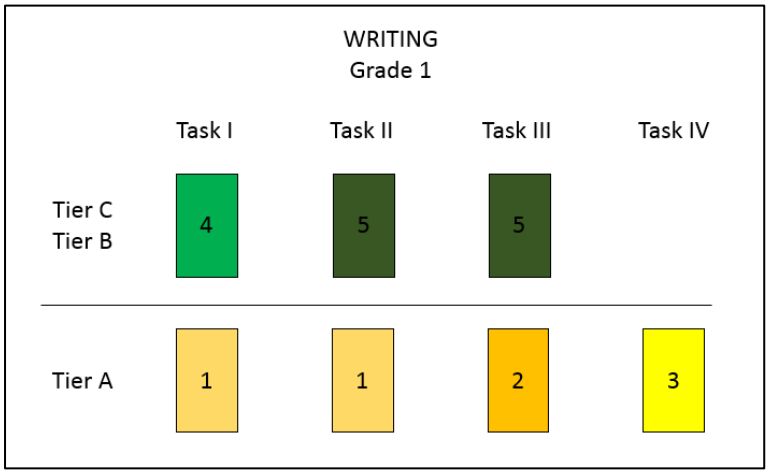


Figure 4. Distribution of tasks by targeted proficiency level for each tier of the Grade 1 Writing test.

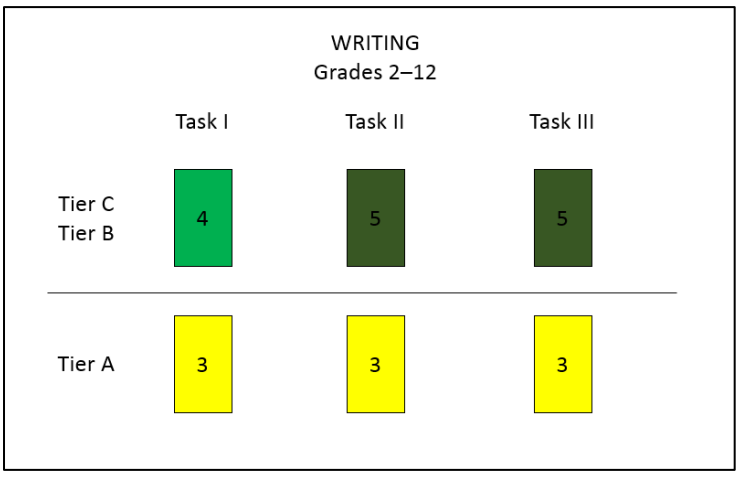


Figure 5. Distribution of tasks by targeted proficiency level for each tier of the Grades 2–12 Writing test.

Although timing guidance is provided to test administrators in the Test Administrator Manual, the Writing subtest is untimed.

2.1.4. Speaking

For the ACCESS Speaking test, Table 5 shows, for each grade-level cluster and tier, the number of tasks, the targeted range of WIDA proficiency levels, the task type, the response format, and the scoring procedure.

Table 5
Number and Types of Items on the Speaking Subtest

Grade-Level Cluster	Tier	Number of Tasks	Targeted PL range	Task Types	Response Formats	Scoring Procedures
1	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
1	B/C	6	PL3 - PL5			
2	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
2	B/C	6	PL3 - PL5			
3	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
3	B/C	6	PL3 - PL5			
4-5	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
4-5	B/C	6	PL3 - PL5			
6-8	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
6-8	B/C	6	PL3 - PL5			
9-12	A	6	PL1 - PL3	Speaking Constructed Response	Polytomous Constructed Response	Human Scored; Scored by Test Administrator
9-12	B/C	6	PL3 - PL5			

Figure 6 shows the format of the Speaking test. The Speaking test includes tasks that target language elicitation at three proficiency levels: 1, 3, and 5. The tasks are grouped into thematic folders, which are aligned to one or two of the WIDA Standards. These folders are generally presented in the same order as the folders in the Listening and Reading subtests; folders aligned to SIL are presented first, then folders aligned to LoLA, and then folders aligned to LoMa.

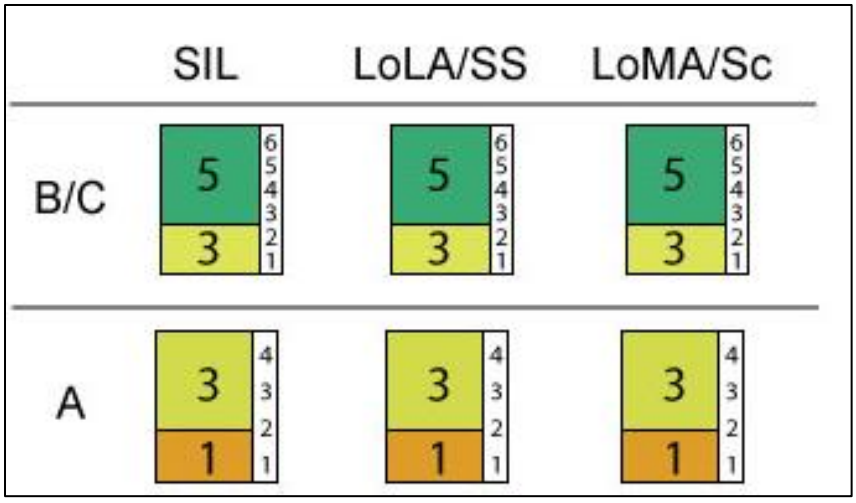


Figure 6. Distribution of tasks for each tier of the Speaking test.

As shown in Figure 5, the Speaking test includes two tiers. 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 A and Tier B/C folders that address Social and Instructional Language in a given grade cluster make up a single thematic panel, with the PL 3 tasks shared across tiered folders in a panel. In other words, within a Social and Instructional Language panel, the same PL 3 task appears on both the Tier A and the Tier B/C form.

Although timing guidance is provided to test administrators in the Test Administrator Manual, the Speaking subtest is untimed.

2.2. Test Construction

2.2.1. Item Development

ACCESS Series 501 Paper is one of two static rotating Paper test forms. The ACCESS testing program transitioned in 2016 from an entirely paper-based program to the launch of ACCESS in both Online and Paper formats.

The Listening and Reading items for ACCESS Paper were developed prior to the launch of ACCESS Online, when ACCESS was entirely paper based. Most Writing tasks were developed for ACCESS when it was entirely paper based; however, a small subset of Writing tasks on ACCESS Series 501 Paper were developed as online tasks that were subsequently reformatted for administration as paper-based tasks. The Speaking tasks were developed and field tested as online tasks before being reformatted for administration as paper tasks.

The general process of item writing and editing, and of item content and bias and sensitivity reviews, remains similar across these transitions. For ACCESS Paper items, trained item writers worked from item specifications to draft items within a thematic folder. After initial development, folders were screened at the Center for Applied Linguistics (CAL), and those that were approved for further development underwent 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, were produced.

After items were internally refined, they were reviewed by two panels: a content review panel and a bias and sensitivity review panel. The panels consisted of specially trained educators with culturally and linguistically diverse backgrounds from WIDA Consortium states. Items were submitted to the content review panel to ensure that the content was accessible and relevant to students in the targeted grade-level cluster and at the targeted proficiency level and that each item or task matched the MPI from the WIDA ELD Standards that it was intended to assess. Content reviewers were educators from WIDA states with relevant ESL and /or content-area teaching experience. The bias and sensitivity review panel ensure 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 were 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 the recommendations of the two panels, the items were revised as necessary.

For Writing and Speaking tasks, after external bias and sensitivity review and content review, tasks were subject to small-scale tryouts, led by CAL staff. In these tryouts, candidate folders were administered to students; student responses, as well as observations and interviews, informed further revisions to the folders. If tasks were deemed appropriate after tryouts, they then moved to the field testing stage.

Note that this section applies to ACCESS Paper Grades 1–12. For detail on Kindergarten, see Section 2.4 below and the technical report on the development of the Kindergarten static form (MacGregor et al., 2009).

2.2.2. Field Testing and Item Selection

2.2.2.1. Listening and Reading

The Listening and Reading items for ACCESS Paper were created prior to the launch of ACCESS Online and were created when ACCESS was entirely paper based. ACCESS was first field tested in 2004, and from 2004 to 2014, development continued for ACCESS, culminating in Series 303, operational in 2014–2015. For further detail on this original field test and on the processes for ongoing item development from 2004 to 2014, see the ACCESS for ELLs Technical Reports, particularly ACCESS for ELLs Technical Report No. 1, *Development and*

Field Test of ACCESS for ELLs (Kenyon, 2006) and *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 303* (Kenyon, 2006).

In all grade clusters, the Tier A Listening and Reading forms are static forms, which were constructed prior to the launch of ACCESS Online.

In all grade clusters, the Tier B/C forms in Listening and Reading are new forms for Series 501. These forms are composed of items that were previously operational in Series 400 and 401 and that were developed, as described above, during the development cycles when ACCESS was entirely paper based. Beginning with Series 403, to streamline operational administration, the ACCESS Paper Listening and Reading Tier B and Tier C tests were combined to create a new Tier B/C test in Listening and in Reading for each grade-level cluster.

In order to select these new forms, the pool of Listening and Reading Paper Tier B and Tier C items that were administered to the Series 401 and Series 400 populations was recalibrated using the population data (see Part 2, Section 2.7 for more information on the recalibration). A forms selection meeting was conducted in early 2018, prior to the operational administration of Series 403. Staff from WIDA and CAL reviewed the pool of items in Series 401 and 400 Listening and Reading Tier B and Tier C and selected two new static Tier B/C forms for each grade-level cluster in Listening and Reading—one for use in Series 403 and the other for use in Series 501. Forms were selected to maintain the coverage of WIDA ELD Standards as called for in the test design and to ensure inclusion of items of sufficient difficulty to measure students in the Tier C range.

2.2.2.2. Writing

There are two static rotating forms for ACCESS Paper Writing. The first of these is composed of the same set of items, across all grade-level clusters and tiers, as the test used the first year of ACCESS Online. The second form is composed of the same set of items, across all grade-level clusters and tiers, as the test used the second year of ACCESS Online.

Tasks on the first of the two rotating static forms were used operationally prior to the launch of the Online test and were re–field tested in the Online mode for the first year of ACCESS Online. Tasks selected for use in the first Online operational test were then reformatted for presentation in the first of the Paper static forms.

The second rotating static form uses continuing tasks from the first form, as well as tasks newly field tested for the second year of ACCESS Online and then reformatted for Paper presentation. For further detail on this field test, see the Series 401 Online ACCESS technical report (Center for Applied Linguistics, 2018).

ACCESS Paper 501 is the first of the two rotating static forms.

2.2.2.3. Speaking

The Speaking test for ACCESS Paper is likewise one of two static rotating forms. The first of these forms is composed of the same set of items, across all grade-level clusters and tiers, as the second year of the ACCESS Online Speaking test; the second form is composed of the same set of items, across all grade-level clusters and tiers, as the third year of the ACCESS Online Speaking test. Speaking tasks have some differences in presentation between Online and Paper. In addition, the Paper test does not include the Speaking tier Pre-A, which is included in the Online test.[1]

Tasks for these two rotating forms were field tested during the initial ACCESS Online field test, as well as embedded during the first and second years of the ACCESS Online assessments. These speaking tasks went through both quantitative and qualitative analyses following the field test to determine their appropriateness for inclusion in the next year's operational test. After field testing, the Speaking tasks were then produced in the paper-based format.

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. Note that this section applies to ACCESS Paper Grades 1–12. For detail on the item and task design for Kindergarten, see Section 2.4 below and the technical report on the development of the Kindergarten static form (MacGregor et al., 2009).

2.3.1. Listening Items

All Listening items are multiple choice and are designed to be group administered. They 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.

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.

[1] Students with very low ability levels in the Listening and Reading domains are routed to the Pre-A tier for Speaking in the Online test. The purpose of the Pre-A tier is to reduce the affective impact of the test on these students. As the Paper test is not adaptive, there is no way to route these students to Pre-A for Paper.

In ACCESS Paper, the Listening tests have a Tier A and a Tier B/C form for each grade-level cluster; students are placed into the tier based on a decision made at the school or district level as local EL teachers judge students' abilities based on their classroom performance.

Listening items are developed so that each folder appears on a 2-page spread in a test booklet, although some folders go onto a third page. 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 done via audio CD, and explicit instructions on starting and pausing the CD are provided in the Test Administrator's Script and the Test Administrator Manual.

Listening items are centrally scored by Data Recognition Corporation (DRC) via an automated process.

2.3.2. Reading Items

All reading items are multiple choice and are designed to be group administered. They are similar in format to Listening items. Reading items are selected-response items, with one key and either two or three distractors, depending on grade-level cluster and targeted proficiency level. For Grades 1 and 2, all items have a key and two distractors. For Grades 3, 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.

The stimulus for Reading items is written text, and answer choices primarily are also written text, though for Grades 1–12 response options for items targeting PLs 1, 2, and 3 may be illustrations rather than text. As with Listening items, Reading items are grouped into thematic folders of three test items each. In ACCESS Paper, the Reading tests have a Tier A and a Tier B/C form for each grade-level cluster; students are placed into the tier based on a decision made at the school or district level.

Reading items are centrally scored by DRC via an automated process.

2.3.3. Writing Tasks

All writing tasks are constructed response tasks and are designed to be group administered. Students write responses by hand in paper booklets.

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.

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 written 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. In Grades 1–5, a word box may be provided, depending on the grade level, targeted proficiency level, and task.

For all grade clusters and tiers, the Writing test is group administered by a live test administrator. The test administrator reads instructions aloud from the Test Administrator's Script and monitors student progress through the test. For all grade clusters and tiers, the students hand-write their answers in the same test booklet containing the Listening and Reading tests.

2.3.4. Speaking Tasks

The Speaking test is administered individually to each test-taker. The test is media delivered. Students listen to an audio recording of the test input while following along in a test booklet.

Stimuli on the Speaking test include graphics, audio, and text, presented in a test booklet as a series of "speech bubbles" from the perspective of the virtual test administrator (VTA) and virtual model student. All text is multimodal, presented both in the test booklet and read aloud on the audio CD. Scripts containing the task content are audio recorded with professional voice actors and produced by a professional recording studio. Audio playback of test item content is done via audio CD, and explicit instructions on starting and pausing the CD are provided in the Test Administrator's Script and the Test Administrator Manual.

The CD audio stimuli are presented in terms of a 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 the test administrator, who scores the Speaking test. Students respond orally to the tasks, with their responses scored immediately by the test administrator using a scoring scale. The test administrator records scores on the Speaking test in the same booklet the student used for the Listening, Reading, and Writing tests.

2.4. Kindergarten

The Kindergarten test is a static form and is not refreshed from year to year.

2.4.1. Test Design

The design of K ACCESS is intended to be engaging for very young children, and the test design was informed by consultation with Kindergarten teachers and a panel of early childhood assessment experts. The test design incorporates a high-interest, age-appropriate storybook format, using child-friendly graphics, and includes manipulatives for students to demonstrate comprehension. The test is built on two thematic texts in a storybook format, one narrative and one expository. The storybook is read aloud by the test administrator. There are Listening, Speaking, Reading, and Writing assessment tasks related to each text. In order to minimize testing times and to ensure that students are presented with assessment tasks appropriate to their abilities, the test includes stopping rules (designed to ensure that children of beginning proficiency are not overchallenged) and skipping rules (designed so that children of more advanced proficiency can skip forward to more challenging tasks).

The test is administered one-on-one by trained test administrators, who mark up responses in the Student Response Booklet.

Table 6 provides, for each domain, the number of items, the targeted range of WIDA proficiency levels, the item types, the response format, and the scoring procedure.

Table 6
Number and Types of Items on Kindergarten ACCESS

Domain	Number of Items	Targeted PL range	Item Types	Response Formats	Scoring Procedures
Listening	30	P1-P5	Dichotomous	Student points to picture or manipulates cards	Administrator records response (correct/incorrect) in Student Response Booklet
Speaking	10	P1-P5	Dichotomous	Oral response	Administrator records response (correct/incorrect) in Student Response Booklet
Writing	6	P1-P5	Dichotomous and Polytomous	Student handwrites in booklet	Administrator records response (correct/incorrect) for dichotomous tasks. Administrator rates responses and records rating for polytomous tasks
Reading	30	P1-P5	Dichotomous	Student reads aloud or matches picture cards with text cards	Administrator records response (correct/incorrect) in Student Response Booklet

2.4.2. Test Construction

Field testing for Kindergarten ACCESS was conducted in 2008. A full description of item development, field testing, final forms selection, and initial standard setting for Kindergarten can be found in the technical brief *Development and Field Test of Kindergarten ACCESS for ELLs* (MacGregor et al., 2009). Cut scores for Kindergarten were most recently updated in the 2016 ACCESS standard setting (Cook & MacGregor, 2017); see Part 2 Section 2.1 for more information.

2.4.3. Item and Task Design

As noted above, the Kindergarten ACCESS test is composed of two thematic texts. The items and tasks are designed to build upon the content of these texts.

In the domain of Listening, the test administrator reads the prompt aloud to the student, and the student responds by either pointing to an item in a picture or manipulating a picture card. The test administrator records the response (correct or incorrect) in the Student Response Booklet.

Students respond to Writing tasks in the Student Response Booklet. The initial Writing tasks for each thematic text are dichotomously scored by the test administrator. The Test Administrator Script indicates the level required for a task to meet expectations and to be scored correct. The final Writing task in each thematic text section is scored on a rating scale. The test administrator rates the student's Writing on a scale of 0 to 6.

Speaking tasks are read aloud, and students respond orally. Tasks are dichotomously scored by the test administrator. The Test Administrator Script indicates the level required for a task to meet expectations and to be scored correct.

To administer Reading tasks, test administrators ask students to identify letters or read text. Students respond by manipulating picture cards or by pointing at pictures. Students may also read aloud. The test administrator records the response (correct or incorrect) in the Student Response Booklet.

The items on Kindergarten ACCESS were developed to collectively assess all five WIDA Standards in all domains across the proficiency levels, as shown in Table 7. In order to keep the test an appropriate length for the population, it was not possible to assess each Standard at each proficiency in each domain. Therefore, tasks were distributed by Standard across the proficiency levels and domains in order to achieve appropriate coverage.

Although the average time per test is provided to test administrators in the Test Administrator Manual, Kindergarten ACCESS is untimed.

Student Response Booklets are centrally scanned at DRC.

Table 7

Number of Items by WIDA Standard and Targeted Proficiency Level on Kindergarten ACCESS

Listening										
WIDA Standard	Narrative Storyline					Expository Storyline				
	Number of items at targeted PL range					Number of items at targeted PL range				
	1	2	3	4	5	1	2	3	4	5
SI	3		3		3	3				
LA									3	
MA							3			
SC										
SS		3		3				3		3
Speaking										
WIDA Standard	Narrative Storyline					Expository Storyline				
	Number of items at targeted PL range					Number of items at targeted PL range				
	1	2	3	4	5	1	2	3	4	5
SI			3						3	
LA					3					
MA							3	3		3
SC						3				
SS	3	3		3						
Writing										
WIDA Standard	Narrative Storyline					Expository Storyline				
	Number of items at targeted PL range					Number of items at targeted PL range				
	1	2-5				1	2	3	4/5	
SI	1					1				
LA										
MA							3			
SC										
SS								4		
IT (SIL, LoLA, LoSS)		1							1	
Reading										
WIDA Standard	Narrative Storyline					Expository Storyline				
	Number of items at targeted PL range					Number of items at targeted PL range				
	1	2	3	4	5	1	2	3	4	5
SI	3				3	3			3	
LA										
MA										
SC		3	3				3	3		3
SS				3						

3. Assessment Performance: The Implementation of ACCESS

3.1. Test Delivery

Administration of ACCESS Paper takes place between December and April of the academic year, with testing windows determined at the state level. The domain tests may be administered in any order. The test may be administered in several sessions within 1 day or over a series of days.

The Listening and Reading tests may be group or individually administered. Students are administered the Listening and Reading test forms using paper test booklets, and students record their answers directly in the test booklets. For the Listening test, the audio stimuli are played aloud via an audio CD.

The Writing test may be group or individually administered. Students are administered the Writing test via paper test booklets. Students record their responses directly in the test booklet.

The Speaking test is individually administered. Students listen to an audio recording and follow along in an accompanying test booklet. Each task also includes a model student response, which serves as an exemplar to the student and also as a benchmark to the test administrator who will score the task. All audio stimuli are presented via audio CD.

3.2. Scoring Procedures

3.2.1. Multiple-Choice Scoring: Listening and Reading

Listening and Reading items are scored dichotomously, as correct or incorrect. Students mark their answers directly in their test booklets, and each page is scanned into an electronic database. Scale scores for each domain are calculated based on the items that are administered to the test-taker and the number 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 Writing

Performance-based tasks in the domain of Writing are scored by trained raters. DRC retains a number of raters from year to year. This pool of experienced raters was drawn from to staff the scoring of the ACCESS for ELLs. To complete the rater staffing, recruiting events were held and applications for rater positions were screened by DRC's recruiting staff. Candidates were personally interviewed by DRC staff. In addition, each candidate was required to provide an on-demand writing sample, an on-demand math sample, references, and proof of a 4-year college degree. In this screening process, preference was given to candidates with previous experience scoring large-scale assessments and degrees emphasizing expertise in English language arts. The rater pool consisted of educators, writers, editors, and other professionals with content-specific

backgrounds. These individuals were valued for their content-specific knowledge, but they were required to set aside their own biases about student performance and accept the scoring standards outlined in the training for scoring the ACCESS for ELLs.

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 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 10 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 below 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.

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.

- The purpose of the first and second ratings is to monitor interrater reliability by comparing the scores given by two separate raters to the same response. When calculating final scores, the first score given is the score of record.

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 used both recalibration sets and validity responses to monitor handscoring quality control. Recalibration sets and validity responses were developed in conjunction with DRC, CAL, and WIDA. CAL developed an initial pool of responses for use as recalibration and validity by selecting responses from a previous administration of the tasks (e.g., a field test). This pool of responses and their scores were reviewed and approved by WIDA staff. DRC supervisors supplemented this pool of responses as needed by selecting additional responses; these responses and their scores were reviewed and approved by CAL and WIDA before use. For each of the first 5 days that raters scored a task, they took one recalibration set of five responses per task. The recalibration sets did not differ from rater to rater. For example, a recalibration set was specified for the first day that a rater scored a specific task; every rater who scored that task took this same recalibration set on the first day that they scored that task. After the raters took the recalibration sets, the scoring director or team leader reviewed the set 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 that a rater was scoring a task, DRC used validity responses to continue monitoring rater performance. The validity responses were seeded into operational scoring; the raters did not know which responses were operational and which were validity responses. Reports generated on a daily basis compared the scores given by each rater to the "true" score for each validity response. When a rater was working on a task, the validity responses were dealt to that rater in a random order. Each validity response was dealt to multiple raters over the course of the project (i.e., given enough time, every rater working on a task would score every validity response for that task), but the validity responses were not dealt in the same order to each rater.

Handling Unusual Responses

The following processes were in place to manage specific types of "unusual" responses:

- **Scoring questions**. If raters had questions about the application of the scoring guidelines to a response (e.g., if they were uncertain as to the proper score that should be assigned), the raters forwarded the response to team leaders for assistance. The team leaders then reviewed the response and applied the proper score. If anything about the response and the rater's question indicated that the rater needed any clarifications about the scoring guidelines, the team leaders met with raters to review the response and to explain how to score it based on the scoring guidelines.
- **Nonscore codes.** Unusual or aberrant responses that could not be assigned a score based on the scoring guidelines received a nonscorable code (e.g., Writing responses that are entirely blank or consist entirely of scribbles or pictures). DRC's handscoring team collaborated with WIDA and CAL to define what specifically constitutes a nonscorable response in order to ensure consistency of nonscorable codes, and this information was provided from CAL to DRC along with other item-specific training materials that were used to train DRC's raters. During scoring, when scorers apply a nonscoreable code (with the exception of Blank), the response was automatically forwarded to a handscoring supervisor for review and approval. If the handscoring supervisors had any questions about the application of nonscore codes to specific responses, DRC contacted WIDA and CAL representatives for further review and discussion.
- **Alerts.** To handle possible alert papers (i.e., student responses indicating potential issues related to the student's safety and/or well-being that may require attention at the local level, potential plagiarism, or potential teacher interference), DRC's imaging system gave scorers the ability to alert questionable student responses. When a response was flagged with the alert status, it was automatically routed to handscoring supervisors for review. When the handscoring supervisors concurred with the "alert" status of the response, the response was then passed on to WIDA's project management team who provided the response to the appropriate local education agency.
- **Request for originals.** When raters came across a scanned student response that was difficult to read (for example, having some partially erased text), the rater would flag the response with a "request original" status. When a response was flagged as "request original," it was automatically forwarded to a handscoring supervisor. If the handscoring supervisor agreed that the original student response needed to be reviewed in order to properly apply the scoring guidelines, the request was forwarded to staff in DRC's Operations Services, who located the original student response so that it could be reviewed by handscoring supervisors in order to score the response.

Changes in Scoring Procedures Due to the COVID-19 Pandemic

During the second half of March 2020, DRC pivoted from site-based scoring to remote scoring in order to continue handscoring operations in the safest manner. DRC's remote scoring was designed to very closely emulate the work done in the physical scoring locations. The platform, content, and expectations for quality remained the same, and interactive technology and content

training and discussions were conducted live (virtually). The differences came with the method through which training was delivered (online) and in the modes of communication used (web screen sharing, webcast, video chat, and chat). Scoring leaders were equipped with a variety of tools to ensure every rater was successful in understanding and applying scoring criteria to student responses.

Remote scoring began with a training session to guide supervisors and raters through the use of the tools that DRC utilized for remote scoring. These training sessions took place in late March and were completed by early April. Once supervisors and raters were trained on the remote scoring process, handscoring resumed for the ACCESS assessments. A description of DRC's remote scoring process follows.

- **System tools—scoring, training, chat.** ScoreBoard is DRC's secure, web-based scoring application that is designed to be used in a distributed environment. The platform is used within DRC's scoring centers and in remote locations (e.g., in a rater's home). Integrated training resources provide the capability to securely maintain digital training materials within the scoring platform itself.

  Live, interactive training was conducted via Moodle Learning Management System, which mirrors aspects of the scoring room and provides a versatile platform for training. It also served as a place to share files of important documents, including daily scoring statistics and platform user guides. Through embedded communication tools, Scoring Directors, Assistant Scoring Directors, and Team Leaders facilitated group and one-on-one training sessions and discussions using audio and video.

  To facilitate instant communication between supervisors and raters, DRC utilized a chat tool called Zulip in conjunction with ScoreBoard and Moodle. Zulip provided a tool for raters to directly ask supervisors questions about responses and allowed supervisors to direct individuals or groups of raters to join Moodle training rooms for important discussions and retraining.

- **Security.** Security is essential to the handscoring process. When users logged into ScoreBoard, they were required to read and accept the security policy before they were allowed to access the project. Raters were also required to read and sign nondisclosure agreements. During training and large-group discussions, emphasis was always given to what security means, the importance of maintaining security, and how this is accomplished. In the remote environment, these security reminders were given daily. Raters working remotely were required to work in a private environment away from other people (including family members). Restrictions built into ScoreBoard defined the hours during the day raters were able to log into the system, ensuring that raters were only scoring responses while supervisors were in place to monitor handscoring and answer any questions.

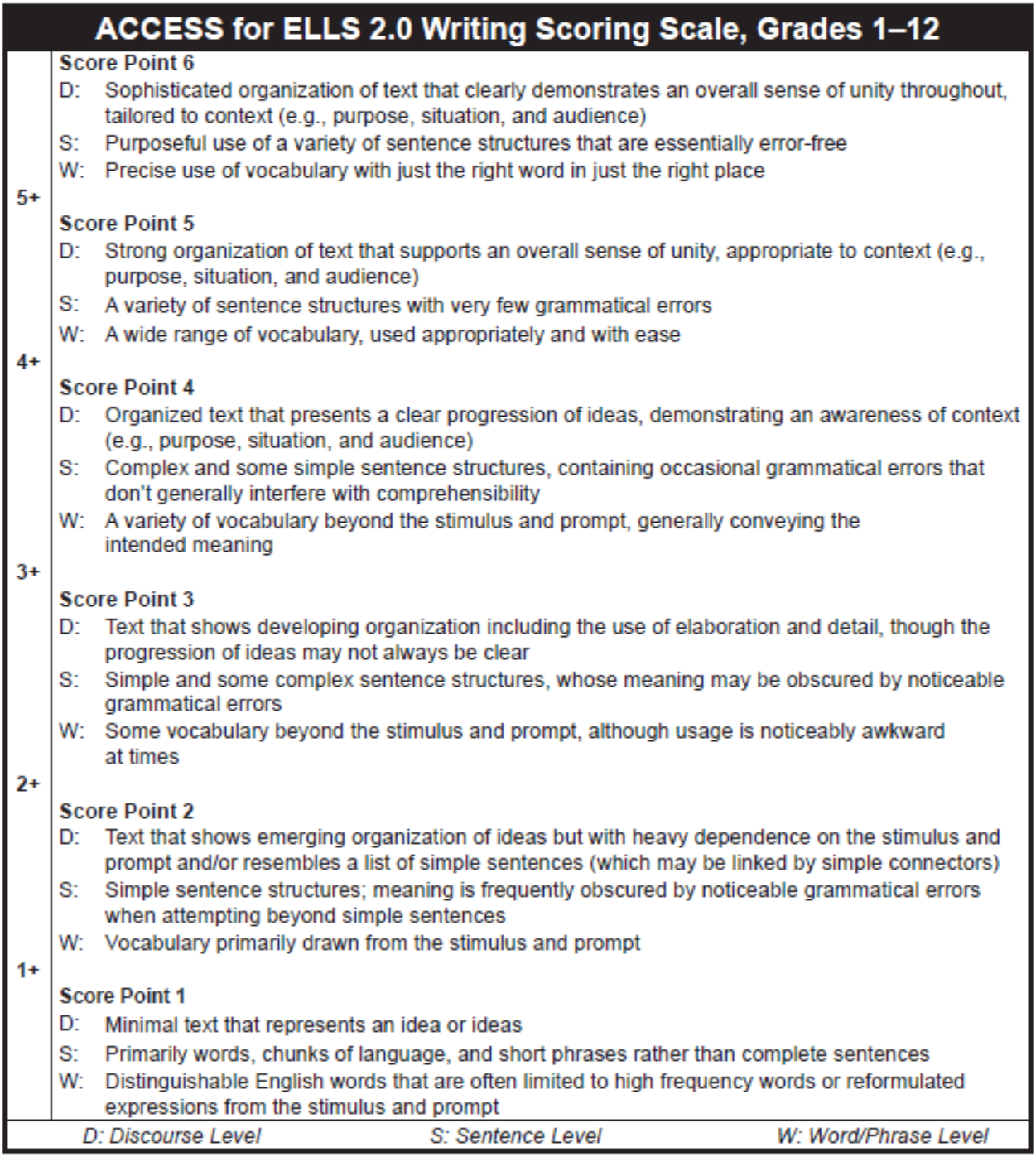
- **Content training with Moodle.** While DRC enabled capacity for remote content training, for Paper ACCESS, all content training for operational items was already completed while raters were onsite.
- **Quality control.** DRC's robust quality control processes and handscoring metrics were identical for onsite and remote scoring sessions. During remote scoring, scored responses were monitored with second reads exactly as they were at the scoring sites. Read-behinds were also conducted in the exact same manner; however, any conversations and/or retraining needed as a result of the monitoring were held in one-on-one video chat sessions. Handscoring quality reports continued to be available daily and on demand for handscoring supervisors and DRC's project leadership, and DRC continued to provide WIDA staffing with handscoring reports on the same schedule as when handscoring was onsite.

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).

Figure 7 shows the Writing Scoring Scale.



ACCESS for ELLS 2.0 Writing Scoring Scale, Grades 1–12

	Score Point 6 D: Sophisticated organization of text that clearly demonstrates an overall sense of unity throughout, tailored to context (e.g., purpose, situation, and audience) S: Purposeful use of a variety of sentence structures that are essentially error-free W: Precise use of vocabulary with just the right word in just the right place
5+	
	Score Point 5 D: Strong organization of text that supports an overall sense of unity, appropriate to context (e.g., purpose, situation, and audience) S: A variety of sentence structures with very few grammatical errors W: A wide range of vocabulary, used appropriately and with ease
4+	
	Score Point 4 D: Organized text that presents a clear progression of ideas, demonstrating an awareness of context (e.g., purpose, situation, and audience) S: Complex and some simple sentence structures, containing occasional grammatical errors that don't generally interfere with comprehensibility W: A variety of vocabulary beyond the stimulus and prompt, generally conveying the intended meaning
3+	
	Score Point 3 D: Text that shows developing organization including the use of elaboration and detail, though the progression of ideas may not always be clear S: Simple and some complex sentence structures, whose meaning may be obscured by noticeable grammatical errors W: Some vocabulary beyond the stimulus and prompt, although usage is noticeably awkward at times
2+	
	Score Point 2 D: Text that shows emerging organization of ideas but with heavy dependence on the stimulus and prompt and/or resembles a list of simple sentences (which may be linked by simple connectors) S: Simple sentence structures; meaning is frequently obscured by noticeable grammatical errors when attempting beyond simple sentences W: Vocabulary primarily drawn from the stimulus and prompt
1+	
	Score Point 1 D: Minimal text that represents an idea or ideas S: Primarily words, chunks of language, and short phrases rather than complete sentences W: Distinguishable English words that are often limited to high frequency words or reformulated expressions from the stimulus and prompt

D: Discourse Level *S: Sentence Level* *W: Word/Phrase Level*

Figure 7. Writing Scoring Scale.

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, the rater awards a plus score point between the two applicable whole score points.

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, as defined below.

> **Nonscorable**: The response is blank; consists only of verbatim copied text; consists only of text that is completely off task; or is entirely in a language other than English.
>
> **Completely off-task response**: The entire response shows no understanding of or interaction with the prompt. It may be a memorized, previously practiced response or appear to answer another, unrelated prompt. A response that is entirely off task is nonscorable.
>
> **Completely off-topic response**: The entire response shows a misinterpretation or misunderstanding of the prompt. An off-topic response is related to the prompt, but does not seem to address it as intended. However, the response is clearly not a memorized, previously practiced response. These responses are scored in their entirety using the scoring scale; however, the maximum holistic score for a completely off-topic response is 2+.
>
> **Partially off-task response**: The response contains both off-task and on-task writing. These responses are scored by ignoring the off-task portion (which may be memorized and previously practiced) and scoring only the on-task portion using the scoring scale.
>
> **Partially off-topic response**: The response contains both off-topic and on-topic writing (i.e., a portion of the response shows a misinterpretation or misunderstanding of the prompt). These responses are scored in their entirety using the scoring scale.

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 8.

Table 8
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

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 to 1 and 0 to 3, respectively. The third and fourth tasks 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.

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

The Speaking test is scored using a scoring scale that is designed to evaluate student responses relative to the model student's response. (See Section 2.3.4 above for more information about the role of the model student in the design of the Speaking tasks.) As part of test administration, the test administrators hear the model student response before each student response, which supports them in assigning an appropriate score relative to the model response. Speaking responses are

immediately scored by the administrator while the test is administered. After listening to the student's responses, the administrator assigns a score.

The Speaking Test is the only portion of ACCESS Paper that is scored locally. Test administrators must complete the relevant virtual ACCESS Paper test administrator training module for the Speaking test and pass the accompanying quiz (either Grades 1–5 or Grades 6–12). The training focuses on developing the test administrators' ability to score the test reliably. Separate training materials are available that address test administration and monitoring procedures. To help ensure that test administrators reliably score the test, they are trained on the Speaking Scoring Scale. Training materials are available for each grade-level cluster, and raters listen to anchor samples and view score justifications that provide detailed explanations for scores based on the scoring scale. Practice samples are also available so that raters can practice assigning scores. The course includes both required training material for each grade-level cluster as well as optional training material. Raters are required to complete training sections for each grade-level cluster they will administer and score. However, if a rater will score more than three grade-level clusters, they may complete rater training for only three. The quizzes include 12 Speaking rating tasks in which raters listen to and assign a score to a task response. The pass rate for the quiz is 80% correct.

The Speaking Scoring Scale defines five score points: *Exemplary*, *Strong*, *Adequate*, *Attempted*, and *No Response (in English).* The *No Response* score point only applies if the examinee refuses to respond, or if the examinee responds in a language other than English.

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 2.3.4). 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.

Figure 8 shows the Speaking Scoring Scale.

ACCESS for ELLs 2.0 Speaking Scoring Scale	
Score point	**Response characteristics**
Exemplary use of oral language to provide an elaborated response	• Language use comparable to or going beyond the model in sophistication • Clear, automatic, and fluent delivery • Precise and appropriate word choice
Strong use of oral language to provide a detailed response	• Language use approaching that of model in sophistication, though not as rich • Clear delivery • Appropriate word choice
Adequate use of oral language to provide a satisfactory response	• Language use not as sophisticated as that of model • Generally comprehensible use of oral language • Adequate word choice
Attempted use of oral language to provide a response in English	• Language use does not support an adequate response • Comprehensibility may be compromised • Word choice may not be fully adequate
No response (in English)	• Does not respond (in English)

Figure 8. Speaking Scoring Scale.

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 9. 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 student who had been routed to the B/C test would easily achieve a score of *Adequate and Above* on these 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 9
Rating to Raw Score Conversion (Speaking)

Rating	Raw score
No Response (in English)	0
Attempted	1
Adequate/Adequate and Above	2
Strong	3
Exemplary	4

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. Listening Test Administration

The ACCESS for ELLs Paper Listening test is media delivered. Listening test items are delivered via CD.

3.3.1.1. Listening Test Materials

Test materials include the following items:

- Test Administrator's Script
- Student Test Booklet(s)
- Listening and Speaking Test CD (a separate CD for each grade-level cluster and tiered test form). In the rare event that a student requires a human reader as an accommodation, the Recording Script is required to administer the Listening section individually for that particular student.
- At least one sharpened number 2 pencil for each student to mark responses
- Speakers
- A CD player or desktop/laptop computer (to play the CD)

3.3.1.2. Organization and Timing of the Listening Test

The Listening test is designed to take approximately 25 to 40 minutes, depending on the grade-level cluster and tier. The test administration time does not include time for convening students, taking attendance, distributing and collecting test materials, explaining test directions, or completing practice items. The length of test items increases with students' language proficiency

and grade level. For example, the Tier B/C Listening test takes longer to administer than the Tier A Listening test, and the Listening test for Grades 9–12 may take slightly longer than the test for Grades 4–5.

3.3.2. Reading Test Administration

The ACCESS for ELLs Reading test is completed within Student Test Booklets after a scripted introduction by the Test Administrator.

3.3.2.1. Reading Test Materials

Reading test materials include the following items:

- Test Administrator's Script
- Student Test Booklet(s)
- At least one sharpened number 2 pencil for each student to mark responses

3.3.2.2. Organization and Timing of the Reading Test

The Reading test is designed to take no more than 35 to 45 minutes. The test administration time does not include time for convening students, taking attendance, distributing and collecting test materials, explaining test directions, or completing practice items.

3.3.3. Writing Test Administration

Students respond to a set of tasks, writing their responses in their Student Test Booklets.

3.3.3.1. Writing Test Materials

Writing test materials include the following items:

- Test Administrator's Script
- Student Test Booklet(s)
- At least one sharpened number 2 pencil for each student to write responses
- Scratch paper

3.3.3.2. Organization and Timing of the Writing Test

There are three tasks (Parts A, B, and C) on each Tier (Tiers A and B/C) of the Writing test for all grade levels except Tier A for Grade 1, which contains four tasks. For grade-level clusters 2, 3, 4–5, 6–8, and 9–12, the Tier A Writing tests have recommended guidelines for Parts A, B, and C of 15 minutes each, with up to 5 additional minutes for each part if needed for students to finish writing, for a total of 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.

3.3.4. Speaking Test Administration

The ACCESS for ELLs Speaking test is an individually administered test that standardizes test administration across students. Speaking test items are media delivered. Speaking test audio is provided on the same CD as the Listening test. The Speaking test provides ELLs with the opportunity to demonstrate their academic English language proficiency in speaking across the WIDA ELD Standards through a set of constructed-response tasks. The Speaking test is tiered. Students will either take the Tier A form or the Tier B/C form; both are included in the same Speaking Test Booklet.

3.3.4.1. Audio Format of the Speaking Test

The Speaking test is multimodal. The student hears audio input and also sees the input as text in the Speaking Test Booklet. This presentation format supports the student in understanding test input. Media delivery of the Speaking test means that an audio recording will guide the student through the Speaking test. The audio recording includes two voices: a model student and a virtual test administrator.

Each task on the Speaking test is preceded by a model student task and response. The questions posed to the model student are at the same proficiency level as the tasks to which the student will respond, allowing the model student to demonstrate the expected language use at a given proficiency level. In most cases the model questions are designed to be parallel to but not exactly the same as the examinee questions. The model student also has an important function in scoring, since the scoring scale is designed to evaluate student responses relative to the model student's response.

The virtual test administrator guides the student through the test and asks the student questions designed to elicit language at targeted proficiency levels. While the virtual test administrator will instruct and guide the student through the Speaking test, the administrator may also need to assist the student in navigating test materials (e.g., turning the page when prompted). The Speaking test includes standardized, built-in response time for every task. The amount of time varies according to the grade-level cluster, tier, and proficiency level of the task and ranges from 15 to 50 seconds in Grades 1–3 and from 15 to 45 seconds in Grades 4–12. Students may not require the entire time allotted. After the response time has ended, the test audio will automatically continue to the next Speaking task.

3.3.4.2. Speaking Test Materials

Speaking test materials include the following items:

- Test Administrator's Script

- Speaking Test Booklet (contains test graphics and prompts)
- Student Test Booklet (contains Speaking test scoring sheet and scoring scale)
- Listening and Speaking test CD (a separate CD for each grade-level cluster and tiered test form). In the rare event that a student requires a human reader as an accommodation, the Recording Script is required to administer the Speaking section.
- A CD player or desktop/laptop computer (to play the CD)
- Speakers

3.3.4.3. Organization and Timing of the Speaking Test

Speaking tasks on the Speaking test are contained within three parts: A, B, and C. As in other domains of ACCESS for ELLs, tasks on the Speaking test are grouped thematically. Each part addresses one or more of the WIDA ELD Standards and contains two tasks. In all, the Speaking test contains six individual tasks across the three parts. Each task is associated with a proficiency level (1, 3, or 5) and includes one or two questions to which the student responds. Student questions are indicated by a blue speech bubble in the test booklet.

The Speaking test is designed to take approximately 15 to 35 minutes per student, but the actual time will depend on the grade-level cluster and tier of the test administered. Note that the approximate test administration time does not include setting up the test session or explaining test directions. An additional 10 minutes should be allocated to set up the Speaking test.

3.3.5. Test Administrator Training

To prepare individuals to serve as test administrators, test administrator training for ACCESS Series 403 Paper is conducted through online training modules hosted on the WIDA website. Three certifications are offered to participants: a group test administration certification pertaining to the Listening, Reading, and Writing portions of ACCESS; a certification for the Speaking test; and a certification for the Kindergarten test. In order to receive any of the three certifications, participants have to complete the relevant online course and pass a qualifying exam after completing the course.

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 Nondisclosure and User Agreement upon their first login. Use of the WIDA Assessment Management System and INSIGHT test engine are also

subject to the terms of use outlined in the WIDA Assessment Management System. 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. Accessibility and Fairness

The WIDA Accessibility and Accommodations Framework provides support for all ELLs, as well as targeted accommodations for students with individualized education plans (IEPs) or 504 plans. These supports are intended to increase accessibility to the assessments for all ELLs. (Please see the Accessibility and Accommodations Supplement for detailed information: https://wida.wisc.edu/resources/accessibility-and-accommodations-supplement.)

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 and Kindergarten 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.

Audio aids, color contrast, color overlay, highlighters, colored pencils or crayons, line guide or tracking tool, low-vision aids or magnification devices, sticky notes, and scratch paper are the universal tools used in the Paper administration.

3.4.2. Support Provided to ELLs with IEPs 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 who have 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. Accessibility features are provided to ELLs by test administrators for paper-based tests. 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. Features available during paper-based test administration include the following:

- Audio amplification device (provided by student)
- Highlighter, colored pencils, or crayons
- Place marker (blank)
- Low-vision aids or magnification device
- Color overlay
- Equipment or technology that the student uses for other tests and schoolwork, e.g., adapted pencil (altered size or grip), slant board, wedge, etc.
- Scratch/blank paper (submit with test or dispose 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 English language proficiency 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.

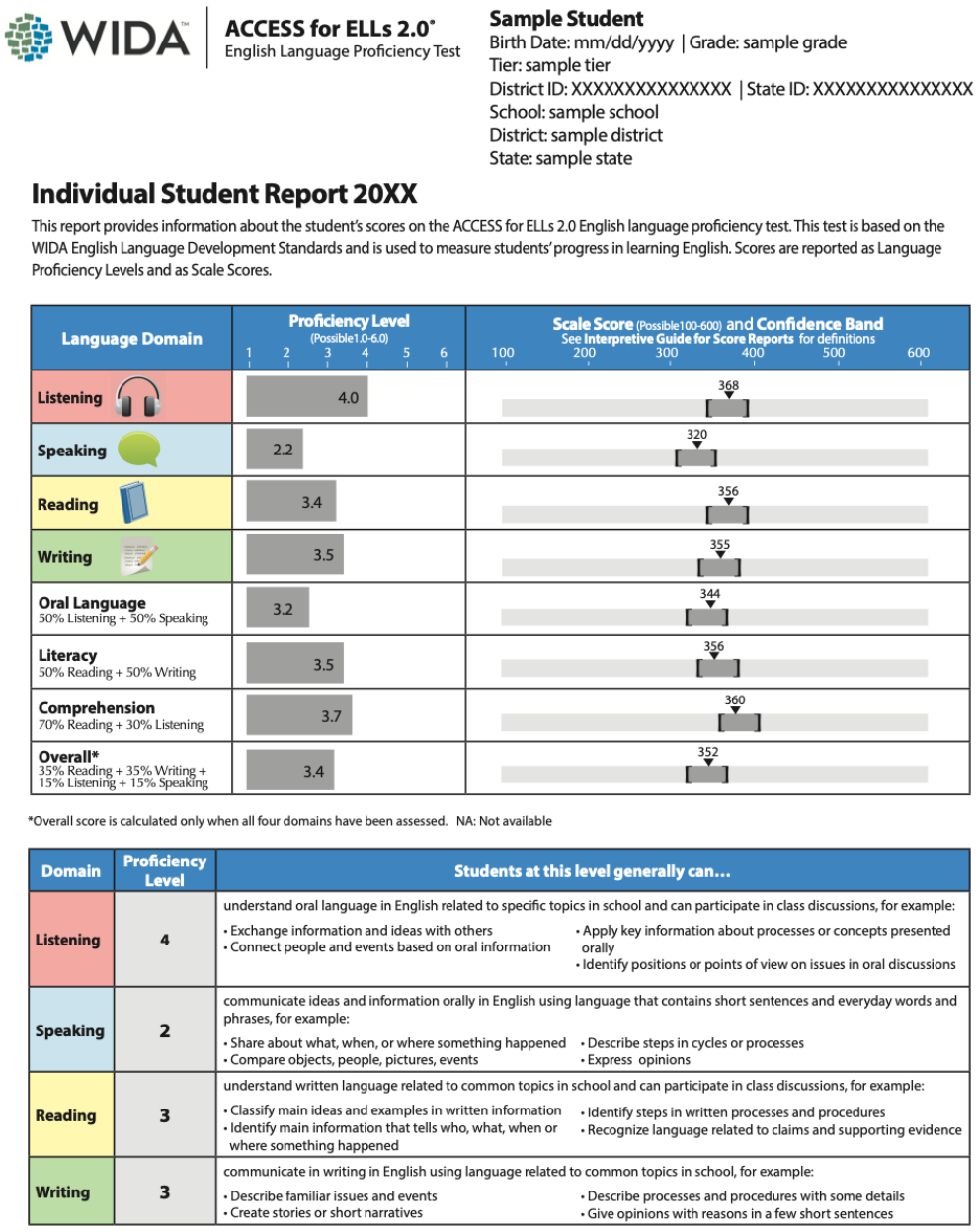
WIDA also offers *Alternate ACCESS for ELLs*. This test is intended only for those ELLs who have cognitive disabilities that are so significant as to prevent meaningful participation in ACCESS testing, even with accommodations. The results of the Alternate ACCESS for ELLs operational administration appear in a separate technical report.

WIDA also offers Braille Test for ELLs and Large Print Test. The Braille test is paper based and the translation and graphics are provided in either contracted or uncontracted Braille for Tier B (Grades 1–12). This test is used to provide access to the test for ELLs who are blind. For students with visual impairments, the Large Print Test is used, where the font size is increased to 18 point. For the online test, the magnification/zoom tool increases the on-screen font size up to 1.5× or 2×, depending on the size of the computer monitor.

4. Summary of Score Reports

4.1. Individual Student Report

The Individual Student Report (Figure 9) 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.



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: XXXXXXXXXXXXXX | State ID: XXXXXXXXXXXXXX
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 (Possible1.0-6.0) 1 2 3 4 5 6	Scale Score (Possible100-600) 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 9. Individual Student Report.

As shown in Figure 9, the score report includes four domain scores (Listening, Speaking, Reading and Writing) and four composite scores (Oral Language, Literacy, Comprehension, and Overall). 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. Composition of single domain scores in composite scores is presented in the individual student report.

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 6-point English Language Proficiency scale. The whole number reflects a student's English language proficiency 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 levels 3.0 and 4.0.

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. For ACCESS Scale Scores, the confidence band is equal to the 95% probability level.

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

The second part of the Student Report provides information about the individual student's proficiency levels as whole numbers and describes what students at the reported proficiency level may typically 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.

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.
- Either scale scores or proficiency level scores can be used to compare test scores from different years, 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 audiences for frequency reports are 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

Each shows 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

In this section of the report, detail is provided on students' participation in the assessment and on scale score and proficiency level results. These data are disaggregated in several ways, including by grade-level cluster, grade and tier, and also by gender, ethnicity, and race.

Analyses use the Census Bureau approach to reporting race and ethnicity (https://www.census.gov/topics/population/race/about.html). Ethnicity is conceptualized as a binary category (Hispanic or non-Hispanic). There are five categories for race: American Indian/Alaskan Native, Asian, Black/African American, Pacific Islander/Hawaiian, and White. The race and ethnicity categories 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. Starting with Series 202, 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 as one of the racial categories (e.g., Asian) and 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, then they are labeled non-Hispanic multiracial.

A total of 11 students were excluded from the analyses due to mismatches in students' tiers across domains. In addition, 12,152 students taking Paper ACCESS tests in Colorado used equated scores to the Online ACCESS tests; therefore, their score analyses were not included in this 501 Paper Annual Technical Report. For the equated scoring procedure, please refer to the WIDA mode-adjustment procedure report.

1.1 Participation

Participation in ACCESS Paper is shown in three ways: by grade-level cluster, by grade, and by tier. Participation data are reported by state, by gender, and ethnicity.

1.1.1 Grade-Level Cluster

Table 1.1.1.1 shows participation across the 40 WIDA states and U.S. territories that participated in the operational testing program of ACCESS Paper in 2019–2020 by grade level. The rows provide data for the number of students in that grade-level cluster who took the test by state, with the final row showing the total number of participants across all 40 states and territories. Some states' sample sizes are small except for Kindergarten, which is only in Paper form, since most students take the Online form of the tests. The biggest state was Florida, which constitutes about 53% of the students who take Paper ACCESS. Illinois, Georgia, and South Carolina were the next largest states. The full names of acronyms of U.S. territories are the following: BI, Bureau of Indian Education; DC, District of Columbia; DD, Department of Defense Education Activity; MP, Northern Mariana Islands; and VI, U.S. Virgin Islands.

Table 1.1.1.1
Participation by Grade-Level Cluster by State, S501 Paper

State	Cluster							Total
	K	1	2	3	4–5	6–8	9–12	
AK	1,107	9	10	13	29	47	144	1,359
AL	3,402	2	2	6	10	4	1	3,427
BI	608	256	246	255	516	581	317	2,779
CO	9,814	418	367	330	509	473	254	12,165
DC	780	1	1	1	4	0	0	787
DD	693	868	874	814	1,282	1,092	603	6,226
DE	1,554	1	3	7	8	2	1	1,576
FL	32,172	33,234	33,108	30,625	44,264	46,685	44,881	264,969
GA	15,424	1,875	1,833	1,770	52	24	30	21,008
HI	1,767	5	1	1	5	4	2	1,785
ID	1,937	4	7	3	8	8	8	1,975
IL	24,204	296	290	292	480	585	259	26,406
IN	7,877	33	31	21	30	31	16	8,039
KY	4,121	11	5	4	5	5	5	4,156
MA	11,281	78	78	61	136	66	95	11,795
MD	10,426	9	9	13	31	23	19	10,530
ME	496	2	2	2	3	12	8	525
MI	8,896	118	112	128	204	235	298	9,991
MN	8,147	42	73	52	76	94	53	8,537
MO	4,317	8	9	7	17	9	2	4,369
MP	50	0	0	0	0	0	0	50
MT	295	0	0	0	0	0	0	295
NC	11,376	11	18	24	40	17	11	11,497
ND	413	0	0	0	0	0	1	414
NH	449	35	33	30	39	42	29	657
NJ	7,435	68	51	22	13	8	13	7,610
NM	3,705	1	0	2	3	4	41	3,756
NV	6,443	0	0	0	1	2	23	6,469
OK	6,178	43	37	31	70	80	20	6,459
PA	5,749	428	379	296	556	503	508	8,419
RI	1,452	0	4	3	3	6	9	1,477
SC	3,479	1,382	1,504	1,610	3,148	3,898	4,415	19,436
SD	846	28	25	34	48	26	0	1,007
TN	5,441	0	1	1	2	1	3	5,449
UT	4,258	1	0	1	0	0	1	4,261
VA	14,210	1,440	660	630	752	77	116	17,885
VI	48	36	29	22	15	91	0	241
VT	165	2	4	4	2	3	1	181
WI	4,922	28	27	23	37	34	18	5,089
WY	275	1	3	6	6	10	8	309
Total	226,212	40,774	39,836	37,144	52,404	54,782	52,213	503,365

Table 1.1.1.2 shows participation by grade-level cluster and by gender across all states and territories for the population of students who participated in ACCESS Paper, while Table 1.1.1.3

shows participation by grade-level cluster and by ethnicity. The gender ratio was 46% female and 51% male in Clusters 1–3 and 44% female and 52% male in clusters 4–12. The Hispanic ethnicity percentage was about 76% in all clusters except Kindergarten, which was 64%.

Table 1.1.1.2
Participation by Grade-Level Cluster by Gender, S501 Paper

Cluster		Gender			Total
		F	M	Missing	
K	Count	102,872	117,005	6,335	226,212
	% within Cluster	45.5%	51.7%	2.8%	100.0%
1	Count	18,203	21,032	1,539	40,774
	% within Cluster	44.6%	51.6%	3.8%	100.0%
2	Count	18,165	20,575	1,096	39,836
	% within Cluster	45.6%	51.6%	2.8%	100.0%
3	Count	16,318	19,719	1,107	37,144
	% within Cluster	43.9%	53.1%	3.0%	100.0%
4–5	Count	23,145	27,413	1,846	52,404
	% within Cluster	44.2%	52.3%	3.5%	100.0%
6–8	Count	23,752	28,659	2,371	54,782
	% within Cluster	43.4%	52.3%	4.3%	100.0%
9–12	Count	22,854	26,736	2,623	52,213
	% within Cluster	43.8%	51.2%	5.0%	100.0%
Total	Count	225,309	261,139	16,917	503,365
	% within Cluster	44.8%	51.9%	3.4%	100.0%

Table 1.1.1.3
Participation by Grade-Level Cluster by Ethnicity, S501 Paper

Cluster		Hispanic/Non-Hispanic			Total
		Hispanic	Other	Unknown	
K	Count	144,175	67,993	14,044	226,212
	% within Cluster	63.7%	30.1%	6.2%	100.0%
1	Count	30,813	8,513	1,448	40,774
	% within Cluster	75.6%	20.9%	3.6%	100.0%
2	Count	30,284	8,156	1,396	39,836
	% within Cluster	76.0%	20.5%	3.5%	100.0%
3	Count	28,412	7,409	1,323	37,144
	% within Cluster	76.5%	19.9%	3.6%	100.0%
4–5	Count	40,009	10,165	2,230	52,404
	% within Cluster	76.3%	19.4%	4.3%	100.0%
6–8	Count	42,224	10,417	2,141	54,782
	% within Cluster	77.1%	19.0%	3.9%	100.0%
9–12	Count	39,312	11,238	1,663	52,213
	% within Cluster	75.3%	21.5%	3.2%	100.0%

Total	Count	355,229	123,891	24,245	503,365
	% within Cluster	70.6	24.6%	4.8%	100.0%

1.1.2 Grade

This section provides data similar to that in the previous section, but broken out by grade rather than by grade-level cluster. As shown in Table 1.1.2.1, the largest grade was Kindergarten, which comprised almost 45% of the Paper ACCESS population.

Table 1.1.2.1
Participation by Grade by State, S501 Paper

State	Grade													Total
	K	1	2	3	4	5	6	7	8	9	10	11	12	
AK	1,107	9	10	13	8	21	14	15	18	36	39	38	31	1,359
AL	3,402	2	2	6	4	6	1	3	0	0	0	1	0	3,427
BI	608	256	246	255	256	260	227	182	172	101	78	66	72	2,779
CO	9,814	418	367	330	283	226	161	145	167	107	57	46	44	12,165
DC	780	1	1	1	3	1	0	0	0	0	0	0	0	787
DD	693	868	874	814	672	610	412	395	285	212	178	131	82	6,226
DE	1,554	1	3	7	6	2	1	0	1	1	0	0	0	1,576
FL	32,172	33,234	33,108	30,625	23,287	20,977	17,482	14,836	14,367	13,960	12,713	11,056	7,152	264,969
GA	15,424	1,875	1,833	1,770	25	27	11	7	6	11	6	8	5	21,008
HI	1,767	5	1	1	3	2	0	3	1	0	0	2	0	1,785
ID	1,937	4	7	3	6	2	6	0	2	7	1	0	0	1,975
IL	24,204	296	290	292	255	225	222	195	168	73	73	55	58	26,406
IN	7,877	33	31	21	18	12	13	14	4	4	7	2	3	8,039
KY	4,121	11	5	4	4	1	3	1	1	3	2	0	0	4,156
MA	11,281	78	78	61	78	58	29	24	13	36	21	28	10	11,795
MD	10,426	9	9	13	18	13	11	6	6	9	6	3	1	10,530
ME	496	2	2	2	0	3	3	4	5	0	2	4	2	525
MI	8,896	118	112	128	108	96	68	76	91	84	77	84	53	9,991
MN	8,147	42	73	52	44	32	29	28	37	20	10	13	10	8,537
MO	4,317	8	9	7	4	13	6	2	1	0	1	1	0	4,369
MP	50	0	0	0	0	0	0	0	0	0	0	0	0	50
MT	295	0	0	0	0	0	0	0	0	0	0	0	0	295
NC	11,376	11	18	24	24	16	7	6	4	3	3	3	2	11,497
ND	413	0	0	0	0	0	0	0	0	0	1	0	0	414
NH	449	35	33	30	23	16	18	14	10	11	5	11	2	657
NJ	7,435	68	51	22	10	3	4	2	2	5	3	2	3	7,610
NM	3,705	1	0	2	2	1	0	2	2	20	9	11	1	3,756
NV	6,443	0	0	0	1	0	0	0	2	2	6	10	5	6,469
OK	6,178	43	37	31	35	35	43	21	16	8	8	3	1	6,459
PA	5,749	428	379	296	321	235	174	175	154	170	143	117	78	8,419
RI	1,452	0	4	3	0	3	2	2	2	4	1	1	3	1,477
SC	3,479	1,382	1,504	1,610	1,740	1,408	1,312	1,353	1,233	1,632	1,124	912	747	19,436
SD	846	28	25	34	36	12	17	8	1	0	0	0	0	1,007
TN	5,441	0	1	1	1	1	0	1	0	1	0	1	1	5,449
UT	4,258	1	0	1	0	0	0	0	0	0	0	1	0	4,261
VA	14,210	1,440	660	630	520	232	29	23	25	43	31	26	16	17,885
VI	48	36	29	22	9	6	33	30	28	0	0	0	0	241
VT	165	2	4	4	1	1	3	0	0	0	0	1	0	181
WI	4,922	28	27	23	18	19	14	8	12	5	5	3	5	5,089

WY	275	1	3	6	3	3	3	2	5	2	1	1	4	309
Total	226,212	40,774	39,836	37,144	27,826	24,578	20,358	17,583	16,841	16,570	14,611	12,641	8,391	503,365

Table 1.1.2.2
Participation by Grade by Gender, S501 Paper

Grade		Gender			Total
		F	M	Missing	
K	Count	102,872	117,005	6,335	226,212
	% within Grade	45.5%	51.7%	2.8%	100.0%
1	Count	18,203	21,032	1,539	40,774
	% within Grade	44.6%	51.6%	3.8%	100.0%
2	Count	18,165	20,575	1,096	39,836
	% within Grade	45.6%	51.6%	2.8%	100.0%
3	Count	16,318	19,719	1,107	37,144
	% within Grade	43.9%	53.1%	3.0%	100.0%
4	Count	12,395	14,458	973	27,826
	% within Grade	44.5%	52.0%	3.5%	100.0%
5	Count	10,750	12,955	873	24,578
	% within Grade	43.7%	52.7%	3.6%	100.0%
6	Count	8,831	10,673	854	20,358
	% within Grade	43.4%	52.4%	4.2%	100.0%
7	Count	7,676	9,146	761	17,583
	% within Grade	43.7%	52.0%	4.3%	100.0%
8	Count	7,245	8,840	756	16,841
	% within Grade	43.0%	52.5%	4.5%	100.0%
9	Count	7,082	8,751	737	16,570
	% within Grade	42.7%	52.8%	4.4%	100.0%
10	Count	6,410	7,490	711	14,611
	% within Grade	43.9%	51.3%	4.9%	100.0%
11	Count	5,526	6,407	708	12,641
	% within Grade	43.7%	50.7%	5.6%	100.0%
12	Count	3,836	4,088	467	8,391
	% within Grade	45.7%	48.7%	5.6%	100.0%
Total	Count	225,309	261,139	16,917	503,365
	% within Grade	44.8%	51.9%	3.4%	100.0%

Table 1.1.2.3
Participation by Grade by Ethnicity, S501 Paper

Grade		Hispanic/Non-Hispanic			Total
		Hispanic	Other	Unknown	
K	Count	144,175	67,993	14,044	226,212
	% within Grade	63.7%	30.1%	6.2%	100.0%
1	Count	30,813	8,513	1,448	40,774
	% within Grade	75.6%	20.9%	3.6%	100.0%
2	Count	30,284	8,156	1,396	39,836
	% within Grade	76.0%	20.5%	3.5%	100.0%
3	Count	28,412	7,409	1,323	37,144
	% within Grade	76.5%	19.9%	3.6%	100.0%
4	Count	21,202	5,412	1,212	27,826
	% within Grade	76.2%	19.4%	4.4%	100.0%
5	Count	18,807	4,753	1,018	24,578
	% within Grade	76.5%	19.3%	4.1%	100.0%
6	Count	15,687	3,834	837	20,358
	% within Grade	77.1%	18.8%	4.1%	100.0%
7	Count	13,554	3,323	706	17,583
	% within Grade	77.1%	18.9%	4.0%	100.0%
8	Count	12,983	3,260	598	16,841
	% within Grade	77.1%	19.4%	3.6%	100.0%
9	Count	12,682	3,338	550	16,570
	% within Grade	76.5%	20.1%	3.3%	100.0%
10	Count	11,212	2,945	454	14,611
	% within Grade	76.7%	20.2%	3.1%	100.0%
11	Count	9,427	2,833	381	12,641
	% within Grade	74.6%	22.4%	3.0%	100.0%
12	Count	5,991	2,122	278	8,391
	% within Grade	71.4%	25.3%	3.3%	100.0%
Total	Count	355,229	123,891	24,245	503,365
	% within Grade	70.6%	24.6%	4.8%	100.0%

1.1.3 Tier

This section provides information on participation by tier. The tables show this information in several ways:

- By grade-level cluster, tier, and domain
- By grade, tier, and domain
- By grade-level cluster and tier for gender
- By grade-level cluster and tier for ethnicity

Table 1.1.3.1 shows the number of students in each tier per cluster. In Grade 1, 49% of students were in Tier A and 51% in Tier B/C. In Grade 2, 23% of students were in Tier A and 77% in Tier B/C. In Grade 3 and Grades 4–5, 20% were in Tier A and 80% in Tier B/C. In Grades 6–8 and 9–12, there were about 25% of students in Tier A and 75% in Tier B/C. In all domains these percentages remained the same since students were placed in one tier throughout the test.

Table 1.1.3.1
Participation by Grade-Level Cluster by Tier by Domain, S501 Paper

Grade			Domain			
			Listening	**Reading**	**Speaking**	**Writing**
K	Tier	-	226,204	226,198	226,203	226,197
1	Tier	A	19,974	19,973	19,974	19,973
		B	20,788	20,793	20,783	20,791
	Total		40,762	40,766	40,757	40,764
2	Tier	A	9,131	9,131	9,130	9,130
		B	30,692	30,704	30,702	30,704
	Total		39,823	39,835	39,832	39,834
3	Tier	A	7,466	7,468	7,468	7,468
		B	29,673	29,672	29,674	29,675
	Total		37,139	37,140	37,142	37,143
4–5	Tier	A	10,479	10,479	10,478	10,479
		B	41,925	41,923	41,925	41,923
	Total		52,404	52,402	52,403	52,402
6–8	Tier	A	13,519	13,520	13,519	13,519
		B	41,260	41,258	41,248	41,257
	Total		54,779	54,778	54,767	54,776
9–12	Tier	A	13,605	13,605	13,605	13,600
		B	38,602	38,603	38,604	38,600
	Total		52,207	52,208	52,209	52,200

Table 1.1.3.2
Participation by Grade by Tier by Domain, S501 Paper

Grade			Domain			
			Listening	**Reading**	**Speaking**	**Writing**
K	Tier	-	226,204	226,198	226,203	226,197
1	Tier	A	19,974	19,973	19,974	19,973
		B	20,788	20,793	20,783	20,791
	Total		40,762	40,766	40,757	40,764
2	Tier	A	9,131	9,131	9,130	9,130
		B	30,692	30,704	30,702	30,704
	Total		39,823	39,835	39,832	39,834
3	Tier	A	7,466	7,468	7,468	7,468
		B	29,673	29,672	29,674	29,675
	Total		37,139	37,140	37,142	37,143
4		A	5,499	5,499	5,498	5,499
		B	22,327	22,326	22,327	22,325
	Total		27,826	27,825	27,825	27,824
5	Tier	A	4,980	4,980	4,980	4,980
		B	19,598	19,597	19,598	19,598
	Total		24,578	24,577	24,578	24,578
6	Tier	A	4,612	4,612	4,612	4,611
		B	15,746	15,745	15,737	15,744
	Total		20,358	20,357	20,349	20,355
7	Tier	A	4,399	4,400	4,399	4,400
		B	13,183	13,181	13,178	13,182
	Total		17,582	17,581	17,577	17,582
8	Tier	A	4,508	4,508	4,508	4,508
		B	12,331	12,332	12,333	12,331
	Total		16,839	16,840	16,841	16,839
9	Tier	A	5,314	5,314	5,314	5,310
		B	11,255	11,255	11,255	11,252
	Total		16,569	16,569	16,569	16,562
10	Tier	A	3,818	3,818	3,818	3,818
		B	10,792	10,792	10,793	10,792
	Total		14,610	14,610	14,611	14,610
11	Tier	A	2,893	2,893	2,893	2,893
		B	9,745	9,746	9,746	9,746
	Total		12,638	12,639	12,639	12,639
12	Tier	A	1,580	1,580	1,580	1,579
		B	6,810	6,810	6,810	6,810
	Total		8,390	8,390	8,390	8,389

Table 1.1.3.3
Participation by Grade-Level Cluster by Tier by Gender

Cluster	Tier		Gender			Total
			F	M	Missing	
K	-	Count	102,872	117,005	6,335	226,212
		% within Tier	45.5%	51.7%	2.8%	100.0%
1	A	Count	8,517	10,536	926	19,979
		% within Tier	42.6%	52.7%	4.6%	100.0%
	BC	Count	9,686	10,496	613	20,795
		% within Tier	46.6%	50.5%	2.9%	100.0%
2	A	Count	3,929	4,902	300	9,131
		% within Tier	43.0%	53.7%	3.3%	100.0%
	BC	Count	14,236	15,673	796	30,705
		% within Tier	46.4%	51.0%	2.6%	100.0%
3	A	Count	3,162	4,056	250	7,468
		% within Tier	42.3%	54.3%	3.3%	100.0%
	BC	Count	13,156	15,663	857	29,676
		% within Tier	44.3%	52.8%	2.9%	100.0%
4–5	A	Count	4,676	5,444	359	10,479
		% within Tier	44.6%	52.0%	3.4%	100.0%
	BC	Count	18,469	21,969	1,487	41,925
		% within Tier	44.1%	52.4%	3.5%	100.0%
6–8	A	Count	5,815	7,173	532	13,520
		% within Tier	43.0%	53.1%	3.9%	100.0%
	BC	Count	17,937	21,486	1,839	41,262
		% within Tier	43.5%	52.1%	4.5%	100.0%
9–12	A	Count	5,978	6,817	811	13,606
		% within Tier	43.9%	50.1%	6.0%	100.0%
	BC	Count	16,876	19,919	1,812	38,607
		% within Tier	43.7%	51.6%	4.7%	100.0%

Table 1.1.3.4 presents percentages of Hispanic and other ethnic groups in tiers. Overall, the percentages of Hispanic students in Tier A were 4% to 5% higher than in Tier B/C except in Grades 2 and 3.

Table 1.1.3.4
Participation by Grade-Level Cluster by Tier by Ethnicity

Cluster	Tier		Ethnicity			Total
			Hispanic	Other	Unknown	
K	-	Count	144,175	67,993	14,044	226,212
		% within Tier	63.7%	30.1%	6.2%	100.0%
1	A	Count	15,504	3,857	618	19,979
		% within Tier	77.6%	19.3%	3.1%	100.0%
	BC	Count	15,309	4,656	830	20,795
		% within Tier	73.6%	22.4%	4.0%	100.0%
2	A	Count	6,951	1,729	451	9,131
		% within Tier	76.1%	18.9%	4.9%	100.0%
	BC	Count	23,333	6,427	945	30,705
		% within Tier	76.0%	20.9%	3.1%	100.0%
3	A	Count	5,774	1,342	352	7,468
		% within Tier	77.3%	18.0%	4.7%	100.0%
	BC	Count	22,638	6,067	971	29,676
		% within Tier	76.3%	20.4%	3.3%	100.0%
4–5	A	Count	8,193	1,765	521	10,479
		% within Tier	78.2%	16.8%	5.0%	100.0%
	BC	Count	31,816	8,400	1,709	41,925
		% within Tier	75.9%	20.0%	4.1%	100.0%
6–8	A	Count	10,895	2,155	470	13,520
		% within Tier	80.6%	15.9%	3.5%	100.0%
	BC	Count	31,329	8,262	1,671	41,262
		% within Tier	75.9%	20.0%	4.0%	100.0%
9–12	A	Count	10,695	2,552	359	13,606
		% within Tier	78.6%	18.8%	2.6%	100.0%
	BC	Count	28,617	8,686	1,304	38,607
		% within Tier	74.1%	22.5%	3.4%	100.0%

1.2 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 on ACCESS, first for the four domains (Listening, Speaking, Reading, and Writing) and then for the four composites (Oral Language, Literacy, Comprehension, and Overall). 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. Tables are provided for the total student population, for the student population by gender, and for the student population by race and ethnicity. In Table 1.2.1.1, the order of average scale scores

among single domains in descending order were Listening, Reading, Speaking, and then Writing in clusters of 1, 2–3, 4–5, and 6–8. Kindergarten had the average scale scores in the order of Speaking, Listening, Writing, and then Reading. Cluster 9–12 had the order of Listening, Reading, Writing, and then Speaking. Cluster 6–8 and 9–12 showed the highest average scale scores in all single domains across all clusters.

Table 1.2.1.1
Mean Scale Scores by Grade-Level Cluster, S501 Paper

Cluster		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
K	Mean	264.50	183.84	199.83	268.65	266.84	192.08	208.03	214.28
	N	226,001	225,994	225,987	226,000	225,997	225,982	225,990	225,978
1	Mean	307.91	289.40	254.50	271.48	291.29	272.95	295.64	279.11
	N	34,887	31,223	40,744	40,426	34,612	31,208	27,978	27,755
2	Mean	332.52	316.46	285.26	291.61	313.15	301.95	321.89	305.87
	N	37,084	33,414	39,818	39,528	36,808	33,402	31,795	31,553
3	Mean	358.21	337.52	300.11	303.34	331.56	319.47	344.08	323.40
	N	34,834	31,495	37,133	36,848	34,575	31,490	30,088	29,864
4–5	Mean	379.19	352.42	331.14	331.17	355.99	342.44	360.85	346.76
	N	50,331	46,614	52,398	51,988	49,939	46,610	45,302	44,957
6–8	Mean	387.84	357.59	332.18	352.13	371.08	345.29	366.90	353.13
	N	52,283	47,103	54,744	54,257	51,805	47,080	45,586	45,171
9–12	Mean	384.78	382.98	359.62	349.43	368.33	371.99	384.00	371.35
	N	49,373	45,822	52,154	51,590	48,812	45,784	44,041	43,514

Table 1.2.1.2 demonstrated that female groups performed higher than male groups in general.

Table 1.2.1.2
Mean Scale Scores by Grade-Level Cluster by Gender, S501 Paper

Cluster	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
K	F	Mean	270.61	185.64	204.98	278.02	274.58	195.56	211.12	219.04
		N	102,782	102,780	102,780	102,783	102,782	102,779	102,779	102,778
	M	Mean	260.00	183.04	196.31	261.67	261.10	189.91	206.12	211.05
		N	116,888	116,883	116,876	116,886	116,884	116,872	116,880	116,869
	Missing	Mean	248.24	169.41	181.25	245.38	247.06	175.54	193.05	196.79
		N	6,331	6,331	6,331	6,331	6,331	6,331	6,331	6,331
1	F	Mean	309.99	290.73	260.12	274.00	293.54	276.29	297.11	282.02
		N	15,849	14,054	18,193	18,040	15,722	14,050	12,724	12,622
	M	Mean	306.53	288.65	250.20	270.18	289.99	270.50	294.76	277.06
		N	17,740	16,017	21,012	20,853	17,598	16,006	14,226	14,110
	Missing	Mean	301.39	283.43	246.83	259.51	281.59	266.18	289.52	271.46
		N	1,298	1,152	1,539	1,533	1,292	1,152	1,028	1,023
2	F	Mean	334.63	318.32	292.36	295.51	315.95	306.27	323.74	309.63
		N	17,101	15,470	18,159	18,031	16,981	15,466	14,828	14,722
	M	Mean	331.23	315.21	279.87	289.28	311.46	298.70	320.66	303.11
		N	19,001	17,094	20,564	20,413	18,854	17,086	16,180	16,049
	Missing	Mean	320.87	307.84	268.80	270.52	297.00	288.89	312.30	291.85
		N	982	850	1,095	1,084	973	850	787	782
3	F	Mean	358.30	338.19	307.35	306.22	332.93	323.23	344.50	326.27
		N	15,486	14,083	16,317	16,189	15,370	14,083	13,541	13,443
	M	Mean	358.65	337.21	294.94	301.77	331.06	316.83	344.04	321.50
		N	18,347	16,503	19,709	19,564	18,215	16,498	15,705	15,589
	Missing	Mean	348.65	332.90	285.59	288.68	319.33	309.20	337.99	312.57
		N	1,001	909	1,107	1,095	990	909	842	832
4–5	F	Mean	378.43	352.68	336.85	331.45	355.64	345.32	360.76	348.59
		N	22,289	20,805	23,142	22,965	22,120	20,804	20,254	20,104
	M	Mean	380.28	352.38	326.90	331.71	356.90	340.36	361.20	345.64
		N	26,274	24,194	27,410	27,186	26,059	24,191	23,483	23,296
	Missing	Mean	372.57	349.49	322.54	319.75	347.03	336.47	356.84	339.92
		N	1,768	1,615	1,846	1,837	1,760	1,615	1,565	1,557
6–8	F	Mean	389.40	359.64	338.34	351.59	371.67	349.26	368.77	356.07
		N	22,769	20,628	23,738	23,515	22,554	20,620	20,024	19,833
	M	Mean	386.73	355.96	327.35	353.02	370.94	342.10	365.42	350.85
		N	27,276	24,379	28,635	28,403	27,043	24,364	23,546	23,348
	Missing	Mean	385.44	356.26	328.72	346.69	366.74	343.30	365.49	350.50
		N	2,238	2,096	2,371	2,339	2,208	2,096	2,016	1,990
9–12	F	Mean	384.66	385.13	364.81	348.73	367.80	375.56	385.41	373.59
		N	21,826	20,418	22,829	22,612	21,602	20,403	19,745	19,529
	M	Mean	385.74	381.95	356.42	351.66	369.87	369.92	383.60	370.37
		N	25,147	23,181	26,702	26,434	24,881	23,158	22,204	21,957
	Missing	Mean	375.86	373.99	347.00	332.57	356.75	360.73	375.04	360.54
		N	2,400	2,223	2,623	2,544	2,329	2,223	2,092	2,028

Table 1.2.1.3 presents scale score performance by ethnic groups. The top three performing ethnic groups were Asian students, White students, and multiracial students.

Table 1.2.1.3
Mean Scale Scores by Grade-Level Cluster by Ethnicity, S501 Paper

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
K	Non-Hispanic Asian	Mean	284.38	218.55	231.98	287.56	286.23	225.53	238.29	243.50
		N	28,528	28,525	28,526	28,527	28,527	28,524	28,525	28,524
	Non-Hispanic Pacific Islander	Mean	259.69	165.04	185.81	266.63	263.41	175.65	193.43	201.76
		N	1,451	1,451	1,451	1,451	1,451	1,451	1,451	1,451
	Non-Hispanic Black	Mean	273.01	197.54	208.10	287.61	280.57	203.06	220.17	226.10
		N	13,038	13,037	13,036	13,039	13,038	13,036	13,036	13,035
	Hispanic (Of Any Race)	Mean	258.59	175.11	191.60	262.16	260.63	183.59	200.14	206.48
		N	144,100	144,098	144,094	144,099	144,098	144,091	144,096	144,089
	Non-Hispanic American Indian	Mean	269.11	173.77	187.29	267.25	268.46	180.76	202.36	206.85
		N	1,879	1,880	1,880	1,880	1,879	1,880	1,879	1,879
	Non-Hispanic Multiracial	Mean	286.10	207.33	218.02	297.43	292.03	212.93	230.94	236.44
		N	1,209	1,209	1,209	1,209	1,209	1,209	1,209	1,209
	Non-Hispanic White	Mean	280.11	196.59	215.84	287.86	284.25	206.46	221.63	229.57
		N	21,803	21,802	21,801	21,803	21,803	21,801	21,802	21,801
	Unknown	Mean	250.61	171.68	187.97	247.30	249.20	180.05	195.35	200.59
		N	13,993	13,992	13,990	13,992	13,992	13,990	13,992	13,990
1	Non-Hispanic Asian	Mean	311.03	301.76	267.61	281.29	297.80	285.59	305.15	289.92
		N	1,807	1,642	2,048	2,041	1,800	1,640	1,492	1,484
	Non-Hispanic Pacific Islander	Mean	301.82	289.59	264.54	282.96	296.22	277.98	292.50	284.05
		N	45	41	50	50	45	41	38	38
	Non-Hispanic Black	Mean	304.87	288.83	249.89	278.56	293.08	271.10	294.41	278.76
		N	2,487	2,223	3,062	3,026	2,460	2,223	1,913	1,888
	Hispanic (Of Any Race)	Mean	307.27	287.65	253.11	268.26	289.36	271.26	294.20	277.29
		N	26,413	23,566	30,790	30,558	26,204	23,555	21,154	20,991
	Non-Hispanic American Indian	Mean	304.49	290.22	259.04	273.09	291.46	275.17	294.63	279.84
		N	221	203	269	265	217	203	175	173
	Non-Hispanic Multiracial	Mean	312.91	301.89	263.04	288.37	303.29	284.14	306.38	291.23
		N	258	236	295	295	258	235	215	214
	Non-Hispanic White	Mean	312.35	294.24	260.18	282.98	299.19	278.28	300.40	285.20
		N	2,380	2,174	2,783	2,756	2,360	2,173	1,952	1,934
	Unknown	Mean	314.23	296.83	261.61	284.86	301.12	280.24	302.69	286.89
		N	1,276	1,138	1,447	1,435	1,268	1,138	1,039	1,033

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
2	Non-Hispanic Asian	Mean	335.48	327.37	298.55	299.92	318.21	313.49	330.19	315.20
		N	1,799	1,641	1,913	1,905	1,792	1,641	1,575	1,570
	Non-Hispanic Pacific Islander	Mean	327.94	318.37	291.71	292.02	310.25	305.54	322.11	308.13
		N	52	41	56	56	52	41	38	38
	Non-Hispanic Black	Mean	330.40	313.25	281.54	293.68	313.22	298.82	319.08	303.72
		N	2,895	2,547	3,163	3,140	2,874	2,545	2,374	2,356
	Hispanic (Of Any Race)	Mean	331.87	315.12	283.89	289.59	311.84	300.61	320.74	304.55
		N	28,240	25,397	30,269	30,047	28,024	25,387	24,221	24,030
	Non-Hispanic American	Mean	333.33	315.26	292.18	283.08	309.75	304.38	320.86	306.90
		N	233	210	261	258	232	210	198	197
	Non-Hispanic Multiracial	Mean	340.29	327.23	297.32	310.01	325.43	312.91	331.57	316.14
		N	219	202	234	232	218	202	192	191
	Non-Hispanic White	Mean	338.75	323.10	292.01	303.13	322.10	308.19	328.54	313.13
		N	2,331	2,162	2,526	2,508	2,314	2,162	2,033	2,020
	Unknown	Mean	334.94	323.08	289.47	297.01	316.97	307.43	327.26	310.53
		N	1,315	1,214	1,396	1,382	1,302	1,214	1,164	1,151
3	Non-Hispanic Asian	Mean	367.81	345.27	311.56	311.00	339.84	328.83	352.52	332.44
		N	1,457	1,339	1,548	1,538	1,449	1,339	1,284	1,276
	Non-Hispanic Pacific Islander	Mean	354.84	335.95	310.40	312.74	335.07	323.63	340.76	325.16
		N	43	41	53	53	43	41	37	37
	Non-Hispanic Black	Mean	356.48	335.62	296.31	305.20	331.56	316.82	342.07	321.58
		N	2,971	2,663	3,213	3,173	2,935	2,662	2,521	2,490
	Hispanic (Of Any Race)	Mean	357.01	336.71	299.00	301.27	329.93	318.47	343.14	322.18
		N	26,709	24,058	28,404	28,202	26,525	24,055	23,023	22,866
	Non-Hispanic American Indian	Mean	354.17	336.28	309.30	295.65	324.73	322.50	341.72	322.81
		N	223	193	231	231	223	193	188	188
	Non-Hispanic Multiracial	Mean	360.15	338.14	303.69	313.66	338.53	321.20	344.43	326.11
		N	175	174	194	194	175	174	160	160
	Non-Hispanic White	Mean	364.61	341.98	305.81	315.60	341.06	324.56	349.05	329.85
		N	2,025	1,882	2,167	2,149	2,008	1,881	1,796	1,780
	Unknown	Mean	366.93	342.77	308.05	313.57	341.54	326.47	350.89	331.70
		N	1,231	1,145	1,323	1,308	1,217	1,145	1,079	1,067

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
4–5	Non-Hispanic Asian	Mean	386.34	361.10	337.80	340.27	364.30	350.18	369.14	354.62
		N	1,644	1,547	1,724	1,714	1,635	1,547	1,499	1,492
	Non-Hispanic Pacific Islander	Mean	382.16	358.45	340.92	343.42	364.73	353.21	366.85	357.88
		N	63	53	65	64	62	53	52	51
	Non-Hispanic Black	Mean	380.24	351.76	329.84	337.60	359.40	341.60	360.72	347.30
		N	4,260	3,873	4,493	4,450	4,219	3,873	3,726	3,693
	Hispanic (Of Any Race)	Mean	377.84	351.21	330.24	328.43	353.97	341.34	359.57	345.33
		N	38,539	35,644	40,004	39,697	38,241	35,640	34,705	34,442
	Non-Hispanic American Indian	Mean	378.22	352.04	334.11	333.27	356.03	344.14	360.06	347.86
		N	425	403	438	433	420	403	391	387
	Non-Hispanic Multiracial	Mean	387.52	359.72	337.81	348.23	369.20	349.36	368.56	355.83
		N	256	237	268	264	252	237	232	228
	Non-Hispanic White	Mean	385.74	357.72	335.37	342.88	365.56	347.19	366.75	353.34
		N	3,009	2,829	3,176	3,154	2,991	2,829	2,723	2,705
	Unknown	Mean	385.71	359.89	337.08	341.01	364.04	349.41	368.41	354.30
		N	2,135	2,028	2,230	2,212	2,119	2,028	1,974	1,959
6–8	Non-Hispanic Asian	Mean	396.43	366.55	340.14	364.95	381.62	353.90	375.95	362.46
		N	1,662	1,509	1,729	1,719	1,653	1,509	1,471	1,464
	Non-Hispanic Pacific Islander	Mean	386.32	360.14	335.55	357.38	371.95	348.36	368.06	355.64
		N	94	87	100	99	93	86	81	80
	Non-Hispanic Black	Mean	390.46	357.05	332.17	360.99	376.48	345.24	367.23	354.67
		N	4,399	3,764	4,670	4,566	4,303	3,759	3,620	3,550
	Hispanic (Of Any Race)	Mean	385.91	356.15	330.81	348.39	368.30	343.82	365.29	351.24
		N	40,384	36,471	42,199	41,866	40,050	36,457	35,324	35,024
	Non-Hispanic American Indian	Mean	390.26	359.77	336.25	355.80	374.46	348.24	369.64	356.76
		N	382	334	395	392	379	334	323	320
	Non-Hispanic Multiracial	Mean	400.09	366.71	340.33	368.07	385.53	353.89	376.82	363.30
		N	280	253	297	297	280	253	244	244
	Non-Hispanic White	Mean	396.91	364.76	338.67	370.35	384.86	352.35	374.89	362.48
		N	3,051	2,799	3,216	3,193	3,031	2,798	2,705	2,686
	Unknown	Mean	397.80	366.83	340.92	365.70	382.70	354.53	376.45	363.04
		N	2,031	1,886	2,138	2,125	2,016	1,884	1,818	1,803

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
9–12	Non-Hispanic Asian	Mean	397.50	395.39	374.05	366.35	383.11	385.64	396.64	385.59
		N	1,915	1,778	1,996	1,971	1,889	1,777	1,735	1,710
	Non-Hispanic Pacific Islander	Mean	395.66	387.93	370.57	363.90	381.01	380.28	390.65	380.53
		N	90	90	101	100	89	90	82	81
	Non-Hispanic Black	Mean	379.31	378.20	351.17	348.55	364.81	365.53	378.72	365.43
		N	5,000	4,341	5,412	5,367	4,955	4,336	4,103	4,060
	Hispanic (Of Any Race)	Mean	383.30	382.00	358.98	346.92	366.37	371.05	382.83	370.10
		N	37,265	34,788	39,266	38,883	36,882	34,759	33,471	33,114
	Non-Hispanic American	Mean	390.45	382.44	363.13	346.12	369.04	372.73	385.21	371.81
		N	305	273	322	312	298	273	263	256
	Non-Hispanic Multiracial	Mean	398.67	389.44	371.34	367.47	383.52	380.60	392.25	381.34
		N	256	235	269	267	255	235	230	229
	Non-Hispanic White	Mean	395.29	390.36	365.82	363.58	380.91	379.01	392.53	380.20
		N	2,968	2,820	3,127	3,093	2,933	2,817	2,716	2,681
	Unknown	Mean	398.09	389.85	370.03	362.00	381.42	381.00	393.19	381.67
		N	1,574	1,497	1,661	1,597	1,511	1,497	1,441	1,383

1.2.2 Mean Scale Score Across Domain and Composite Score by Grade

This section shows the mean scale scores broken down by grade rather than by grade-level cluster. Tables are provided for the total student population, for the student population by gender, and for the student population by race and ethnicity. Table 1.2.2.1 shows increment of scale scores by grade. Listening domain peaked at Grade 8. Reading and Writing domains had the highest mean scale scores in Grade 11. Speaking had the highest mean scale score in Grade 12.

Table 1.2.2.1
Mean Scale Scores by Grade, S501 Paper

Grade		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
K	Mean	264.50	183.84	199.83	268.65	266.84	192.08	208.03	214.28
	N	226,001	225,994	225,987	226,000	225,997	225,982	225,990	225,978
1	Mean	307.91	289.40	254.50	271.48	291.29	272.95	295.64	279.11
	N	34,887	31,223	40,744	40,426	34,612	31,208	27,978	27,755
2	Mean	332.52	316.46	285.26	291.61	313.15	301.95	321.89	305.87
	N	37,084	33,414	39,818	39,528	36,808	33,402	31,795	31,553
3	Mean	358.21	337.52	300.11	303.34	331.56	319.47	344.08	323.40
	N	34,834	31,495	37,133	36,848	34,575	31,490	30,088	29,864
4	Mean	374.91	348.65	326.59	328.53	352.53	338.24	356.91	342.76
	N	26,690	24,463	27,821	27,616	26,494	24,459	23,751	23,575
5	Mean	384.02	356.57	336.29	334.17	359.91	347.07	365.20	351.17
	N	23,641	22,151	24,577	24,372	23,445	22,151	21,551	21,382
6	Mean	384.02	353.07	329.82	352.80	369.43	341.95	362.59	350.23
	N	19,450	17,344	20,351	20,150	19,263	17,340	16,793	16,635
7	Mean	388.34	357.44	332.30	351.14	370.66	345.21	366.91	352.86
	N	16,854	15,236	17,561	17,429	16,715	15,220	14,782	14,650
8	Mean	391.96	363.13	334.89	352.33	373.52	349.35	372.05	356.87
	N	15,979	14,523	16,832	16,678	15,827	14,520	14,011	13,886
9	Mean	378.22	376.60	352.40	339.05	360.12	364.98	377.43	363.86
	N	15,573	14,329	16,549	16,372	15,396	14,316	13,709	13,545
10	Mean	385.00	382.97	359.00	347.60	367.47	371.77	383.98	370.92
	N	13,841	12,781	14,598	14,428	13,675	12,773	12,304	12,146
11	Mean	390.04	388.32	366.26	356.84	374.47	377.95	389.37	377.39
	N	11,977	11,196	12,628	12,500	11,852	11,188	10,787	10,672
12	Mean	389.32	387.22	364.93	361.97	376.62	376.83	388.50	377.29
	N	7,982	7,516	8,379	8,290	7,889	7,507	7,241	7,151

Table 1.2.2.2
Mean Scale Scores by Grade by Gender, S501 Paper

Grade	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
K	F	Mean	270.61	185.64	204.98	278.02	274.58	195.56	211.12	219.04
		N	102,782	102,780	102,780	102,783	102,782	102,779	102,779	102,778
	M	Mean	260.00	183.04	196.31	261.67	261.10	189.91	206.12	211.05
		N	116,888	116,883	116,876	116,886	116,884	116,872	116,880	116,869
	Missing	Mean	248.24	169.41	181.25	245.38	247.06	175.54	193.05	196.79
		N	6,331	6,331	6,331	6,331	6,331	6,331	6,331	6,331
1	F	Mean	309.99	290.73	260.12	274.00	293.54	276.29	297.11	282.02
		N	15,849	14,054	18,193	18,040	15,722	14,050	12,724	12,622
	M	Mean	306.53	288.65	250.20	270.18	289.99	270.50	294.76	277.06
		N	17,740	16,017	21,012	20,853	17,598	16,006	14,226	14,110
	Missing	Mean	301.39	283.43	246.83	259.51	281.59	266.18	289.52	271.46
		N	1,298	1,152	1,539	1,533	1,292	1,152	1,028	1,023
2	F	Mean	334.63	318.32	292.36	295.51	315.95	306.27	323.74	309.63
		N	17,101	15,470	18,159	18,031	16,981	15,466	14,828	14,722
	M	Mean	331.23	315.21	279.87	289.28	311.46	298.70	320.66	303.11
		N	19,001	17,094	20,564	20,413	18,854	17,086	16,180	16,049
	Missing	Mean	320.87	307.84	268.80	270.52	297.00	288.89	312.30	291.85
		N	982	850	1,095	1,084	973	850	787	782
3	F	Mean	358.30	338.19	307.35	306.22	332.93	323.23	344.50	326.27
		N	15,486	14,083	16,317	16,189	15,370	14,083	13,541	13,443
	M	Mean	358.65	337.21	294.94	301.77	331.06	316.83	344.04	321.50
		N	18,347	16,503	19,709	19,564	18,215	16,498	15,705	15,589
	Missing	Mean	348.65	332.90	285.59	288.68	319.33	309.20	337.99	312.57
		N	1,001	909	1,107	1,095	990	909	842	832
4	F	Mean	374.28	348.93	332.03	328.94	352.25	341.01	356.85	344.52
		N	11,917	11,030	12,393	12,306	11,832	11,029	10,728	10,651
	M	Mean	375.97	348.66	322.60	329.11	353.48	336.28	357.29	341.75
		N	13,848	12,596	14,455	14,344	13,743	12,593	12,214	12,121
	Missing	Mean	367.28	344.86	316.58	314.78	342.10	331.26	351.97	334.75
		N	925	837	973	966	919	837	809	803
5	F	Mean	383.21	356.91	342.40	334.36	359.55	350.17	365.15	353.18
		N	10,372	9,775	10,749	10,659	10,288	9,775	9,526	9,453
	M	Mean	385.07	356.43	331.69	334.61	360.72	344.79	365.45	349.86
		N	12,426	11,598	12,955	12,842	12,316	11,598	11,269	11,175
	Missing	Mean	378.38	354.46	329.19	325.26	352.43	342.08	362.04	345.43
		N	843	778	873	871	841	778	756	754

Grade	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
6	F	Mean	385.89	354.89	336.83	353.15	370.56	346.18	364.38	353.51
		N	8,480	7,594	8,828	8,745	8,402	7,592	7,375	7,305
	M	Mean	382.72	351.63	324.44	353.04	368.88	338.60	361.20	347.73
		N	10,180	9,018	10,669	10,563	10,083	9,016	8,725	8,647
	Missing	Mean	380.71	352.03	324.57	346.29	364.37	339.27	360.92	346.90
		N	790	732	854	842	778	732	693	683
7	F	Mean	389.52	359.39	338.19	349.87	370.77	349.02	368.61	355.59
		N	7,381	6,732	7,667	7,605	7,320	6,727	6,543	6,485
	M	Mean	387.58	355.97	327.65	352.65	370.97	342.16	365.64	350.81
		N	8,735	7,807	9,133	9,074	8,667	7,796	7,557	7,491
	Missing	Mean	385.57	355.13	328.77	345.85	365.92	342.62	364.66	349.40
		N	738	697	761	750	728	697	682	674
8	F	Mean	393.58	365.63	340.34	351.50	374.00	353.21	374.25	359.69
		N	6,908	6,302	7,243	7,165	6,832	6,301	6,106	6,043
	M	Mean	390.74	361.13	330.55	353.39	373.40	346.20	370.27	354.62
		N	8,361	7,554	8,833	8,766	8,293	7,552	7,264	7,210
	Missing	Mean	390.59	362.08	333.36	347.99	370.22	348.42	371.30	355.55
		N	710	667	756	747	702	667	641	633
9	F	Mean	378.28	379.12	358.23	339.05	359.93	369.06	379.19	366.54
		N	6,728	6,245	7,075	7,015	6,666	6,241	6,012	5,953
	M	Mean	378.48	374.97	348.28	340.07	360.82	362.15	376.37	362.11
		N	8,174	7,484	8,737	8,639	8,076	7,475	7,136	7,047
	Missing	Mean	374.45	370.60	345.30	326.81	353.39	357.74	371.95	357.17
		N	671	600	737	718	654	600	561	545
10	F	Mean	385.70	385.53	364.75	348.20	368.13	375.75	385.96	373.87
		N	6,117	5,711	6,404	6,341	6,055	5,709	5,520	5,457
	M	Mean	385.54	381.67	355.62	348.97	368.28	369.54	383.25	369.52
		N	7,071	6,476	7,483	7,406	6,995	6,470	6,220	6,149
	Missing	Mean	372.58	372.58	342.70	326.97	351.93	357.93	372.73	357.05
		N	653	594	711	681	625	594	564	540
11	F	Mean	389.14	389.68	370.20	354.66	372.88	380.58	389.98	378.59
		N	5,291	4,976	5,521	5,462	5,230	4,973	4,825	4,769
	M	Mean	392.19	388.38	364.62	361.19	377.60	377.19	390.10	377.87
		N	6,034	5,599	6,399	6,345	5,984	5,594	5,372	5,326
	Missing	Mean	377.52	376.88	350.44	334.05	358.18	363.69	377.69	362.97
		N	652	621	708	693	638	621	590	577
12	F	Mean	388.16	388.74	369.29	358.95	374.35	379.71	389.04	378.53
		N	3,690	3,486	3,829	3,794	3,651	3,480	3,388	3,350
	M	Mean	391.38	386.96	362.43	366.39	379.80	375.43	389.02	377.19
		N	3,868	3,622	4,083	4,044	3,826	3,619	3,476	3,435
	Missing	Mean	380.60	376.62	351.00	347.88	367.16	364.71	378.97	366.90
		N	424	408	467	452	412	408	377	366

Table 1.2.2.3
Mean Scale Scores by Grade by Ethnicity, S501 Paper

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
K	Non-Hispanic Asian	Mean	284.38	218.55	231.98	287.56	286.23	225.53	238.29	243.50
		N	28,528	28,525	28,526	28,527	28,527	28,524	28,525	28,524
	Non-Hispanic Pacific Islander	Mean	259.69	165.04	185.81	266.63	263.41	175.65	193.43	201.76
		N	1,451	1,451	1,451	1,451	1,451	1,451	1,451	1,451
	Non-Hispanic Black	Mean	273.01	197.54	208.10	287.61	280.57	203.06	220.17	226.10
		N	13,038	13,037	13,036	13,039	13,038	13,036	13,036	13,035
	Hispanic (Of Any Race)	Mean	258.59	175.11	191.60	262.16	260.63	183.59	200.14	206.48
		N	144,100	144,098	144,094	144,099	144,098	144,091	144,096	144,089
	Non-Hispanic American Indian	Mean	269.11	173.77	187.29	267.25	268.46	180.76	202.36	206.85
		N	1,879	1,880	1,880	1,880	1,879	1,880	1,879	1,879
	Non-Hispanic Multiracial	Mean	286.10	207.33	218.02	297.43	292.03	212.93	230.94	236.44
		N	1,209	1,209	1,209	1,209	1,209	1,209	1,209	1,209
	Non-Hispanic White	Mean	280.11	196.59	215.84	287.86	284.25	206.46	221.63	229.57
		N	21,803	21,802	21,801	21,803	21,803	21,801	21,802	21,801
	Unknown	Mean	250.61	171.68	187.97	247.30	249.20	180.05	195.35	200.59
		N	13,993	13,992	13,990	13,992	13,992	13,990	13,992	13,990
1	Non-Hispanic Asian	Mean	311.03	301.76	267.61	281.29	297.80	285.59	305.15	289.92
		N	1,807	1,642	2,048	2,041	1,800	1,640	1,492	1,484
	Non-Hispanic Pacific Islander	Mean	301.82	289.59	264.54	282.96	296.22	277.98	292.50	284.05
		N	45	41	50	50	45	41	38	38
	Non-Hispanic Black	Mean	304.87	288.83	249.89	278.56	293.08	271.10	294.41	278.76
		N	2,487	2,223	3,062	3,026	2,460	2,223	1,913	1,888
	Hispanic (Of Any Race)	Mean	307.27	287.65	253.11	268.26	289.36	271.26	294.20	277.29
		N	26,413	23,566	30,790	30,558	26,204	23,555	21,154	20,991
	Non-Hispanic American Indian	Mean	304.49	290.22	259.04	273.09	291.46	275.17	294.63	279.84
		N	221	203	269	265	217	203	175	173
	Non-Hispanic Multiracial	Mean	312.91	301.89	263.04	288.37	303.29	284.14	306.38	291.23
		N	258	236	295	295	258	235	215	214
	Non-Hispanic White	Mean	312.35	294.24	260.18	282.98	299.19	278.28	300.40	285.20
		N	2,380	2,174	2,783	2,756	2,360	2,173	1,952	1,934
	Unknown	Mean	314.23	296.83	261.61	284.86	301.12	280.24	302.69	286.89
		N	1,276	1,138	1,447	1,435	1,268	1,138	1,039	1,033

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
2	Non-Hispanic Asian	Mean	335.48	327.37	298.55	299.92	318.21	313.49	330.19	315.20
		N	1,799	1,641	1,913	1,905	1,792	1,641	1,575	1,570
	Non-Hispanic Pacific Islander	Mean	327.94	318.37	291.71	292.02	310.25	305.54	322.11	308.13
		N	52	41	56	56	52	41	38	38
	Non-Hispanic Black	Mean	330.40	313.25	281.54	293.68	313.22	298.82	319.08	303.72
		N	2,895	2,547	3,163	3,140	2,874	2,545	2,374	2,356
	Hispanic (Of Any Race)	Mean	331.87	315.12	283.89	289.59	311.84	300.61	320.74	304.55
		N	28,240	25,397	30,269	30,047	28,024	25,387	24,221	24,030
	Non-Hispanic American Indian	Mean	333.33	315.26	292.18	283.08	309.75	304.38	320.86	306.90
		N	233	210	261	258	232	210	198	197
	Non-Hispanic Multiracial	Mean	340.29	327.23	297.32	310.01	325.43	312.91	331.57	316.14
		N	219	202	234	232	218	202	192	191
	Non-Hispanic White	Mean	338.75	323.10	292.01	303.13	322.10	308.19	328.54	313.13
		N	2,331	2,162	2,526	2,508	2,314	2,162	2,033	2,020
	Unknown	Mean	334.94	323.08	289.47	297.01	316.97	307.43	327.26	310.53
		N	1,315	1,214	1,396	1,382	1,302	1,214	1,164	1,151
3	Non-Hispanic Asian	Mean	367.81	345.27	311.56	311.00	339.84	328.83	352.52	332.44
		N	1,457	1,339	1,548	1,538	1,449	1,339	1,284	1,276
	Non-Hispanic Pacific Islander	Mean	354.84	335.95	310.40	312.74	335.07	323.63	340.76	325.16
		N	43	41	53	53	43	41	37	37
	Non-Hispanic Black	Mean	356.48	335.62	296.31	305.20	331.56	316.82	342.07	321.58
		N	2,971	2,663	3,213	3,173	2,935	2,662	2,521	2,490
	Hispanic (Of Any Race)	Mean	357.01	336.71	299.00	301.27	329.93	318.47	343.14	322.18
		N	26,709	24,058	28,404	28,202	26,525	24,055	23,023	22,866
	Non-Hispanic American Indian	Mean	354.17	336.28	309.30	295.65	324.73	322.50	341.72	322.81
		N	223	193	231	231	223	193	188	188
	Non-Hispanic Multiracial	Mean	360.15	338.14	303.69	313.66	338.53	321.20	344.43	326.11
		N	175	174	194	194	175	174	160	160
	Non-Hispanic White	Mean	364.61	341.98	305.81	315.60	341.06	324.56	349.05	329.85
		N	2,025	1,882	2,167	2,149	2,008	1,881	1,796	1,780
	Unknown	Mean	366.93	342.77	308.05	313.57	341.54	326.47	350.89	331.70
		N	1,231	1,145	1,323	1,308	1,217	1,145	1,079	1,067

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
4	Non-Hispanic Asian	Mean	384.79	359.49	335.87	341.15	363.88	348.52	367.58	353.40
		N	959	884	1,005	1,002	956	884	855	853
	Non-Hispanic Pacific Islander	Mean	378.96	351.95	330.89	341.96	360.70	346.37	361.42	352.37
		N	27	19	28	28	27	19	19	19
	Non-Hispanic Black	Mean	376.29	347.16	324.11	333.13	355.26	336.43	356.30	342.43
		N	2,163	1,956	2,298	2,269	2,135	1,956	1,864	1,842
	Hispanic (Of Any Race)	Mean	373.50	347.60	325.72	325.92	350.53	337.23	355.70	341.38
		N	20,413	18,670	21,198	21,048	20,267	18,666	18,175	18,041
	Non-Hispanic American Indian	Mean	372.40	348.76	330.02	327.33	350.48	341.14	356.26	344.25
		N	236	221	243	239	232	221	215	212
	Non-Hispanic Multiracial	Mean	380.21	354.75	332.60	342.06	362.63	344.56	362.91	350.46
		N	132	124	140	140	132	124	119	119
	Non-Hispanic White	Mean	380.83	353.13	330.34	339.64	361.66	342.32	362.13	348.92
		N	1,612	1,511	1,697	1,685	1,602	1,511	1,459	1,449
	Unknown	Mean	380.75	353.59	332.02	337.72	359.71	343.45	362.40	348.69
		N	1,148	1,078	1,212	1,205	1,143	1,078	1,045	1,040
5	Non-Hispanic Asian	Mean	388.52	363.25	340.50	339.04	364.90	352.39	371.22	356.26
		N	685	663	719	712	679	663	644	639
	Non-Hispanic Pacific Islander	Mean	384.56	362.09	348.51	344.56	367.83	357.03	369.97	361.16
		N	36	34	37	36	35	34	33	32
	Non-Hispanic Black	Mean	384.32	356.47	335.85	342.25	363.65	346.87	365.15	352.15
		N	2,097	1,917	2,195	2,181	2,084	1,917	1,862	1,851
	Hispanic (Of Any Race)	Mean	382.74	355.17	335.32	331.25	357.84	345.86	363.81	349.67
		N	18,126	16,974	18,806	18,649	17,974	16,974	16,530	16,401
	Non-Hispanic American Indian	Mean	385.48	356.03	339.20	340.58	362.89	347.77	364.71	352.22
		N	189	182	195	194	188	182	176	175
	Non-Hispanic Multiracial	Mean	395.30	365.17	343.51	355.20	376.43	354.63	374.52	361.70
		N	124	113	128	124	120	113	113	109
	Non-Hispanic White	Mean	391.40	362.98	341.13	346.60	370.07	352.77	372.08	358.44
		N	1,397	1,318	1,479	1,469	1,389	1,318	1,264	1,256
	Unknown	Mean	391.48	367.03	343.09	344.95	369.11	356.16	375.17	360.64
		N	987	950	1,018	1,007	976	950	929	919

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
6	Non-Hispanic Asian	Mean	389.56	359.67	333.68	361.70	376.50	347.18	368.67	355.85
		N	607	555	634	629	603	555	538	534
	Non-Hispanic Pacific Islander	Mean	381.96	358.89	332.03	357.83	371.33	345.77	366.46	353.63
		N	27	27	29	30	27	26	24	24
	Non-Hispanic Black	Mean	386.75	352.45	329.70	360.94	374.78	341.75	362.99	351.64
		N	1,654	1,403	1,746	1,704	1,617	1,401	1,353	1,325
	Hispanic (Of Any Race)	Mean	382.62	352.01	329.07	350.14	367.42	340.99	361.41	348.98
		N	15,020	13,403	15,683	15,551	14,894	13,402	12,991	12,882
	Non-Hispanic American Indian	Mean	384.66	353.98	332.53	352.29	370.09	343.44	363.82	351.62
		N	143	129	150	148	141	129	123	121
	Non-Hispanic Multiracial	Mean	389.63	357.42	333.37	361.97	376.18	345.82	366.83	353.88
		N	93	83	97	97	93	83	80	80
	Non-Hispanic White	Mean	391.05	358.54	334.17	366.20	379.83	347.08	368.74	357.19
		N	1,116	1,012	1,175	1,161	1,105	1,012	980	970
	Unknown	Mean	389.96	360.37	334.15	359.39	375.55	347.94	369.42	355.98
		N	790	732	837	830	783	732	704	699
7	Non-Hispanic Asian	Mean	397.12	366.43	339.42	363.32	381.42	353.74	376.41	362.36
		N	531	480	546	542	527	480	473	470
	Non-Hispanic Pacific Islander	Mean	381.92	356.41	333.39	351.47	367.17	346.28	365.45	354.07
		N	37	32	38	36	36	32	31	30
	Non-Hispanic Black	Mean	390.39	357.22	332.30	361.19	375.64	345.67	367.24	354.58
		N	1,389	1,204	1,457	1,421	1,355	1,201	1,168	1,141
	Hispanic (Of Any Race)	Mean	386.07	355.69	330.70	346.86	367.50	343.43	364.96	350.65
		N	13,005	11,800	13,538	13,456	12,914	11,789	11,448	11,358
	Non-Hispanic American Indian	Mean	395.36	362.15	341.46	364.34	380.08	351.80	372.01	359.86
		N	137	109	138	138	137	109	109	109
	Non-Hispanic Multiracial	Mean	402.72	368.35	340.82	374.88	390.71	355.62	378.83	366.54
		N	89	78	96	96	89	78	76	76
	Non-Hispanic White	Mean	399.66	365.92	340.16	371.91	387.10	353.47	376.51	363.94
		N	991	908	1,044	1,039	987	908	873	869
	Unknown	Mean	401.30	369.65	342.91	366.91	384.45	356.88	379.48	365.06
		N	675	625	704	701	670	623	604	597

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
8	Non-Hispanic Asian	Mean	403.68	374.72	348.30	370.31	387.73	361.93	383.99	370.23
		N	524	474	549	548	523	474	460	460
	Non-Hispanic Pacific Islander	Mean	395.67	365.61	341.12	363.42	378.23	353.14	372.65	359.31
		N	30	28	33	33	30	28	26	26
	Non-Hispanic Black	Mean	395.05	362.46	334.97	360.85	379.39	349.03	372.43	358.46
		N	1,356	1,157	1,467	1,441	1,331	1,157	1,099	1,084
	Hispanic (Of Any Race)	Mean	389.73	361.57	333.02	347.89	370.21	347.58	370.25	354.57
		N	12,359	11,268	12,978	12,859	12,242	11,266	10,885	10,784
	Non-Hispanic American Indian	Mean	391.26	364.84	334.74	349.58	372.95	350.66	374.68	359.90
		N	102	96	107	106	101	96	91	90
	Non-Hispanic Multiracial	Mean	407.63	373.70	346.37	367.47	389.69	359.70	384.16	369.07
		N	98	92	104	104	98	92	88	88
	Non-Hispanic White	Mean	400.95	370.73	342.42	373.56	388.40	357.27	380.31	367.05
		N	944	879	997	993	939	878	852	847
	Unknown	Mean	404.58	372.44	348.05	373.07	390.58	360.88	382.57	370.41
		N	566	529	597	594	563	529	510	507
9	Non-Hispanic Asian	Mean	392.79	389.95	367.50	360.19	377.66	380.04	391.29	379.89
		N	529	478	554	547	522	478	468	462
	Non-Hispanic Pacific Islander	Mean	384.51	377.13	363.20	346.56	365.70	370.58	377.00	367.40
		N	37	38	41	41	37	38	35	35
	Non-Hispanic Black	Mean	375.44	373.86	346.74	341.51	359.46	360.71	373.88	359.80
		N	1,325	1,144	1,447	1,442	1,321	1,144	1,068	1,064
	Hispanic (Of Any Race)	Mean	376.01	375.00	351.08	335.59	357.29	363.37	375.61	361.87
		N	11,955	11,026	12,665	12,538	11,826	11,016	10,575	10,457
	Non-Hispanic American Indian	Mean	388.95	378.65	360.44	338.46	365.82	369.55	382.53	368.96
		N	141	127	151	145	137	127	120	116
	Non-Hispanic Multiracial	Mean	398.37	389.99	370.17	366.43	382.98	379.38	392.85	380.32
		N	82	71	83	83	82	71	71	71
	Non-Hispanic White	Mean	388.80	385.48	359.75	356.43	374.55	373.57	387.32	374.70
		N	993	953	1,059	1,045	978	950	905	889
	Unknown	Mean	394.88	386.11	362.83	353.36	376.75	375.98	390.09	377.47
		N	511	492	549	531	493	492	467	451

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
10	Non-Hispanic Asian	Mean	395.20	394.46	372.17	363.31	380.77	384.45	395.12	383.83
		N	494	467	516	510	487	466	453	446
	Non-Hispanic Pacific Islander	Mean	402.12	392.46	374.72	362.14	385.38	384.63	399.91	388.00
		N	26	24	29	29	26	24	22	22
	Non-Hispanic Black	Mean	379.49	378.16	349.66	347.16	363.87	365.11	378.59	364.72
		N	1,302	1,077	1,412	1,398	1,288	1,076	1,017	1,007
	Hispanic (Of Any Race)	Mean	383.54	382.03	358.47	345.03	365.52	370.86	382.85	369.71
		N	10,648	9,925	11,202	11,087	10,535	9,919	9,564	9,455
	Non-Hispanic American Indian	Mean	392.43	382.89	365.68	352.93	371.82	374.71	385.29	373.67
		N	69	63	73	71	68	63	62	61
	Non-Hispanic Multiracial	Mean	395.04	385.37	366.77	359.63	378.34	377.78	387.98	377.67
		N	69	63	73	72	68	63	62	61
	Non-Hispanic White	Mean	397.13	390.71	365.25	362.24	381.12	379.05	393.23	380.19
		N	800	748	840	830	791	748	722	713
	Unknown	Mean	399.76	390.20	371.20	364.25	382.87	381.50	393.63	382.29
		N	433	414	453	431	412	414	402	381
11	Non-Hispanic Asian	Mean	403.81	400.30	381.23	369.02	387.34	391.17	402.17	390.99
		N	513	483	534	527	506	483	471	464
	Non-Hispanic Pacific Islander	Mean	402.59	392.88	378.39	375.59	394.88	390.06	396.31	389.87
		N	17	16	18	17	16	16	16	15
	Non-Hispanic Black	Mean	383.01	381.28	356.01	352.36	368.56	369.47	382.36	369.68
		N	1,271	1,117	1,375	1,357	1,253	1,115	1,065	1,047
	Hispanic (Of Any Race)	Mean	389.04	387.88	366.08	355.47	373.33	377.56	388.68	376.73
		N	8,955	8,419	9,418	9,335	8,875	8,413	8,112	8,040
	Non-Hispanic American Indian	Mean	394.50	390.00	364.12	350.20	371.75	376.48	391.47	374.64
		N	66	56	68	66	64	56	55	53
	Non-Hispanic Multiracial	Mean	403.15	394.13	377.04	370.12	386.75	385.95	397.26	386.30
		N	67	64	74	73	67	64	61	61
	Non-Hispanic White	Mean	400.26	394.65	372.52	371.57	387.03	384.08	397.07	385.50
		N	726	700	760	756	721	700	677	672
	Unknown	Mean	395.77	390.69	371.95	357.83	377.30	382.20	392.61	380.51
		N	362	341	381	369	350	341	330	320

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
12	Non-Hispanic Asian	Mean	398.52	397.26	375.98	375.43	388.04	387.26	398.34	388.32
		N	379	350	392	387	374	350	343	338
	Non-Hispanic Pacific Islander	Mean	408.30	406.50	373.77	407.23	404.10	389.25	411.00	397.78
		N	10	12	13	13	10	12	9	9
	Non-Hispanic Black	Mean	379.48	379.76	352.78	354.45	368.09	367.10	380.23	367.80
		N	1,102	1,003	1,178	1,170	1,093	1,001	953	942
	Hispanic (Of Any Race)	Mean	389.12	387.05	365.45	360.96	376.00	376.93	388.30	377.17
		N	5,707	5,418	5,981	5,923	5,646	5,411	5,220	5,162
	Non-Hispanic American	Mean	383.83	383.48	368.20	358.10	371.72	375.22	384.08	374.42
		N	29	27	30	30	29	27	26	26
	Non-Hispanic Multiracial	Mean	398.03	387.22	371.59	379.18	388.29	378.46	389.94	381.17
		N	38	37	39	39	38	37	36	36
	Non-Hispanic White	Mean	398.30	393.68	369.75	369.05	384.61	382.77	395.33	383.46
		N	449	419	468	462	443	419	412	407
	Unknown	Mean	404.66	395.49	379.68	381.41	393.71	388.42	399.23	390.49
		N	268	250	278	266	256	250	242	231

1.2.3 Correlations

The tables in this section show Pearson correlations among the four domain scale scores by grade-level clusters across all tiers, as well as the number of students included in each correlation. Results are provided by grade-level cluster. In earlier grades of K, 1, and 2, the correlation between Listening and Speaking, and the correlation between Reading and Writing were pronounced. In Grades 3 to 12, the highest correlations were the correlation between Listening and Reading and the correlation between Reading and Writing.

Table 1.2.3.1
Correlations Among Scale Scores: K, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.526	0.563	0.784
	N	226,001	225,990	225,983	225,997
Reading	Pearson Correlation		1	0.717	0.486
	N		225,994	225,982	225,992
Writing	Pearson Correlation			1	0.537
	N			225,987	225,984
Speaking	Pearson Correlation				1
	N				226,000

Table 1.2.3.2
Correlations Among Scale Scores: Grade 1, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.499	0.472	0.504
	N	34,887	27,978	34,871	34,612
Reading	Pearson Correlation		1	0.548	0.455
	N		31,223	31,208	30,978
Writing	Pearson Correlation			1	0.491
	N			40,744	40,403
Speaking	Pearson Correlation				1
	N				40,426

Table 1.2.3.3
Correlations Among Scale Scores: Grade 2, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.617	0.559	0.549
	N	37,084	31,795	37,072	36,808
Reading	Pearson Correlation		1	0.678	0.530
	N		33,414	33402	33169
Writing	Pearson Correlation			1	0.549
	N			39,818	39,514
Speaking	Pearson Correlation				1
	N				39,528

Table 1.2.3.4
Correlations Among Scale Scores: Grade 3, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.654	0.540	0.509
	N	34,834	30,088	34,831	34,575
Reading	Pearson Correlation		1	0.627	0.533
	N		31,495	31,490	31,259
Writing	Pearson Correlation			1	0.562
	N			37,133	36,843
Speaking	Pearson Correlation				1
	N				36,848

Table 1.2.3.5
Correlations Among Scale Scores: Grades 4–5, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.728	0.628	0.616
	N	50,331	45,302	50,326	49,939
Reading	Pearson Correlation		1	0.693	0.631
	N		46,614	46,610	46,259
Writing	Pearson Correlation			1	0.638
	N			52,398	51,983
Speaking	Pearson Correlation				1
	N				51,988

Table 1.2.3.6
Correlations Among Scale Scores: Grades 6–8, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.727	0.717	0.685
	N	52,283	45,586	52,253	51,805
Reading	Pearson Correlation		1	0.676	0.618
	N		47,103	47,080	46,688
Writing	Pearson Correlation			1	0.706
	N			54,744	54,224
Speaking	Pearson Correlation				1
	N				54,257

Table 1.2.3.7
Correlations Among Scale Scores: Grades 9–12, S501 Paper

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.728	0.678	0.662
	N	49,373	44,041	49,336	48,812
Reading	Pearson Correlation		1	0.686	0.621
	N		45,822	45,784	45,291
Writing	Pearson Correlation			1	0.675
	N			52,154	51,538
Speaking	Pearson Correlation				1
	N				51,590

1.3 Proficiency Level Results

Proficiency level (PL) results show the distribution of students falling into the six language proficiency levels outlined by the WIDA ELD Standards. The results are presented in eight subsections—four domains and four composites--by count and percentage.

Each table in this section shows either the number or percentage of students classified into each language proficiency level. Results are first presented by grade-level cluster and tier, then by grade and tier, and then by grade alone.

Performance of PL 5 and 6 was observed in the descending order of Listening, Reading, Writing, and Speaking. The percentages of PL 6 in Tier B/C in Listening were as follows: clusters 1 to 6–8 had 40% to 47%; K, 33%; and 9–12, 15%. The percentages of PL 6 in Tier B/C in the Reading domain were as follows: Grade 1, 13%; Grade 2, 21%; Grade 3, 10%; Grades 4–5, 16%; Grades 6–8, 5.5%; and Grades 9–12, 14%. For the Writing domain, less than 1% were in PL 6. In the Speaking domain, 28% of Kindergarten students reached PL 6. In Grades 1–12, percentages in PL 6 were low but increased over the grades (Grade 1, 2%; Grade 2, 4%; Grade 3, 6%; Grades 4–5, 6%; Grades 6–8, 9%; Grades 9–12, 10%).

1.3.1 Domains

1.3.1.1 Listening

1.3.1.1.1 By Cluster by Tier

Table 1.3.1.1.1.1
Proficiency Level by Cluster (Count): Listening, S501 Paper

Cluster	Tier	Listening Proficiency Range						Total
		1	2	3	4	5	6	
K	-	63,378	21,524	19,067	12,715	32,929	76,388	226,001
1	A	1,404	2,110	2,688	1,923	4,887	3,713	16,725
	BC	132	473	3,523	1,693	4,187	8,154	18,162
2	A	1,247	1,974	1,569	1,083	2,340	11	8,224
	BC	63	1,267	3,910	5,879	4,208	13,533	28,860
3	A	274	1,883	1,917	1,124	938	687	6,823
	BC	14	464	3,050	4,973	6,942	12,568	28,011
4–5	A	844	3,177	2,482	1,493	956	820	9,772
	BC	51	499	2,743	4,955	14,482	17,829	40,559
6–8	A	3,676	4,761	2,050	1,017	840	160	12,504
	BC	47	767	3,797	8,725	10,442	16,001	39,779
9–12	A	5,975	4,271	1,551	532	161	0	12,490
	BC	756	3,032	9,538	9,886	8,111	5,560	36,883

Table 1.3.1.1.1.2
Proficiency Level by Cluster (Percent): Listening, S501 Paper

Cluster	Tier	Listening Proficiency Range						Total
		1	2	3	4	5	6	
K	-	28.0%	9.5%	8.4%	5.6%	14.6%	33.8%	100.0%
1	A	8.4%	12.6%	16.1%	11.5%	29.2%	22.2%	100.0%
	BC	0.7%	2.6%	19.4%	9.3%	23.1%	44.9%	100.0%
2	A	15.2%	24.0%	19.1%	13.2%	28.5%	0.1%	100.0%
	BC	0.2%	4.4%	13.5%	20.4%	14.6%	46.9%	100.0%
3	A	4.0%	27.6%	28.1%	16.5%	13.7%	10.1%	100.0%
	BC	0.0%	1.7%	10.9%	17.8%	24.8%	44.9%	100.0%
4–5	A	8.6%	32.5%	25.4%	15.3%	9.8%	8.4%	100.0%
	BC	0.1%	1.2%	6.8%	12.2%	35.7%	44.0%	100.0%
6–8	A	29.4%	38.1%	16.4%	8.1%	6.7%	1.3%	100.0%
	BC	0.1%	1.9%	9.5%	21.9%	26.3%	40.2%	100.0%
9–12	A	47.8%	34.2%	12.4%	4.3%	1.3%	0.0%	100.0%
	BC	2.0%	8.2%	25.9%	26.8%	22.0%	15.1%	100.0%

1.3.1.1.2 By Grade by Tier

Table 1.3.1.1.2.1
Proficiency Level by Grade (Count): Listening, S501 Paper

Grade	Tier	Listening Proficiency Range						Total
		1	2	3	4	5	6	
K	-	63,378	21,524	19,067	12,715	32,929	76,388	226,001
1	A	1,404	2,110	2,688	1,923	4,887	3,713	16,725
	BC	132	473	3,523	1,693	4,187	8,154	18,162
2	A	1,247	1,974	1,569	1,083	2,340	11	8,224
	BC	63	1,267	3,910	5,879	4,208	13,533	28,860
3	A	274	1,883	1,917	1,124	938	687	6,823
	BC	14	464	3,050	4,973	6,942	12,568	28,011
4	A	380	1,623	1,342	784	595	388	5,112
	BC	17	240	1,478	2,669	7,634	9,540	21,578
5	A	464	1,554	1,140	709	361	432	4,660
	BC	34	259	1,265	2,286	6,848	8,289	18,981
6	A	990	1,766	642	469	325	75	4,267
	BC	9	237	1,242	3,689	3,946	6,060	15,183
7	A	1,389	1,295	838	198	300	84	4,104
	BC	25	298	1,180	3,150	3,414	4,683	12,750
8	A	1,297	1,700	570	350	215	1	4,133
	BC	13	232	1,375	1,886	3,082	5,258	11,846
9	A	1,866	2,217	529	120	98	0	4,830
	BC	38	720	2,180	3,582	2,154	2,069	10,743
10	A	1,676	1,152	463	195	33	0	3,519
	BC	120	742	2,910	2,179	3,043	1,328	10,322
11	A	1,470	606	395	181	21	0	2,673
	BC	222	965	2,052	2,862	1,830	1,373	9,304
12	A	963	296	164	36	9	0	1,468
	BC	376	605	2,396	1,263	1,084	790	6,514

Table 1.3.1.1.2.2

Proficiency Level by Grade (Percent): Listening, S501 Paper

Grade	Tier	Listening Proficiency Range						Total
		1	2	3	4	5	6	
K	-	28.0%	9.5%	8.4%	5.6%	14.6%	33.8%	100.0%
1	A	8.4%	12.6%	16.1%	11.5%	29.2%	22.2%	100.0%
	BC	0.7%	2.6%	19.4%	9.3%	23.1%	44.9%	100.0%
2	A	15.2%	24.0%	19.1%	13.2%	28.5%	0.1%	100.0%
	BC	0.2%	4.4%	13.5%	20.4%	14.6%	46.9%	100.0%
3	A	4.0%	27.6%	28.1%	16.5%	13.7%	10.1%	100.0%
	BC	0.0%	1.7%	10.9%	17.8%	24.8%	44.9%	100.0%
4	A	7.4%	31.7%	26.3%	15.3%	11.6%	7.6%	100.0%
	BC	0.1%	1.1%	6.8%	12.4%	35.4%	44.2%	100.0%
5	A	10.0%	33.3%	24.5%	15.2%	7.7%	9.3%	100.0%
	BC	0.2%	1.4%	6.7%	12.0%	36.1%	43.7%	100.0%
6	A	23.2%	41.4%	15.0%	11.0%	7.6%	1.8%	100.0%
	BC	0.1%	1.6%	8.2%	24.3%	26.0%	39.9%	100.0%
7	A	33.8%	31.6%	20.4%	4.8%	7.3%	2.0%	100.0%
	BC	0.2%	2.3%	9.3%	24.7%	26.8%	36.7%	100.0%
8	A	31.4%	41.1%	13.8%	8.5%	5.2%	0.0%	100.0%
	BC	0.1%	2.0%	11.6%	15.9%	26.0%	44.4%	100.0%
9	A	38.6%	45.9%	11.0%	2.5%	2.0%	0.0%	100.0%
	BC	0.4%	6.7%	20.3%	33.3%	20.1%	19.3%	100.0%
10	A	47.6%	32.7%	13.2%	5.5%	0.9%	0.0%	100.0%
	BC	1.2%	7.2%	28.2%	21.1%	29.5%	12.9%	100.0%
11	A	55.0%	22.7%	14.8%	6.8%	0.8%	0.0%	100.0%
	BC	2.4%	10.4%	22.1%	30.8%	19.7%	14.8%	100.0%
12	A	65.6%	20.2%	11.2%	2.5%	0.6%	0.0%	100.0%
	BC	5.8%	9.3%	36.8%	19.4%	16.6%	12.1%	100.0%

1.3.1.1.3 By Grade

Table 1.3.1.1.3.1
Proficiency Level by Grade (Count): Listening

Grade	Listening Proficiency Range						Total
	1	2	3	4	5	6	
K	63,378	21,524	19,067	12,715	32,929	76,388	226,001
1	1,536	2,583	6,211	3,616	9,074	11,867	34,887
2	1,310	3,241	5,479	6,962	6,548	13,544	37,084
3	288	2,347	4,967	6,097	7,880	13,255	34,834
4	397	1,863	2,820	3,453	8,229	9,928	26,690
5	498	1,813	2,405	2,995	7,209	8,721	23,641
6	999	2,003	1,884	4,158	4,271	6,135	19,450
7	1,414	1,593	2,018	3,348	3,714	4,767	16,854
8	1,310	1,932	1,945	2,236	3,297	5,259	15,979
9	1,904	2,937	2,709	3,702	2,252	2,069	15,573
10	1,796	1,894	3,373	2,374	3,076	1,328	13,841
11	1,692	1,571	2,447	3,043	1,851	1,373	11,977
12	1,339	901	2,560	1,299	1,093	790	7,982

Table 1.3.1.1.3.2
Proficiency Level by Grade (Percent): Listening

Grade	Listening Proficiency Range						Total
	1	2	3	4	5	6	
K	28.0%	9.5%	8.4%	5.6%	14.6%	33.8%	100.0%
1	4.4%	7.4%	17.8%	10.4%	26.0%	34.0%	100.0%
2	3.5%	8.7%	14.8%	18.8%	17.7%	36.5%	100.0%
3	0.8%	6.7%	14.3%	17.5%	22.6%	38.1%	100.0%
4	1.5%	7.0%	10.6%	12.9%	30.8%	37.2%	100.0%
5	2.1%	7.7%	10.2%	12.7%	30.5%	36.9%	100.0%
6	5.1%	10.3%	9.7%	21.4%	22.0%	31.5%	100.0%
7	8.4%	9.5%	12.0%	19.9%	22.0%	28.3%	100.0%
8	8.2%	12.1%	12.2%	14.0%	20.6%	32.9%	100.0%
9	12.2%	18.9%	17.4%	23.8%	14.5%	13.3%	100.0%
10	13.0%	13.7%	24.4%	17.2%	22.2%	9.6%	100.0%
11	14.1%	13.1%	20.4%	25.4%	15.5%	11.5%	100.0%
12	16.8%	11.3%	32.1%	16.3%	13.7%	9.9%	100.0%

1.3.1.2 Reading

1.3.1.2.1 By Cluster by Tier

Table 1.3.1.2.1.1
Proficiency Level by Cluster (Count): Reading, S501 Paper

Cluster	Tier	Reading Proficiency Range						Total
		1	2	3	4	5	6	
K	-	166,059	7,055	19,489	12,494	20,897	0	225,994
1	A	6,333	5,342	2,104	542	903	555	15,779
	BC	97	2,452	5,833	2,826	2,215	2,021	15,444
2	A	4,318	1,646	766	679	417	0	7,826
	BC	878	6,663	5,257	2,831	4,659	5,300	25,588
3	A	3,407	1,633	616	187	341	209	6,393
	BC	107	2,138	7,260	5,480	7,585	2,532	25,102
4–5	A	5,380	1,882	1,139	416	521	50	9,388
	BC	258	5,116	10,892	5,854	9,208	5,898	37,226
6–8	A	6,481	3,761	1,117	294	319	137	12,109
	BC	1,235	11,967	10,908	4,331	4,612	1,941	34,994
9–12	A	4,265	5,285	1,495	468	639	270	12,422
	BC	787	10,081	9,187	3,571	5,028	4,746	33,400

Table 1.3.1.2.1.2
Proficiency Level by Cluster (Percent): Reading, S501 Paper

Cluster	Tier	Reading Proficiency Range						Total
		1	2	3	4	5	6	
K	-	73.5%	3.1%	8.6%	5.5%	9.2%	0.0%	100.0%
1	A	40.1%	33.9%	13.3%	3.4%	5.7%	3.5%	100.0%
	BC	0.6%	15.9%	37.8%	18.3%	14.3%	13.1%	100.0%
2	A	55.2%	21.0%	9.8%	8.7%	5.3%	0.0%	100.0%
	BC	3.4%	26.0%	20.5%	11.1%	18.2%	20.7%	100.0%
3	A	53.3%	25.5%	9.6%	2.9%	5.3%	3.3%	100.0%
	BC	0.4%	8.5%	28.9%	21.8%	30.2%	10.1%	100.0%
4–5	A	57.3%	20.0%	12.1%	4.4%	5.5%	0.5%	100.0%
	BC	0.7%	13.7%	29.3%	15.7%	24.7%	15.8%	100.0%
6–8	A	53.5%	31.1%	9.2%	2.4%	2.6%	1.1%	100.0%
	BC	3.5%	34.2%	31.2%	12.4%	13.2%	5.5%	100.0%
9–12	A	34.3%	42.5%	12.0%	3.8%	5.1%	2.2%	100.0%
	BC	2.4%	30.2%	27.5%	10.7%	15.1%	14.2%	100.0%

1.3.1.2.2 By Grade by Tier

Table 1.3.1.2.2.1
Proficiency Level by Grade (Count): Reading, S501 Paper

Grade	Tier	Reading Proficiency Range						Total
		1	2	3	4	5	6	
K	-	166,059	7,055	19,489	12,494	20,897	0	225,994
1	A	6,333	5,342	2,104	542	903	555	15,779
	BC	97	2,452	5,833	2,826	2,215	2,021	15,444
2	A	4,318	1,646	766	679	417	0	7,826
	BC	878	6,663	5,257	2,831	4,659	5,300	25,588
3	A	3,407	1,633	616	187	341	209	6,393
	BC	107	2,138	7,260	5,480	7,585	2,532	25,102
4	A	2,746	1,063	537	214	277	50	4,887
	BC	104	2,320	6,140	3,234	4,809	2,969	19,576
5	A	2,634	819	602	202	244	0	4,501
	BC	154	2,796	4,752	2,620	4,399	2,929	17,650
6	A	1,884	1,510	422	105	120	41	4,082
	BC	292	4,674	4,263	1,630	1,835	568	13,262
7	A	2,264	1,100	397	89	61	71	3,982
	BC	426	3,739	3,815	1,332	1,303	639	11,254
8	A	2,333	1,151	298	100	138	25	4,045
	BC	517	3,554	2,830	1,369	1,474	734	10,478
9	A	1,589	2,140	590	141	248	76	4,784
	BC	146	2,506	2,400	1,676	1,424	1,393	9,545
10	A	1,207	1,436	462	139	123	132	3,499
	BC	106	2,706	2,754	1,071	1,417	1,228	9,282
11	A	928	1,075	296	125	215	42	2,681
	BC	194	2,547	2,378	514	1,394	1,488	8,515
12	A	541	634	147	63	53	20	1,458
	BC	341	2,322	1,655	310	793	637	6,058

Table 1.3.1.2.2.2
Proficiency Level by Grade (Percent): Reading, S501 Paper

Grade	Tier	Reading Proficiency Range						Total
		1	2	3	4	5	6	
K	-	73.5%	3.1%	8.6%	5.5%	9.2%	0%	100.0%
1	A	40.1%	33.9%	13.3%	3.4%	5.7%	3.5%	100.0%
	BC	0.6%	15.9%	37.8%	18.3%	14.3%	13.1%	100.0%
2	A	55.2%	21%	9.8%	8.7%	5.3%	0%	100.0%
	BC	3.4%	26%	20.5%	11.1%	18.2%	20.7%	100.0%
3	A	53.3%	25.5%	9.6%	2.9%	5.3%	3.3%	100.0%
	BC	0.4%	8.5%	28.9%	21.8%	30.2%	10.1%	100.0%
4	A	56.2%	21.8%	11%	4.4%	5.7%	1%	100.0%
	BC	0.5%	11.9%	31.4%	16.5%	24.6%	15.2%	100.0%
5	A	58.5%	18.2%	13.4%	4.5%	5.4%	0%	100.0%
	BC	0.9%	15.8%	26.9%	14.8%	24.9%	16.6%	100.0%
6	A	46.2%	37%	10.3%	2.6%	2.9%	1%	100.0%
	BC	2.2%	35.2%	32.1%	12.3%	13.8%	4.3%	100.0%
7	A	56.9%	27.6%	10%	2.2%	1.5%	1.8%	100.0%
	BC	3.8%	33.2%	33.9%	11.8%	11.6%	5.7%	100.0%
8	A	57.7%	28.5%	7.4%	2.5%	3.4%	0.6%	100.0%
	BC	4.9%	33.9%	27%	13.1%	14.1%	7%	100.0%
9	A	33.2%	44.7%	12.3%	2.9%	5.2%	1.6%	100.0%
	BC	1.5%	26.3%	25.1%	17.6%	14.9%	14.6%	100.0%
10	A	34.5%	41%	13.2%	4%	3.5%	3.8%	100.0%
	BC	1.1%	29.2%	29.7%	11.5%	15.3%	13.2%	100.0%
11	A	34.6%	40.1%	11%	4.7%	8%	1.6%	100.0%
	BC	2.3%	29.9%	27.9%	6%	16.4%	17.5%	100.0%
12	A	37.1%	43.5%	10.1%	4.3%	3.6%	1.4%	100.0%
	BC	5.6%	38.3%	27.3%	5.1%	13.1%	10.5%	100.0%

1.3.1.2.3 By Grade

Table 1.3.1.2.3.1
Proficiency Level by Grade (Count): Reading

Grade	Reading Proficiency Range						Total
	1	2	3	4	5	6	
K	166,059	7,055	19,489	12,494	20,897	0	225,994
1	6,430	7,794	7,937	3,368	3,118	2,576	31,223
2	5,196	8,309	6,023	3,510	5,076	5,300	33,414
3	3,514	3,771	7,876	5,667	7,926	2,741	31,495
4	2,850	3,383	6,677	3,448	5,086	3,019	24,463
5	2,788	3,615	5,354	2,822	4,643	2,929	22,151
6	2,176	6,184	4,685	1,735	1,955	609	17,344
7	2,690	4,839	4,212	1,421	1,364	710	15,236
8	2,850	4,705	3,128	1,469	1,612	759	14,523
9	1,735	4,646	2,990	1,817	1,672	1,469	14,329
10	1,313	4,142	3,216	1,210	1,540	1,360	12,781
11	1,122	3,622	2,674	639	1,609	1,530	11,196
12	882	2,956	1,802	373	846	657	7,516

Table 1.3.1.2.3.2
Proficiency Level by Grade (Percent): Reading

Grade	Reading Proficiency Range						Total
	1	2	3	4	5	6	
K	73.5%	3.1%	8.6%	5.5%	9.2%	0.0%	100.0%
1	20.6%	25.0%	25.4%	10.8%	10.0%	8.3%	100.0%
2	15.6%	24.9%	18.0%	10.5%	15.2%	15.9%	100.0%
3	11.2%	12.0%	25.0%	18.0%	25.2%	8.7%	100.0%
4	11.7%	13.8%	27.3%	14.1%	20.8%	12.3%	100.0%
5	12.6%	16.3%	24.2%	12.7%	21.0%	13.2%	100.0%
6	12.5%	35.7%	27.0%	10.0%	11.3%	3.5%	100.0%
7	17.7%	31.8%	27.6%	9.3%	9.0%	4.7%	100.0%
8	19.6%	32.4%	21.5%	10.1%	11.1%	5.2%	100.0%
9	12.1%	32.4%	20.9%	12.7%	11.7%	10.3%	100.0%
10	10.3%	32.4%	25.2%	9.5%	12.0%	10.6%	100.0%
11	10.0%	32.4%	23.9%	5.7%	14.4%	13.7%	100.0%
12	11.7%	39.3%	24.0%	5.0%	11.3%	8.7%	100.0%

1.3.1.3 Writing

1.3.1.3.1 By Cluster by Tier

Table 1.3.1.3.1.1
Proficiency Level by Cluster (Count): Writing, S501 Paper

Cluster	Tier	Writing Proficiency Range						Total
		1	2	3	4	5	6	
K	-	147,048	40,361	30,054	8,524	0	0	225,987
1	A	8,155	10,851	959	0	0	0	19,965
	BC	2,412	6,427	11,446	486	8	0	20,779
2	A	3,885	2,951	2,285	3	0	0	9,124
	BC	1,442	6,044	20,821	2,359	28	0	30,694
3	A	2,839	3,020	1,600	5	0	0	7,464
	BC	812	2,573	23,193	3,067	23	1	29,669
4–5	A	3,102	2,755	4,558	63	0	0	10,478
	BC	573	1,000	23,658	16,401	277	11	41,920
6–8	A	6,043	4,480	2,904	88	2	0	13,517
	BC	890	1,930	26,004	12,369	34	0	41,227
9–12	A	4,730	4,314	4,096	448	1	0	13,589
	BC	2,094	2,528	17,086	16,452	403	2	38,565

Table 1.3.1.3.1.2
Proficiency Level by Cluster (Percent): Writing, S501 Paper

Cluster	Tier	Writing Proficiency Range						Total
		1	2	3	4	5	6	
K	-	65.1%	17.9%	13.3%	3.8%	0.0%	0.0%	100.0%
1	A	40.8%	54.4%	4.8%	0.0%	0.0%	0.0%	100.0%
	BC	11.6%	30.9%	55.1%	2.3%	0.0%	0.0%	100.0%
2	A	42.6%	32.3%	25.0%	0.0%	0.0%	0.0%	100.0%
	BC	4.7%	19.7%	67.8%	7.7%	0.1%	0.0%	100.0%
3	A	38.0%	40.5%	21.4%	0.1%	0.0%	0.0%	100.0%
	BC	2.7%	8.7%	78.2%	10.3%	0.1%	0.0%	100.0%
4–5	A	29.6%	26.3%	43.5%	0.6%	0.0%	0.0%	100.0%
	BC	1.4%	2.4%	56.4%	39.1%	0.7%	0.0%	100.0%
6–8	A	44.7%	33.1%	21.5%	0.7%	0.0%	0.0%	100.0%
	BC	2.2%	4.7%	63.1%	30.0%	0.1%	0.0%	100.0%
9–12	A	34.8%	31.7%	30.1%	3.3%	0.0%	0.0%	100.0%
	BC	5.4%	6.6%	44.3%	42.7%	1.0%	0.0%	100.0%

1.3.1.3.2 By Grade by Tier

Table 1.3.1.3.2.1
Proficiency Level by Grade (Count): Writing, S501 Paper

Grade	Tier	Writing Proficiency Range						Total
		1	2	3	4	5	6	
K	-	147,048	40,361	30,054	8,524	0	0	225,987
1	A	8,155	10,851	959	0	0	0	19,965
	BC	2,412	6,427	11,446	486	8	0	20,779
2	A	3,885	2,951	2,285	3	0	0	9,124
	BC	1,442	6,044	20,821	2,359	28	0	30,694
3	A	2,839	3,020	1,600	5	0	0	7,464
	BC	812	2,573	23,193	3,067	23	1	29,669
4	A	1,785	1,435	2,237	41	0	0	5,498
	BC	349	587	13,624	7,670	87	6	22,323
5	A	1,317	1,320	2,321	22	0	0	4,980
	BC	224	413	10,034	8,731	190	5	19,597
6	A	1,842	1,522	1,221	25	1	0	4,611
	BC	270	749	9,475	5,227	19	0	15,740
7	A	1,981	1,613	763	42	1	0	4,400
	BC	271	654	8,397	3,829	10	0	13,161
8	A	2,220	1,345	920	21	0	0	4,506
	BC	349	527	8,132	3,313	5	0	12,326
9	A	1,786	1,712	1,562	247	0	0	5,307
	BC	274	663	4,087	6,058	159	1	11,242
10	A	1,210	1,351	1,134	119	1	0	3,815
	BC	563	565	5,058	4,454	143	0	10,783
11	A	1,019	956	864	51	0	0	2,890
	BC	581	596	4,452	4,027	81	1	9,738
12	A	715	295	536	31	0	0	1,577
	BC	676	704	3,489	1,913	20	0	6,802

Table 1.3.1.3.2.2
Proficiency Level by Grade (Percent): Writing, S501 Paper

Grade	Tier	Writing Proficiency Range						Total
		1	2	3	4	5	6	
K	-	65.1%	17.9%	13.3%	3.8%	0.0%	0.0%	100.0%
1	A	40.8%	54.4%	4.8%	0.0%	0.0%	0.0%	100.0%
	BC	11.6%	30.9%	55.1%	2.3%	0.0%	0.0%	100.0%
2	A	42.6%	32.3%	25.0%	0.0%	0.0%	0.0%	100.0%
	BC	4.7%	19.7%	67.8%	7.7%	0.1%	0.0%	100.0%
3	A	38.0%	40.5%	21.4%	0.1%	0.0%	0.0%	100.0%
	BC	2.7%	8.7%	78.2%	10.3%	0.1%	0.0%	100.0%
4	A	32.5%	26.1%	40.7%	0.7%	0.0%	0.0%	100.0%
	BC	1.6%	2.6%	61.0%	34.4%	0.4%	0.0%	100.0%
5	A	26.4%	26.5%	46.6%	0.4%	0.0%	0.0%	100.0%
	BC	1.1%	2.1%	51.2%	44.6%	1.0%	0.0%	100.0%
6	A	39.9%	33.0%	26.5%	0.5%	0.0%	0.0%	100.0%
	BC	1.7%	4.8%	60.2%	33.2%	0.1%	0.0%	100.0%
7	A	45.0%	36.7%	17.3%	1.0%	0.0%	0.0%	100.0%
	BC	2.1%	5.0%	63.8%	29.1%	0.1%	0.0%	100.0%
8	A	49.3%	29.8%	20.4%	0.5%	0.0%	0.0%	100.0%
	BC	2.8%	4.3%	66.0%	26.9%	0.0%	0.0%	100.0%
9	A	33.7%	32.3%	29.4%	4.7%	0.0%	0.0%	100.0%
	BC	2.4%	5.9%	36.4%	53.9%	1.4%	0.0%	100.0%
10	A	31.7%	35.4%	29.7%	3.1%	0.0%	0.0%	100.0%
	BC	5.2%	5.2%	46.9%	41.3%	1.3%	0.0%	100.0%
11	A	35.3%	33.1%	29.9%	1.8%	0.0%	0.0%	100.0%
	BC	6.0%	6.1%	45.7%	41.4%	0.8%	0.0%	100.0%
12	A	45.3%	18.7%	34.0%	2.0%	0.0%	0.0%	100.0%
	BC	9.9%	10.3%	51.3%	28.1%	0.3%	0.0%	100.0%

1.3.1.3.3 By Grade

Table 1.3.1.3.3.1
Proficiency Level by Grade (Count): Writing

Grade	Writing Proficiency Range						Total
	1	2	3	4	5	6	
K	147,048	40,361	30,054	8,524	0	0	225,987
1	10,567	17,278	12,405	486	8	0	40,744
2	5,327	8,995	23,106	2,362	28	0	39,818
3	3,651	5,593	24,793	3,072	23	1	37,133
4	2,134	2,022	15,861	7,711	87	6	27,821
5	1,541	1,733	12,355	8,753	190	5	24,577
6	2,112	2,271	10,696	5,252	20	0	20,351
7	2,252	2,267	9,160	3,871	11	0	17,561
8	2,569	1,872	9,052	3,334	5	0	16,832
9	2,060	2,375	5,649	6,305	159	1	16,549
10	1,773	1,916	6,192	4,573	144	0	14,598
11	1,600	1,552	5,316	4,078	81	1	12,628
12	1,391	999	4,025	1,944	20	0	8,379

Table 1.3.1.3.3.2
Proficiency Level by Grade (Percent): Writing

Grade	Writing Proficiency Range						Total
	1	2	3	4	5	6	
K	65.1%	17.9%	13.3%	3.8%	0.0%	0.0%	100.0%
1	25.9%	42.4%	30.4%	1.2%	0.0%	0.0%	100.0%
2	13.4%	22.6%	58.0%	5.9%	0.1%	0.0%	100.0%
3	9.8%	15.1%	66.8%	8.3%	0.1%	0.0%	100.0%
4	7.7%	7.3%	57.0%	27.7%	0.3%	0.0%	100.0%
5	6.3%	7.1%	50.3%	35.6%	0.8%	0.0%	100.0%
6	10.4%	11.2%	52.6%	25.8%	0.1%	0.0%	100.0%
7	12.8%	12.9%	52.2%	22.0%	0.1%	0.0%	100.0%
8	15.3%	11.1%	53.8%	19.8%	0.0%	0.0%	100.0%
9	12.4%	14.4%	34.1%	38.1%	1.0%	0.0%	100.0%
10	12.1%	13.1%	42.4%	31.3%	1.0%	0.0%	100.0%
11	12.7%	12.3%	42.1%	32.3%	0.6%	0.0%	100.0%
12	16.6%	11.9%	48.0%	23.2%	0.2%	0.0%	100.0%

1.3.1.4 Speaking

1.3.1.4.1 By Cluster by Tier

Table 1.3.1.4.1.1
Proficiency Level by Cluster (Count): Speaking, S501 Paper

Cluster	Tier	Speaking Proficiency Range						Total
		1	2	3	4	5	6	
K	-	54,732	47,538	17,173	19,122	24,937	62,498	226,000
1	A	4,813	6,802	4,770	2,976	452	0	19,813
	BC	405	5,487	6,864	5,174	2,292	391	20,613
2	A	3,713	2,068	2,492	543	223	0	9,039
	BC	1,514	5,638	12,125	7,473	2,405	1,334	30,489
3	A	3,531	1,838	1,321	695	0	0	7,385
	BC	1,210	5,431	12,478	7,253	1,483	1,608	29,463
4–5	A	6,248	1,954	1,589	467	138	0	10,396
	BC	931	4,181	11,763	16,063	6,005	2,649	41,592
6–8	A	8,049	2,651	1,437	979	260	0	13,376
	BC	1,342	3,398	10,617	16,215	5,474	3,835	40,881
9–12	A	9,102	1,472	2,035	683	75	0	13,367
	BC	4,213	5,308	13,592	9,257	2,102	3,751	38,223

Table 1.3.1.4.1.2
Proficiency Level by Cluster (Percent): Speaking, S501 Paper

Cluster	Tier	Speaking Proficiency Range						Total
		1	2	3	4	5	6	
K	-	24.2%	21.0%	7.6%	8.5%	11.0%	27.7%	100.0%
1	A	24.3%	34.3%	24.1%	15.0%	2.3%	0.0%	100.0%
	BC	2.0%	26.6%	33.3%	25.1%	11.1%	1.9%	100.0%
2	A	41.1%	22.9%	27.6%	6.0%	2.5%	0.0%	100.0%
	BC	5.0%	18.5%	39.8%	24.5%	7.9%	4.4%	100.0%
3	A	47.8%	24.9%	17.9%	9.4%	0.0%	0.0%	100.0%
	BC	4.1%	18.4%	42.4%	24.6%	5.0%	5.5%	100.0%
4–5	A	60.1%	18.8%	15.3%	4.5%	1.3%	0.0%	100.0%
	BC	2.2%	10.1%	28.3%	38.6%	14.4%	6.4%	100.0%
6–8	A	60.2%	19.8%	10.7%	7.3%	1.9%	0.0%	100.0%
	BC	3.3%	8.3%	26.0%	39.7%	13.4%	9.4%	100.0%
9–12	A	68.1%	11.0%	15.2%	5.1%	0.6%	0.0%	100.0%
	BC	11.0%	13.9%	35.6%	24.2%	5.5%	9.8%	100.0%

1.3.1.4.2 By Grade by Tier

Table 1.3.1.4.2.1
Proficiency Level by Grade (Count): Speaking, S501 Paper

Grade	Tier	Speaking Proficiency Range						Total
		1	2	3	4	5	6	
K	-	54,732	47,538	17,173	19,122	24,937	62,498	226,000
1	A	4,813	6,802	4,770	2,976	452	0	19,813
	BC	405	5,487	6,864	5,174	2,292	391	20,613
2	A	3,713	2,068	2,492	543	223	0	9,039
	BC	1,514	5,638	12,125	7,473	2,405	1,334	30,489
3	A	3,531	1,838	1,321	695	0	0	7,385
	BC	1,210	5,431	12,478	7,253	1,483	1,608	29,463
4	A	3,081	1,049	1,026	238	68	0	5,462
	BC	430	1,949	6,363	8,599	3,613	1,200	22,154
5	A	3,167	905	563	229	70	0	4,934
	BC	501	2,232	5,400	7,464	2,392	1,449	19,438
6	A	2,493	1,145	487	305	132	0	4,562
	BC	315	1,190	3,557	6,973	1,771	1,782	15,588
7	A	2,789	709	466	334	61	0	4,359
	BC	423	1,291	3,327	4,824	2,206	999	13,070
8	A	2,767	797	484	340	67	0	4,455
	BC	604	917	3,733	4,418	1,497	1,054	12,223
9	A	3,876	455	650	170	75	0	5,226
	BC	1,071	1,443	3,529	3,516	550	1,037	11,146
10	A	2,558	346	587	250	0	0	3,741
	BC	1,137	1,477	4,192	2,360	553	968	10,687
11	A	1,778	336	543	190	0	0	2,847
	BC	1,156	1,412	3,279	2,218	569	1,019	9,653
12	A	890	335	255	73	0	0	1,553
	BC	849	976	2,592	1,163	430	727	6,737

Table 1.3.1.4.2.2
Proficiency Level by Grade (Percent): Speaking, S501 Paper

Grade	Tier	Speaking Proficiency Range						Total
		1	2	3	4	5	6	
K	-	24.2%	21.0%	7.6%	8.5%	11.0%	27.7%	100.0%
1	A	24.3%	34.3%	24.1%	15.0%	2.3%	0.0%	100.0%
	BC	2.0%	26.6%	33.3%	25.1%	11.1%	1.9%	100.0%
2	A	41.1%	22.9%	27.6%	6.0%	2.5%	0.0%	100.0%
	BC	5.0%	18.5%	39.8%	24.5%	7.9%	4.4%	100.0%
3	A	47.8%	24.9%	17.9%	9.4%	0.0%	0.0%	100.0%
	BC	4.1%	18.4%	42.4%	24.6%	5.0%	5.5%	100.0%
4	A	56.4%	19.2%	18.8%	4.4%	1.2%	0.0%	100.0%
	BC	1.9%	8.8%	28.7%	38.8%	16.3%	5.4%	100.0%
5	A	64.2%	18.3%	11.4%	4.6%	1.4%	0.0%	100.0%
	BC	2.6%	11.5%	27.8%	38.4%	12.3%	7.5%	100.0%
6	A	54.6%	25.1%	10.7%	6.7%	2.9%	0.0%	100.0%
	BC	2.0%	7.6%	22.8%	44.7%	11.4%	11.4%	100.0%
7	A	64.0%	16.3%	10.7%	7.7%	1.4%	0.0%	100.0%
	BC	3.2%	9.9%	25.5%	36.9%	16.9%	7.6%	100.0%
8	A	62.1%	17.9%	10.9%	7.6%	1.5%	0.0%	100.0%
	BC	4.9%	7.5%	30.5%	36.1%	12.2%	8.6%	100.0%
9	A	74.2%	8.7%	12.4%	3.3%	1.4%	0.0%	100.0%
	BC	9.6%	12.9%	31.7%	31.5%	4.9%	9.3%	100.0%
10	A	68.4%	9.2%	15.7%	6.7%	0.0%	0.0%	100.0%
	BC	10.6%	13.8%	39.2%	22.1%	5.2%	9.1%	100.0%
11	A	62.5%	11.8%	19.1%	6.7%	0.0%	0.0%	100.0%
	BC	12.0%	14.6%	34.0%	23.0%	5.9%	10.6%	100.0%
12	A	57.3%	21.6%	16.4%	4.7%	0.0%	0.0%	100.0%
	BC	12.6%	14.5%	38.5%	17.3%	6.4%	10.8%	100.0%

1.3.1.4.3 By Grade

Table 1.3.1.4.3.1
Proficiency Level by Grade (Count): Speaking

Grade	Speaking Proficiency Range						Total
	1	2	3	4	5	6	
K	54,732	47,538	17,173	19,122	24,937	62,498	226,000
1	5,218	12,289	11,634	8,150	2,744	391	40,426
2	5,227	7,706	14,617	8,016	2,628	1,334	39,528
3	4,741	7,269	13,799	7,948	1,483	1,608	36,848
4	3,511	2,998	7,389	8,837	3,681	1,200	27,616
5	3,668	3,137	5,963	7,693	2,462	1,449	24,372
6	2,808	2,335	4,044	7,278	1,903	1,782	20,150
7	3,212	2,000	3,793	5,158	2,267	999	17,429
8	3,371	1,714	4,217	4,758	1,564	1,054	16,678
9	4,947	1,898	4,179	3,686	625	1,037	16,372
10	3,695	1,823	4,779	2,610	553	968	14,428
11	2,934	1,748	3,822	2,408	569	1,019	12,500
12	1,739	1,311	2,847	1,236	430	727	8,290

Table 1.3.1.4.3.2
Proficiency Level by Grade (Percent): Speaking

Grade	Speaking Proficiency Range						Total
	1	2	3	4	5	6	
K	24.2%	21.0%	7.6%	8.5%	11.0%	27.7%	100.0%
1	12.9%	30.4%	28.8%	20.2%	6.8%	1.0%	100.0%
2	13.2%	19.5%	37.0%	20.3%	6.6%	3.4%	100.0%
3	12.9%	19.7%	37.4%	21.6%	4.0%	4.4%	100.0%
4	12.7%	10.9%	26.8%	32.0%	13.3%	4.3%	100.0%
5	15.1%	12.9%	24.5%	31.6%	10.1%	5.9%	100.0%
6	13.9%	11.6%	20.1%	36.1%	9.4%	8.8%	100.0%
7	18.4%	11.5%	21.8%	29.6%	13.0%	5.7%	100.0%
8	20.2%	10.3%	25.3%	28.5%	9.4%	6.3%	100.0%
9	30.2%	11.6%	25.5%	22.5%	3.8%	6.3%	100.0%
10	25.6%	12.6%	33.1%	18.1%	3.8%	6.7%	100.0%
11	23.5%	14.0%	30.6%	19.3%	4.6%	8.2%	100.0%
12	21.0%	15.8%	34.3%	14.9%	5.2%	8.8%	100.0%

1.3.2 Composites

Performance of composites is observed in their percentage in PL5 and 6: Comprehension (10-45%), Oral (10-30%), Overall (0-10%), and Literacy (0-5%). In Literacy and Overall, there are fewer students in PL 5 and 6 than Comprehension and Oral.

1.3.2.1 Oral

1.3.2.1.1 By Cluster by Tier

Table 1.3.2.1.1.1
Proficiency Level by Cluster (Count): Oral, S501 Paper

Cluster	Tier	Oral Language Proficiency Range						Total
		1	2	3	4	5	6	
K	-	61,439	32,546	27,636	21,007	37,831	45,538	225,997
1	A	2,261	3,577	6,206	3,271	1,162	123	16,600
	BC	110	1,406	5,714	5,906	3,660	1,216	18,012
2	A	2,315	2,045	2,409	1,133	239	0	8,141
	BC	184	2,159	8,936	10,552	5,455	1,381	28,667
3	A	1,905	2,056	1,698	894	198	0	6,751
	BC	75	1,473	8,388	11,168	5,229	1,491	27,824
4–5	A	3,802	2,815	1,871	995	202	11	9,696
	BC	98	1,156	7,716	15,669	11,346	4,258	40,243
6–8	A	6,235	3,082	1,921	939	189	7	12,373
	BC	161	1,454	6,877	15,978	10,448	4,514	39,432
9–12	A	7,653	2,454	1,652	491	34	0	12,284
	BC	1,259	4,277	12,310	12,052	4,834	1,796	36,528

Table 1.3.2.1.1.2
Proficiency Level by Cluster (Percent): Oral, S501 Paper

Cluster	Tier	Oral Language Proficiency Range						Total
		1	2	3	4	5	6	
K	-	27.2%	14.4%	12.2%	9.3%	16.7%	20.1%	100.0%
1	A	13.6%	21.5%	37.4%	19.7%	7.0%	0.7%	100.0%
	BC	0.6%	7.8%	31.7%	32.8%	20.3%	6.8%	100.0%
2	A	28.4%	25.1%	29.6%	13.9%	2.9%	0.0%	100.0%
	BC	0.6%	7.5%	31.2%	36.8%	19.0%	4.8%	100.0%
3	A	28.2%	30.5%	25.2%	13.2%	2.9%	0.0%	100.0%
	BC	0.3%	5.3%	30.1%	40.1%	18.8%	5.4%	100.0%
4–5	A	39.2%	29.0%	19.3%	10.3%	2.1%	0.1%	100.0%
	BC	0.2%	2.9%	19.2%	38.9%	28.2%	10.6%	100.0%
6–8	A	50.4%	24.9%	15.5%	7.6%	1.5%	0.1%	100.0%
	BC	0.4%	3.7%	17.4%	40.5%	26.5%	11.4%	100.0%

9–12	A	62.3%	20.0%	13.4%	4.0%	0.3%	0.0%	100.0%
	BC	3.4%	11.7%	33.7%	33.0%	13.2%	4.9%	100.0%

1.3.2.1.2 By Grade by Tier

Table 1.3.2.1.2.1
Proficiency Level by Grade (Count): Oral, S501 Paper

Grade	Tier	Oral Language Proficiency Range						Total
		1	2	3	4	5	6	
K	-	61,439	32,546	27,636	21,007	37,831	45,538	225,997
1	A	2,261	3,577	6,206	3,271	1,162	123	16,600
	BC	110	1,406	5,714	5,906	3,660	1,216	18,012
2	A	2,315	2,045	2,409	1,133	239	0	8,141
	BC	184	2,159	8,936	10,552	5,455	1,381	28,667
3	A	1,905	2,056	1,698	894	198	0	6,751
	BC	75	1,473	8,388	11,168	5,229	1,491	27,824
4	A	1,845	1,544	989	565	125	11	5,079
	BC	40	561	4,075	8,278	5,927	2,534	21,415
5	A	1,957	1,271	882	430	77	0	4,617
	BC	58	595	3,641	7,391	5,419	1,724	18,828
6	A	1,844	1,184	700	390	97	7	4,222
	BC	26	424	2,509	6,077	4,207	1,798	15,041
7	A	2,154	944	610	290	69	0	4,067
	BC	63	503	2,256	4,979	3,431	1,416	12,648
8	A	2,237	954	611	259	23	0	4,084
	BC	72	527	2,112	4,922	2,810	1,300	11,743
9	A	2,947	1,076	530	181	19	0	4,753
	BC	162	1,046	3,287	3,837	1,629	682	10,643
10	A	2,203	601	461	176	11	0	3,452
	BC	336	1,216	3,454	3,331	1,402	484	10,223
11	A	1,615	509	408	96	4	0	2,632
	BC	375	1,149	3,109	2,987	1,187	413	9,220
12	A	888	268	253	38	0	0	1,447
	BC	386	866	2,460	1,897	616	217	6,442

Table 1.3.2.1.2.2
Proficiency Level by Grade (Percent): Oral, S501 Paper

Grade	Tier	Oral Language Proficiency Range						Total
		1	2	3	4	5	6	
K	-	27.2%	14.4%	12.2%	9.3%	16.7%	20.1%	100.0%
1	A	13.6%	21.5%	37.4%	19.7%	7.0%	0.7%	100.0%
	BC	0.6%	7.8%	31.7%	32.8%	20.3%	6.8%	100.0%
2	A	28.4%	25.1%	29.6%	13.9%	2.9%	0.0%	100.0%
	BC	0.6%	7.5%	31.2%	36.8%	19.0%	4.8%	100.0%
3	A	28.2%	30.5%	25.2%	13.2%	2.9%	0.0%	100.0%
	BC	0.3%	5.3%	30.1%	40.1%	18.8%	5.4%	100.0%
4	A	36.3%	30.4%	19.5%	11.1%	2.5%	0.2%	100.0%
	BC	0.2%	2.6%	19.0%	38.7%	27.7%	11.8%	100.0%
5	A	42.4%	27.5%	19.1%	9.3%	1.7%	0.0%	100.0%
	BC	0.3%	3.2%	19.3%	39.3%	28.8%	9.2%	100.0%
6	A	43.7%	28.0%	16.6%	9.2%	2.3%	0.2%	100.0%
	BC	0.2%	2.8%	16.7%	40.4%	28.0%	12.0%	100.0%
7	A	53.0%	23.2%	15.0%	7.1%	1.7%	0.0%	100.0%
	BC	0.5%	4.0%	17.8%	39.4%	27.1%	11.2%	100.0%
8	A	54.8%	23.4%	15.0%	6.3%	0.6%	0.0%	100.0%
	BC	0.6%	4.5%	18.0%	41.9%	23.9%	11.1%	100.0%
9	A	62.0%	22.6%	11.2%	3.8%	0.4%	0.0%	100.0%
	BC	1.5%	9.8%	30.9%	36.1%	15.3%	6.4%	100.0%
10	A	63.8%	17.4%	13.4%	5.1%	0.3%	0.0%	100.0%
	BC	3.3%	11.9%	33.8%	32.6%	13.7%	4.7%	100.0%
11	A	61.4%	19.3%	15.5%	3.6%	0.2%	0.0%	100.0%
	BC	4.1%	12.5%	33.7%	32.4%	12.9%	4.5%	100.0%
12	A	61.4%	18.5%	17.5%	2.6%	0.0%	0.0%	100.0%
	BC	6.0%	13.4%	38.2%	29.4%	9.6%	3.4%	100.0%

1.3.2.1.3 By Grade

Table 1.3.2.1.3.1
Proficiency Level by Grade (Count): Oral

Grade	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
K	61,439	32,546	27,636	21,007	37,831	45,538	225,997
1	2,371	4,983	11,920	9,177	4,822	1,339	34,612
2	2,499	4,204	11,345	11,685	5,694	1,381	36,808
3	1,980	3,529	10,086	12,062	5,427	1,491	34,575
4	1,885	2,105	5,064	8,843	6,052	2,545	26,494
5	2,015	1,866	4,523	7,821	5,496	1,724	23,445
6	1,870	1,608	3,209	6,467	4,304	1,805	19,263
7	2,217	1,447	2,866	5,269	3,500	1,416	16,715
8	2,309	1,481	2,723	5,181	2,833	1,300	15,827
9	3,109	2,122	3,817	4,018	1,648	682	15,396
10	2,539	1,817	3,915	3,507	1,413	484	13,675
11	1,990	1,658	3,517	3,083	1,191	413	11,852
12	1,274	1,134	2,713	1,935	616	217	7,889

Table 1.3.2.1.3.2
Proficiency Level by Grade (Percent): Oral

Grade	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
K	27.2%	14.4%	12.2%	9.3%	16.7%	20.1%	100.0%
1	6.9%	14.4%	34.4%	26.5%	13.9%	3.9%	100.0%
2	6.8%	11.4%	30.8%	31.7%	15.5%	3.8%	100.0%
3	5.7%	10.2%	29.2%	34.9%	15.7%	4.3%	100.0%
4	7.1%	7.9%	19.1%	33.4%	22.8%	9.6%	100.0%
5	8.6%	8.0%	19.3%	33.4%	23.4%	7.4%	100.0%
6	9.7%	8.3%	16.7%	33.6%	22.3%	9.4%	100.0%
7	13.3%	8.7%	17.1%	31.5%	20.9%	8.5%	100.0%
8	14.6%	9.4%	17.2%	32.7%	17.9%	8.2%	100.0%
9	20.2%	13.8%	24.8%	26.1%	10.7%	4.4%	100.0%
10	18.6%	13.3%	28.6%	25.6%	10.3%	3.5%	100.0%
11	16.8%	14.0%	29.7%	26.0%	10.0%	3.5%	100.0%
12	16.1%	14.4%	34.4%	24.5%	7.8%	2.8%	100.0%

1.3.2.2 Literacy

1.3.2.2.1 By Cluster by Tier

Table 1.3.2.2.1.1
Proficiency Level by Cluster (Count): Literacy, S501 Paper

Cluster	Tier	Literacy Proficiency Range						Total
		1	2	3	4	5	6	
K	-	163,715	28,119	23,733	10,415	0	0	225,982
1	A	6,091	6,942	2,722	14	0	0	15,769
	BC	805	3,891	9,038	1,513	173	19	15,439
2	A	3,706	2,354	1,732	30	0	0	7,822
	BC	782	4,831	14,015	5,301	625	26	25,580
3	A	2,653	2,295	1,383	61	0	0	6,392
	BC	177	1,861	16,825	5,822	377	36	25,098
4–5	A	3,941	2,749	2,470	225	2	0	9,387
	BC	224	1,062	18,224	15,553	1,970	190	37,223
6–8	A	5,800	4,239	1,920	145	5	0	12,109
	BC	427	3,127	21,899	9,058	449	11	34,971
9–12	A	4,265	4,599	2,982	548	17	0	12,411
	BC	829	3,713	14,502	11,834	2,444	51	33,373

Table 1.3.2.2.1.2
Proficiency Level by Cluster (Percent): Literacy, S501 Paper

Cluster	Tier	Literacy Proficiency Range						Total
		1	2	3	4	5	6	
K	-	72.4%	12.4%	10.5%	4.6%	0.0%	0.0%	100.0%
1	A	38.6%	44.0%	17.3%	0.1%	0.0%	0.0%	100.0%
	BC	5.2%	25.2%	58.5%	9.8%	1.1%	0.1%	100.0%
2	A	47.4%	30.1%	22.1%	0.4%	0.0%	0.0%	100.0%
	BC	3.1%	18.9%	54.8%	20.7%	2.4%	0.1%	100.0%
3	A	41.5%	35.9%	21.6%	1.0%	0.0%	0.0%	100.0%
	BC	0.7%	7.4%	67.0%	23.2%	1.5%	0.1%	100.0%
4–5	A	42.0%	29.3%	26.3%	2.4%	0.0%	0.0%	100.0%
	BC	0.6%	2.9%	49.0%	41.8%	5.3%	0.5%	100.0%
6–8	A	47.9%	35.0%	15.9%	1.2%	0.0%	0.0%	100.0%
	BC	1.2%	8.9%	62.6%	25.9%	1.3%	0.0%	100.0%
9–12	A	34.4%	37.1%	24.0%	4.4%	0.1%	0.0%	100.0%
	BC	2.5%	11.1%	43.5%	35.5%	7.3%	0.2%	100.0%

1.3.2.2.2 By Grade by Tier

Table 1.3.2.2.2.1
Proficiency Level by Grade (Count): Literacy, S501 Paper

Grade	Tier	Literacy Proficiency Range						Total
		1	2	3	4	5	6	
K	-	163,715	28,119	23,733	10,415	0	0	225,982
1	A	6,091	6,942	2,722	14	0	0	15,769
	BC	805	3,891	9,038	1,513	173	19	15,439
2	A	3,706	2,354	1,732	30	0	0	7,822
	BC	782	4,831	14,015	5,301	625	26	25,580
3	A	2,653	2,295	1,383	61	0	0	6,392
	BC	177	1,861	16,825	5,822	377	36	25,098
4	A	2,035	1,432	1,294	123	2	0	4,886
	BC	125	555	10,282	7,649	857	105	19,573
5	A	1,906	1,317	1,176	102	0	0	4,501
	BC	99	507	7,942	7,904	1,113	85	17,650
6	A	1,825	1,448	752	54	3	0	4,082
	BC	124	1,054	8,469	3,475	130	6	13,258
7	A	1,940	1,390	598	53	1	0	3,982
	BC	128	1,045	7,083	2,822	155	5	11,238
8	A	2,035	1,401	570	38	1	0	4,045
	BC	175	1,028	6,347	2,761	164	0	10,475
9	A	1,628	1,834	1,099	209	9	0	4,779
	BC	94	747	3,895	3,966	804	31	9,537
10	A	1,184	1,266	863	179	5	0	3,497
	BC	161	909	4,005	3,474	714	13	9,276
11	A	901	968	688	119	3	0	2,679
	BC	206	998	3,653	2,984	662	6	8,509
12	A	552	531	332	41	0	0	1,456
	BC	368	1,059	2,949	1,410	264	1	6,051

Table 1.3.2.2.2.2
Proficiency Level by Grade (Percent): Literacy, S501 Paper

Grade	Tier	Literacy Proficiency Range						Total
		1	2	3	4	5	6	
K	-	72.4%	12.4%	10.5%	4.6%	0.0%	0.0%	100.0%
1	A	38.6%	44.0%	17.3%	0.1%	0.0%	0.0%	100.0%
	BC	5.2%	25.2%	58.5%	9.8%	1.1%	0.1%	100.0%
2	A	47.4%	30.1%	22.1%	0.4%	0.0%	0.0%	100.0%
	BC	3.1%	18.9%	54.8%	20.7%	2.4%	0.1%	100.0%
3	A	41.5%	35.9%	21.6%	1.0%	0.0%	0.0%	100.0%
	BC	0.7%	7.4%	67.0%	23.2%	1.5%	0.1%	100.0%
4	A	41.6%	29.3%	26.5%	2.5%	0.0%	0.0%	100.0%
	BC	0.6%	2.8%	52.5%	39.1%	4.4%	0.5%	100.0%
5	A	42.3%	29.3%	26.1%	2.3%	0.0%	0.0%	100.0%
	BC	0.6%	2.9%	45.0%	44.8%	6.3%	0.5%	100.0%
6	A	44.7%	35.5%	18.4%	1.3%	0.1%	0.0%	100.0%
	BC	0.9%	7.9%	63.9%	26.2%	1.0%	0.0%	100.0%
7	A	48.7%	34.9%	15.0%	1.3%	0.0%	0.0%	100.0%
	BC	1.1%	9.3%	63.0%	25.1%	1.4%	0.0%	100.0%
8	A	50.3%	34.6%	14.1%	0.9%	0.0%	0.0%	100.0%
	BC	1.7%	9.8%	60.6%	26.4%	1.6%	0.0%	100.0%
9	A	34.1%	38.4%	23.0%	4.4%	0.2%	0.0%	100.0%
	BC	1.0%	7.8%	40.8%	41.6%	8.4%	0.3%	100.0%
10	A	33.9%	36.2%	24.7%	5.1%	0.1%	0.0%	100.0%
	BC	1.7%	9.8%	43.2%	37.5%	7.7%	0.1%	100.0%
11	A	33.6%	36.1%	25.7%	4.4%	0.1%	0.0%	100.0%
	BC	2.4%	11.7%	42.9%	35.1%	7.8%	0.1%	100.0%
12	A	37.9%	36.5%	22.8%	2.8%	0.0%	0.0%	100.0%
	BC	6.1%	17.5%	48.7%	23.3%	4.4%	0.0%	100.0%

1.3.2.2.3 By Grade

Table 1.3.2.2.3.1
Proficiency Level by Grade (Count): Literacy

Grade	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
K	163,715	28,119	23,733	10,415	0	0	225,982
1	6,896	10,833	11,760	1,527	173	19	31,208
2	4,488	7,185	15,747	5,331	625	26	33,402
3	2,830	4,156	18,208	5,883	377	36	31,490
4	2,160	1,987	11,576	7,772	859	105	24,459
5	2,005	1,824	9,118	8,006	1,113	85	22,151
6	1,949	2,502	9,221	3,529	133	6	17,340
7	2,068	2,435	7,681	2,875	156	5	15,220
8	2,210	2,429	6,917	2,799	165	0	14,520
9	1,722	2,581	4,994	4,175	813	31	14,316
10	1,345	2,175	4,868	3,653	719	13	12,773
11	1,107	1,966	4,341	3,103	665	6	11,188
12	920	1,590	3,281	1,451	264	1	7,507

Table 1.3.2.2.3.2
Proficiency Level by Grade (Percent): Literacy

Grade	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
K	72.4%	12.4%	10.5%	4.6%	0.0%	0.0%	100.0%
1	22.1%	34.7%	37.7%	4.9%	0.6%	0.1%	100.0%
2	13.4%	21.5%	47.1%	16.0%	1.9%	0.1%	100.0%
3	9.0%	13.2%	57.8%	18.7%	1.2%	0.1%	100.0%
4	8.8%	8.1%	47.3%	31.8%	3.5%	0.4%	100.0%
5	9.1%	8.2%	41.2%	36.1%	5.0%	0.4%	100.0%
6	11.2%	14.4%	53.2%	20.4%	0.8%	0.0%	100.0%
7	13.6%	16.0%	50.5%	18.9%	1.0%	0.0%	100.0%
8	15.2%	16.7%	47.6%	19.3%	1.1%	0.0%	100.0%
9	12.0%	18.0%	34.9%	29.2%	5.7%	0.2%	100.0%
10	10.5%	17.0%	38.1%	28.6%	5.6%	0.1%	100.0%
11	9.9%	17.6%	38.8%	27.7%	5.9%	0.1%	100.0%
12	12.3%	21.2%	43.7%	19.3%	3.5%	0.0%	100.0%

1.3.2.3 Comprehension

1.3.2.3.1 By Cluster by Tier

Table 1.3.2.3.1.1
Proficiency Level by Cluster (Count): Comprehension, S501 Paper

Cluster	Tier	Comprehension Proficiency Range						Total
		1	2	3	4	5	6	
K	-	145,534	18,326	21,214	10,761	24,456	5,699	225,990
1	A	2,448	4,115	4,177	1,383	1,373	459	13,955
	BC	17	518	4,180	3,291	3,657	2,360	14,023
2	A	2,456	2,349	1,242	642	600	0	7,289
	BC	77	2,527	6,618	4,107	5,737	5,440	24,506
3	A	1,414	2,558	965	394	410	234	5,975
	BC	15	350	5,347	6,069	8,163	4,169	24,113
4–5	A	3,319	2,901	1,209	608	737	151	8,925
	BC	18	1,087	7,436	7,848	11,939	8,049	36,377
6–8	A	5,061	4,170	1,371	418	349	81	11,450
	BC	143	4,108	10,538	7,876	7,892	3,579	34,136
9–12	A	4,632	4,777	1,403	494	360	17	11,683
	BC	273	5,988	10,121	5,788	6,110	4,078	32,358

Table 1.3.2.3.1.2
Proficiency Level by Cluster (Percent): Comprehension, S501 Paper

Cluster	Tier	Comprehension Proficiency Range						Total
		1	2	3	4	5	6	
K	-	64.4%	8.1%	9.4%	4.8%	10.8%	2.5%	100.0%
1	A	17.5%	29.5%	29.9%	9.9%	9.8%	3.3%	100.0%
	BC	0.1%	3.7%	29.8%	23.5%	26.1%	16.8%	100.0%
2	A	33.7%	32.2%	17.0%	8.8%	8.2%	0.0%	100.0%
	BC	0.3%	10.3%	27.0%	16.8%	23.4%	22.2%	100.0%
3	A	23.7%	42.8%	16.2%	6.6%	6.9%	3.9%	100.0%
	BC	0.1%	1.5%	22.2%	25.2%	33.9%	17.3%	100.0%
4–5	A	37.2%	32.5%	13.5%	6.8%	8.3%	1.7%	100.0%
	BC	0.0%	3.0%	20.4%	21.6%	32.8%	22.1%	100.0%
6–8	A	44.2%	36.4%	12.0%	3.7%	3.0%	0.7%	100.0%
	BC	0.4%	12.0%	30.9%	23.1%	23.1%	10.5%	100.0%
9–12	A	39.6%	40.9%	12.0%	4.2%	3.1%	0.1%	100.0%
	BC	0.8%	18.5%	31.3%	17.9%	18.9%	12.6%	100.0%

1.3.2.3.2 By Grade by Tier

Table 1.3.2.3.2.1
Proficiency Level by Grade (Count): Comprehension, S501 Paper

Grade	Tier	Comprehension Proficiency Range						Total
		1	2	3	4	5	6	
K	-	145,534	18,326	21,214	10,761	24,456	5,699	225,990
1	A	2,448	4,115	4,177	1,383	1,373	459	13,955
	BC	17	518	4,180	3,291	3,657	2,360	14,023
2	A	2,456	2,349	1,242	642	600	0	7,289
	BC	77	2,527	6,618	4,107	5,737	5,440	24,506
3	A	1,414	2,558	965	394	410	234	5,975
	BC	15	350	5,347	6,069	8,163	4,169	24,113
4	A	1,569	1,647	616	307	368	126	4,633
	BC	8	405	3,865	4,380	6,400	4,060	19,118
5	A	1,750	1,254	593	301	369	25	4,292
	BC	10	682	3,571	3,468	5,539	3,989	17,259
6	A	1,435	1,573	517	169	129	41	3,864
	BC	27	1,321	4,202	3,169	2,979	1,231	12,929
7	A	1,760	1,297	458	126	114	30	3,785
	BC	37	1,438	3,432	2,423	2,445	1,222	10,997
8	A	1,866	1,300	396	123	106	10	3,801
	BC	79	1,349	2,904	2,284	2,468	1,126	10,210
9	A	1,645	2,029	489	150	142	10	4,465
	BC	18	1,089	2,982	1,861	1,987	1,307	9,244
10	A	1,356	1,260	414	149	115	7	3,301
	BC	39	1,418	2,926	1,679	1,760	1,181	9,003
11	A	1,032	951	342	132	80	0	2,537
	BC	84	1,756	2,360	1,399	1,581	1,070	8,250
12	A	599	537	158	63	23	0	1,380
	BC	132	1,725	1,853	849	782	520	5,861

Table 1.3.2.3.2.2
Proficiency Level by Grade (Percent): Comprehension, S501 Paper

Grade	Tier	Comprehension Proficiency Range						Total
		1	2	3	4	5	6	
K	-	64.4%	8.1%	9.4%	4.8%	10.8%	2.5%	100.0%
1	A	17.5%	29.5%	29.9%	9.9%	9.8%	3.3%	100.0%
	BC	0.1%	3.7%	29.8%	23.5%	26.1%	16.8%	100.0%
2	A	33.7%	32.2%	17.0%	8.8%	8.2%	0.0%	100.0%
	BC	0.3%	10.3%	27.0%	16.8%	23.4%	22.2%	100.0%
3	A	23.7%	42.8%	16.2%	6.6%	6.9%	3.9%	100.0%
	BC	0.1%	1.5%	22.2%	25.2%	33.9%	17.3%	100.0%
4	A	33.9%	35.5%	13.3%	6.6%	7.9%	2.7%	100.0%
	BC	0.0%	2.1%	20.2%	22.9%	33.5%	21.2%	100.0%
5	A	40.8%	29.2%	13.8%	7.0%	8.6%	0.6%	100.0%
	BC	0.1%	4.0%	20.7%	20.1%	32.1%	23.1%	100.0%
6	A	37.1%	40.7%	13.4%	4.4%	3.3%	1.1%	100.0%
	BC	0.2%	10.2%	32.5%	24.5%	23.0%	9.5%	100.0%
7	A	46.5%	34.3%	12.1%	3.3%	3.0%	0.8%	100.0%
	BC	0.3%	13.1%	31.2%	22.0%	22.2%	11.1%	100.0%
8	A	49.1%	34.2%	10.4%	3.2%	2.8%	0.3%	100.0%
	BC	0.8%	13.2%	28.4%	22.4%	24.2%	11.0%	100.0%
9	A	36.8%	45.4%	11.0%	3.4%	3.2%	0.2%	100.0%
	BC	0.2%	11.8%	32.3%	20.1%	21.5%	14.1%	100.0%
10	A	41.1%	38.2%	12.5%	4.5%	3.5%	0.2%	100.0%
	BC	0.4%	15.8%	32.5%	18.6%	19.5%	13.1%	100.0%
11	A	40.7%	37.5%	13.5%	5.2%	3.2%	0.0%	100.0%
	BC	1.0%	21.3%	28.6%	17.0%	19.2%	13.0%	100.0%
12	A	43.4%	38.9%	11.4%	4.6%	1.7%	0.0%	100.0%
	BC	2.3%	29.4%	31.6%	14.5%	13.3%	8.9%	100.0%

1.3.2.3.3 By Grade

Table 1.3.2.3.3.1
Proficiency Level by Grade (Count): Comprehension

Grade	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
K	145,534	18,326	21,214	10,761	24,456	5,699	225,990
1	2,465	4,633	8,357	4,674	5,030	2,819	27,978
2	2,533	4,876	7,860	4,749	6,337	5,440	31,795
3	1,429	2,908	6,312	6,463	8,573	4,403	30,088
4	1,577	2,052	4,481	4,687	6,768	4,186	23,751
5	1,760	1,936	4,164	3,769	5,908	4,014	21,551
6	1,462	2,894	4,719	3,338	3,108	1,272	16,793
7	1,797	2,735	3,890	2,549	2,559	1,252	14,782
8	1,945	2,649	3,300	2,407	2,574	1,136	14,011
9	1,663	3,118	3,471	2,011	2,129	1,317	13,709
10	1,395	2,678	3,340	1,828	1,875	1,188	12,304
11	1,116	2,707	2,702	1,531	1,661	1,070	10,787
12	731	2,262	2,011	912	805	520	7,241

Table 1.3.2.3.3.2
Proficiency Level by Grade (Percent): Comprehension

Grade	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
K	64.4%	8.1%	9.4%	4.8%	10.8%	2.5%	100.0%
1	8.8%	16.6%	29.9%	16.7%	18.0%	10.1%	100.0%
2	8.0%	15.3%	24.7%	14.9%	19.9%	17.1%	100.0%
3	4.7%	9.7%	21.0%	21.5%	28.5%	14.6%	100.0%
4	6.6%	8.6%	18.9%	19.7%	28.5%	17.6%	100.0%
5	8.2%	9.0%	19.3%	17.5%	27.4%	18.6%	100.0%
6	8.7%	17.2%	28.1%	19.9%	18.5%	7.6%	100.0%
7	12.2%	18.5%	26.3%	17.2%	17.3%	8.5%	100.0%
8	13.9%	18.9%	23.6%	17.2%	18.4%	8.1%	100.0%
9	12.1%	22.7%	25.3%	14.7%	15.5%	9.6%	100.0%
10	11.3%	21.8%	27.1%	14.9%	15.2%	9.7%	100.0%
11	10.3%	25.1%	25.0%	14.2%	15.4%	9.9%	100.0%
12	10.1%	31.2%	27.8%	12.6%	11.1%	7.2%	100.0%

1.3.2.4 Overall

1.3.2.4.1 By Cluster by Tier

Table 1.3.2.4.1.1
Proficiency Level by Grade-Level Cluster (Count): Overall, S501 Paper

Cluster	Tier	Overall Proficiency Range						Total
		1	2	3	4	5	6	
K	-	126,488	39,355	33,870	22,566	3,699	0	225,978
1	A	3,084	5,747	4,859	162	0	0	13,852
	BC	398	1,591	8,813	2,674	398	29	13,903
2	A	2,614	2,516	1,950	130	0	0	7,210
	BC	235	2,797	12,678	7,491	1,101	41	24,343
3	A	2,027	2,199	1,509	177	1	0	5,913
	BC	65	1,027	13,269	8,680	858	52	23,951
4–5	A	3,519	2,688	2,273	374	1	0	8,855
	BC	143	648	11,949	19,466	3,632	264	36,102
6–8	A	5,453	3,552	2,056	270	4	0	11,335
	BC	196	1,502	14,355	16,219	1,538	26	33,836
9–12	A	5,096	3,649	2,342	396	8	0	11,491
	BC	605	3,131	13,326	12,264	2,620	77	32,023

Table 1.3.2.4.1.2
Proficiency Level by Grade-Level Cluster (Percent): Overall, S501 Paper

Cluster	Tier	Overall Proficiency Range						Total
		1	2	3	4	5	6	
K	-	56.0%	17.4%	15.0%	10.0%	1.6%	0.0%	100.0%
1	A	22.3%	41.5%	35.1%	1.2%	0.0%	0.0%	100.0%
	BC	2.9%	11.4%	63.4%	19.2%	2.9%	0.2%	100.0%
2	A	36.3%	34.9%	27.0%	1.8%	0.0%	0.0%	100.0%
	BC	1.0%	11.5%	52.1%	30.8%	4.5%	0.2%	100.0%
3	A	34.3%	37.2%	25.5%	3.0%	0.0%	0.0%	100.0%
	BC	0.3%	4.3%	55.4%	36.2%	3.6%	0.2%	100.0%
4–5	A	39.7%	30.4%	25.7%	4.2%	0.0%	0.0%	100.0%
	BC	0.4%	1.8%	33.1%	53.9%	10.1%	0.7%	100.0%
6–8	A	48.1%	31.3%	18.1%	2.4%	0.0%	0.0%	100.0%
	BC	0.6%	4.4%	42.4%	47.9%	4.5%	0.1%	100.0%
9–12	A	44.3%	31.8%	20.4%	3.4%	0.1%	0.0%	100.0%
	BC	1.9%	9.8%	41.6%	38.3%	8.2%	0.2%	100.0%

1.3.2.4.2 By Grade by Tier

Table 1.3.2.4.2.1
Proficiency Level by Grade (Count): Overall, S501 Paper

Grade	Tier	Overall Proficiency Range						Total
		1	2	3	4	5	6	
K	-	126,488	39,355	33,870	22,566	3,699	0	225,978
1	A	3,084	5,747	4,859	162	0	0	13,852
	BC	398	1,591	8,813	2,674	398	29	13,903
2	A	2,614	2,516	1,950	130	0	0	7,210
	BC	235	2,797	12,678	7,491	1,101	41	24,343
3	A	2,027	2,199	1,509	177	1	0	5,913
	BC	65	1,027	13,269	8,680	858	52	23,951
4	A	1,772	1,418	1,196	215	1	0	4,602
	BC	76	326	6,598	10,069	1,743	161	18,973
5	A	1,747	1,270	1,077	159	0	0	4,253
	BC	67	322	5,351	9,397	1,889	103	17,129
6	A	1,628	1,269	812	115	2	0	3,826
	BC	51	459	5,380	6,355	551	13	12,809
7	A	1,882	1,141	633	96	1	0	3,753
	BC	63	504	4,709	5,099	513	9	10,897
8	A	1,943	1,142	611	59	1	0	3,756
	BC	82	539	4,266	4,765	474	4	10,130
9	A	1,956	1,461	821	155	5	0	4,398
	BC	73	588	3,527	3,973	936	50	9,147
10	A	1,443	972	698	121	2	0	3,236
	BC	146	807	3,609	3,594	734	20	8,910
11	A	1,089	768	553	87	1	0	2,498
	BC	150	837	3,439	3,070	671	7	8,174
12	A	608	448	270	33	0	0	1,359
	BC	236	899	2,751	1,627	279	0	5,792

Table 1.3.2.4.2.2
Proficiency Level by Grade (Percent): Overall, S501 Paper

Grade	Tier	Overall Proficiency Range						Total
		1	2	3	4	5	6	
K	-	56.0%	17.4%	15.0%	10.0%	1.6%	0.0%	100.0%
1	A	22.3%	41.5%	35.1%	1.2%	0.0%	0.0%	100.0%
	BC	2.9%	11.4%	63.4%	19.2%	2.9%	0.2%	100.0%
2	A	36.3%	34.9%	27.0%	1.8%	0.0%	0.0%	100.0%
	BC	1.0%	11.5%	52.1%	30.8%	4.5%	0.2%	100.0%
3	A	34.3%	37.2%	25.5%	3.0%	0.0%	0.0%	100.0%
	BC	0.3%	4.3%	55.4%	36.2%	3.6%	0.2%	100.0%
4	A	38.5%	30.8%	26.0%	4.7%	0.0%	0.0%	100.0%
	BC	0.4%	1.7%	34.8%	53.1%	9.2%	0.8%	100.0%
5	A	41.1%	29.9%	25.3%	3.7%	0.0%	0.0%	100.0%
	BC	0.4%	1.9%	31.2%	54.9%	11.0%	0.6%	100.0%
6	A	42.6%	33.2%	21.2%	3.0%	0.1%	0.0%	100.0%
	BC	0.4%	3.6%	42.0%	49.6%	4.3%	0.1%	100.0%
7	A	50.1%	30.4%	16.9%	2.6%	0.0%	0.0%	100.0%
	BC	0.6%	4.6%	43.2%	46.8%	4.7%	0.1%	100.0%
8	A	51.7%	30.4%	16.3%	1.6%	0.0%	0.0%	100.0%
	BC	0.8%	5.3%	42.1%	47.0%	4.7%	0.0%	100.0%
9	A	44.5%	33.2%	18.7%	3.5%	0.1%	0.0%	100.0%
	BC	0.8%	6.4%	38.6%	43.4%	10.2%	0.5%	100.0%
10	A	44.6%	30.0%	21.6%	3.7%	0.1%	0.0%	100.0%
	BC	1.6%	9.1%	40.5%	40.3%	8.2%	0.2%	100.0%
11	A	43.6%	30.7%	22.1%	3.5%	0.0%	0.0%	100.0%
	BC	1.8%	10.2%	42.1%	37.6%	8.2%	0.1%	100.0%
12	A	44.7%	33.0%	19.9%	2.4%	0.0%	0.0%	100.0%
	BC	4.1%	15.5%	47.5%	28.1%	4.8%	0.0%	100.0%

1.3.2.4.3 By Grade

Table 1.3.2.4.3.1
Proficiency Level by Grade (Count): Overall

Grade	Overall Proficiency Range						Total
	1	2	3	4	5	6	
K	126,488	39,355	33,870	22,566	3,699	0	225,978
1	3,482	7,338	13,672	2,836	398	29	27,755
2	2,849	5,313	14,628	7,621	1,101	41	31,553
3	2,092	3,226	14,778	8,857	859	52	29,864
4	1,848	1,744	7,794	10,284	1,744	161	23,575
5	1,814	1,592	6,428	9,556	1,889	103	21,382
6	1,679	1,728	6,192	6,470	553	13	16,635
7	1,945	1,645	5,342	5,195	514	9	14,650
8	2,025	1,681	4,877	4,824	475	4	13,886
9	2,029	2,049	4,348	4,128	941	50	13,545
10	1,589	1,779	4,307	3,715	736	20	12,146
11	1,239	1,605	3,992	3,157	672	7	10,672
12	844	1,347	3,021	1,660	279	0	7,151

Table 1.3.2.4.3.2
Proficiency Level by Grade (Percent): Overall

Grade	Overall Proficiency Range						Total
	1	2	3	4	5	6	
K	56.0%	17.4%	15.0%	10.0%	1.6%	0.0%	100.0%
1	12.5%	26.4%	49.3%	10.2%	1.4%	0.1%	100.0%
2	9.0%	16.8%	46.4%	24.2%	3.5%	0.1%	100.0%
3	7.0%	10.8%	49.5%	29.7%	2.9%	0.2%	100.0%
4	7.8%	7.4%	33.1%	43.6%	7.4%	0.7%	100.0%
5	8.5%	7.4%	30.1%	44.7%	8.8%	0.5%	100.0%
6	10.1%	10.4%	37.2%	38.9%	3.3%	0.1%	100.0%
7	13.3%	11.2%	36.5%	35.5%	3.5%	0.1%	100.0%
8	14.6%	12.1%	35.1%	34.7%	3.4%	0.0%	100.0%
9	15.0%	15.1%	32.1%	30.5%	6.9%	0.4%	100.0%
10	13.1%	14.6%	35.5%	30.6%	6.1%	0.2%	100.0%
11	11.6%	15.0%	37.4%	29.6%	6.3%	0.1%	100.0%
12	11.8%	18.8%	42.2%	23.2%	3.9%	0.0%	100.0%

2 Analysis of Domains

The measurement model that forms the basis of the analysis for the development of ACCESS for ELLs 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 developers used 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. 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 was 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 providing a correct response "1" by student "n" to item "i"
P_{ni0} = probability of providing an incorrect response "0" by student "n" to item "i"
B_n = ability of student "n"
D_i = difficulty of item "i"

When the probability of a student providing a correct answer to an item equals the probability of a student providing an incorrect answer (i.e., 50% probability of getting it right and 50% probability of getting it wrong), Pni1/Pni0 is equal to 1. The log of 1 is 0. This is the point at which a student's ability equals the difficulty of an item. For example, a student whose ability estimate 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 providing a correct answer to that item.

Rasch Model for Polytomous Scoring

For the Writing and Speaking tasks, a Rasch-grouped rating scale model, which is an extension of Andrich's rating scale model (Andrich, 1978), is used. Mathematically, this can be represented as

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

where

P_{ngik} = probability of student "n" on task "i" receiving a rating at level "k" on rating scale "g"
$P_{ngi(k-1)}$ = probability of student "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 student "n"
D_{gi} = difficulty of task "i" specific to rating scale "g"
F_{gk} = step calibration value of category "k" relative to category "k-1" 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 rating scale that was used for the group of tasks. There is only one rating scale (g = 1) in the Writing domain and two grouped rating scales (g = 2) in the Speaking domain. As with the dichotomous Rasch model, there is an item difficulty parameter (D_{gi}) for each item for rating scale "g" modeled by the Rasch rating scale model (Andrich, 1978). In addition, there is a step calibration value or *step measure* (F_{gk}) that corresponds to the location on the latent variable where the probability of being observed in the "k" and "k – 1" category for rating scale "g*"* is equal relative to the difficulty measure of the task. The step measures are also the points where adjacent category probability "k – 1" and "k" curves for rating scale "g" intercept. All tasks that belong to the same rating scale group have the same step measures.

As described in Part 1, Section 3.2.2, 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 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 of 0–9. We conducted a study in the summer of 2016 to reconstruct the logit scales. Detailed information about the derivation of the Writing rating scales as well as the psychometric properties of Writing rating scales are available in the scaling report (see Center for Applied Linguistics, 2017).

For Speaking, we model Proficiency Level 1 tasks as a group on a 0–2 scale, and PL 3 and PL 5 tasks as a group on a 0–4 scale (see Part 1, Section 3.2.4). We conducted a study in the summer of 2016 to reconstruct the logit scales and detailed information about the derivation as well as the psychometric properties of Speaking rating scales are available in the scaling report (Center for Applied Linguistics, 2017).

Scale Scores and Proficiency Level Scores

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

For Paper ACCESS Grades 1–12, 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, we established the current ACCESS scale for the original paper-only version of the test and maintained this scale through the transition to an online and paper delivered test in the 2015–2016 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, we conducted a study in the summer of 2016 to reconstruct the logit scale (see Center for Applied Linguistics, 2017).

Note that these new scales were not applied to the Kindergarten test, which is a static form. The following scaling equations are used for the Kindergarten test:

- L: (Ability Measure in Logits * 37.571) + 316.637
- R: (Ability Measure in Logits * 26.000) + 323.272
- W: (Ability Measure in Logits * 31.097) + 317.068
- S: (Ability Measure in Logits * 20.084) + 322.686

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 2015–2016 administration (Series 400) of ACCESS for ELLs. WIDA and CAL conducted a third cut score study 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 (2016–2017 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 the student's 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, PL scores are dependent upon the grade a student was in when the student took the assessment. 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, PL scores do not form an interval scale. Only scale scores should be used as interval measures. PL 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 information on the psychometric qualities of the items and tasks. We provide values for 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 Writing and Speaking, we also provide raw score distributions by task.

Tables in this section 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.

All Rasch analyses were conducted using the Rasch measurement software program *Winsteps* (Linacre, 2006). When speaking of the measure of student 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 previously presented. 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 the Rasch model would be expected to produce if the data fit the model perfectly. Outfit mean square statistics for items and tasks are influenced by outlier responses for machine-scored dichotomous items or outlier ratings for rater-scored performance tasks. For example, a difficult item that some low-ability students get correct—for reasons unknown—will have a high outfit mean square statistic. Similarly, an easy item that some high-ability students get wrong will also have a high outfit mean square statistics. Infit mean square statistics are influenced by unexpected patterns of students' responses and ratings on items and tasks 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 response and rating patterns 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 wrote:

- Values greater than 2.0 "distort or degrade[1] the measurement system."
- Values between 1.5 and 2.0 are "unproductive for construction of measurement, but not degrading."

[1] We interpret "degrade" here in the sense of lowering the quality of the measurement system.

- 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 we followed conservative guidelines 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 to 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 raters scored polytomously on a rubric scale. We design some performance tasks to elicit a restricted range of performances (for example, very easy tasks where we expect that most students will get the highest rating), and these tasks can cause the model to predict the data too well (overfitting). Conversely, when raters score performance tasks using a very wide rubric scale such as the ACCESS for ELLs Writing rubric, sometimes unmodeled noise or other sources of variance in the ratings of the students' responses to the task will cause the model to underpredict those ratings (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, according to Linacre (2002), including the rating of the student's performance on the task when calculating that student's score does not degrade the measurement of the student's performance.

Tables in this section are presented by test form (i.e., by grade cluster and tier) for Listening, Reading, and Writing. For the Speaking test, due to the design of the test, a number of items are shared between tiers. In order to best present the results of the Speaking task analysis, all Speaking items in a grade-level cluster are presented in one single table.

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 measure (in logits), the number of items, the average *p* value (for Listening and Reading only), the average infit mean square statistic, and the average outfit mean square statistic.

The second section of these tables presents results from 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 or task difficulty measure in logits. For dichotomously scored items (Listening and Reading), the next column shows the *p* value (percentage of correct answers on that item). The final two columns show the Rasch fit statistics for the item or task. Folders with items that have fit statistics greater than 2.0 are evaluated by the test development team to determine whether and when the folders can be refreshed in the next test refreshment cycle.

In addition, Writing and Speaking tables have a section at the bottom of the table that provides raw score distributions by task.

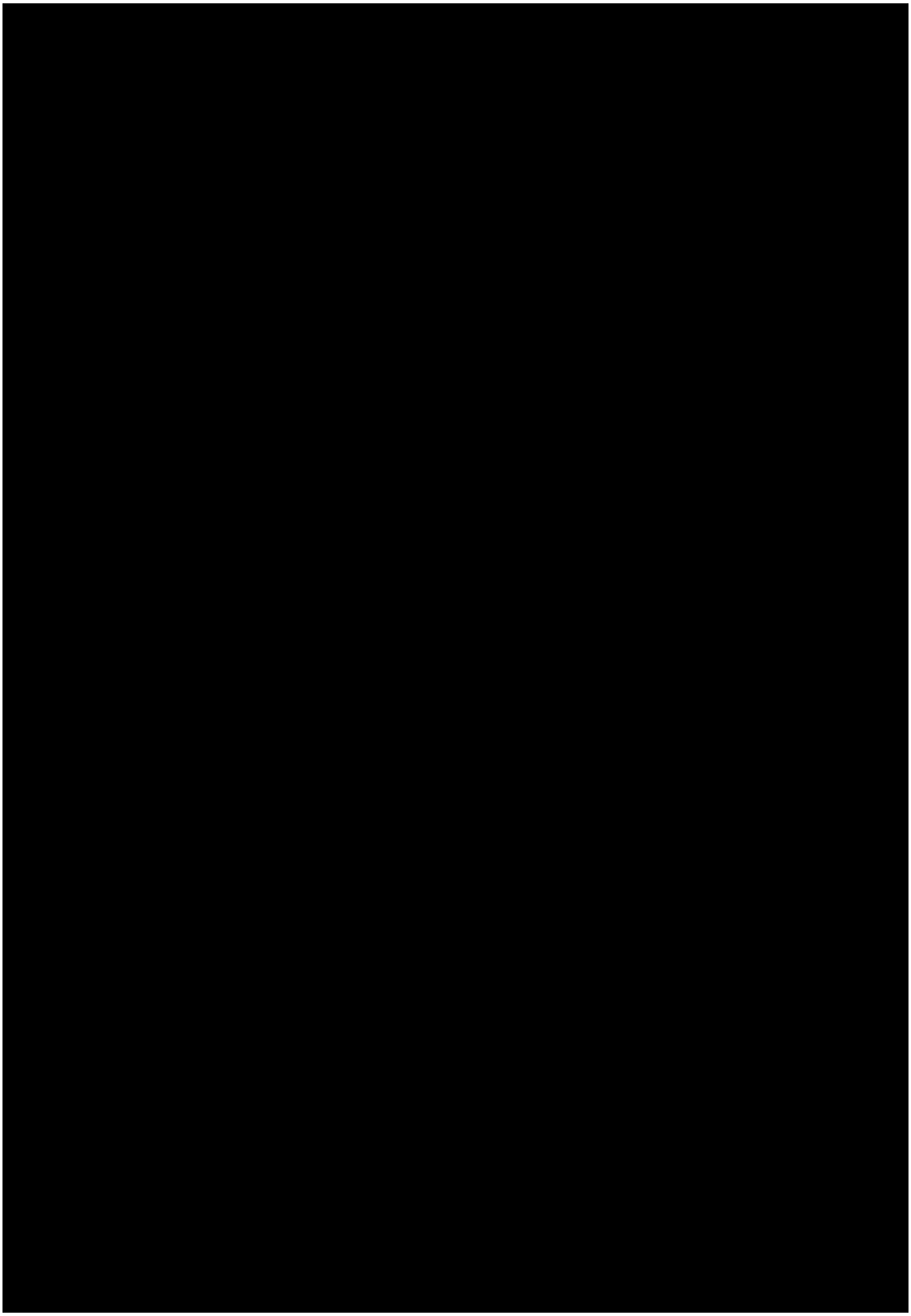
For the Grades 1–12 tests, all items and tasks across domains have infit mean square statistics less than 2, indicating that the items and tasks provide good measurement for students around the ability range that the items and tasks are targeting. One task in Writing Grade 1 has an outfit mean square statistic greater than 2. This is the easiest task for this test form, and there might be some high-ability students receiving a low rating, causing the outfit mean square statistics to be inflated.

The results show that for the Kindergarten test, all items and tasks across domains have infit mean square statistics less than 2, except for the fifth task in the Writing domain, indicating that most items and tasks provide good measurement for students around the ability range that the items and tasks are targeting. As discussed earlier, the outfit mean square statistic is sensitive to outlier responses and ratings that are not close to the ability range that the items and tasks are targeting. Four items in the Listening domain, 11 items in the Reading domain, one task in the Writing domain, and two tasks in the Speaking domain have outfit mean square statistics greater than 2. For the most part, these are very easy items or tasks (with p values > 0.85), early in the test. These outfit values are likely due to high-ability students getting these early test items incorrect. The test design includes multiple easy items at the onset of the test in order to ensure that Kindergarten students, who are often unfamiliar with standardized testing, are not presented with discouraging difficult items at the beginning of their test administration.

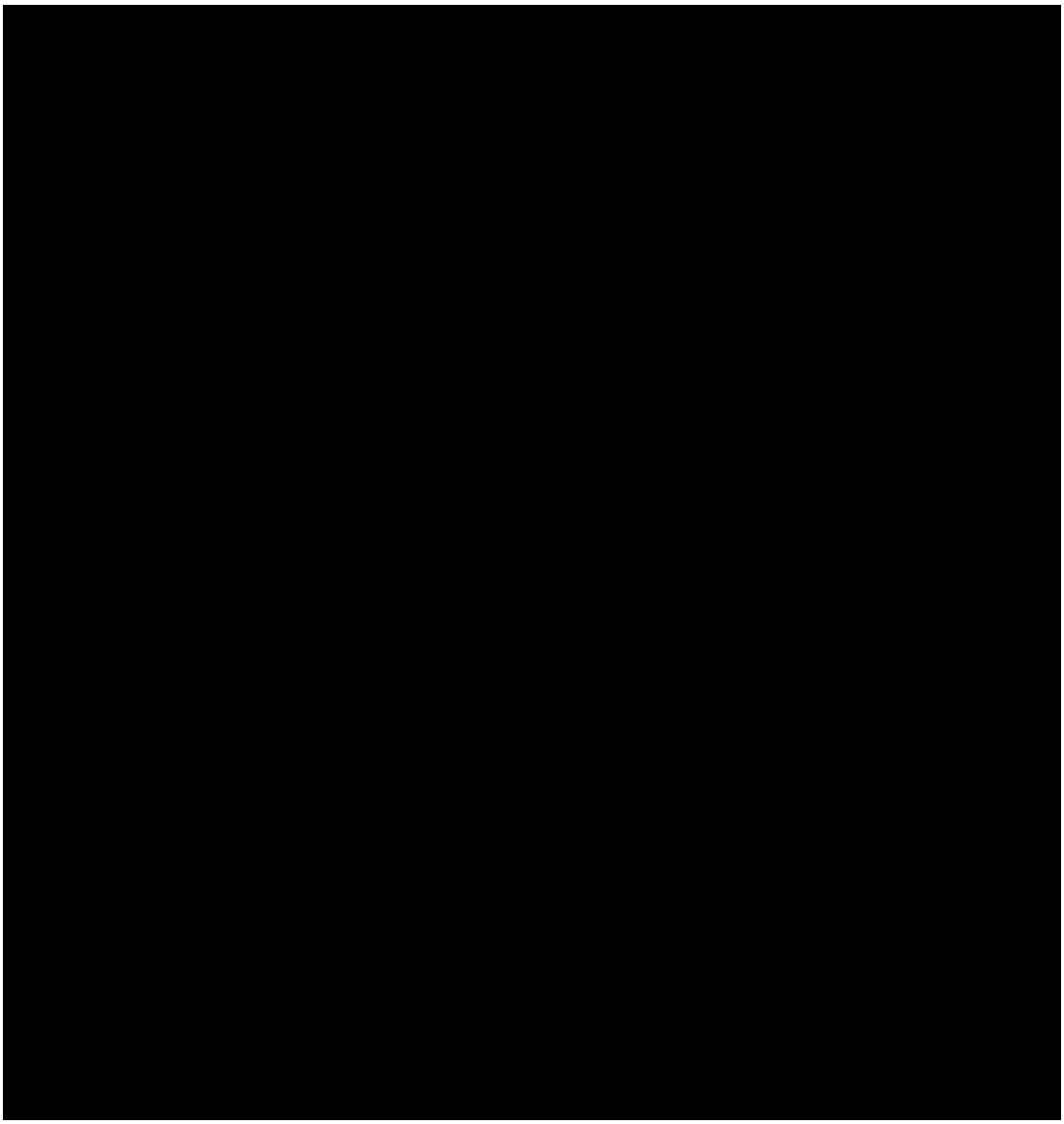
Outfit values are exceedingly high (9.90) for the first three Reading items. The Kindergarten ACCESS technical brief notes that the items in this folder are prereading items and that children with high reading ability who are not familiar with these items may not answer correctly, leading to high outfit values.

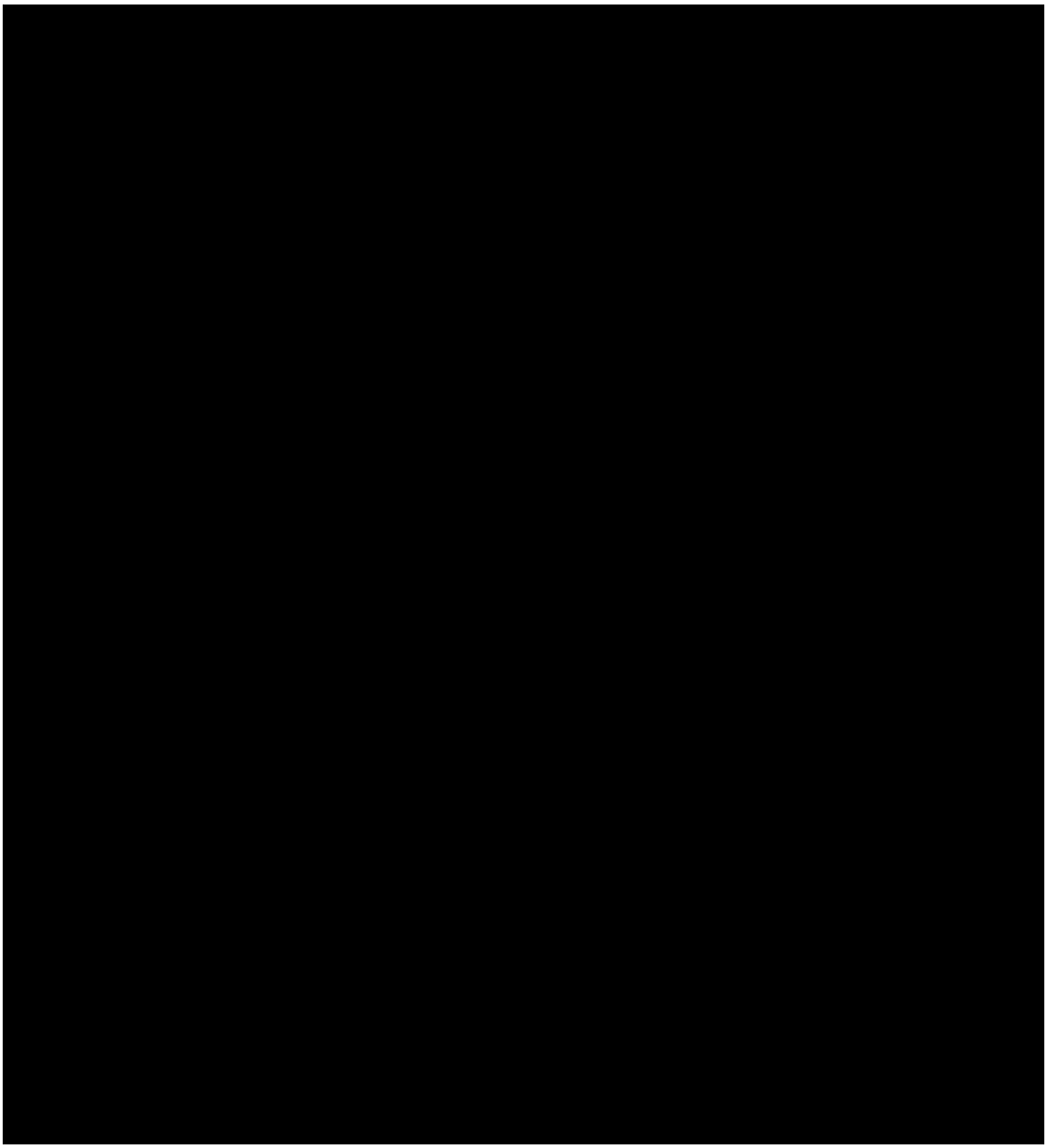
2.1.1 Listening

2.1.1.0 Kindergarten

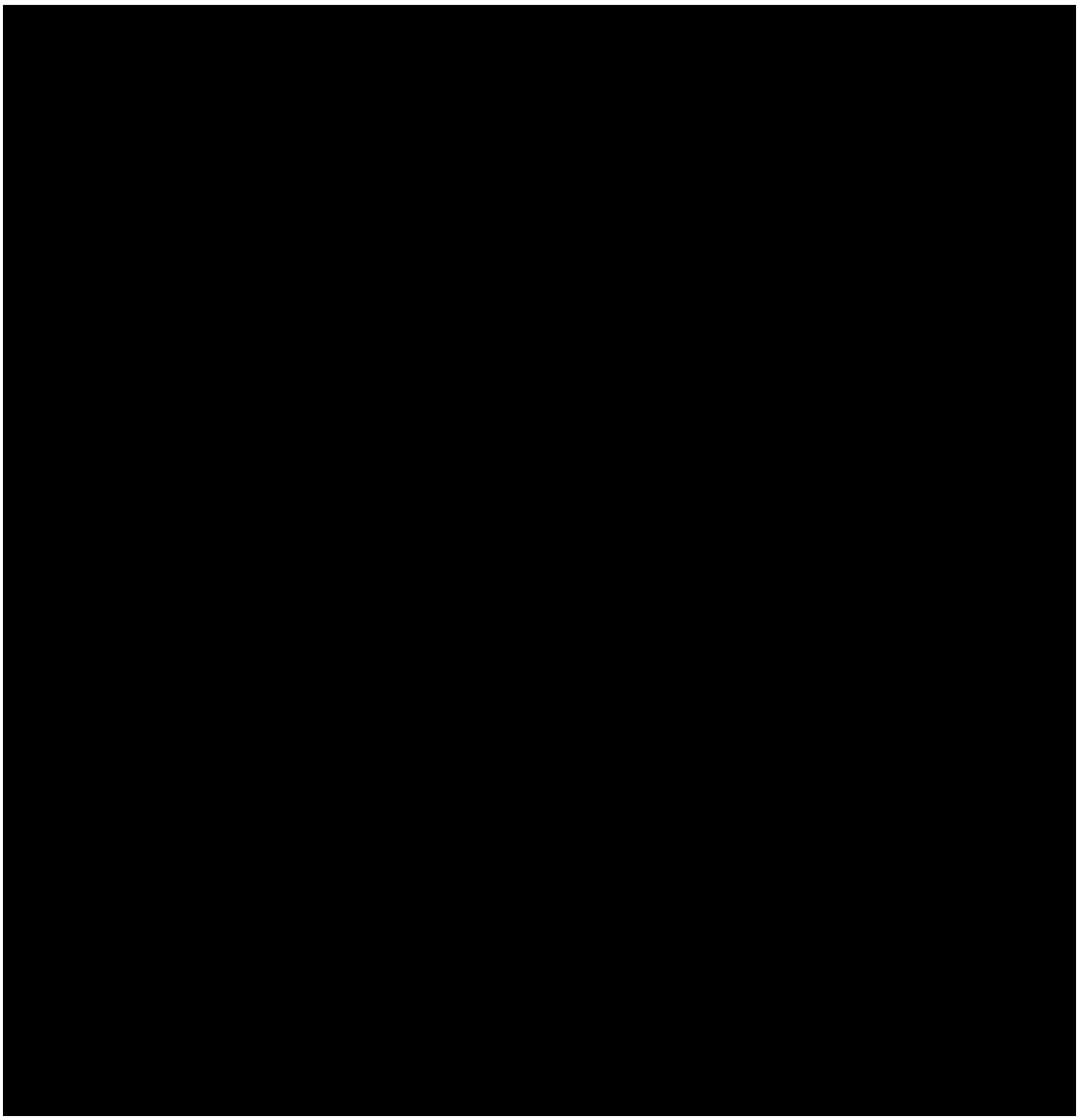


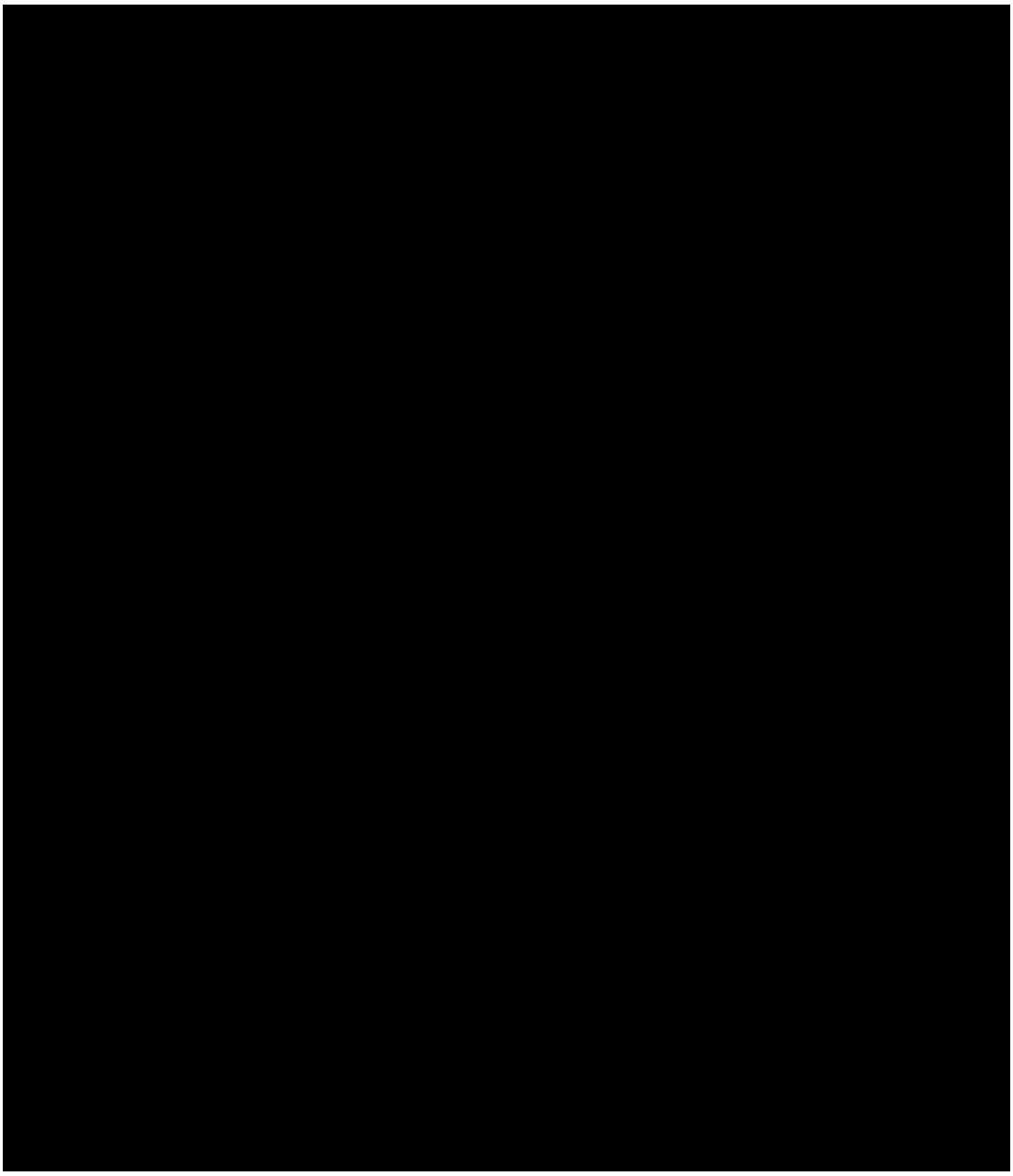
2.1.1.1 Grade 1



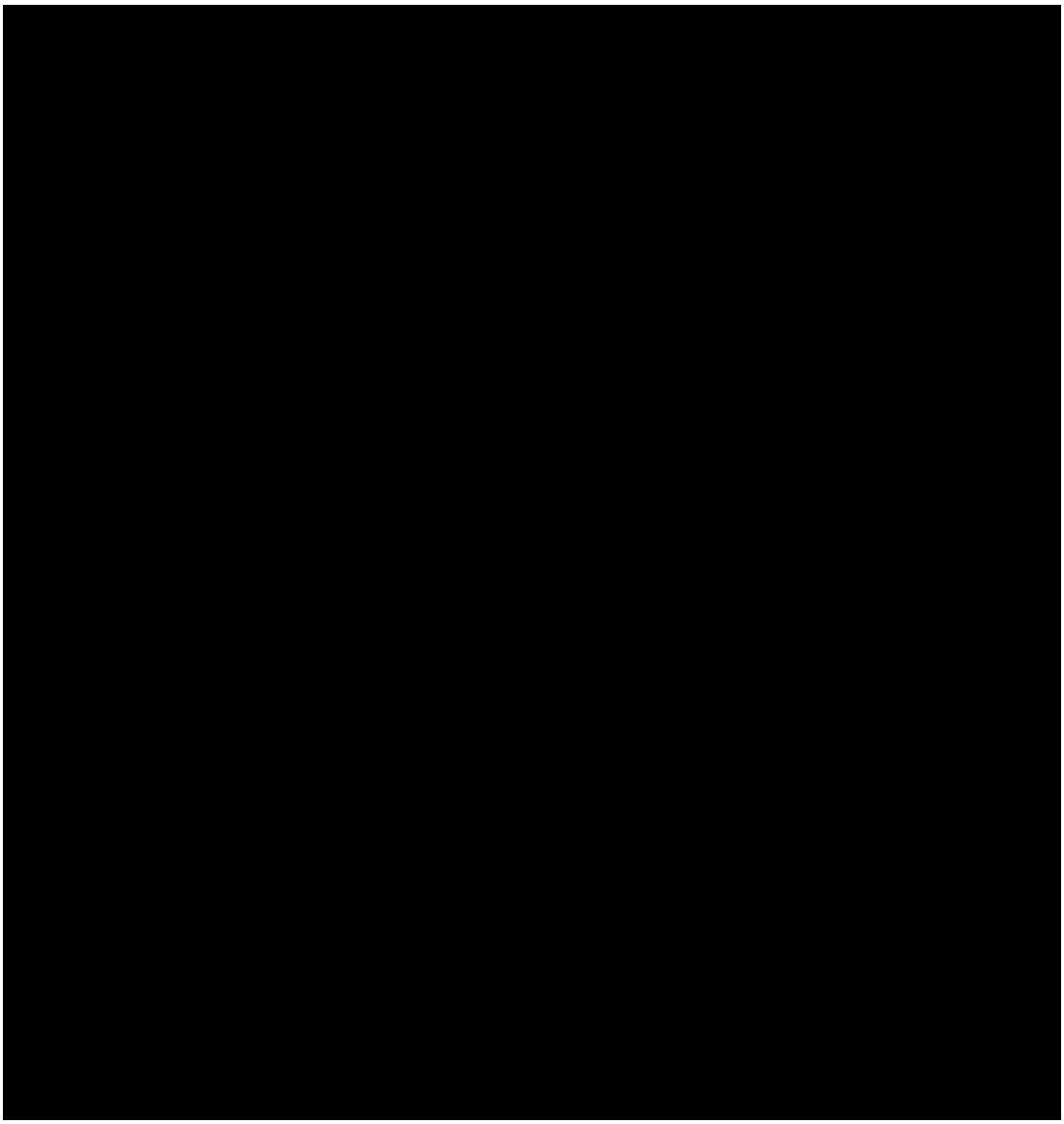


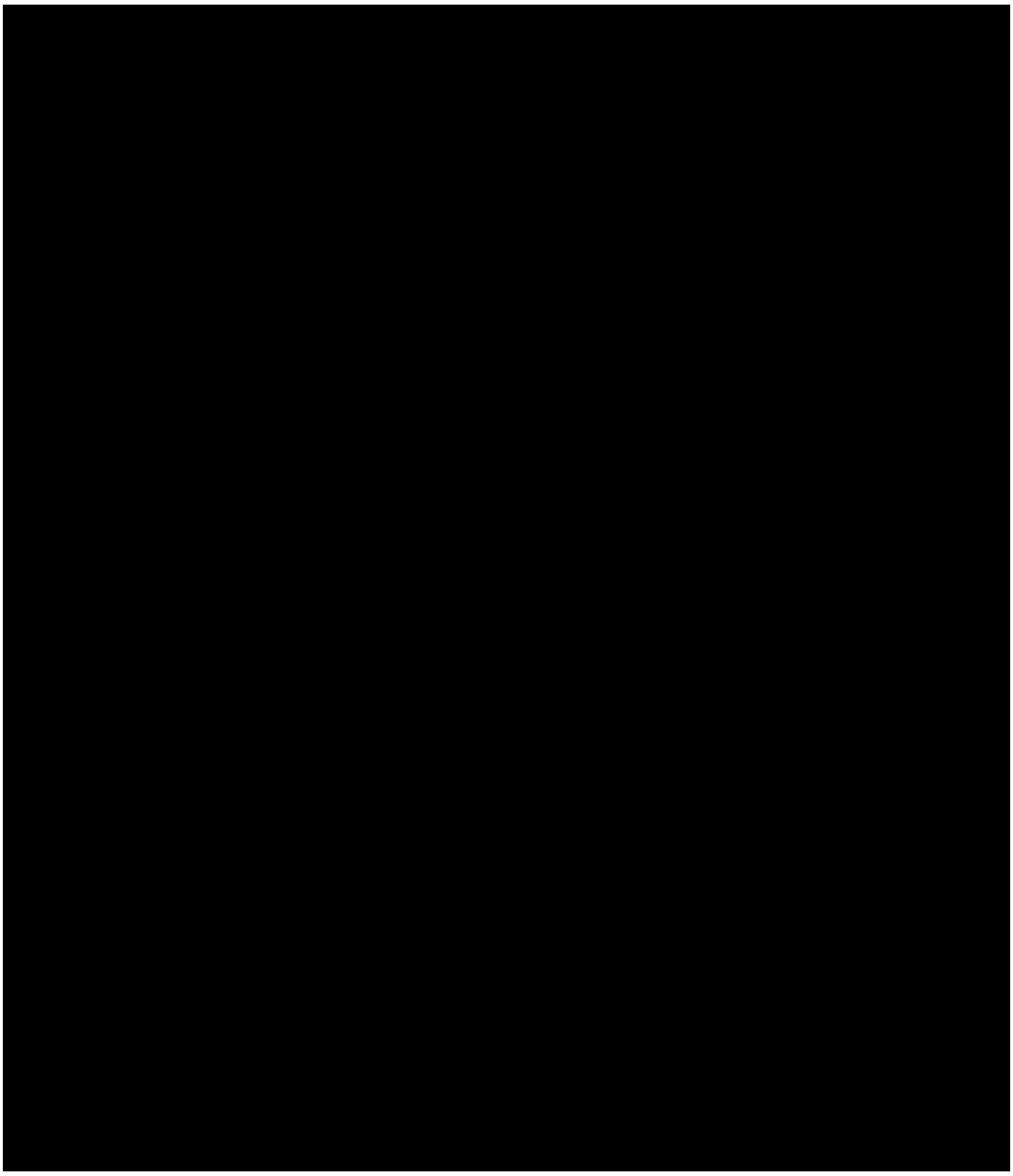
2.1.1.2 Grade 2



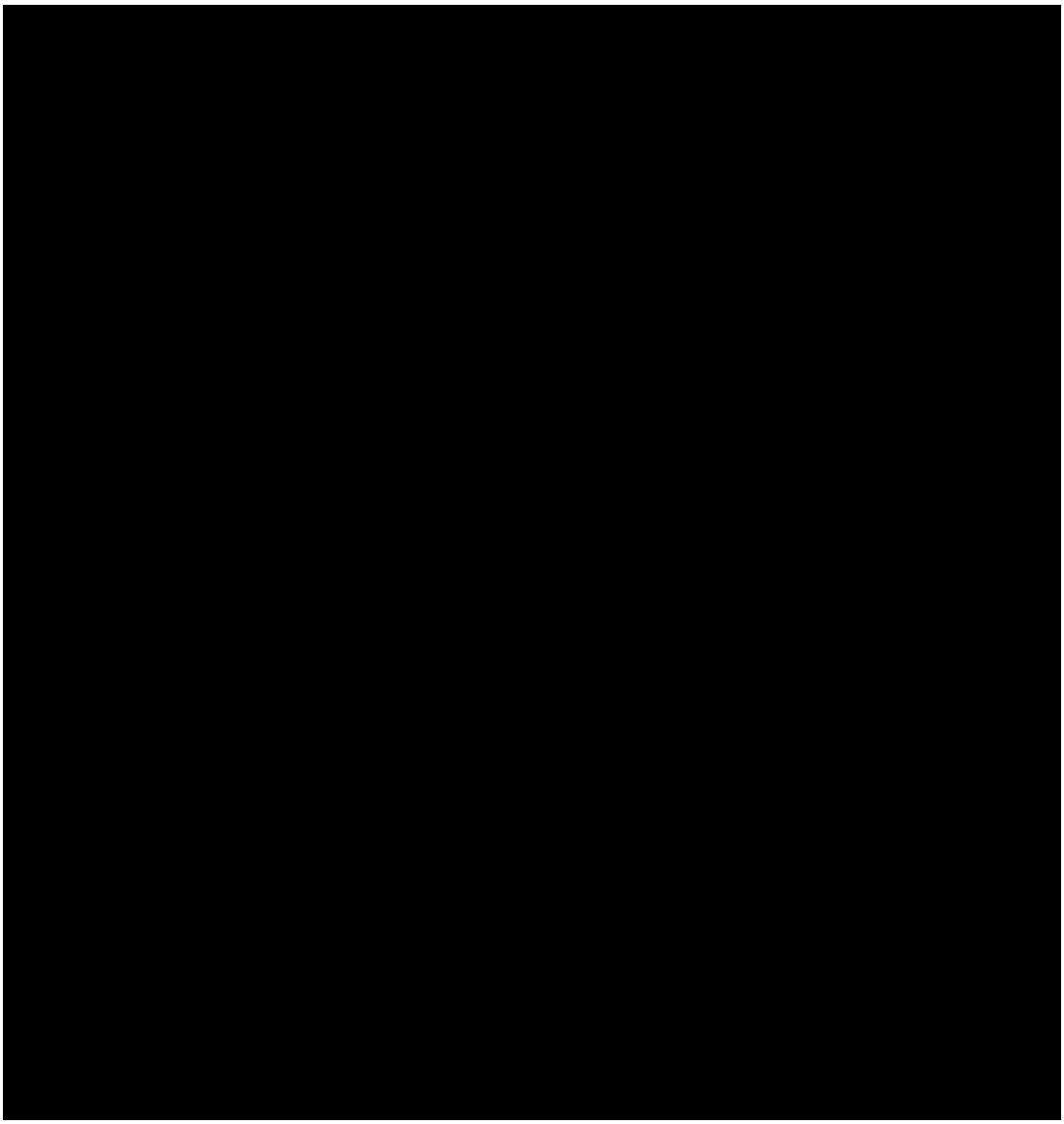


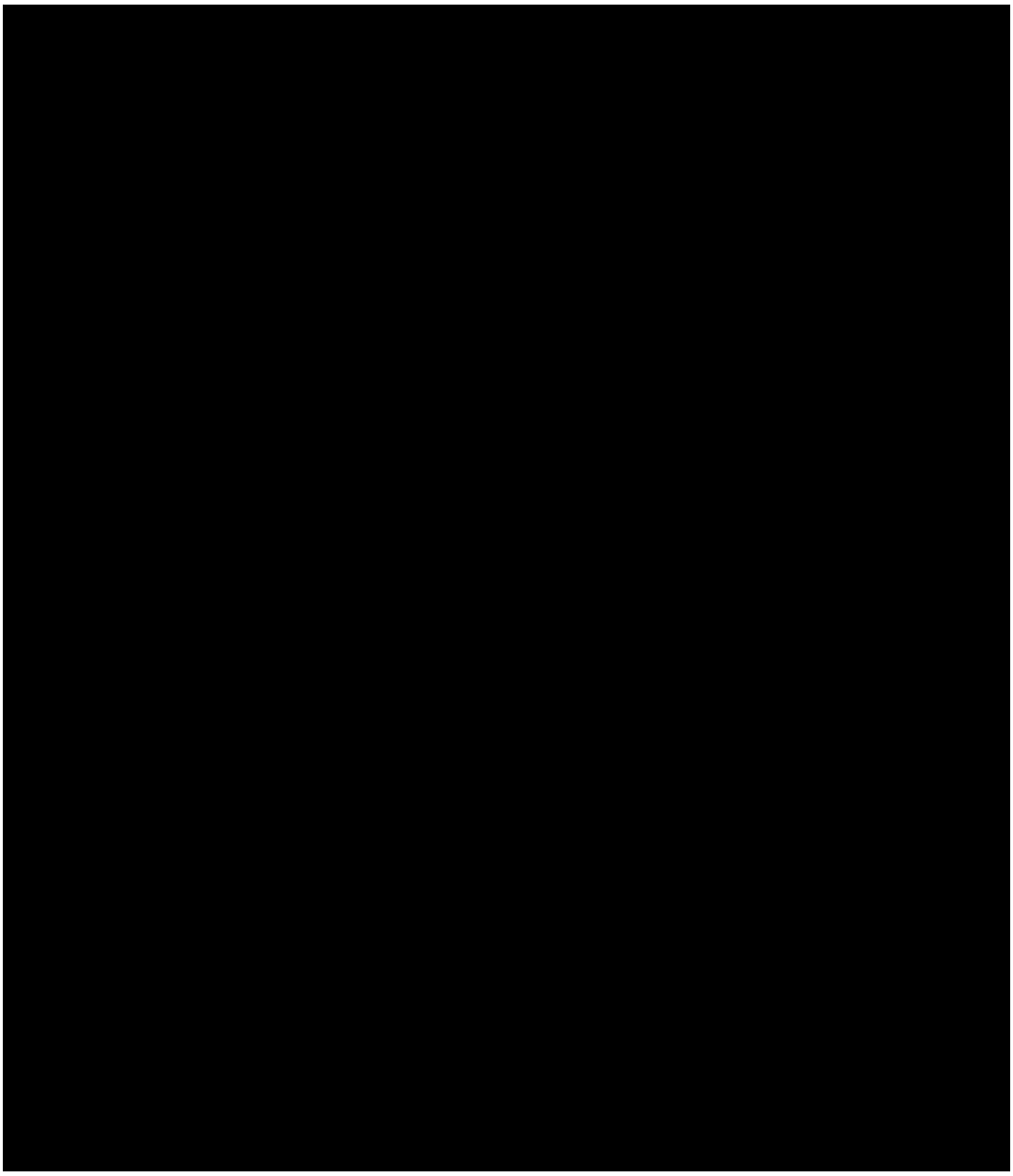
2.1.1.3 Grade 3



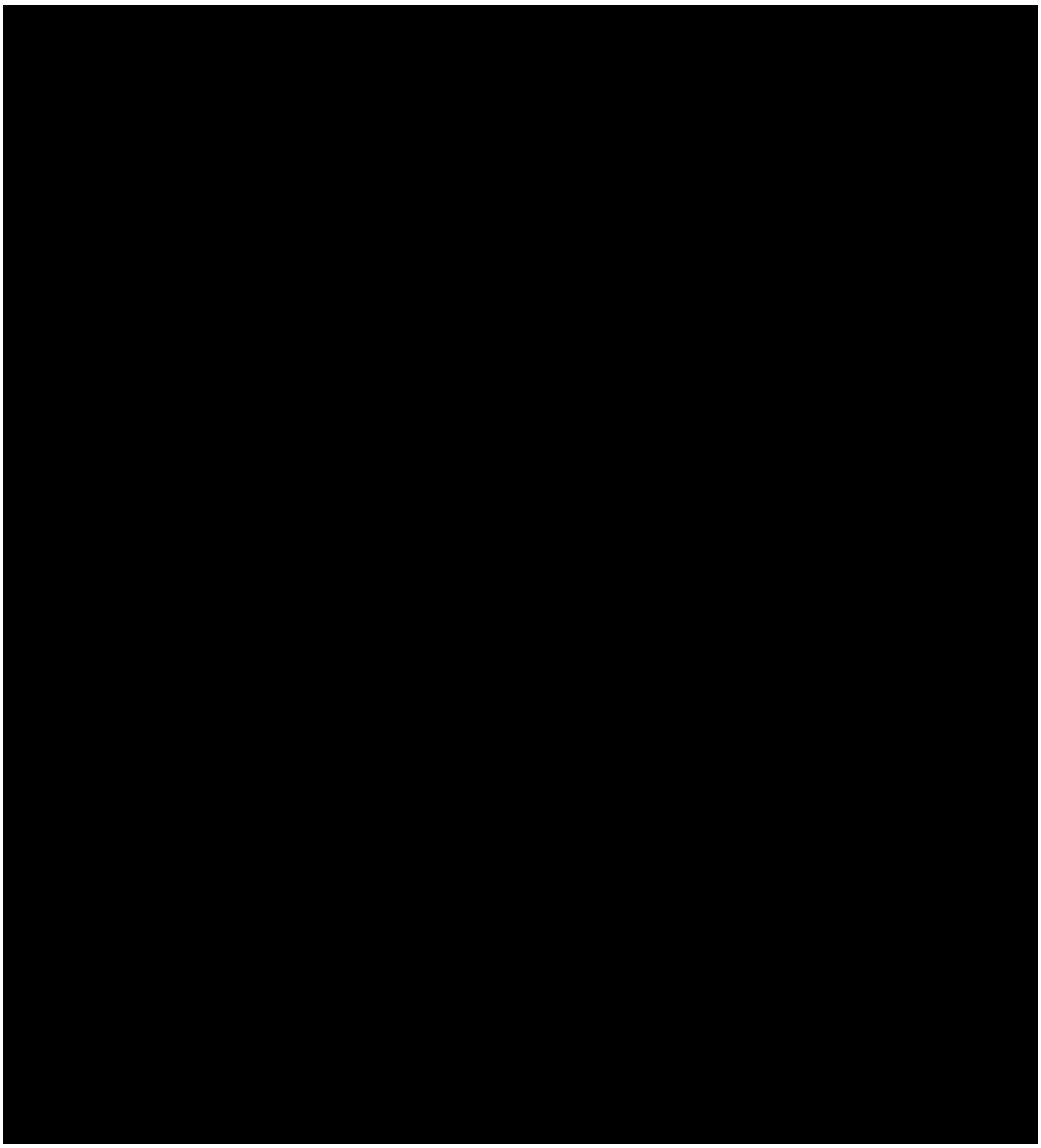


2.1.1.4 Grades 4–5

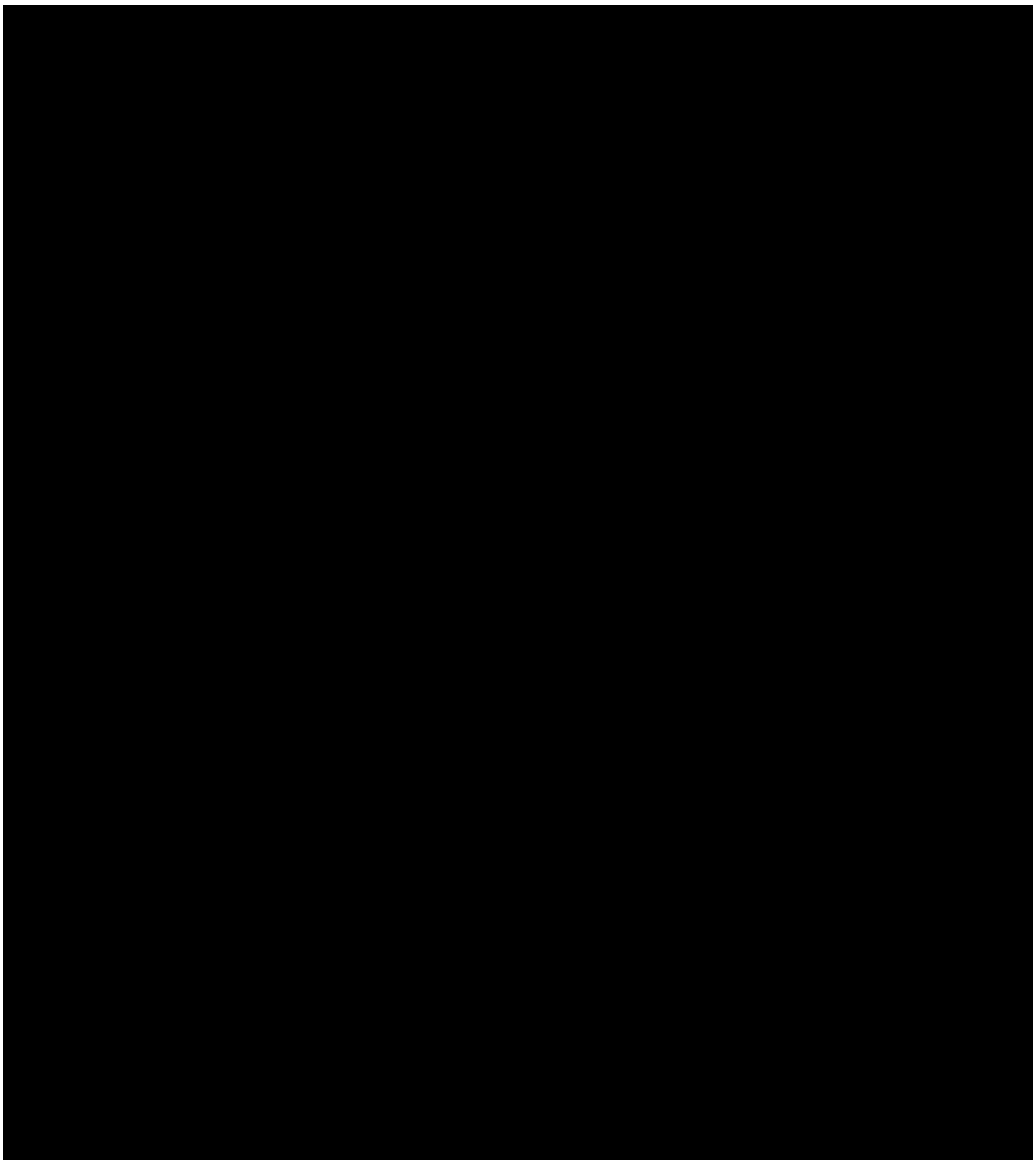




2.1.1.5 Grades 6–8

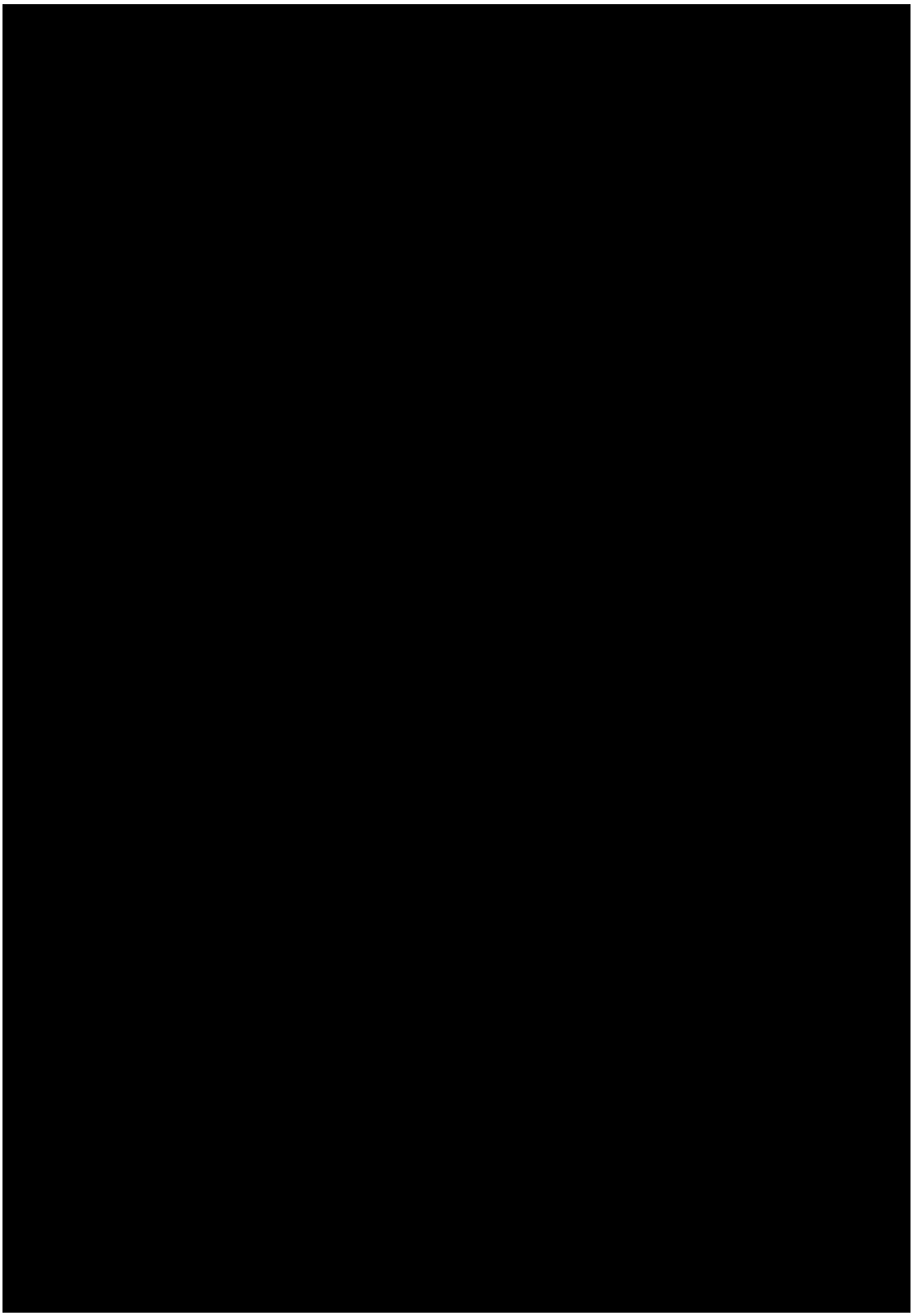


2.1.1.6 Grades 9–12

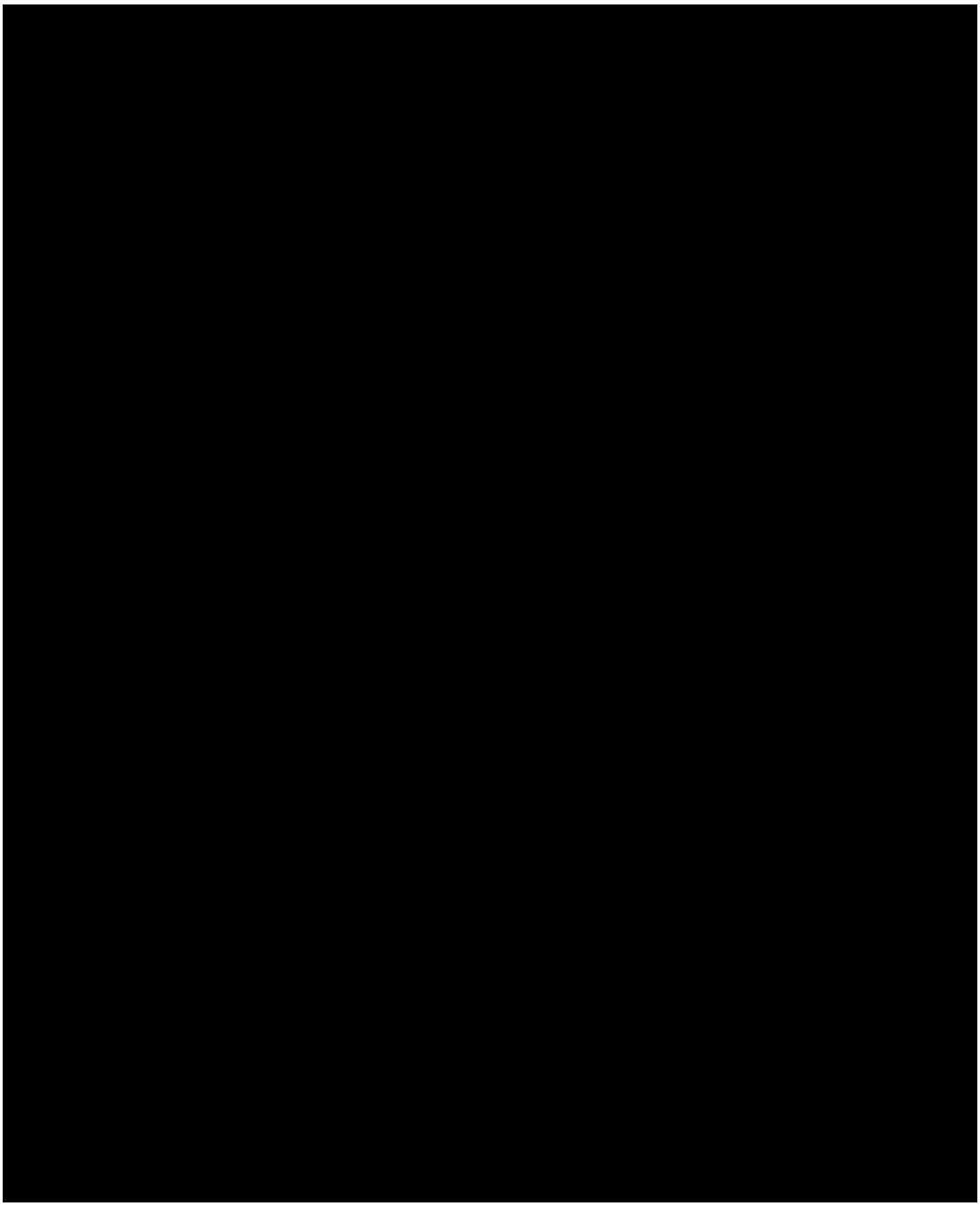


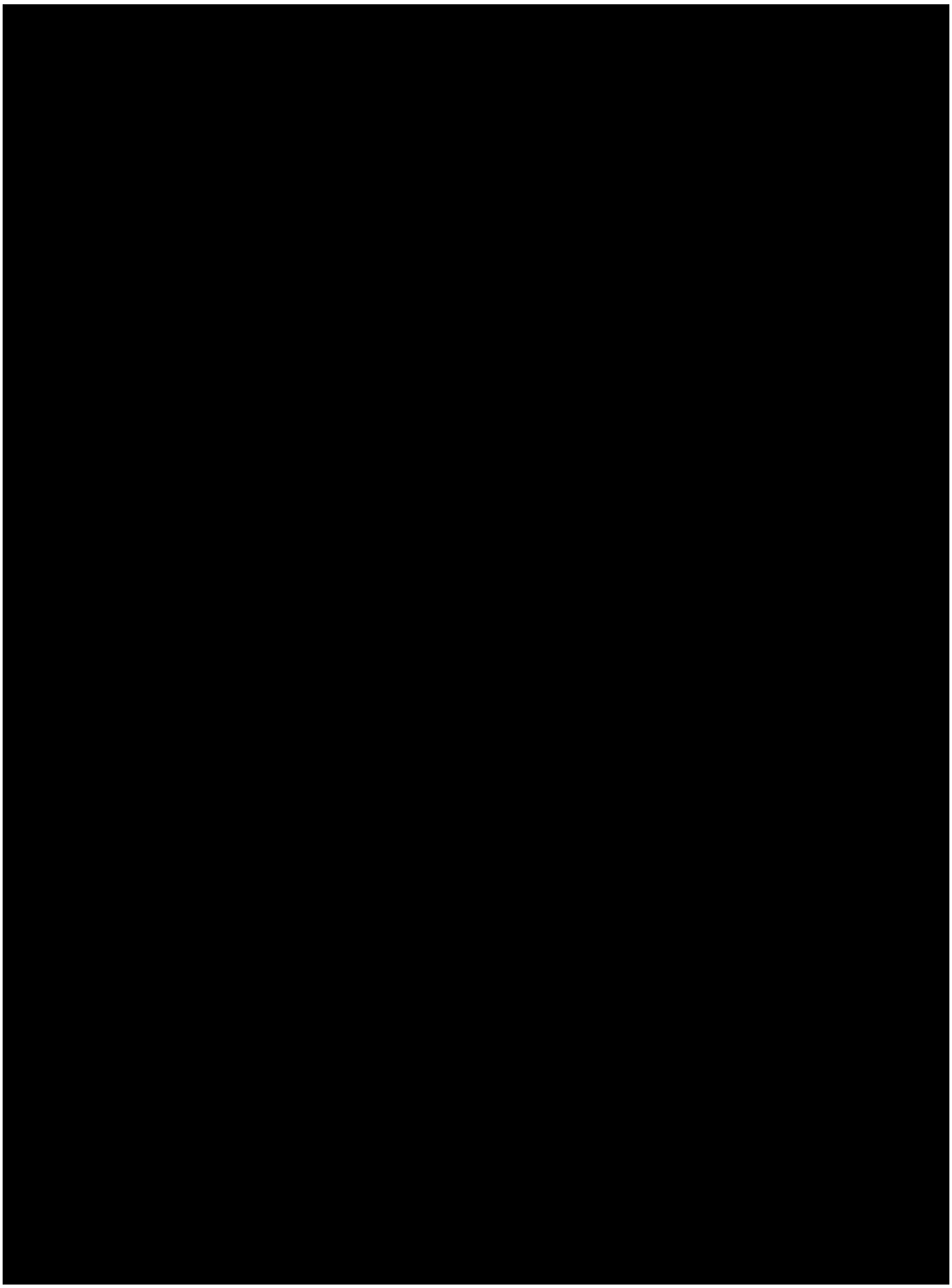
2.1.2 Reading

2.1.2.0 Kindergarten

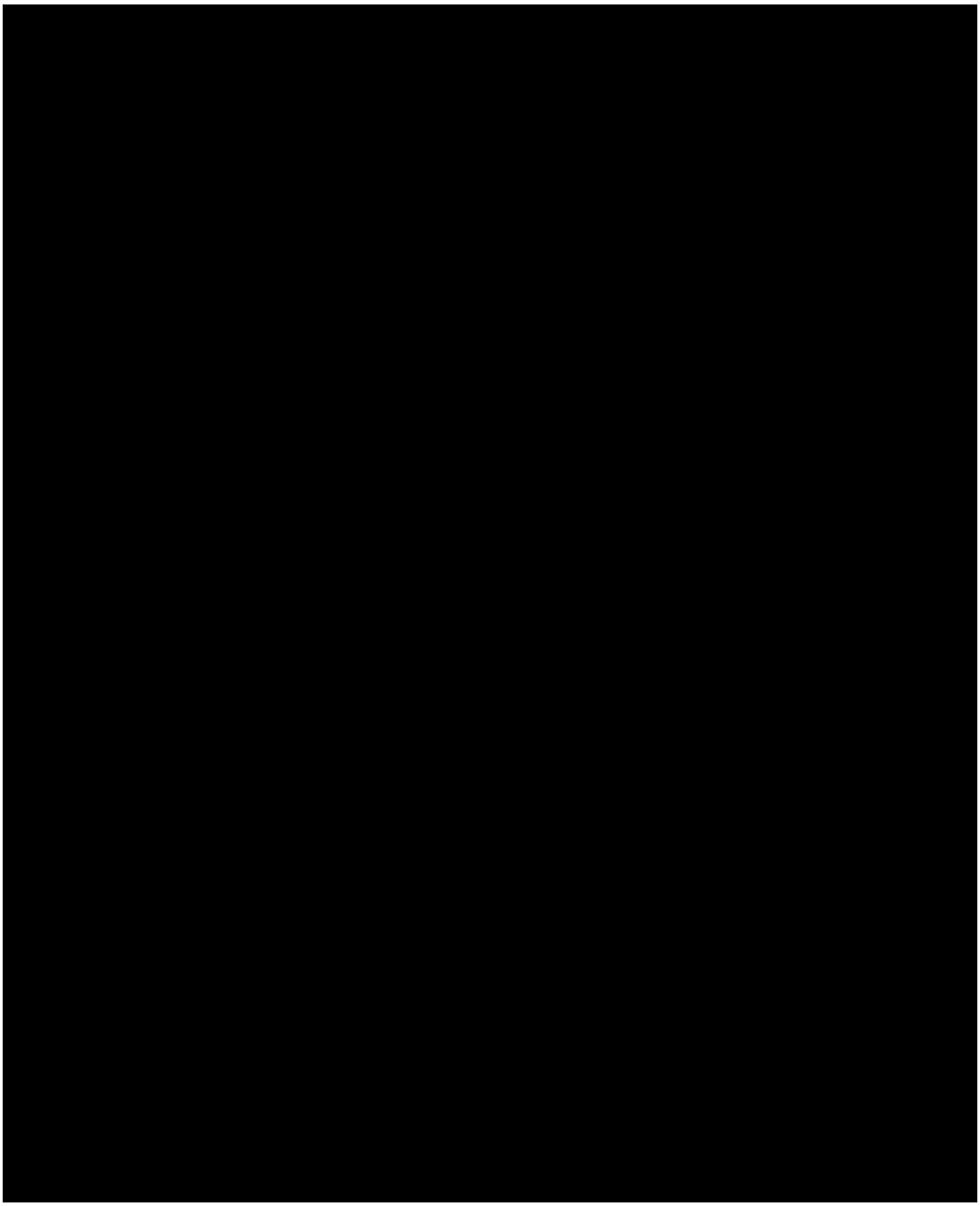


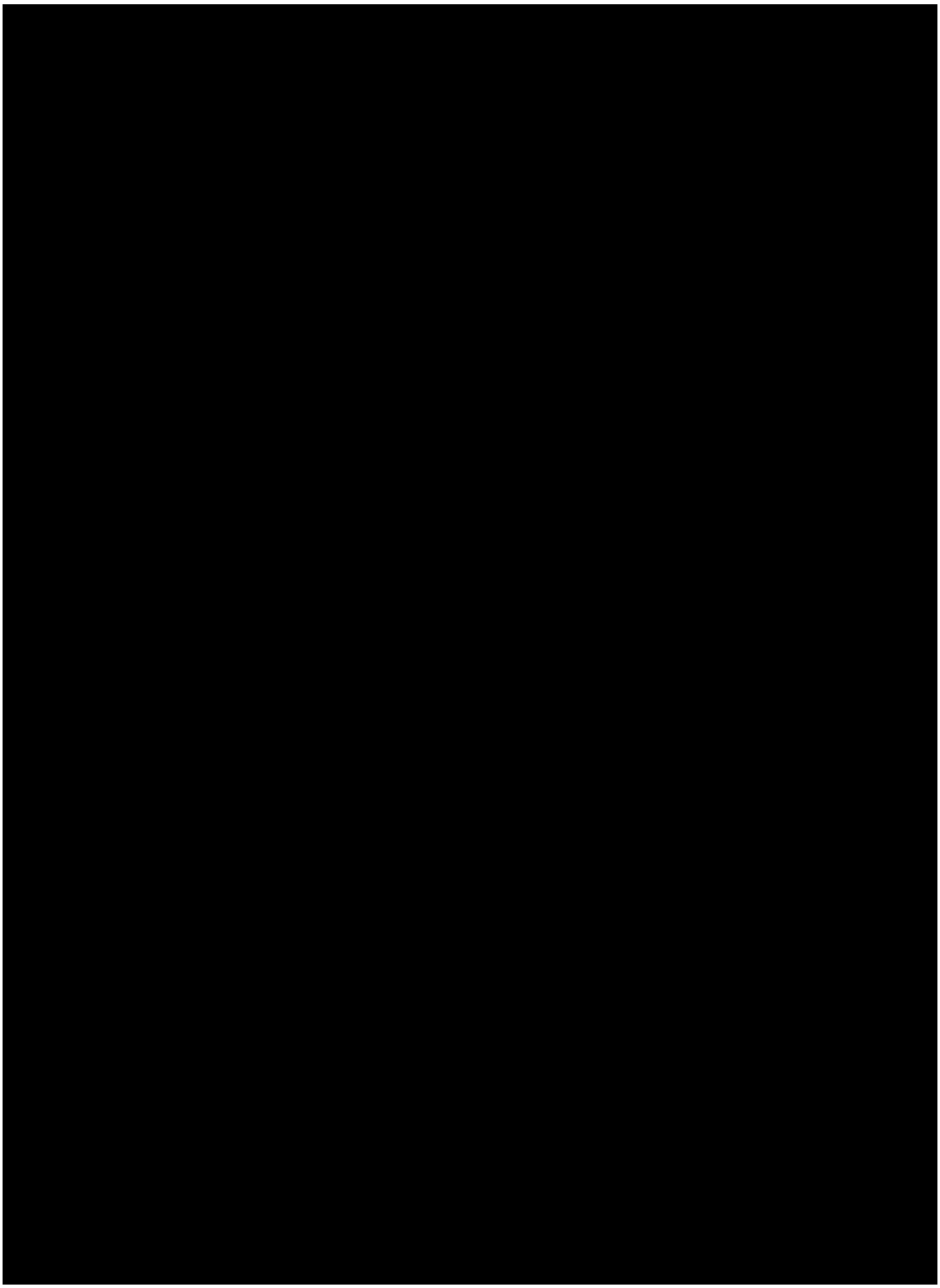
2.1.2.1 Grade 1



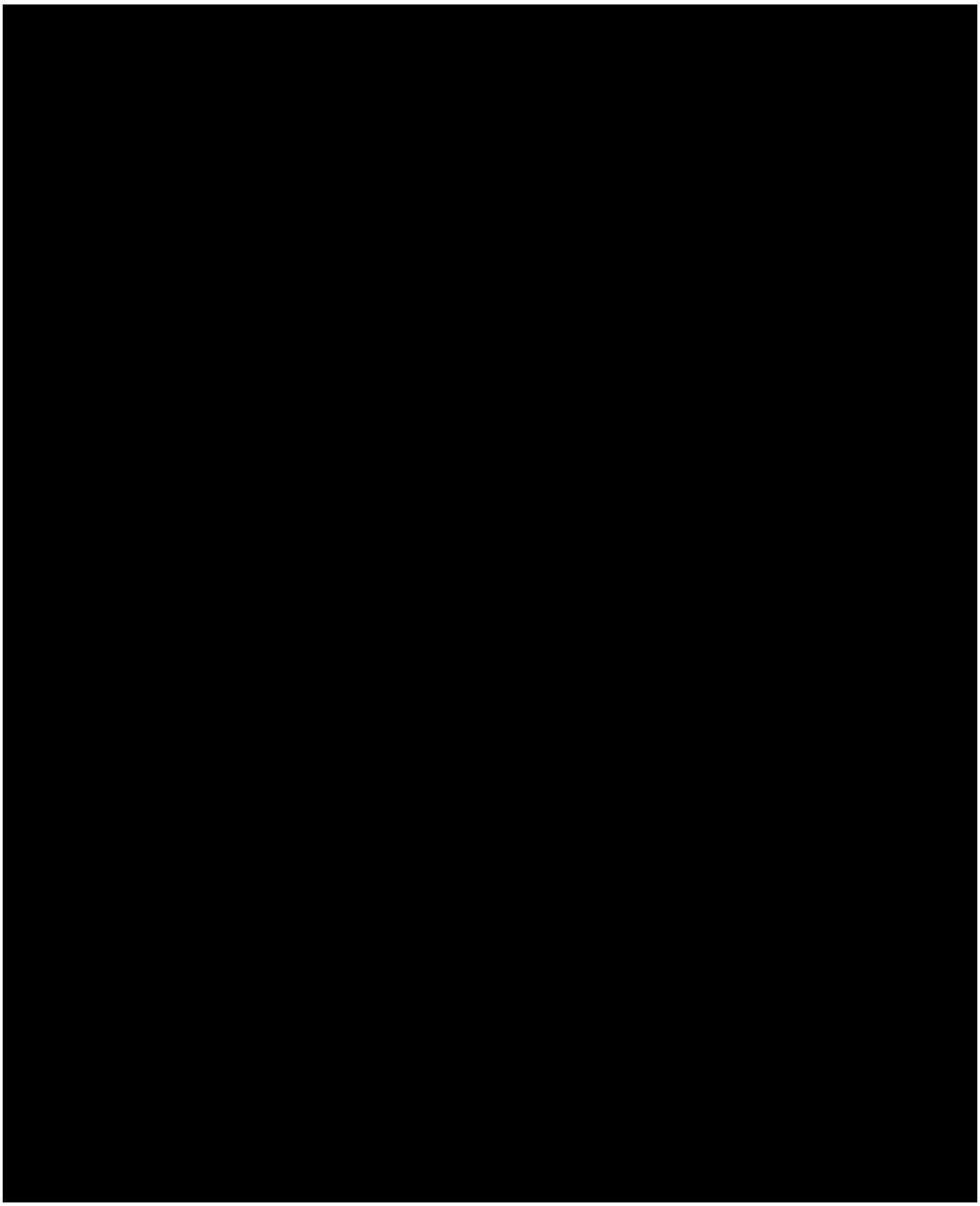


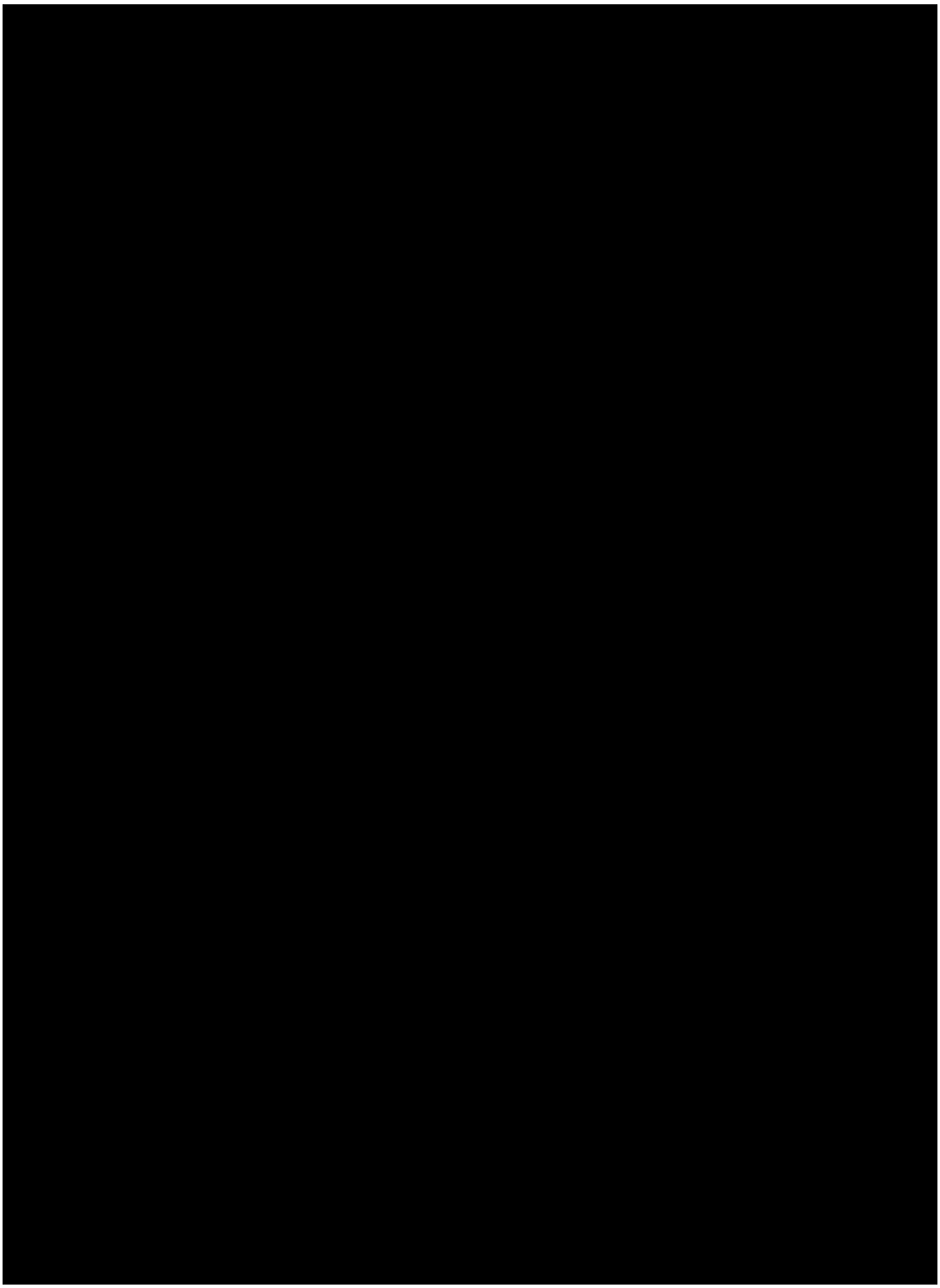
2.1.2.2 Grade 2



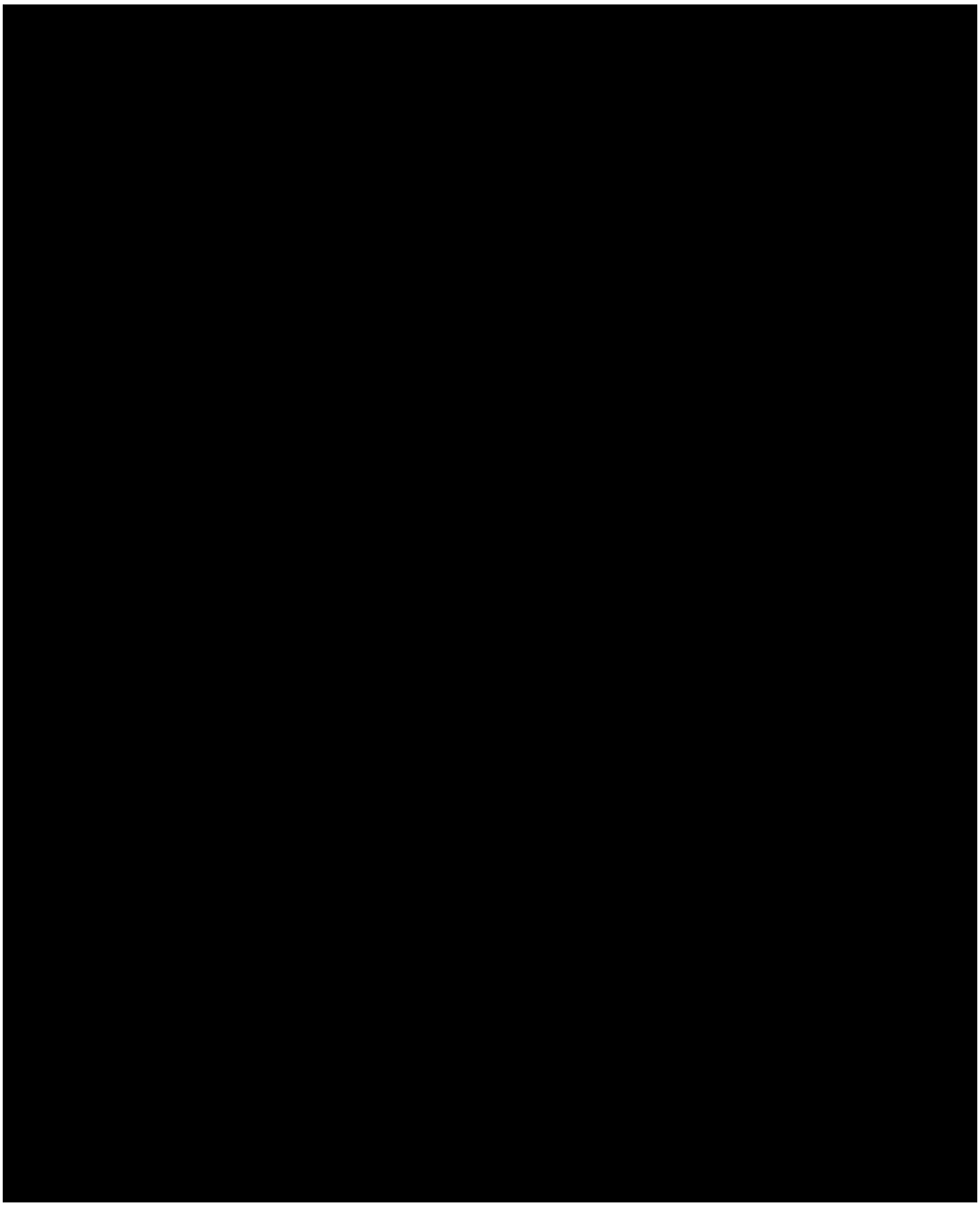


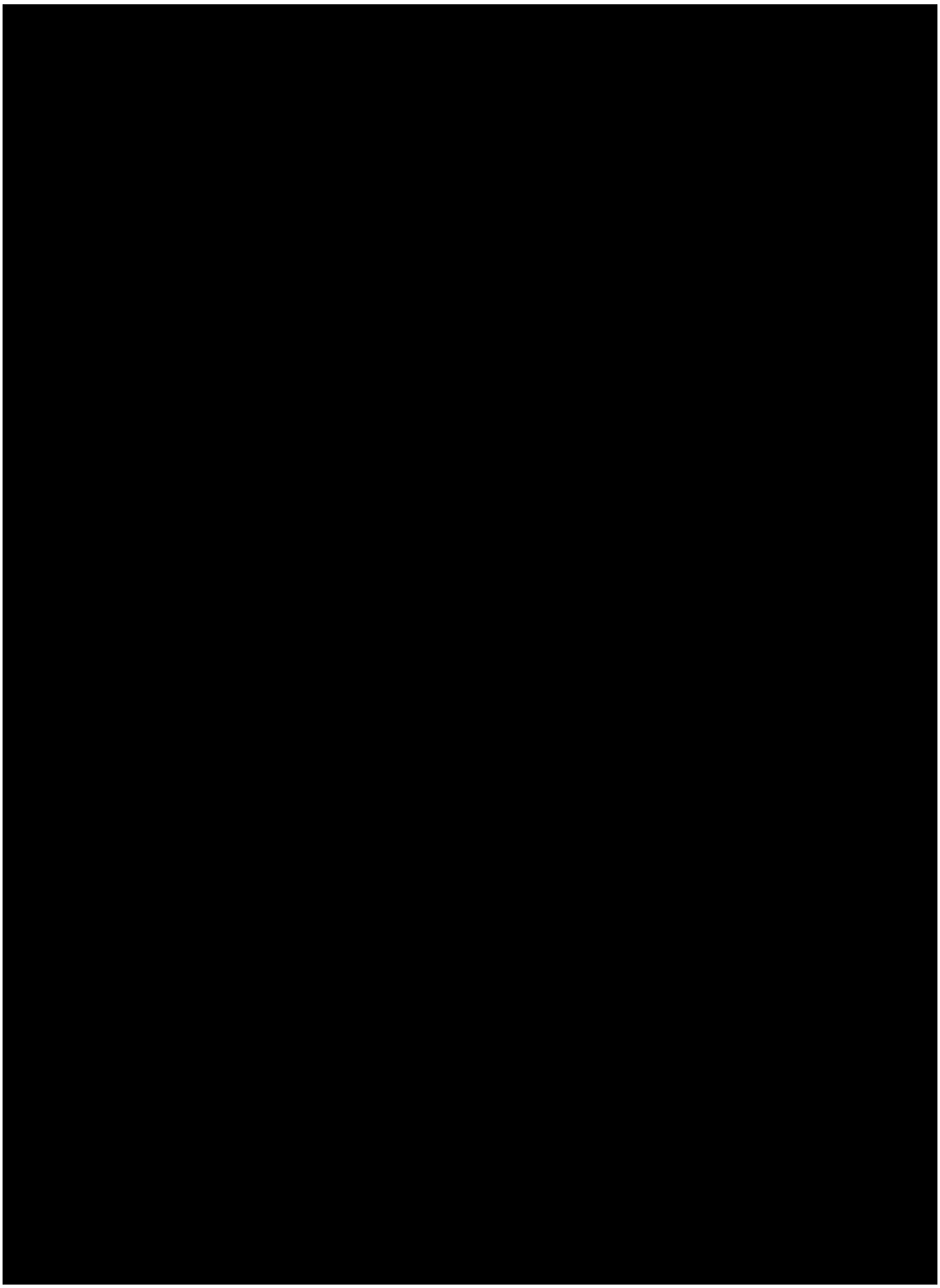
2.1.2.3 Grade 3



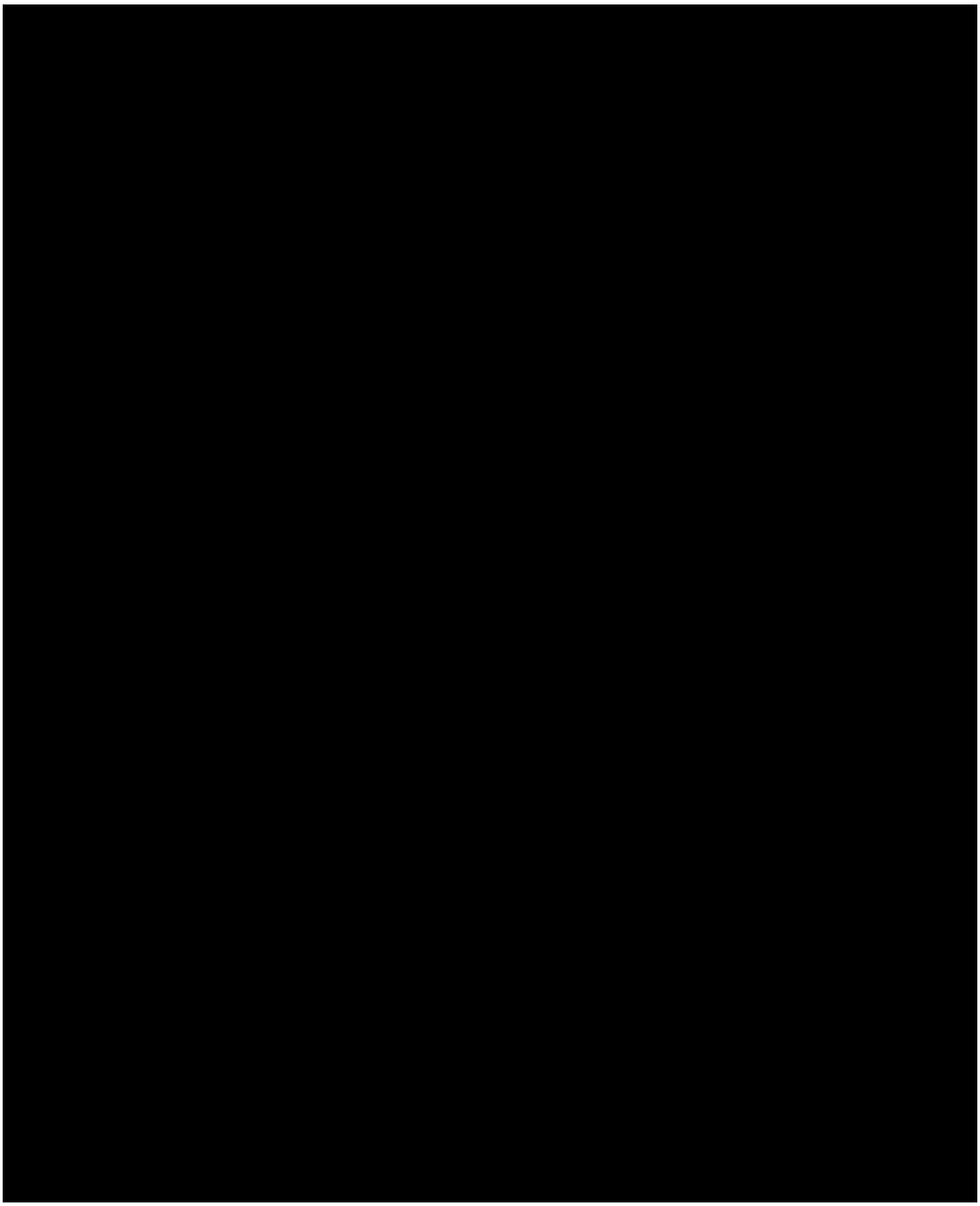


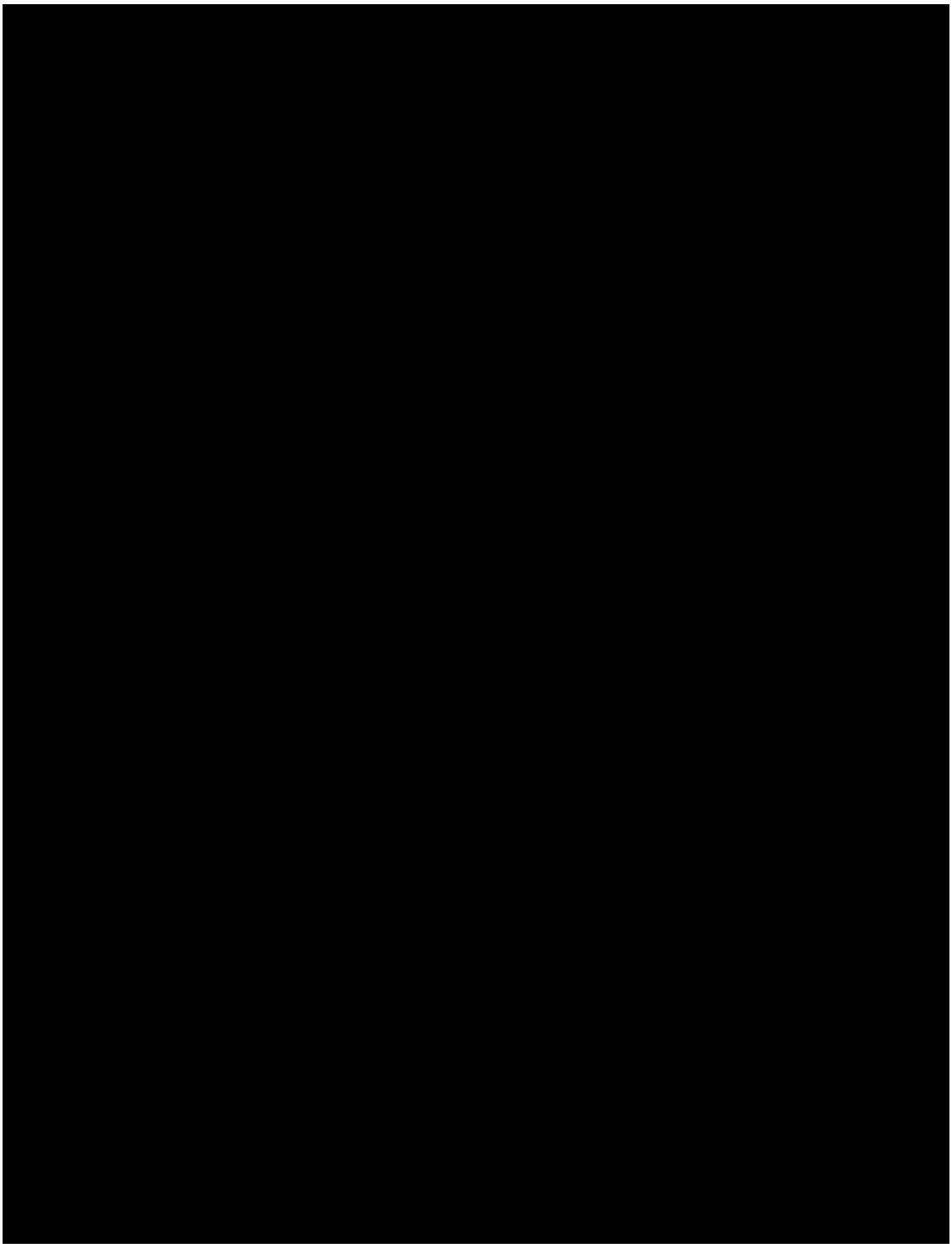
2.1.2.4 Grades 4–5



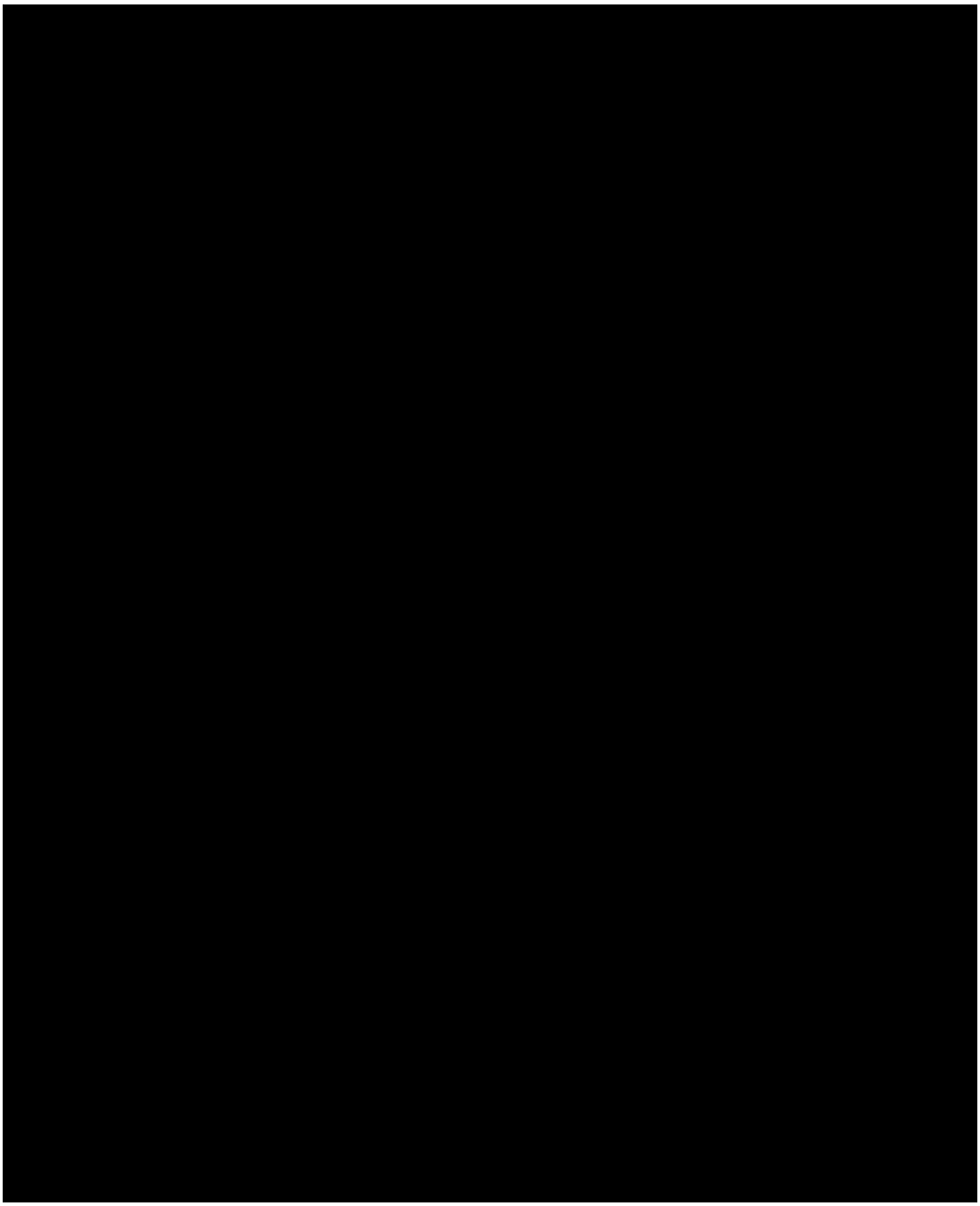


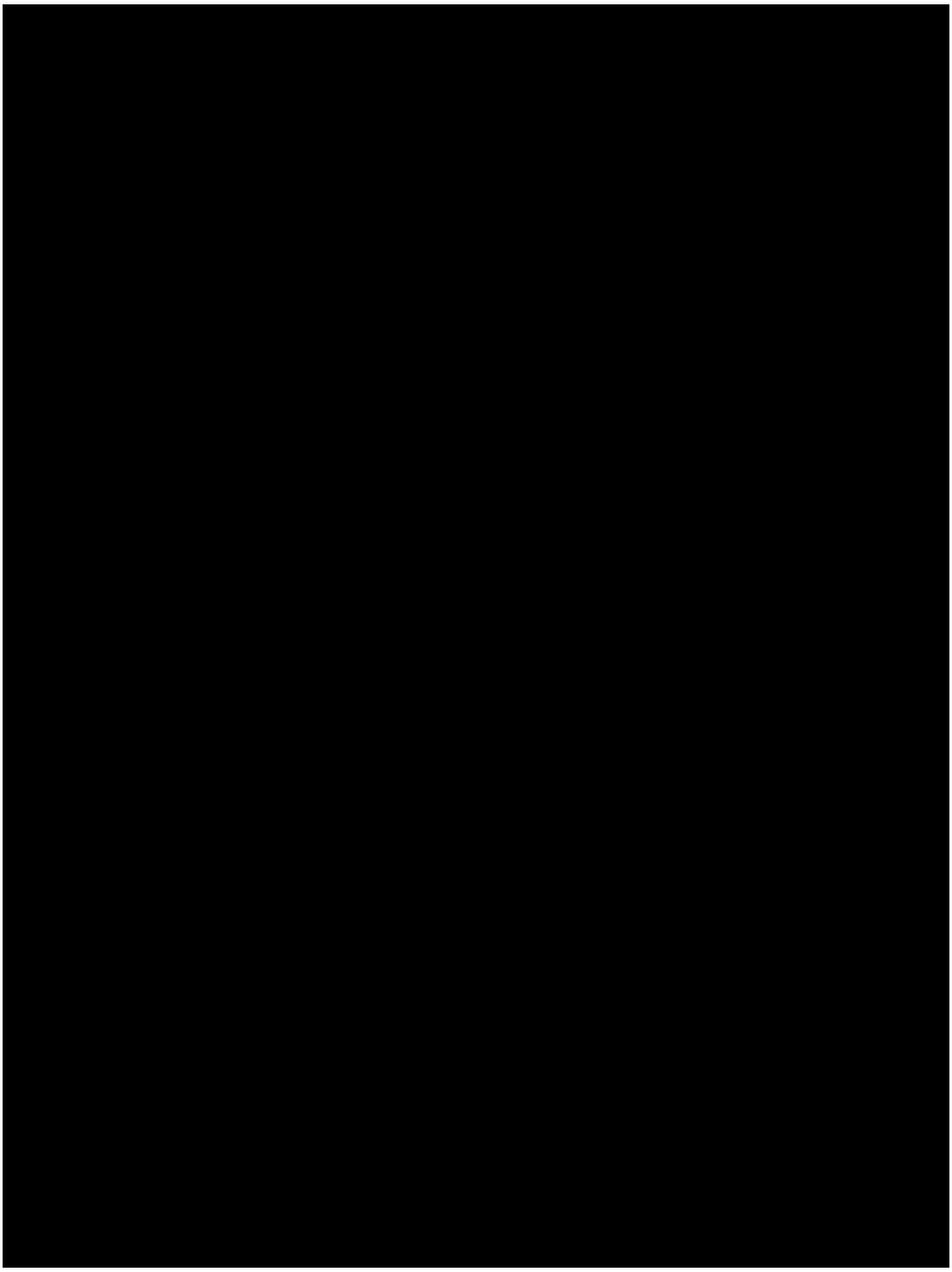
2.1.2.5 Grades 6–8





2.1.2.6 Grades 9–12





2.1.3 Writing

2.1.3.0 Kindergarten

2.1.3.1 Grade 1

2.1.3.2 Grade 2

2.1.3.3 Grade 3

2.1.3.4 Grades 4–5

2.1.3.5 Grades 6–8

2.1.3.6 Grades 9–12

2.1.4 Speaking

2.1.4.0 Kindergarten

2.1.4.1 Grade 1

2.1.4.2 Grade 2

2.1.4.3 Grade 3

2.1.4.4 Grades 4–5

2.1.4.5 Grades 6–8

2.1.4.6 Grades 9–12

2.2 DIF Analysis and Summary

Differential item functioning (DIF) analysis investigates whether factors extraneous to English language proficiency (i.e., the construct being measured on the test) may have influenced some students' performances on items. DIF attempts to find items that may be functioning differently for different groups based on criteria irrelevant to the construct that is purportedly being measured. We compare the performance of students on ACCESS for ELLs Paper items and tasks by dividing students into two different groupings: first, males versus females; second, students of Hispanic ethnic background versus students of all other backgrounds. We exclude students for whom gender or ethnicity[2] was unknown from both analyses. We used two commonly used procedures for detecting DIF: one for dichotomously scored items (Listening and Reading) and one for polytomously scored items (Writing and Speaking).

Dichotomous Items

We used the Mantel-Haenszel (M-H) chi-square statistic (Mantel & Haenszel, 1959) procedure for dichotomous items, originally proposed by the Educational Testing Service (ETS). 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. We assume 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). We use the M-H chi-square statistic 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.

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

We used the software program *EZDIF* (Waller, n.d.) to run the DIF analyses for all forms containing dichotomous items. For each test form, the greatest number of ability-level groupings is used; however, for many test forms, students scoring some of the lowest and highest raw scores need to be grouped together in order to have enough cases in each cell for the statistic to

[2] 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.

be appropriately calculated. (Note that this software program uses a two-step purification process; that is, items with C-level DIF in the first pass are removed from the matching variable in the second stage, and the DIF is then recalculated for the remaining items.)

Polytomous Items

For polytomous items (i.e., Writing and Speaking tasks), we take a similar approach. Our approach is based on the M-H chi-square statistic and the standardized mean difference following procedures that ETS developed (Allen, Carlson, & Zalanak, 1999; Zwick, Donoghue, & Grima, 1993). These DIF procedures for polytomous items were used to identify tasks that exhibit DIF. We used JMetrik (Meyer, 2018), an open source computer program for psychometric analysis, to conduct 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, we calculate the standardized mean difference between the performances of the two comparison groups. 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 that ETS proposed (Allen et al., 1999) is also rescaled to the standardized P-DIF* metric (Meyer, 2018).

Following guidance that ETS proposed for the National Assessment of Educational Progress (Allen et al., 1999), we classify ACCESS for ELLs Writing and Speaking tasks 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 ethnicities. The top part of the table summarizes the number of items that 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.

Paper ACCESS is administered as two rotating static forms. Bias and sensitivity panels reviewed these items prior to any field testing, as described in Section 2.2.1. We conducted DIF analysis prior to the final selection of the two static forms. For any items or tasks that showed C-level (or CC-level) DIF, an additional DIF review panel was convened to re-examine the item for bias concerns.

Panel members were drawn from CAL staff members who have expertise in instruction and/or professional development for English learner students. The panel included a mix of women and men and included CAL staff who have a language other than English as a first language, with attention paid to ensuring representation of individuals from Spanish-language backgrounds and non–Spanish-language backgrounds. The facilitator asked the panel to discuss the item and come to consensus on whether the item demonstrates bias against a particular group and is appropriate to place on the operational test. The facilitator does not disclose to the panel which subgroup the DIF analysis indicates is favored by the item.

Two items showed C-level DIF. The first item, on the Grade 3 and Grades 4–5 Listening Tier A test, exhibited C-level DIF favoring non-Hispanic students. The panel concluded that this item did not show bias and is appropriate for operational testing. The second item, on the Grades 9–12 Listening Tier A test, showed C-level DIF favoring Hispanic students. The panel concluded that this item showed bias against Hispanic students.

2.2.1 Listening

2.2.1.0 Kindergarten

2.2.1.1 Grade 1

2.2.1.2 Grade 2

2.2.1.3 Grade 3

2.2.1.4 Grades 4–5

2.2.1.5 Grades 6–8

2.2.1.6 Grades 9–12

[redacted]

2.2.2 Reading

2.2.2.0 Kindergarten

[redacted]

2.2.2.1 Grade 1

2.2.2.2 Grade 2

2.2.2.3 Grade 3

2.2.2.4 Grades 4–5

2.2.2.5 Grades 6–8

2.2.2.6 Grades 9–12

2.2.3 Writing

2.2.3.0 Kindergarten

Table 2.2.3.0

DIF Analysis and Summary: Writ K S501 Paper

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

2.2.3.1 Grade 1

Table 2.2.3.1.1

DIF Analysis and Summary: Writ 1 A S501 Paper

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

Table 2.2.3.1.2

DIF Analysis and Summary: Writ 1 B/C S501 Paper

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 Grade 2

Table 2.2.3.2.1

DIF Analysis and Summary: Writ 2 A S501 Paper

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

Note: the test form is shared between 2A and 3A.

Table 2.2.3.2.2

DIF Analysis and Summary: Writ 2 B/C S501 Paper

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

Note: the test form is shared between 2BC and 3BC.

2.2.3.3 Grade 3

Table 2.2.3.3.1

DIF Analysis and Summary: Writ 3 A S501 Paper

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

Note: the test form is shared between 2A and 3A.

Table 2.2.3.3.2

DIF Analysis and Summary: Writ 3 B/C S501 Paper

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

Note: the test form is shared between 2BC and 3BC.

2.2.3.4 Grades 4–5

Table 2.2.3.4.1

DIF Analysis and Summary: Writ 4-5 A S501 Paper

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 2.2.3.4.2

DIF Analysis and Summary: Writ 4-5 B/C S501 Paper

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 6–8

Table 2.2.3.5.1

DIF Analysis and Summary: Writ 6-8 A S501 Paper

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 2.2.3.5.2

DIF Analysis and Summary: Writ 6-8 B/C S501 Paper

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

2.2.3.6 Grades 9–12

Table 2.2.3.6.1

DIF Analysis and Summary: Writ 9-12 A S501 Paper

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 2.2.3.6.2

DIF Analysis and Summary: Writ 9-12 B/C S501 Paper

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

2.2.4 Speaking

2.2.4.0 Kindergarten

Table 2.2.4.0

DIF Analysis and Summary: Spek K S501 Paper

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
AA	6	4	4	6
BB	0	0	0	0
CC	0	0	0	0

2.2.4.1 Grade 1

Table 2.2.4.1.1

DIF Analysis and Summary: Spek 1 A S501 Paper

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 2.2.4.1.2

DIF Analysis and Summary: Spek 1 B/C S501 Paper

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

2.2.4.2 Grade 2

Table 2.2.4.2.1

DIF Analysis and Summary: Spek 2 A S501 Paper

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

Note: the test form is shared between 2A and 3A.

Table 2.2.4.2.2

DIF Analysis and Summary: Spek 2 B/C S501 Paper

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
CC	0	0	0	0

Note: the test form is shared between 2BC and 3BC.

2.2.4.3 Grade 3

Table 2.2.4.3.1

DIF Analysis and Summary: Spek 3 A S501 Paper

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

Note: the test form is shared between 2A and 3A.

Table 2.2.4.3.2

DIF Analysis and Summary: Spek 3 B/C S501 Paper

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
CC	0	0	0	0

Note: the test form is shared between 2BC and 3BC.

2.2.4.4 Grades 4–5

Table 2.2.4.4.1

DIF Analysis and Summary: Spek 4-5 A S501 Paper

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

Table 2.2.4.4.2

DIF Analysis and Summary: Spek 4-5 B/C S501 Paper

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

2.2.4.5 Grades 6–8

Table 2.2.4.5.1

DIF Analysis and Summary: Spek 6-8 A S501 Paper

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

Table 2.2.4.5.2

DIF Analysis and Summary: Spek 6-8 B/C S501 Paper

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

2.2.4.6 Grades 9–12

Table 2.2.4.6.1

DIF Analysis and Summary: Spek 9-12 A S501 Paper

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

Table 2.2.4.6.2

DIF Analysis and Summary: Spek 9-12 B/C S501 Paper

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

2.3 Raw Score Distribution

Figures and tables in this section provide detail on the distribution of raw scores. 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

Test design and student population impact the distribution of raw scores. In general, raw score distributions tend to be smoothly distributed with a single peak; however, there are a number of exceptions. Understanding these distributions supports the understanding of other statistical properties of the test forms.

In the domain of Writing, in Tier B/C, the three tasks are weighted once, twice, and three times, respectively. The impact of this weighting is that the raw scores are not smoothly distributed.

In the domain of Speaking, on Tier A forms, three of the six tasks are scored on a restricted portion of the rubric (with possible raw scores of 0 to 2). Most students score all six of these points; however, less proficient students may score only one or two points consistently on the remaining tasks. On Tier B/C, students are automatically awarded these six points (as it is assumed they would have the ability to achieve the maximum possible points on the easiest tasks). These aspects of the test design impact raw score distribution.

The Kindergarten test design includes skipping and stopping rules intended to reduce testing time for young children; these rules also have an impact on the distribution of raw scores, leading to less smooth distributions.

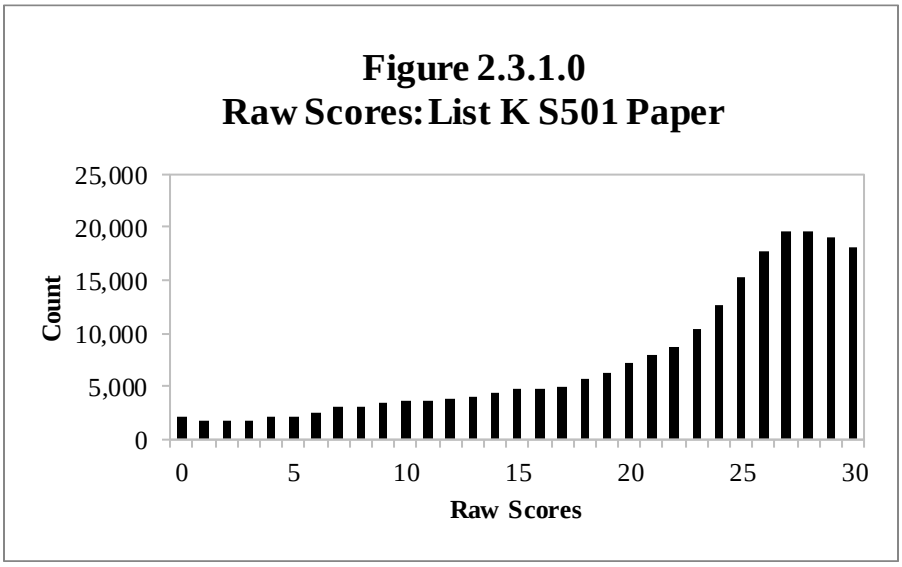
2.3.1 Listening

2.3.1.0 Kindergarten

Table 2.3.1.0

Raw Score Descriptive Statistics: List K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	226,001	0	30	21.57	7.68
Total	226,001	0	30	21.57	7.68



2.3.1.1 Grade 1

Table 2.3.1.1.1

Raw Score Descriptive Statistics: List 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	16,725	1	18	13.87	3.17
Total	16,725	1	18	13.87	3.17

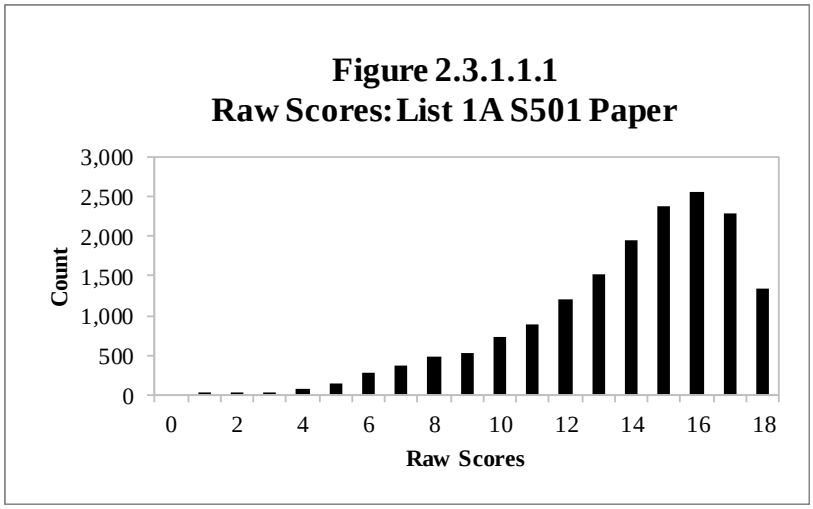
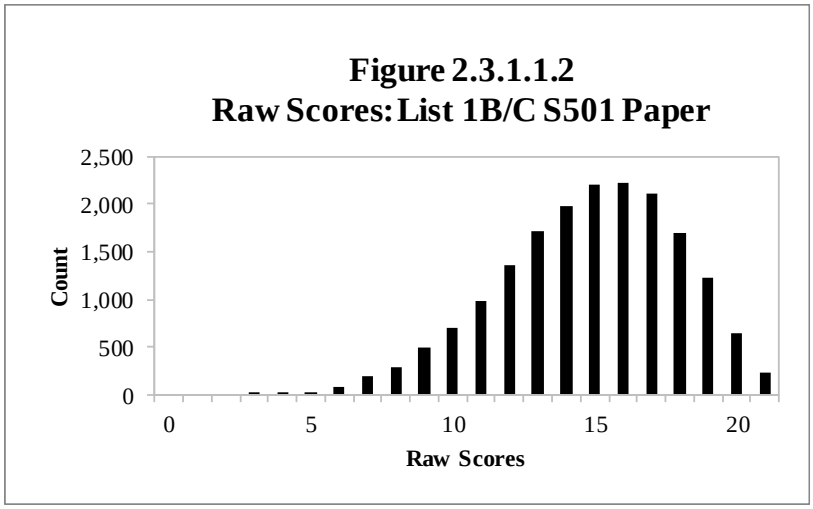


Table 2.3.1.1.2

Raw Score Descriptive Statistics: List 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	18,162	3	21	14.80	3.14
Total	18,162	3	21	14.80	3.14



2.3.1.2 Grade 2

Table 2.3.1.2.1

Raw Score Descriptive Statistics: List 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	8,224	2	18	13.82	3.54
Total	8,224	2	18	13.82	3.54

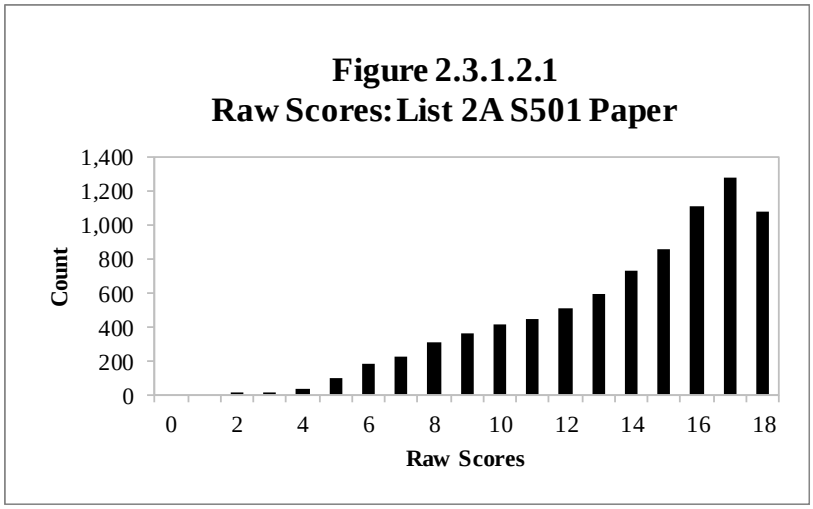
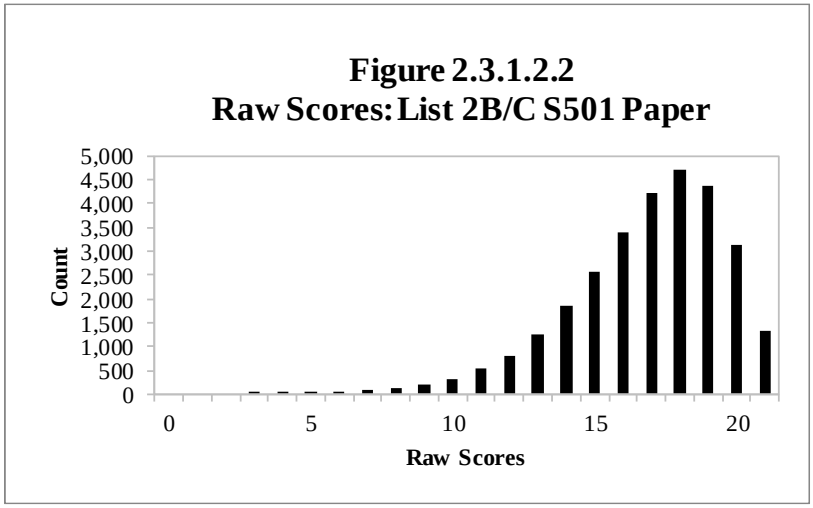


Table 2.3.1.2.2

Raw Score Descriptive Statistics: List 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	28,860	3	21	16.85	2.71
Total	28,860	3	21	16.85	2.71



2.3.1.3 Grade 3

Table 2.3.1.3.1

Raw Score Descriptive Statistics: List 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	6,823	0	18	10.57	3.60
Total	6,823	0	18	10.57	3.60

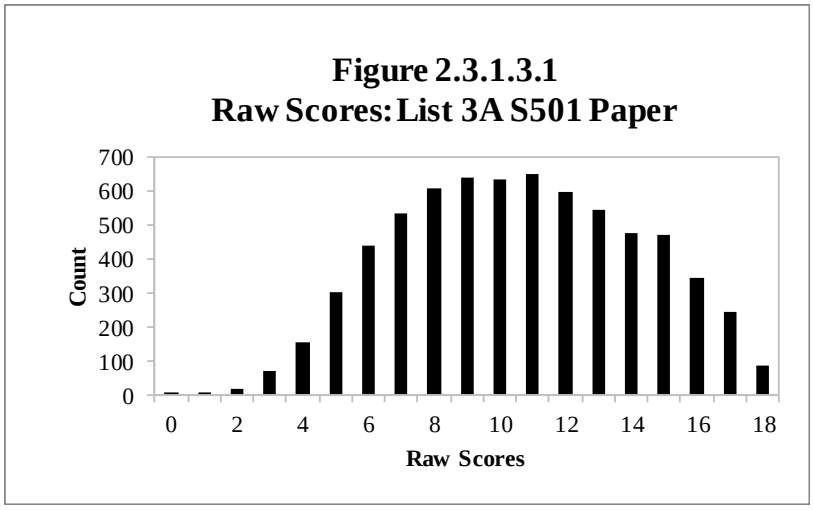
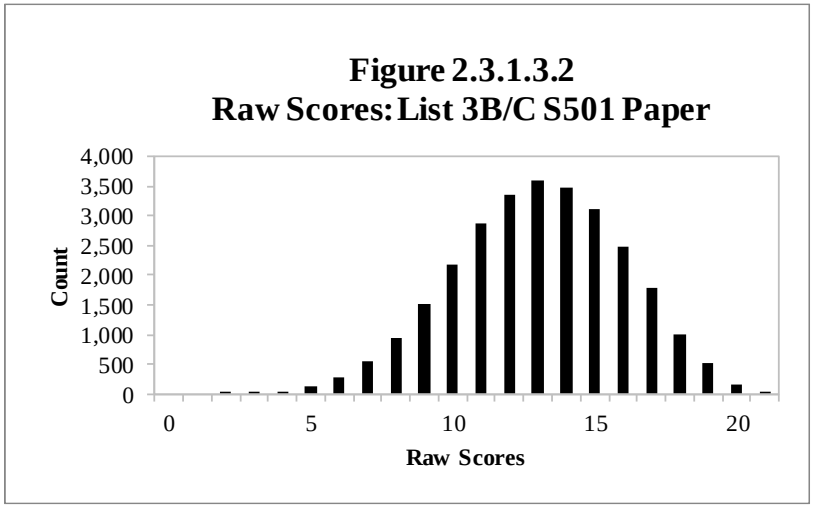


Table 2.3.1.3.2

Raw Score Descriptive Statistics: List 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	28,011	2	21	13.03	2.98
Total	28,011	2	21	13.03	2.98



2.3.1.4 Grades 4–5

Table 2.3.1.4.1

Raw Score Descriptive Statistics: List 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,112	0	18	10.83	3.68
5	4,660	1	18	11.36	3.69
Total	9,772	0	18	11.08	3.70

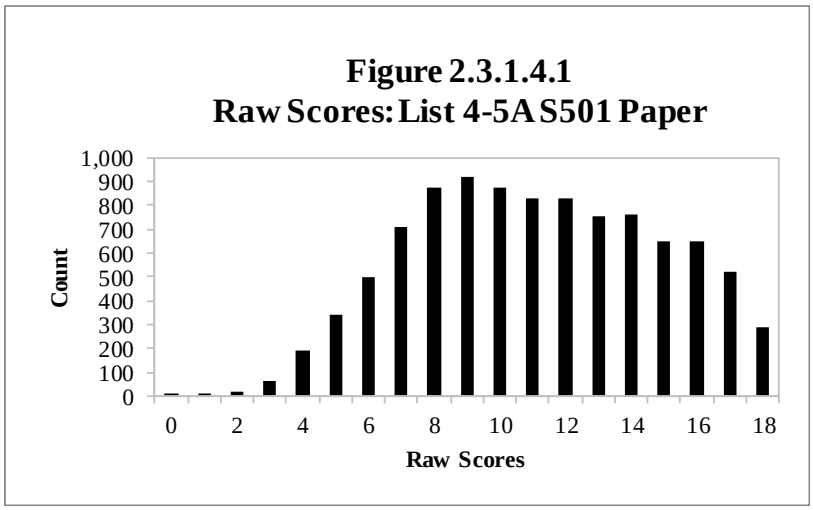
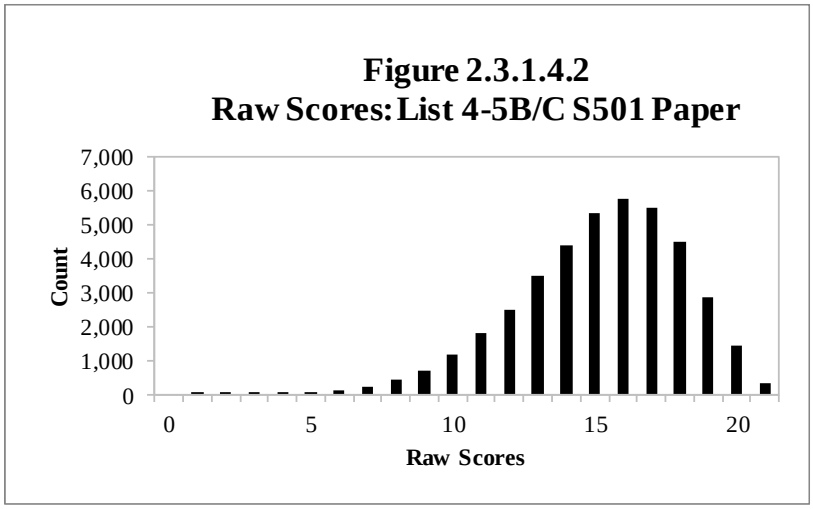


Table 2.3.1.4.2

Raw Score Descriptive Statistics: List 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	21,578	1	21	14.78	2.89
5	18,981	2	21	15.68	2.81
Total	40,559	1	21	15.20	2.89



2.3.1.5 Grades 6–8

Table 2.3.1.5.1

Raw Score Descriptive Statistics: List 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,267	1	18	10.29	3.50
7	4,104	0	18	10.27	3.49
8	4,133	1	18	10.41	3.41
Total	12,504	0	18	10.32	3.46

Figure 2.3.1.5.1
Raw Scores: List 6-8A S501 Paper

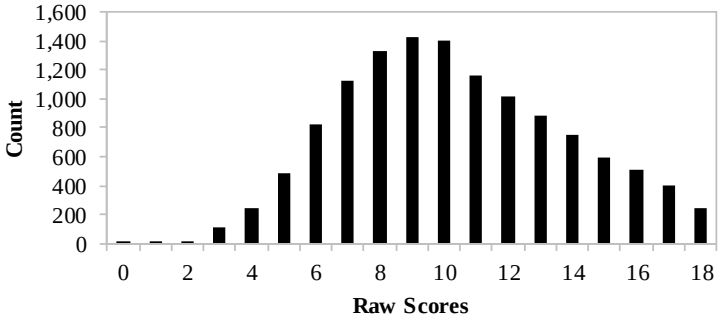
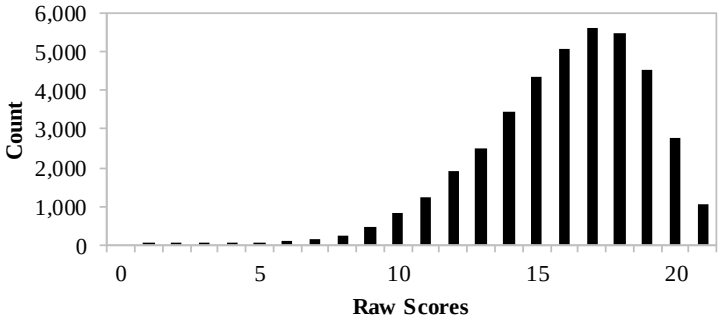


Table 2.3.1.5.2

Raw Score Descriptive Statistics: List 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,183	1	21	15.45	2.88
7	12,750	3	21	16.13	2.92
8	11,846	3	21	16.63	2.84
Total	39,779	1	21	16.02	2.92

Figure 2.3.1.5.2
Raw Scores: List 6-8B/C S501 Paper



2.3.1.6 Grades 9–12

Table 2.3.1.6.1

Raw Score Descriptive Statistics: List 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	4,830	1	18	10.41	2.94
10	3,519	1	18	10.83	2.95
11	2,673	3	18	11.24	2.86
12	1,468	1	18	11.38	2.78
Total	12,490	1	18	10.82	2.93

Figure 2.3.1.6.1
Raw Scores: List 9-12AS501 Paper

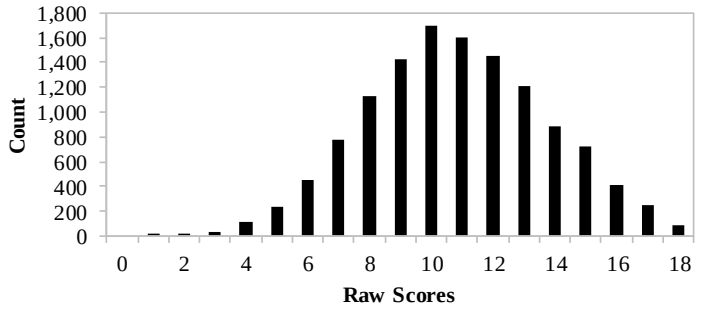
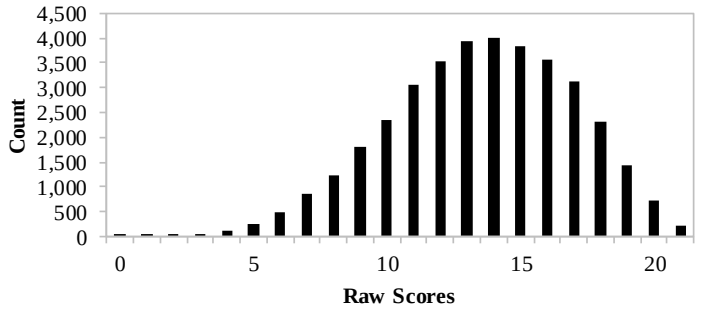


Table 2.3.1.6.2

Raw Score Descriptive Statistics: List 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	10,743	2	21	13.48	3.29
10	10,322	2	21	13.57	3.42
11	9,304	0	21	13.75	3.48
12	6,514	2	21	13.27	3.51
Total	36,883	0	21	13.53	3.42

Figure 2.3.1.6.2
Raw Scores: List 9-12B/C S501 Paper



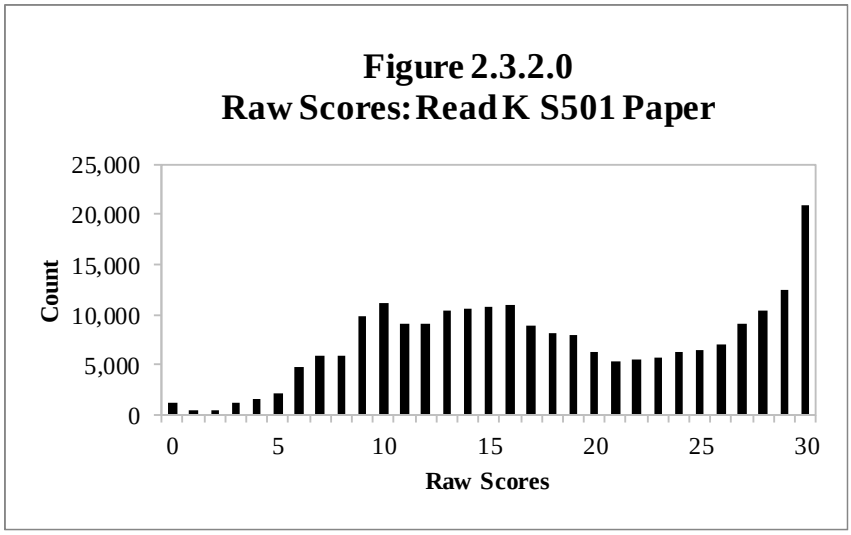
2.3.2 Reading

2.3.2.0 Kindergarten

Table 2.3.2.0

Raw Score Descriptive Statistics: Read K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,994	0	30	18.26	7.88
Total	225,994	0	30	18.26	7.88



2.3.2.1 Grade 1

Table 2.3.2.1.1

Raw Score Descriptive Statistics: Read 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	15,779	1	24	12.40	4.63
Total	15,779	1	24	12.40	4.63

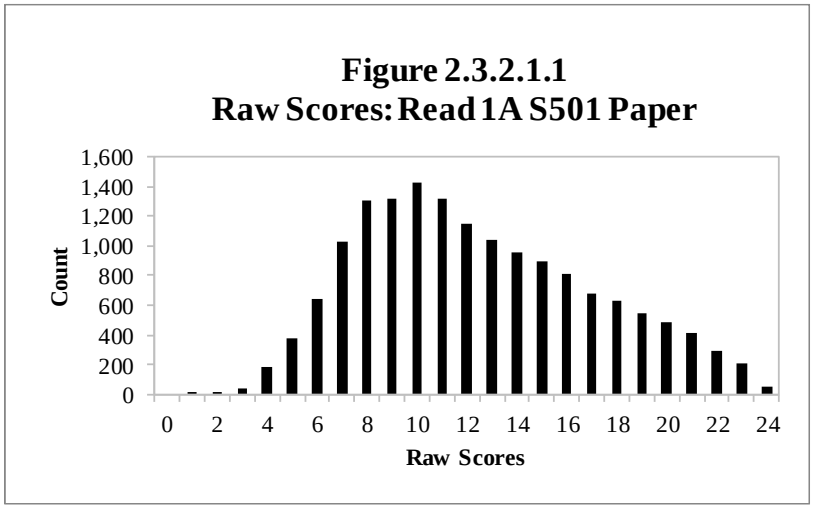
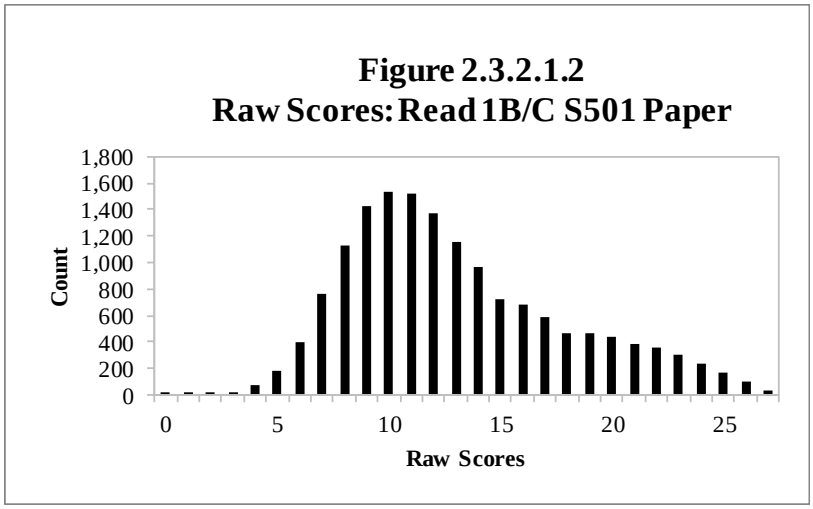


Table 2.3.2.1.2

Raw Score Descriptive Statistics: Read 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	15,444	0	27	13.04	4.87
Total	15,444	0	27	13.04	4.87



2.3.2.2 Grade 2

Table 2.3.2.2.1

Raw Score Descriptive Statistics: Read 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	7,826	1	24	14.01	5.15
Total	7,826	1	24	14.01	5.15

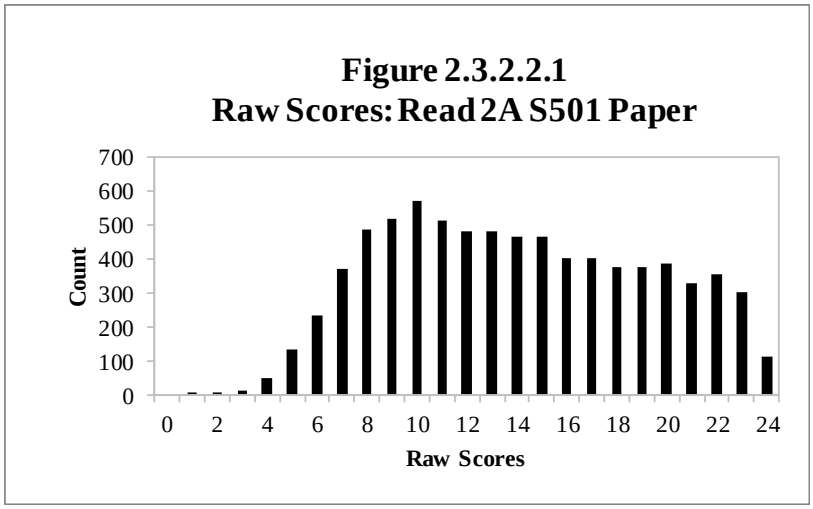
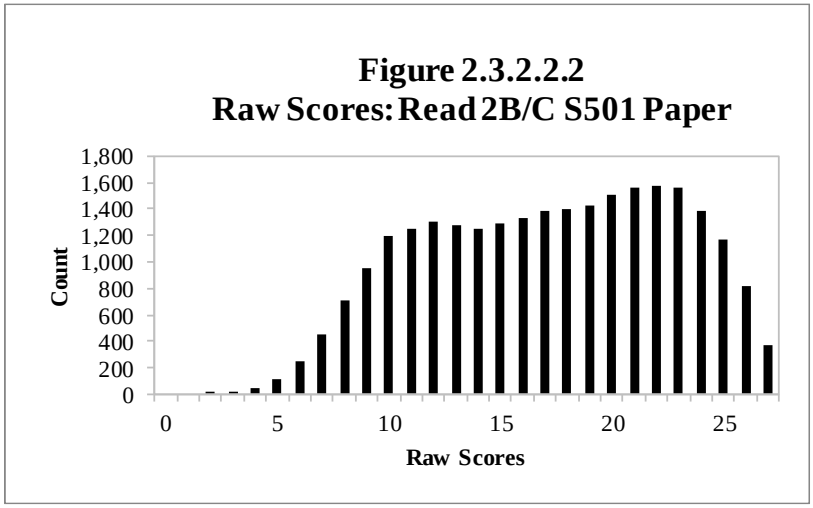


Table 2.3.2.2.2

Raw Score Descriptive Statistics: Read 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	25,588	2	27	17.16	5.52
Total	25,588	2	27	17.16	5.52



2.3.2.3 Grade 3

Table 2.3.2.3.1

Raw Score Descriptive Statistics: Read 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	6,393	2	24	11.37	4.77
Total	6,393	2	24	11.37	4.77

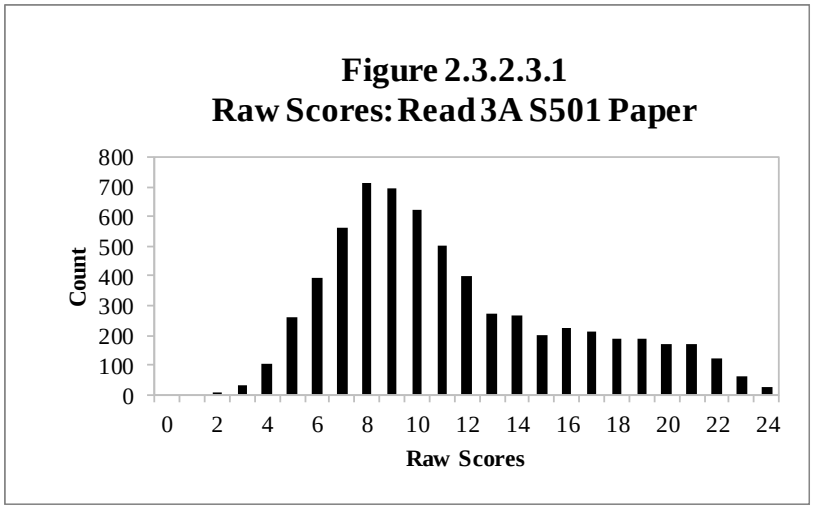
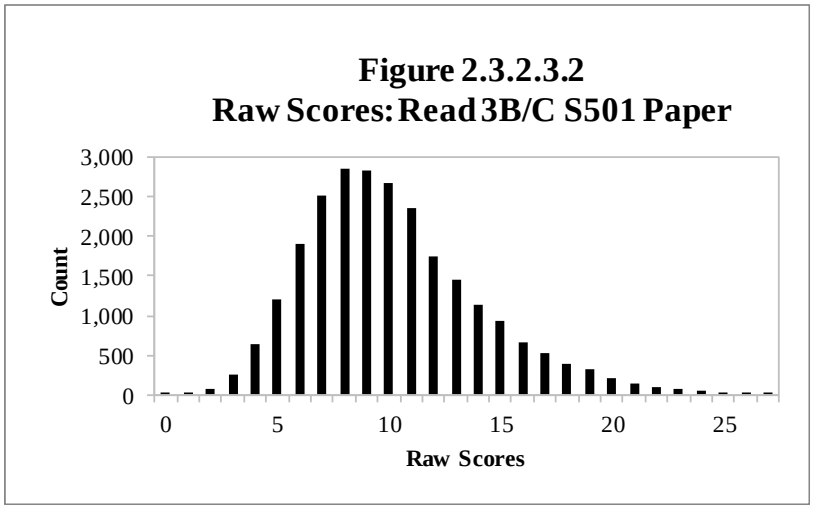


Table 2.3.2.3.2

Raw Score Descriptive Statistics: Read 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	25,102	0	27	10.16	3.93
Total	25,102	0	27	10.16	3.93



2.3.2.4 Grades 4–5

Table 2.3.2.4.1

Raw Score Descriptive Statistics: Read 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	4,887	2	24	12.61	5.21
5	4,501	1	24	13.61	5.36
Total	9,388	1	24	13.09	5.31

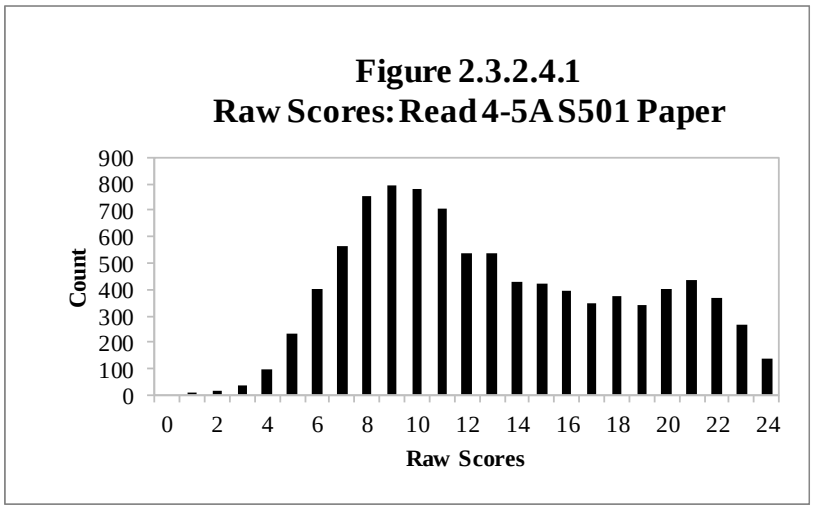
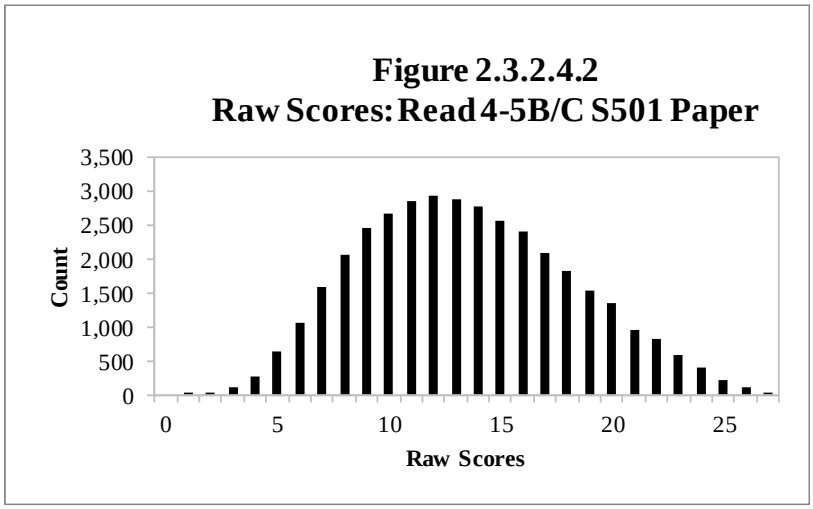


Table 2.3.2.4.2

Raw Score Descriptive Statistics: Read 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	19,576	1	27	12.62	4.51
5	17,650	1	27	14.46	4.81
Total	37,226	1	27	13.49	4.75



2.3.2.5 Grades 6–8

Table 2.3.2.5.1

Raw Score Descriptive Statistics: Read 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,082	2	24	11.71	4.51
7	3,982	1	24	12.19	4.65
8	4,045	2	24	12.92	4.59
Total	12,109	1	24	12.27	4.61

Figure 2.3.2.5.1
Raw Scores: Read 6-8A S501 Paper

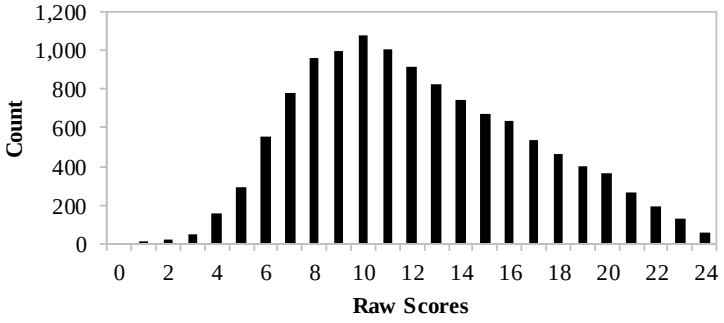
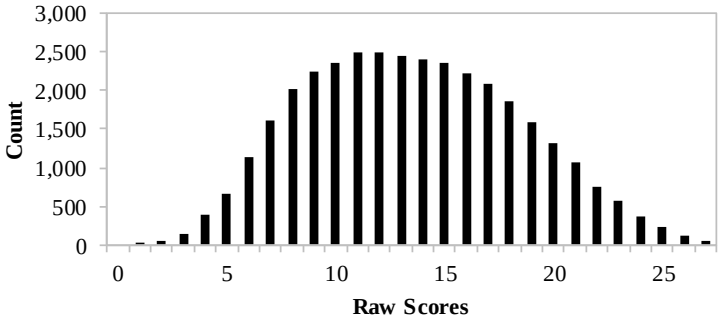


Table 2.3.2.5.2

Raw Score Descriptive Statistics: Read 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	13,262	1	27	12.31	4.44
7	11,254	1	27	13.57	4.89
8	10,478	1	27	15.02	5.10
Total	34,994	1	27	13.52	4.92

Figure 2.3.2.5.2
Raw Scores: Read 6-8B/C S501 Paper



2.3.2.6 Grades 9–12

Table 2.3.2.6.1

Raw Score Descriptive Statistics: Read 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	4,784	1	24	12.86	4.49
10	3,499	2	24	13.87	4.71
11	2,681	3	24	14.84	4.64
12	1,458	3	24	15.15	4.44
Total	12,422	1	24	13.84	4.66

Figure 2.3.2.6.1
Raw Scores: Read 9-12A S501 Paper

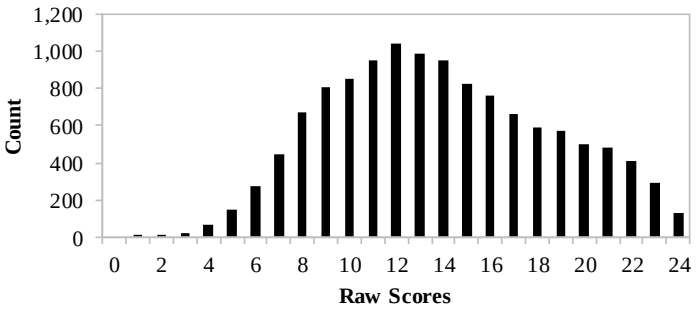
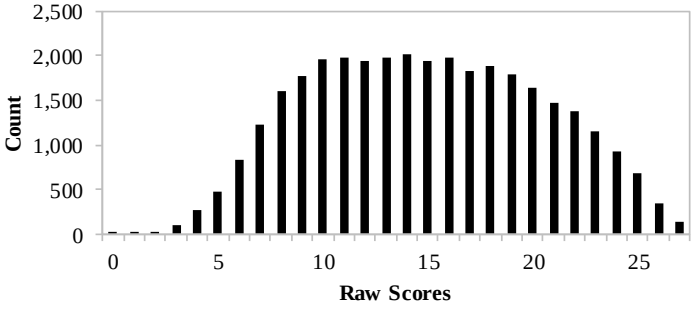


Table 2.3.2.6.2

Raw Score Descriptive Statistics: Read 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	9,545	0	27	14.30	5.28
10	9,282	1	27	14.91	5.38
11	8,515	1	27	15.60	5.55
12	6,058	0	27	14.84	5.58
Total	33,400	0	27	14.90	5.46

Figure 2.3.2.6.2
Raw Scores: Read 9-12B/C S501 Paper



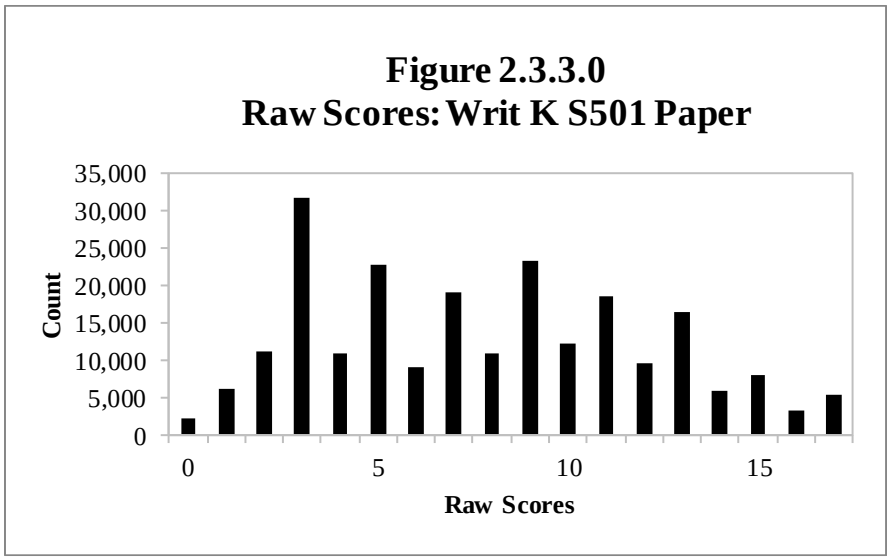
2.3.3 Writing

2.3.3.0 Kindergarten

Table 2.3.3.0

Raw Score Descriptive Statistics: Writ K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,987	0	17	7.79	4.25
Total	225,987	0	17	7.79	4.25



2.3.3.1 Grade 1

Table 2.3.3.1.1

Raw Score Descriptive Statistics: Writ 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	19,965	0	25	11.88	5.87
Total	19,965	0	25	11.88	5.87

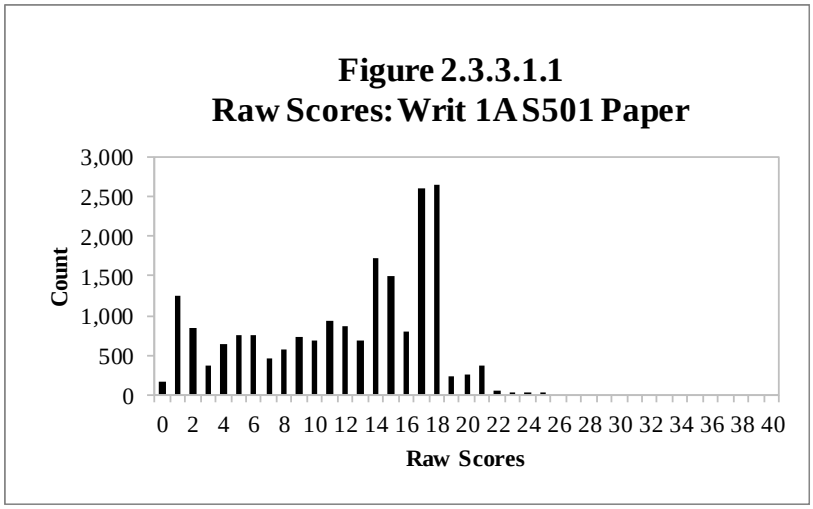
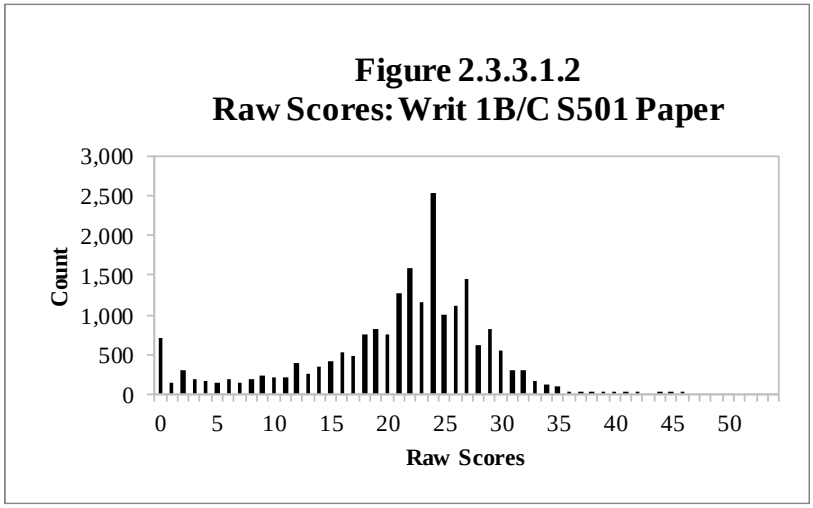


Table 2.3.3.1.2

Raw Score Descriptive Statistics: Writ 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	20,779	0	46	20.64	8.01
Total	20,779	0	46	20.64	8.01



2.3.3.2 Grade 2

Table 2.3.3.2.1

Raw Score Descriptive Statistics: Writ 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	9,124	0	18	7.30	4.76
Total	9,124	0	18	7.30	4.76

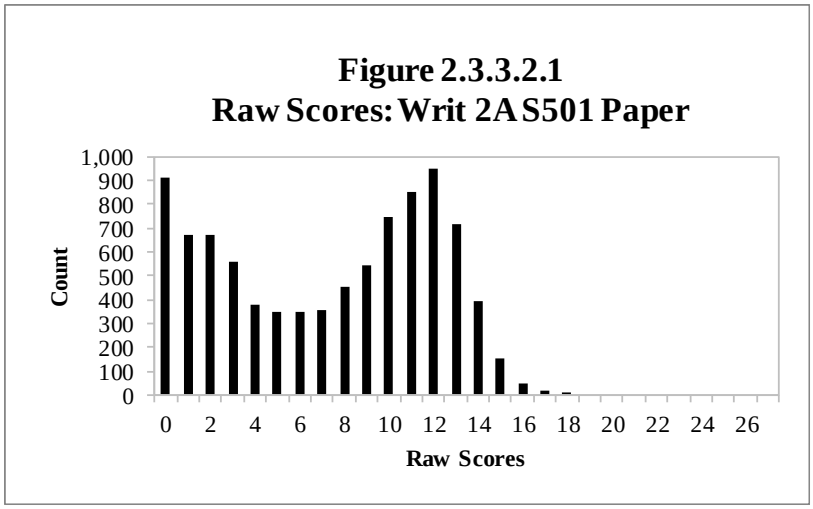
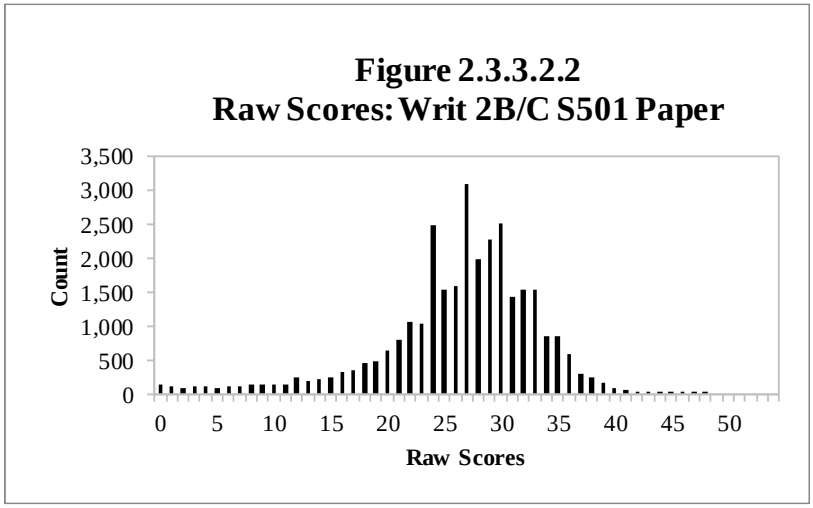


Table 2.3.3.2.2

Raw Score Descriptive Statistics: Writ 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	30,694	0	48	26.25	7.05
Total	30,694	0	48	26.25	7.05



2.3.3.3 Grade 3

Table 2.3.3.3.1

Raw Score Descriptive Statistics: Writ 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	7,464	0	18	8.35	4.64
Total	7,464	0	18	8.35	4.64

Figure 2.3.3.3.1
Raw Scores: Writ 3A S501 Paper

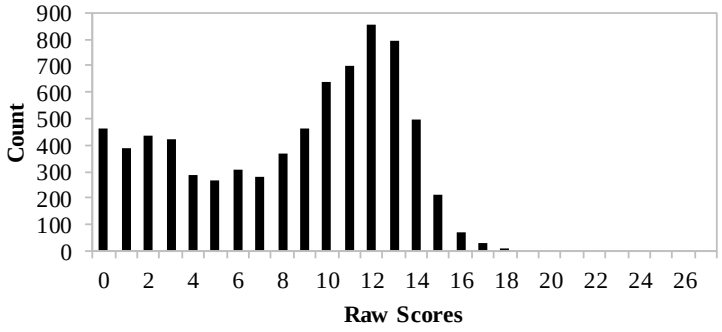
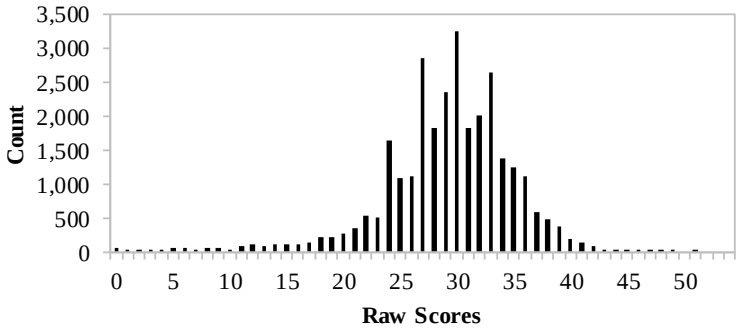


Table 2.3.3.3.2

Raw Score Descriptive Statistics: Writ 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,669	0	51	29.04	6.10
Total	29,669	0	51	29.04	6.10

Figure 2.3.3.3.2
Raw Scores: Writ 3B/C S501 Paper



2.3.3.4 Grades 4–5

Table 2.3.3.4.1

Raw Score Descriptive Statistics: Writ 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,498	0	20	7.62	4.29
5	4,980	0	18	8.27	4.20
Total	10,478	0	20	7.93	4.26

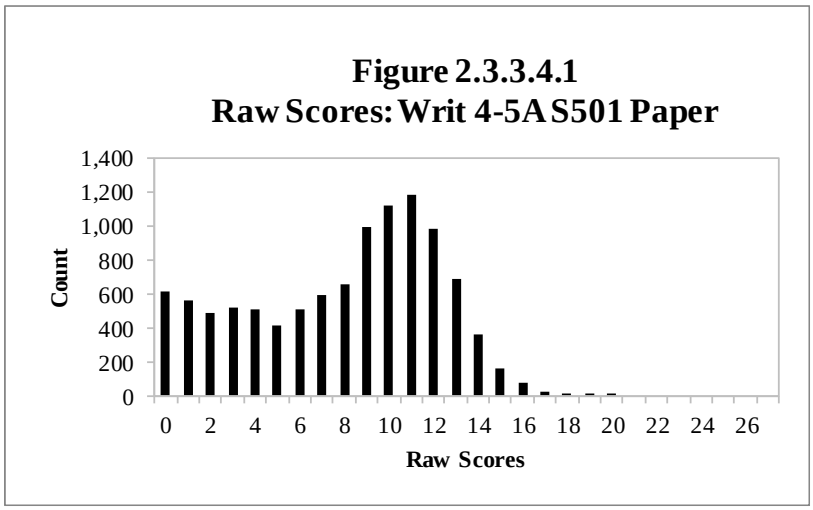
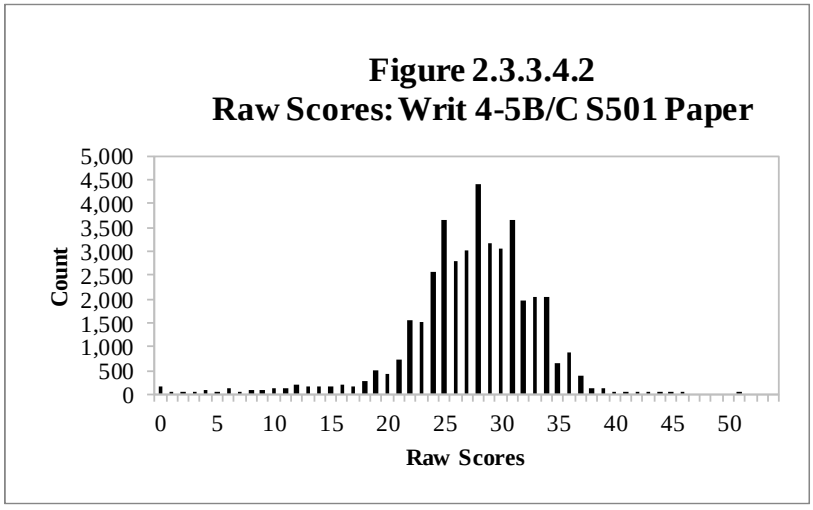


Table 2.3.3.4.2

Raw Score Descriptive Statistics: Writ 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	22,323	0	45	26.27	5.76
5	19,597	0	51	28.39	5.63
Total	41,920	0	51	27.27	5.80



2.3.3.5 Grades 6–8

Table 2.3.3.5.1

Raw Score Descriptive Statistics: Writ 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,611	0	23	8.37	4.32
7	4,400	0	24	8.64	4.19
8	4,506	0	23	9.06	4.15
Total	13,517	0	24	8.69	4.23

Figure 2.3.3.5.1
Raw Scores: Writ 6-8A S501 Paper

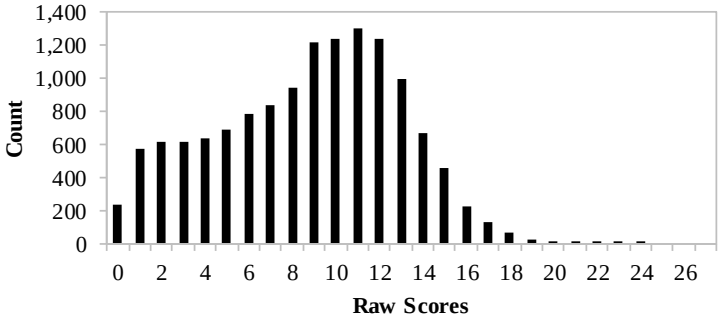
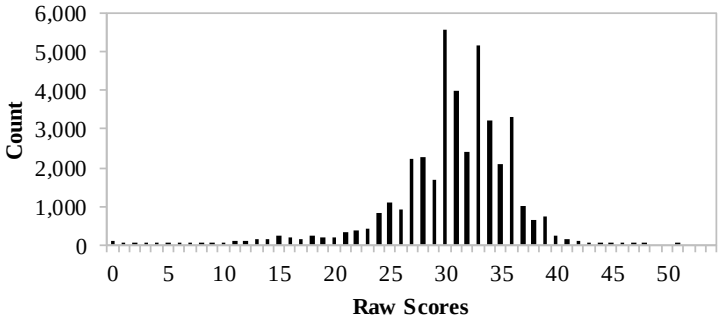


Table 2.3.3.5.2

Raw Score Descriptive Statistics: Writ 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,740	0	46	29.67	5.85
7	13,161	0	48	30.55	5.85
8	12,326	0	51	31.29	5.81
Total	41,227	0	51	30.44	5.87

Figure 2.3.3.5.2
Raw Scores: Writ 6-8B/C S501 Paper



2.3.3.6 Grades 9–12

Table 2.3.3.6.1

Raw Score Descriptive Statistics: Writ 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,307	0	20	8.03	4.65
10	3,815	0	22	8.93	4.57
11	2,890	0	21	9.94	4.41
12	1,577	0	20	10.48	4.14
Total	13,589	0	22	8.97	4.61

Figure 2.3.3.6.1
Raw Scores: Writ 9-12AS501 Paper

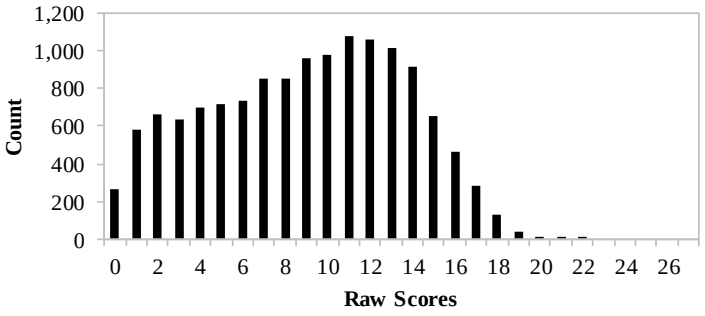
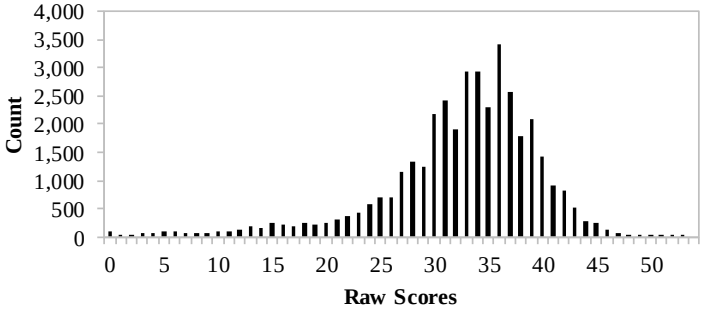


Table 2.3.3.6.2

Raw Score Descriptive Statistics: Writ 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	11,242	0	51	31.99	7.08
10	10,783	0	51	32.15	7.60
11	9,738	0	53	33.04	7.38
12	6,802	0	52	31.94	8.07
Total	38,565	0	53	32.29	7.49

Figure 2.3.3.6.2
Raw Scores: Writ 9-12B/C S501 Paper



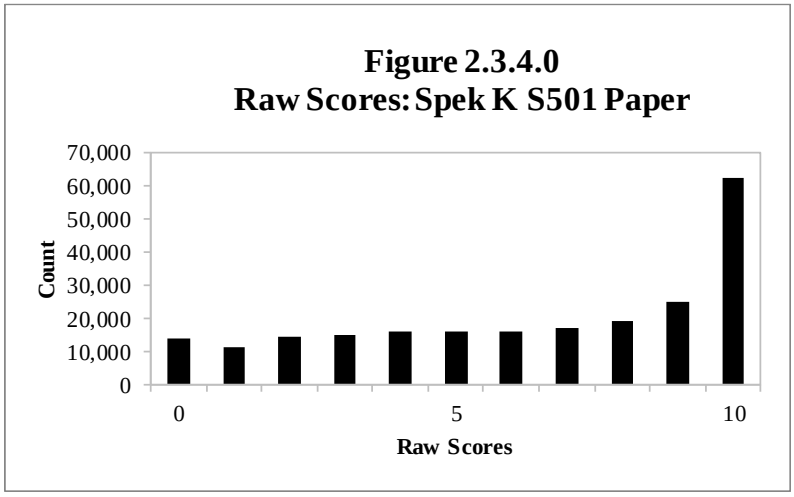
2.3.4 Speaking

2.3.4.0 Kindergarten

Table 2.3.4.0

Raw Score Descriptive Statistics: Spek K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	226,000	0	10	6.40	3.33
Total	226,000	0	10	6.40	3.33



2.3.4.1 Grade 1

Table 2.3.4.1.1

Raw Score Descriptive Statistics: Spek 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	19,813	0	18	10.20	3.88
Total	19,813	0	18	10.20	3.88

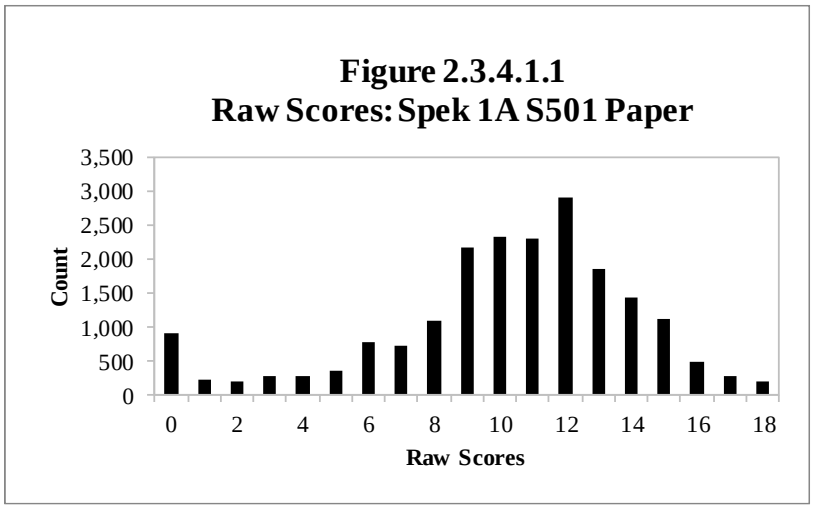
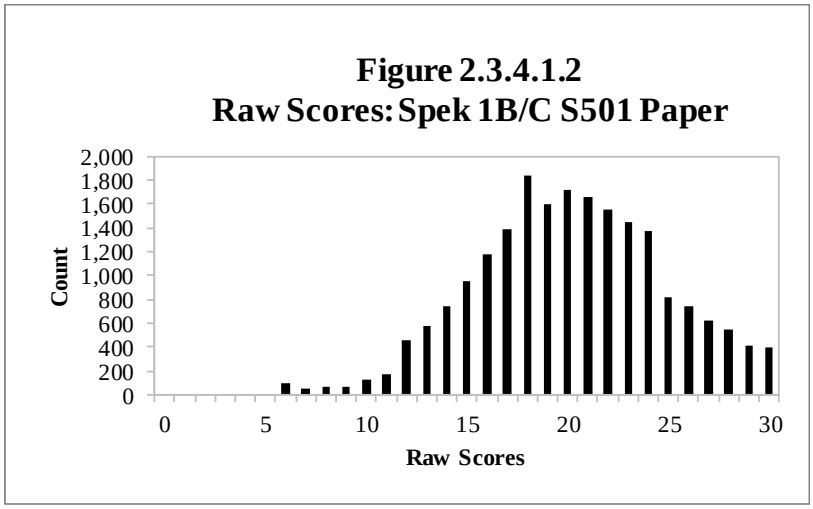


Table 2.3.4.1.2

Raw Score Descriptive Statistics: Spek 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	20,613	6	30	20.10	4.66
Total	20,613	6	30	20.10	4.66



2.3.4.2 Grade 2

Table 2.3.4.2.1

Raw Score Descriptive Statistics: Spek 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	9,039	0	18	10.40	4.59
Total	9,039	0	18	10.40	4.59

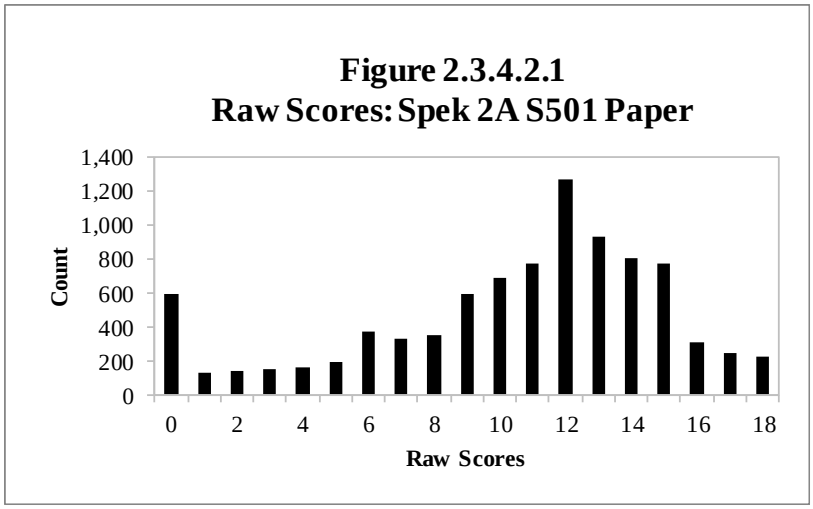
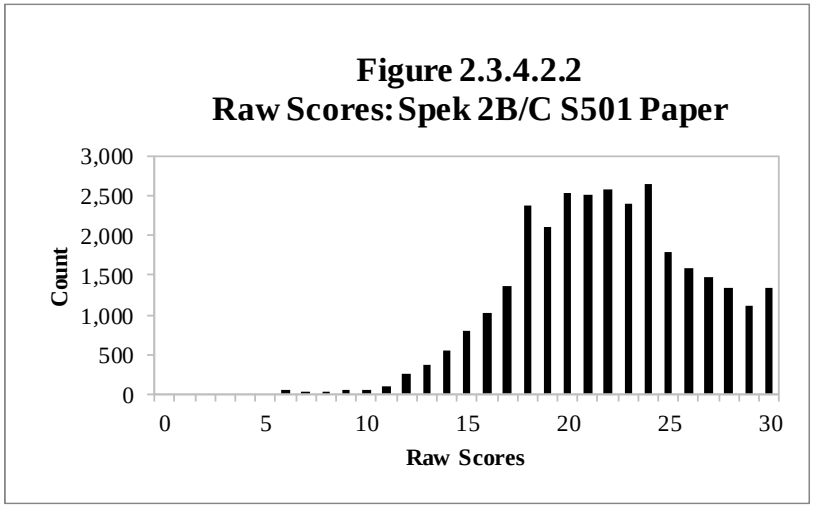


Table 2.3.4.2.2

Raw Score Descriptive Statistics: Spek 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	30,489	6	30	21.90	4.46
Total	30,489	6	30	21.90	4.46



2.3.4.3 Grade 3

Table 2.3.4.3.1

Raw Score Descriptive Statistics: Spek 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	7,385	0	18	10.57	4.56
Total	7,385	0	18	10.57	4.56

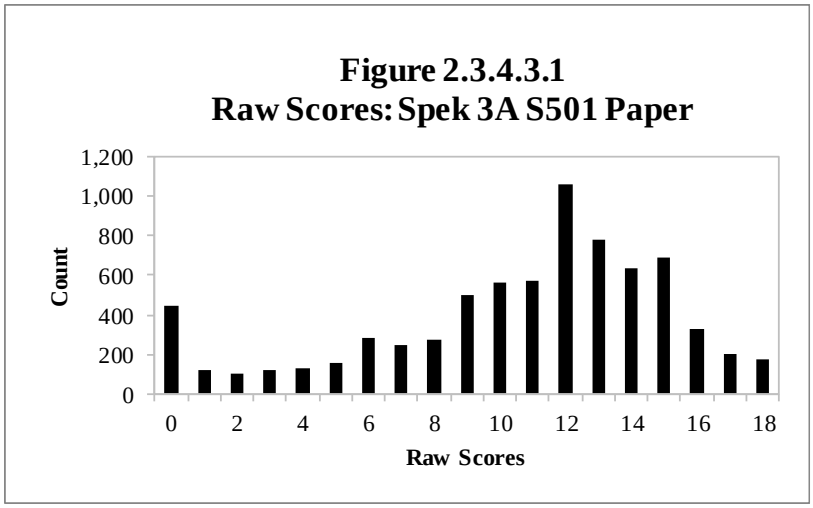
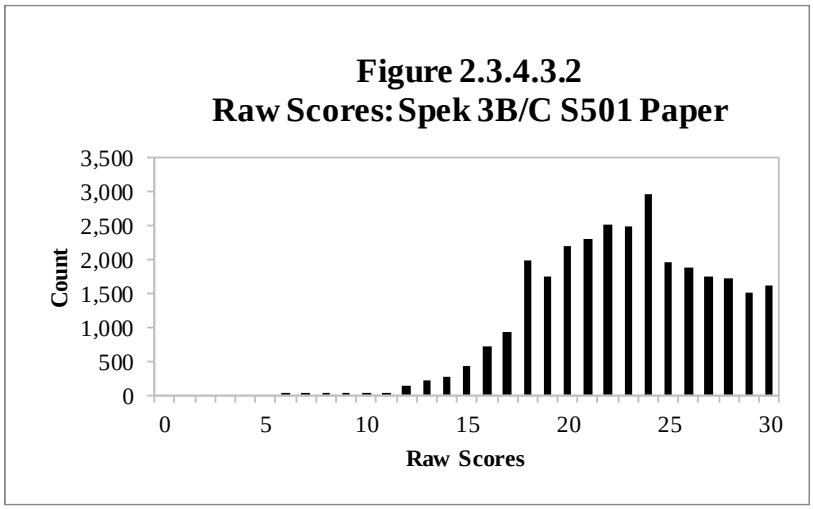


Table 2.3.4.3.2

Raw Score Descriptive Statistics: Spek 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,463	6	30	22.85	4.23
Total	29,463	6	30	22.85	4.23



2.3.4.4 Grades 4–5

Table 2.3.4.4.1

Raw Score Descriptive Statistics: Spek 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,462	0	18	9.18	4.57
5	4,934	0	18	9.28	4.52
Total	10,396	0	18	9.23	4.55

Figure 2.3.4.4.1
Raw Scores: Spek 4-5A S501 Paper

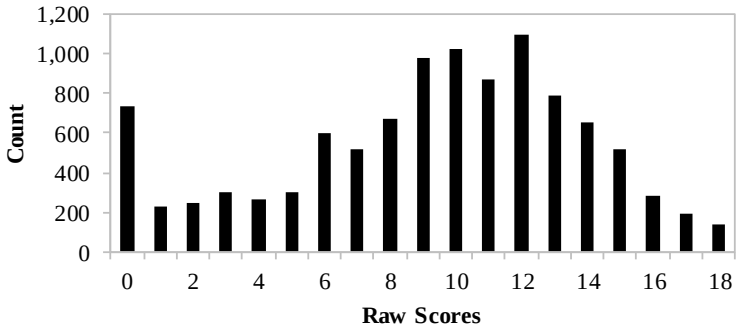
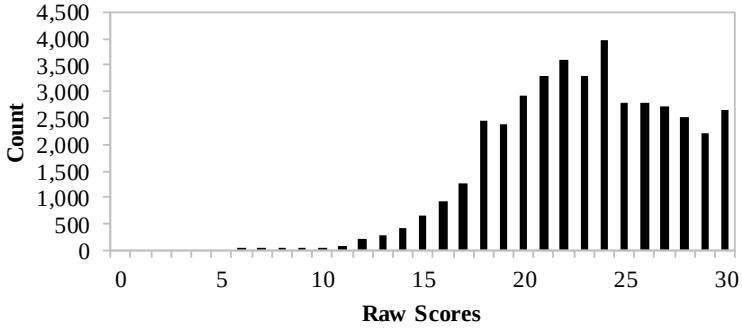


Table 2.3.4.4.2

Raw Score Descriptive Statistics: Spek 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	22,154	6	30	22.72	4.31
5	19,438	6	30	23.33	4.30
Total	41,592	6	30	23.01	4.32

Figure 2.3.4.4.2
Raw Scores: Spek 4-5B/C S501 Paper



2.3.4.5 Grades 6–8

Table 2.3.4.5.1

Raw Score Descriptive Statistics: Spek 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,562	0	18	8.71	4.52
7	4,359	0	18	8.61	4.52
8	4,455	0	18	8.80	4.44
Total	13,376	0	18	8.71	4.49

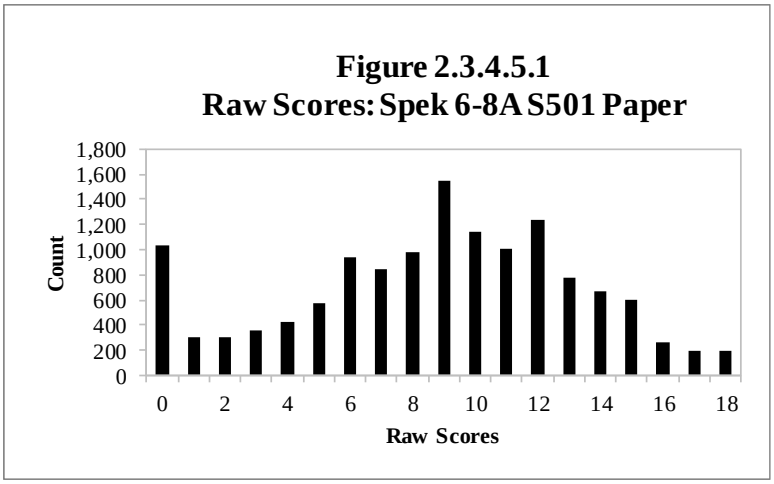
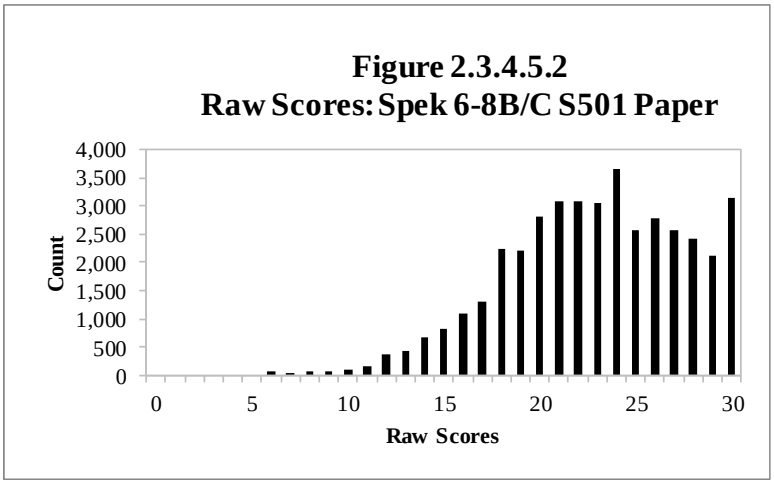


Table 2.3.4.5.2

Raw Score Descriptive Statistics: Spek 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,588	6	30	22.62	4.61
7	13,070	6	30	22.80	4.72
8	12,223	6	30	23.10	4.76
Total	40,881	6	30	22.82	4.69



2.3.4.6 Grades 9–12

Table 2.3.4.6.1

Raw Score Descriptive Statistics: Spek 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,226	0	18	8.54	4.41
10	3,741	0	18	9.37	4.23
11	2,847	0	18	9.97	4.05
12	1,553	0	18	10.44	3.86
Total	13,367	0	18	9.30	4.28

Figure 2.3.4.6.1
Raw Scores: Spek 9-12A S501 Paper

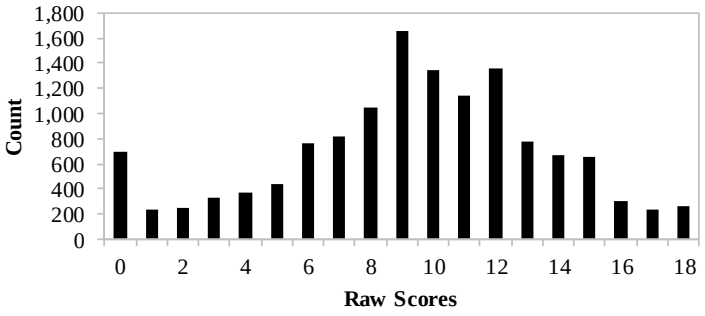
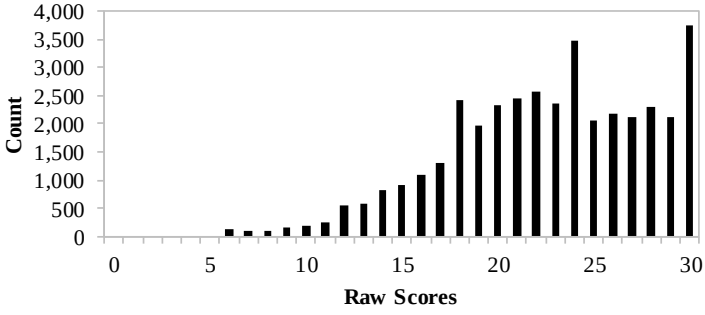


Table 2.3.4.6.2

Raw Score Descriptive Statistics: Spek 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	11,146	6	30	22.53	5.09
10	10,687	6	30	22.34	5.28
11	9,653	6	30	22.77	5.16
12	6,737	6	30	22.76	5.25
Total	38,223	6	30	22.58	5.19

Figure 2.3.4.6.2
Raw Scores: Spek 9-12B/C S501 Paper



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, we converted raw scores to vertically equated scale scores. Scale score distribution is presented by grade-level cluster and tier, and also by grade-level cluster, combining tiers.

For each test form, the figure shows the distribution of the scale scores. 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. ACCESS Paper is tiered; therefore, depending on the tiers the students were placed in, their range of possible scale scales will vary.

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 score

As is the case for raw scores, scale score distributions are impacted by the test design and student population. Scale score distribution figures for the grade-level cluster incorporate distributions from Tier A and Tier B/C test forms and so will not appear smooth.

In the domain of Writing, task weighting results in raw scores that are not smoothly distributed. This distribution is also apparent in the distribution of scale scores.

The Kindergarten test design includes skipping and stopping rules intended to reduce testing time for young children; these rules also have an impact on the distribution of raw scores and subsequently on the distribution of scale scores, leading to less smooth distributions.

2.4.1 Listening

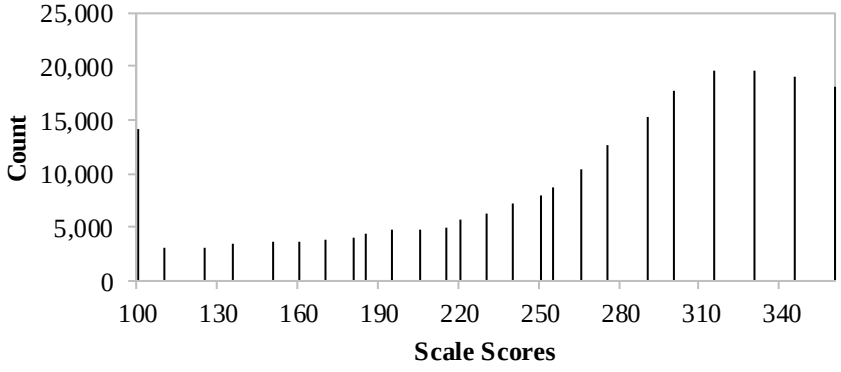
2.4.1.0 Kindergarten

Table 2.4.1.0

Scale Score Descriptive Statistics: List K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	226,001	100	363	264.50	75.97
Total	226,001	100	363	264.50	75.97

Figure 2.4.1.0
Scale Scores: List K S501 Paper



2.4.1.1 Grade 1

Table 2.4.1.1.1

Scale Score Descriptive Statistics: List 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	16,455	121	352	295.20	39.10
Total	16,455	121	352	295.20	39.10

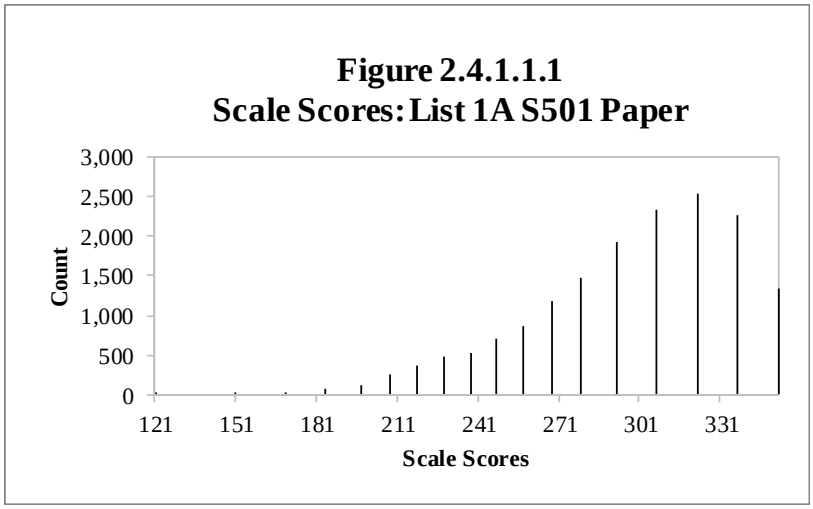


Table 2.4.1.1.2

Scale Score Descriptive Statistics: List 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	18,000	197	401	319.40	34.35
Total	18,000	197	401	319.40	34.35

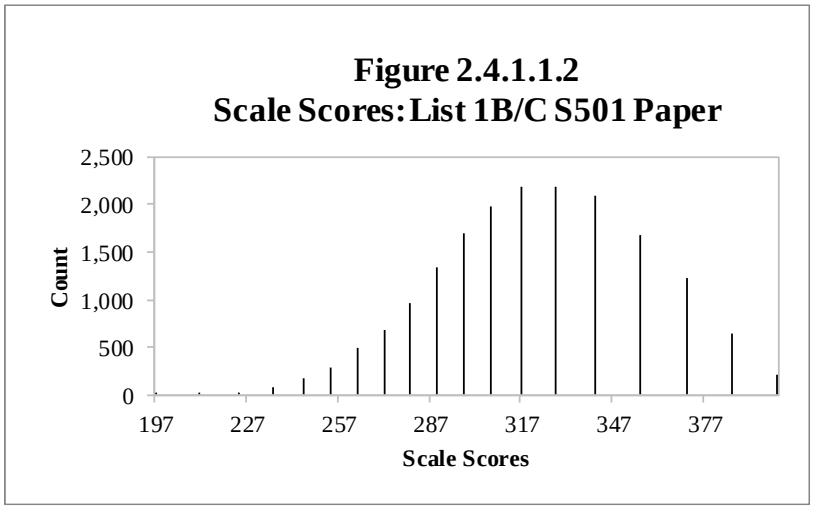
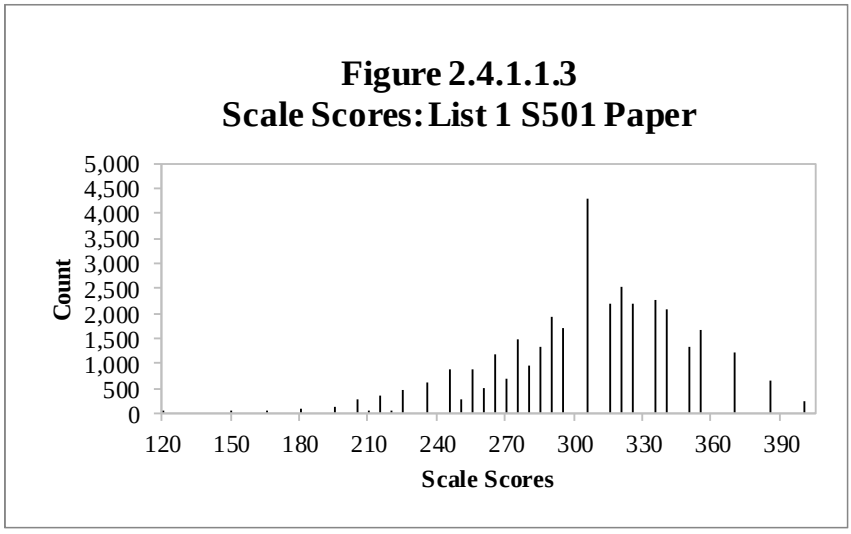


Table 2.4.1.1.3

Scale Score Descriptive Statistics: List 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	34,455	121	401	307.84	38.64
Total	34,455	121	401	307.84	38.64



2.4.1.2 Grade 2

Table 2.4.1.2.1

Scale Score Descriptive Statistics: List 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	8,088	150	352	295.44	43.66
Total	8,088	150	352	295.44	43.66

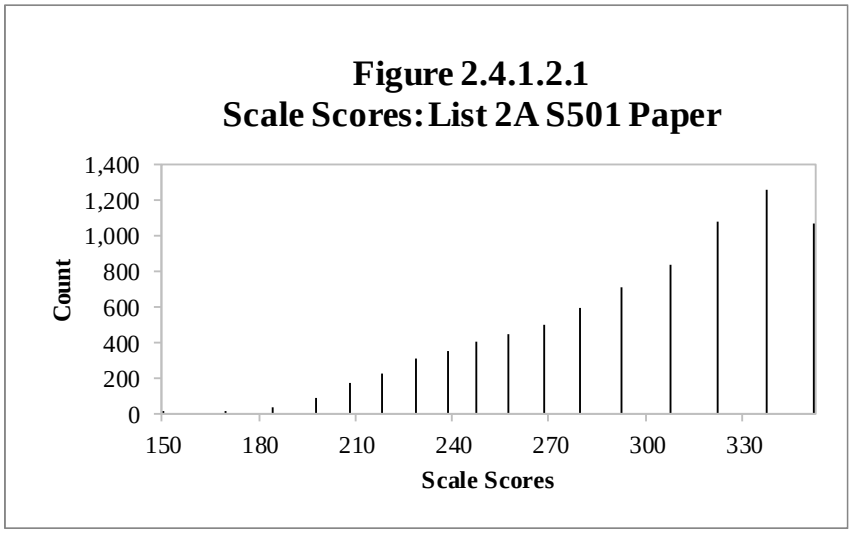


Table 2.4.1.2.2

Scale Score Descriptive Statistics: List 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	28,566	197	401	343.32	33.04
Total	28,566	197	401	343.32	33.04

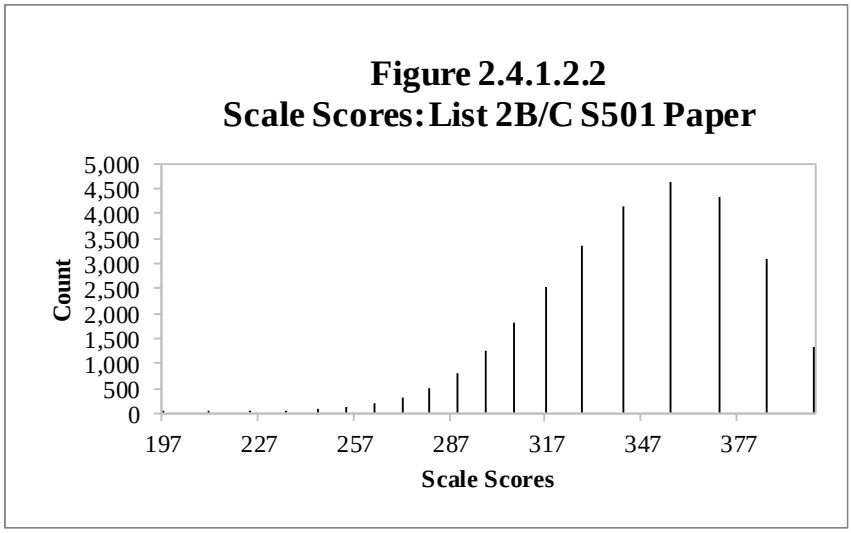
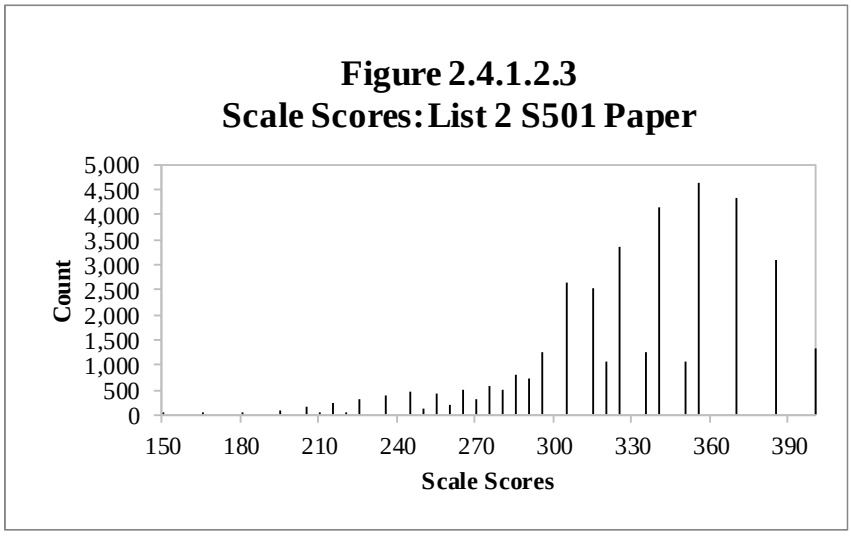


Table 2.4.1.2.3

Scale Score Descriptive Statistics: List 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	36,654	150	401	332.76	40.81
Total	36,654	150	401	332.76	40.81



2.4.1.3 Grade 3

Table 2.4.1.3.1

Scale Score Descriptive Statistics: List 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	6,723	112	416	321.97	39.70
Total	6,723	112	416	321.97	39.70

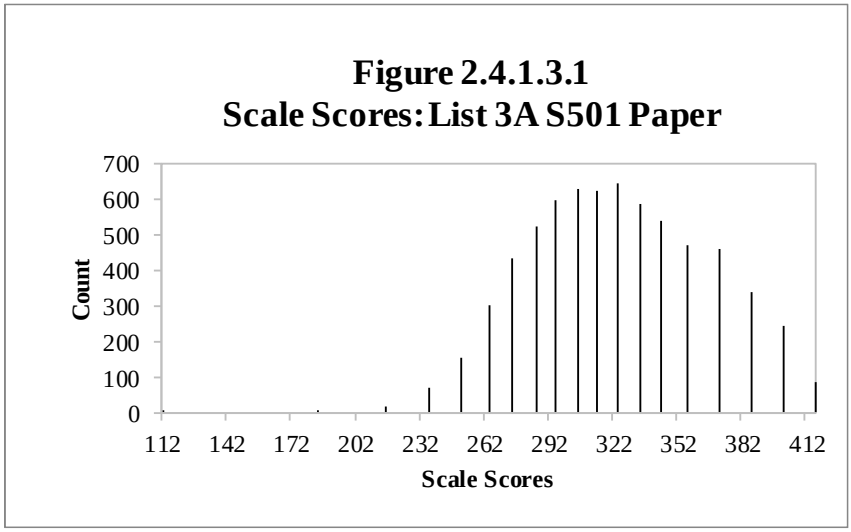


Table 2.4.1.3.2

Scale Score Descriptive Statistics: List 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	27,730	231	469	367.21	31.10
Total	27,730	231	469	367.21	31.10

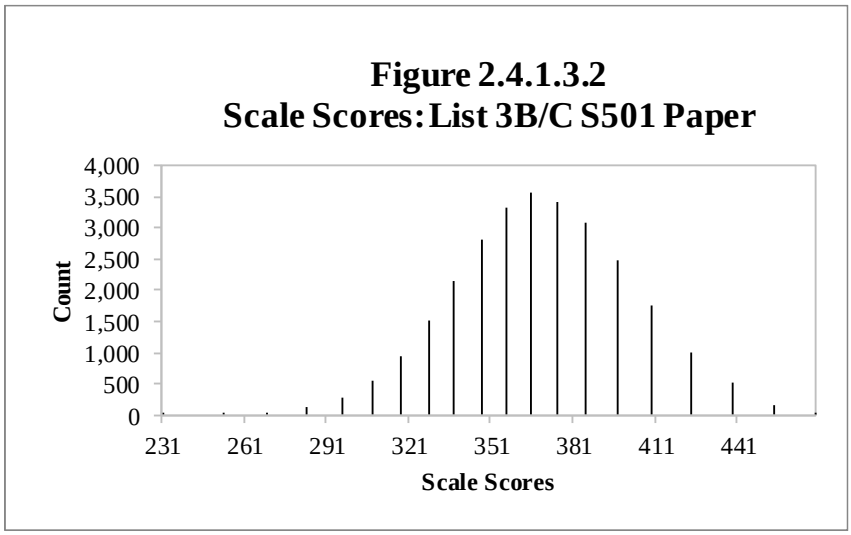
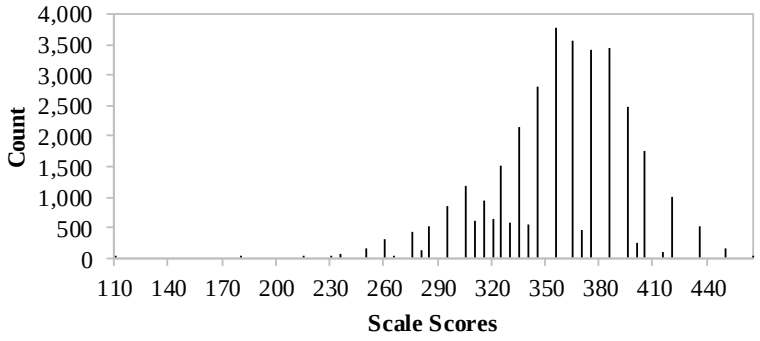


Table 2.4.1.3.3

Scale Score Descriptive Statistics: List 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	34,453	112	469	358.38	37.52
Total	34,453	112	469	358.38	37.52

Figure 2.4.1.3.3
Scale Scores: List 3 S501 Paper



2.4.1.4 Grades 4–5

Table 2.4.1.4.1

Scale Score Descriptive Statistics: List 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,036	120	416	325.27	41.20
5	4,618	184	416	331.19	41.67
Total	9,654	120	416	328.10	41.53

Figure 2.4.1.4.1
Scale Scores: List 4-5A S501 Paper

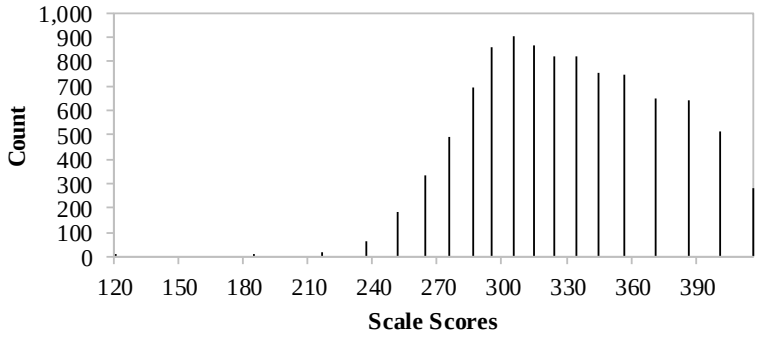


Table 2.4.1.4.2

Scale Score Descriptive Statistics: List 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	21,303	198	469	386.29	32.16
5	18,747	231	469	396.70	32.57
Total	40,050	198	469	391.17	32.77

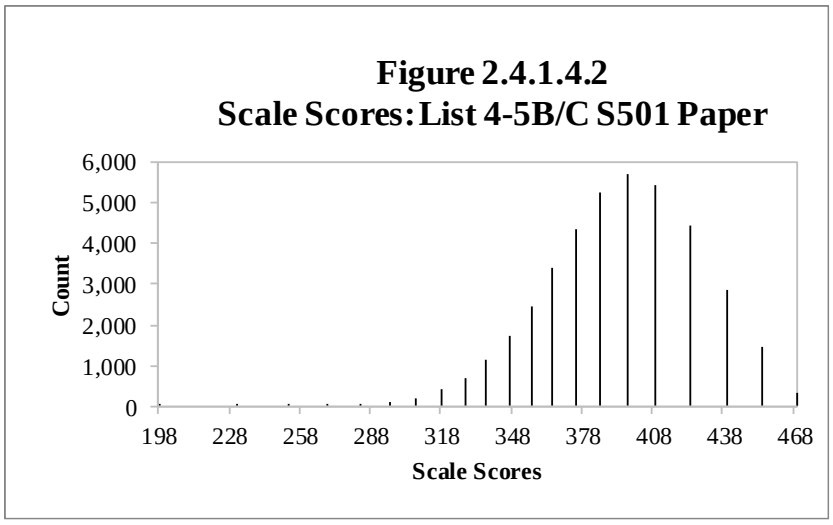
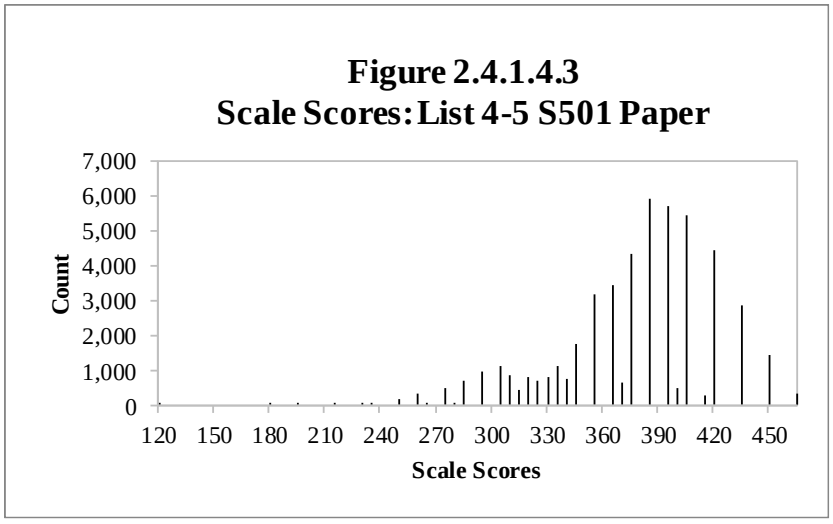


Table 2.4.1.4.3

Scale Score Descriptive Statistics: List 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	26,339	120	469	374.63	41.68
5	23,365	184	469	383.75	43.30
Total	49,704	120	469	378.92	42.69



2.4.1.5 Grades 6–8

Table 2.4.1.5.1

Scale Score Descriptive Statistics: List 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,240	181	424	323.68	41.22
7	4,080	132	424	323.47	41.20
8	4,107	181	424	324.88	40.20
Total	12,427	132	424	324.01	40.88

Figure 2.4.1.5.1
Scale Scores: List 6-8A S501 Paper

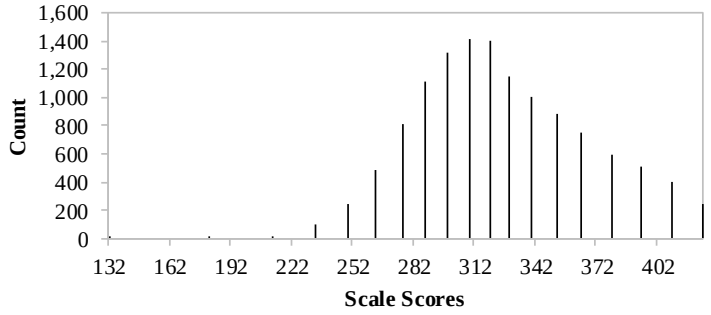


Table 2.4.1.5.2

Scale Score Descriptive Statistics: List 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,022	226	475	401.03	31.98
7	12,610	275	475	409.31	33.51
8	11,694	275	475	415.30	33.53
Total	39,326	226	475	407.93	33.46

Figure 2.4.1.5.2
Scale Scores: List 6-8B/C S501 Paper

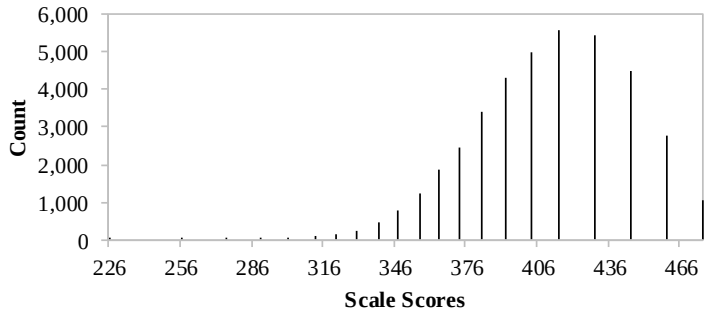
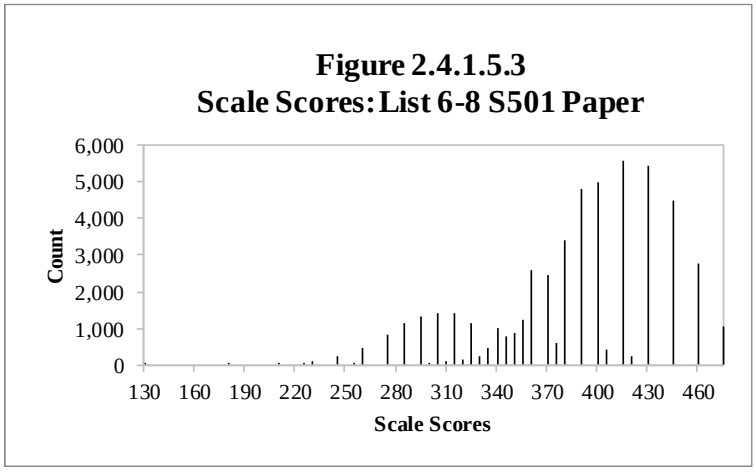


Table 2.4.1.5.3

Scale Score Descriptive Statistics: List 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	19,262	181	475	384.00	46.89
7	16,690	132	475	388.33	51.23
8	15,801	181	475	391.80	53.15
Total	51,753	132	475	387.78	50.37



2.4.1.6 Grades 9–12

Table 2.4.1.6.1

Scale Score Descriptive Statistics: List 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	4,804	169	428	321.04	36.55
10	3,507	169	428	326.20	37.01
11	2,656	223	428	331.32	36.08
12	1,462	169	428	332.78	35.10
Total	12,429	169	428	326.08	36.70

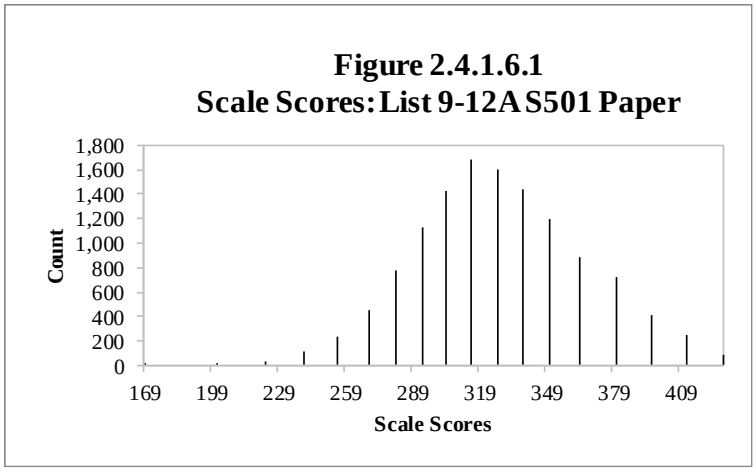


Table 2.4.1.6.2

Scale Score Descriptive Statistics: List 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	10,631	269	499	403.84	34.14
10	10,258	269	499	404.94	35.59
11	9,248	148	499	406.89	36.56
12	6,467	269	499	401.98	36.54
Total	36,604	148	499	404.59	35.63

Figure 2.4.1.6.2
Scale Scores: List 9-12B/C S501 Paper

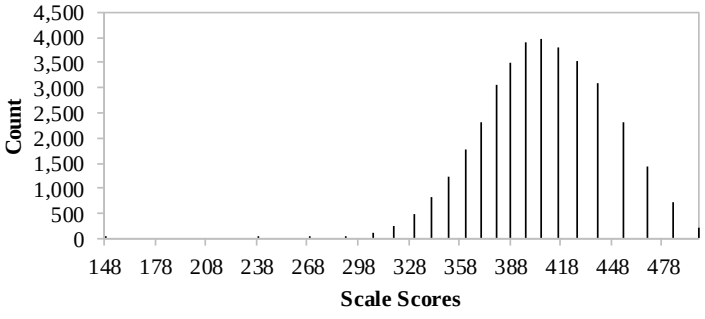
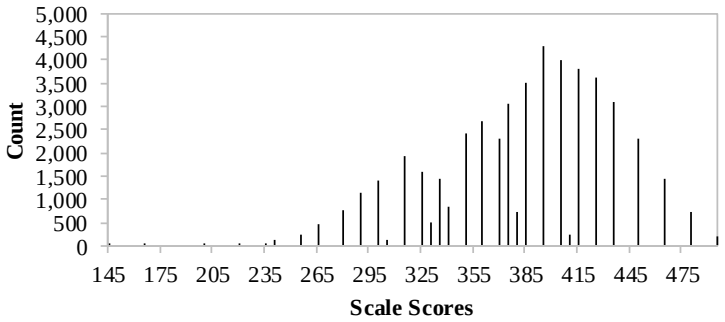


Table 2.4.1.6.3

Scale Score Descriptive Statistics: List 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	15,435	169	499	378.07	51.85
10	13,765	169	499	384.88	49.70
11	11,904	148	499	390.03	48.15
12	7,929	169	499	389.22	45.12
Total	49,033	148	499	384.69	49.55

Figure 2.4.1.6.3
Scale Scores: List 9-12 S501 Paper



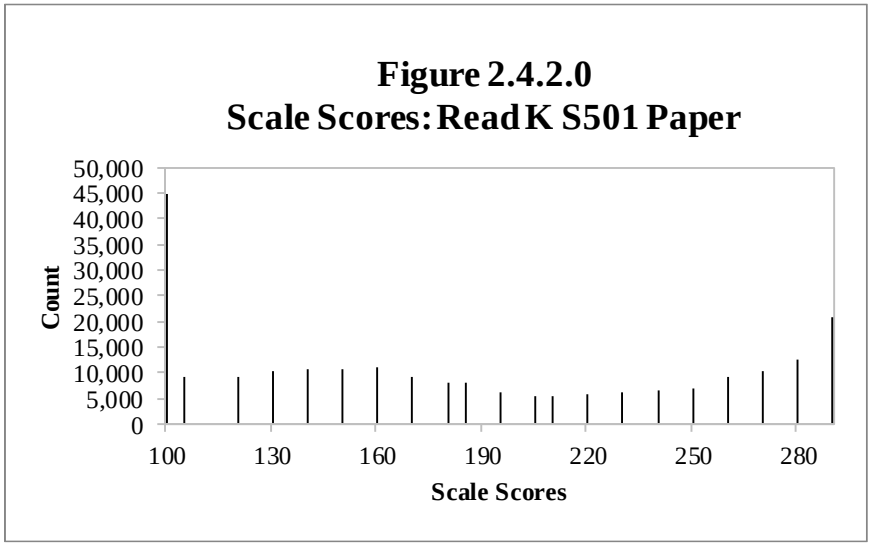
2.4.2 Reading

2.4.2.0 Kindergarten

Table 2.4.2.0

Scale Score Descriptive Statistics: Read K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,994	100	290	183.84	67.44
Total	225,994	100	290	183.84	67.44



2.4.2.1 Grade 1

Table 2.4.2.1.1

Scale Score Descriptive Statistics: Read 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	15,533	180	353	272.87	26.76
Total	15,533	180	353	272.87	26.76

Figure 2.4.2.1.1
Scale Scores: Read 1A S501 Paper

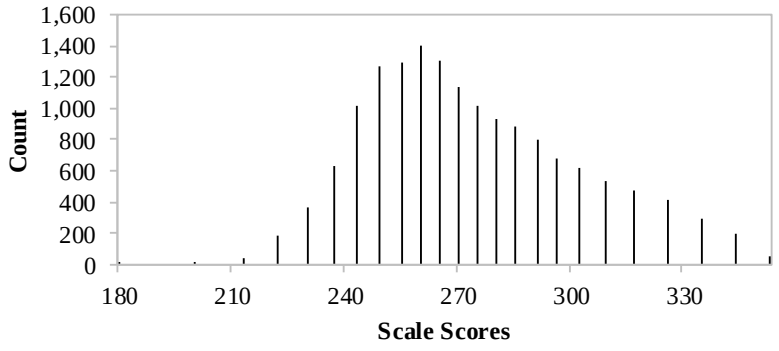


Table 2.4.2.1.2

Scale Score Descriptive Statistics: Read 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	15,304	141	392	306.27	23.83
Total	15,304	141	392	306.27	23.83

Figure 2.4.2.1.2
Scale Scores: Read 1B/C S501 Paper

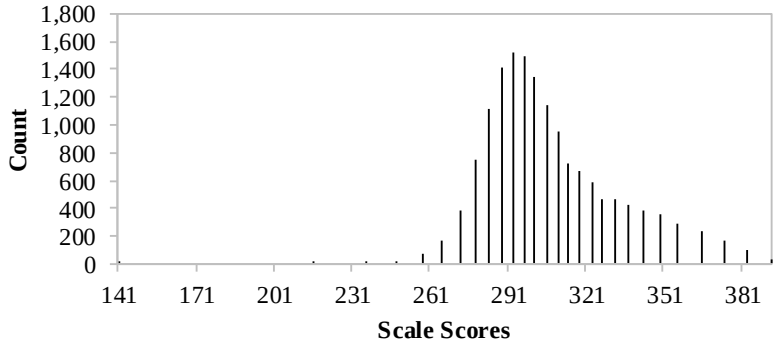
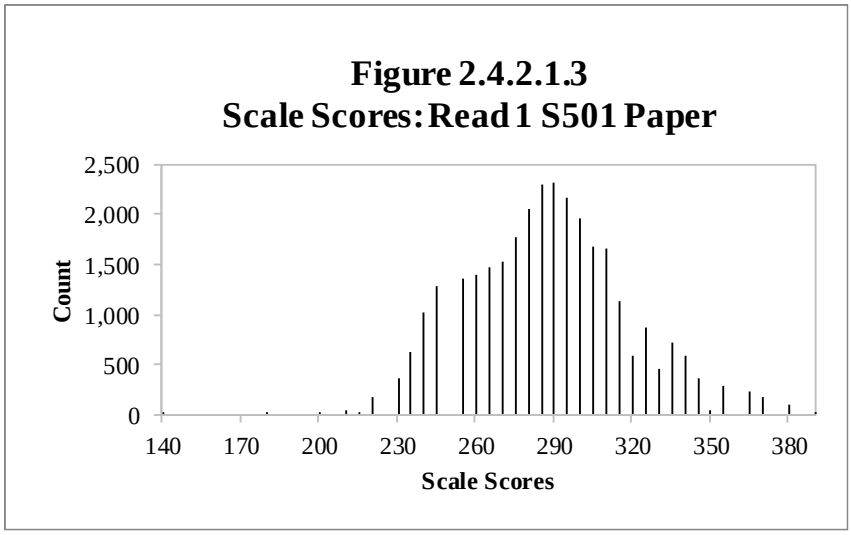


Table 2.4.2.1.3

Scale Score Descriptive Statistics: Read 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	30,837	141	392	289.45	30.35
Total	30,837	141	392	289.45	30.35



2.4.2.2 Grade 2

Table 2.4.2.2.1

Scale Score Descriptive Statistics: Read 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	7,695	180	353	282.66	30.95
Total	7,695	180	353	282.66	30.95

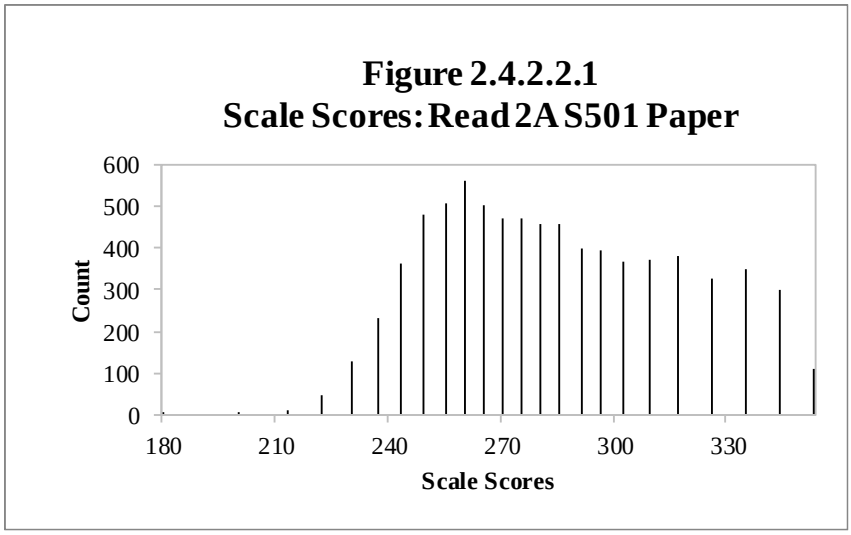


Table 2.4.2.2.2

Scale Score Descriptive Statistics: Read 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	25,317	236	392	326.79	28.95
Total	25,317	236	392	326.79	28.95

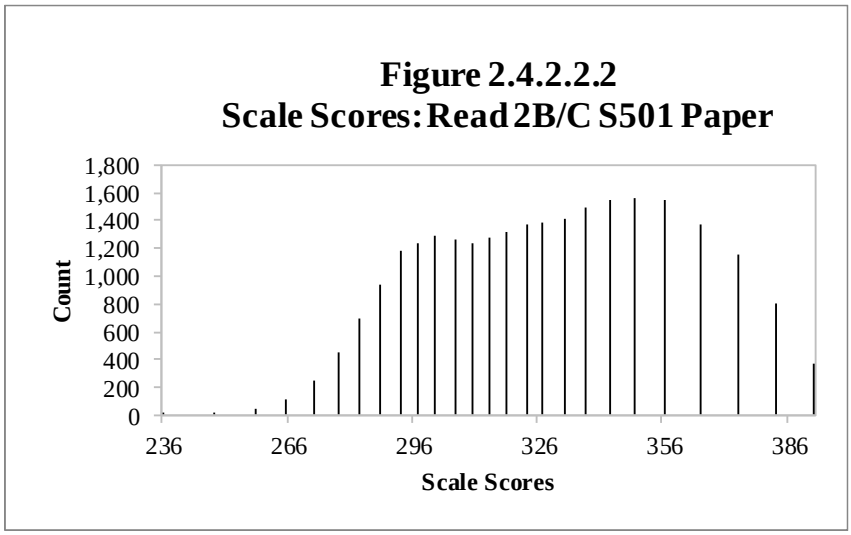
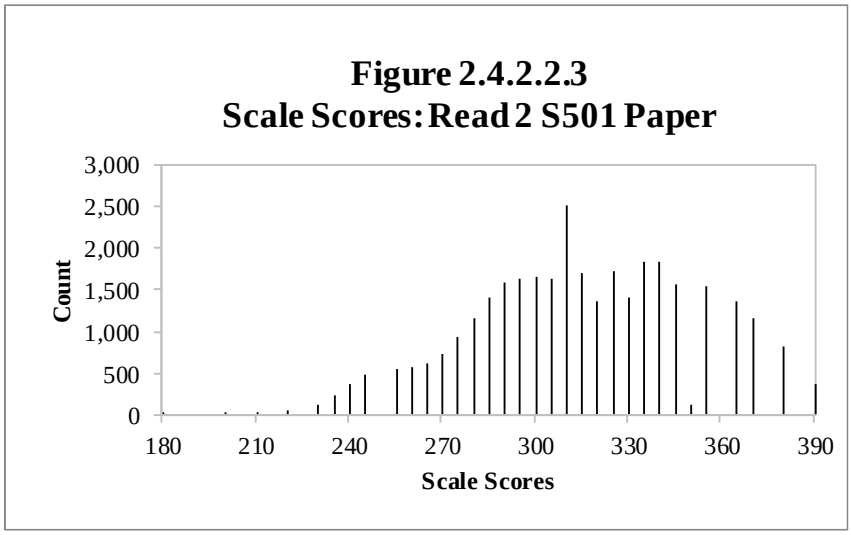


Table 2.4.2.2.3

Scale Score Descriptive Statistics: Read 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	33,012	180	392	316.50	34.85
Total	33,012	180	392	316.50	34.85



2.4.2.3 Grade 3

Table 2.4.2.3.1

Scale Score Descriptive Statistics: Read 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	6,301	234	389	301.84	28.40
Total	6,301	234	389	301.84	28.40

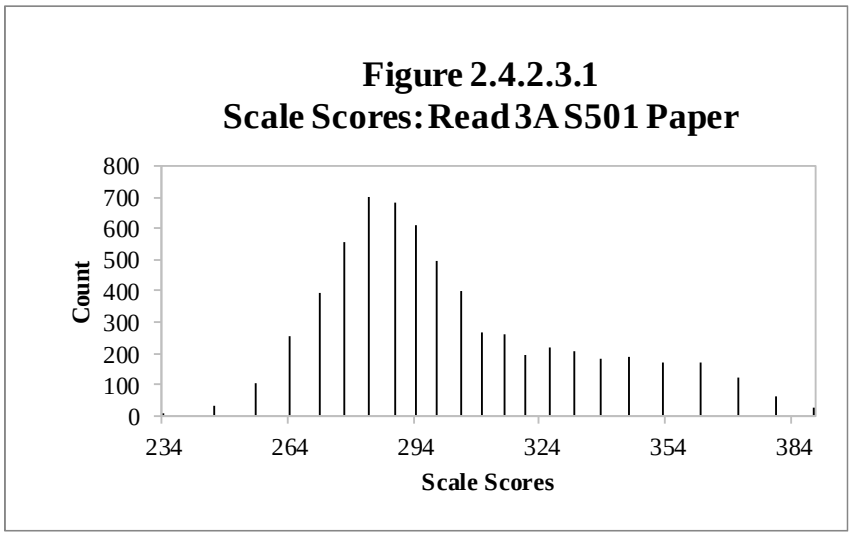


Table 2.4.2.3.2

Scale Score Descriptive Statistics: Read 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	24,834	158	447	346.68	19.26
Total	24,834	158	447	346.68	19.26

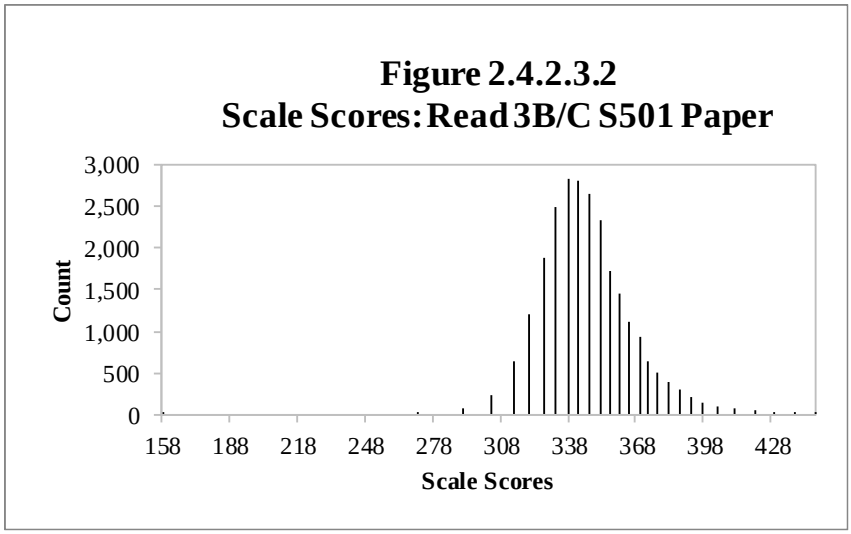
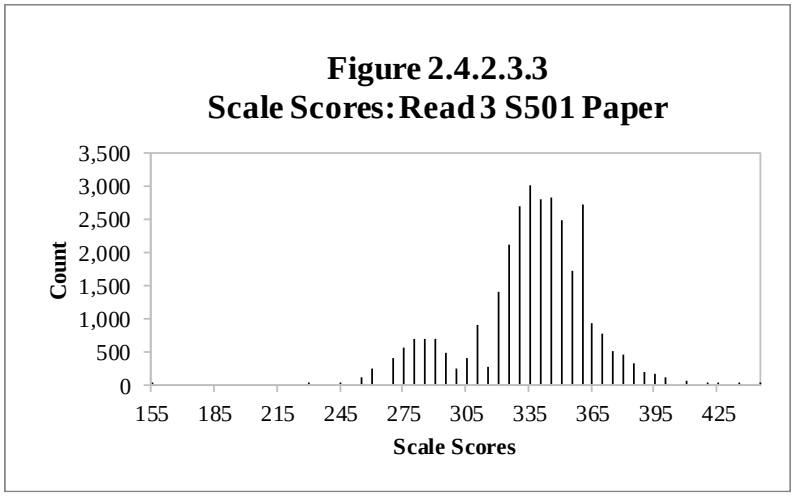


Table 2.4.2.3.3

Scale Score Descriptive Statistics: Read 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	31,135	158	447	337.61	27.99
Total	31,135	158	447	337.61	27.99



2.4.2.4 Grades 4–5

Table 2.4.2.4.1

Scale Score Descriptive Statistics: Read 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	4,814	234	389	309.44	31.77
5	4,461	213	389	315.45	33.09
Total	9,275	213	389	312.33	32.55

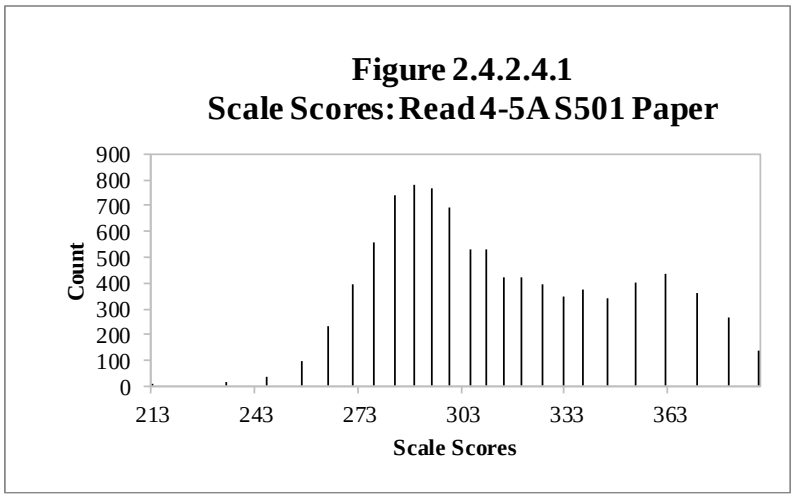


Table 2.4.2.4.2

Scale Score Descriptive Statistics: Read 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	19,318	271	447	358.45	21.59
5	17,420	271	447	367.20	23.33
Total	36,738	271	447	362.60	22.86

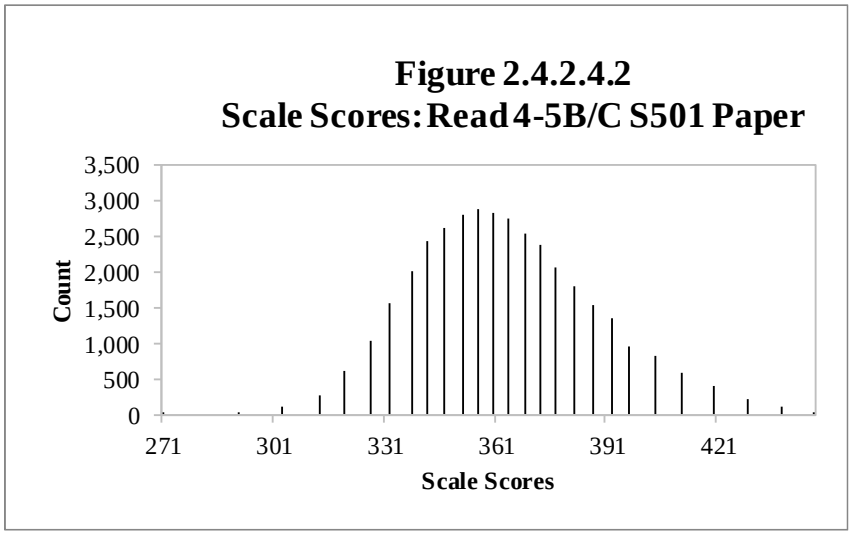
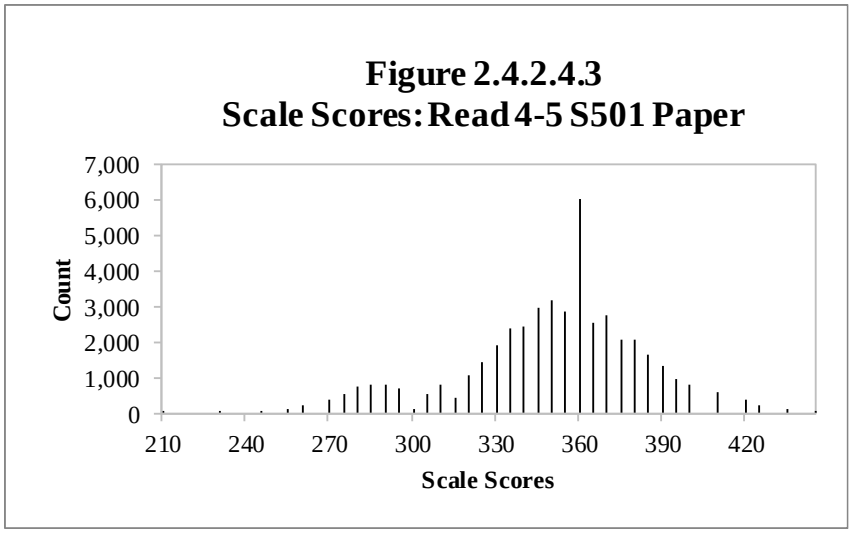


Table 2.4.2.4.3

Scale Score Descriptive Statistics: Read 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	24,132	234	447	348.67	30.95
5	21,881	213	447	356.65	33.04
Total	46,013	213	447	352.47	32.21



2.4.2.5 Grades 6–8

Table 2.4.2.5.1

Scale Score Descriptive Statistics: Read 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,057	256	416	326.95	26.53
7	3,961	235	416	329.86	27.75
8	4,025	256	416	334.06	27.45
Total	12,043	235	416	330.28	27.40

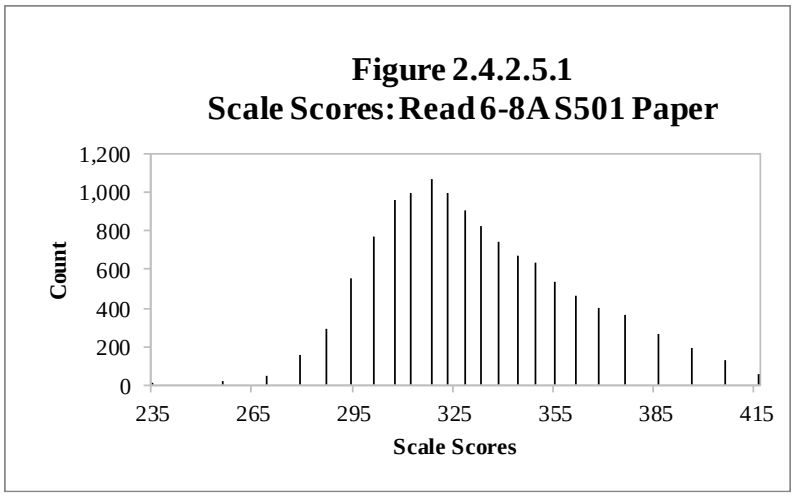


Table 2.4.2.5.2

Scale Score Descriptive Statistics: Read 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	13,110	276	452	361.41	21.35
7	11,123	276	452	367.50	23.92
8	10,334	276	452	374.63	25.45
Total	34,567	276	452	367.32	24.08

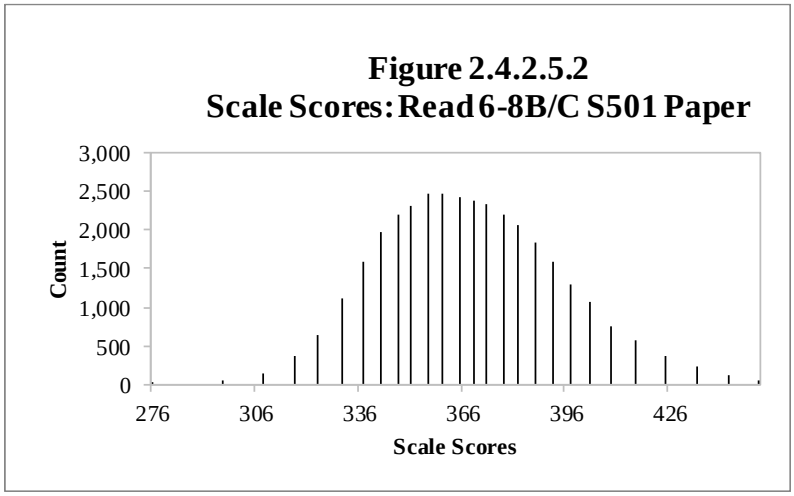
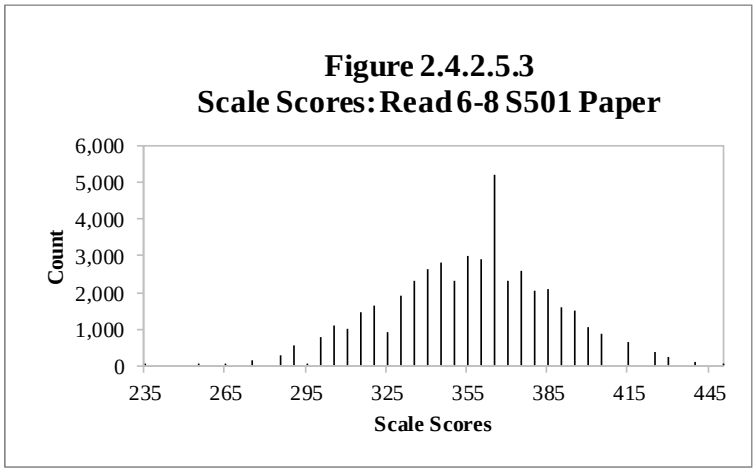


Table 2.4.2.5.3

Scale Score Descriptive Statistics: Read 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,167	256	452	353.27	27.00
7	15,084	235	452	357.62	29.97
8	14,359	256	452	363.26	31.77
Total	46,610	235	452	357.75	29.78



2.4.2.6 Grades 9–12

Table 2.4.2.6.1

Scale Score Descriptive Statistics: Read 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	4,759	249	433	350.37	27.40
10	3,488	270	433	356.83	29.45
11	2,666	283	433	362.71	29.48
12	1,452	283	433	364.29	28.26
Total	12,365	249	433	356.49	29.06

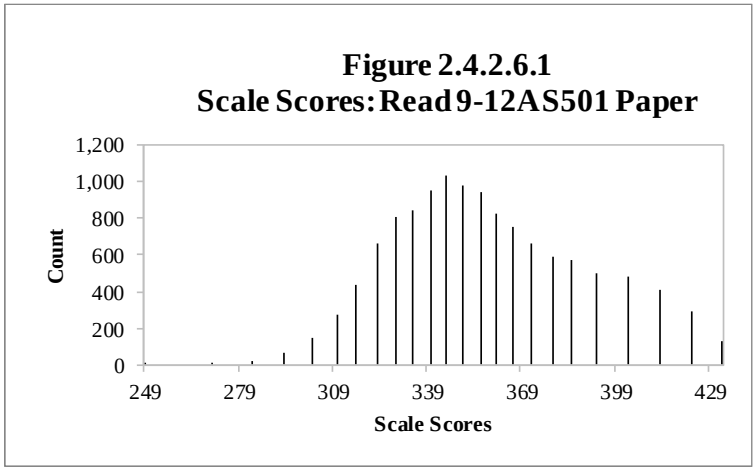


Table 2.4.2.6.2

Scale Score Descriptive Statistics: Read 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	9,437	233	470	389.86	25.91
10	9,217	297	470	392.90	26.55
11	8,463	297	470	396.44	27.82
12	6,011	233	470	392.80	27.97
Total	33,128	233	470	392.92	27.07

Figure 2.4.2.6.2
Scale Scores: Read 9-12B/C S501 Paper

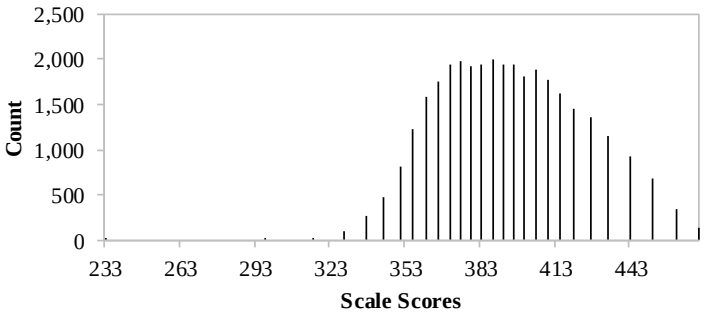
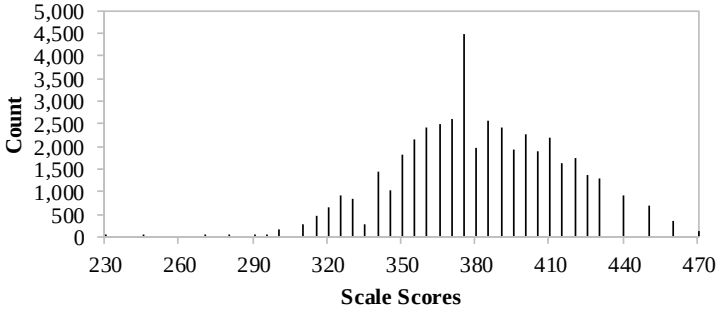


Table 2.4.2.6.3

Scale Score Descriptive Statistics: Read 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	14,196	233	470	376.62	32.33
10	12,705	270	470	383.00	31.76
11	11,129	283	470	388.36	31.69
12	7,463	233	470	387.25	30.21
Total	45,493	233	470	383.02	32.03

Figure 2.4.2.6.3
Scale Scores: Read 9-12 S501 Paper



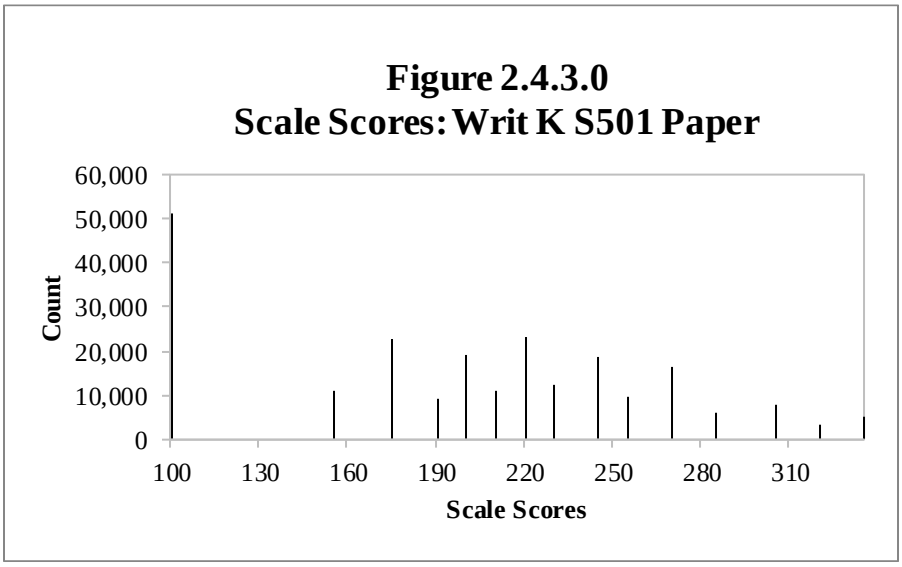
2.4.3 Writing

2.4.3.0 Kindergarten

Table 2.4.3.0

Scale Score Descriptive Statistics: Writ K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,987	100	339	199.83	66.87
Total	225,987	100	339	199.83	66.87



2.4.3.1 Grade 1

Table 2.4.3.1.1

Scale Score Descriptive Statistics: Writ 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	19,965	111	321	235.77	36.36
Total	19,965	111	321	235.77	36.36

Figure 2.4.3.1.1
Scale Scores: Writ 1A S501 Paper

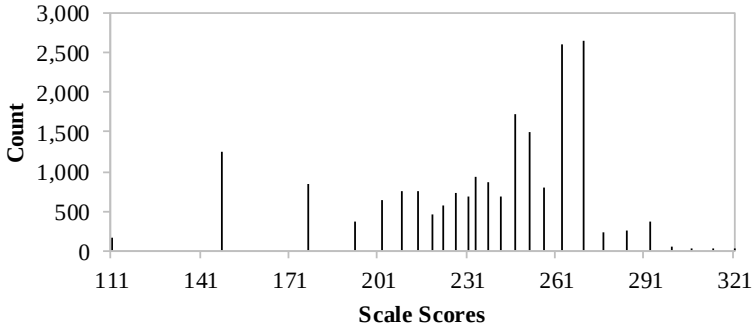


Table 2.4.3.1.2

Scale Score Descriptive Statistics: Writ 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	20,779	111	400	272.50	41.33
Total	20,779	111	400	272.50	41.33

Figure 2.4.3.1.2
Scale Scores: Writ 1B/C S501 Paper

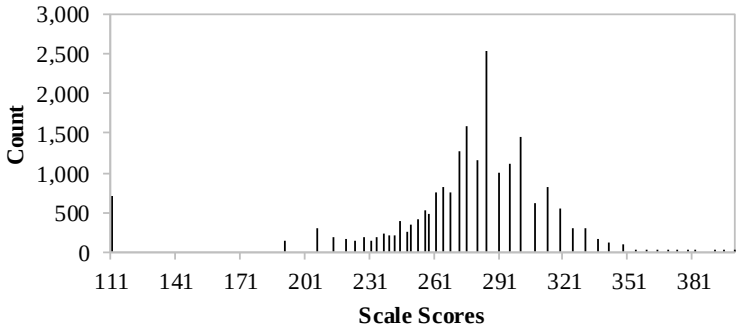
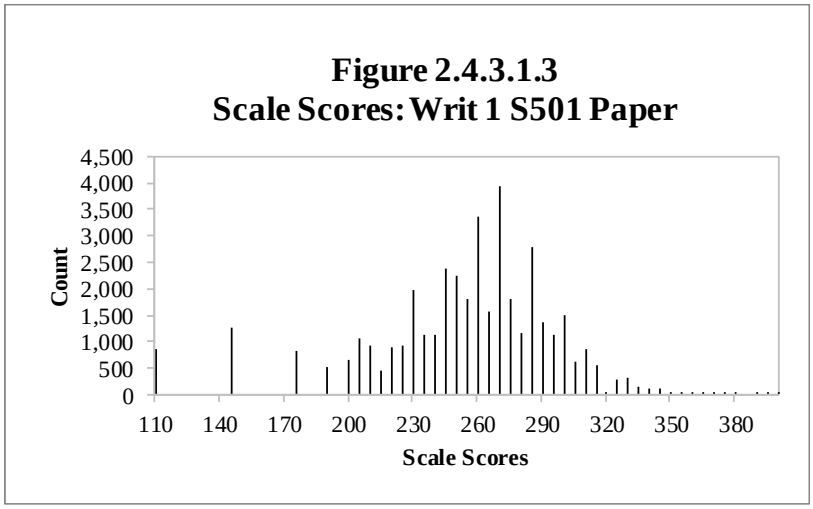


Table 2.4.3.1.3

Scale Score Descriptive Statistics: Writ 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	40,744	111	400	254.50	43.08
Total	40,744	111	400	254.50	43.08



2.4.3.2 Grade 2

Table 2.4.3.2.1

Scale Score Descriptive Statistics: Writ 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	9,124	133	347	241.07	46.43
Total	9,124	133	347	241.07	46.43

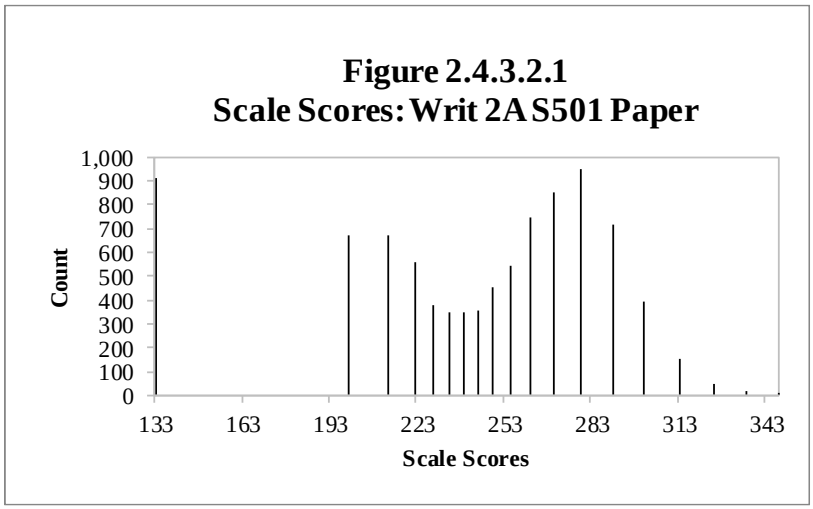


Table 2.4.3.2.2

Scale Score Descriptive Statistics: Writ 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	30,694	133	407	298.40	33.43
Total	30,694	133	407	298.40	33.43

Figure 2.4.3.2.2
Scale Scores: Writ B/C S501 Paper

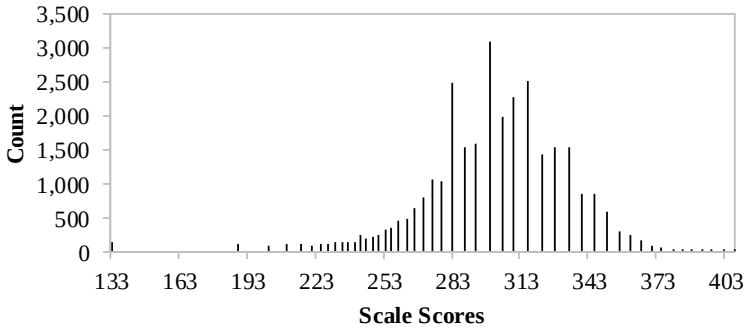
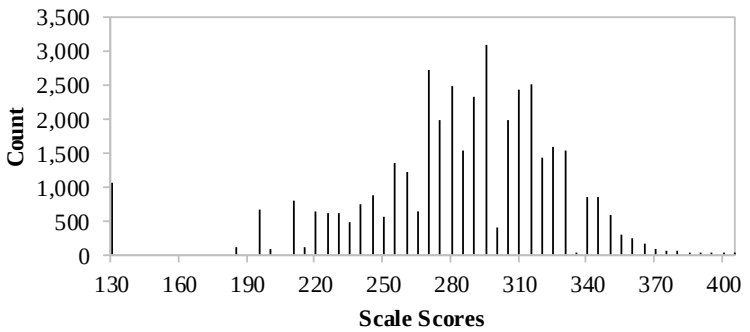


Table 2.4.3.2.3

Scale Score Descriptive Statistics: Writ 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	39,818	133	407	285.26	44.00
Total	39,818	133	407	285.26	44.00

Figure 2.4.3.2.3
Scale Scores: Writ 2 S501 Paper



2.4.3.3 Grade 3

Table 2.4.3.3.1

Scale Score Descriptive Statistics: Writ 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	7,464	133	347	251.32	42.76
Total	7,464	133	347	251.32	42.76

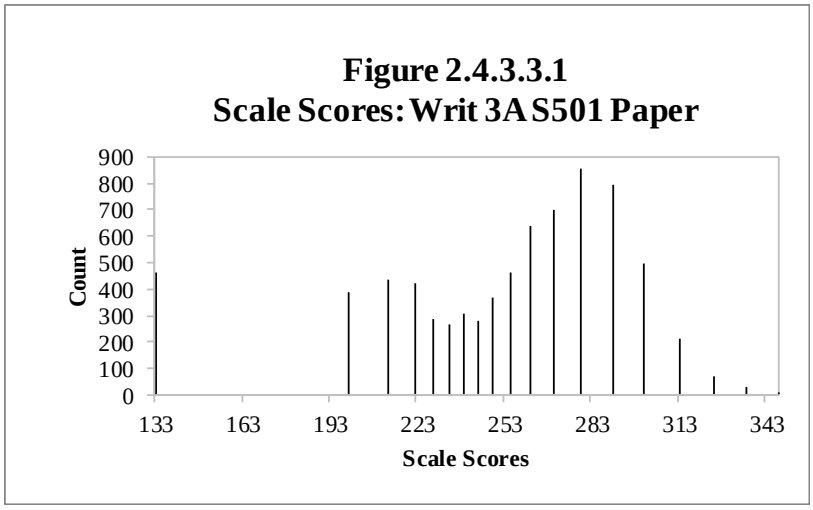


Table 2.4.3.3.2

Scale Score Descriptive Statistics: Writ 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,669	133	428	312.39	30.45
Total	29,669	133	428	312.39	30.45

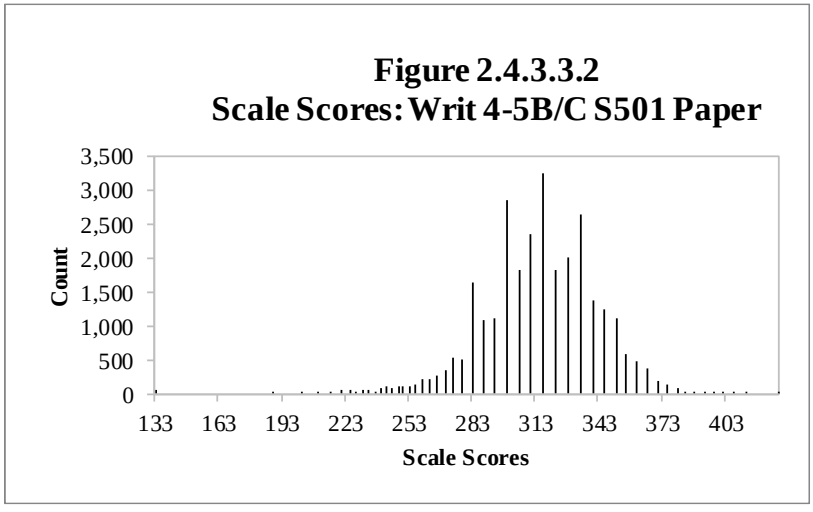
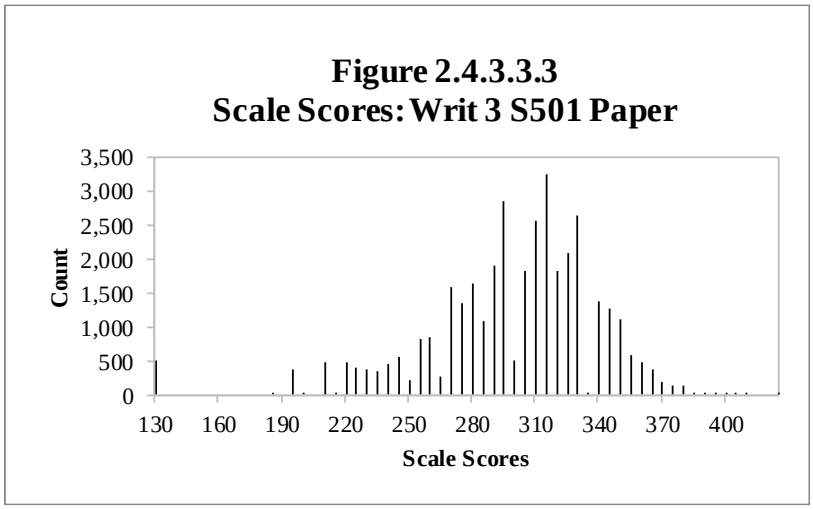


Table 2.4.3.3.3

Scale Score Descriptive Statistics: Writ 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	37,133	133	428	300.11	41.32
Total	37,133	133	428	300.11	41.32



2.4.3.4 Grades 4–5

Table 2.4.3.4.1

Scale Score Descriptive Statistics: Writ 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,418	155	398	275.40	41.66
5	4,936	155	378	281.36	39.42
Total	10,354	155	398	278.24	40.71

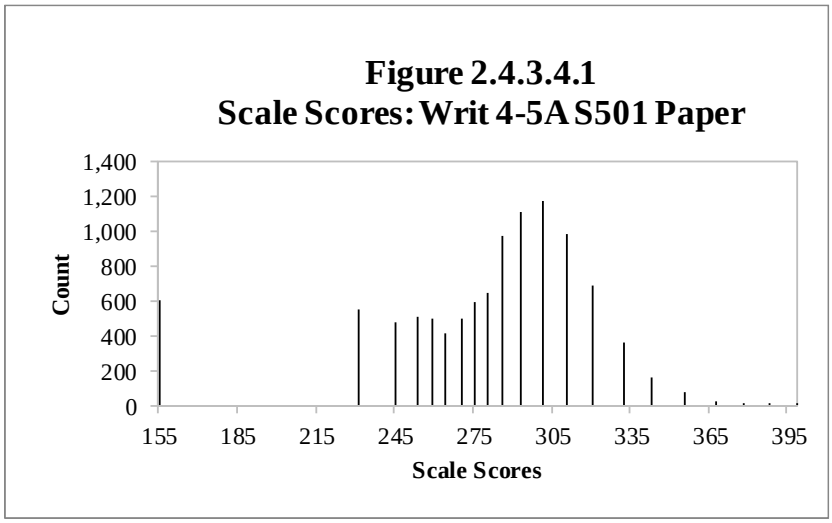


Table 2.4.3.4.2

Scale Score Descriptive Statistics: Writ 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	22,042	155	435	339.19	28.13
5	19,357	155	439	350.22	28.96
Total	41,399	155	439	344.35	29.05

Figure 2.4.3.4.2
Scale Scores: Writ 4-5B/C S501 Paper

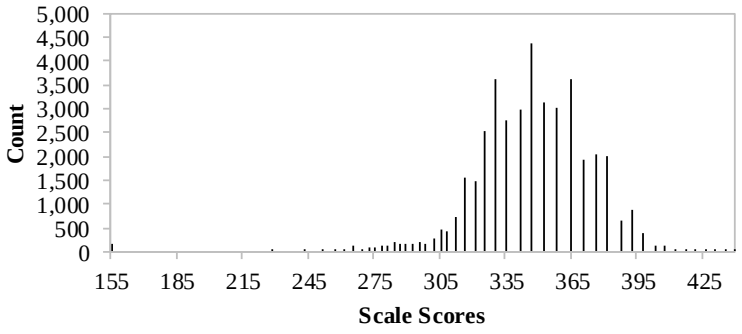
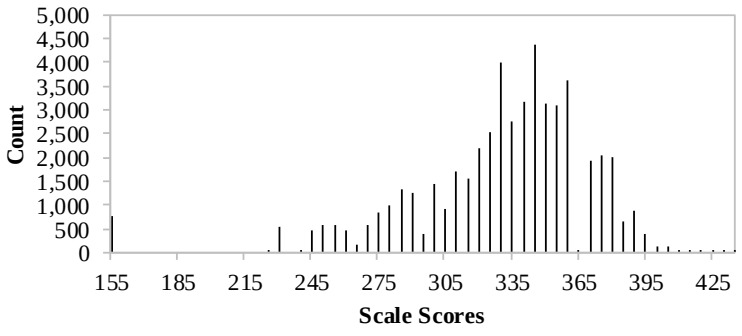


Table 2.4.3.4.3

Scale Score Descriptive Statistics: Writ 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	27,460	155	435	326.60	40.27
5	24,293	155	439	336.23	41.85
Total	51,753	155	439	331.12	41.30

Figure 2.4.3.4.3
Scale Scores: Writ 4-5 S501 Paper



2.4.3.5 Grades 6–8

Table 2.4.3.5.1

Scale Score Descriptive Statistics: Writ 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,583	188	416	276.44	34.04
7	4,376	188	425	278.60	32.71
8	4,480	188	416	281.78	32.92
Total	13,439	188	425	278.92	33.31

Figure 2.4.3.5.1
Scale Scores: Writ 6-8A S501 Paper

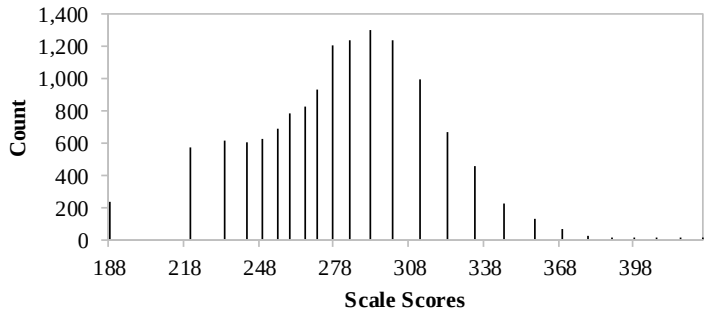


Table 2.4.3.5.2

Scale Score Descriptive Statistics: Writ 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,578	188	427	345.58	28.74
7	13,017	188	436	350.40	29.02
8	12,172	188	456	354.30	29.09
Total	40,767	188	456	349.72	29.16

Figure 2.4.3.5.2
Scale Scores: Writ 6-8B/C S501 Paper

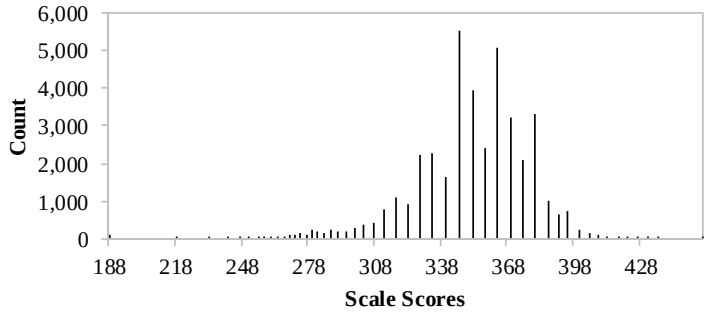
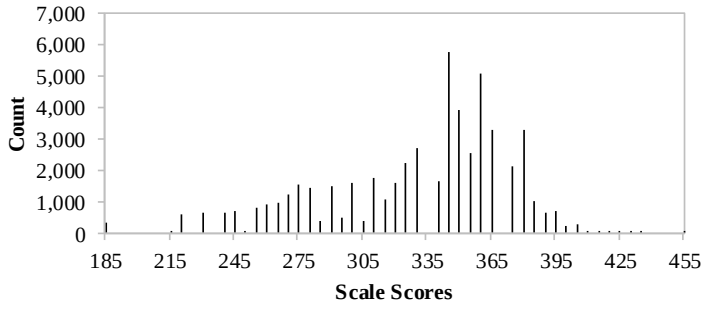


Table 2.4.3.5.3

Scale Score Descriptive Statistics: Writ 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	20,161	188	427	329.86	41.73
7	17,393	188	436	332.34	43.25
8	16,652	188	456	334.79	44.10
Total	54,206	188	456	332.17	43.00

Figure 2.4.3.5.3
Scale Scores: Writ 6-8 S501 Paper



2.4.3.6 Grades 9–12

Table 2.4.3.6.1

Scale Score Descriptive Statistics: Writ 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,279	232	421	307.44	36.52
10	3,803	232	439	314.58	36.25
11	2,873	232	430	322.53	35.96
12	1,571	232	421	326.53	34.22
Total	13,526	232	439	314.87	36.74

Figure 2.4.3.6.1
Scale Scores: Writ 9-12A S501 Paper

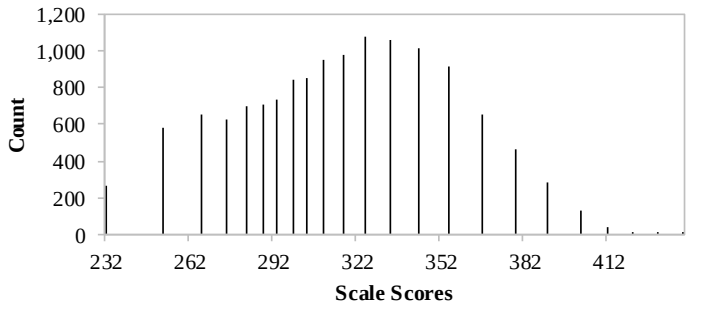


Table 2.4.3.6.2

Scale Score Descriptive Statistics: Writ 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	11,129	232	471	373.65	35.04
10	10,717	232	471	374.73	37.09
11	9,681	232	500	379.29	36.32
12	6,754	232	482	373.79	39.10
Total	38,281	232	500	375.40	36.75

Figure 2.4.3.6.2
Scale Scores: Writ 9-12B/C S501 Paper

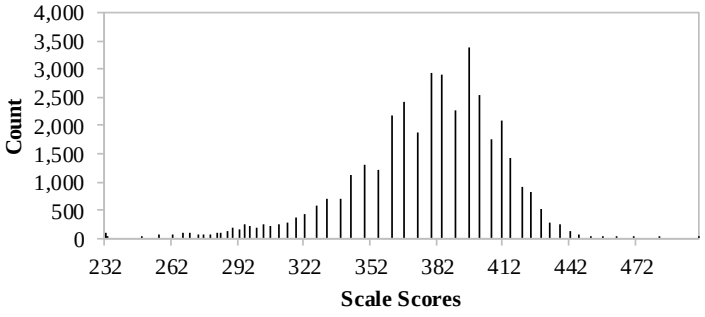
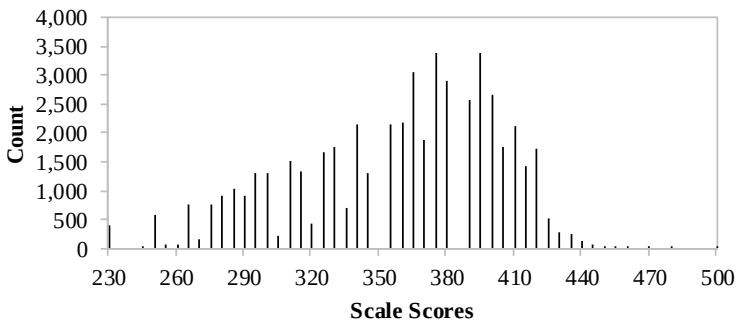


Table 2.4.3.6.3

Scale Score Descriptive Statistics: Writ 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	16,408	232	471	352.35	47.10
10	14,520	232	471	358.98	45.38
11	12,554	232	500	366.30	43.38
12	8,325	232	482	364.87	42.46
Total	51,807	232	500	359.60	45.36

Figure 2.4.3.6.3
Scale Scores: Writ 9-12 S501 Paper



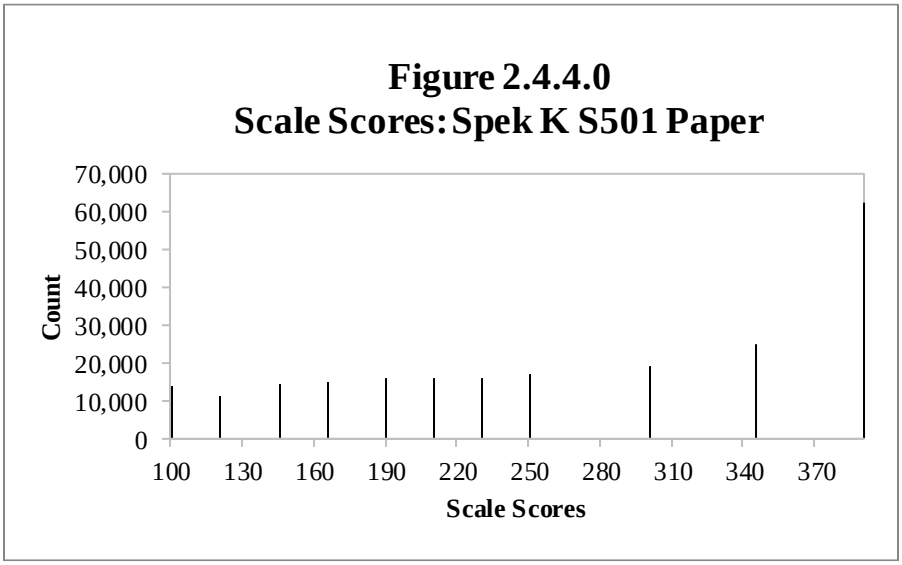
2.4.4 Speaking

2.4.4.0 Kindergarten

Table 2.4.4.0

Scale Score Descriptive Statistics: Spek K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	226,000	100	392	268.65	101.08
Total	226,000	100	392	268.65	101.08



2.4.4.1 Grade 1

Table 2.4.4.1.1

Scale Score Descriptive Statistics: Spek 1 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	19,503	106	401	244.39	63.83
Total	19,503	106	401	244.39	63.83

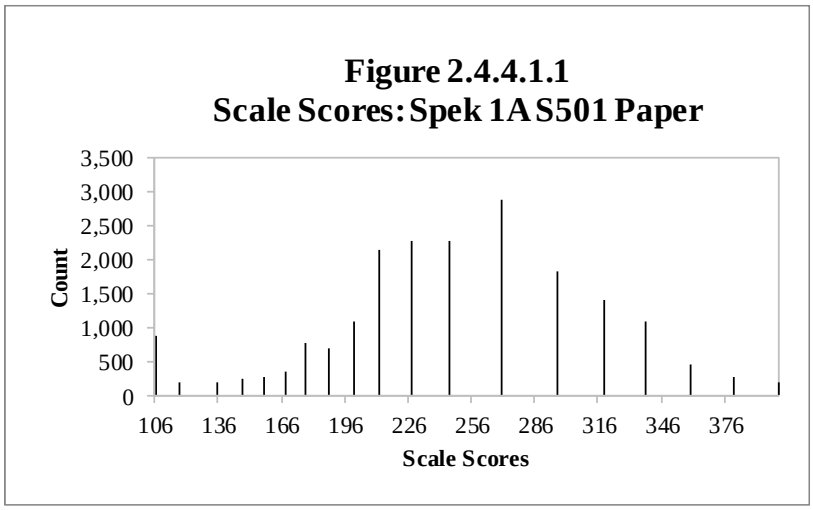


Table 2.4.4.1.2

Scale Score Descriptive Statistics: Spek 1 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	20,432	106	414	297.86	52.67
Total	20,432	106	414	297.86	52.67

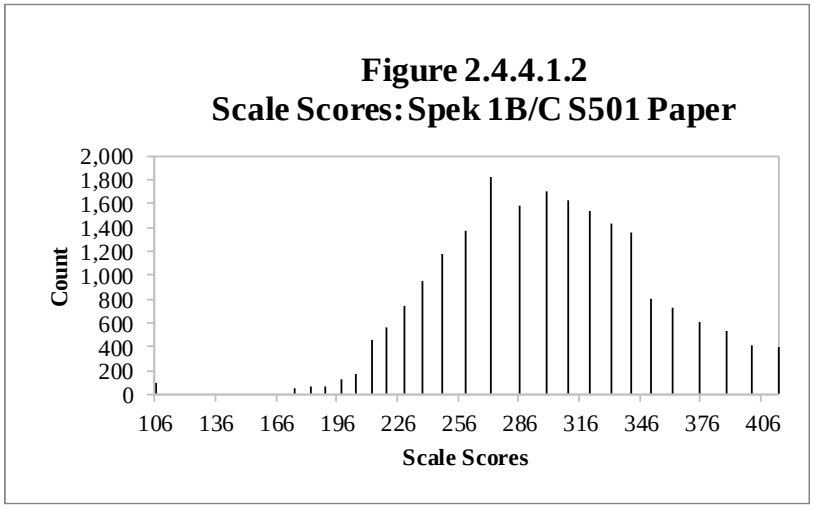
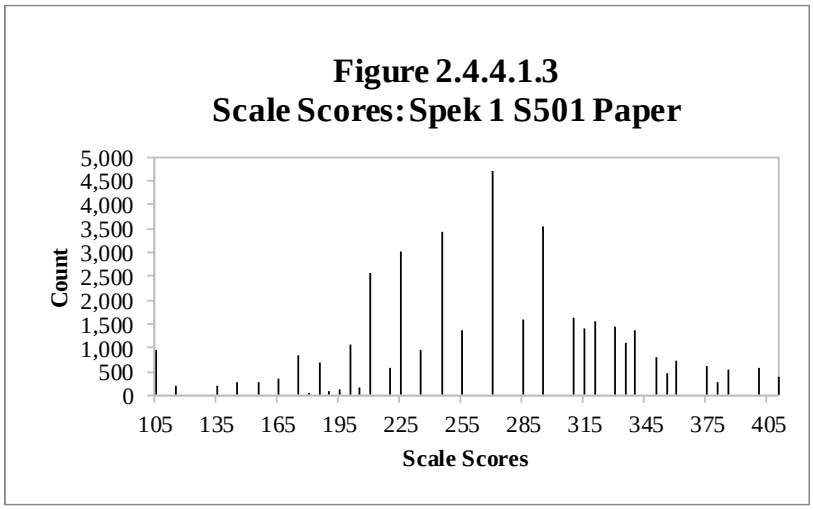


Table 2.4.4.1.3

Scale Score Descriptive Statistics: Spek 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	39,935	106	414	271.74	64.21
Total	39,935	106	414	271.74	64.21



2.4.4.2 Grade 2

Table 2.4.4.2.1

Scale Score Descriptive Statistics: Spek 2 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	8,894	118	380	232.76	71.16
Total	8,894	118	380	232.76	71.16

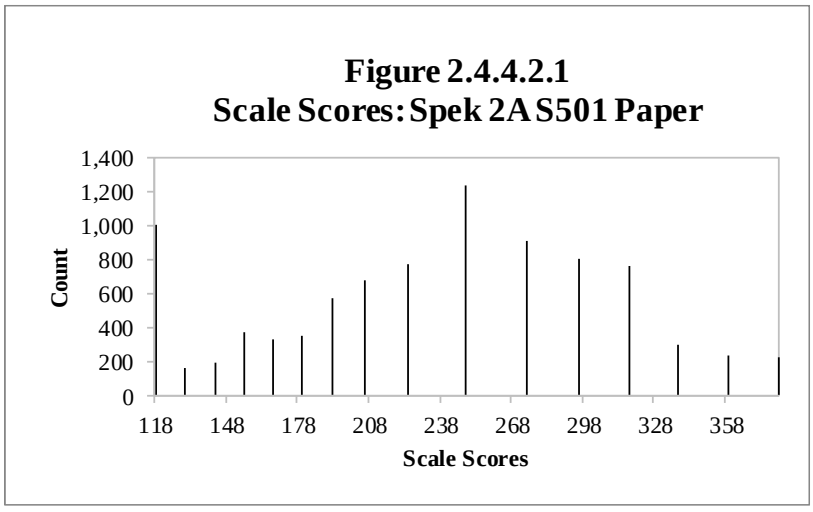


Table 2.4.4.2.2

Scale Score Descriptive Statistics: Spek 2 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	30,190	118	425	309.48	53.30
Total	30,190	118	425	309.48	53.30

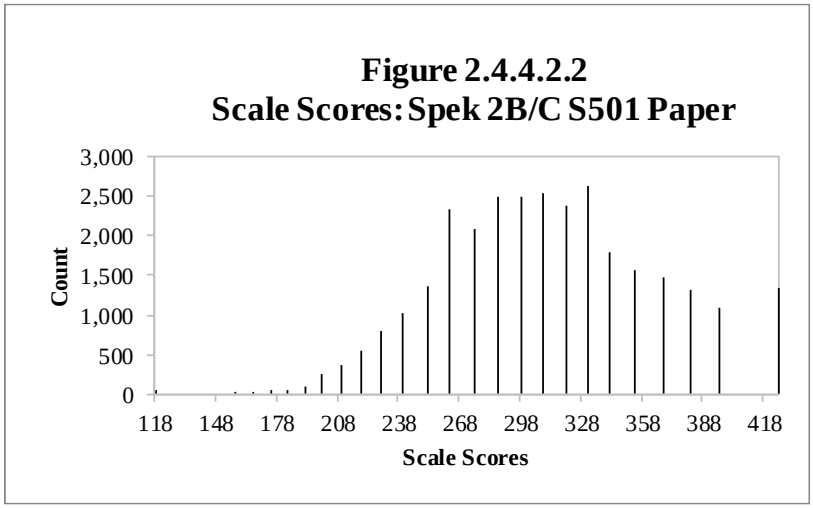
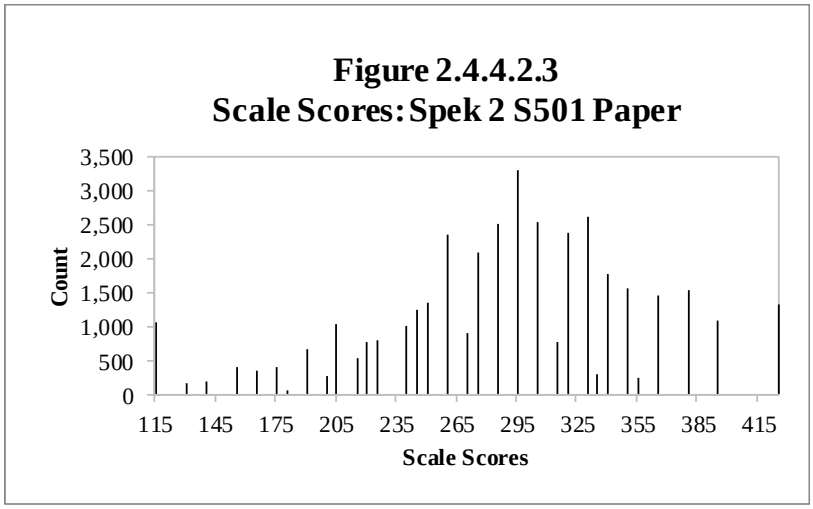


Table 2.4.4.2.3

Scale Score Descriptive Statistics: Spek 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	39,084	118	425	292.02	66.19
Total	39,084	118	425	292.02	66.19



2.4.4.3 Grade 3

Table 2.4.4.3.1

Scale Score Descriptive Statistics: Spek 3 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	7,284	118	380	235.54	71.26
Total	7,284	118	380	235.54	71.26

Figure 2.4.4.3.1
Scale Scores: Spek 3A S501 Paper

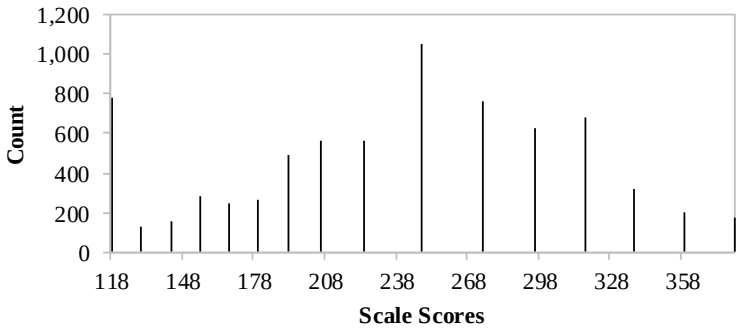


Table 2.4.4.3.2

Scale Score Descriptive Statistics: Spek 3 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,175	118	425	320.68	51.47
Total	29,175	118	425	320.68	51.47

Figure 2.4.4.3.2
Scale Scores: Spek 3B/C S501 Paper

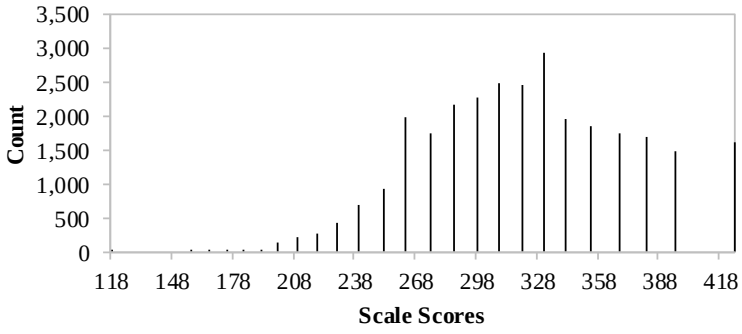
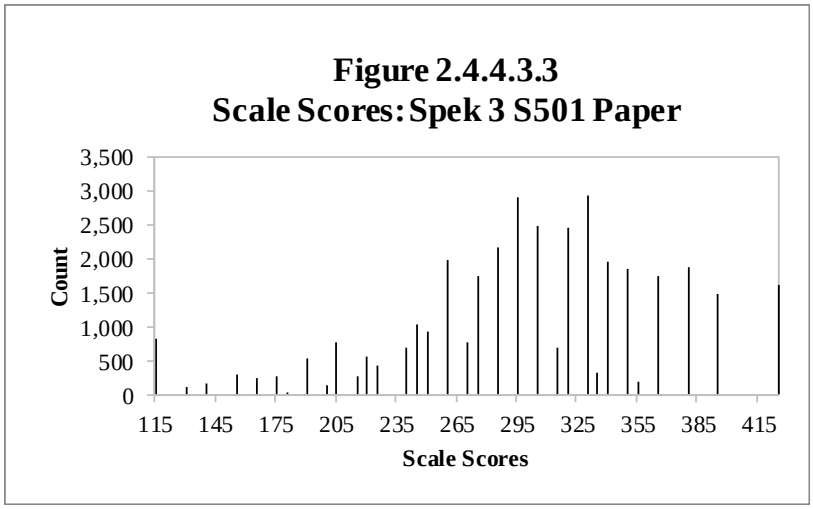


Table 2.4.4.3.3

Scale Score Descriptive Statistics: Spek 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	36,459	118	425	303.67	65.52
Total	36,459	118	425	303.67	65.52



2.4.4.4 Grades 4–5

Table 2.4.4.4.1

Scale Score Descriptive Statistics: Spek 4-5 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	5,382	130	408	236.92	67.04
5	4,891	130	408	238.44	67.16
Total	10,273	130	408	237.65	67.10

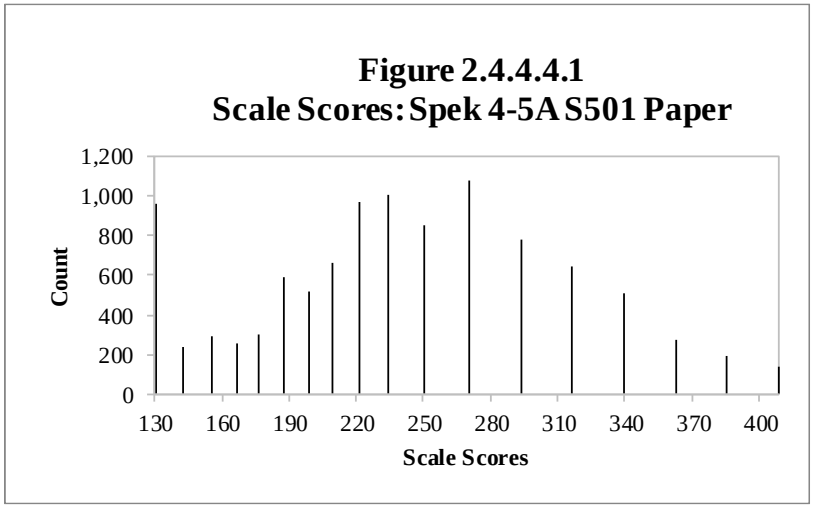


Table 2.4.4.4.2

Scale Score Descriptive Statistics: Spek 4-5 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	21,873	130	443	351.46	51.81
5	19,198	130	443	358.93	52.00
Total	41,071	130	443	354.95	52.03

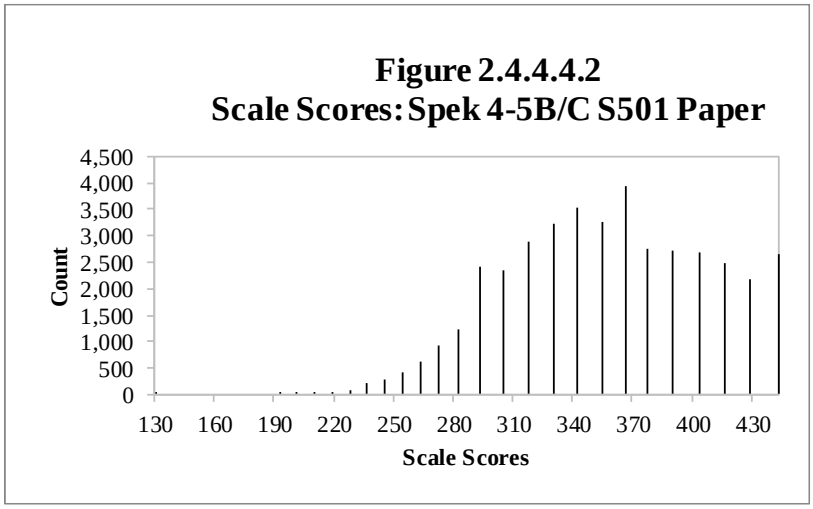
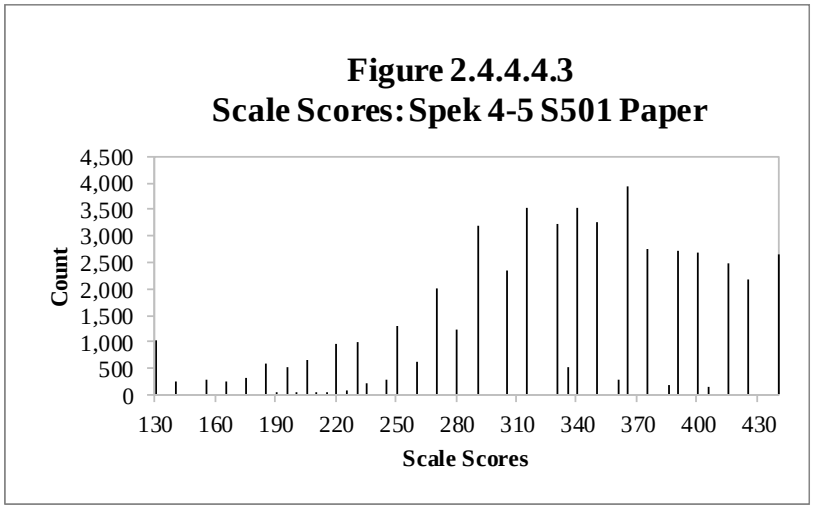


Table 2.4.4.4.3

Scale Score Descriptive Statistics: Spek 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	27,255	130	443	328.85	71.56
5	24,089	130	443	334.46	73.62
Total	51,344	130	443	331.48	72.59



2.4.4.5 Grades 6–8

Table 2.4.4.5.1

Scale Score Descriptive Statistics: Spek 6-8 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	4,534	148	438	262.73	71.05
7	4,335	148	438	261.19	70.81
8	4,429	148	438	263.76	69.97
Total	13,298	148	438	262.57	70.62

Figure 2.4.4.5.1
Scale Scores: Spek 6-8A S501 Paper

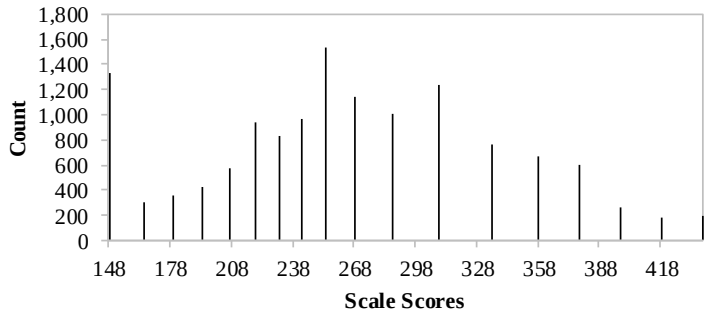


Table 2.4.4.5.2

Scale Score Descriptive Statistics: Spek 6-8 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	15,426	148	468	379.71	53.04
7	12,926	148	468	381.72	54.49
8	12,069	148	468	385.14	54.83
Total	40,421	148	468	381.97	54.09

Figure 2.4.4.5.2
Scale Scores: Spek 6-8B/C S501 Paper

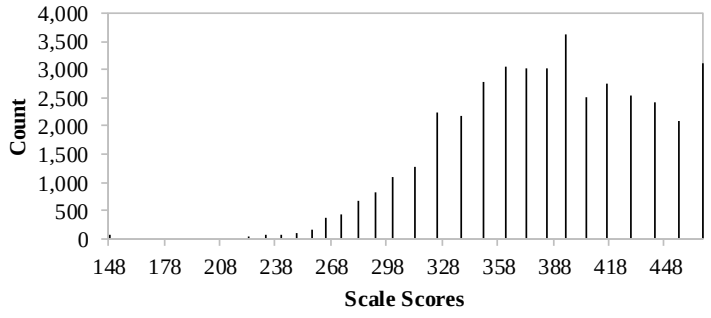
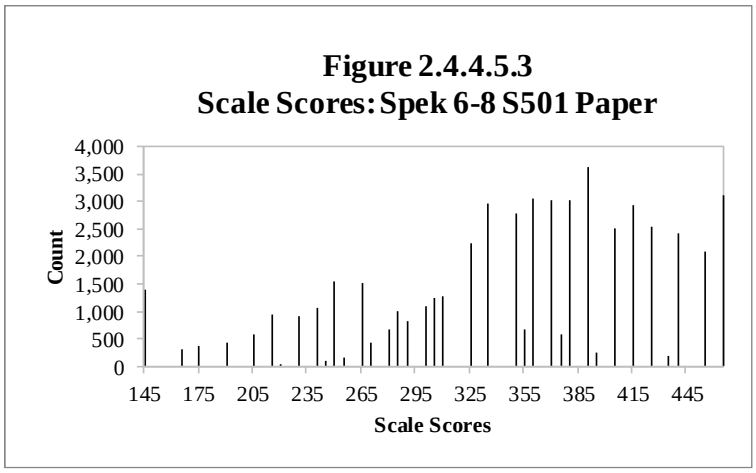


Table 2.4.4.5.3

Scale Score Descriptive Statistics: Spek 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	19,960	148	468	353.14	75.65
7	17,261	148	468	351.45	78.83
8	16,498	148	468	352.56	80.05
Total	53,719	148	468	352.42	78.05



2.4.4.6 Grades 9–12

Table 2.4.4.6.1

Scale Score Descriptive Statistics: Spek 9-12 A S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,198	172	443	261.47	66.32
10	3,729	172	443	273.53	68.38
11	2,830	172	443	282.73	67.35
12	1,547	172	443	289.93	66.36
Total	13,304	172	443	272.68	67.90

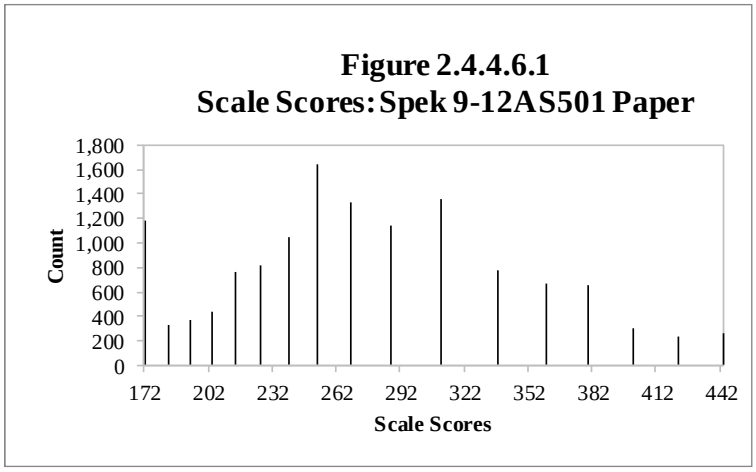


Table 2.4.4.6.2

Scale Score Descriptive Statistics: Spek 9-12 B/C S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	11,034	172	476	376.05	60.44
10	10,621	172	476	373.87	62.33
11	9,596	172	476	379.06	61.42
12	6,689	172	476	378.98	62.58
Total	37,940	172	476	376.72	61.63

Figure 2.4.4.6.2
Scale Scores: Spek 9-12B/C S501 Paper

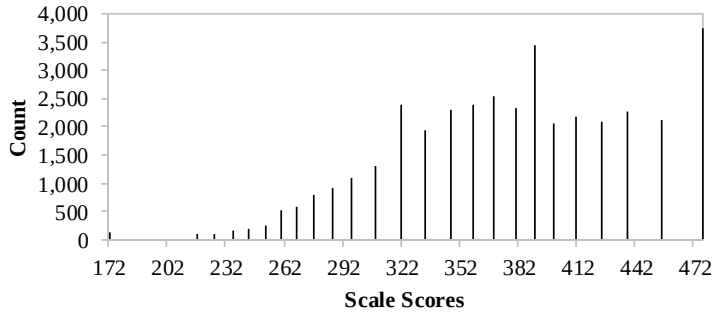
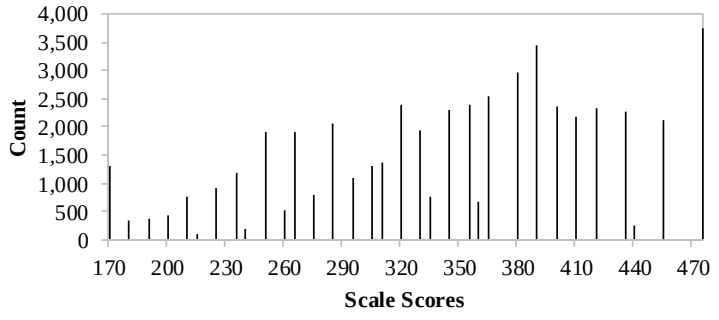


Table 2.4.4.6.3

Scale Score Descriptive Statistics: Spek 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	16,232	172	476	339.36	82.16
10	14,350	172	476	347.79	77.63
11	12,426	172	476	357.12	74.69
12	8,236	172	476	362.25	72.23
Total	51,244	172	476	349.71	78.04

Figure 2.4.4.6.3
Scale Scores: Spek 9-12 S501 Paper



2.5 Proficiency Level Distributions

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

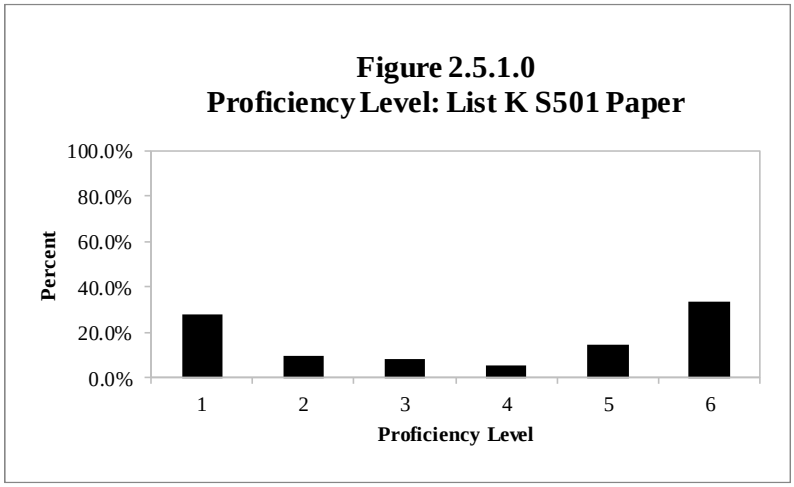
2.5.1 Listening

2.5.1.0 Kindergarten

Table 2.5.1.0

Proficiency Level Distribution: List K S501 Paper

Level	Grade K		Total	
	Count	Percent	Count	Percent
1	63,378	28.04%	63,378	28.04%
2	21,524	9.52%	21,524	9.52%
3	19,067	8.44%	19,067	8.44%
4	12,715	5.63%	12,715	5.63%
5	32,929	14.57%	32,929	14.57%
6	76,388	33.80%	76,388	33.80%
Total	226,001	100.00%	226,001	100.00%



2.5.1.1 Grade 1

Table 2.5.1.1.1

Proficiency Level Distribution: List 1 A S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,325	8.05%	1,325	8.05%
2	2,091	12.71%	2,091	12.71%
3	2,659	16.16%	2,659	16.16%
4	1,923	11.69%	1,923	11.69%
5	4,860	29.54%	4,860	29.54%
6	3,597	21.86%	3,597	21.86%
Total	16,455	100.00%	16,455	100.00%

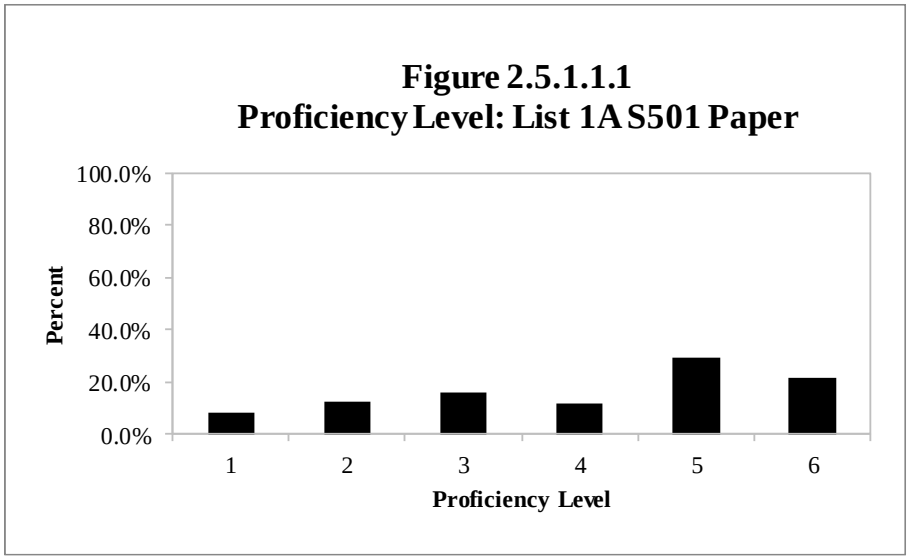


Table 2.5.1.1.2

Proficiency Level Distribution: List 1 B/C S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	128	0.71%	128	0.71%
2	471	2.62%	471	2.62%
3	3,487	19.37%	3,487	19.37%
4	1,693	9.41%	1,693	9.41%
5	4,170	23.17%	4,170	23.17%
6	8,051	44.73%	8,051	44.73%
Total	18,000	100.00%	18,000	100.00%

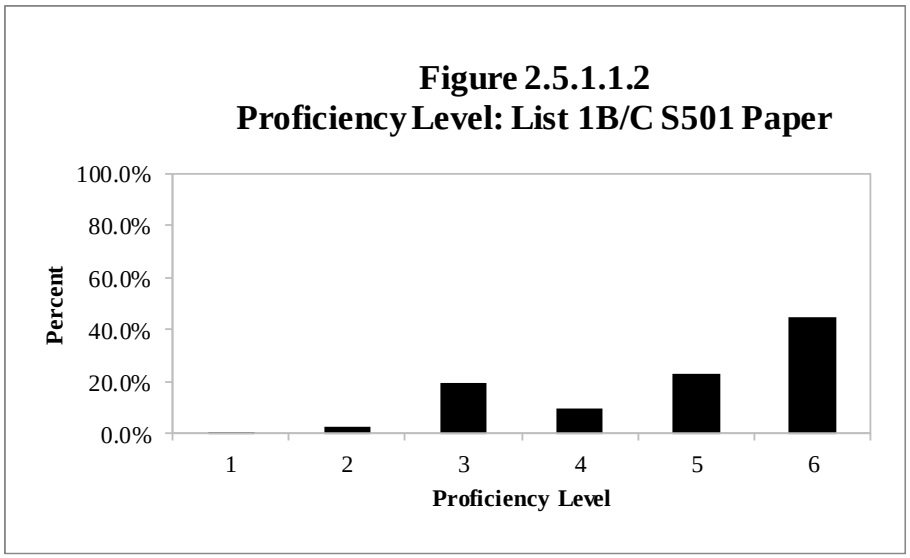
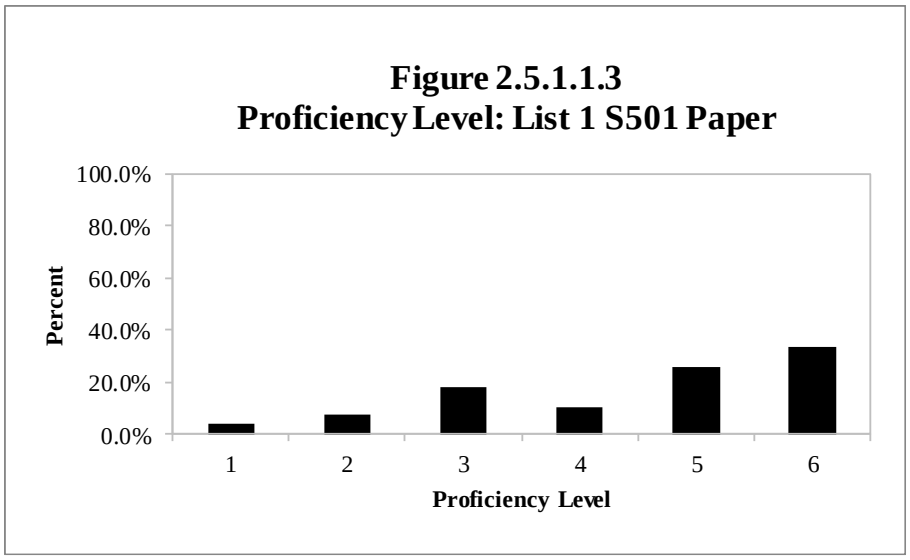


Table 2.5.1.1.3

Proficiency Level Distribution: List 1 S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,453	4.22%	1,453	4.22%
2	2,562	7.44%	2,562	7.44%
3	6,146	17.84%	6,146	17.84%
4	3,616	10.49%	3,616	10.49%
5	9,030	26.21%	9,030	26.21%
6	11,648	33.81%	11,648	33.81%
Total	34,455	100.00%	34,455	100.00%



2.5.1.2 Grade 2

Table 2.5.1.2.1

Proficiency Level Distribution: List 2 A S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,198	14.81%	1,198	14.81%
2	1,939	23.97%	1,939	23.97%
3	1,547	19.13%	1,547	19.13%
4	1,083	13.39%	1,083	13.39%
5	2,321	28.70%	2,321	28.70%
6	0	0.00%	0	0.00%
Total	8,088	100.00%	8,088	100.00%

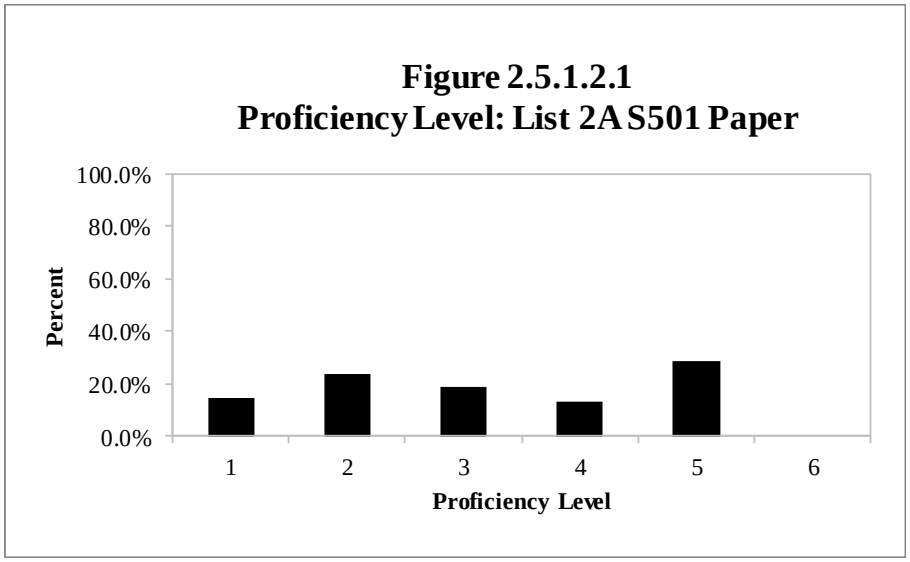


Table 2.5.1.2.2

Proficiency Level Distribution: List 2 B/C S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	48	0.17%	48	0.17%
2	1,219	4.27%	1,219	4.27%
3	3,855	13.50%	3,855	13.50%
4	5,879	20.58%	5,879	20.58%
5	4,158	14.56%	4,158	14.56%
6	13,407	46.93%	13,407	46.93%
Total	28,566	100.00%	28,566	100.00%

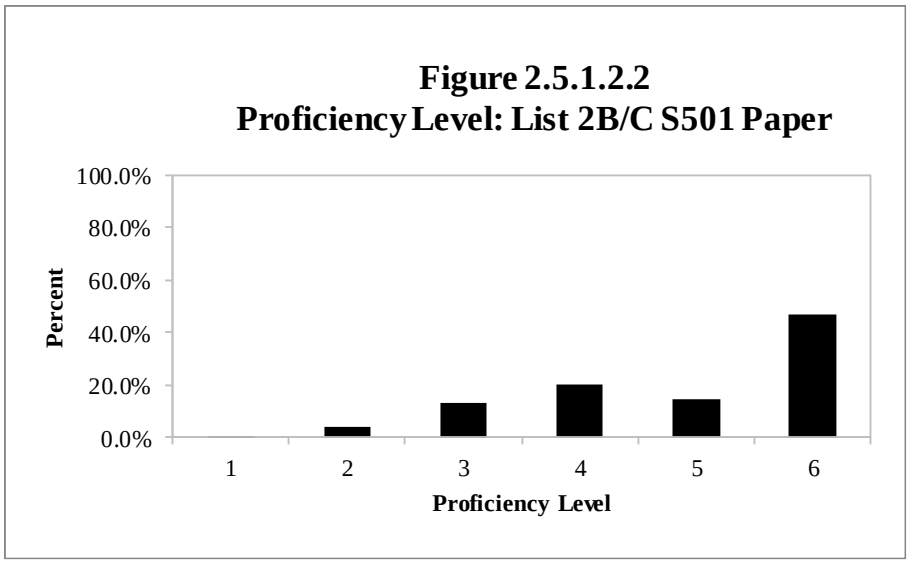
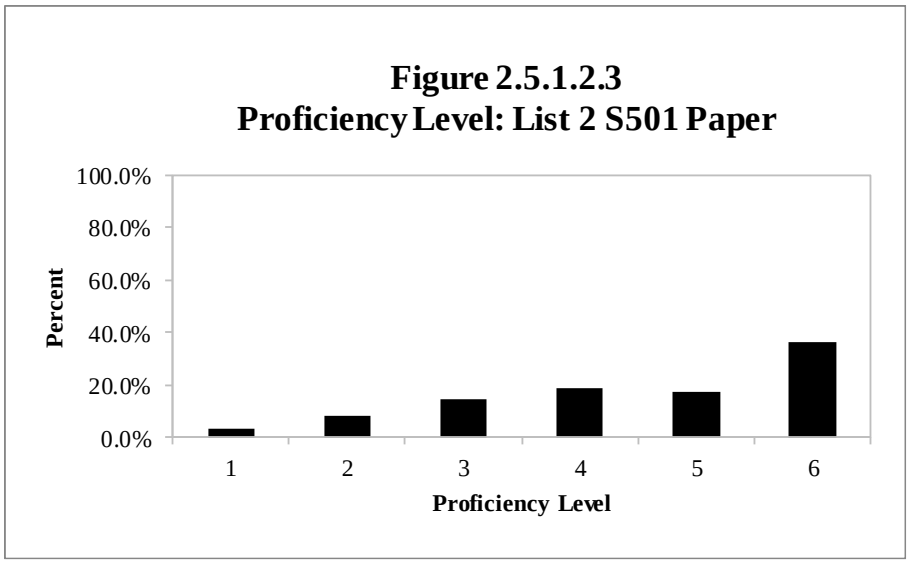


Table 2.5.1.2.3

Proficiency Level Distribution: List 2 S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,246	3.40%	1,246	3.40%
2	3,158	8.62%	3,158	8.62%
3	5,402	14.74%	5,402	14.74%
4	6,962	18.99%	6,962	18.99%
5	6,479	17.68%	6,479	17.68%
6	13,407	36.58%	13,407	36.58%
Total	36,654	100.00%	36,654	100.00%



2.5.1.3 Grade 3

Table 2.5.1.3.1

Proficiency Level Distribution: List 3 A S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	245	3.64%	245	3.64%
2	1,857	27.62%	1,857	27.62%
3	1,896	28.20%	1,896	28.20%
4	1,124	16.72%	1,124	16.72%
5	931	13.85%	931	13.85%
6	670	9.97%	670	9.97%
Total	6,723	100.00%	6,723	100.00%

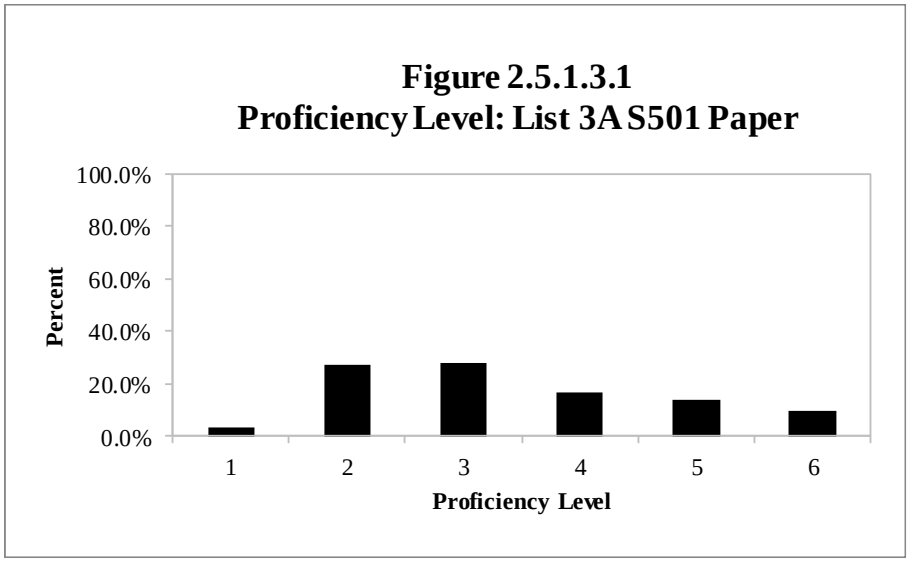


Table 2.5.1.3.2

Proficiency Level Distribution: List 3 B/C S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	12	0.04%	12	0.04%
2	432	1.56%	432	1.56%
3	2,985	10.76%	2,985	10.76%
4	4,973	17.93%	4,973	17.93%
5	6,879	24.81%	6,879	24.81%
6	12,449	44.89%	12,449	44.89%
Total	27,730	100.00%	27,730	100.00%

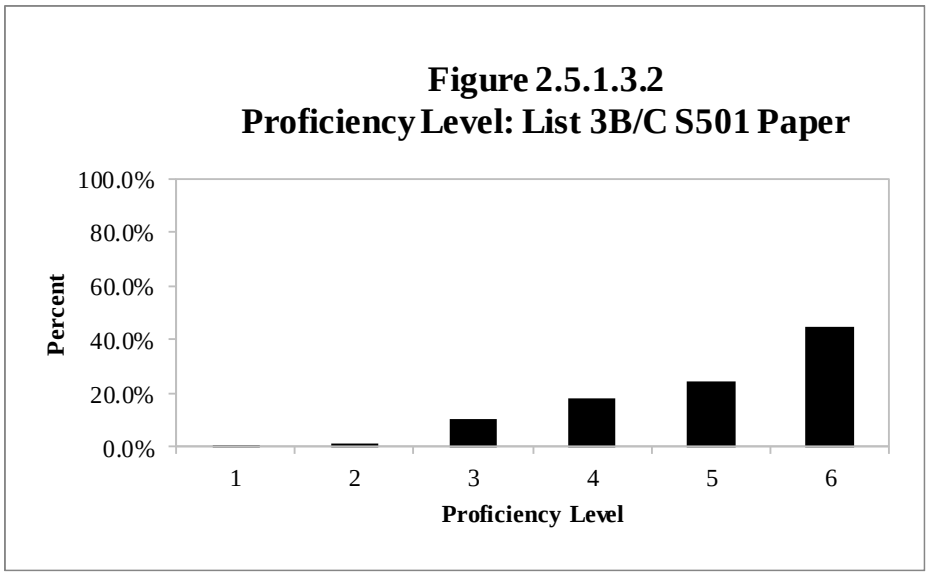
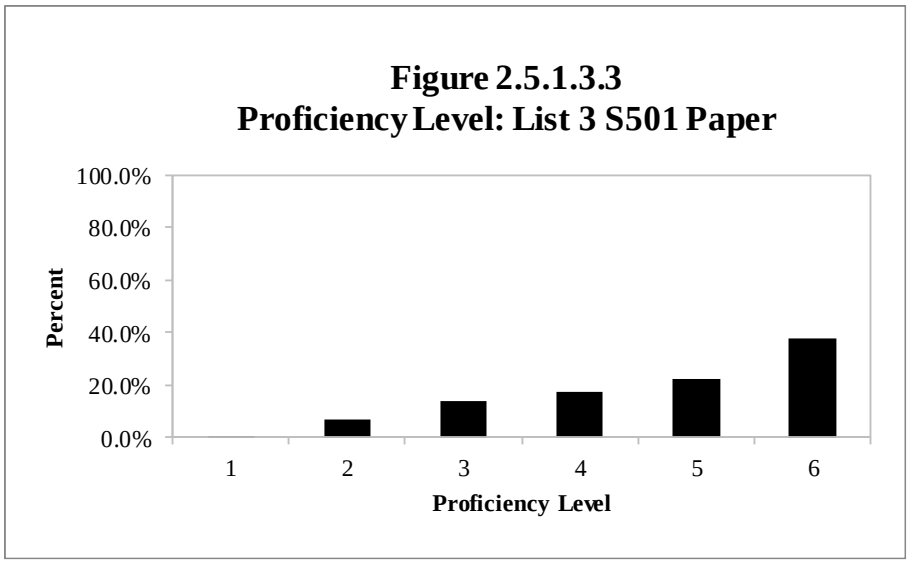


Table 2.5.1.3.3

Proficiency Level Distribution: List 3 S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	257	0.75%	257	0.75%
2	2,289	6.64%	2,289	6.64%
3	4,881	14.17%	4,881	14.17%
4	6,097	17.70%	6,097	17.70%
5	7,810	22.67%	7,810	22.67%
6	13,119	38.08%	13,119	38.08%
Total	34,453	100.00%	34,453	100.00%



2.5.1.4 Grades 4–5

Table 2.5.1.4.1

Proficiency Level Distribution: List 4-5 A S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	370	7.35%	449	9.72%	819	8.48%
2	1,612	32.01%	1,547	33.50%	3,159	32.72%
3	1,312	26.05%	1,133	24.53%	2,445	25.33%
4	781	15.51%	709	15.35%	1,490	15.43%
5	591	11.74%	356	7.71%	947	9.81%
6	370	7.35%	424	9.18%	794	8.22%
Total	5,036	100.00%	4,618	100.00%	9,654	100.00%

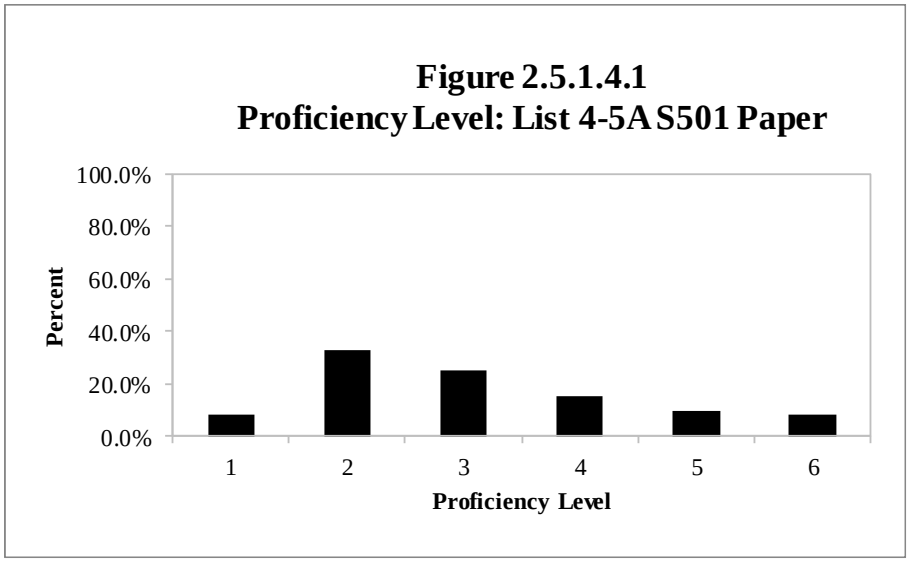


Table 2.5.1.4.2

Proficiency Level Distribution: List 4-5 B/C S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	16	0.08%	33	0.18%	49	0.12%
2	239	1.12%	257	1.37%	496	1.24%
3	1,464	6.87%	1,252	6.68%	2,716	6.78%
4	2,655	12.46%	2,265	12.08%	4,920	12.28%
5	7,600	35.68%	6,803	36.29%	14,403	35.96%
6	9,329	43.79%	8,137	43.40%	17,466	43.61%
Total	21,303	100.00%	18,747	100.00%	40,050	100.00%

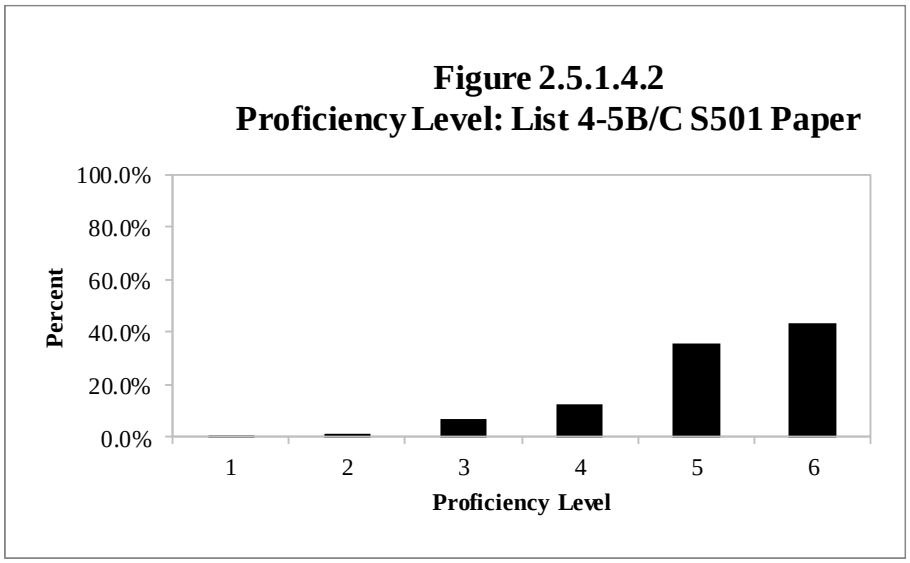
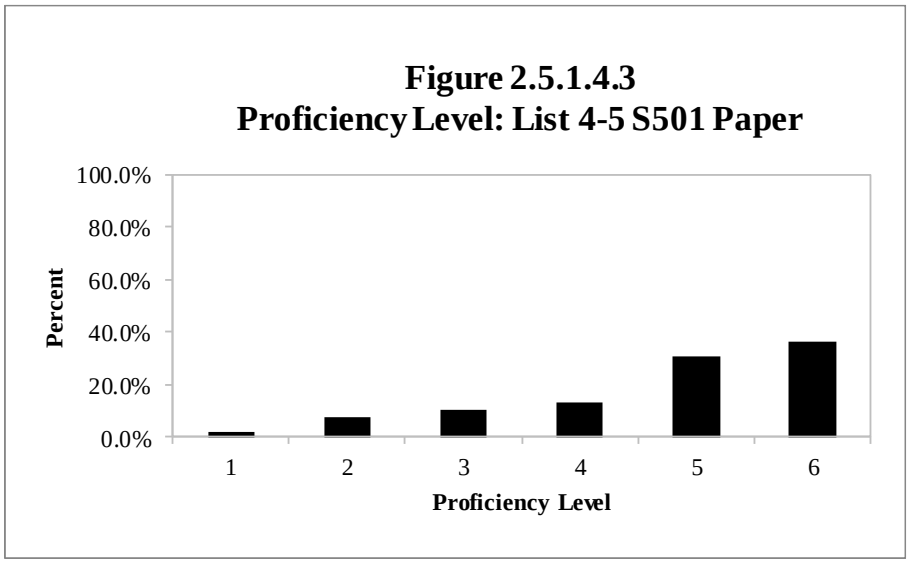


Table 2.5.1.4.3

Proficiency Level Distribution: List 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	386	1.47%	482	2.06%	868	1.75%
2	1,851	7.03%	1,804	7.72%	3,655	7.35%
3	2,776	10.54%	2,385	10.21%	5,161	10.38%
4	3,436	13.05%	2,974	12.73%	6,410	12.90%
5	8,191	31.10%	7,159	30.64%	15,350	30.88%
6	9,699	36.82%	8,561	36.64%	18,260	36.74%
Total	26,339	100.00%	23,365	100.00%	49,704	100.00%



2.5.1.5 Grades 6–8

Table 2.5.1.5.1

Proficiency Level Distribution: List 6-8 A S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	986	23.25%	1,383	33.90%	1,291	31.43%	3,660	29.45%
2	1,750	41.27%	1,288	31.57%	1,694	41.25%	4,732	38.08%
3	638	15.05%	829	20.32%	559	13.61%	2,026	16.30%
4	467	11.01%	197	4.83%	349	8.50%	1,013	8.15%
5	324	7.64%	300	7.35%	214	5.21%	838	6.74%
6	75	1.77%	83	2.03%	0	0.00%	158	1.27%
Total	4,240	100.00%	4,080	100.00%	4,107	100.00%	12,427	100.00%

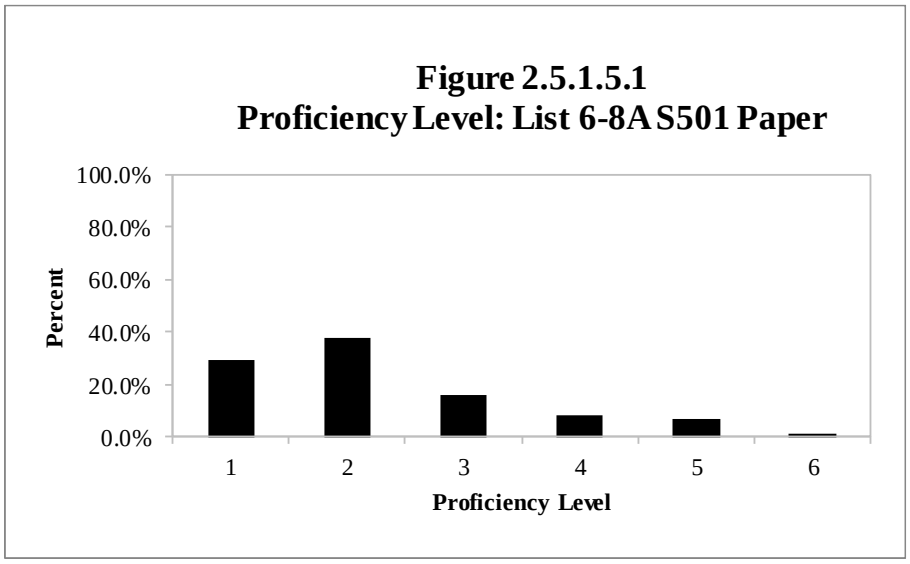


Table 2.5.1.5.2

Proficiency Level Distribution: List 6-8 B/C S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	9	0.06%	25	0.20%	13	0.11%	47	0.12%
2	237	1.58%	297	2.36%	232	1.98%	766	1.95%
3	1,209	8.05%	1,156	9.17%	1,348	11.53%	3,713	9.44%
4	3,661	24.37%	3,104	24.62%	1,865	15.95%	8,630	21.94%
5	3,905	26.00%	3,400	26.96%	3,042	26.01%	10,347	26.31%
6	6,001	39.95%	4,628	36.70%	5,194	44.42%	15,823	40.24%
Total	15,022	100.00%	12,610	100.00%	11,694	100.00%	39,326	100.00%

Figure 2.5.1.5.2
Proficiency Level: List 6-8B/C S501 Paper

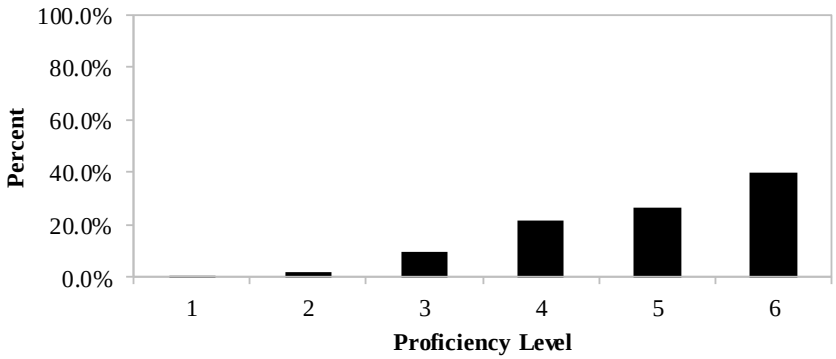
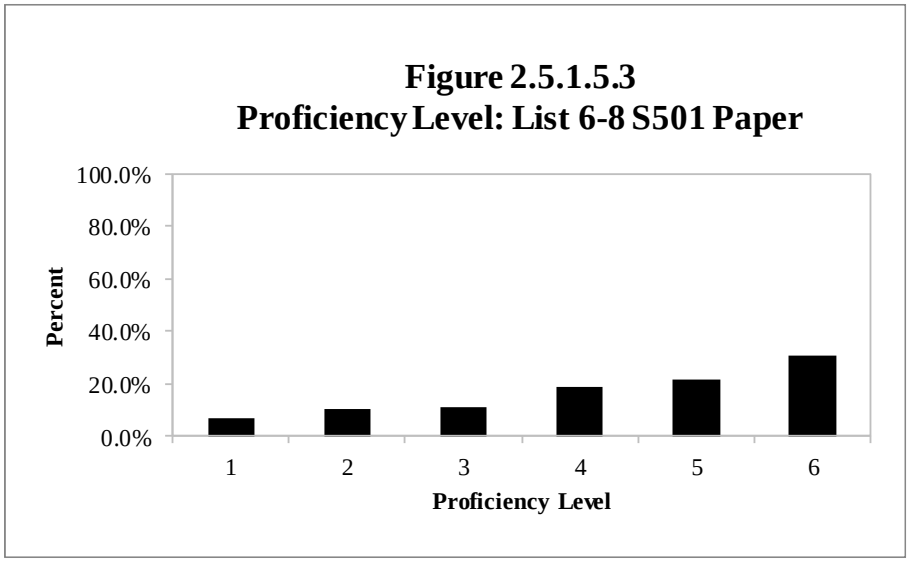


Table 2.5.1.5.3

Proficiency Level Distribution: List 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	995	5.17%	1,408	8.44%	1,304	8.25%	3,707	7.16%
2	1,987	10.32%	1,585	9.50%	1,926	12.19%	5,498	10.62%
3	1,847	9.59%	1,985	11.89%	1,907	12.07%	5,739	11.09%
4	4,128	21.43%	3,301	19.78%	2,214	14.01%	9,643	18.63%
5	4,229	21.96%	3,700	22.17%	3,256	20.61%	11,185	21.61%
6	6,076	31.54%	4,711	28.23%	5,194	32.87%	15,981	30.88%
Total	19,262	100.00%	16,690	100.00%	15,801	100.00%	51,753	100.00%



2.5.1.6 Grades 9–12

Table 2.5.1.6.1

Proficiency Level Distribution: List 9-12 A S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,860	38.72%	1,674	47.73%	1,458	54.89%	960	65.66%	5,952	47.89%
2	2,205	45.90%	1,146	32.68%	604	22.74%	295	20.18%	4,250	34.19%
3	522	10.87%	461	13.15%	392	14.76%	162	11.08%	1,537	12.37%
4	119	2.48%	193	5.50%	181	6.81%	36	2.46%	529	4.26%
5	98	2.04%	33	0.94%	21	0.79%	9	0.62%	161	1.30%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	4,804	100.00%	3,507	100.00%	2,656	100.00%	1,462	100.00%	12,429	100.00%

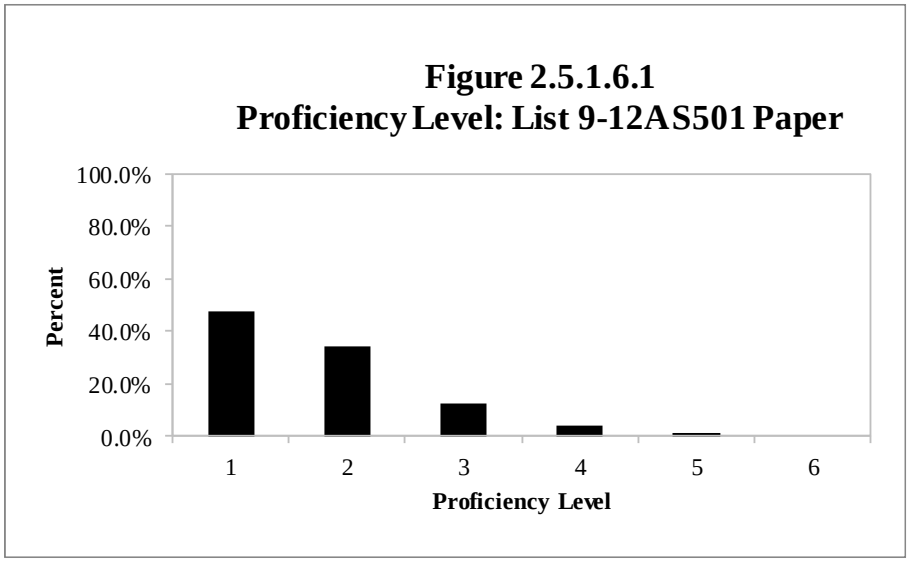


Table 2.5.1.6.2

Proficiency Level Distribution: List 9-12 B/C S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	38	0.36%	120	1.17%	220	2.38%	375	5.80%	753	2.06%
2	719	6.76%	737	7.18%	963	10.41%	600	9.28%	3,019	8.25%
3	2,150	20.22%	2,894	28.21%	2,034	21.99%	2,378	36.77%	9,456	25.83%
4	3,541	33.31%	2,166	21.12%	2,853	30.85%	1,256	19.42%	9,816	26.82%
5	2,146	20.19%	3,036	29.60%	1,820	19.68%	1,079	16.68%	8,081	22.08%
6	2,037	19.16%	1,305	12.72%	1,358	14.68%	779	12.05%	5,479	14.97%
Total	10,631	100.00%	10,258	100.00%	9,248	100.00%	6,467	100.00%	36,604	100.00%

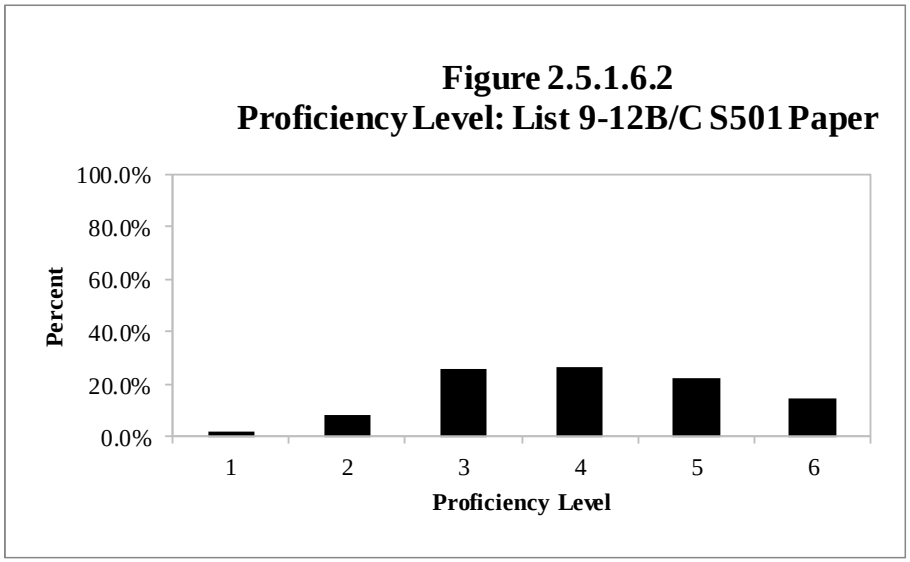
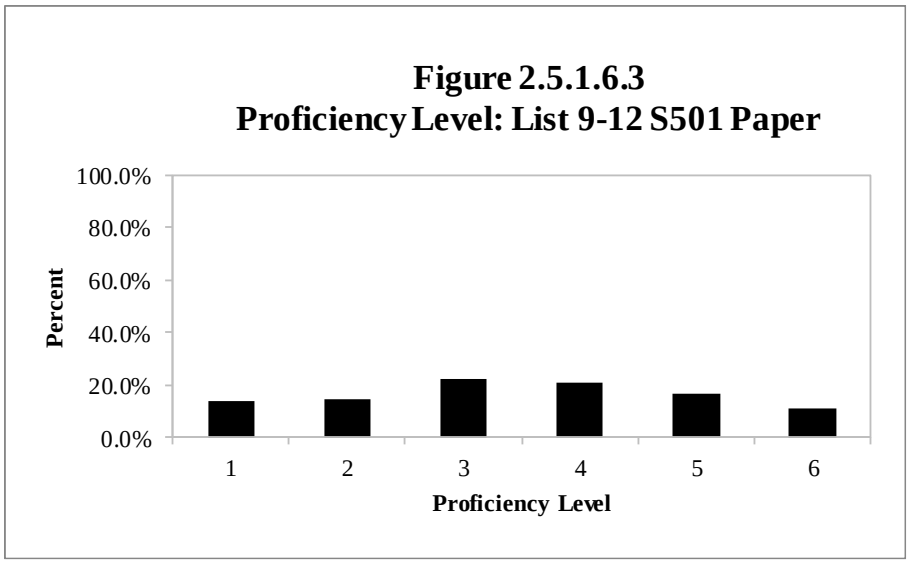


Table 2.5.1.6.3

Proficiency Level Distribution: List 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,898	12.30%	1,794	13.03%	1,678	14.10%	1,335	16.84%	6,705	13.67%
2	2,924	18.94%	1,883	13.68%	1,567	13.16%	895	11.29%	7,269	14.82%
3	2,672	17.31%	3,355	24.37%	2,426	20.38%	2,540	32.03%	10,993	22.42%
4	3,660	23.71%	2,359	17.14%	3,034	25.49%	1,292	16.29%	10,345	21.10%
5	2,244	14.54%	3,069	22.30%	1,841	15.47%	1,088	13.72%	8,242	16.81%
6	2,037	13.20%	1,305	9.48%	1,358	11.41%	779	9.82%	5,479	11.17%
Total	15,435	100.00%	13,765	100.00%	11,904	100.00%	7,929	100.00%	49,033	100.00%



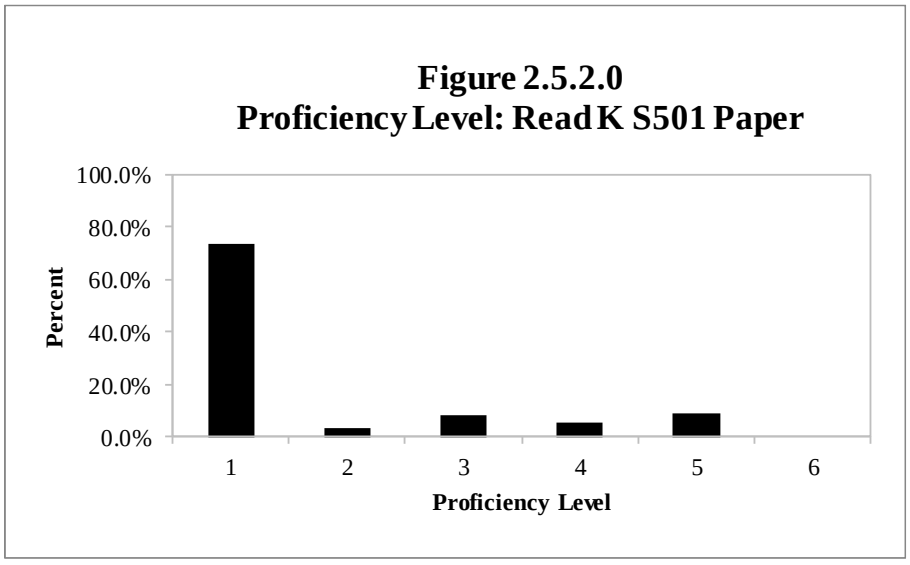
2.5.2 Reading

2.5.2.0 Kindergarten

Table 2.5.2.0

Proficiency Level Distribution: Read K S501 Paper

Level	Grade K		Total	
	Count	**Percent**	**Count**	**Percent**
1	166,059	73.48%	166,059	73.48%
2	7,055	3.12%	7,055	3.12%
3	19,489	8.62%	19,489	8.62%
4	12,494	5.53%	12,494	5.53%
5	20,897	9.25%	20,897	9.25%
6	0	0.00%	0	0.00%
Total	225,994	100.00%	225,994	100.00%



2.5.2.1 Grade 1

Table 2.5.2.1.1

Proficiency Level Distribution: Read 1 A S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	6,204	39.94%	6,204	39.94%
2	5,278	33.98%	5,278	33.98%
3	2,090	13.46%	2,090	13.46%
4	532	3.42%	532	3.42%
5	884	5.69%	884	5.69%
6	545	3.51%	545	3.51%
Total	15,533	100.00%	15,533	100.00%

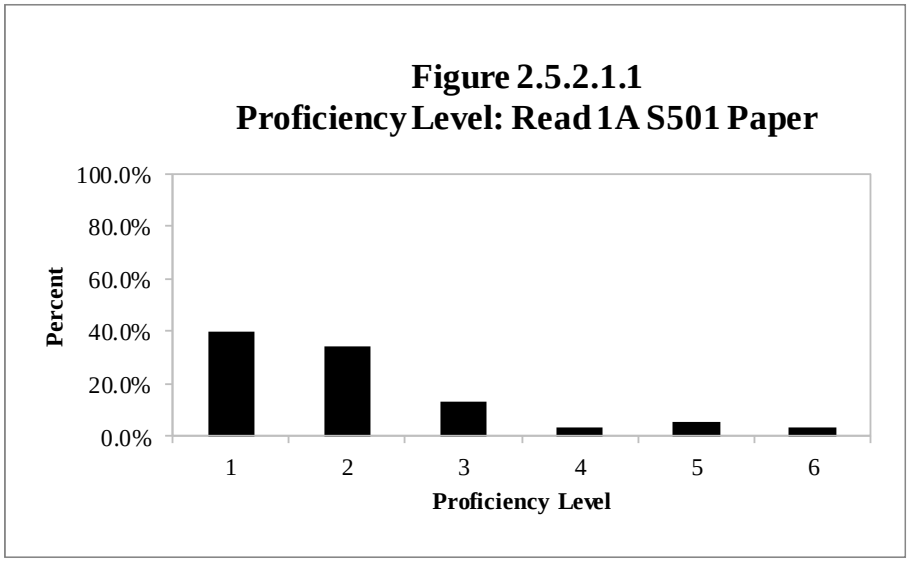


Table 2.5.2.1.2

Proficiency Level Distribution: Read 1 B/C S501 Paper

Level	Grade 1		Total	
	Count	Percent	Count	Percent
1	97	0.63%	97	0.63%
2	2,431	15.88%	2,431	15.88%
3	5,794	37.86%	5,794	37.86%
4	2,811	18.37%	2,811	18.37%
5	2,184	14.27%	2,184	14.27%
6	1,987	12.98%	1,987	12.98%
Total	15,304	100.00%	15,304	100.00%

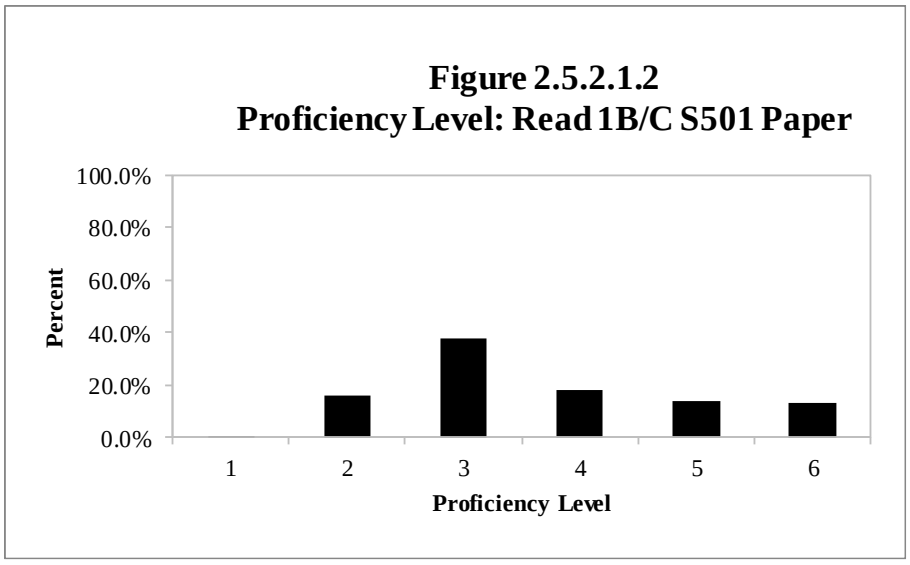
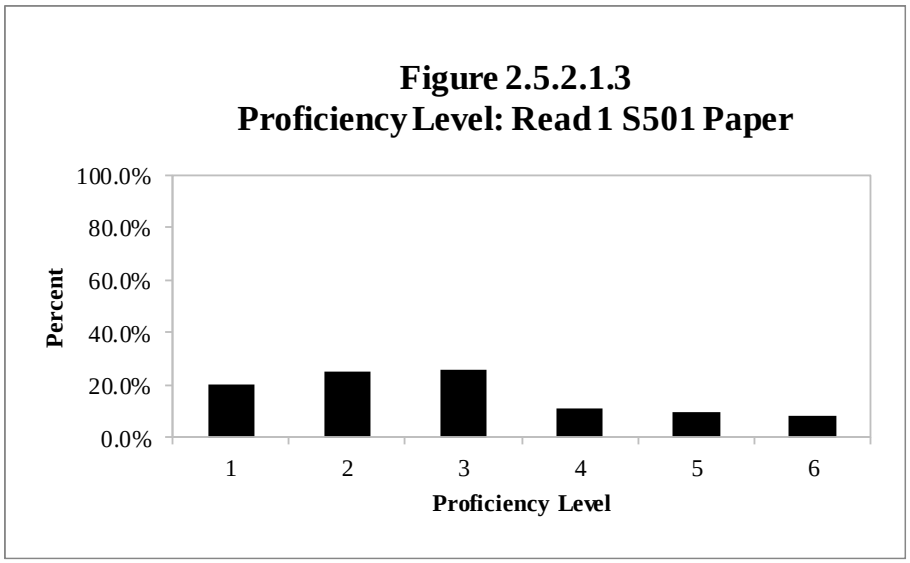


Table 2.5.2.1.3

Proficiency Level Distribution: Read 1 S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	6,301	20.43%	6,301	20.43%
2	7,709	25.00%	7,709	25.00%
3	7,884	25.57%	7,884	25.57%
4	3,343	10.84%	3,343	10.84%
5	3,068	9.95%	3,068	9.95%
6	2,532	8.21%	2,532	8.21%
Total	30,837	100.00%	30,837	100.00%



2.5.2.2 Grade 2

Table 2.5.2.2.1

Proficiency Level Distribution: Read 2 A S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	4,237	55.06%	4,237	55.06%
2	1,617	21.01%	1,617	21.01%
3	751	9.76%	751	9.76%
4	677	8.80%	677	8.80%
5	413	5.37%	413	5.37%
6	0	0.00%	0	0.00%
Total	7,695	100.00%	7,695	100.00%

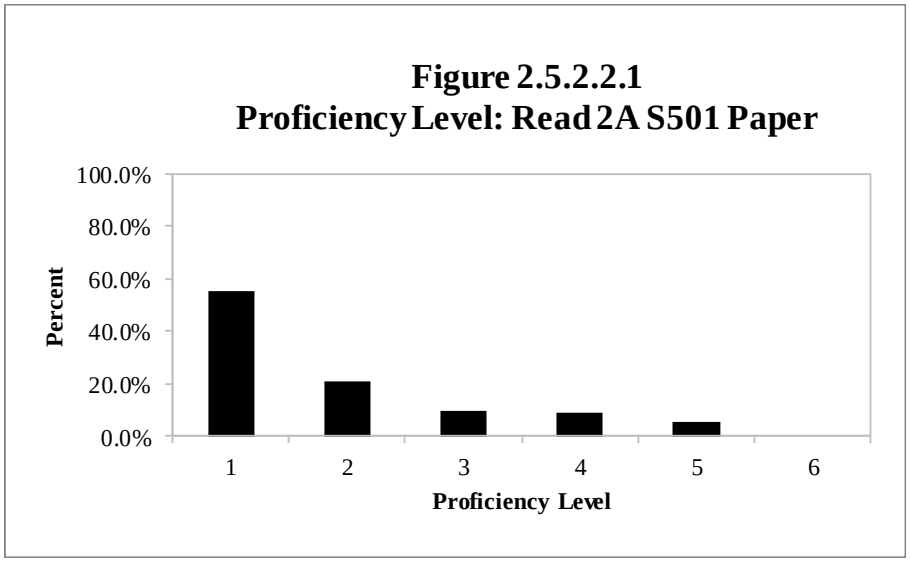


Table 2.5.2.2.2

Proficiency Level Distribution: Read 2 B/C S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	872	3.44%	872	3.44%
2	6,609	26.10%	6,609	26.10%
3	5,202	20.55%	5,202	20.55%
4	2,796	11.04%	2,796	11.04%
5	4,593	18.14%	4,593	18.14%
6	5,245	20.72%	5,245	20.72%
Total	25,317	100.00%	25,317	100.00%

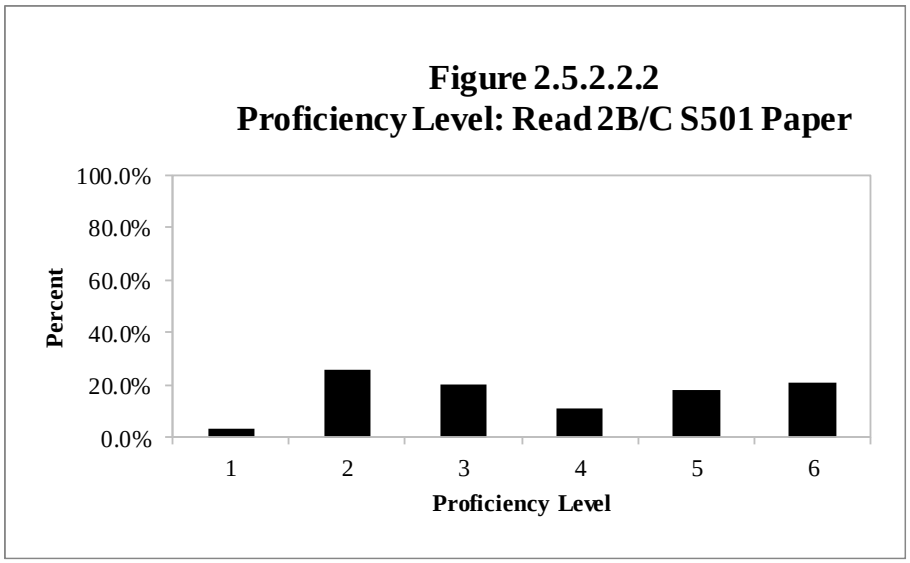
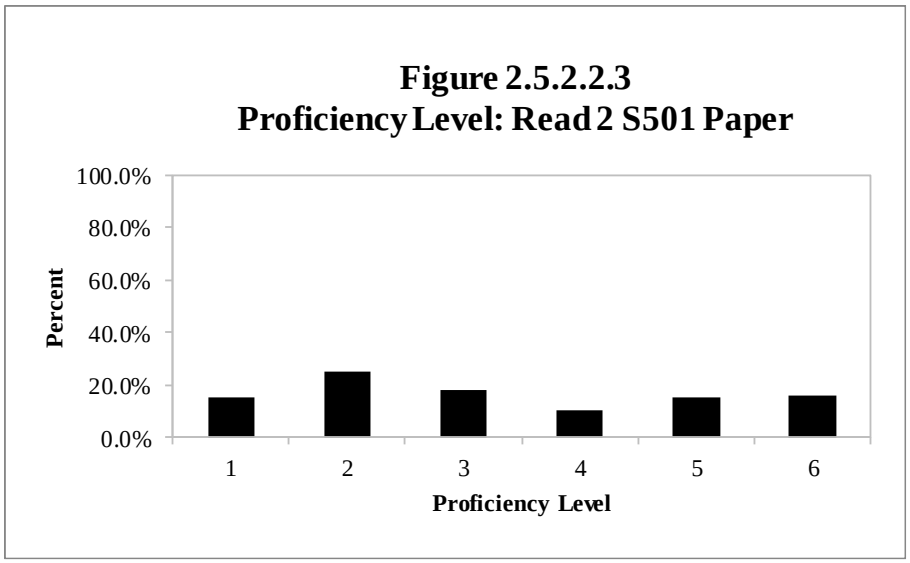


Table 2.5.2.2.3

Proficiency Level Distribution: Read 2 S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	5,109	15.48%	5,109	15.48%
2	8,226	24.92%	8,226	24.92%
3	5,953	18.03%	5,953	18.03%
4	3,473	10.52%	3,473	10.52%
5	5,006	15.16%	5,006	15.16%
6	5,245	15.89%	5,245	15.89%
Total	33,012	100.00%	33,012	100.00%



2.5.2.3 Grade 3

Table 2.5.2.3.1

Proficiency Level Distribution: Read 3 A S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	3,342	53.04%	3,342	53.04%
2	1,615	25.63%	1,615	25.63%
3	610	9.68%	610	9.68%
4	187	2.97%	187	2.97%
5	339	5.38%	339	5.38%
6	208	3.30%	208	3.30%
Total	6,301	100.00%	6,301	100.00%

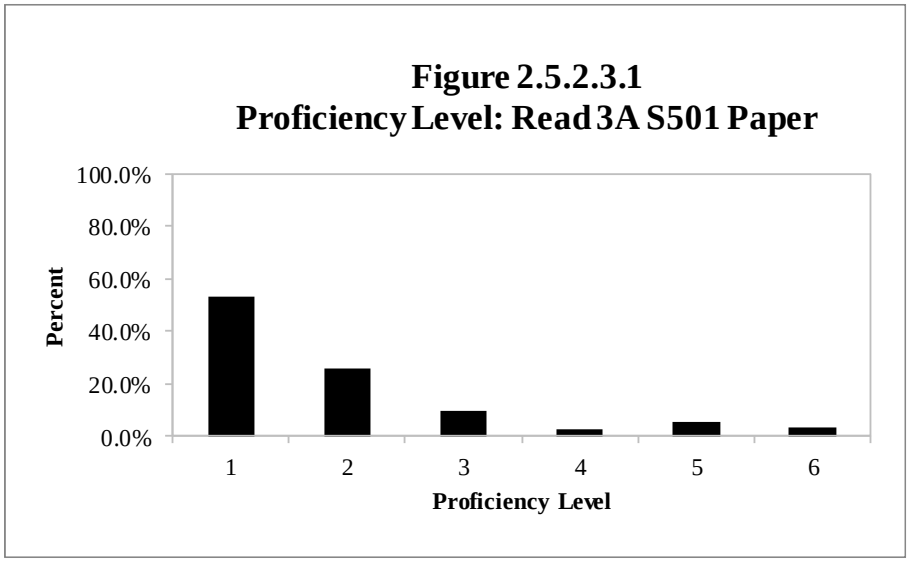


Table 2.5.2.3.2

Proficiency Level Distribution: Read 3 B/C S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	102	0.41%	102	0.41%
2	2,077	8.36%	2,077	8.36%
3	7,194	28.97%	7,194	28.97%
4	5,448	21.94%	5,448	21.94%
5	7,529	30.32%	7,529	30.32%
6	2,484	10.00%	2,484	10.00%
Total	24,834	100.00%	24,834	100.00%

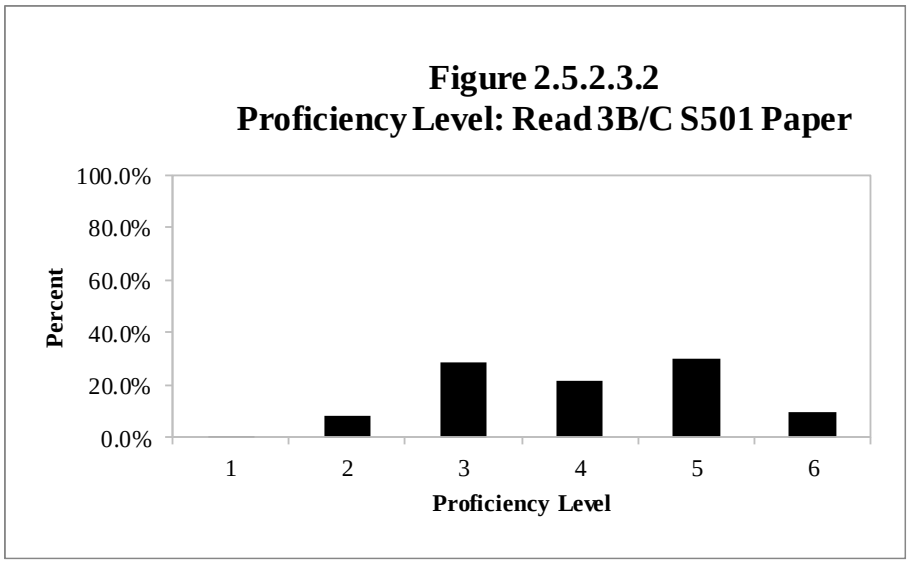
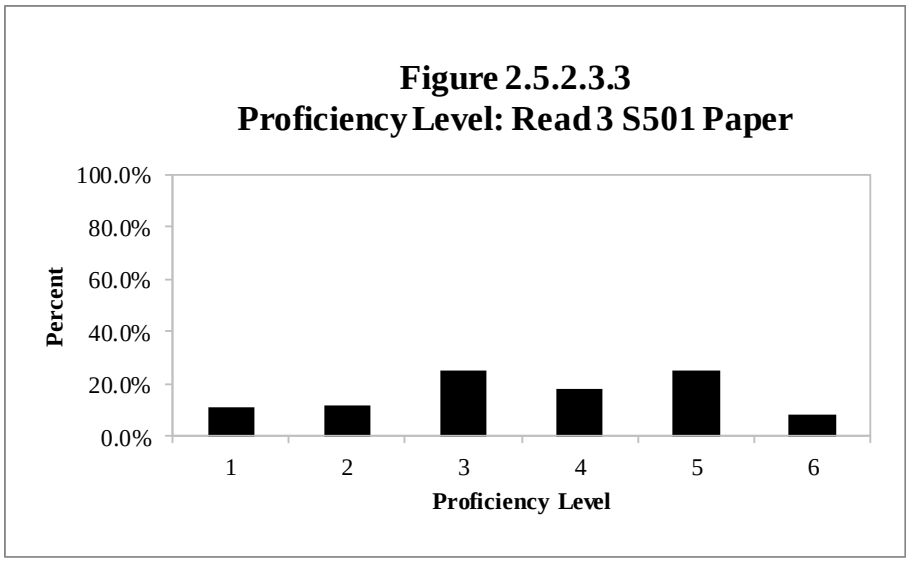


Table 2.5.2.3.3

Proficiency Level Distribution: Read 3 S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	3,444	11.06%	3,444	11.06%
2	3,692	11.86%	3,692	11.86%
3	7,804	25.07%	7,804	25.07%
4	5,635	18.10%	5,635	18.10%
5	7,868	25.27%	7,868	25.27%
6	2,692	8.65%	2,692	8.65%
Total	31,135	100.00%	31,135	100.00%



2.5.2.4 Grades 4–5

Table 2.5.2.4.1

Proficiency Level Distribution: Read 4-5 A S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,701	56.11%	2,607	58.44%	5,308	57.23%
2	1,043	21.67%	810	18.16%	1,853	19.98%
3	532	11.05%	601	13.47%	1,133	12.22%
4	214	4.45%	202	4.53%	416	4.49%
5	275	5.71%	241	5.40%	516	5.56%
6	49	1.02%	0	0.00%	49	0.53%
Total	4,814	100.00%	4,461	100.00%	9,275	100.00%

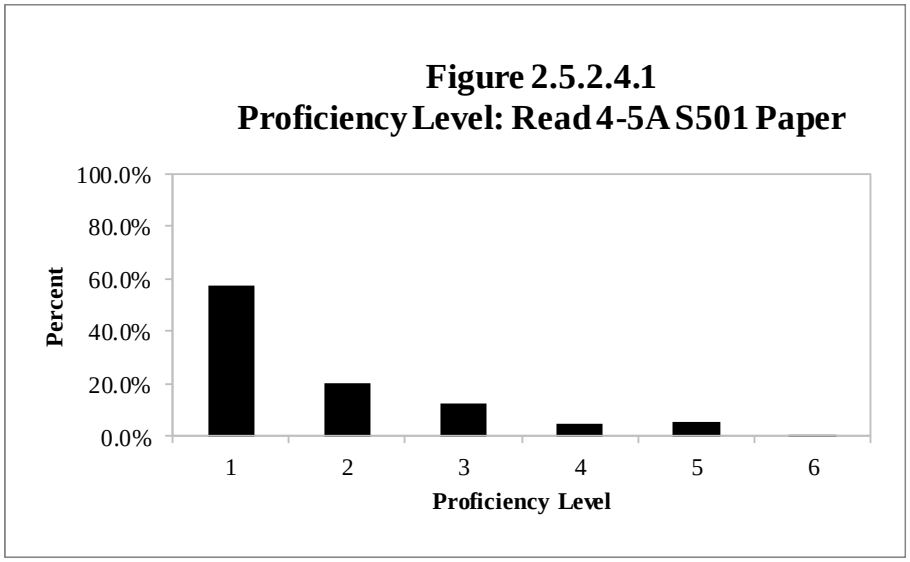


Table 2.5.2.4.2

Proficiency Level Distribution: Read 4-5 B/C S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	101	0.52%	151	0.87%	252	0.69%
2	2,276	11.78%	2,717	15.60%	4,993	13.59%
3	6,065	31.40%	4,683	26.88%	10,748	29.26%
4	3,215	16.64%	2,607	14.97%	5,822	15.85%
5	4,739	24.53%	4,359	25.02%	9,098	24.76%
6	2,922	15.13%	2,903	16.66%	5,825	15.86%
Total	19,318	100.00%	17,420	100.00%	36,738	100.00%

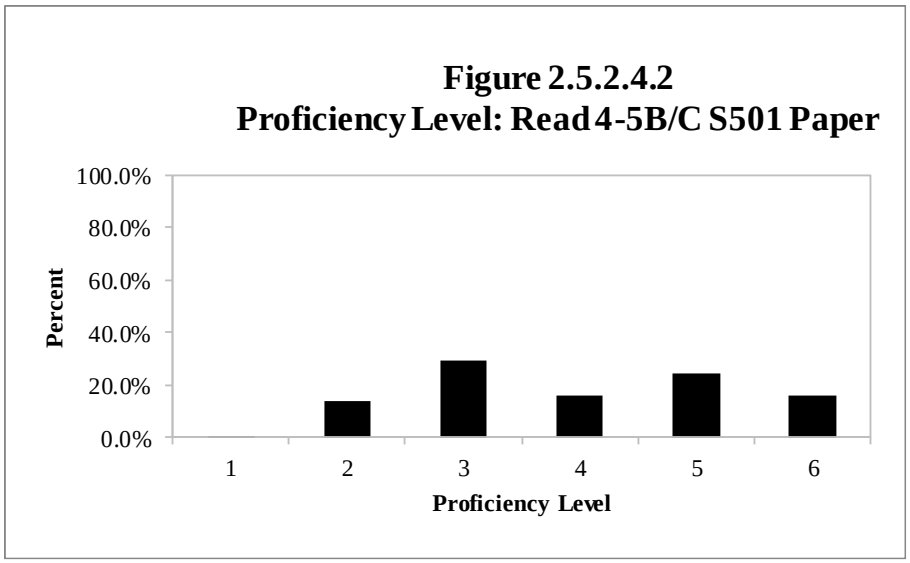
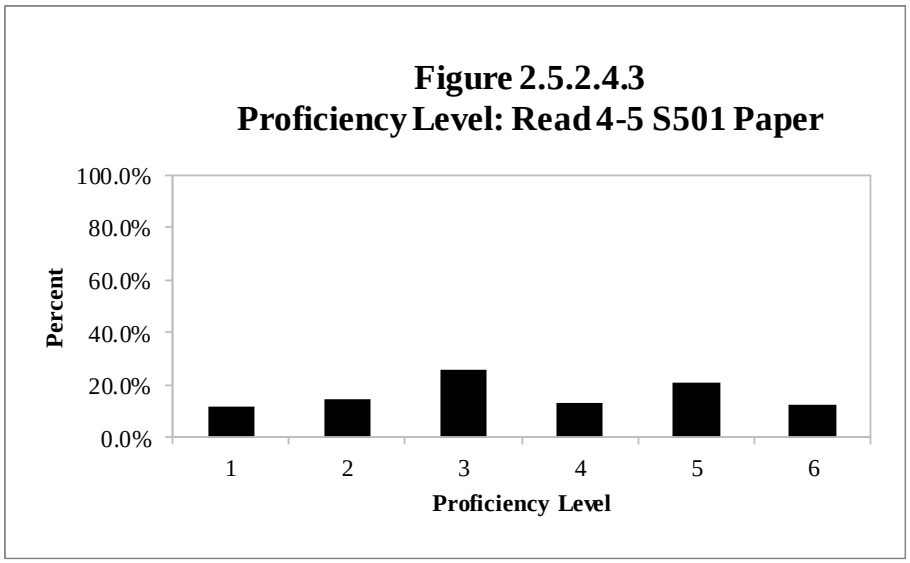


Table 2.5.2.4.3

Proficiency Level Distribution: Read 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,802	11.61%	2,758	12.60%	5,560	12.08%
2	3,319	13.75%	3,527	16.12%	6,846	14.88%
3	6,597	27.34%	5,284	24.15%	11,881	25.82%
4	3,429	14.21%	2,809	12.84%	6,238	13.56%
5	5,014	20.78%	4,600	21.02%	9,614	20.89%
6	2,971	12.31%	2,903	13.27%	5,874	12.77%
Total	24,132	100.00%	21,881	100.00%	46,013	100.00%



2.5.2.5 Grades 6–8

Table 2.5.2.5.1

Proficiency Level Distribution: Read 6-8 A S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,861	45.87%	2,244	56.65%	2,314	57.49%	6,419	53.30%
2	1,508	37.17%	1,099	27.75%	1,150	28.57%	3,757	31.20%
3	422	10.40%	397	10.02%	298	7.40%	1,117	9.28%
4	105	2.59%	89	2.25%	100	2.48%	294	2.44%
5	120	2.96%	61	1.54%	138	3.43%	319	2.65%
6	41	1.01%	71	1.79%	25	0.62%	137	1.14%
Total	4,057	100.00%	3,961	100.00%	4,025	100.00%	12,043	100.00%

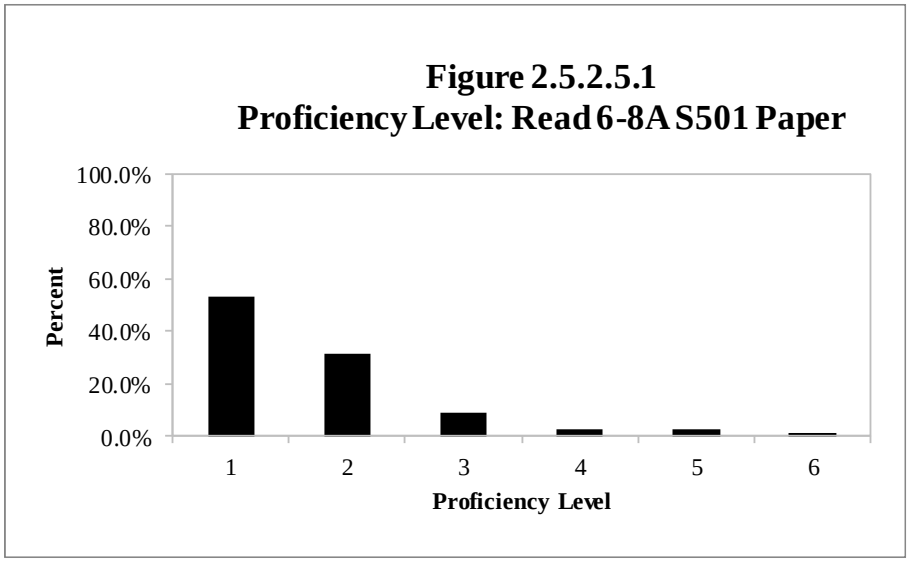


Table 2.5.2.5.2

Proficiency Level Distribution: Read 6-8 B/C S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	255	1.95%	392	3.52%	481	4.65%	1,128	3.26%
2	4,610	35.16%	3,689	33.17%	3,506	33.93%	11,805	34.15%
3	4,229	32.26%	3,787	34.05%	2,808	27.17%	10,824	31.31%
4	1,627	12.41%	1,325	11.91%	1,353	13.09%	4,305	12.45%
5	1,828	13.94%	1,295	11.64%	1,455	14.08%	4,578	13.24%
6	561	4.28%	635	5.71%	731	7.07%	1,927	5.57%
Total	13,110	100.00%	11,123	100.00%	10,334	100.00%	34,567	100.00%

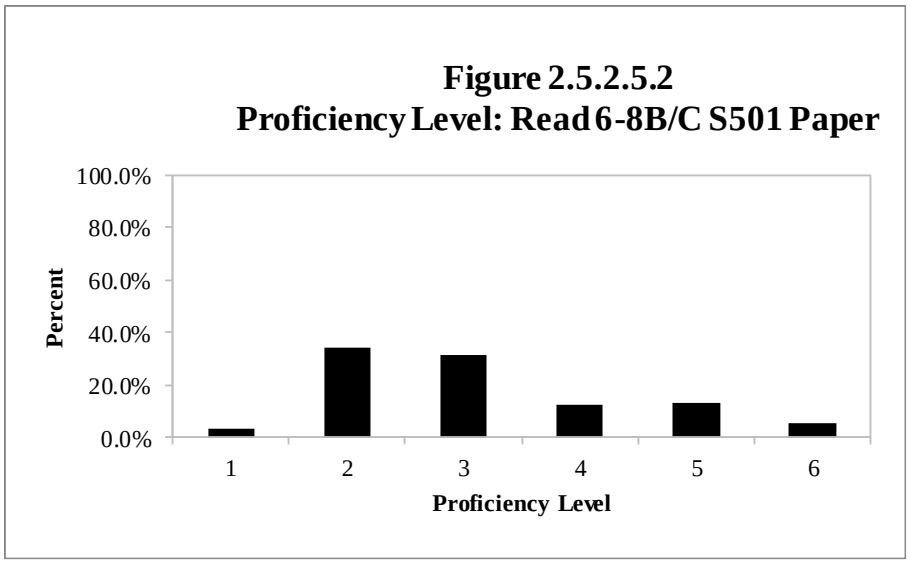
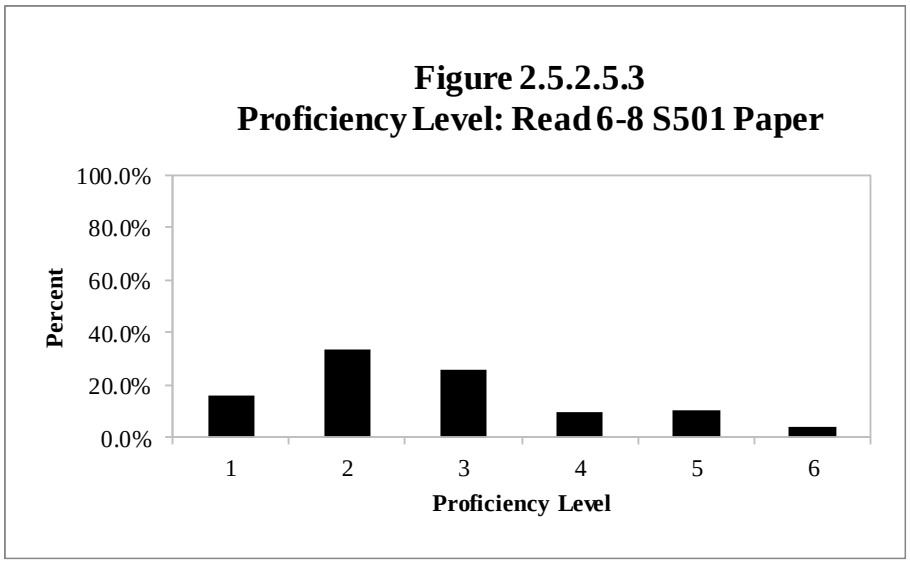


Table 2.5.2.5.3

Proficiency Level Distribution: Read 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,116	12.33%	2,636	17.48%	2,795	19.47%	7,547	16.19%
2	6,118	35.64%	4,788	31.74%	4,656	32.43%	15,562	33.39%
3	4,651	27.09%	4,184	27.74%	3,106	21.63%	11,941	25.62%
4	1,732	10.09%	1,414	9.37%	1,453	10.12%	4,599	9.87%
5	1,948	11.35%	1,356	8.99%	1,593	11.09%	4,897	10.51%
6	602	3.51%	706	4.68%	756	5.26%	2,064	4.43%
Total	17,167	100.00%	15,084	100.00%	14,359	100.00%	46,610	100.00%



2.5.2.6 Grades 9–12

Table 2.5.2.6.1

Proficiency Level Distribution: Read 9-12 A S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,572	33.03%	1,202	34.46%	922	34.58%	538	37.05%	4,234	34.24%
2	2,133	44.82%	1,431	41.03%	1,069	40.10%	633	43.60%	5,266	42.59%
3	590	12.40%	461	13.22%	295	11.07%	147	10.12%	1,493	12.07%
4	140	2.94%	139	3.99%	123	4.61%	62	4.27%	464	3.75%
5	248	5.21%	123	3.53%	215	8.06%	52	3.58%	638	5.16%
6	76	1.60%	132	3.78%	42	1.58%	20	1.38%	270	2.18%
Total	4,759	100.00%	3,488	100.00%	2,666	100.00%	1,452	100.00%	12,365	100.00%

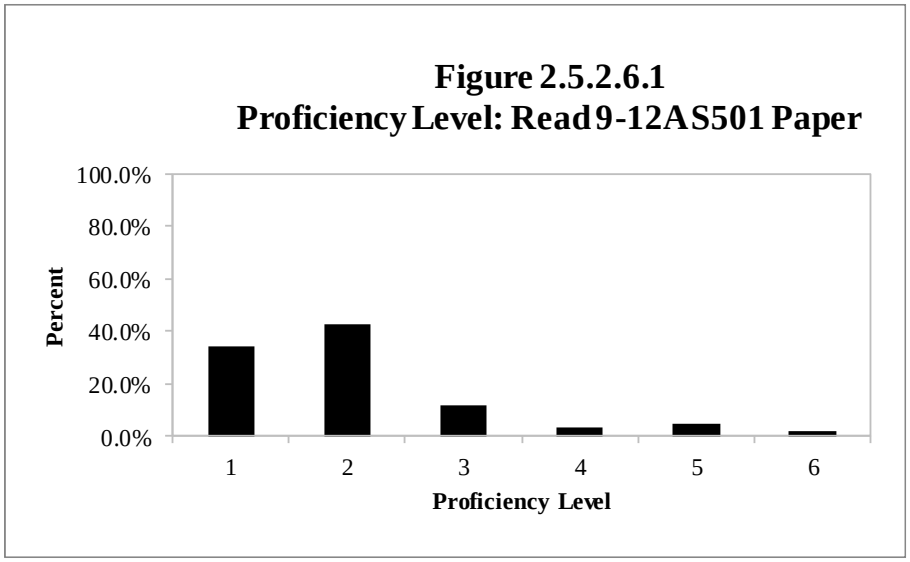


Table 2.5.2.6.2

Proficiency Level Distribution: Read 9-12 B/C S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	146	1.55%	101	1.10%	189	2.23%	333	5.54%	769	2.32%
2	2,460	26.07%	2,688	29.16%	2,525	29.84%	2,308	38.40%	9,981	30.13%
3	2,374	25.16%	2,740	29.73%	2,374	28.05%	1,643	27.33%	9,131	27.56%
4	1,662	17.61%	1,058	11.48%	508	6.00%	307	5.11%	3,535	10.67%
5	1,415	14.99%	1,406	15.25%	1,386	16.38%	783	13.03%	4,990	15.06%
6	1,380	14.62%	1,224	13.28%	1,481	17.50%	637	10.60%	4,722	14.25%
Total	9,437	100.00%	9,217	100.00%	8,463	100.00%	6,011	100.00%	33,128	100.00%

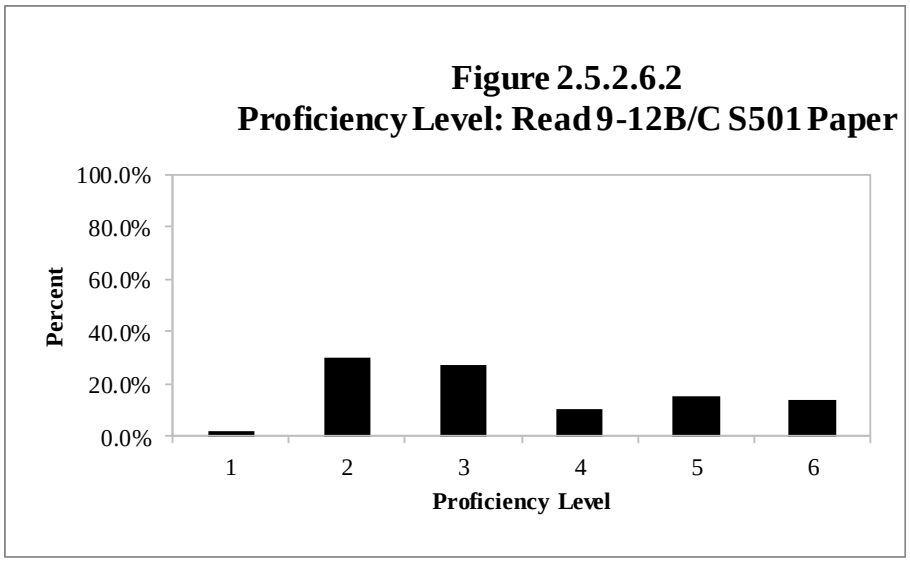
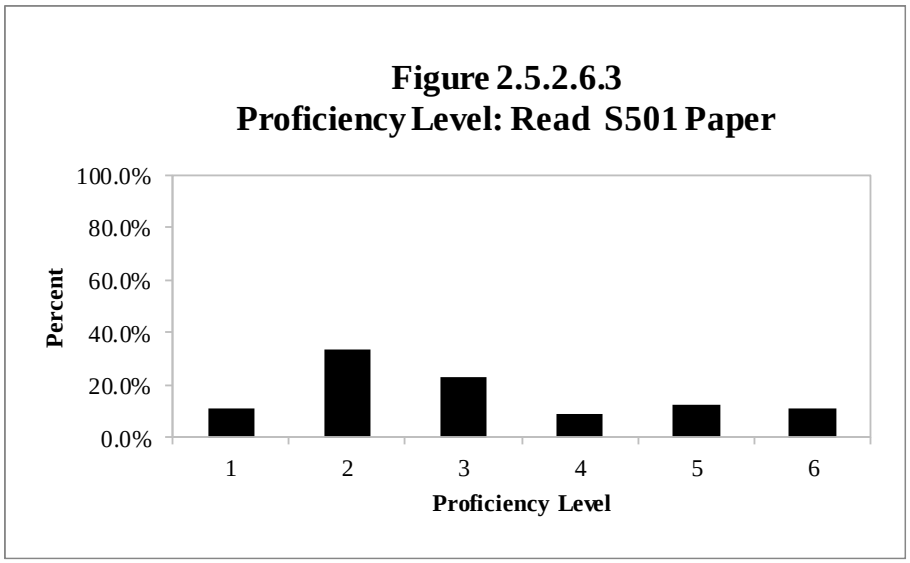


Table 2.5.2.6.3

Proficiency Level Distribution: Read 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,718	12.10%	1,303	10.26%	1,111	9.98%	871	11.67%	5,003	11.00%
2	4,593	32.35%	4,119	32.42%	3,594	32.29%	2,941	39.41%	15,247	33.52%
3	2,964	20.88%	3,201	25.19%	2,669	23.98%	1,790	23.98%	10,624	23.35%
4	1,802	12.69%	1,197	9.42%	631	5.67%	369	4.94%	3,999	8.79%
5	1,663	11.71%	1,529	12.03%	1,601	14.39%	835	11.19%	5,628	12.37%
6	1,456	10.26%	1,356	10.67%	1,523	13.68%	657	8.80%	4,992	10.97%
Total	14,196	100.00%	12,705	100.00%	11,129	100.00%	7,463	100.00%	45,493	100.00%



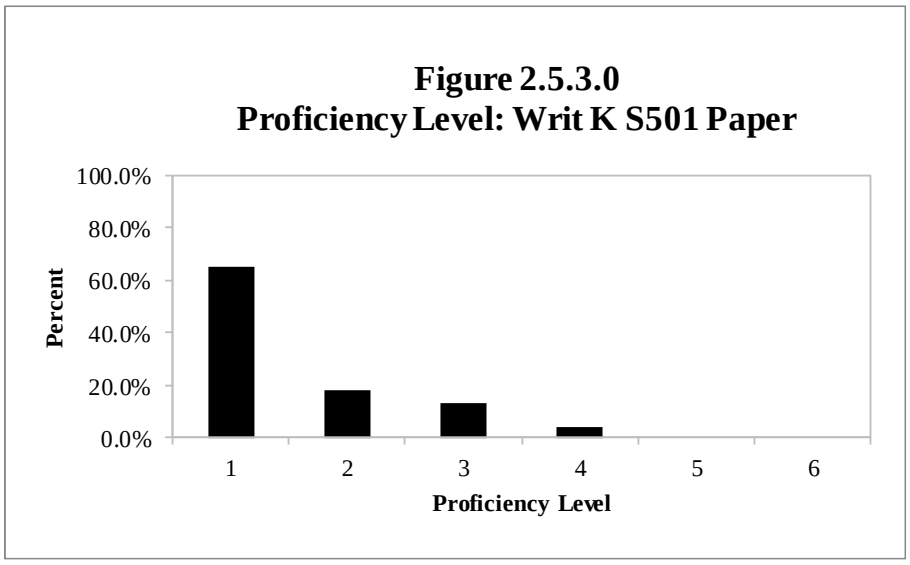
2.5.3 Writing

2.5.3.0 Kindergarten

Table 2.5.3.0

Proficiency Level Distribution: Writ K S501 Paper

Level	Grade K		Total	
	Count	Percent	Count	Percent
1	147,048	65.07%	147,048	65.07%
2	40,361	17.86%	40,361	17.86%
3	30,054	13.30%	30,054	13.30%
4	8,524	3.77%	8,524	3.77%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	225,987	100.00%	225,987	100.00%



2.5.3.1 Grade 1

Table 2.5.3.1.1

Proficiency Level Distribution: Writ 1 A S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	8,155	40.85%	8,155	40.85%
2	10,851	54.35%	10,851	54.35%
3	959	4.80%	959	4.80%
4	0	0.00%	0	0.00%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	19,965	100.00%	19,965	100.00%

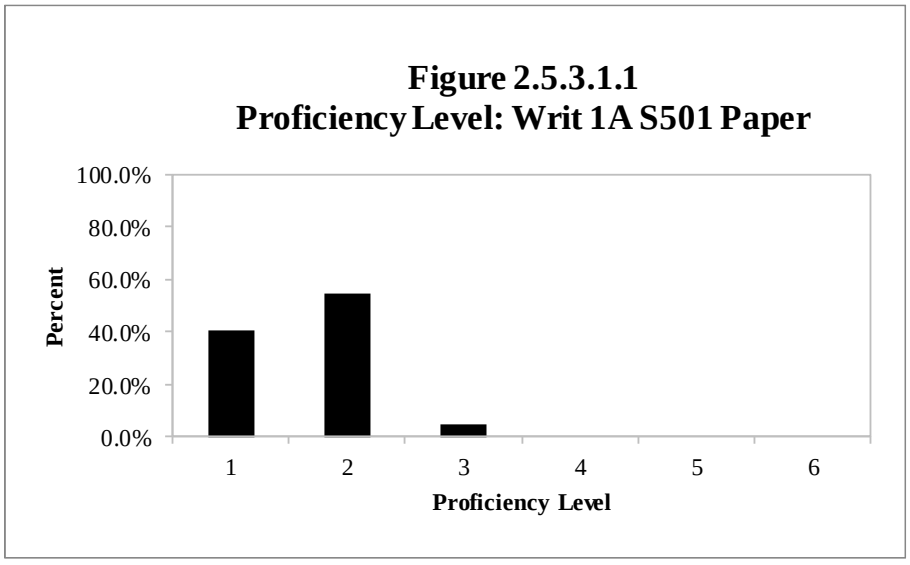


Table 2.5.3.1.2

Proficiency Level Distribution: Writ 1 B/C S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	2,412	11.61%	2,412	11.61%
2	6,427	30.93%	6,427	30.93%
3	11,446	55.08%	11,446	55.08%
4	486	2.34%	486	2.34%
5	8	0.04%	8	0.04%
6	0	0.00%	0	0.00%
Total	20,779	100.00%	20,779	100.00%

Figure 2.5.3.1.2
Proficiency Level: Writ 1B/C S501 Paper

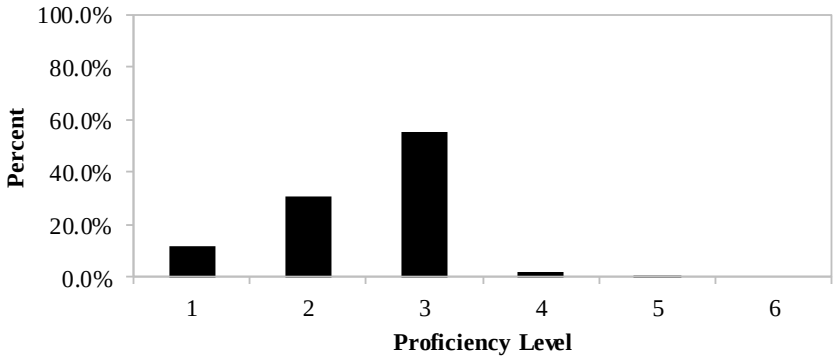
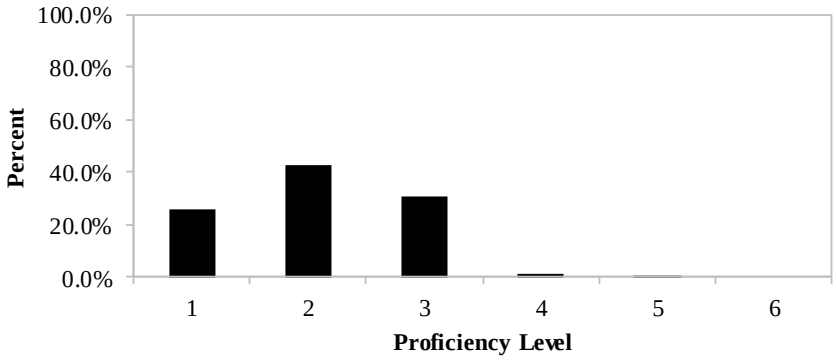


Table 2.5.3.1.3

Proficiency Level Distribution: Writ 1 S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	10,567	25.94%	10,567	25.94%
2	17,278	42.41%	17,278	42.41%
3	12,405	30.45%	12,405	30.45%
4	486	1.19%	486	1.19%
5	8	0.02%	8	0.02%
6	0	0.00%	0	0.00%
Total	40,744	100.00%	40,744	100.00%

Figure 2.5.3.1.3
Proficiency Level: Writ 1 S501 Paper



2.5.3.2 Grade 2

Table 2.5.3.2.1

Proficiency Level Distribution: Writ 2 A S501 Paper

Level	Grade 2		Total	
	Count	**Percent**	**Count**	**Percent**
1	3,885	42.58%	3,885	42.58%
2	2,951	32.34%	2,951	32.34%
3	2,285	25.04%	2,285	25.04%
4	3	0.03%	3	0.03%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	9,124	100.00%	9,124	100.00%

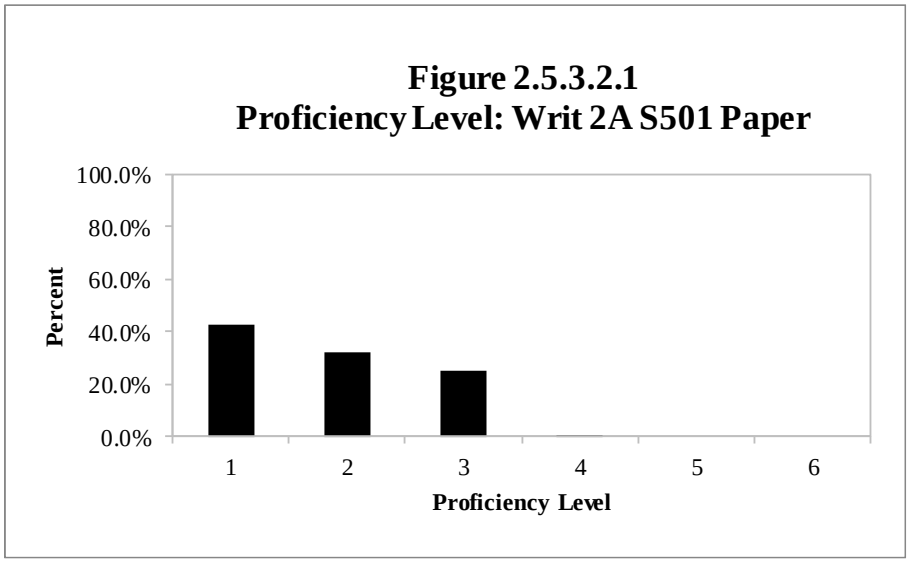


Table 2.5.3.2.2

Proficiency Level Distribution: Writ 2 B/C S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,442	4.70%	1,442	4.70%
2	6,044	19.69%	6,044	19.69%
3	20,821	67.83%	20,821	67.83%
4	2,359	7.69%	2,359	7.69%
5	28	0.09%	28	0.09%
6	0	0.00%	0	0.00%
Total	30,694	100.00%	30,694	100.00%

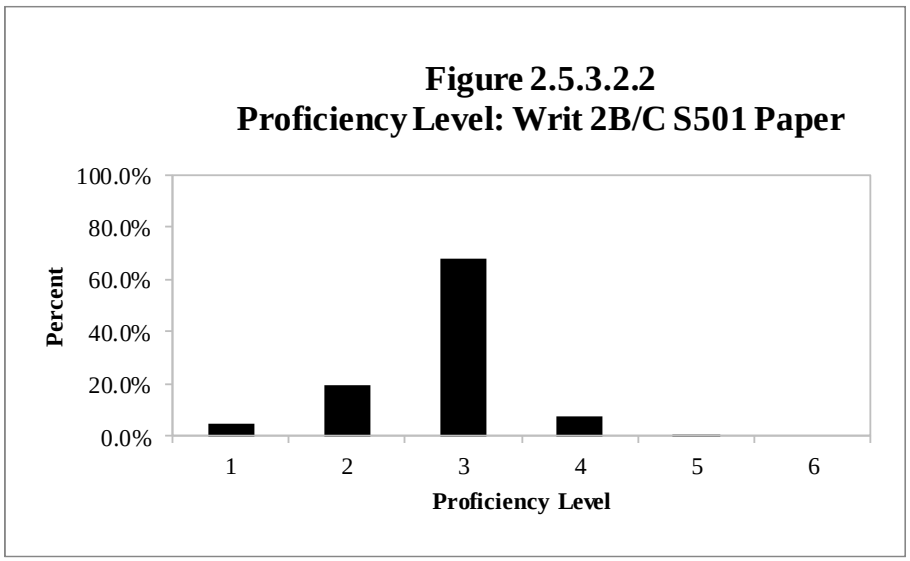
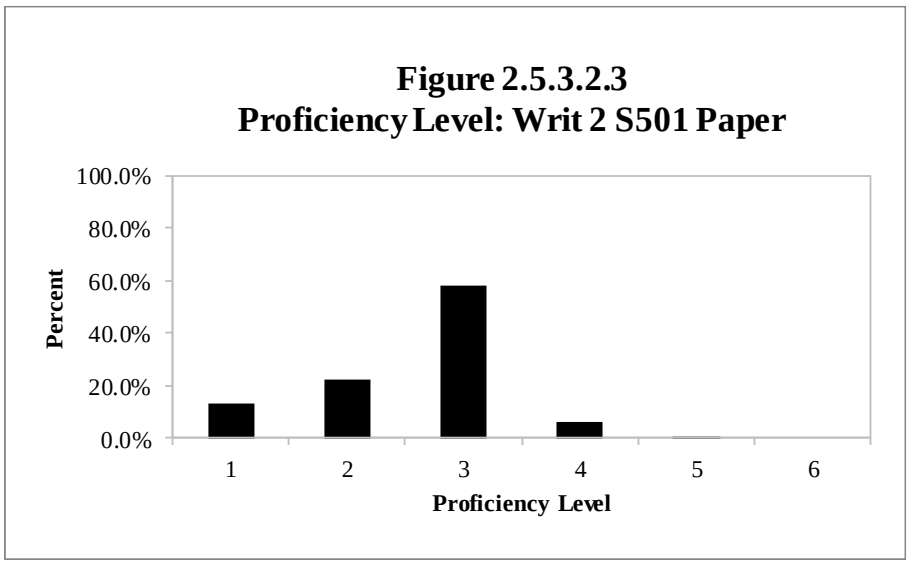


Table 2.5.3.2.3

Proficiency Level Distribution: Writ 2 S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	5,327	13.38%	5,327	13.38%
2	8,995	22.59%	8,995	22.59%
3	23,106	58.03%	23,106	58.03%
4	2,362	5.93%	2,362	5.93%
5	28	0.07%	28	0.07%
6	0	0.00%	0	0.00%
Total	39,818	100.00%	39,818	100.00%



2.5.3.3 Grade 3

Table 2.5.3.3.1

Proficiency Level Distribution: Writ 3 A S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	2,839	38.04%	2,839	38.04%
2	3,020	40.46%	3,020	40.46%
3	1,600	21.44%	1,600	21.44%
4	5	0.07%	5	0.07%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	7,464	100.00%	7,464	100.00%

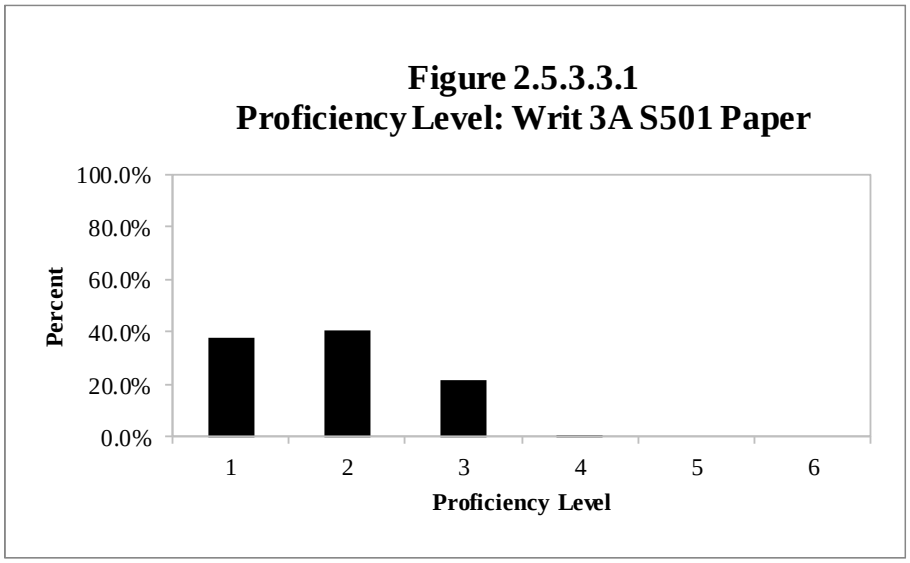


Table 2.5.3.3.2

Proficiency Level Distribution: Writ 3 B/C S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	812	2.74%	812	2.74%
2	2,573	8.67%	2,573	8.67%
3	23,193	78.17%	23,193	78.17%
4	3,067	10.34%	3,067	10.34%
5	23	0.08%	23	0.08%
6	1	0.00%	1	0.00%
Total	29,669	100.00%	29,669	100.00%

Figure 2.5.3.3.2
Proficiency Level: Writ 3B/C S501 Paper

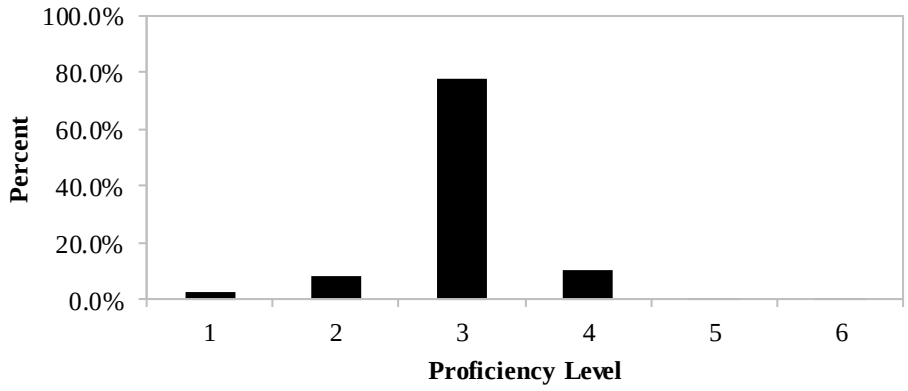
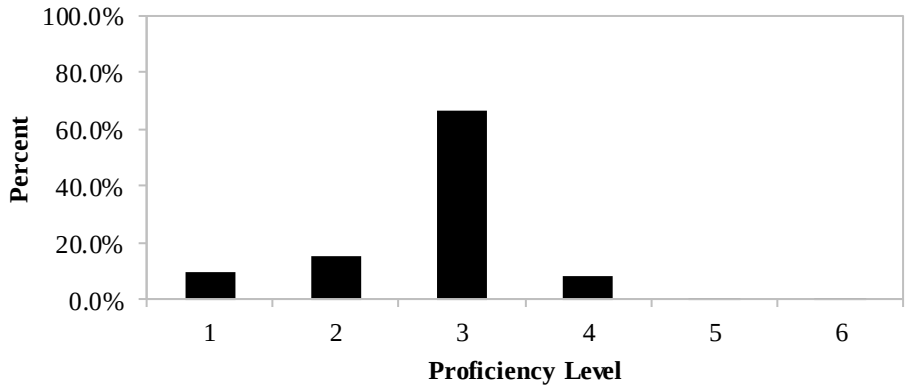


Table 2.5.3.3.3

Proficiency Level Distribution: Writ 3 S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	3,651	9.83%	3,651	9.83%
2	5,593	15.06%	5,593	15.06%
3	24,793	66.77%	24,793	66.77%
4	3,072	8.27%	3,072	8.27%
5	23	0.06%	23	0.06%
6	1	0.00%	1	0.00%
Total	37,133	100.00%	37,133	100.00%

Figure 2.5.3.3.3
Proficiency Level: Writ 3 S501 Paper



2.5.3.4 Grades 4–5

Table 2.5.3.4.1

Proficiency Level Distribution: Writ 4-5 A S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,752	32.34%	1,301	26.36%	3,053	29.49%
2	1,416	26.14%	1,304	26.42%	2,720	26.27%
3	2,209	40.77%	2,309	46.78%	4,518	43.64%
4	41	0.76%	22	0.45%	63	0.61%
5	0	0.00%	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%	0	0.00%
Total	5,418	100.00%	4,936	100.00%	10,354	100.00%

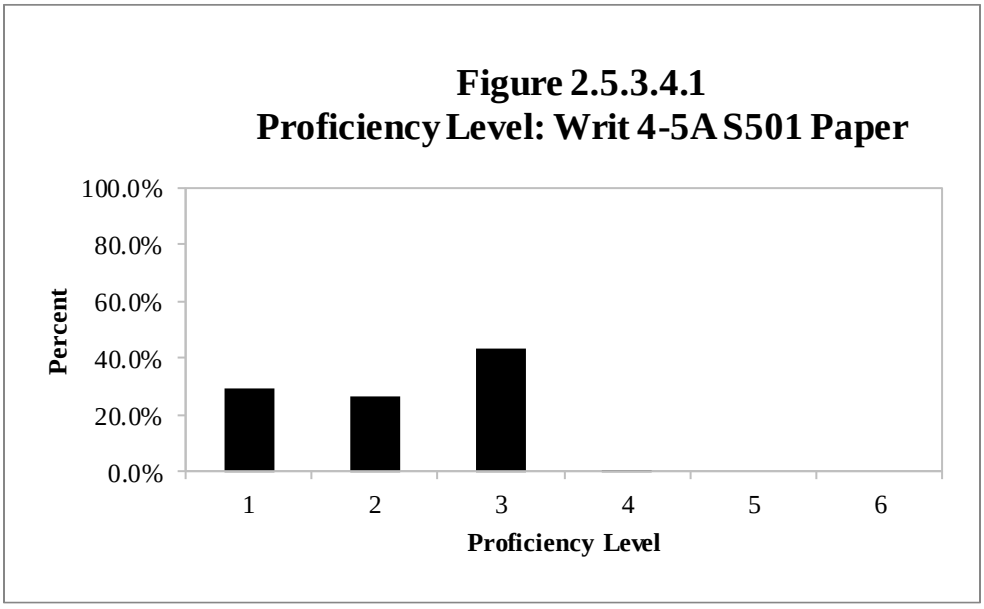


Table 2.5.3.4.2

Proficiency Level Distribution: Writ 4-5 B/C S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	345	1.57%	220	1.14%	565	1.36%
2	580	2.63%	409	2.11%	989	2.39%
3	13,483	61.17%	9,924	51.27%	23,407	56.54%
4	7,546	34.23%	8,620	44.53%	16,166	39.05%
5	82	0.37%	180	0.93%	262	0.63%
6	6	0.03%	4	0.02%	10	0.02%
Total	22,042	100.00%	19,357	100.00%	41,399	100.00%

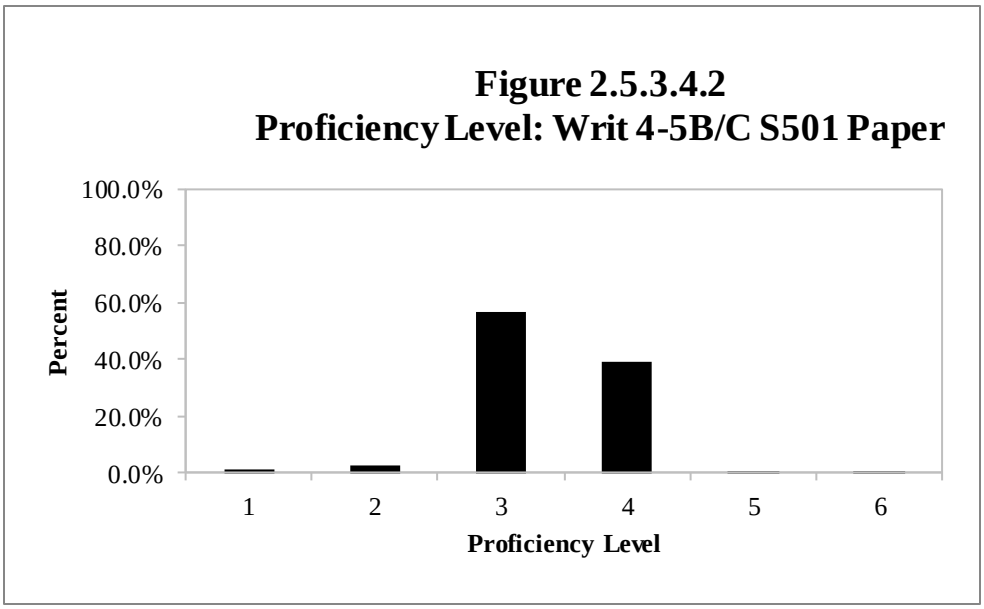
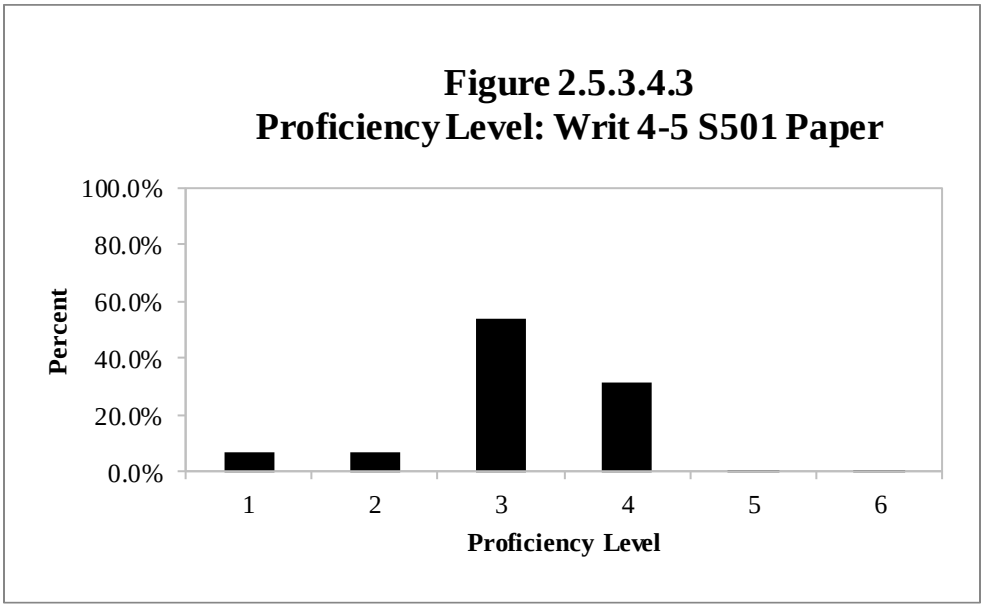


Table 2.5.3.4.3

Proficiency Level Distribution: Writ 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,097	7.64%	1,521	6.26%	3,618	6.99%
2	1,996	7.27%	1,713	7.05%	3,709	7.17%
3	15,692	57.14%	12,233	50.36%	27,925	53.96%
4	7,587	27.63%	8,642	35.57%	16,229	31.36%
5	82	0.30%	180	0.74%	262	0.51%
6	6	0.02%	4	0.02%	10	0.02%
Total	27,460	100.00%	24,293	100.00%	51,753	100.00%



2.5.3.5 Grades 6–8

Table 2.5.3.5.1

Proficiency Level Distribution: Writ 6-8 A S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,825	39.82%	1,966	44.93%	2,203	49.17%	5,994	44.60%
2	1,515	33.06%	1,604	36.65%	1,339	29.89%	4,458	33.17%
3	1,217	26.55%	763	17.44%	917	20.47%	2,897	21.56%
4	25	0.55%	42	0.96%	21	0.47%	88	0.65%
5	1	0.02%	1	0.02%	0	0.00%	2	0.01%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	4,583	100.00%	4,376	100.00%	4,480	100.00%	13,439	100.00%

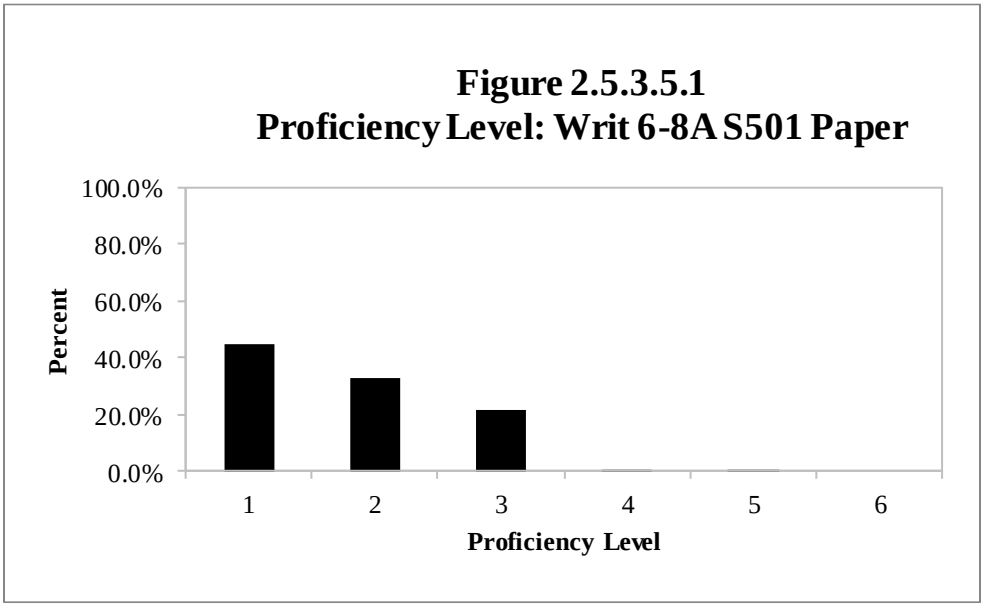


Table 2.5.3.5.2

Proficiency Level Distribution: Writ 6-8 B/C S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	268	1.72%	267	2.05%	346	2.84%	881	2.16%
2	734	4.71%	632	4.86%	513	4.21%	1,879	4.61%
3	9,360	60.08%	8,306	63.81%	8,041	66.06%	25,707	63.06%
4	5,197	33.36%	3,802	29.21%	3,269	26.86%	12,268	30.09%
5	19	0.12%	10	0.08%	3	0.02%	32	0.08%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	15,578	100.00%	13,017	100.00%	12,172	100.00%	40,767	100.00%

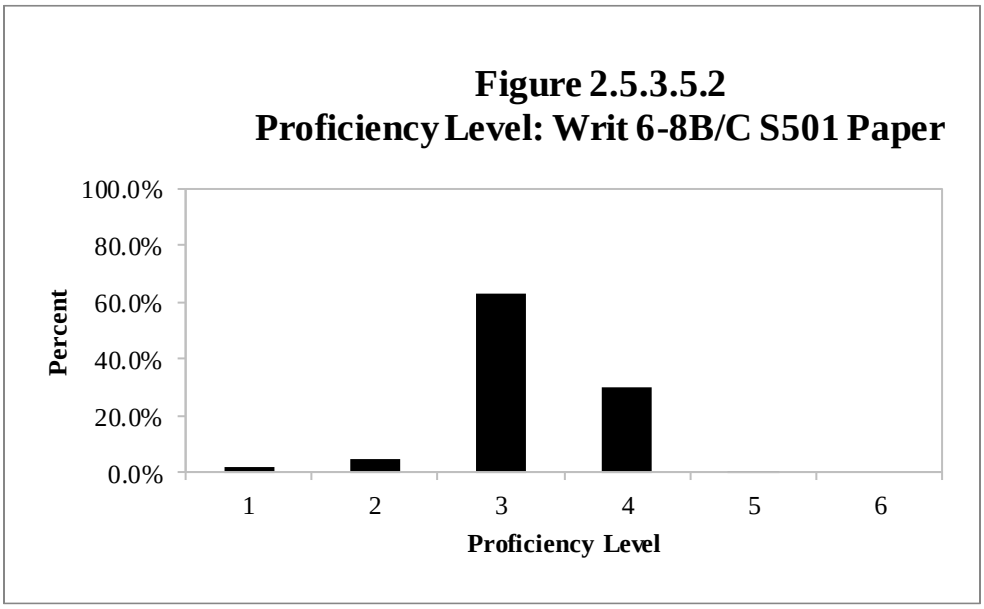
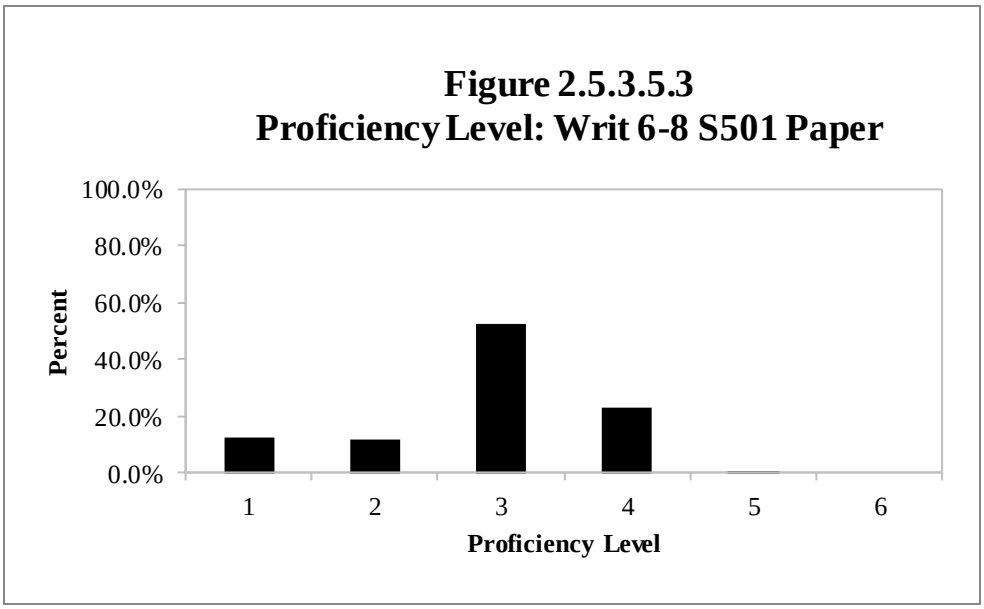


Table 2.5.3.5.3

Proficiency Level Distribution: Writ 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,093	10.38%	2,233	12.84%	2,549	15.31%	6,875	12.68%
2	2,249	11.16%	2,236	12.86%	1,852	11.12%	6,337	11.69%
3	10,577	52.46%	9,069	52.14%	8,958	53.80%	28,604	52.77%
4	5,222	25.90%	3,844	22.10%	3,290	19.76%	12,356	22.79%
5	20	0.10%	11	0.06%	3	0.02%	34	0.06%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	20,161	100.00%	17,393	100.00%	16,652	100.00%	54,206	100.00%



2.5.3.6 Grades 9–12

Table 2.5.3.6.1

Proficiency Level Distribution: Writ 9-12 A S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,782	33.76%	1,209	31.79%	1,016	35.36%	713	45.39%	4,720	34.90%
2	1,695	32.11%	1,345	35.37%	946	32.93%	292	18.59%	4,278	31.63%
3	1,556	29.48%	1,131	29.74%	860	29.93%	535	34.05%	4,082	30.18%
4	246	4.66%	117	3.08%	51	1.78%	31	1.97%	445	3.29%
5	0	0.00%	1	0.03%	0	0.00%	0	0.00%	1	0.01%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	5,279	100.00%	3,803	100.00%	2,873	100.00%	1,571	100.00%	13,526	100.00%

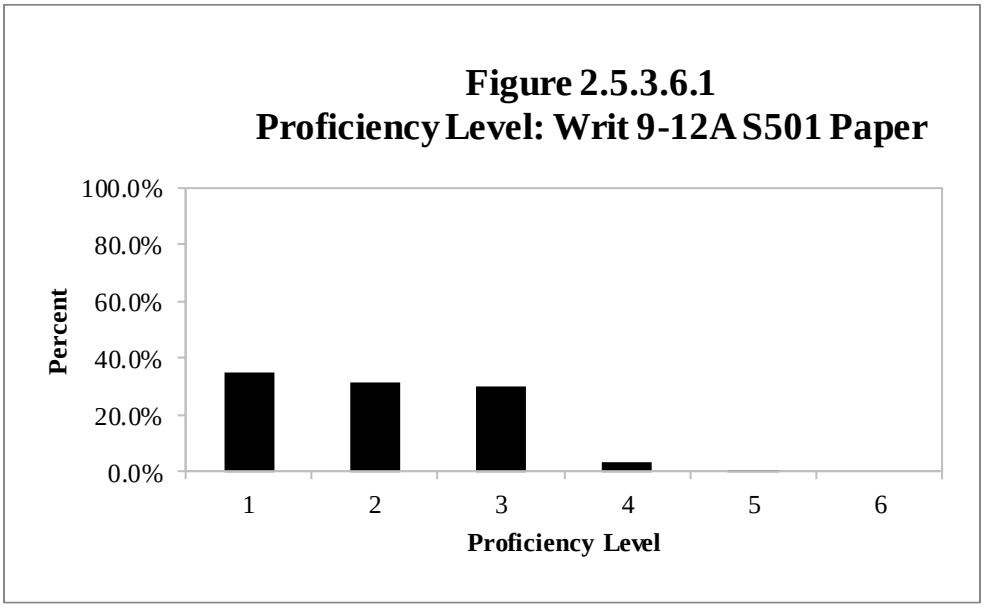


Table 2.5.3.6.2

Proficiency Level Distribution: Writ 9-12 B/C S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	272	2.44%	557	5.20%	575	5.94%	674	9.98%	2,078	5.43%
2	660	5.93%	564	5.26%	592	6.12%	701	10.38%	2,517	6.58%
3	4,026	36.18%	5,025	46.89%	4,425	45.71%	3,464	51.29%	16,940	44.25%
4	6,016	54.06%	4,428	41.32%	4,008	41.40%	1,895	28.06%	16,347	42.70%
5	154	1.38%	143	1.33%	80	0.83%	20	0.30%	397	1.04%
6	1	0.01%	0	0.00%	1	0.01%	0	0.00%	2	0.01%
Total	11,129	100.00%	10,717	100.00%	9,681	100.00%	6,754	100.00%	38,281	100.00%

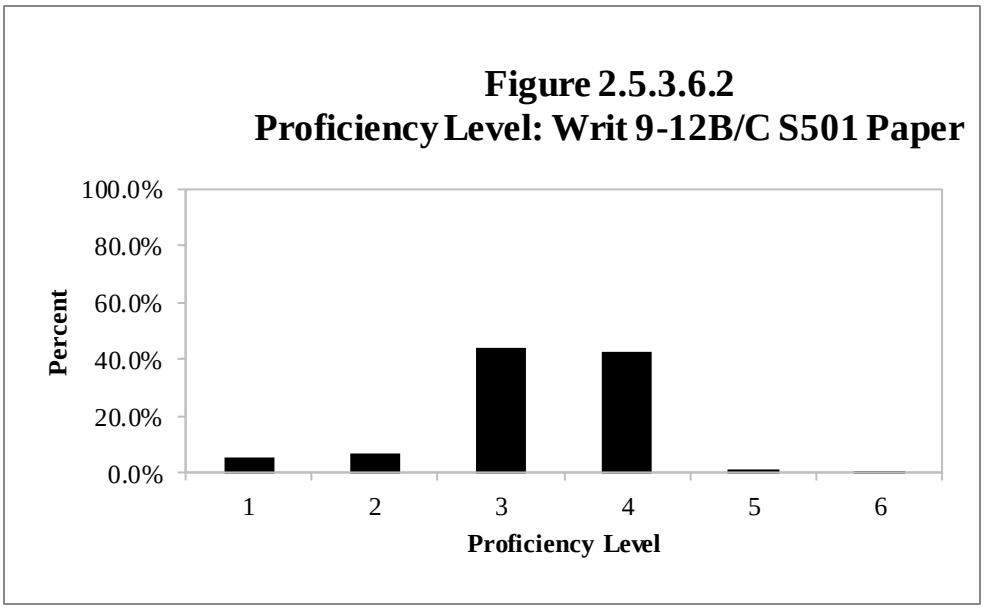
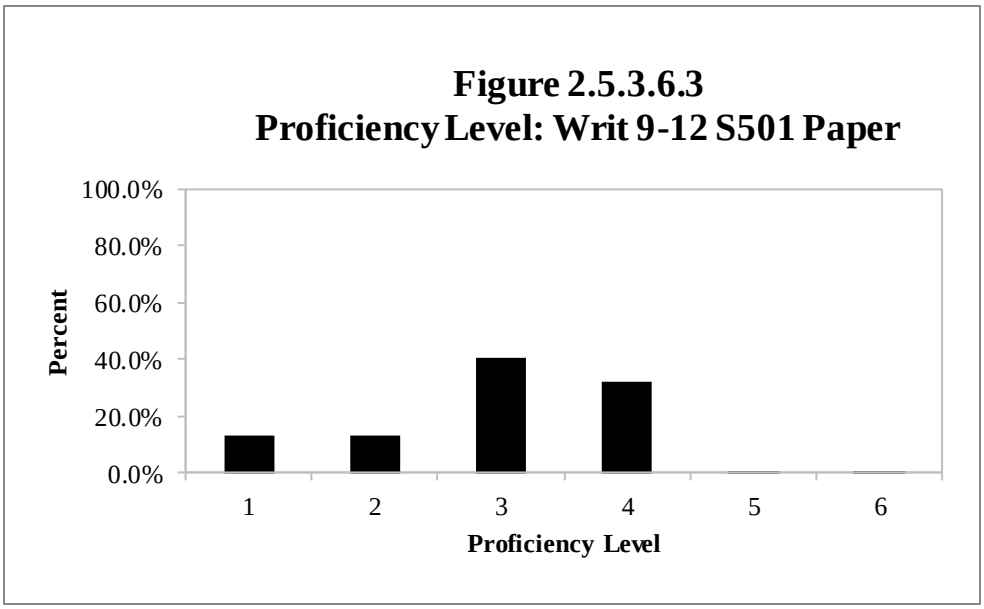


Table 2.5.3.6.3

Proficiency Level Distribution: Writ 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,054	12.52%	1,766	12.16%	1,591	12.67%	1,387	16.66%	6,798	13.12%
2	2,355	14.35%	1,909	13.15%	1,538	12.25%	993	11.93%	6,795	13.12%
3	5,582	34.02%	6,156	42.40%	5,285	42.10%	3,999	48.04%	21,022	40.58%
4	6,262	38.16%	4,545	31.30%	4,059	32.33%	1,926	23.14%	16,792	32.41%
5	154	0.94%	144	0.99%	80	0.64%	20	0.24%	398	0.77%
6	1	0.01%	0	0.00%	1	0.01%	0	0.00%	2	0.00%
Total	16,408	100.00%	14,520	100.00%	12,554	100.00%	8,325	100.00%	51,807	100.00%



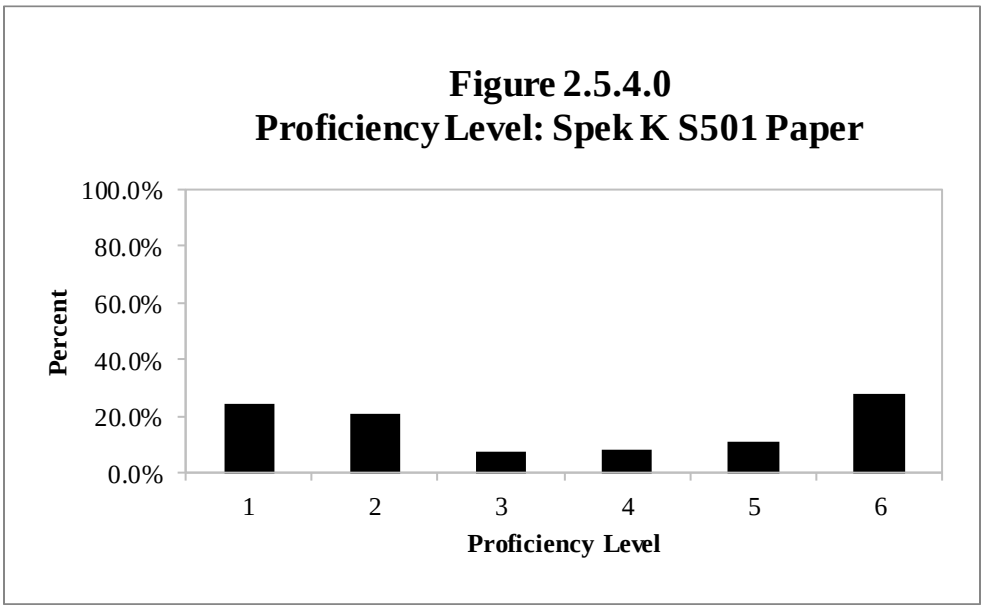
2.5.4 Speaking

2.5.4.0 Kindergarten

Table 2.5.4.0

Proficiency Level Distribution: Spek K S501 Paper

	Grade K		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	54,732	24.22%	54,732	24.22%
2	47,538	21.03%	47,538	21.03%
3	17,173	7.60%	17,173	7.60%
4	19,122	8.46%	19,122	8.46%
5	24,937	11.03%	24,937	11.03%
6	62,498	27.65%	62,498	27.65%
Total	226,000	100.00%	226,000	100.00%



2.5.4.1 Grade 1

Table 2.5.4.1.1

Proficiency Level Distribution: Spek 1 A S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	4,711	24.16%	4,711	24.16%
2	6,684	34.27%	6,684	34.27%
3	4,693	24.06%	4,693	24.06%
4	2,963	15.19%	2,963	15.19%
5	452	2.32%	452	2.32%
6	0	0.00%	0	0.00%
Total	19,503	100.00%	19,503	100.00%

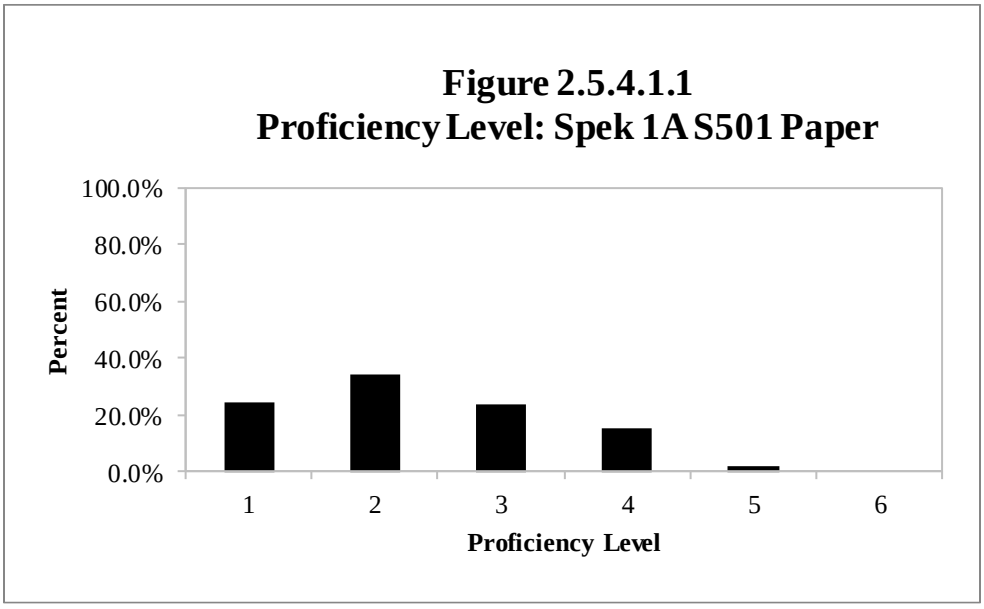


Table 2.5.4.1.2

Proficiency Level Distribution: Spek 1 B/C S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	405	1.98%	405	1.98%
2	5,443	26.64%	5,443	26.64%
3	6,760	33.09%	6,760	33.09%
4	5,143	25.17%	5,143	25.17%
5	2,290	11.21%	2,290	11.21%
6	391	1.91%	391	1.91%
Total	20,432	100.00%	20,432	100.00%

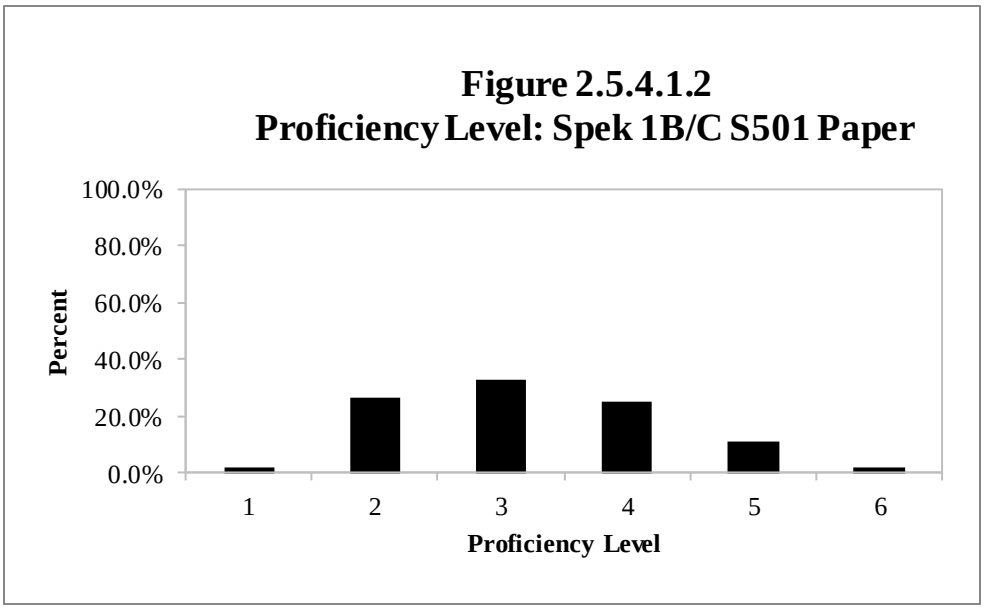
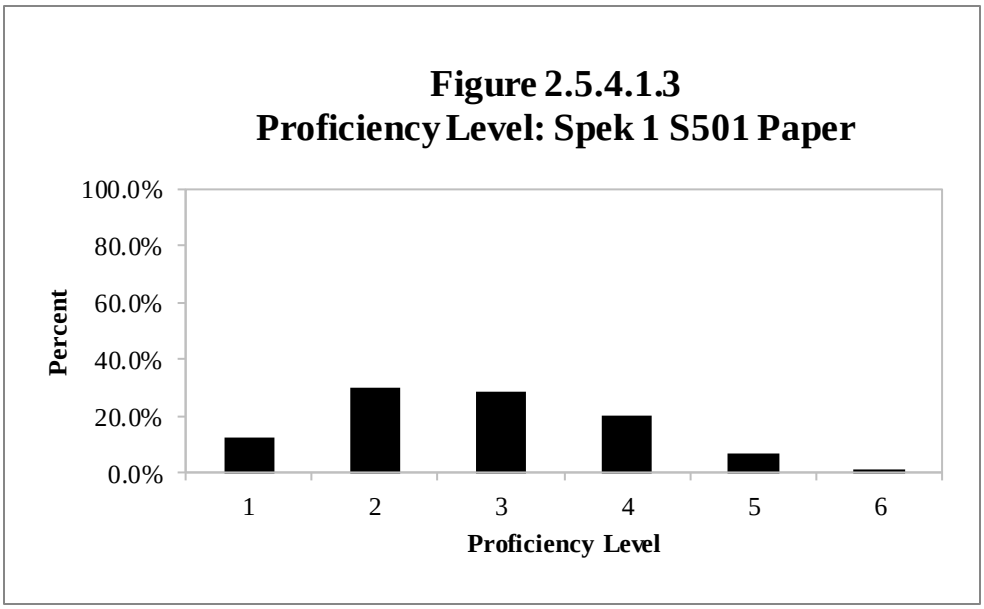


Table 2.5.4.1.3

Proficiency Level Distribution: Spek 1 S501 Paper

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	5,116	12.81%	5,116	12.81%
2	12,127	30.37%	12,127	30.37%
3	11,453	28.68%	11,453	28.68%
4	8,106	20.30%	8,106	20.30%
5	2,742	6.87%	2,742	6.87%
6	391	0.98%	391	0.98%
Total	39,935	100.00%	39,935	100.00%



2.5.4.2 Grade 2

Table 2.5.4.2.1

Proficiency Level Distribution: Spek 2 A S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	3,650	41.04%	3,650	41.04%
2	2,009	22.59%	2,009	22.59%
3	2,469	27.76%	2,469	27.76%
4	543	6.11%	543	6.11%
5	223	2.51%	223	2.51%
6	0	0.00%	0	0.00%
Total	8,894	100.00%	8,894	100.00%

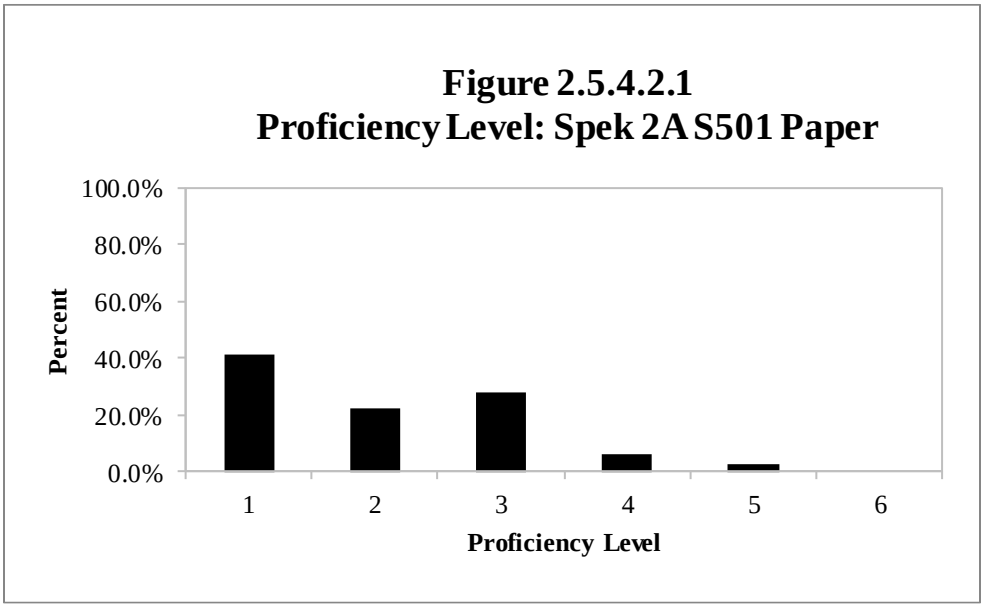


Table 2.5.4.2.2

Proficiency Level Distribution: Spek 2 B/C S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,490	4.94%	1,490	4.94%
2	5,506	18.24%	5,506	18.24%
3	12,012	39.79%	12,012	39.79%
4	7,443	24.65%	7,443	24.65%
5	2,405	7.97%	2,405	7.97%
6	1,334	4.42%	1,334	4.42%
Total	30,190	100.00%	30,190	100.00%

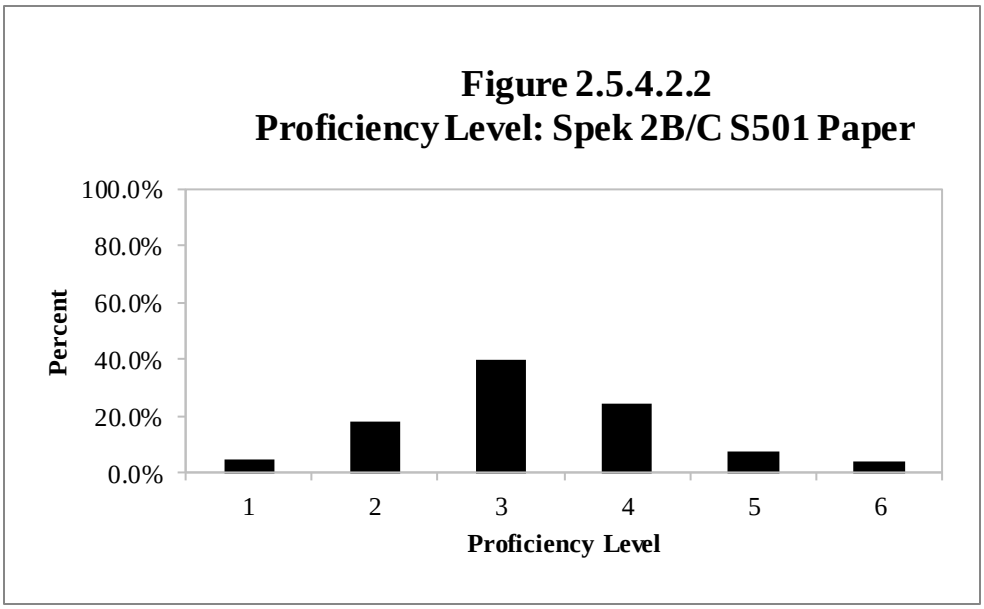
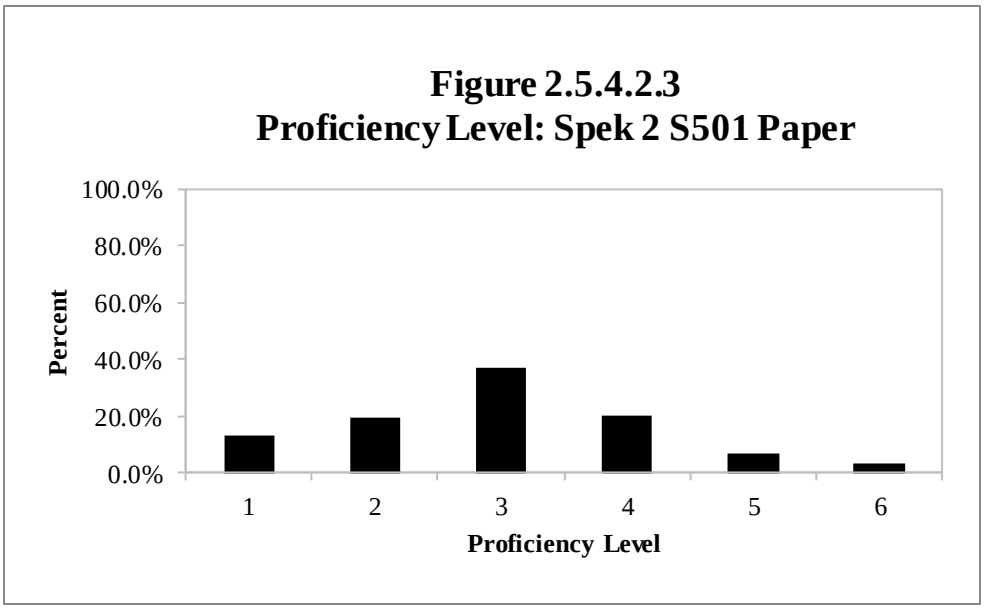


Table 2.5.4.2.3

Proficiency Level Distribution: Spek 2 S501 Paper

	Grade 2		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	5,140	13.15%	5,140	13.15%
2	7,515	19.23%	7,515	19.23%
3	14,481	37.05%	14,481	37.05%
4	7,986	20.43%	7,986	20.43%
5	2,628	6.72%	2,628	6.72%
6	1,334	3.41%	1,334	3.41%
Total	39,084	100.00%	39,084	100.00%



2.5.4.3 Grade 3

Table 2.5.4.3.1

Proficiency Level Distribution: Spek 3 A S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	3,474	47.69%	3,474	47.69%
2	1,807	24.81%	1,807	24.81%
3	1,308	17.96%	1,308	17.96%
4	695	9.54%	695	9.54%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	7,284	100.00%	7,284	100.00%

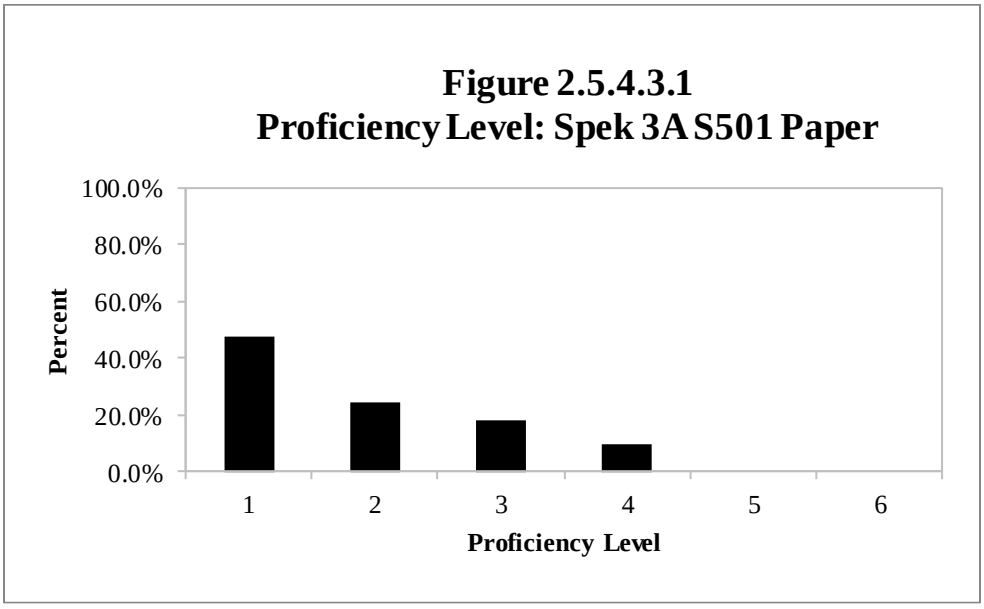


Table 2.5.4.3.2

Proficiency Level Distribution: Spek 3 B/C S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	1,191	4.08%	1,191	4.08%
2	5,346	18.32%	5,346	18.32%
3	12,313	42.20%	12,313	42.20%
4	7,234	24.80%	7,234	24.80%
5	1,483	5.08%	1,483	5.08%
6	1,608	5.51%	1,608	5.51%
Total	29,175	100.00%	29,175	100.00%

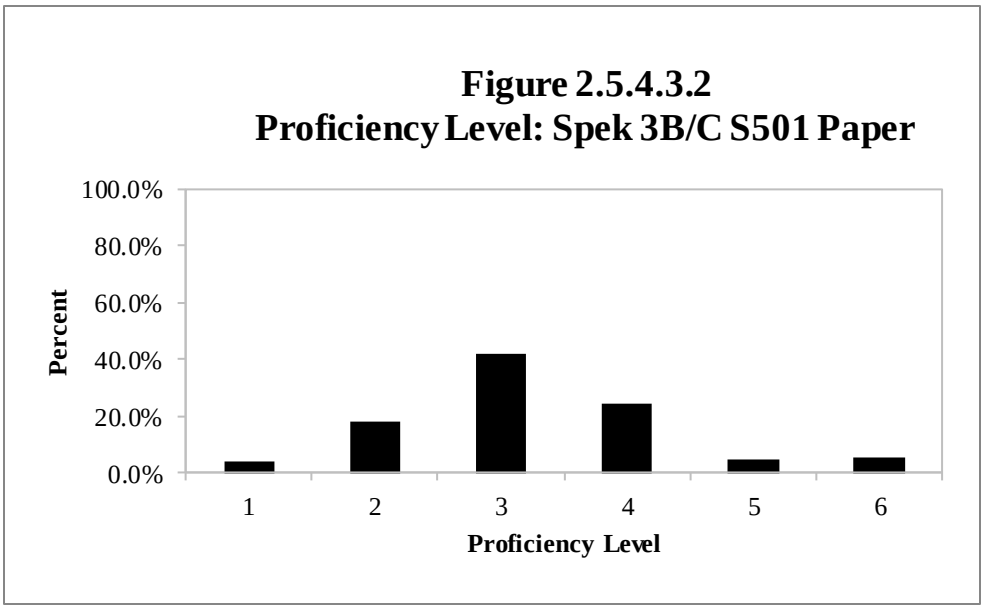
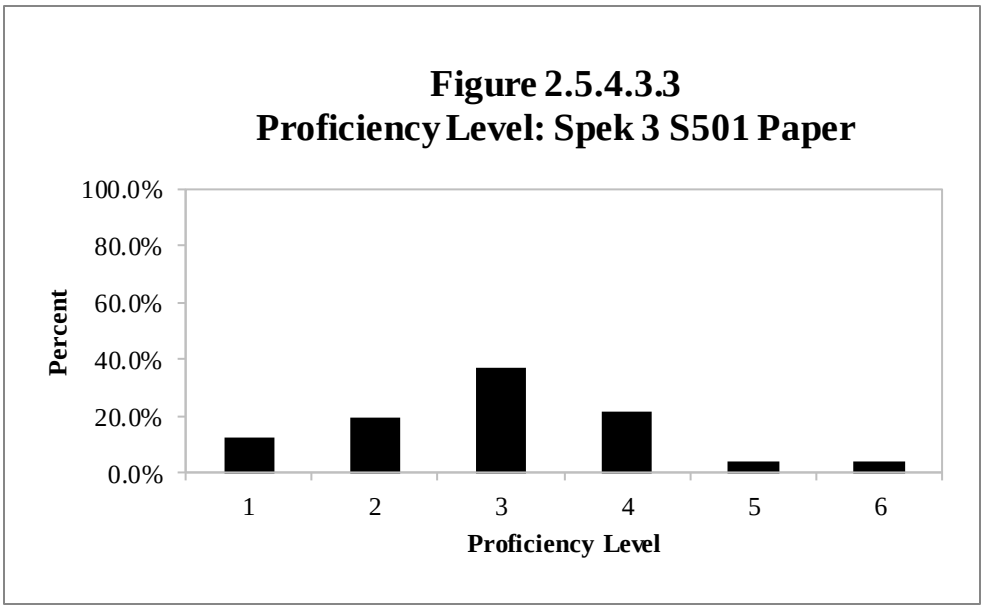


Table 2.5.4.3.3

Proficiency Level Distribution: Spek 3 S501 Paper

	Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	4,665	12.80%	4,665	12.80%
2	7,153	19.62%	7,153	19.62%
3	13,621	37.36%	13,621	37.36%
4	7,929	21.75%	7,929	21.75%
5	1,483	4.07%	1,483	4.07%
6	1,608	4.41%	1,608	4.41%
Total	36,459	100.00%	36,459	100.00%



2.5.4.4 Grades 4–5

Table 2.5.4.4.1

Proficiency Level Distribution: Spek 4-5 A S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,051	56.69%	3,143	64.26%	6,194	60.29%
2	1,012	18.80%	891	18.22%	1,903	18.52%
3	1,013	18.82%	558	11.41%	1,571	15.29%
4	238	4.42%	229	4.68%	467	4.55%
5	68	1.26%	70	1.43%	138	1.34%
6	0	0.00%	0	0.00%	0	0.00%
Total	5,382	100.00%	4,891	100.00%	10,273	100.00%

Figure 2.5.4.4.1
Proficiency Level: Spek 4-5A S501 Paper

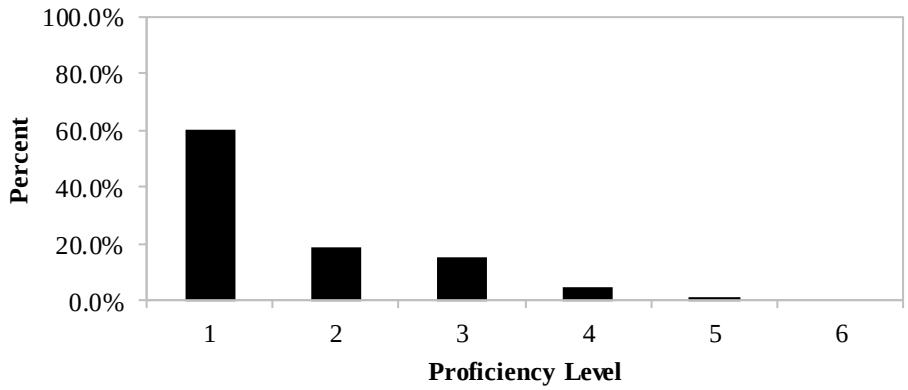


Table 2.5.4.4.2

Proficiency Level Distribution: Spek 4-5 B/C S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	430	1.97%	496	2.58%	926	2.25%
2	1,894	8.66%	2,178	11.34%	4,072	9.91%
3	6,237	28.51%	5,274	27.47%	11,511	28.03%
4	8,499	38.86%	7,409	38.59%	15,908	38.73%
5	3,613	16.52%	2,392	12.46%	6,005	14.62%
6	1,200	5.49%	1,449	7.55%	2,649	6.45%
Total	21,873	100.00%	19,198	100.00%	41,071	100.00%

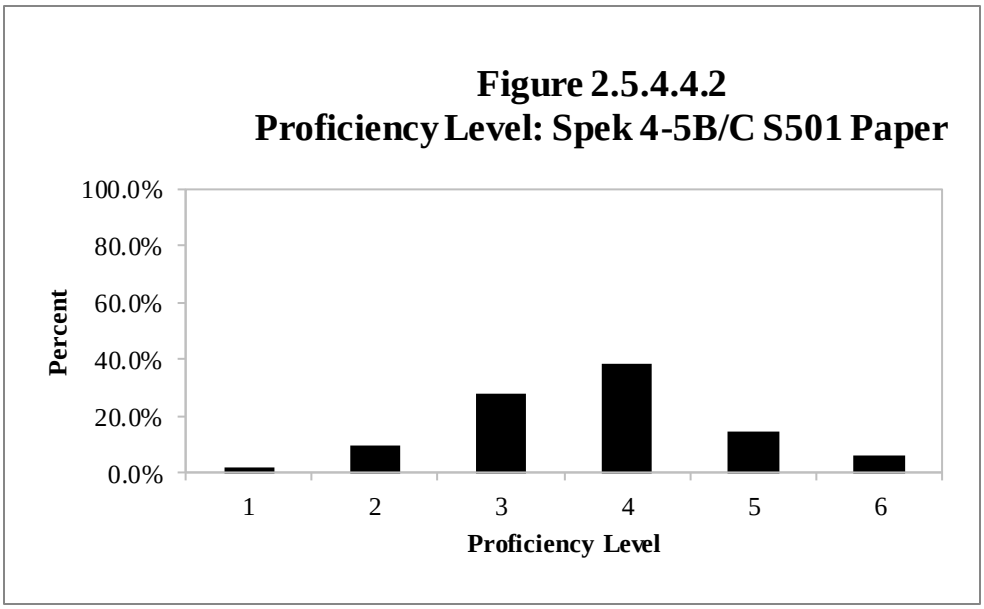
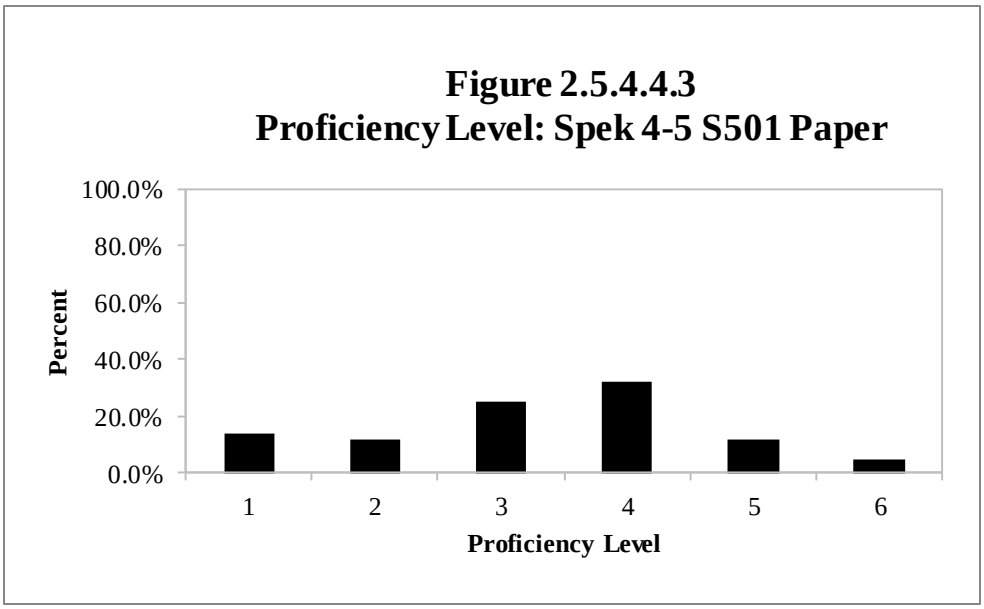


Table 2.5.4.4.3

Proficiency Level Distribution: Spek 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,481	12.77%	3,639	15.11%	7,120	13.87%
2	2,906	10.66%	3,069	12.74%	5,975	11.64%
3	7,250	26.60%	5,832	24.21%	13,082	25.48%
4	8,737	32.06%	7,638	31.71%	16,375	31.89%
5	3,681	13.51%	2,462	10.22%	6,143	11.96%
6	1,200	4.40%	1,449	6.02%	2,649	5.16%
Total	27,255	100.00%	24,089	100.00%	51,344	100.00%



2.5.4.5 Grades 6–8

Table 2.5.4.5.1

Proficiency Level Distribution: Spek 6-8 A S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,471	54.50%	2,772	63.94%	2,752	62.14%	7,995	60.12%
2	1,141	25.17%	702	16.19%	789	17.81%	2,632	19.79%
3	485	10.70%	466	10.75%	481	10.86%	1,432	10.77%
4	305	6.73%	334	7.70%	340	7.68%	979	7.36%
5	132	2.91%	61	1.41%	67	1.51%	260	1.96%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	4,534	100.00%	4,335	100.00%	4,429	100.00%	13,298	100.00%

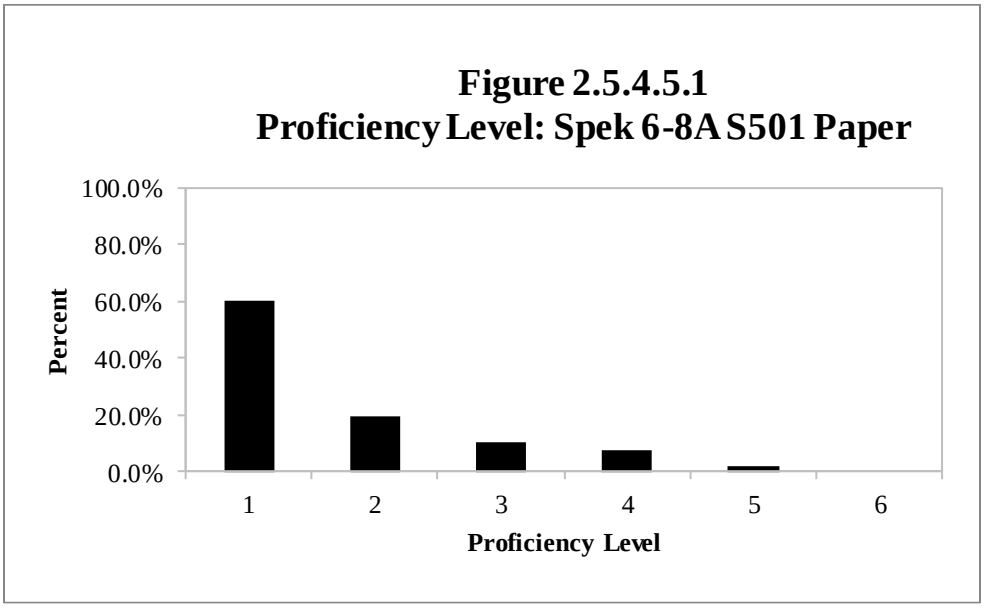


Table 2.5.4.5.2

Proficiency Level Distribution: Spek 6-8 B/C S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	313	2.03%	421	3.26%	599	4.96%	1,333	3.30%
2	1,155	7.49%	1,243	9.62%	884	7.32%	3,282	8.12%
3	3,459	22.42%	3,259	25.21%	3,641	30.17%	10,359	25.63%
4	6,946	45.03%	4,798	37.12%	4,394	36.41%	16,138	39.92%
5	1,771	11.48%	2,206	17.07%	1,497	12.40%	5,474	13.54%
6	1,782	11.55%	999	7.73%	1,054	8.73%	3,835	9.49%
Total	15,426	100.00%	12,926	100.00%	12,069	100.00%	40,421	100.00%

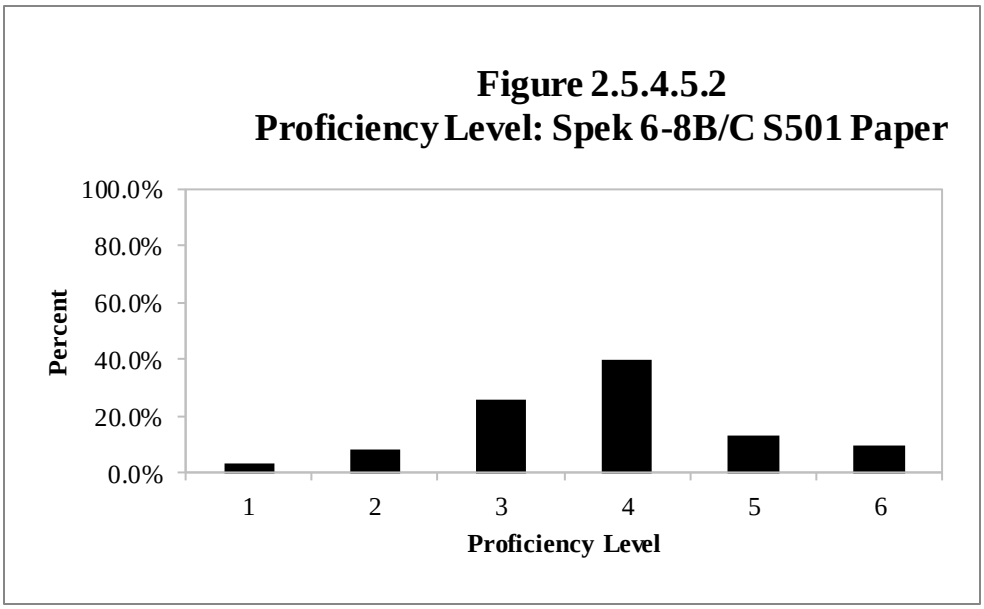
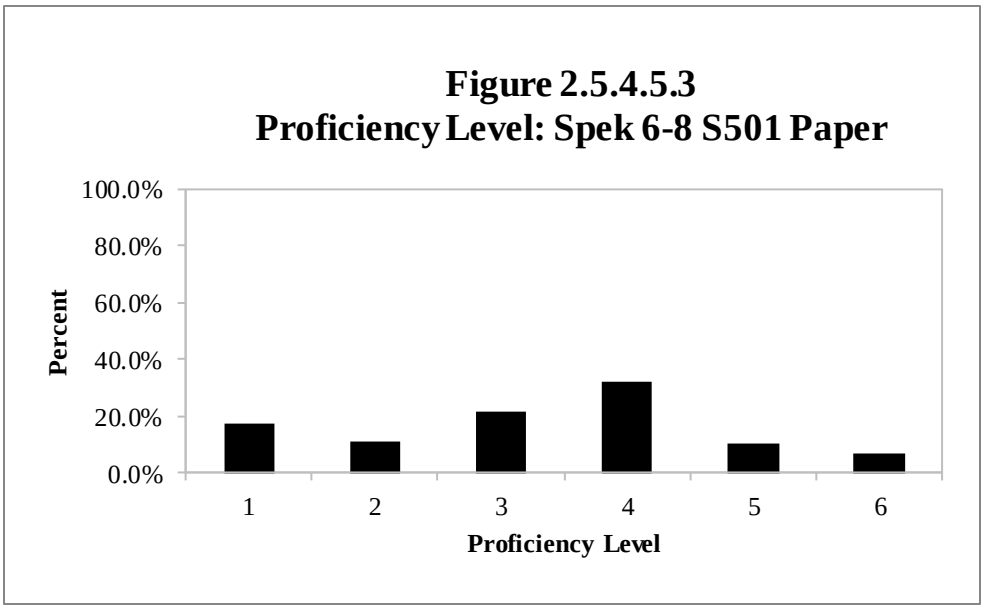


Table 2.5.4.5.3

Proficiency Level Distribution: Spek 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,784	13.95%	3,193	18.50%	3,351	20.31%	9,328	17.36%
2	2,296	11.50%	1,945	11.27%	1,673	10.14%	5,914	11.01%
3	3,944	19.76%	3,725	21.58%	4,122	24.98%	11,791	21.95%
4	7,251	36.33%	5,132	29.73%	4,734	28.69%	17,117	31.86%
5	1,903	9.53%	2,267	13.13%	1,564	9.48%	5,734	10.67%
6	1,782	8.93%	999	5.79%	1,054	6.39%	3,835	7.14%
Total	19,960	100.00%	17,261	100.00%	16,498	100.00%	53,719	100.00%



2.5.4.6 Grades 9–12

Table 2.5.4.6.1

Proficiency Level Distribution: Spek 9-12 A S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,849	74.05%	2,550	68.38%	1,766	62.40%	886	57.27%	9,051	68.03%
2	454	8.73%	342	9.17%	331	11.70%	333	21.53%	1,460	10.97%
3	650	12.50%	587	15.74%	543	19.19%	255	16.48%	2,035	15.30%
4	170	3.27%	250	6.70%	190	6.71%	73	4.72%	683	5.13%
5	75	1.44%	0	0.00%	0	0.00%	0	0.00%	75	0.56%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	5,198	100.00%	3,729	100.00%	2,830	100.00%	1,547	100.00%	13,304	100.00%

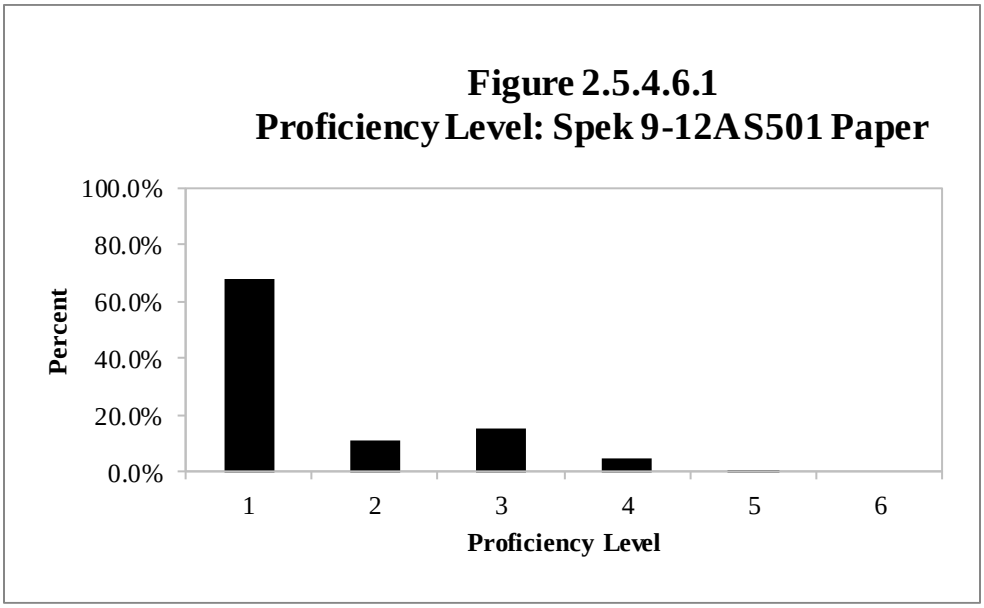


Table 2.5.4.6.2

Proficiency Level Distribution: Spek 9-12 B/C S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,057	9.58%	1,124	10.58%	1,150	11.98%	843	12.60%	4,174	11.00%
2	1,374	12.45%	1,448	13.63%	1,375	14.33%	943	14.10%	5,140	13.55%
3	3,500	31.72%	4,170	39.26%	3,266	34.04%	2,584	38.63%	13,520	35.64%
4	3,516	31.87%	2,358	22.20%	2,217	23.10%	1,162	17.37%	9,253	24.39%
5	550	4.98%	553	5.21%	569	5.93%	430	6.43%	2,102	5.54%
6	1,037	9.40%	968	9.11%	1,019	10.62%	727	10.87%	3,751	9.89%
Total	11,034	100.00%	10,621	100.00%	9,596	100.00%	6,689	100.00%	37,940	100.00%

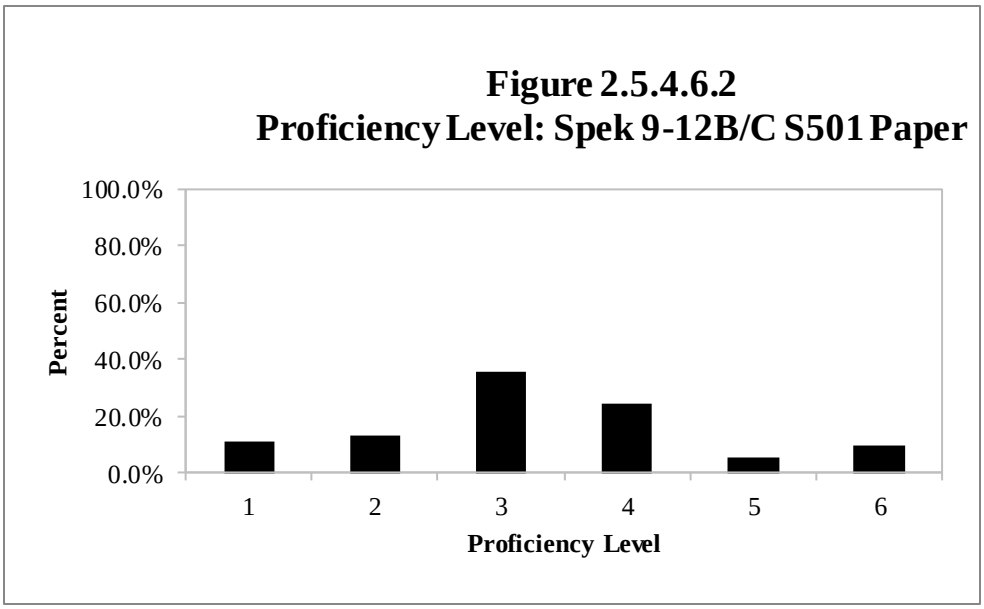
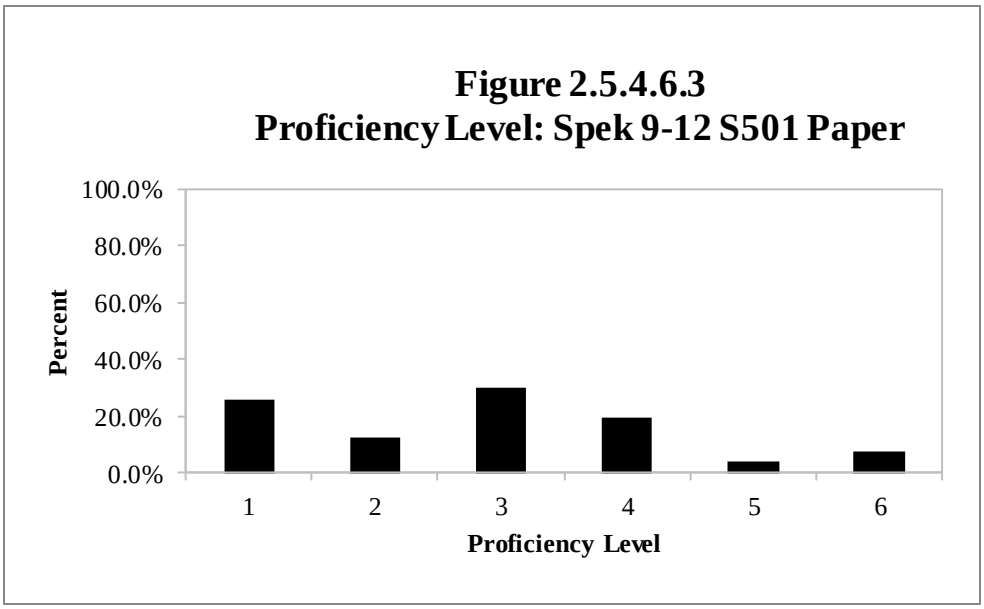


Table 2.5.4.6.3

Proficiency Level Distribution: Spek 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	4,906	30.22%	3,674	25.60%	2,916	23.47%	1,729	20.99%	13,225	25.81%
2	1,828	11.26%	1,790	12.47%	1,706	13.73%	1,276	15.49%	6,600	12.88%
3	4,150	25.57%	4,757	33.15%	3,809	30.65%	2,839	34.47%	15,555	30.35%
4	3,686	22.71%	2,608	18.17%	2,407	19.37%	1,235	15.00%	9,936	19.39%
5	625	3.85%	553	3.85%	569	4.58%	430	5.22%	2,177	4.25%
6	1,037	6.39%	968	6.75%	1,019	8.20%	727	8.83%	3,751	7.32%
Total	16,232	100.00%	14,350	100.00%	12,426	100.00%	8,236	100.00%	51,244	100.00%



2.6 Raw Score to Scale Score to Proficiency Level Conversion

This section presents raw score to scale score conversions and associated proficiency levels for each test form. 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.

Performances that gain very few score points, and performances from students who gain all or almost all of the score points, will have high CSEM values. The model does not precisely estimate these students' abilities; they may be well below or well above the range that is measured by the test, and therefore the error of measurement is large. We provide further detail on the CSEM as it relates to the interpretation of student performance in Section 5.3, which provides CSEM values for proficiency level cuts.

Note that we truncate raw scores of zero 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

2.6.1.0 Kindergarten

Table 2.6.1.0
Raw Score to Scale Score to Proficiency Level Conversion: List K S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for K
0	100	45	1.0
1	100	45	1.0
2	100	45	1.0
3	100	45	1.0
4	100	45	1.0
5	100	45	1.0
6	100	45	1.0
7	114	44	1.1
8	127	41	1.2
9	139	40	1.3
10	150	39	1.3
11	160	38	1.4
12	170	37	1.5
13	180	36	1.6
14	189	36	1.6
15	198	35	1.7
16	207	35	1.8
17	215	35	1.8
18	224	35	1.9
19	232	35	2.1
20	241	35	2.5
21	250	36	2.9
22	259	36	3.2
23	269	37	3.6
24	279	39	4.1
25	290	41	5.1
26	303	44	5.7
27	318	49	6.0
28	333	55	6.0
29	348	64	6.0
30	363	74	6.0

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

2.6.1.1 Grade 1

Table 2.6.1.1.1

Raw Score to Scale Score to Proficiency Level Conversion: List 1 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	104	93	1.0
1	121	77	1.1
2	150	56	1.3
3	169	48	1.4
4	184	43	1.6
5	197	41	1.7
6	208	39	1.7
7	218	38	1.8
8	228	37	1.9
9	238	37	2.0
10	247	37	2.4
11	257	38	2.9
12	268	39	3.2
13	279	41	3.6
14	292	44	4.0
15	307	49	5.1
16	322	55	5.7
17	337	63	6.0
18	352	73	6.0

Note: The test form is shared between 1A and 2A.

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

Table 2.6.1.1.2

Raw Score to Scale Score to Proficiency Level Conversion: List 1 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	104	129	1.0
1	147	77	1.3
2	177	57	1.5
3	197	48	1.7
4	211	43	1.8
5	224	40	1.9
6	235	38	1.9
7	245	37	2.3
8	254	36	2.7
9	263	35	3.1
10	272	35	3.4
11	280	35	3.6
12	289	35	3.9
13	298	36	4.5
14	307	37	5.1
15	317	38	5.5
16	328	41	6.0
17	341	44	6.0
18	356	49	6.0
19	371	55	6.0
20	386	64	6.0
21	401	74	6.0

Note: The test form is shared between 1B/C and 2B/C.

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

2.6.1.2 Grade 2

Table 2.6.1.2.1
Raw Score to Scale Score to Proficiency Level Conversion: List 2 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	112	84	1.0
1	121	77	1.0
2	150	56	1.3
3	169	48	1.4
4	184	43	1.5
5	197	41	1.6
6	208	39	1.7
7	218	38	1.8
8	228	37	1.8
9	238	37	1.9
10	247	37	2.0
11	257	38	2.3
12	268	39	2.6
13	279	41	2.8
14	292	44	3.2
15	307	49	3.7
16	322	55	4.5
17	337	63	5.2
18	352	73	5.9

Note: The test form is shared between 1A and 2A.

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

Table 2.6.1.2.2
Raw Score to Scale Score to Proficiency Level Conversion: List 2 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	112	117	1.0
1	147	77	1.2
2	177	57	1.5
3	197	48	1.6
4	211	43	1.7
5	224	40	1.8
6	235	38	1.9
7	245	37	2.0
8	254	36	2.2
9	263	35	2.4
10	272	35	2.7
11	280	35	2.9
12	289	35	3.1
13	298	36	3.4
14	307	37	3.7
15	317	38	4.1
16	328	41	4.8
17	341	44	5.4
18	356	49	6.0
19	371	55	6.0
20	386	64	6.0
21	401	74	6.0

Note: The test form is shared between 1B/C and 2B/C.

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

2.6.1.3 Grade 3

Table 2.6.1.3.1
Raw Score to Scale Score to Proficiency Level Conversion: List 3 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	112	181	1.0
1	184	79	1.4
2	216	58	1.6
3	236	49	1.8
4	251	44	1.9
5	264	41	2.0
6	275	39	2.3
7	286	37	2.6
8	295	37	2.8
9	305	36	3.1
10	314	36	3.4
11	324	37	3.7
12	334	38	4.1
13	344	40	4.7
14	356	43	5.2
15	371	47	5.8
16	386	54	6.0
17	401	63	6.0
18	416	74	6.0

Note: The test form is shared between 3A and 4-5A.

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

Table 2.6.1.3.2
Raw Score to Scale Score to Proficiency Level Conversion: List 3 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	112	215	1.0
1	198	80	1.5
2	231	60	1.7
3	253	51	1.9
4	269	46	2.1
5	283	43	2.5
6	296	41	2.8
7	307	39	3.2
8	318	38	3.5
9	328	37	3.9
10	337	36	4.3
11	347	36	4.8
12	356	36	5.2
13	365	37	5.6
14	375	37	6.0
15	385	39	6.0
16	397	40	6.0
17	409	43	6.0
18	424	48	6.0
19	439	55	6.0
20	454	63	6.0
21	469	74	6.0

Note: The test form is shared between 3B/C and 4-5B/C.

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

2.6.1.4 Grades 4–5

Table 2.6.1.4.1
Raw Score to Scale Score to Proficiency Level Conversion: List 4-5 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	120	164	1.0	1.0
1	184	79	1.4	1.3
2	216	58	1.6	1.5
3	236	49	1.7	1.7
4	251	44	1.8	1.7
5	264	41	1.9	1.8
6	275	39	2.0	1.9
7	286	37	2.2	2.0
8	295	37	2.5	2.2
9	305	36	2.7	2.5
10	314	36	3.0	2.7
11	324	37	3.3	3.0
12	334	38	3.7	3.3
13	344	40	4.0	3.6
14	356	43	4.6	4.0
15	371	47	5.3	4.8
16	386	54	5.9	5.4
17	401	63	6.0	6.0
18	416	74	6.0	6.0

Note: The test form is shared between 3A and 4-5A.

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

Table 2.6.1.4.2
Raw Score to Scale Score to Proficiency Level Conversion: List 4-5 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	120	195	1.0	1.0
1	198	80	1.5	1.4
2	231	60	1.7	1.6
3	253	51	1.8	1.8
4	269	46	1.9	1.9
5	283	43	2.2	1.9
6	296	41	2.5	2.2
7	307	39	2.8	2.5
8	318	38	3.1	2.8
9	328	37	3.5	3.1
10	337	36	3.8	3.4
11	347	36	4.2	3.7
12	356	36	4.6	4.0
13	365	37	5.0	4.5
14	375	37	5.4	5.0
15	385	39	5.8	5.3
16	397	40	6.0	5.8
17	409	43	6.0	6.0
18	424	48	6.0	6.0
19	439	55	6.0	6.0
20	454	63	6.0	6.0
21	469	74	6.0	6.0

Note: The test form is shared between 3B/C and 4-5B/C.

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

2.6.1.5 Grades 6–8

Table 2.6.1.5.1
Raw Score to Scale Score to Proficiency Level Conversion: List 6-8 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	132	137	1.0	1.0	1.0
1	181	78	1.3	1.3	1.2
2	212	59	1.5	1.4	1.4
3	233	51	1.6	1.6	1.5
4	249	46	1.7	1.6	1.6
5	263	43	1.8	1.7	1.7
6	276	41	1.8	1.8	1.8
7	287	40	1.9	1.9	1.8
8	298	39	2.1	1.9	1.9
9	309	38	2.3	2.1	2.0
10	319	38	2.6	2.4	2.2
11	329	39	2.9	2.7	2.5
12	340	39	3.2	3.0	2.8
13	352	41	3.6	3.4	3.1
14	364	44	4.0	3.8	3.5
15	379	48	4.7	4.3	4.0
16	394	55	5.3	5.0	4.6
17	409	63	5.9	5.5	5.2
18	424	73	6.0	6.0	5.8

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

Table 2.6.1.5.2
Raw Score to Scale Score to Proficiency Level Conversion: List 6-8 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	132	248	1.0	1.0	1.0
1	226	76	1.6	1.5	1.5
2	256	56	1.7	1.7	1.7
3	275	48	1.8	1.8	1.8
4	289	43	1.9	1.9	1.8
5	301	40	2.1	1.9	1.9
6	312	38	2.4	2.2	2.1
7	321	36	2.7	2.5	2.3
8	330	35	2.9	2.7	2.5
9	339	35	3.2	2.9	2.7
10	347	34	3.4	3.2	3.0
11	356	34	3.7	3.5	3.3
12	364	35	4.0	3.8	3.5
13	373	35	4.4	4.1	3.8
14	382	36	4.8	4.5	4.2
15	392	38	5.2	4.9	4.6
16	403	40	5.6	5.3	5.0
17	415	43	6.0	5.8	5.5
18	430	48	6.0	6.0	6.0
19	445	55	6.0	6.0	6.0
20	460	64	6.0	6.0	6.0
21	475	74	6.0	6.0	6.0

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

2.6.1.6 Grades 9–12

Table 2.6.1.6.1
Raw Score to Scale Score to Proficiency Level Conversion: List 9-12 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	148	98	1.0	1.0	1.0	1.0
1	169	79	1.1	1.1	1.1	1.1
2	201	60	1.3	1.3	1.2	1.2
3	223	52	1.4	1.4	1.4	1.3
4	240	47	1.5	1.5	1.5	1.4
5	255	44	1.6	1.6	1.5	1.5
6	269	42	1.7	1.6	1.6	1.6
7	281	41	1.8	1.7	1.7	1.6
8	293	40	1.8	1.8	1.7	1.7
9	304	40	1.9	1.8	1.8	1.8
10	315	40	2.0	1.9	1.8	1.8
11	327	40	2.3	2.0	1.9	1.9
12	338	41	2.6	2.3	2.1	1.9
13	350	43	2.9	2.7	2.5	2.3
14	364	45	3.3	3.1	3.0	2.8
15	380	50	3.9	3.7	3.5	3.4
16	396	56	4.5	4.2	4.0	3.9
17	412	64	5.1	4.8	4.6	4.5
18	428	76	5.7	5.5	5.2	5.0

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

Table 2.6.1.6.2
Raw Score to Scale Score to Proficiency Level Conversion: List 9-12 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	148	230	1.0	1.0	1.0	1.0
1	238	78	1.5	1.5	1.4	1.4
2	269	58	1.7	1.6	1.6	1.6
3	290	50	1.8	1.8	1.7	1.7
4	306	45	1.9	1.8	1.8	1.8
5	319	42	2.1	1.9	1.9	1.8
6	331	39	2.4	2.1	1.9	1.9
7	341	38	2.6	2.4	2.2	1.9
8	351	37	2.9	2.7	2.5	2.3
9	361	36	3.2	3.0	2.8	2.7
10	370	36	3.5	3.3	3.2	3.0
11	379	35	3.8	3.6	3.5	3.3
12	388	36	4.1	3.9	3.8	3.6
13	397	36	4.5	4.3	4.1	3.9
14	406	37	4.8	4.6	4.4	4.2
15	416	38	5.2	5.0	4.8	4.6
16	427	40	5.7	5.4	5.2	5.0
17	439	43	6.0	5.9	5.7	5.5
18	454	48	6.0	6.0	6.0	6.0
19	469	55	6.0	6.0	6.0	6.0
20	484	64	6.0	6.0	6.0	6.0
21	499	75	6.0	6.0	6.0	6.0

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

2.6.2 Reading

2.6.2.0 Kindergarten

Table 2.6.2.0
Raw Score to Scale Score to Proficiency Level Conversion: Read K S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for K
0	100	33	1.0
1	100	33	1.0
2	100	33	1.0
3	100	33	1.0
4	100	33	1.0
5	100	33	1.0
6	100	33	1.0
7	100	33	1.0
8	100	33	1.0
9	100	33	1.0
10	100	33	1.0
11	109	34	1.0
12	120	33	1.1
13	132	33	1.2
14	142	32	1.2
15	152	31	1.3
16	162	30	1.4
17	171	29	1.5
18	180	29	1.5
19	188	29	1.6
20	196	29	1.6
21	205	29	1.7
22	213	29	1.8
23	222	29	1.8
24	230	30	1.9
25	240	31	1.9
26	250	32	2.5
27	260	35	3.0
28	270	38	3.5
29	280	43	4.1
30	290	49	5.0

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

2.6.2.1 Grade 1

Table 2.6.2.1.1

Raw Score to Scale Score to Proficiency Level Conversion: Read 1 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	141	105	1.0
1	180	52	1.3
2	200	38	1.4
3	213	32	1.5
4	222	29	1.6
5	230	27	1.7
6	237	25	1.7
7	243	24	1.8
8	249	23	1.8
9	255	23	1.9
10	260	22	1.9
11	265	22	2.0
12	270	22	2.2
13	275	22	2.5
14	280	22	2.7
15	285	22	2.9
16	291	23	3.2
17	296	24	3.5
18	302	25	3.8
19	309	26	4.4
20	317	28	5.1
21	326	32	5.5
22	335	36	6.0
23	344	41	6.0
24	353	48	6.0

Note: The test form is shared between 1A and 2A.

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

Table 2.6.2.1.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 1 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	141	203	1.0
1	216	53	1.6
2	236	38	1.7
3	248	32	1.8
4	258	28	1.9
5	265	26	2.0
6	272	24	2.3
7	278	23	2.6
8	283	22	2.8
9	288	21	3.1
10	293	21	3.3
11	297	20	3.6
12	301	20	3.8
13	306	20	4.1
14	310	20	4.5
15	314	20	4.9
16	318	20	5.1
17	323	21	5.4
18	327	21	5.6
19	332	22	5.8
20	337	23	6.0
21	343	24	6.0
22	349	25	6.0
23	356	28	6.0
24	365	31	6.0
25	374	36	6.0
26	383	41	6.0
27	392	48	6.0

Note: The test form is shared between 1B/C and 2B/C.

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

2.6.2.2 Grade 2

Table 2.6.2.2.1

Raw Score to Scale Score to Proficiency Level Conversion: Read 2 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	158	77	1.0
1	180	52	1.2
2	200	38	1.3
3	213	32	1.4
4	222	29	1.5
5	230	27	1.6
6	237	25	1.6
7	243	24	1.6
8	249	23	1.7
9	255	23	1.7
10	260	22	1.8
11	265	22	1.8
12	270	22	1.9
13	275	22	1.9
14	280	22	1.9
15	285	22	2.0
16	291	23	2.3
17	296	24	2.5
18	302	25	2.7
19	309	26	3.1
20	317	28	3.5
21	326	32	4.0
22	335	36	4.8
23	344	41	5.3
24	353	48	5.8

Note: The test form is shared between 1A and 2A.

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

Table 2.6.2.2.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 2 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	158	149	1.0
1	216	53	1.4
2	236	38	1.6
3	248	32	1.7
4	258	28	1.8
5	265	26	1.8
6	272	24	1.9
7	278	23	1.9
8	283	22	2.0
9	288	21	2.2
10	293	21	2.4
11	297	20	2.5
12	301	20	2.7
13	306	20	2.9
14	310	20	3.1
15	314	20	3.3
16	318	20	3.5
17	323	21	3.8
18	327	21	4.0
19	332	22	4.5
20	337	23	5.0
21	343	24	5.3
22	349	25	5.6
23	356	28	6.0
24	365	31	6.0
25	374	36	6.0
26	383	41	6.0
27	392	48	6.0

Note: The test form is shared between 1B/C and 2B/C.

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

2.6.2.3 Grade 3

Table 2.6.2.3.1

Raw Score to Scale Score to Proficiency Level Conversion: Read 3 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	158	144	1.0
1	213	52	1.3
2	234	38	1.5
3	246	32	1.6
4	256	29	1.7
5	264	27	1.7
6	271	25	1.8
7	277	24	1.8
8	283	23	1.8
9	289	23	1.9
10	294	22	1.9
11	299	22	2.0
12	305	22	2.3
13	310	22	2.5
14	315	22	2.6
15	320	23	2.8
16	326	23	3.1
17	332	24	3.4
18	338	25	3.7
19	345	27	4.3
20	353	29	5.0
21	362	32	5.5
22	371	36	6.0
23	380	42	6.0
24	389	48	6.0

Note: The test form is shared between 3B/C and 4-5B/C.

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

Table 2.6.2.3.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 3 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	158	438	1.0
1	271	52	1.8
2	291	38	1.9
3	303	32	2.2
4	313	28	2.6
5	320	26	2.8
6	327	24	3.2
7	332	23	3.4
8	338	22	3.7
9	342	21	4.0
10	347	21	4.5
11	352	20	5.0
12	356	20	5.2
13	360	20	5.4
14	364	20	5.6
15	369	20	5.9
16	373	20	6.0
17	377	21	6.0
18	382	21	6.0
19	387	22	6.0
20	392	23	6.0
21	397	24	6.0
22	404	25	6.0
23	411	28	6.0
24	420	31	6.0
25	429	36	6.0
26	438	42	6.0
27	447	48	6.0

Note: The test form is shared between 3B/C and 4-5B/C.

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

2.6.2.4 Grades 4–5

Table 2.6.2.4.1

Raw Score to Scale Score to Proficiency Level Conversion: Read 4-5 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	175	104	1.0	1.0
1	213	52	1.3	1.2
2	234	38	1.4	1.4
3	246	32	1.5	1.5
4	256	29	1.6	1.5
5	264	27	1.6	1.6
6	271	25	1.7	1.6
7	277	24	1.7	1.7
8	283	23	1.8	1.7
9	289	23	1.8	1.8
10	294	22	1.9	1.8
11	299	22	1.9	1.8
12	305	22	1.9	1.9
13	310	22	2.1	1.9
14	315	22	2.2	1.9
15	320	23	2.4	2.1
16	326	23	2.6	2.3
17	332	24	2.8	2.5
18	338	25	3.1	2.7
19	345	27	3.5	3.0
20	353	29	3.9	3.4
21	362	32	4.8	3.8
22	371	36	5.3	4.7
23	380	42	5.8	5.3
24	389	48	6.0	5.8

Note: The test form is shared between 3B/C and 4-5B/C.

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

Table 2.6.2.4.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 4-5 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	175	315	1.0	1.0
1	271	52	1.7	1.6
2	291	38	1.8	1.8
3	303	32	1.9	1.9
4	313	28	2.2	1.9
5	320	26	2.4	2.1
6	327	24	2.7	2.3
7	332	23	2.8	2.5
8	338	22	3.1	2.7
9	342	21	3.3	2.8
10	347	21	3.6	3.1
11	352	20	3.8	3.3
12	356	20	4.2	3.5
13	360	20	4.6	3.7
14	364	20	5.0	4.0
15	369	20	5.2	4.5
16	373	20	5.5	5.0
17	377	21	5.7	5.2
18	382	21	6.0	5.5
19	387	22	6.0	5.7
20	392	23	6.0	6.0
21	397	24	6.0	6.0
22	404	25	6.0	6.0
23	411	28	6.0	6.0
24	420	31	6.0	6.0
25	429	36	6.0	6.0
26	438	42	6.0	6.0
27	447	48	6.0	6.0

Note: The test form is shared between 3B/C and 4-5B/C.

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

2.6.2.5 Grades 6–8

Table 2.6.2.5.1
Raw Score to Scale Score to Proficiency Level Conversion: Read 6-8 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	200	99	1.1	1.0	1.0
1	235	53	1.3	1.3	1.2
2	256	39	1.5	1.4	1.4
3	269	33	1.6	1.5	1.5
4	279	29	1.6	1.6	1.5
5	287	27	1.7	1.6	1.6
6	294	25	1.7	1.7	1.6
7	301	24	1.8	1.7	1.7
8	307	23	1.8	1.8	1.7
9	312	23	1.9	1.8	1.8
10	318	22	1.9	1.9	1.8
11	323	22	2.0	1.9	1.9
12	328	22	2.1	1.9	1.9
13	333	22	2.3	2.1	1.9
14	338	22	2.5	2.2	2.0
15	344	23	2.7	2.4	2.2
16	349	23	2.8	2.6	2.4
17	355	24	3.1	2.8	2.6
18	361	25	3.4	3.0	2.8
19	368	27	3.7	3.4	3.1
20	376	29	4.3	3.8	3.5
21	386	32	5.2	4.6	4.0
22	396	37	5.8	5.4	5.0
23	406	43	6.0	6.0	5.6
24	416	51	6.0	6.0	6.0

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

Table 2.6.2.5.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 6-8 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	200	214	1.1	1.0	1.0
1	276	52	1.6	1.6	1.5
2	296	38	1.8	1.7	1.7
3	308	32	1.8	1.8	1.8
4	317	28	1.9	1.9	1.8
5	324	26	2.0	1.9	1.9
6	331	24	2.2	2.0	1.9
7	337	23	2.4	2.2	2.0
8	342	22	2.6	2.4	2.2
9	347	21	2.8	2.5	2.3
10	351	21	2.9	2.7	2.5
11	356	21	3.1	2.8	2.6
12	360	20	3.3	3.0	2.8
13	365	20	3.6	3.2	2.9
14	369	20	3.8	3.4	3.1
15	373	20	4.0	3.6	3.3
16	378	20	4.5	3.9	3.6
17	382	21	5.0	4.2	3.8
18	387	21	5.2	4.7	4.1
19	392	22	5.5	5.1	4.6
20	397	23	5.8	5.4	5.1
21	403	24	6.0	5.8	5.4
22	409	26	6.0	6.0	5.8
23	416	28	6.0	6.0	6.0
24	425	31	6.0	6.0	6.0
25	434	36	6.0	6.0	6.0
26	443	41	6.0	6.0	6.0
27	452	48	6.0	6.0	6.0

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

2.6.2.6 Grades 9–12

Table 2.6.2.6.1

Raw Score to Scale Score to Proficiency Level Conversion: Read 9-12 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	233	69	1.1	1.1	1.0	1.0
1	249	53	1.3	1.2	1.2	1.1
2	270	39	1.4	1.4	1.3	1.3
3	283	33	1.5	1.5	1.4	1.4
4	293	30	1.6	1.6	1.5	1.5
5	302	28	1.7	1.6	1.6	1.5
6	310	26	1.7	1.7	1.6	1.6
7	316	25	1.8	1.7	1.7	1.6
8	323	24	1.8	1.8	1.7	1.7
9	329	24	1.9	1.8	1.8	1.8
10	334	23	1.9	1.9	1.8	1.8
11	340	23	2.0	1.9	1.9	1.8
12	345	22	2.1	2.0	1.9	1.9
13	350	22	2.3	2.1	2.0	1.9
14	356	23	2.5	2.3	2.2	2.1
15	361	23	2.6	2.5	2.3	2.2
16	366	23	2.8	2.6	2.5	2.4
17	372	24	3.0	2.8	2.7	2.5
18	379	25	3.3	3.1	2.9	2.7
19	385	26	3.6	3.4	3.1	2.9
20	393	29	4.1	3.8	3.5	3.3
21	403	32	5.1	4.6	4.1	3.8
22	413	37	5.7	5.4	5.1	4.8
23	423	43	6.0	6.0	5.7	5.5
24	433	51	6.0	6.0	6.0	6.0

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

Table 2.6.2.6.2

Raw Score to Scale Score to Proficiency Level Conversion: Read 9-12 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	233	171	1.1	1.1	1.0	1.0
1	297	52	1.6	1.6	1.5	1.5
2	316	37	1.8	1.7	1.7	1.6
3	328	31	1.9	1.8	1.8	1.7
4	337	28	1.9	1.9	1.9	1.8
5	344	25	2.1	2.0	1.9	1.9
6	351	24	2.3	2.2	2.0	1.9
7	356	23	2.5	2.3	2.2	2.1
8	361	22	2.6	2.5	2.3	2.2
9	366	21	2.8	2.6	2.5	2.4
10	371	21	2.9	2.8	2.6	2.5
11	375	20	3.1	2.9	2.7	2.6
12	379	20	3.3	3.1	2.9	2.7
13	383	20	3.5	3.3	3.0	2.9
14	388	20	3.8	3.5	3.3	3.0
15	392	20	4.0	3.7	3.5	3.2
16	396	20	4.4	3.9	3.7	3.4
17	400	21	4.8	4.3	3.9	3.6
18	405	21	5.2	4.8	4.3	3.9
19	410	22	5.5	5.2	5.0	4.4
20	415	23	5.8	5.5	5.2	5.0
21	420	24	6.0	5.8	5.5	5.3
22	427	25	6.0	6.0	6.0	5.7
23	434	28	6.0	6.0	6.0	6.0
24	443	31	6.0	6.0	6.0	6.0
25	452	36	6.0	6.0	6.0	6.0
26	461	42	6.0	6.0	6.0	6.0
27	470	48	6.0	6.0	6.0	6.0

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

2.6.3 Writing

2.6.3.0 Kindergarten

Table 2.6.3.0
Raw Score to Scale Score to Proficiency Level Conversion: Writ K S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for K
0	100	107	1.0
1	100	107	1.0
2	100	107	1.0
3	100	107	1.0
4	155	60	1.4
5	177	44	1.5
6	191	37	1.6
7	202	35	1.7
8	213	34	1.8
9	223	35	1.9
10	234	37	2.0
11	246	37	2.3
12	258	39	2.6
13	271	41	3.0
14	288	48	3.4
15	305	57	3.8
16	322	65	4.1
17	339	71	4.5

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

2.6.3.1 Grade 1

Table 2.6.3.1.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 1 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1	Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	111	97	1.0	34	381	25	4.9
1	148	63	1.2	35	387	26	5.2
2	177	46	1.5	36	395	28	5.5
3	193	34	1.6	37	403	31	5.9
4	202	28	1.7	38	414	37	6.0
5	209	24	1.7	39	433	51	6.0
6	214	22	1.8	40	464	95	6.0
7	219	20	1.8				
8	223	20	1.8				
9	227	19	1.9				
10	231	19	1.9				
11	234	19	1.9				
12	238	20	2.0				
13	242	20	2.1				
14	247	21	2.2				
15	252	23	2.3				
16	257	24	2.5				
17	263	25	2.6				
18	270	27	2.8				
19	277	27	3.0				
20	285	28	3.1				
21	293	27	3.2				
22	300	27	3.4				
23	307	26	3.5				
24	314	26	3.6				
25	321	26	3.7				
26	328	26	3.8				
27	334	26	3.9				
28	341	26	4.0				
29	348	26	4.2				
30	355	26	4.4				
31	362	25	4.5				
32	368	25	4.6				
33	374	25	4.8				

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 S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1	Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	111	259	1.0	34	342	24	4.1
1	191	47	1.6	35	348	23	4.2
2	206	32	1.7	36	354	23	4.3
3	214	26	1.8	37	359	23	4.4
4	220	22	1.8	38	364	22	4.6
5	224	20	1.8	39	369	22	4.7
6	228	19	1.9	40	373	21	4.8
7	231	18	1.9	41	378	21	4.9
8	234	17	1.9	42	382	21	5.0
9	237	16	1.9	43	387	21	5.2
10	240	16	2.0	44	391	21	5.3
11	242	16	2.1	45	395	21	5.5
12	245	16	2.1	46	400	21	5.7
13	248	16	2.2	47	405	22	6.0
14	250	16	2.3	48	410	23	6.0
15	253	16	2.4	49	415	24	6.0
16	256	16	2.4	50	422	26	6.0
17	258	17	2.5	51	429	30	6.0
18	261	18	2.6	52	440	36	6.0
19	265	18	2.7	53	458	51	6.0
20	268	19	2.8	54	490	95	6.0
21	272	20	2.9				
22	276	21	3.0				
23	281	21	3.0				
24	285	22	3.1				
25	291	23	3.2				
26	296	23	3.3				
27	301	24	3.4				
28	307	24	3.5				
29	313	24	3.6				
30	319	24	3.7				
31	325	24	3.8				
32	331	24	3.9				
33	337	24	4.0				

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

2.6.3.2 Grade 2

Table 2.6.3.2.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 2 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	133	213	1.0
1	199	45	1.6
2	213	32	1.7
3	222	26	1.8
4	228	24	1.8
5	234	23	1.9
6	239	22	1.9
7	244	23	2.0
8	249	24	2.1
9	255	25	2.3
10	262	27	2.5
11	270	29	2.7
12	279	31	3.0
13	290	33	3.1
14	301	34	3.3
15	313	35	3.5
16	325	34	3.7
17	336	34	3.9
18	347	33	4.1
19	358	31	4.3
20	367	30	4.5
21	376	30	4.7
22	385	29	4.9
23	394	30	5.2
24	403	32	5.6
25	415	38	6.0
26	434	51	6.0
27	465	94	6.0

Note: The test form is shared between 2A and 3A

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 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2	Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	133	159	1.0	34	340	24	3.9
1	188	47	1.5	35	345	23	4.0
2	202	32	1.6	36	351	23	4.2
3	210	26	1.7	37	356	23	4.3
4	216	22	1.7	38	361	22	4.4
5	221	20	1.8	39	366	22	4.5
6	225	19	1.8	40	371	21	4.6
7	228	18	1.8	41	375	21	4.7
8	231	17	1.9	42	380	21	4.8
9	234	17	1.9	43	384	21	4.9
10	237	16	1.9	44	388	21	5.0
11	240	16	1.9	45	393	21	5.2
12	242	16	2.0	46	397	21	5.3
13	245	16	2.0	47	402	22	5.6
14	248	16	2.1	48	407	23	5.8
15	250	16	2.2	49	413	24	6.0
16	253	17	2.2	50	420	27	6.0
17	256	17	2.3	51	428	30	6.0
18	259	18	2.4	52	438	36	6.0
19	263	18	2.5	53	456	51	6.0
20	266	19	2.6	54	488	95	6.0
21	270	20	2.7				
22	274	20	2.8				
23	278	21	2.9				
24	283	22	3.0				
25	288	23	3.1				
26	293	23	3.2				
27	299	24	3.3				
28	305	24	3.4				
29	310	24	3.5				
30	316	24	3.5				
31	322	24	3.6				
32	328	24	3.7				
33	334	24	3.8				

Note: The test form is shared between 2BC and 3BC.

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

2.6.3.3 Grade 3

Table 2.6.3.3.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 3 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	133	213	1.0
1	199	45	1.5
2	213	32	1.7
3	222	26	1.7
4	228	24	1.8
5	234	23	1.8
6	239	22	1.9
7	244	23	1.9
8	249	24	2.0
9	255	25	2.2
10	262	27	2.4
11	270	29	2.6
12	279	31	2.8
13	290	33	3.1
14	301	34	3.2
15	313	35	3.4
16	325	34	3.6
17	336	34	3.8
18	347	33	4.0
19	358	31	4.2
20	367	30	4.4
21	376	30	4.6
22	385	29	4.8
23	394	30	5.0
24	403	32	5.3
25	415	38	5.8
26	434	51	6.0
27	465	94	6.0

Note: The test form is shared between 2A and 3A.

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 3 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3	Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	133	159	1.0	34	340	24	3.9
1	188	47	1.4	35	345	23	3.9
2	202	32	1.6	36	351	23	4.1
3	210	26	1.6	37	356	23	4.2
4	216	22	1.7	38	361	22	4.3
5	221	20	1.7	39	366	22	4.4
6	225	19	1.8	40	371	21	4.5
7	228	18	1.8	41	375	21	4.6
8	231	17	1.8	42	380	21	4.7
9	234	17	1.8	43	384	21	4.7
10	237	16	1.9	44	388	21	4.8
11	240	16	1.9	45	393	21	4.9
12	242	16	1.9	46	397	21	5.1
13	245	16	1.9	47	402	22	5.3
14	248	16	2.0	48	407	23	5.5
15	250	16	2.0	49	413	24	5.7
16	253	17	2.1	50	420	27	6.0
17	256	17	2.2	51	428	30	6.0
18	259	18	2.3	52	438	36	6.0
19	263	18	2.4	53	456	51	6.0
20	266	19	2.5	54	488	95	6.0
21	270	20	2.6				
22	274	20	2.7				
23	278	21	2.8				
24	283	22	3.0				
25	288	23	3.0				
26	293	23	3.1				
27	299	24	3.2				
28	305	24	3.3				
29	310	24	3.4				
30	316	24	3.5				
31	322	24	3.6				
32	328	24	3.7				
33	334	24	3.8				

Note: The test form is shared between 2BC and 3BC.

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

2.6.3.4 Grades 4–5

Table 2.6.3.4.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 4-5 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	155	253	1.0	1.0
1	231	45	1.7	1.6
2	245	32	1.8	1.8
3	253	26	1.8	1.8
4	259	24	1.9	1.9
5	264	23	1.9	1.9
6	270	22	2.1	2.1
7	275	23	2.4	2.3
8	280	23	2.6	2.5
9	286	25	2.9	2.7
10	293	27	3.0	3.0
11	301	29	3.2	3.1
12	310	31	3.3	3.2
13	320	33	3.5	3.4
14	332	34	3.6	3.6
15	343	35	3.8	3.7
16	355	34	4.0	3.9
17	367	34	4.3	4.2
18	378	33	4.5	4.4
19	388	31	4.7	4.6
20	398	30	4.9	4.8
21	407	29	5.2	5.0
22	416	29	5.6	5.3
23	424	30	5.9	5.6
24	434	32	6.0	6.0
25	446	37	6.0	6.0
26	464	51	6.0	6.0
27	496	94	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 4-5 B/C S501 Paper

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	229	1.0	1.0	34	381	24	4.6	4.4
1	228	47	1.6	1.6	35	387	23	4.7	4.6
2	243	32	1.8	1.7	36	392	23	4.8	4.7
3	251	26	1.8	1.8	37	397	22	4.9	4.8
4	257	22	1.9	1.9	38	403	22	5.0	4.9
5	261	20	1.9	1.9	39	407	22	5.2	5.0
6	265	19	1.9	1.9	40	412	21	5.4	5.1
7	269	18	2.1	2.0	41	417	21	5.6	5.3
8	272	17	2.2	2.1	42	421	21	5.8	5.5
9	275	17	2.4	2.3	43	426	21	6.0	5.7
10	278	17	2.5	2.4	44	430	21	6.0	5.8
11	281	17	2.6	2.5	45	435	21	6.0	6.0
12	284	17	2.8	2.6	46	439	22	6.0	6.0
13	286	17	2.9	2.7	47	444	22	6.0	6.0
14	289	17	3.0	2.8	48	450	23	6.0	6.0
15	292	17	3.0	2.9	49	455	25	6.0	6.0
16	295	17	3.1	3.0	50	462	27	6.0	6.0
17	298	17	3.1	3.0	51	470	30	6.0	6.0
18	302	18	3.2	3.1	52	481	36	6.0	6.0
19	305	18	3.2	3.1	53	499	51	6.0	6.0
20	308	19	3.3	3.2	54	531	95	6.0	6.0
21	312	19	3.3	3.3					
22	316	20	3.4	3.3					
23	321	21	3.5	3.4					
24	325	22	3.5	3.5					
25	330	22	3.6	3.5					
26	335	23	3.7	3.6					
27	341	23	3.8	3.7					
28	346	24	3.9	3.8					
29	352	24	4.0	3.9					
30	358	24	4.1	4.0					
31	364	24	4.2	4.1					
32	370	24	4.3	4.2					
33	376	24	4.5	4.3					

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

2.6.3.5 Grades 6–8

Table 2.6.3.5.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 6-8 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	188	103	1.2	1.1	1.0
1	220	45	1.5	1.4	1.3
2	234	32	1.6	1.5	1.4
3	243	27	1.7	1.6	1.5
4	249	24	1.8	1.7	1.6
5	255	23	1.8	1.8	1.7
6	260	23	1.9	1.8	1.7
7	266	23	1.9	1.9	1.8
8	271	24	2.1	1.9	1.8
9	277	25	2.3	2.1	1.9
10	284	27	2.5	2.3	2.1
11	292	29	2.8	2.5	2.3
12	301	31	3.0	2.8	2.6
13	312	33	3.2	3.1	3.0
14	323	34	3.3	3.2	3.1
15	334	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	30	5.1	4.9	4.8
24	425	33	5.4	5.1	5.0
25	438	38	5.8	5.6	5.4
26	457	52	6.0	6.0	5.9
27	488	94	6.0	6.0	6.0

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

Table 2.6.3.5.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 6-8 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	188	96	1.2	1.1	1.0
1	218	47	1.5	1.4	1.3
2	233	32	1.6	1.5	1.4
3	241	26	1.7	1.6	1.5
4	247	22	1.7	1.7	1.6
5	251	20	1.8	1.7	1.6
6	255	18	1.8	1.8	1.7
7	258	17	1.9	1.8	1.7
8	261	17	1.9	1.8	1.7
9	264	16	1.9	1.9	1.8
10	267	16	1.9	1.9	1.8
11	269	16	2.0	1.9	1.8
12	272	16	2.1	1.9	1.9
13	274	16	2.2	2.0	1.9
14	277	16	2.3	2.1	1.9
15	279	16	2.3	2.1	1.9
16	282	16	2.4	2.2	2.0
17	285	17	2.5	2.3	2.1
18	288	17	2.6	2.4	2.2
19	291	18	2.7	2.5	2.3
20	295	19	2.9	2.6	2.4
21	299	20	3.0	2.8	2.6
22	303	21	3.0	2.9	2.7
23	307	21	3.1	3.0	2.8
24	312	22	3.2	3.1	3.0
25	317	23	3.3	3.1	3.0
26	323	23	3.3	3.2	3.1
27	328	24	3.4	3.3	3.2
28	334	24	3.5	3.4	3.3
29	340	24	3.6	3.5	3.4
30	346	24	3.7	3.6	3.5
31	352	24	3.8	3.7	3.6
32	358	24	3.9	3.8	3.7
33	363	24	4.0	3.9	3.8

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
34	369	24	4.1	4.0	3.9
35	375	23	4.2	4.1	4.0
36	380	23	4.3	4.2	4.1
37	386	23	4.4	4.3	4.2
38	391	22	4.5	4.4	4.3
39	395	22	4.6	4.5	4.4
40	400	21	4.7	4.6	4.5
41	405	21	4.8	4.7	4.6
42	409	21	4.9	4.8	4.7
43	413	21	5.0	4.8	4.7
44	418	21	5.1	4.9	4.8
45	422	21	5.3	5.0	4.9
46	427	21	5.5	5.2	5.0
47	431	22	5.6	5.3	5.2
48	436	23	5.8	5.5	5.3
49	442	24	6.0	5.7	5.5
50	448	26	6.0	5.9	5.6
51	456	30	6.0	6.0	5.9
52	467	36	6.0	6.0	6.0
53	485	51	6.0	6.0	6.0
54	516	95	6.0	6.0	6.0

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

2.6.3.6 Grades 9–12

Table 2.6.3.6.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 9-12 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	232	77	1.3	1.2	1.1	1.0
1	252	45	1.5	1.4	1.3	1.2
2	266	32	1.7	1.6	1.5	1.3
3	275	27	1.8	1.7	1.6	1.5
4	282	25	1.9	1.8	1.7	1.5
5	288	24	1.9	1.8	1.7	1.6
6	293	23	2.1	1.9	1.8	1.7
7	299	23	2.3	2.0	1.8	1.7
8	304	24	2.5	2.2	1.9	1.8
9	310	25	2.7	2.4	2.0	1.9
10	317	27	2.9	2.6	2.3	1.9
11	325	29	3.1	2.9	2.6	2.2
12	334	31	3.2	3.1	2.9	2.6
13	344	33	3.4	3.3	3.1	3.0
14	355	34	3.6	3.4	3.3	3.2
15	367	34	3.8	3.6	3.5	3.4
16	379	34	4.0	3.8	3.7	3.6
17	390	34	4.2	4.0	3.9	3.8
18	402	33	4.4	4.3	4.2	4.0
19	412	31	4.6	4.5	4.4	4.2
20	421	30	4.8	4.7	4.6	4.4
21	430	30	5.0	4.8	4.7	4.6
22	439	30	5.2	5.0	4.9	4.8
23	448	31	5.4	5.2	5.1	5.0
24	458	33	5.7	5.5	5.3	5.2
25	471	38	6.0	5.8	5.6	5.4
26	489	52	6.0	6.0	5.9	5.7
27	521	94	6.0	6.0	6.0	6.0

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

Table 2.6.3.6.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 9-12 B/C S501 Paper

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	233	47	1.3	1.2	1.1	1.0
2	248	32	1.5	1.4	1.3	1.1
3	256	26	1.6	1.5	1.4	1.2
4	262	22	1.7	1.5	1.4	1.3
5	267	20	1.7	1.6	1.5	1.4
6	270	18	1.7	1.6	1.5	1.4
7	274	17	1.8	1.7	1.6	1.5
8	276	17	1.8	1.7	1.6	1.5
9	279	16	1.8	1.7	1.6	1.5
10	282	16	1.9	1.8	1.7	1.5
11	284	16	1.9	1.8	1.7	1.6
12	287	16	1.9	1.8	1.7	1.6
13	289	16	2.0	1.8	1.7	1.6
14	292	16	2.1	1.9	1.8	1.7
15	295	16	2.2	1.9	1.8	1.7
16	297	16	2.2	1.9	1.8	1.7
17	300	17	2.3	2.0	1.9	1.8
18	303	17	2.4	2.1	1.9	1.8
19	306	18	2.5	2.2	1.9	1.8
20	310	19	2.7	2.4	2.0	1.9
21	314	20	2.8	2.5	2.2	1.9
22	318	21	2.9	2.7	2.3	2.0
23	322	21	3.0	2.8	2.5	2.2
24	327	22	3.1	3.0	2.7	2.3
25	332	23	3.2	3.1	2.8	2.6
26	338	23	3.3	3.2	3.0	2.8
27	343	24	3.4	3.2	3.1	3.0
28	349	24	3.5	3.3	3.2	3.1
29	355	24	3.6	3.4	3.3	3.2
30	361	24	3.7	3.5	3.4	3.3
31	367	24	3.8	3.6	3.5	3.4
32	373	24	3.9	3.7	3.6	3.5
33	379	24	4.0	3.8	3.7	3.6

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
34	384	24	4.1	3.9	3.8	3.7
35	390	23	4.2	4.0	3.9	3.8
36	396	23	4.3	4.2	4.1	3.9
37	401	23	4.4	4.3	4.2	4.0
38	406	22	4.5	4.4	4.3	4.1
39	411	22	4.6	4.5	4.4	4.2
40	415	21	4.7	4.5	4.4	4.3
41	420	21	4.8	4.6	4.5	4.4
42	424	21	4.8	4.7	4.6	4.5
43	429	21	4.9	4.8	4.7	4.6
44	433	21	5.0	4.9	4.8	4.7
45	437	21	5.1	5.0	4.9	4.8
46	442	21	5.3	5.1	5.0	4.9
47	446	22	5.4	5.2	5.1	5.0
48	451	23	5.5	5.3	5.2	5.1
49	457	24	5.6	5.4	5.3	5.2
50	463	26	5.8	5.6	5.4	5.3
51	471	30	6.0	5.8	5.6	5.4
52	482	36	6.0	6.0	5.8	5.6
53	500	51	6.0	6.0	6.0	6.0
54	531	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

2.6.4.0 Kindergarten

Table 2.6.4.0
Raw Score to Scale Score to Proficiency Level Conversion: Spek K S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for K
0	100	183	1.0
1	123	133	1.2
2	147	83	1.5
3	169	63	1.7
4	191	55	2.0
5	211	52	2.3
6	230	48	2.6
7	250	41	3.0
8	301	32	4.0
9	349	44	5.0
10	392	105	6.0

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

2.6.4.1 Grade 1

Table 2.6.4.1.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 1 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	106	65	1.0
1	117	52	1.1
2	135	39	1.2
3	147	35	1.4
4	157	33	1.5
5	167	33	1.6
6	177	34	1.7
7	188	35	1.8
8	200	36	1.9
9	212	39	2.1
10	227	42	2.3
11	245	48	2.7
12	270	54	3.1
13	296	52	3.7
14	318	48	4.1
15	338	47	4.5
16	359	50	4.9
17	380	59	5.4
18	401	75	5.9

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 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G1
6	106	60	1.0
7	175	29	1.6
8	183	28	1.7
9	190	28	1.8
10	198	28	1.9
11	205	28	2.0
12	213	29	2.1
13	220	29	2.2
14	229	31	2.4
15	238	32	2.5
16	248	34	2.7
17	259	37	2.9
18	272	38	3.2
19	286	38	3.5
20	299	37	3.7
21	310	35	3.9
22	321	34	4.2
23	331	33	4.4
24	341	33	4.6
25	351	34	4.8
26	362	35	5.0
27	375	39	5.3
28	388	44	5.6
29	401	51	5.9
30	414	60	6.0

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

2.6.4.2 Grade 2

Table 2.6.4.2.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 2 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
0	118	38	1.0
1	118	38	1.0
2	118	38	1.0
3	118	38	1.0
4	130	37	1.1
5	143	37	1.2
6	155	36	1.3
7	167	36	1.5
8	179	37	1.6
9	192	38	1.7
10	206	42	1.8
11	224	48	2.0
12	248	54	2.5
13	274	52	3.0
14	296	48	3.4
15	317	47	3.8
16	338	50	4.3
17	359	60	4.7
18	380	76	5.1

Note: The test form is shared between 2A and 3A.

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 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G2
6	118	36	1.0
7	157	32	1.4
8	166	31	1.5
9	175	30	1.5
10	183	30	1.6
11	192	31	1.7
12	200	31	1.8
13	209	32	1.8
14	219	33	1.9
15	229	34	2.1
16	240	35	2.3
17	252	36	2.6
18	263	36	2.8
19	275	36	3.0
20	287	35	3.2
21	298	35	3.5
22	309	34	3.7
23	320	34	3.9
24	331	35	4.1
25	342	36	4.3
26	354	38	4.6
27	368	41	4.8
28	382	46	5.1
29	396	52	5.5
30	425	74	6.0

Note: The test form is shared between 2B/C and 3B/C.

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

2.6.4.3 Grade 3

Table 2.6.4.3.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 3 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
0	118	38	1.0
1	118	38	1.0
2	118	38	1.0
3	118	38	1.0
4	130	37	1.1
5	143	37	1.2
6	155	36	1.3
7	167	36	1.4
8	179	37	1.5
9	192	38	1.6
10	206	42	1.7
11	224	48	1.9
12	248	54	2.2
13	274	52	2.8
14	296	48	3.2
15	317	47	3.6
16	338	50	4.1
17	359	60	4.5
18	380	76	4.8

Note: The test form is shared between 2A and 3A.

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 3 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G3
6	118	36	1.0
7	157	32	1.3
8	166	31	1.4
9	175	30	1.4
10	183	30	1.5
11	192	31	1.6
12	200	31	1.7
13	209	32	1.7
14	219	33	1.8
15	229	34	1.9
16	240	35	2.1
17	252	36	2.3
18	263	36	2.5
19	275	36	2.8
20	287	35	3.0
21	298	35	3.3
22	309	34	3.5
23	320	34	3.7
24	331	35	3.9
25	342	36	4.1
26	354	38	4.4
27	368	41	4.6
28	382	46	4.9
29	396	52	5.2
30	425	74	6.0

Note: The test form is shared between 2B/C and 3B/C.

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

2.6.4.4 Grades 4–5

Table 2.6.4.4.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 4-5 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	130	48	1.0	1.0
1	130	48	1.0	1.0
2	142	40	1.1	1.0
3	155	36	1.2	1.1
4	166	34	1.3	1.2
5	176	34	1.4	1.3
6	187	34	1.5	1.4
7	198	34	1.6	1.5
8	209	35	1.6	1.6
9	221	37	1.7	1.7
10	234	40	1.9	1.8
11	250	44	2.0	1.9
12	270	49	2.5	2.2
13	293	51	3.0	2.7
14	316	50	3.4	3.2
15	339	50	3.9	3.7
16	362	53	4.3	4.2
17	385	61	4.7	4.6
18	408	77	5.2	5.0

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 4-5 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
6	130	47	1.0	1.0
7	192	32	1.5	1.4
8	201	31	1.6	1.5
9	210	31	1.7	1.6
10	219	31	1.7	1.6
11	228	31	1.8	1.7
12	236	31	1.9	1.8
13	245	31	1.9	1.8
14	254	31	2.1	1.9
15	263	32	2.3	2.1
16	272	33	2.5	2.3
17	282	34	2.7	2.5
18	293	35	3.0	2.7
19	305	36	3.2	3.0
20	317	37	3.4	3.3
21	330	37	3.7	3.5
22	342	36	4.0	3.8
23	354	36	4.2	4.0
24	366	36	4.4	4.2
25	377	36	4.6	4.4
26	390	37	4.8	4.7
27	403	40	5.1	4.9
28	416	45	5.5	5.2
29	429	51	5.8	5.6
30	443	60	6.0	6.0

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

2.6.4.5 Grades 6–8

Table 2.6.4.5.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 6-8 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	148	51	1.0	1.0	1.0
1	148	51	1.0	1.0	1.0
2	165	42	1.2	1.1	1.1
3	179	39	1.3	1.2	1.2
4	193	39	1.4	1.3	1.3
5	207	38	1.5	1.4	1.4
6	219	36	1.6	1.5	1.5
7	231	35	1.7	1.6	1.6
8	242	36	1.8	1.7	1.6
9	254	37	1.8	1.8	1.7
10	268	41	2.0	1.9	1.8
11	286	47	2.4	2.2	2.0
12	309	54	2.9	2.8	2.6
13	335	52	3.5	3.3	3.2
14	358	48	3.9	3.7	3.6
15	378	47	4.3	4.1	4.0
16	398	50	4.6	4.5	4.3
17	418	59	5.0	4.8	4.7
18	438	73	5.6	5.4	5.1

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 6-8 B/C S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
6	148	50	1.0	1.0	1.0
7	223	32	1.6	1.6	1.5
8	232	31	1.7	1.6	1.6
9	241	30	1.7	1.7	1.6
10	249	30	1.8	1.7	1.7
11	257	30	1.9	1.8	1.8
12	265	30	1.9	1.9	1.8
13	273	30	2.1	1.9	1.9
14	282	31	2.3	2.1	1.9
15	291	32	2.5	2.3	2.1
16	301	34	2.7	2.6	2.4
17	313	36	3.0	2.9	2.7
18	325	37	3.3	3.1	3.0
19	338	37	3.5	3.4	3.2
20	350	36	3.8	3.6	3.5
21	362	35	4.0	3.8	3.7
22	373	34	4.2	4.0	3.9
23	384	34	4.4	4.2	4.1
24	394	34	4.5	4.4	4.3
25	405	35	4.7	4.6	4.5
26	416	36	4.9	4.8	4.6
27	429	39	5.3	5.1	4.9
28	442	44	5.7	5.5	5.3
29	455	51	6.0	5.9	5.7
30	468	60	6.0	6.0	6.0

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

2.6.4.6 Grades 9–12

Table 2.6.4.6.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 9-12 A S501 Paper

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	172	38	1.1	1.0	1.0	1.0
1	172	38	1.1	1.0	1.0	1.0
2	172	38	1.1	1.0	1.0	1.0
3	183	34	1.2	1.1	1.1	1.0
4	193	33	1.2	1.2	1.2	1.1
5	203	33	1.3	1.3	1.2	1.2
6	214	35	1.4	1.4	1.3	1.3
7	226	37	1.5	1.4	1.4	1.4
8	239	38	1.6	1.5	1.5	1.5
9	253	40	1.7	1.6	1.6	1.6
10	268	43	1.8	1.8	1.7	1.7
11	287	48	1.9	1.9	1.9	1.8
12	311	54	2.5	2.4	2.3	2.2
13	337	52	3.1	3.0	3.0	2.9
14	360	48	3.5	3.4	3.3	3.3
15	380	47	3.9	3.7	3.6	3.6
16	401	50	4.2	4.1	4.0	3.9
17	422	59	4.6	4.5	4.4	4.3
18	443	75	5.1	4.9	4.8	4.7

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

Table 2.6.4.6.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 9-12 B/C S501 Paper

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	217	31	1.4	1.4	1.3	1.3
8	226	31	1.5	1.4	1.4	1.4
9	235	31	1.5	1.5	1.5	1.4
10	243	30	1.6	1.6	1.5	1.5
11	252	30	1.7	1.6	1.6	1.6
12	260	30	1.7	1.7	1.7	1.6
13	268	30	1.8	1.8	1.7	1.7
14	277	31	1.9	1.8	1.8	1.8
15	286	33	1.9	1.9	1.9	1.8
16	296	34	2.1	2.0	1.9	1.9
17	308	36	2.4	2.3	2.2	2.1
18	321	38	2.8	2.6	2.5	2.5
19	334	38	3.1	3.0	2.9	2.8
20	347	37	3.3	3.2	3.1	3.1
21	358	35	3.5	3.4	3.3	3.2
22	369	34	3.7	3.6	3.5	3.4
23	380	33	3.9	3.7	3.6	3.6
24	390	33	4.0	3.9	3.8	3.7
25	400	34	4.2	4.1	4.0	3.9
26	411	36	4.4	4.3	4.2	4.1
27	424	39	4.7	4.5	4.4	4.3
28	437	44	4.9	4.8	4.7	4.6
29	455	54	5.5	5.3	5.1	5.0
30	476	72	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 and Recalibration Summary

All ACCESS Series 501 Paper test forms are static forms, and we did not conduct annual equating analyses.

For technical details on the Kindergarten test, see MacGregor, Kenyon, Gibson, and Evans (2009). For the ACCESS Series 501 Grades 1–12, we provide detail below on prior years that test forms have been used, where relevant, and on equating processes that were in place.

Listening and Reading

For ACCESS Paper Listening and Reading Grades 1–12 Tier A, all forms have been used in prior years. For ACCESS Paper Listening and Reading Grades 1–12 Tier B/C, we constructed new forms for Series 501 using the Series 302 and Series 303 Tier B and Tier C item pools (see Part 1, Section 2.3). Table 2.7.1 summarizes the sources of Listening and Reading forms for Paper Series 501.

Table 2.7.1
Sources of Series 501 Paper Listening and Reading Forms

	Listening		**Reading**	
Tier A	Years previously used: Series 403 Paper Series 402 Paper Series 401 Paper Series 400 Paper Series 303	 2018-19 2017-18 2016-17 2015-16 2014-15	Years previously used: Series 402 Paper Series 400 Paper Series 302	 2017-18 2015-16 2013–14
Tier B/C	Newly constructed for Series 501 using Series 302 and Series 303 Tier B and Tier C item pools		Newly constructed for Series 501 using Series 302 and Series 303 Tier B and Tier C item pools	

The newly constructed Tier B/C forms were drawn from the pool of Series 302 and 303 ACCESS. These forms were operational in 2013–2014 and 2014–2015, which were the 2 years prior to the launch of ACCESS Online. To mitigate concerns that there might be systematic differences between the population of students who took ACCESS 302 and 303 and the population of students who currently take Paper ACCESS, we conducted a series of recalibration studies using Series 400 and Series 401 Paper population data to refine Series 302 and Series 303 Listening and Reading item parameters.

Since Series 401 Paper, Series 400 Paper, and Series 303 Listening Grades 1–12 test forms are identical, and since the Series 401 Paper population is more current than the Series 400 Paper population, we refined the item parameters for the Series 303 Listening Grades 1–12 forms using

Series 401 Paper population data. In the recalibration analyses, we initially anchored the difficulty measures of the Series 303 test items to their previously calibrated values from the Series 303 annual equating study. After the first calibration run, some items that were initially anchored 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 the difficulty value would have been had it not been anchored. If this value was large (i.e., above 0.30 or below -0.30), we unanchored that item in the final calibration run (i.e., its parameter was re-estimated). For Series 501 Paper Reading Grades 1–12 forms, a similar process was used to refine Series 302 and Series 303 item parameters using Series 400 and 401 Paper student population data, respectively.

For Listening Tier A, we applied these refined parameters to the intact Tier A forms from Series 303. For Reading Tier A, we applied these refined parameters to the intact Tier A forms from Series 302.

For Listening and Reading Tier B/C, we used the refined parameters derived from the recalibration studies to conduct a form selection meeting. We constructed the Series 501 Paper Listening and Reading Grades 1–12 Tier B/C forms at this meeting.

Writing and Speaking

Writing and Speaking are also static forms. Table 2.7.2 summarizes prior uses of these forms. Please see the Annual Technical Report for ACCESS for ELLs Paper Series 402 (Center for Applied Linguistics, 2019) for equating summaries for Writing and Speaking.

Table 2.7.2
Sources of Series 501 Paper Writing and Speaking Forms

	Writing		**Speaking**	
Tier A	Years previously used: Series 402 Paper Series 400 Paper	 2017-18 2015-16	Years previously used: Series 402 Paper	 2017-18
Tier B/C	Years previously used: Series 402 Paper Series 400 Paper	 2017-18 2015-16	Years previously used: Series 402 Paper	 2017-18

2.8 Test Characteristic Curve

Test characteristic curves (TCCs) graphically show the relationship between the ability measure (in logits) on the horizontal axis and the expected raw score or the estimated true score on the vertical axis. For a given ability measure, the corresponding expected raw score can be found via the TCC. For reporting purposes, ability measures are used to determine students' proficiency levels. Since TCC transforms ability measures to expected raw scores, this representation allows test users to relate student performance to the number of items on the test.

Mathematically, TCC is the sum of all item characteristic functions on the test form (Lord, 1980). Thus, the TCC depends on the item characteristic functions (Lord, 1980) of the items on the test form. The shape of TCC depends on several factors, including the number and the characteristics of items, the item response theory model used, and the values of the item parameters. Because of this, there is no explicit formula for TCC, and there are no parameters for the curve. The general form of the TCC is monotonically increasing. In most cases when the test form consists of multiple-choice items, such as in the Listening and Reading domains, the TCC curve is a smooth S shape. It is flat in the lower ability range, rises steeply in the middle, and becomes flat again on the right, at the level of proficiency above which students are expected to respond correctly to all items. In other cases, however, it will increase smoothly and then have a small plateau before increasing again. In all cases, it will be asymptotic to the value of the total number of items or total expected raw score points in the upper tail. The area where the TCC is the steepest is the area where the test provides higher discrimination and better measurement as compared to the area where the TCC is flat.

For tests consisting of polytomous tasks, the shape of the TCC is also affected by the values of the item category parameters. Since polytomous tasks have more score categories than multiple-choice items, each task has a wide range of values on the proficiency scale. The adjacent category boundaries are sometimes far apart as a result. In this situation, the TCC will have a less smooth curve or a small plateau in the area between the adjacent category boundaries. This pattern can be observed in Writing and Speaking, where the TCC may not form a perfect "S" shape. Such a pattern is also observed in other tests with polytomous items such as the National Assessment of Educational Progress Writing assessment (Muraki, 1993). Conversely, the closer the adjacent category boundaries are, the smoother the rise of the TCC will be along the ability levels.

There are five vertical lines in each of the TCC plots indicating the five 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 (PLs 1–6) for the domain being tested. (Note that for Kindergarten and Tier A tests in some domains, it was not possible to place into all six proficiency levels.) As would be expected, higher raw scores are required for placement in higher proficiency levels. The relative width of each section between the cut score lines, however, gives an indication of how many items on that form must be answered correctly (for Listening or Reading) or how

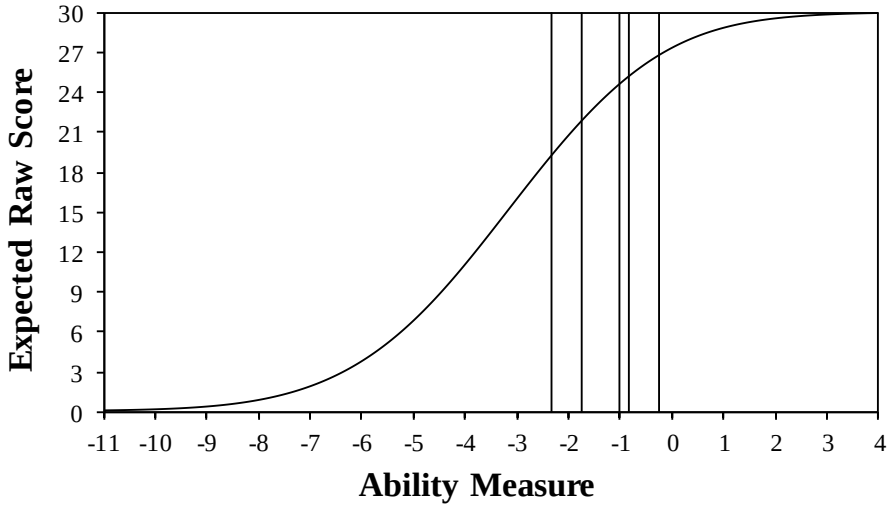
many points must be earned (for Writing or Speaking) to be placed into a WIDA proficiency level.

In addition to the TCC by tier, TCCs across tier for the grade-cluster are plotted on the same graph. Since each tier has different numbers of expected raw score points, it is not appropriate to compare the expected raw score points for the same proficiency measure between tiers. It is, however, informative to compare where the slopes are the steepest, which corresponds to the ability range where the best measurement information is provided. For example, the across-tier TCC for Listening Grade 1 showed that the Listening Grade 1 Tier A form provides the best measurement at around an ability measure of -1.0, or around PL 3, while the Listening Grade 1 Tier B/C form provides the best measurement at a higher proficiency level (an ability measure of 0.3 or around PL 5), as expected. In addition, it is informative to compare the area under the curve for the TCC of each tier form. For example, the Grade 1 Tier A curve covers an area of lower ability range than the Grade 1 Tier B/C curve. As expected, there is also considerable overlap between the areas covered by the two forms.

2.8.1 Listening

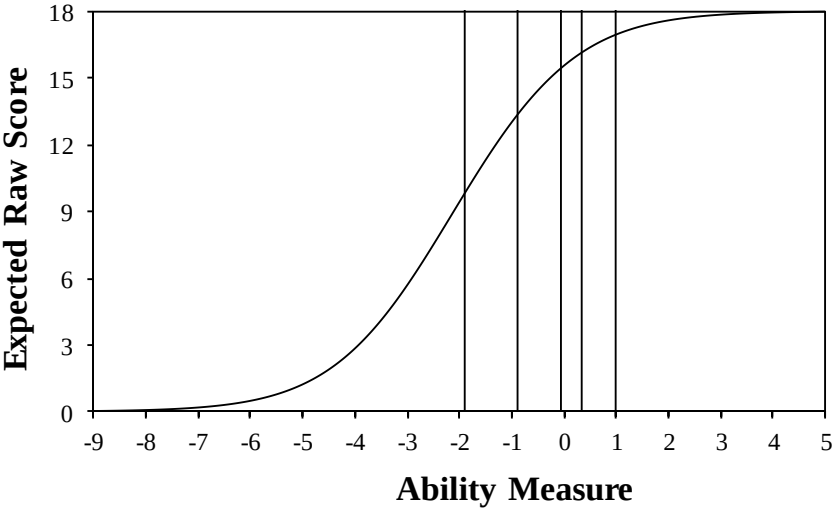
2.8.1.0 Kindergarten

Figure 2.8.1.0
Test Characteristic Curve: List K S501 Paper



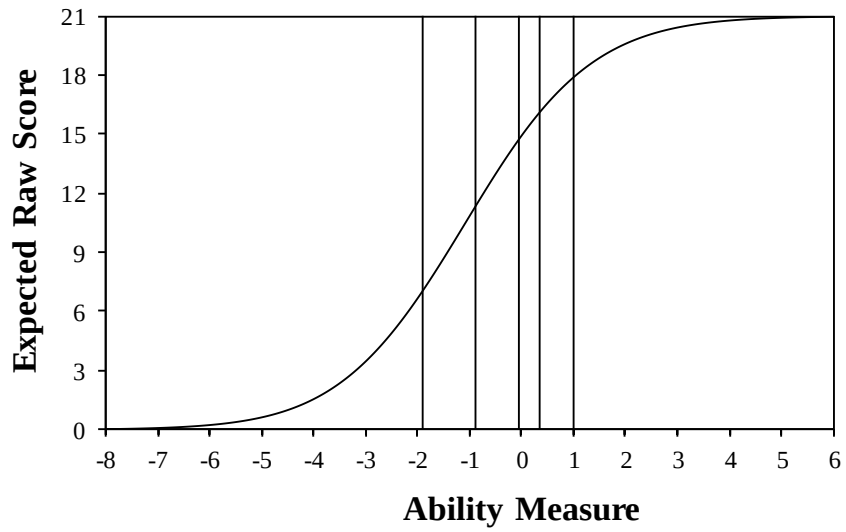
2.8.1.1 Grade 1

Figure 2.8.1.1.1
Test Characteristic Curve: List 1A S501 Paper



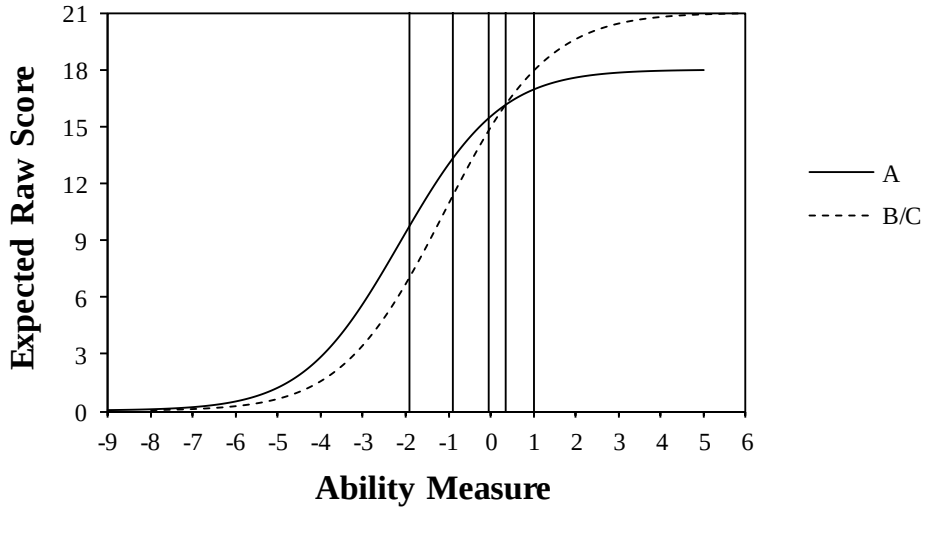
Note: The test form is shared between 1A and 2A.

Figure 2.8.1.1.2
Test Characteristic Curve: List 1B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

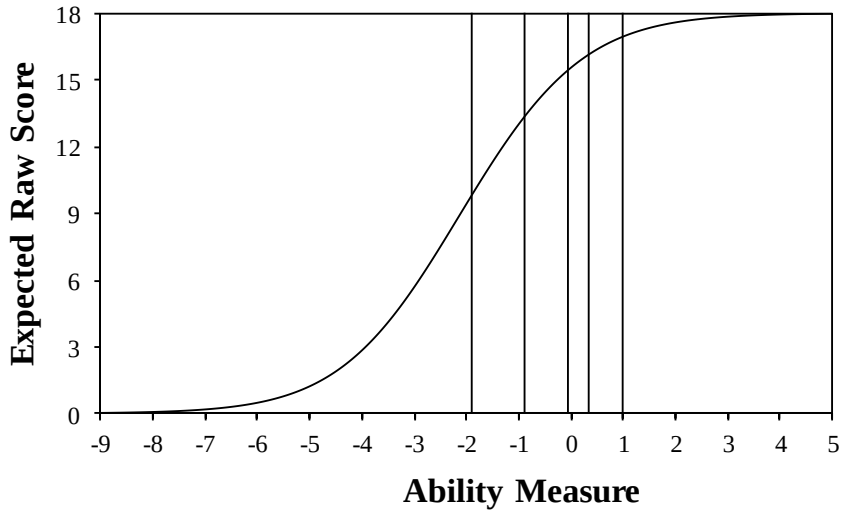
Figure 2.8.1.1.3
Test Characteristic Curve: List 1 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

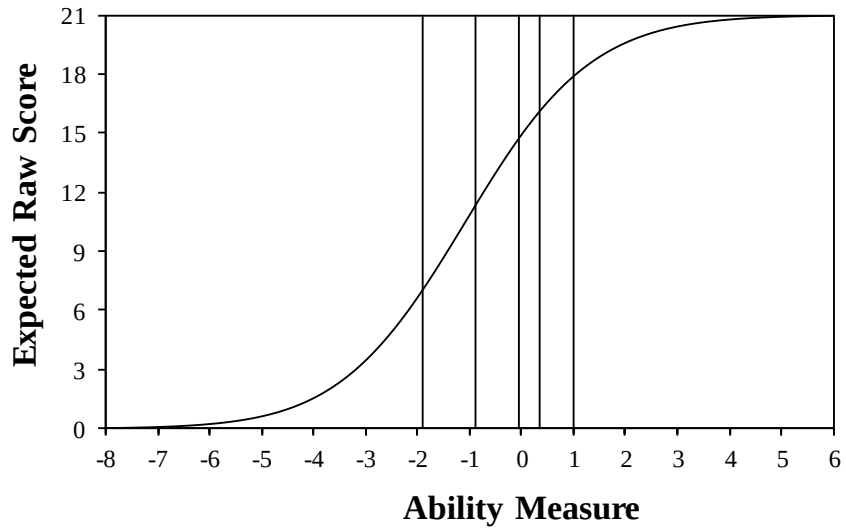
2.8.1.2 Grade 2

Figure 2.8.1.2.1
Test Characteristic Curve: List 2A S501 Paper



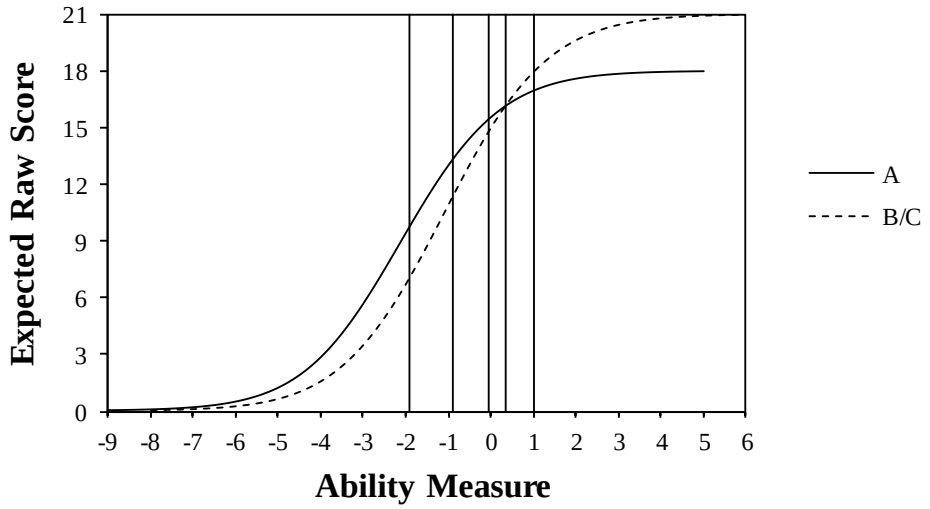
Note: The test form is shared between 1A and 2A.

Figure 2.8.1.2.2
Test Characteristic Curve: List 2B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

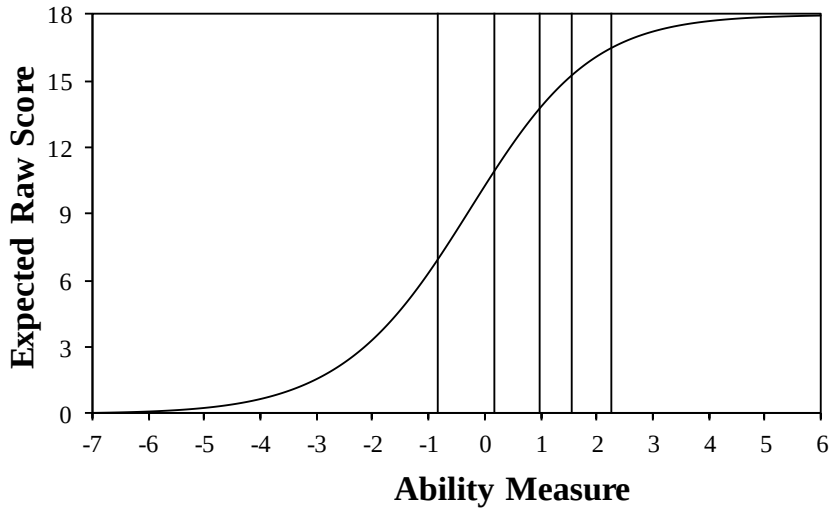
Figure 2.8.1.2.3
Test Characteristic Curve: List 2 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

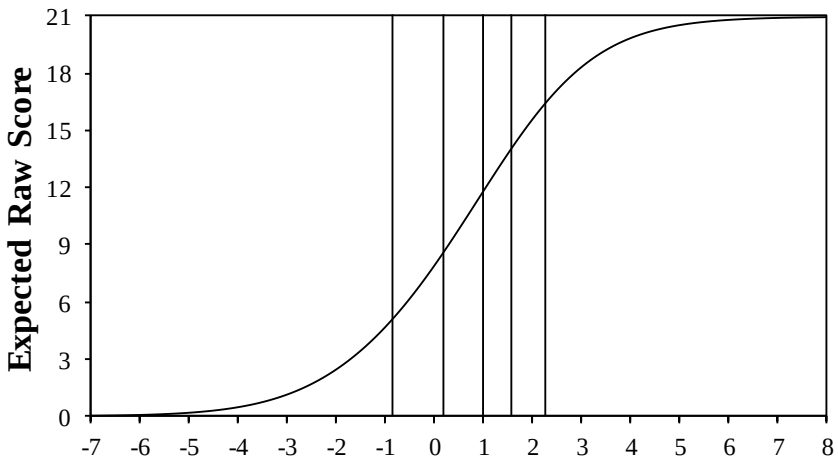
2.8.1.3 Grade 3

Figure 2.8.1.3.1
Test Characteristic Curve: List 3A S501 Paper



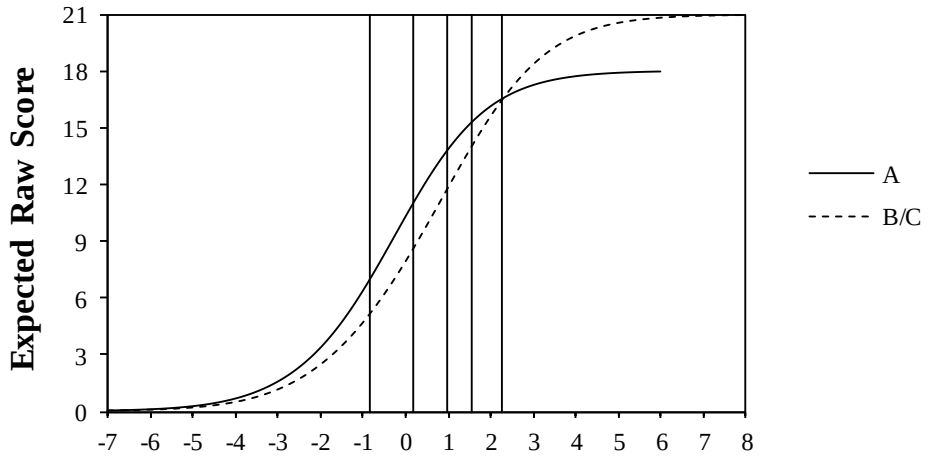
Note: The test form is shared between 3A and 4–5A.

Figure 2.8.1.3.2
Test Characteristic Curve: List 3B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

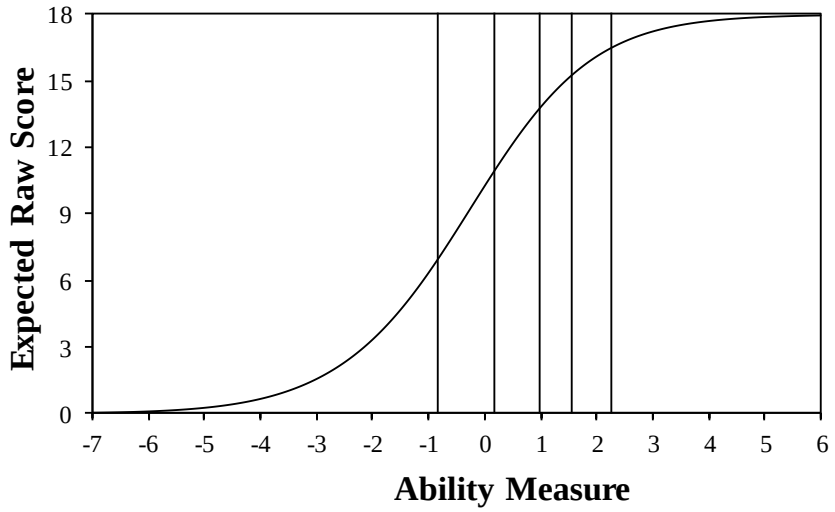
Figure 2.8.1.3.3
Test Characteristic Curve: List 3 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

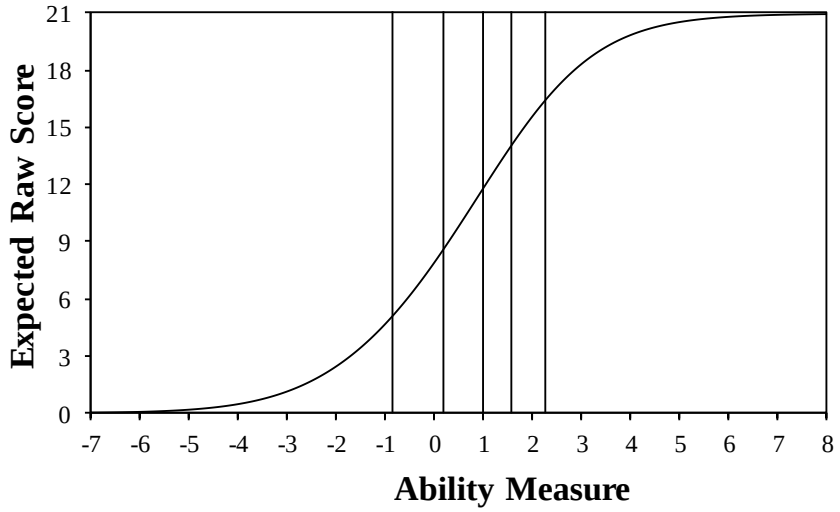
2.8.1.4 Grades 4–5

Figure 2.8.1.4.1
Test Characteristic Curve: List 4-5A S501 Paper



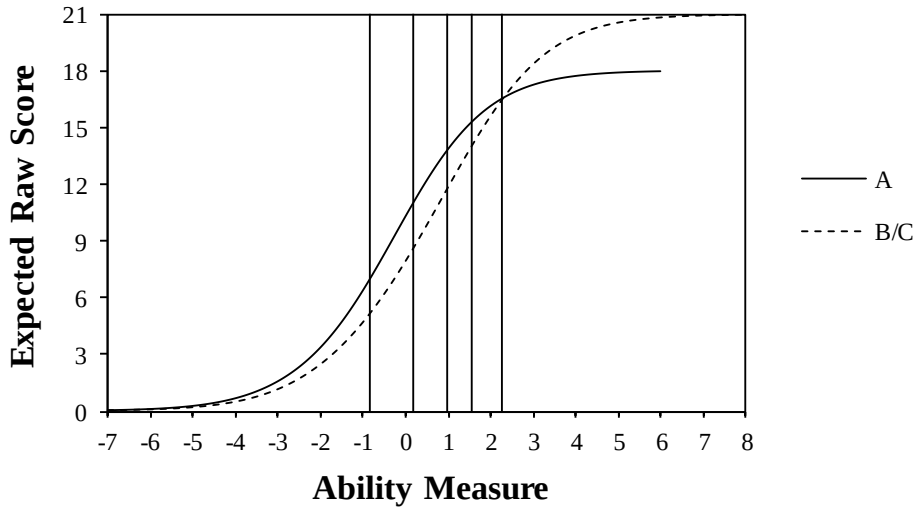
Note: The test form is shared between 3A and 4–5A.

Figure 2.8.1.4.2
Test Characteristic Curve: List 4-5B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

Figure 2.8.1.4.3
Test Characteristic Curve: List 4-5 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

2.8.1.5 Grades 6–8

Figure 2.8.1.5.1
Test Characteristic Curve: List 6-8A S501 Paper

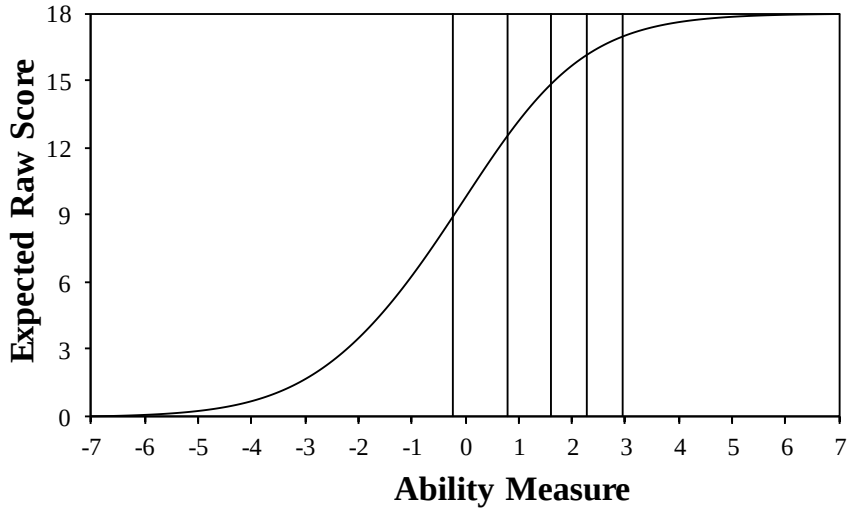


Figure 2.8.1.5.2
Test Characteristic Curve: List 6-8B/C S501 Paper

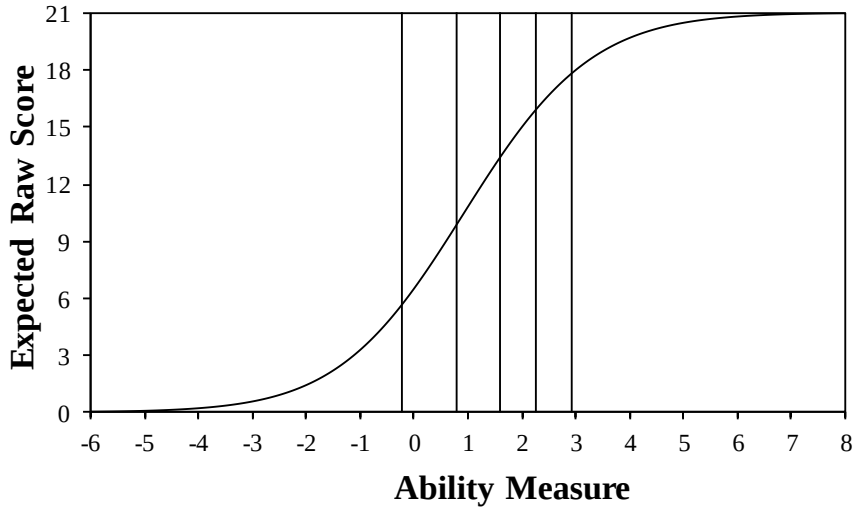
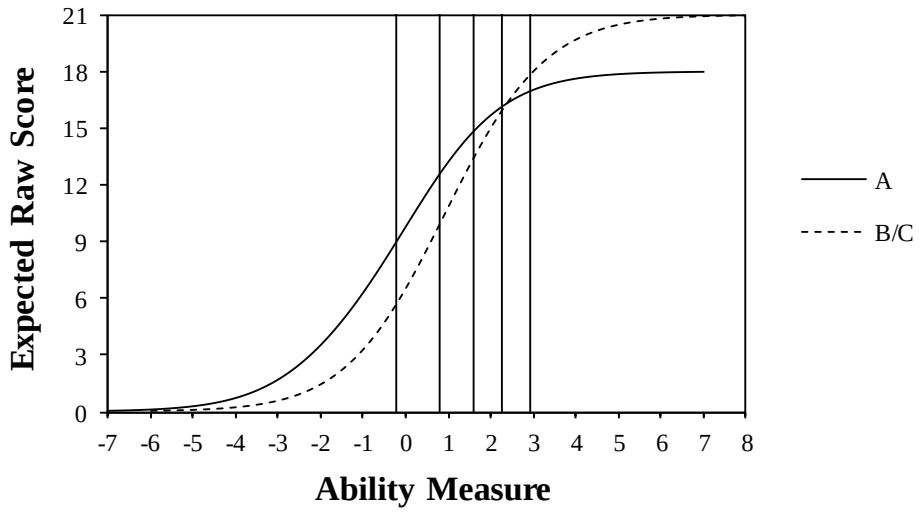


Figure 2.8.1.5.3
Test Characteristic Curve: List 6-8 S501 Paper



2.8.1.6 Grades 9–12

Figure 2.8.1.6.1
Test Characteristic Curve: List 9-12A S501 Paper

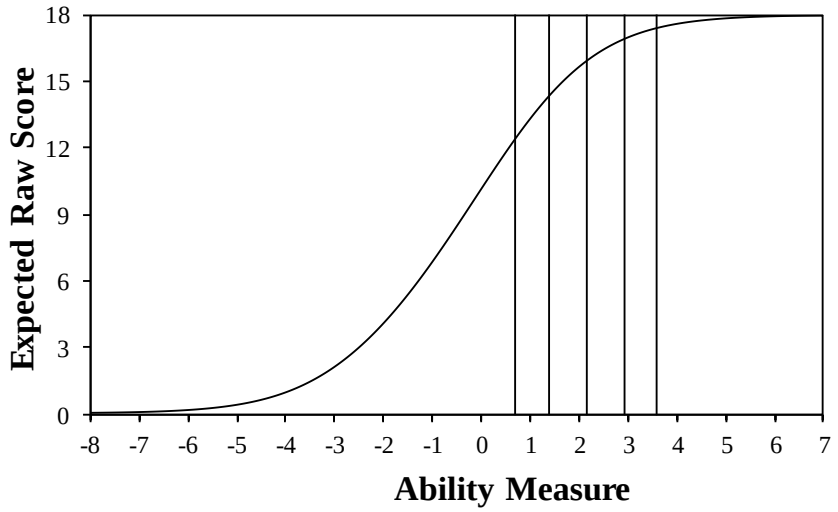


Figure 2.8.1.6.2
Test Characteristic Curve: List 9-12B/C S501 Paper

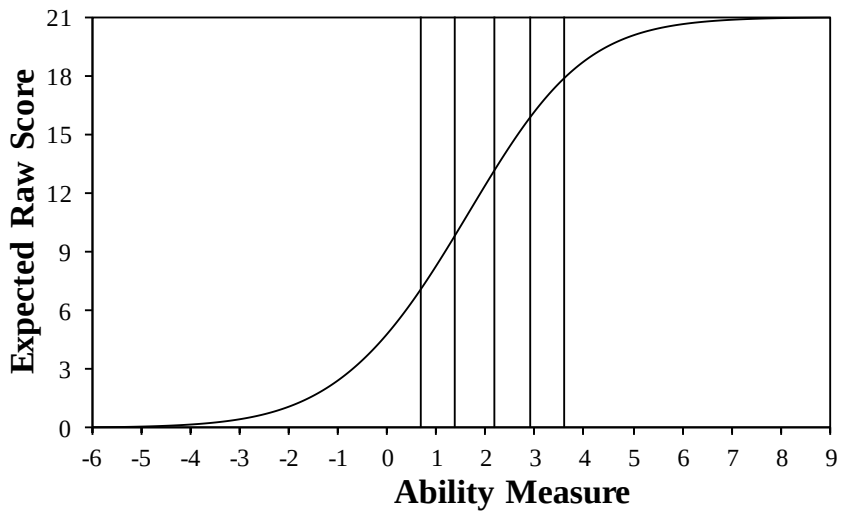
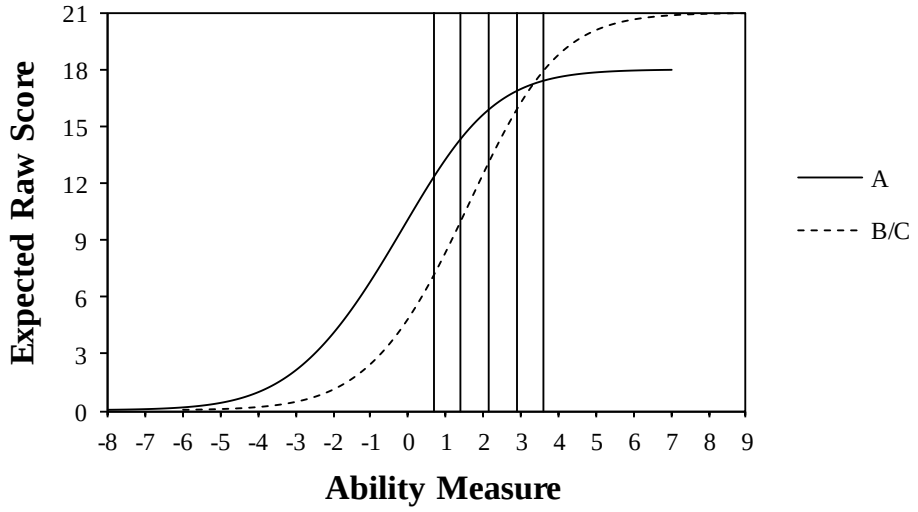


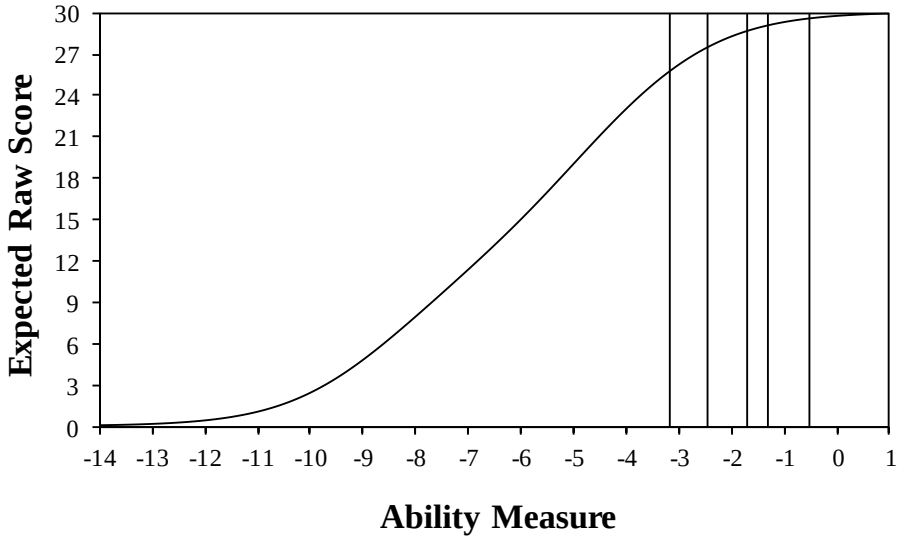
Figure 2.8.1.6.3
Test Characteristic Curve: List 9-12 S501 Paper



2.8.2 Reading

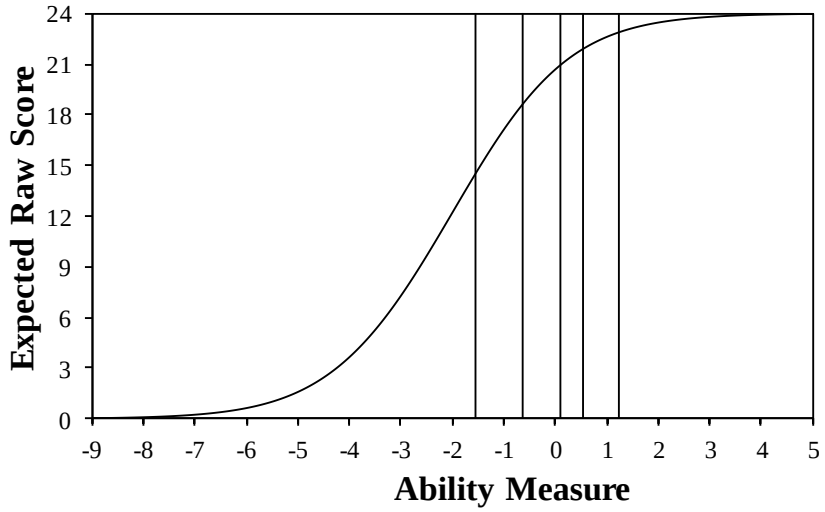
2.8.2.0 Kindergarten

Figure 2.8.2.0
Test Characteristic Curve: Read K S501 Paper



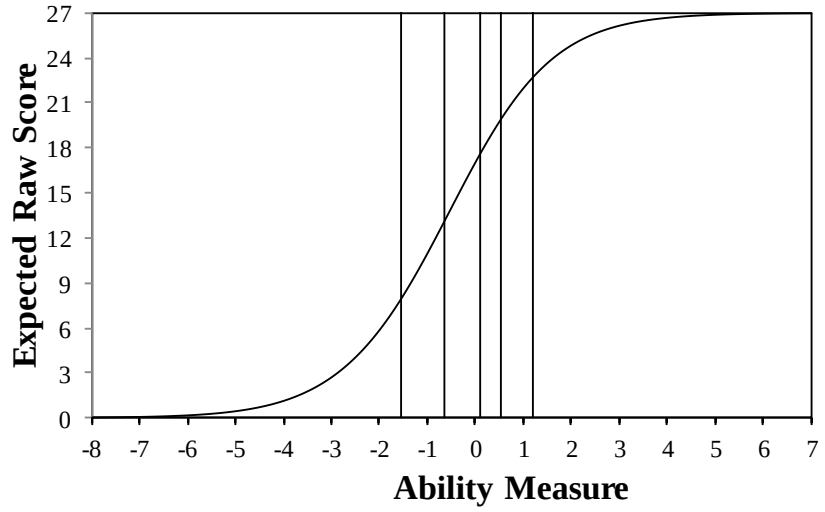
2.8.2.1 Grade 1

Figure 2.8.2.1.1
Test Characteristic Curve: Read 1A S501 Paper



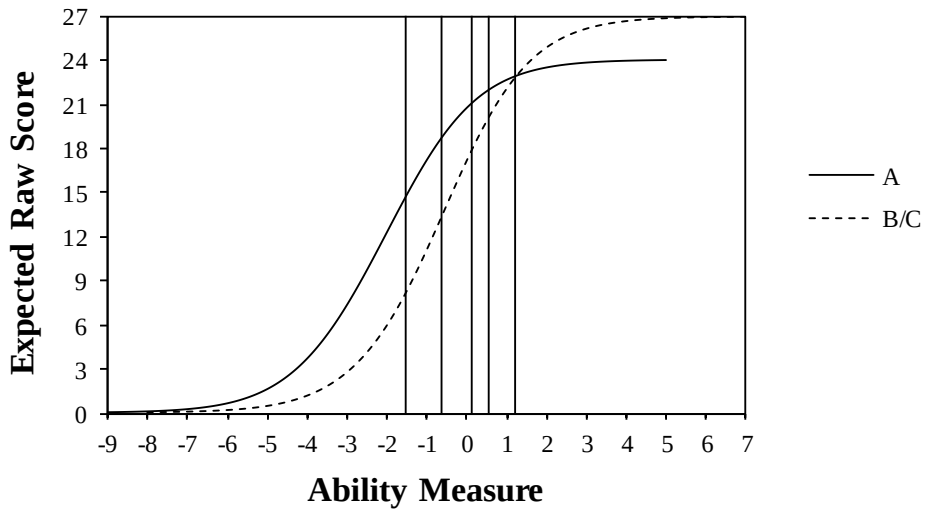
Note: The test form is shared between 1A and 2A.

Figure 2.8.2.1.2
Test Characteristic Curve: Read 1B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

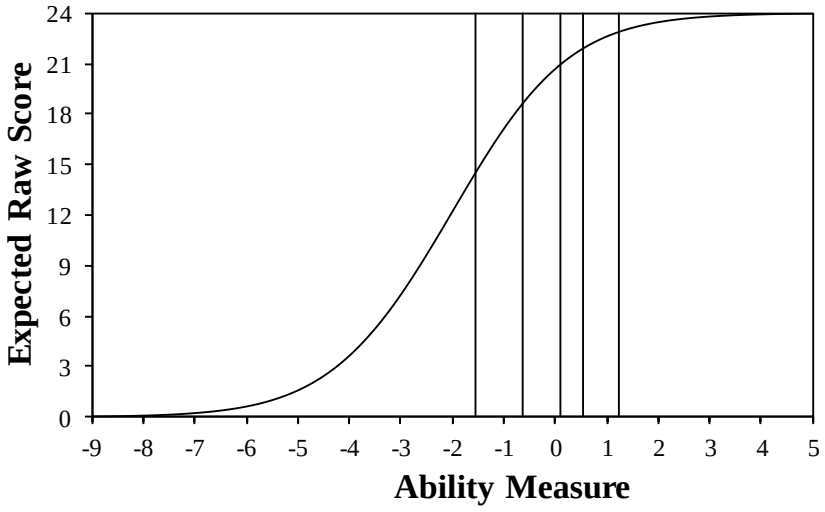
Figure 2.8.2.1.3
Test Characteristic Curve: Read 1 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

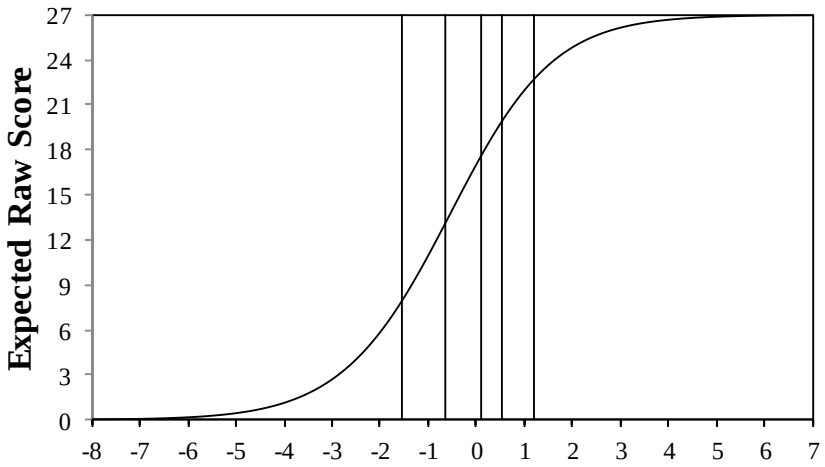
2.8.2.2 Grade 2

Figure 2.8.2.2.1
Test Characteristic Curve: Read 2A S501 Paper



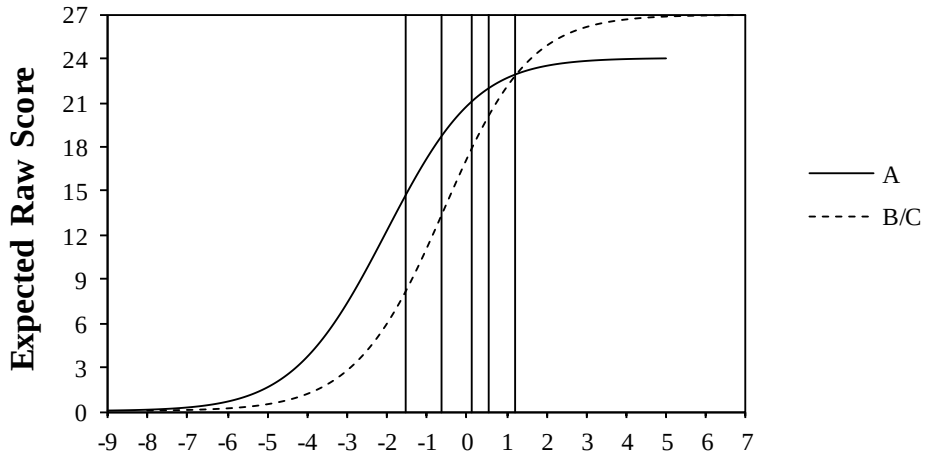
Note: The test form is shared between 1A and 2A.

Figure 2.8.2.2.2
Test Characteristic Curve: Read 2B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

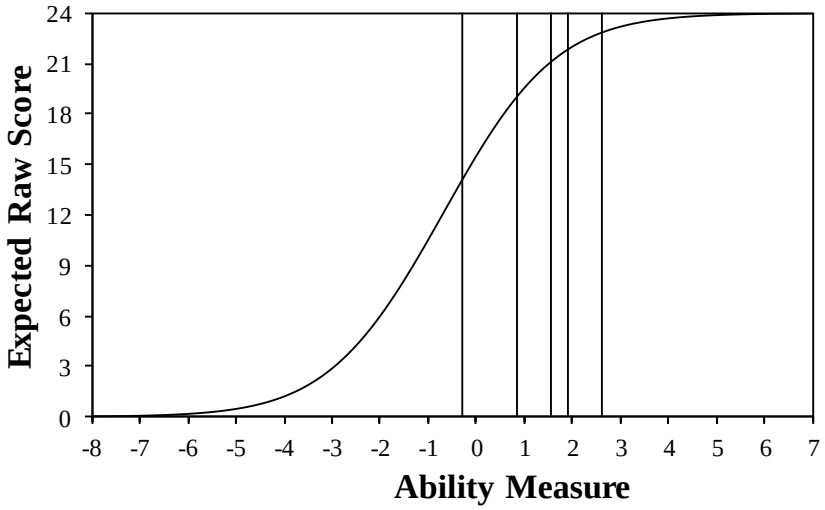
Figure 2.8.2.2.3
Test Characteristic Curve: Read 2 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

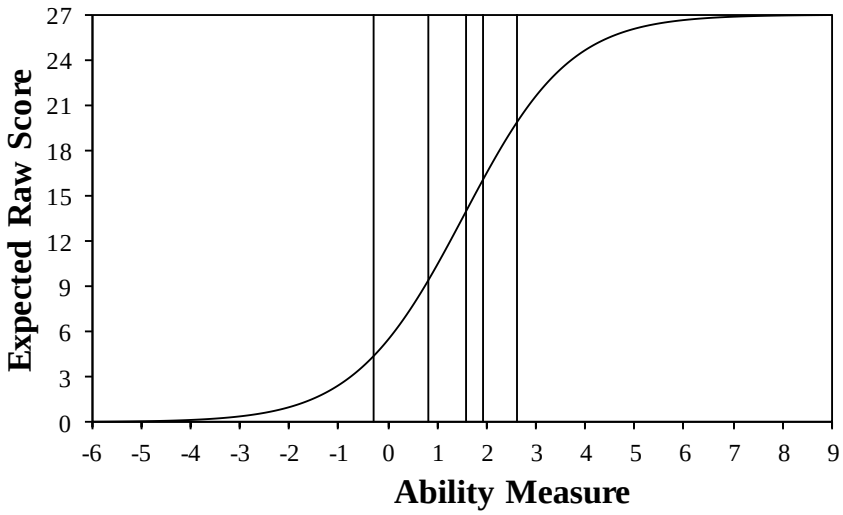
2.8.2.3 Grade 3

Figure 2.8.2.3.1
Test Characteristic Curve: Read 3A S501 Paper



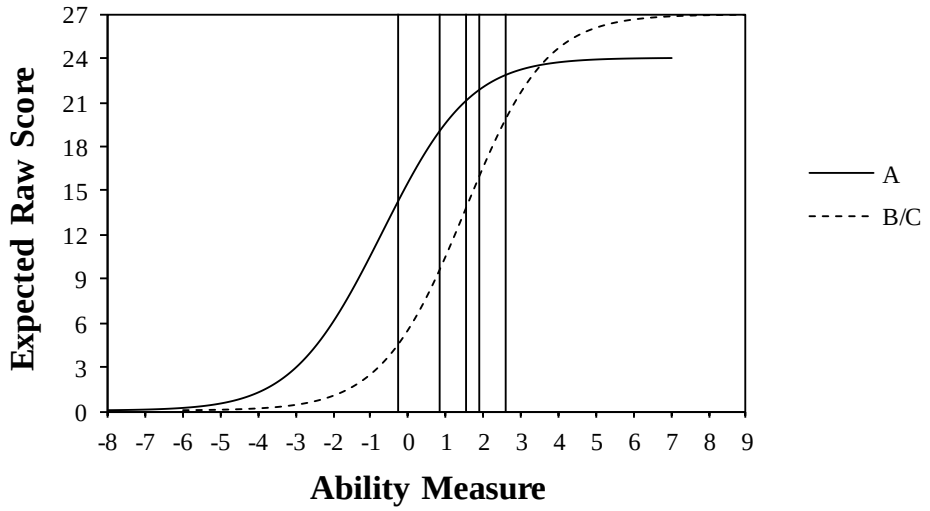
Note: The test form is shared between 3A and 4–5A.

Figure 2.8.2.3.2
Test Characteristic Curve: Read 3B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

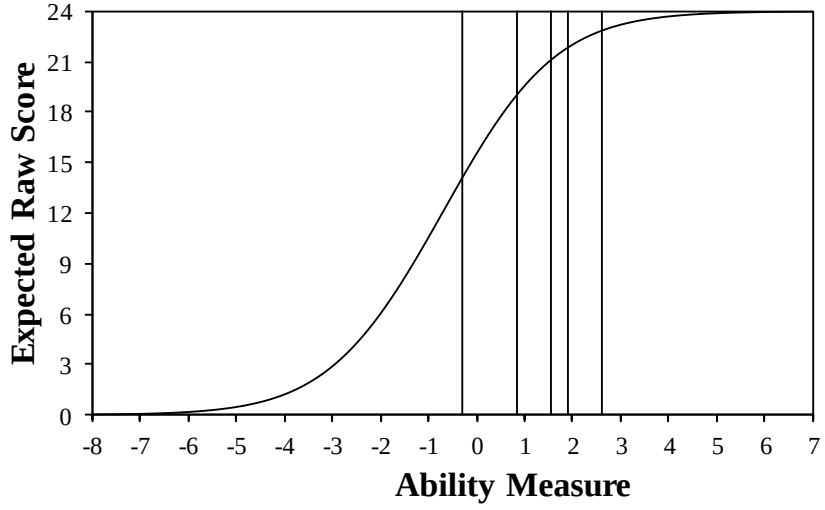
Figure 2.8.2.3.3
Test Characteristic Curve: Read 3 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

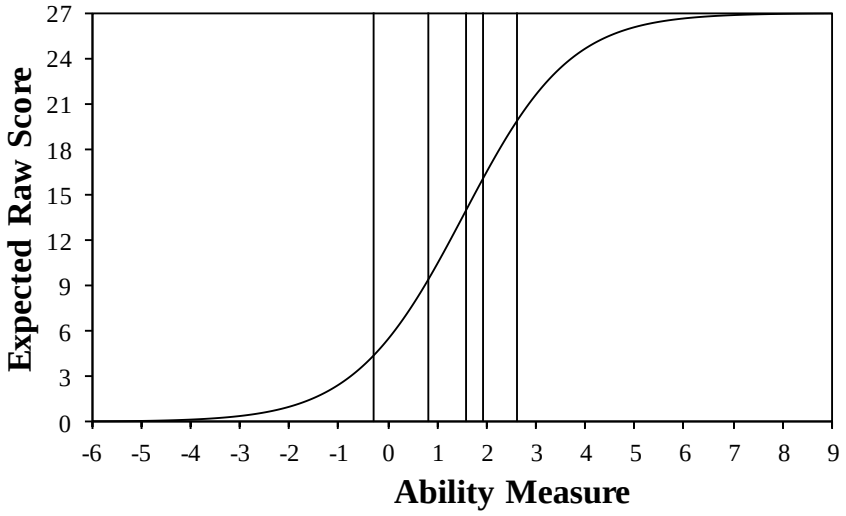
2.8.2.4 Grades 4–5

Figure 2.8.2.4.1
Test Characteristic Curve: Read 4-5A S501 Paper



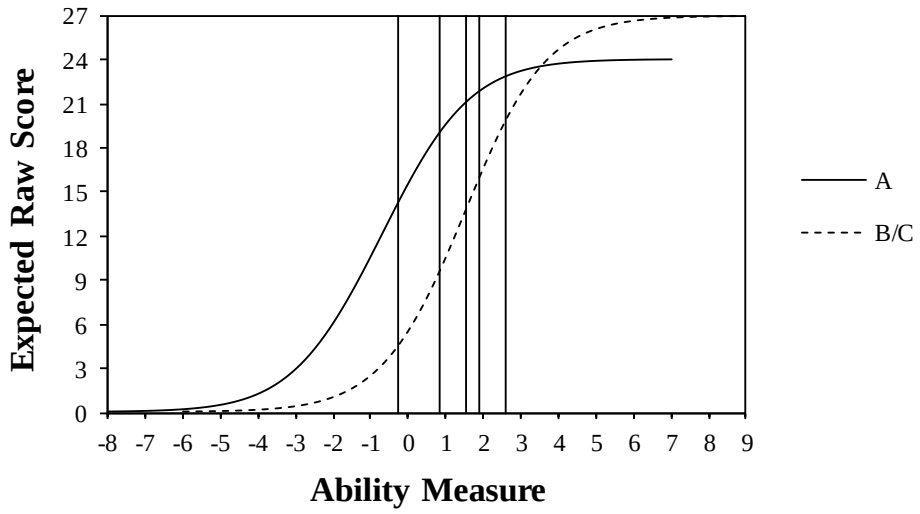
Note: The test form is shared between 3A and 4–5A.

Figure 2.8.2.4.2
Test Characteristic Curve: Read 4-5B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

Figure 2.8.2.4.3
Test Characteristic Curve: Read 4-5 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

2.8.2.5 Grades 6–8

Figure 2.8.2.5.1
Test Characteristic Curve: Read 6-8A S501 Paper

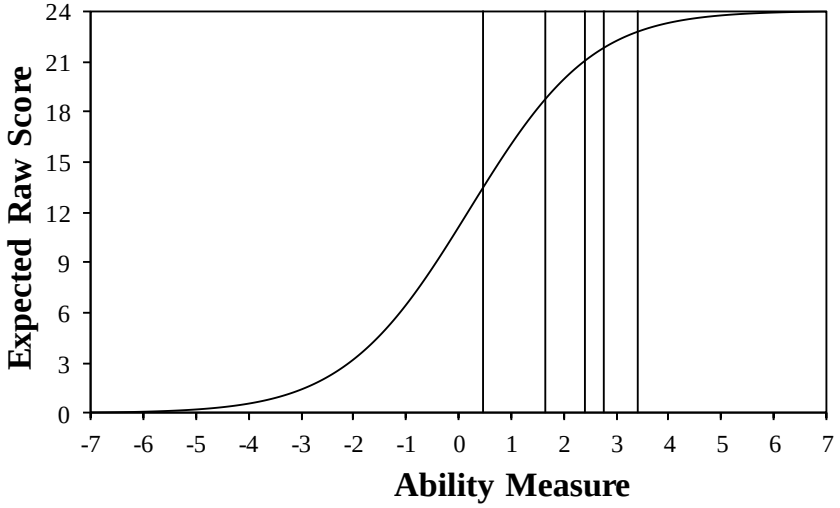


Figure 2.8.2.5.2
Test Characteristic Curve: Read 6-8B/C S501 Paper

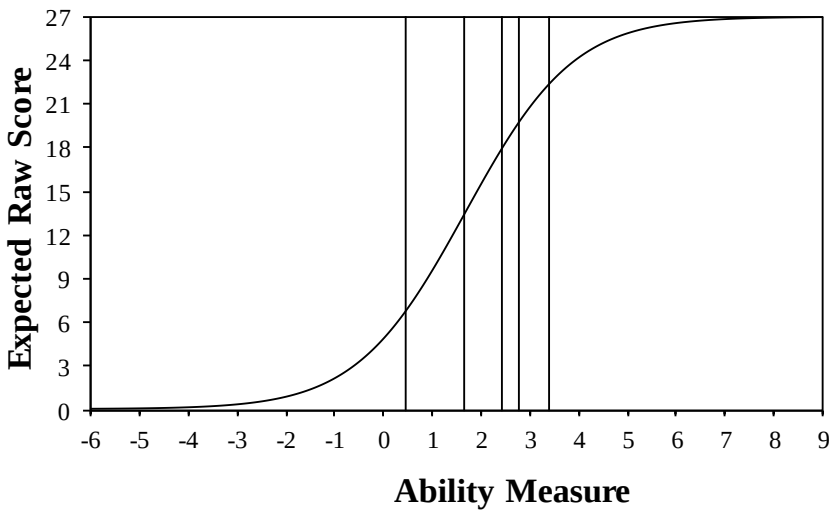
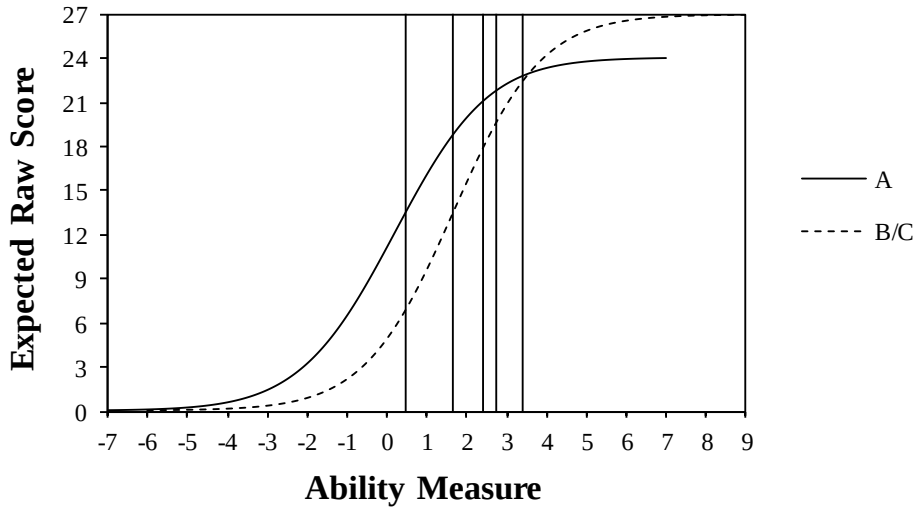


Figure 2.8.2.5.3
Test Characteristic Curve: Read 6-8 S501 Paper



2.8.2.6 Grades 9–12

Figure 2.8.2.6.1
Test Characteristic Curve: Read 9-12A S501 Paper

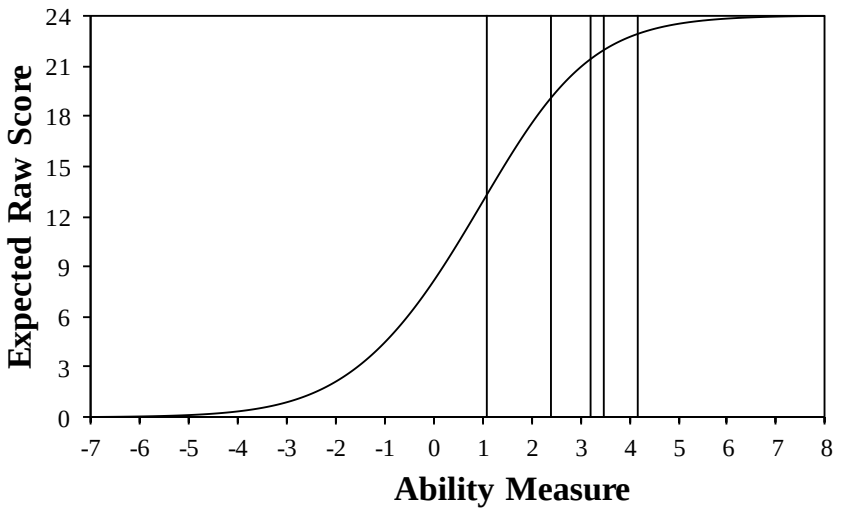


Figure 2.8.2.6.2
Test Characteristic Curve: Read 9-12B/C S501 Paper

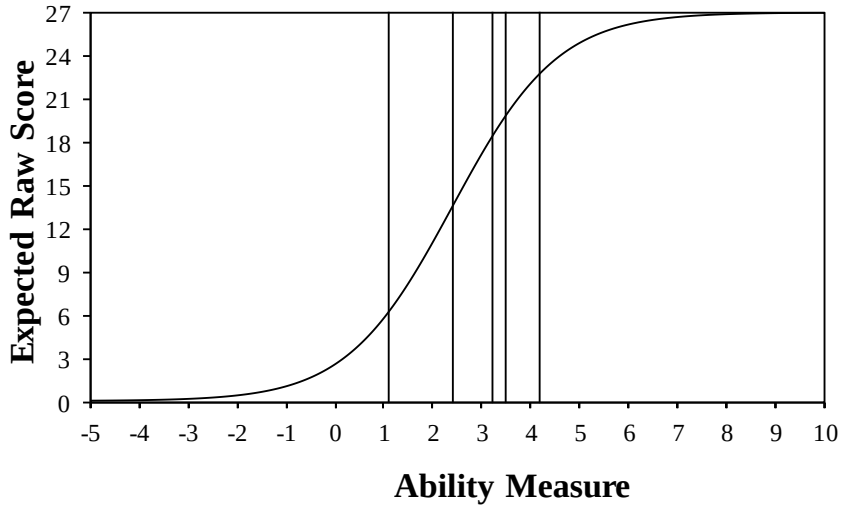
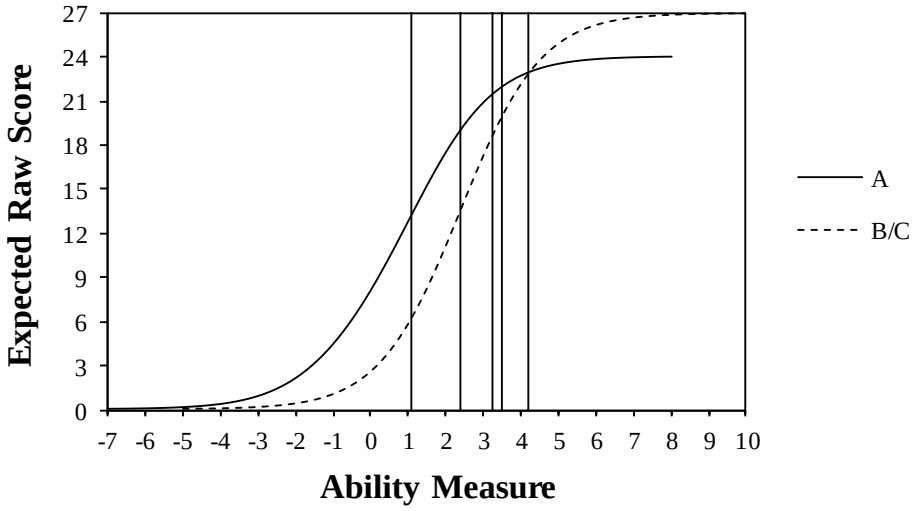


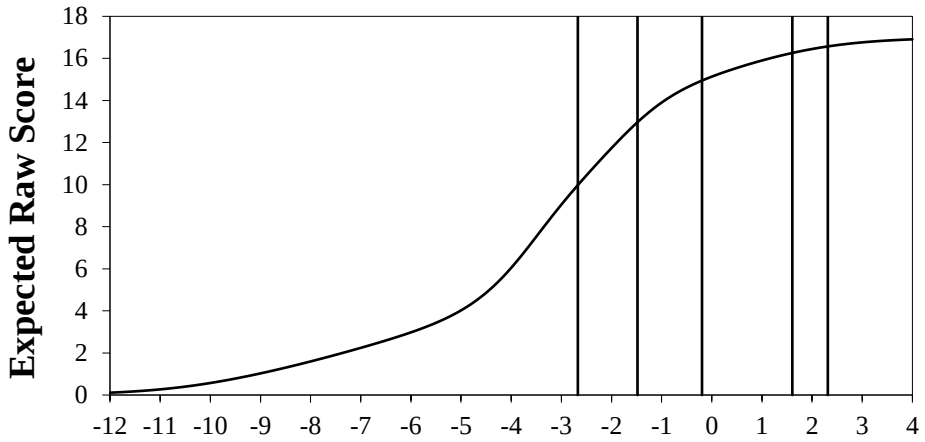
Figure 2.8.2.6.3
Test Characteristic Curve: Read 9-12 S501 Paper



2.8.3 Writing

2.8.3.0 Kindergarten

Figure 2.8.3.0
Test Characteristic Curve: Writ K S501 Paper



2.8.3.1 Grade 1

Figure 2.8.3.1.1
Test Characteristic Curve: Writ 1 A S501 Paper

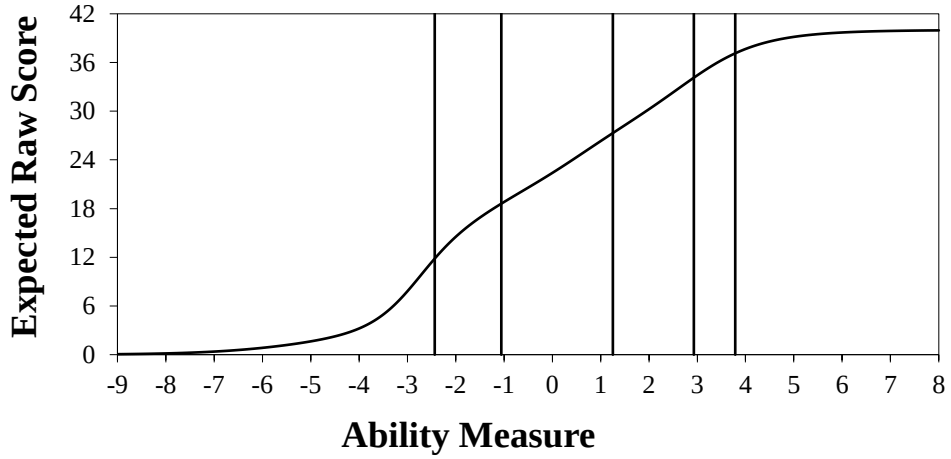


Figure 2.8.3.1.2
Test Characteristic Curve: Writ 1 B/C S501 Paper

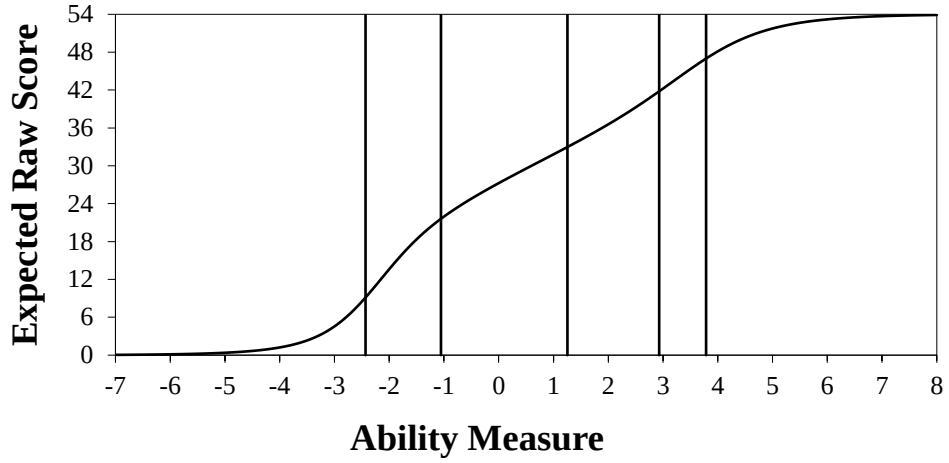
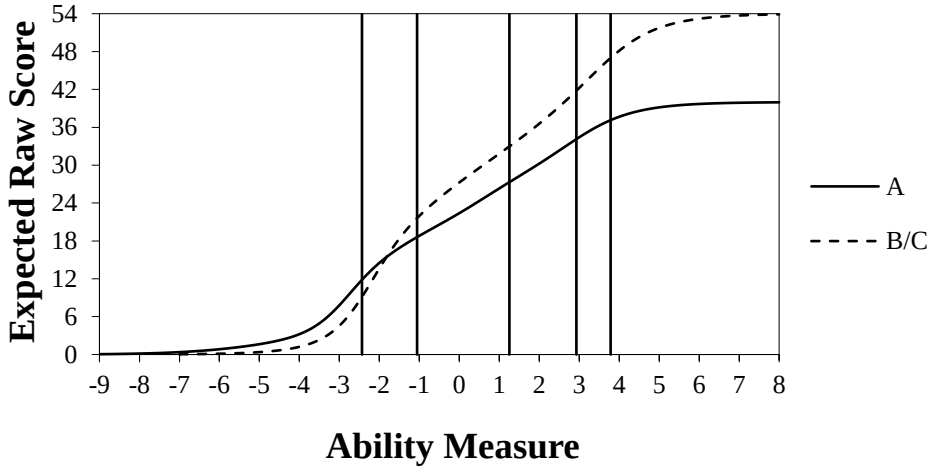
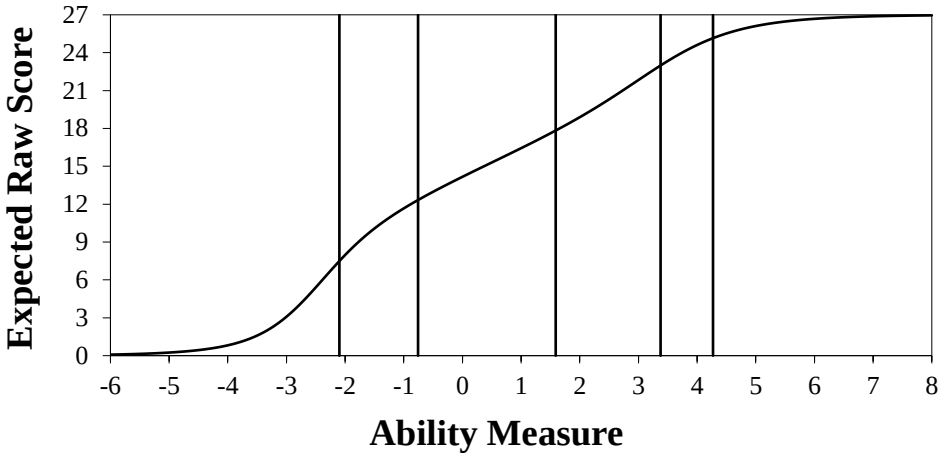


Figure 2.8.3.1.3
Test Characteristic Curve: Writ 1 S501 Paper



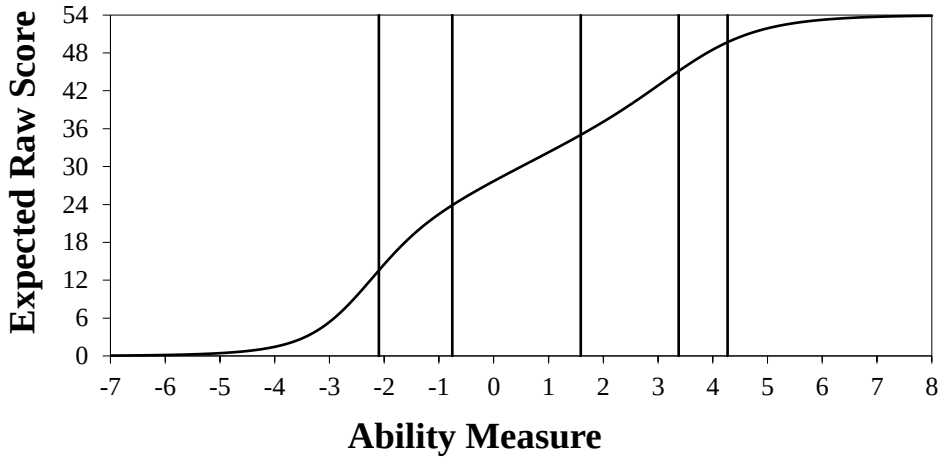
2.8.3.2 Grade 2

Figure 2.8.3.2.1
Test Characteristic Curve: Writ 2 A S501 Paper



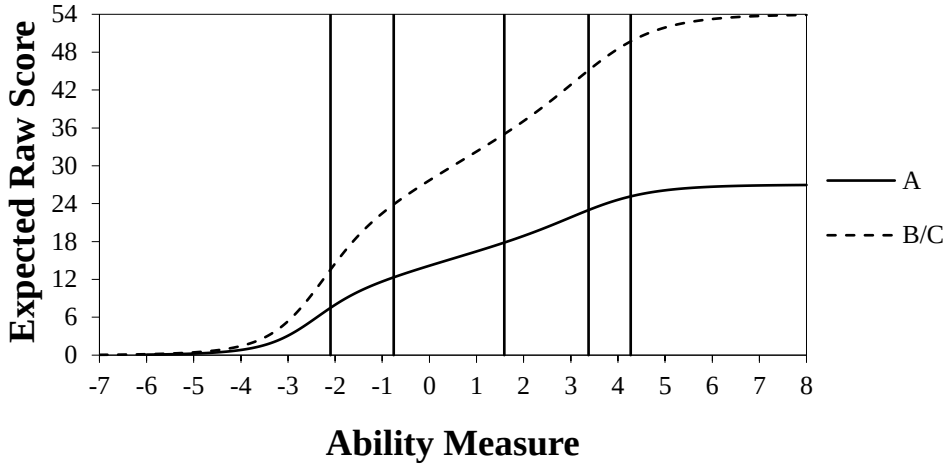
Note: The test form is shared between 2A and 3A.

Figure 2.8.3.2.2
Test Characteristic Curve: Writ 2 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

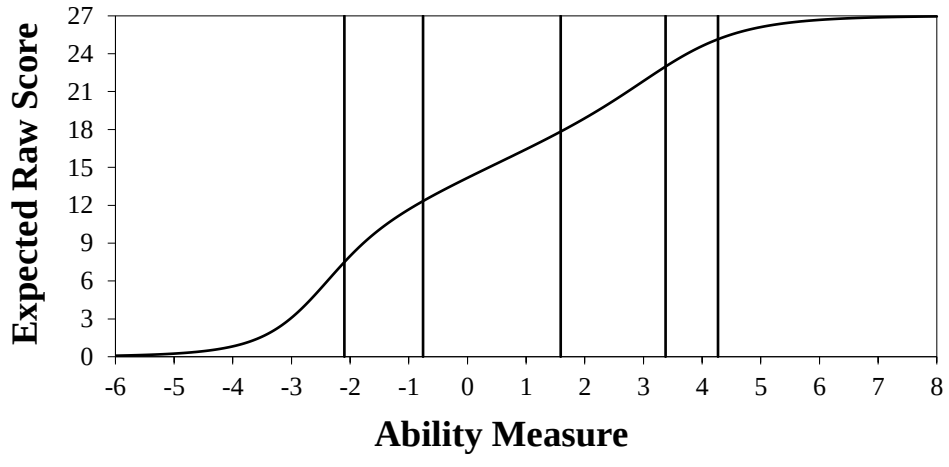
Figure 2.8.3.2.3
Test Characteristic Curve: Writ 2 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

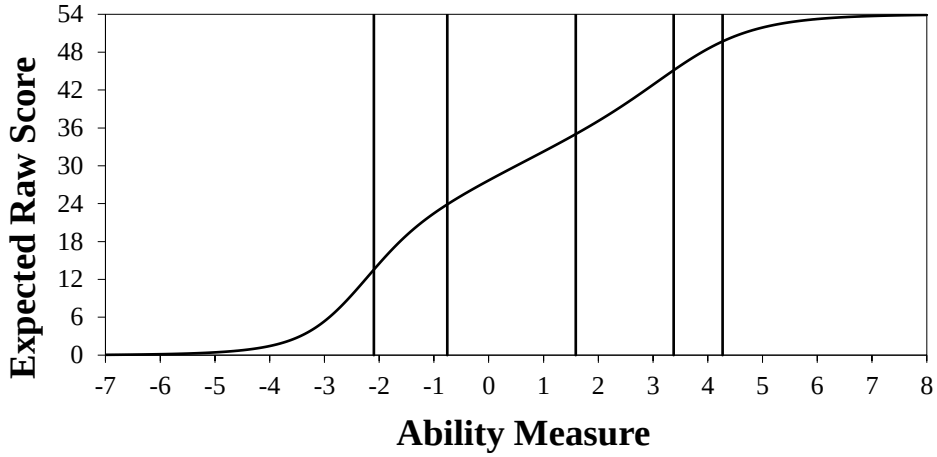
2.8.3.3 Grade 3

Figure 2.8.3.3.1
Test Characteristic Curve: Writ 3 A S501 Paper



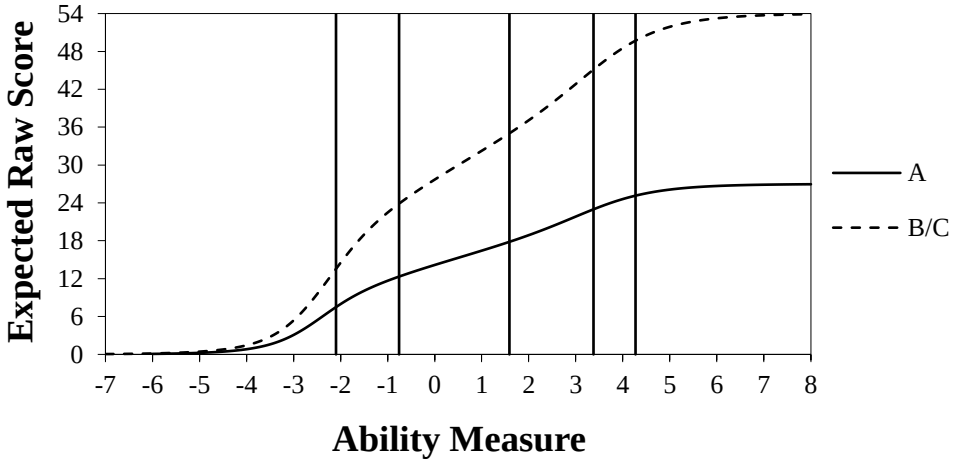
Note: The test form is shared between 2A and 3A.

Figure 2.8.3.3.2
Test Characteristic Curve: Writ 3 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

Figure 2.8.3.3.3
Test Characteristic Curve: Writ 3 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

2.8.3.4 Grades 4–5

Figure 2.8.3.4.1
Test Characteristic Curve: Writ 4-5 A S501 Paper

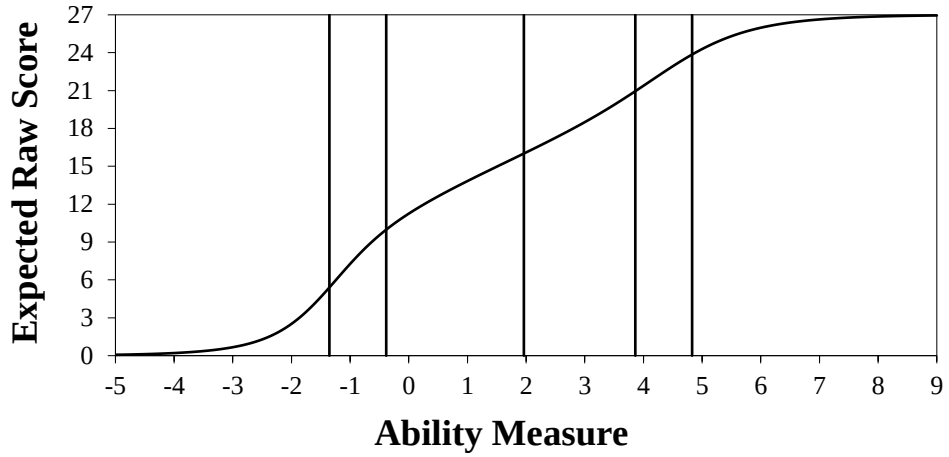


Figure 2.8.3.4.2
Test Characteristic Curve: Writ 4-5 B/C S501 Paper

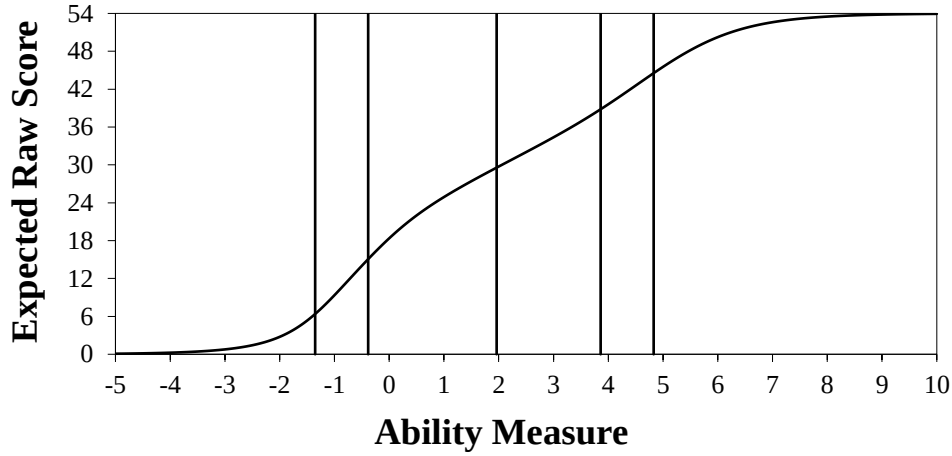
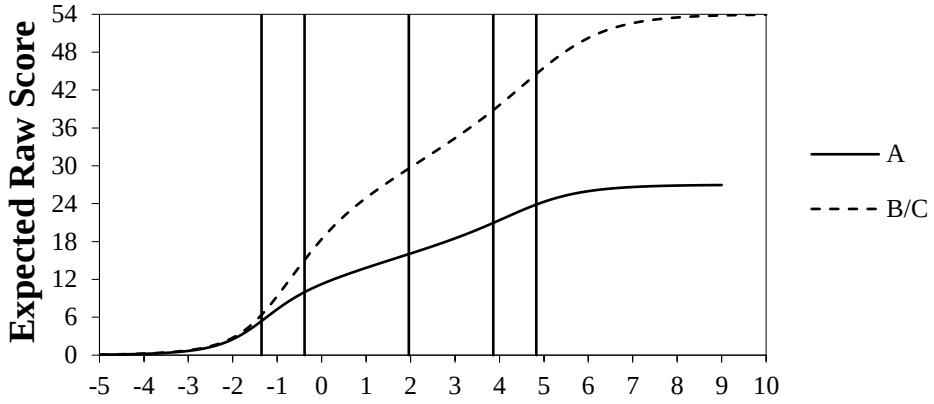


Figure 2.8.3.4.3
Test Characteristic Curve: Writ 4-5 S501 Paper



2.8.3.5 Grades 6–8

Figure 2.8.3.5.1
Test Characteristic Curve: Writ 6-8 A S501 Paper

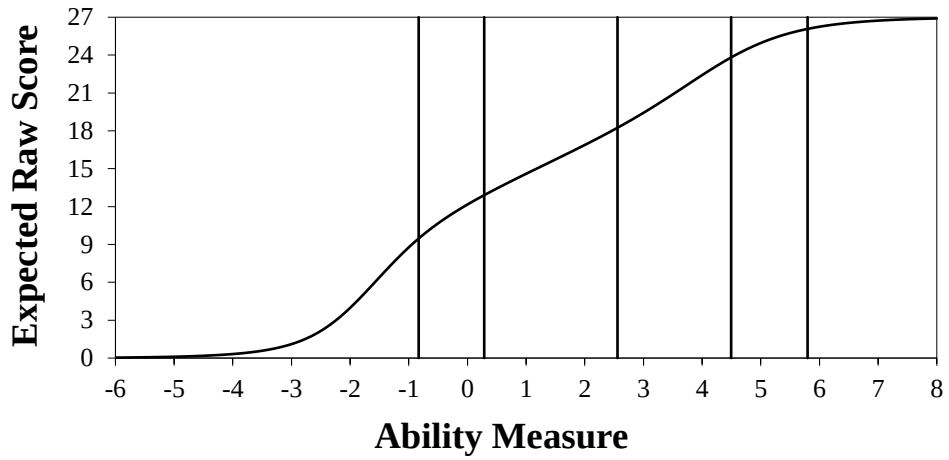


Figure 2.8.3.5.2
Test Characteristic Curve: Writ 6-8 B/C S501 Paper

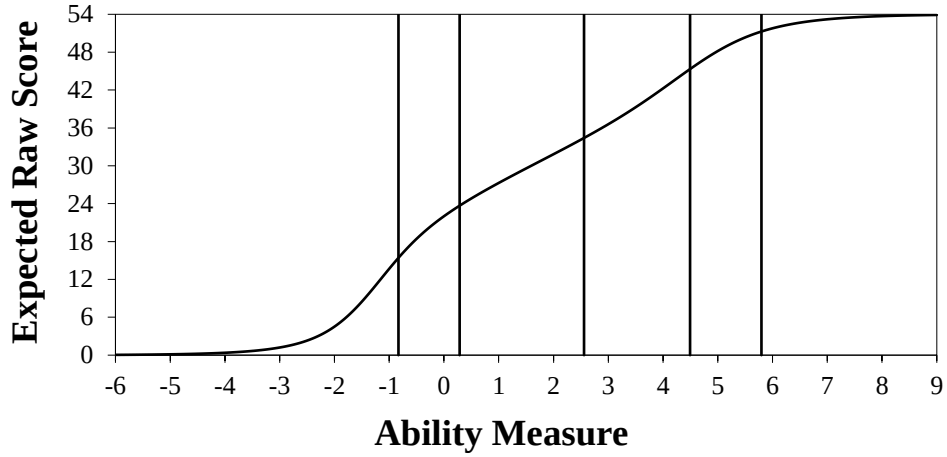
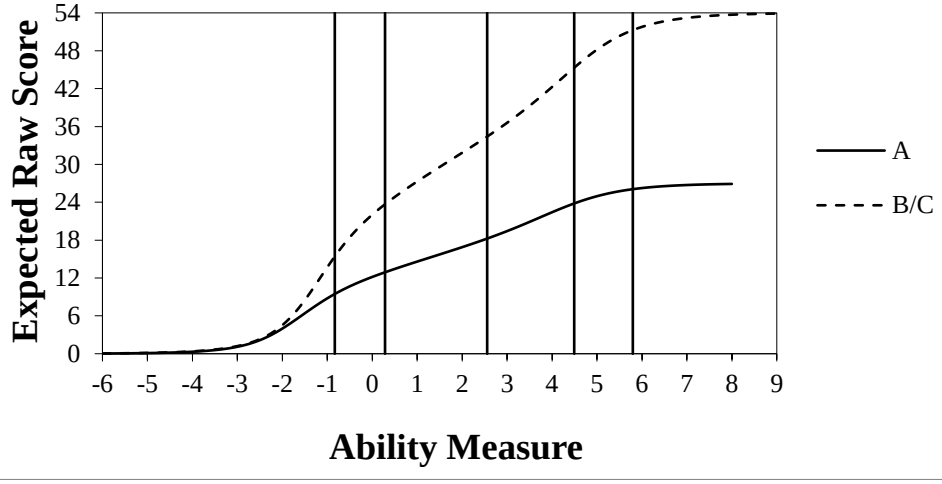


Figure 2.8.3.5.3
Test Characteristic Curve: Writ 6-8 S501 Paper



2.8.3.6 Grades 9–12

Figure 2.8.3.6.1
Test Characteristic Curve: Writ 9-12 A S501 Paper

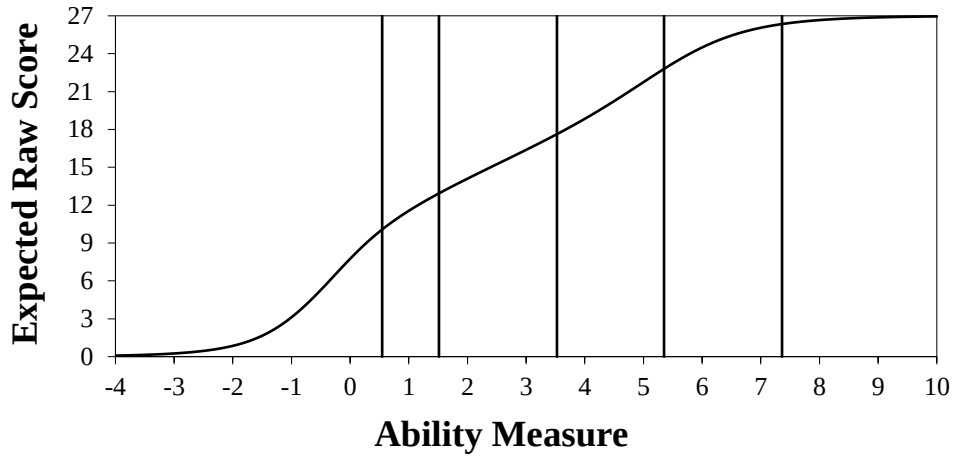


Figure 2.8.3.6.2
Test Characteristic Curve: Writ 9-12 B/C S501 Paper

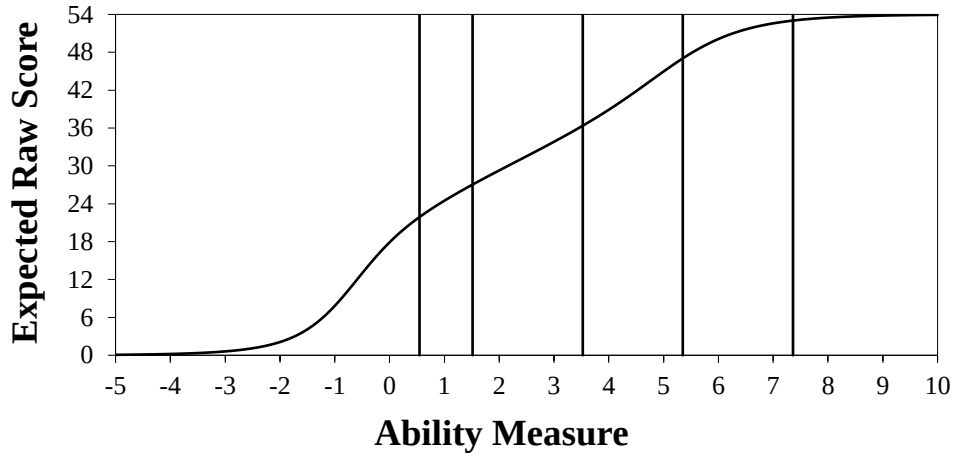
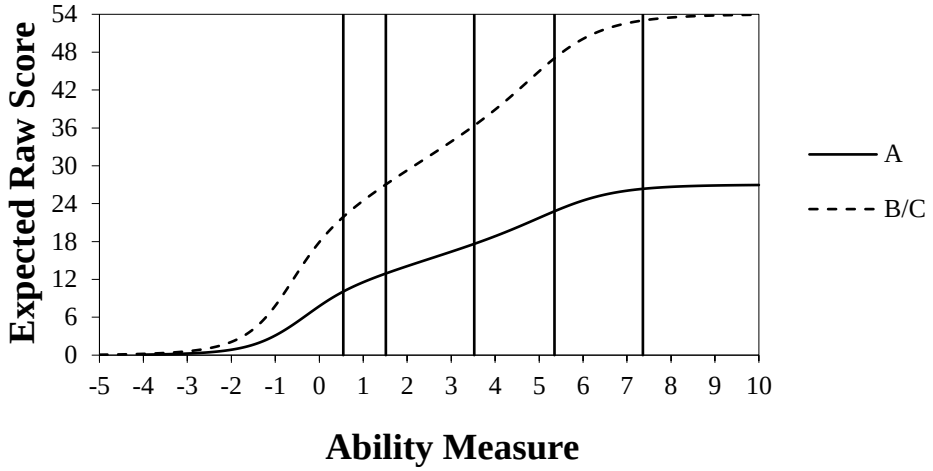


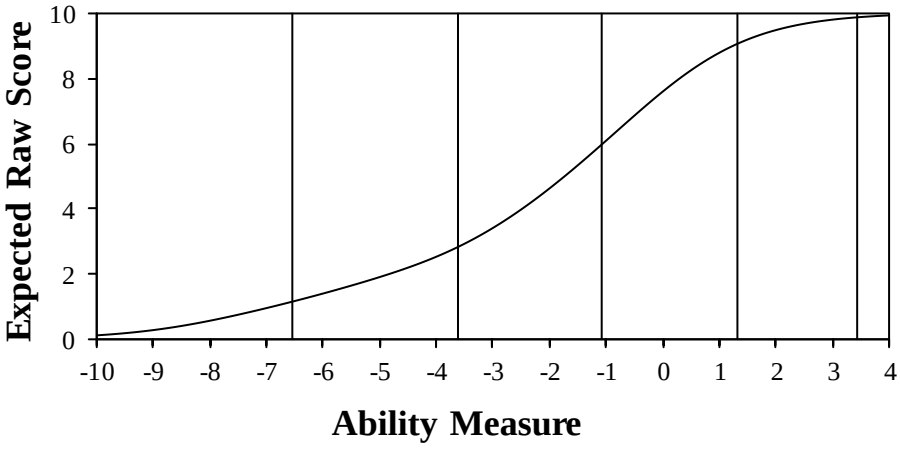
Figure 2.8.3.6.3
Test Characteristic Curve: Writ 9-12 S501 Paper



2.8.4 Speaking

2.8.4.0 Kindergarten

Figure 2.8.4.0
Test Characteristic Curve: Spek K S501 Paper



2.8.4.1 Grade 1

Figure 2.8.4.1.1
Test Characteristic Curve: Spek 1 A S501 Paper

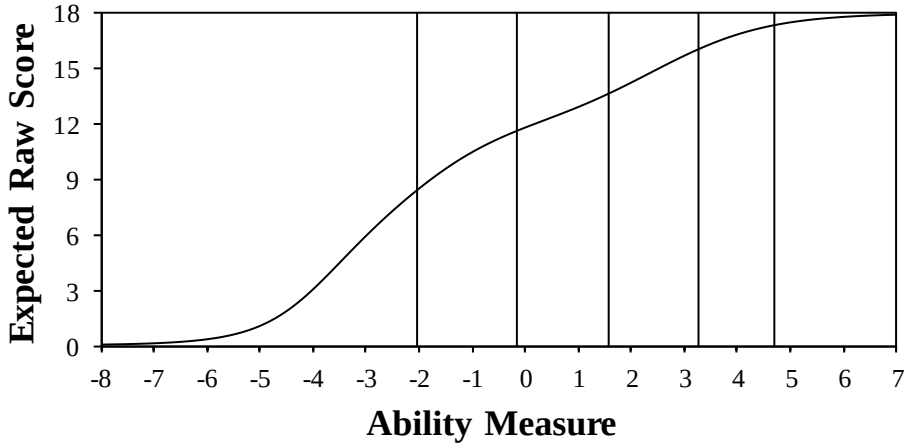


Figure 2.8.4.1.2
Test Characteristic Curve: Spek 1 B/C S501 Paper

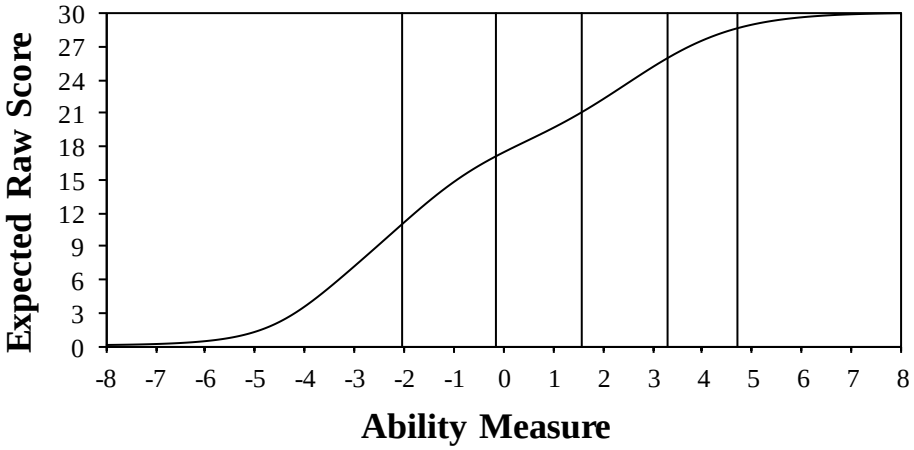
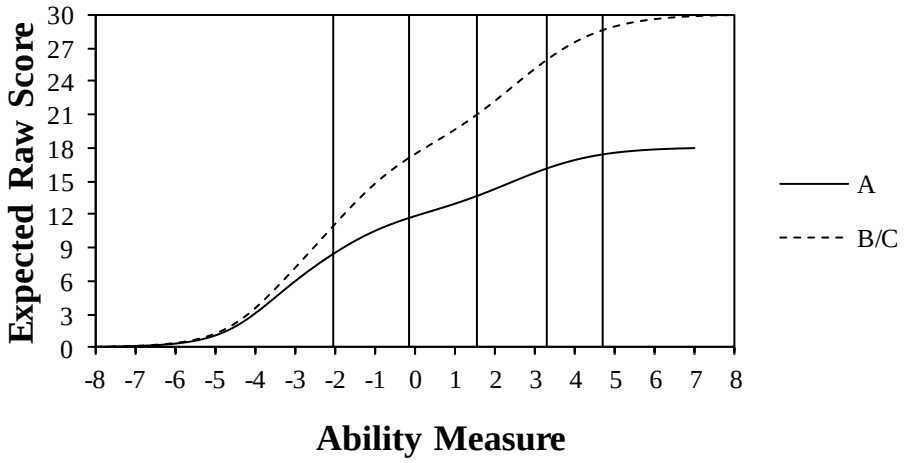
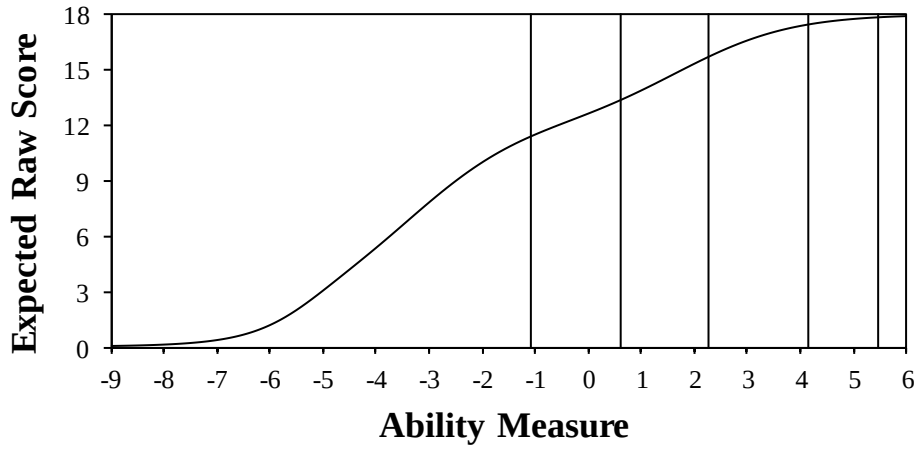


Figure 2.8.4.1.3
Test Characteristic Curve: Spek 1 S501 Paper



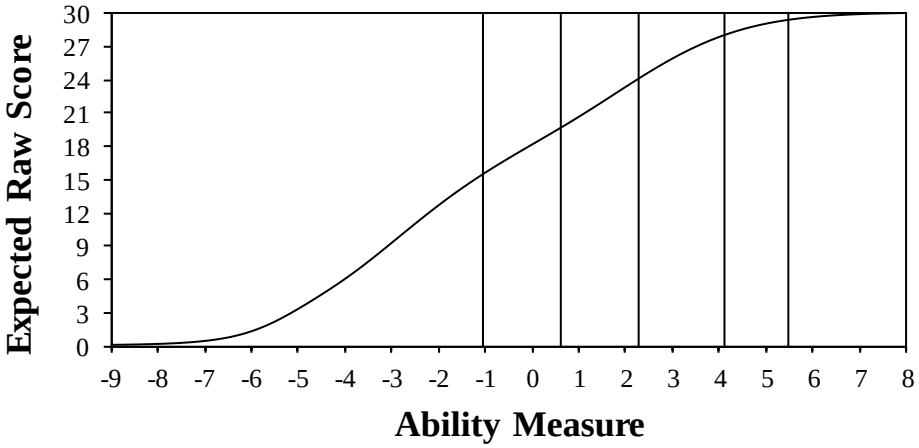
2.8.4.2 Grade 2

Figure 2.8.4.2.1
Test Characteristic Curve: Spek 2 A S501 Paper



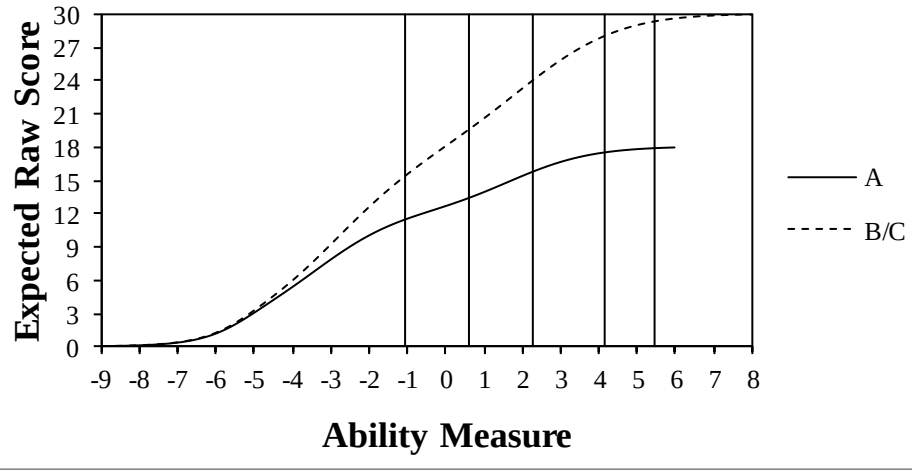
Note: The test form is shared between 2A and 3A.

Figure 2.8.4.2.2
Test Characteristic Curve: Spek 2 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

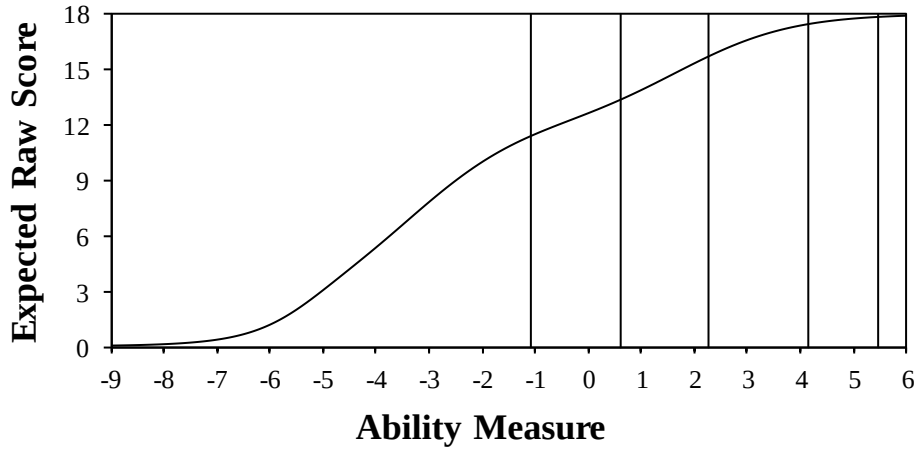
Figure 2.8.4.2.3
Test Characteristic Curve: Spek 2 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

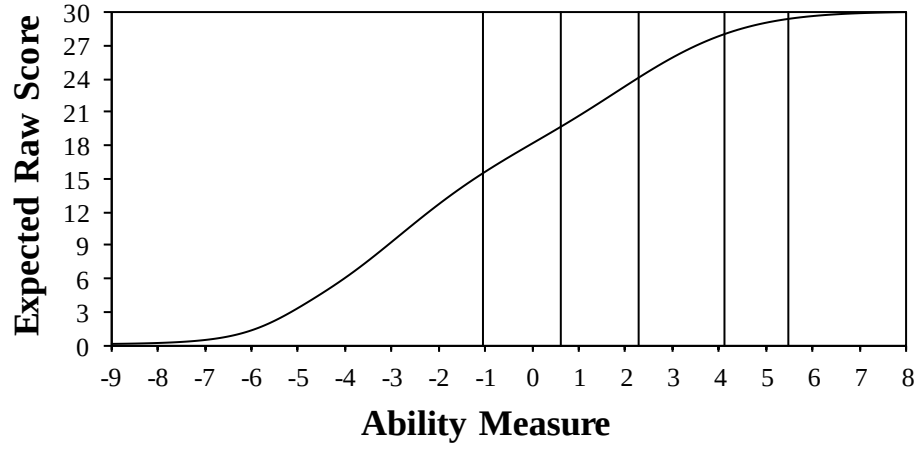
2.8.4.3 Grade 3

Figure 2.8.4.3.1
Test Characteristic Curve: Spek 3 A S501 Paper



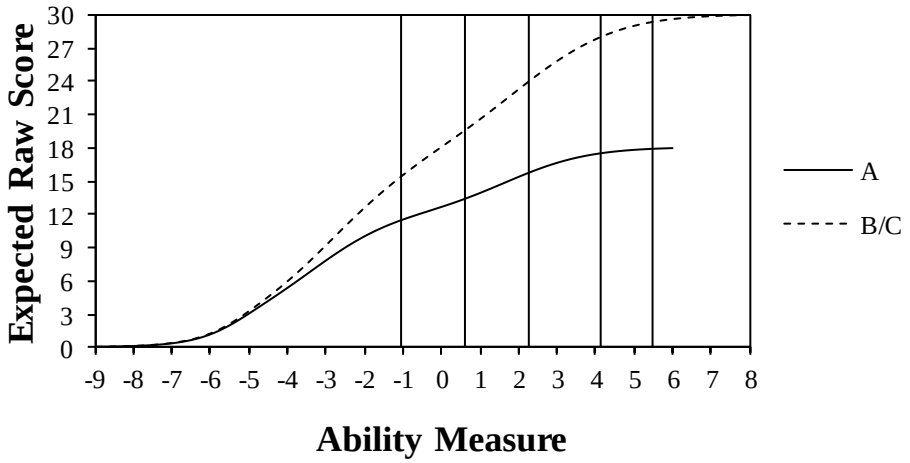
Note: The test form is shared between 2A and 3A.

Figure 2.8.4.3.2
Test Characteristic Curve: Spek 3 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

Figure 2.8.4.3.3
Test Characteristic Curve: Spek 3 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

2.8.4.4 Grades 4–5

Figure 2.8.4.4.1
Test Characteristic Curve: Spek 4-5 A S501 Paper

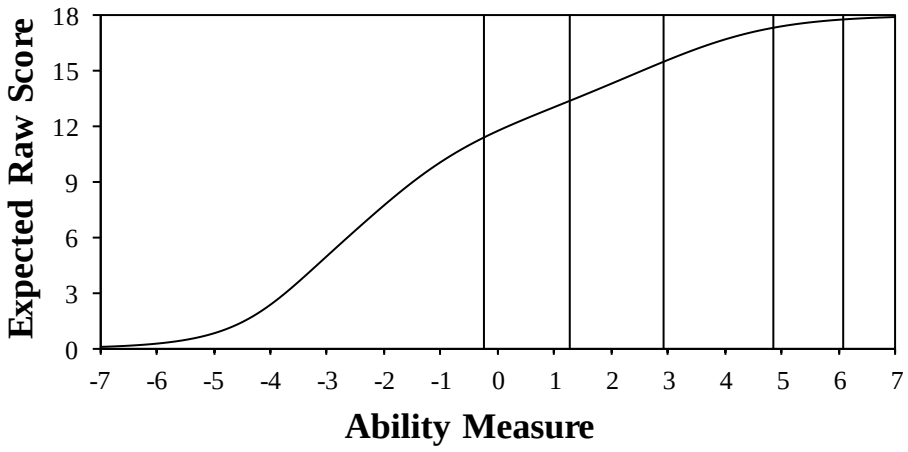


Figure 2.8.4.4.2
Test Characteristic Curve: Spek 4-5 B/C S501 Paper

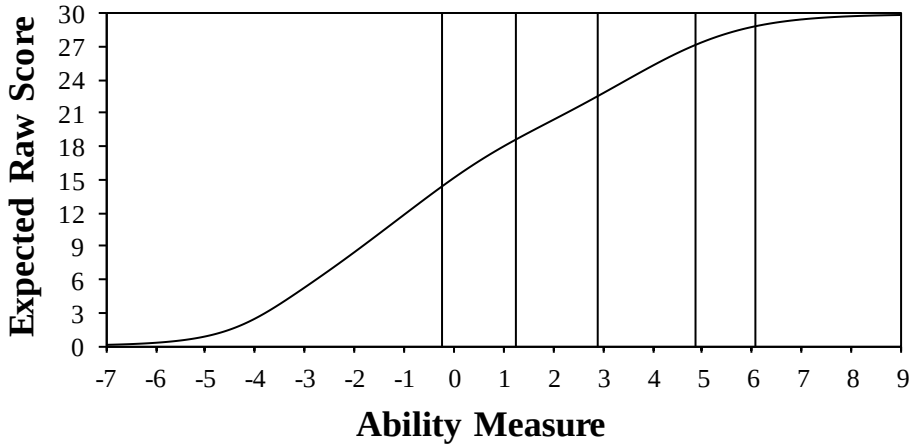
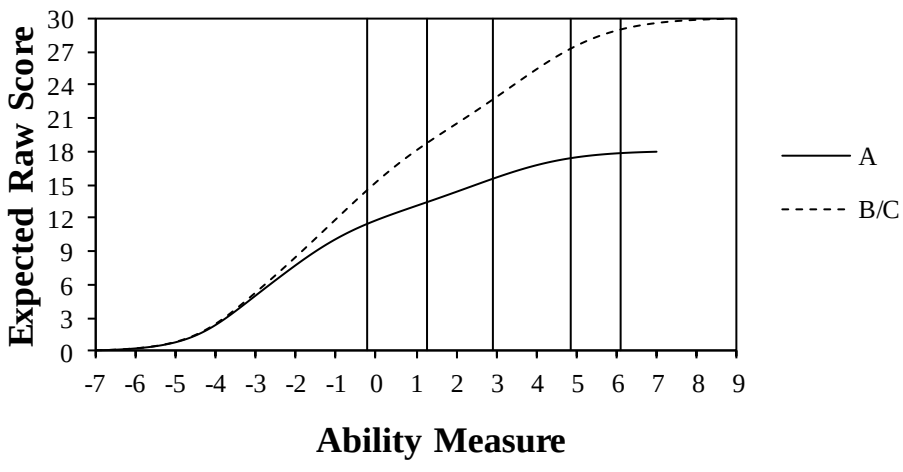


Figure 2.8.4.4.3
Test Characteristic Curve: Spek 4-5 S501 Paper



2.8.4.5 Grades 6–8

Figure 2.8.4.5.1
Test Characteristic Curve: Spek 6-8 A S501 Paper

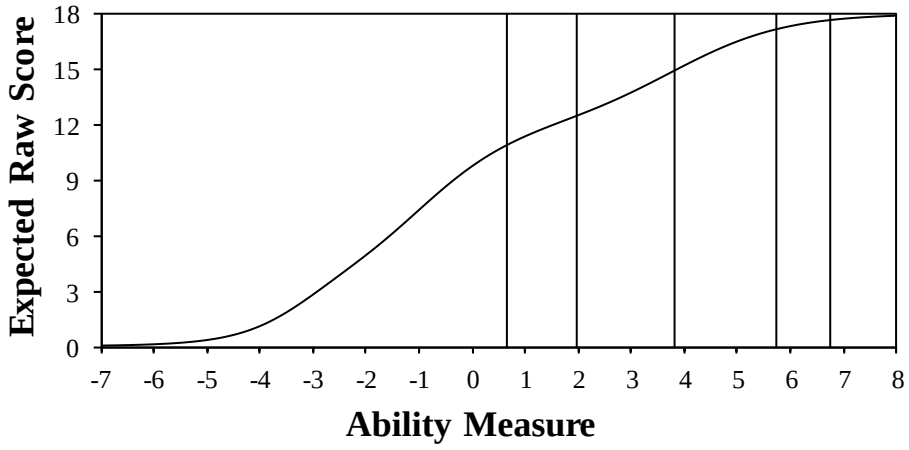


Figure 2.8.4.5.2
Test Characteristic Curve: Spek 6-8 B/C S501 Paper

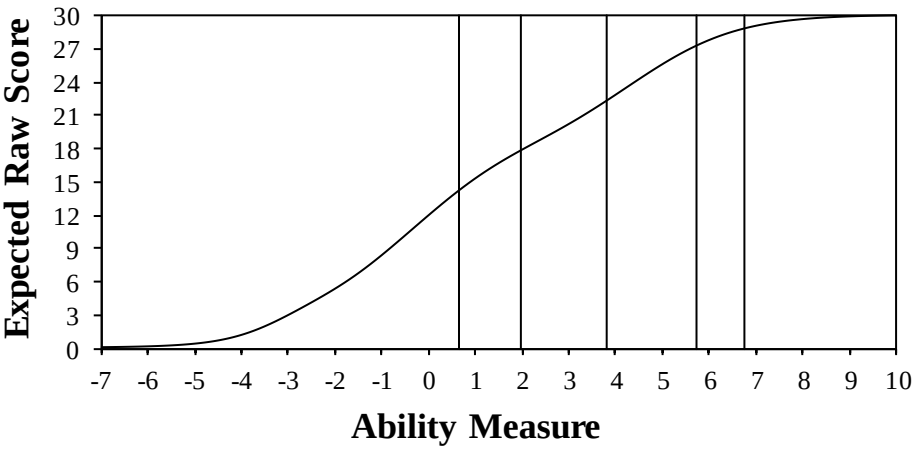
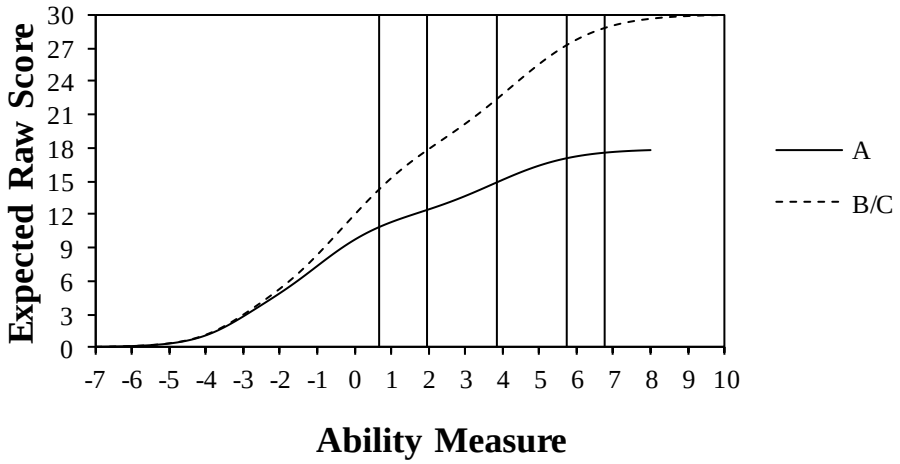


Figure 2.8.4.5.3
Test Characteristic Curve: Spek 6-8 S501 Paper



2.8.4.6 Grades 9–12

Figure 2.8.4.6.1
Test Characteristic Curve: Spek 9-12 A S501 Paper

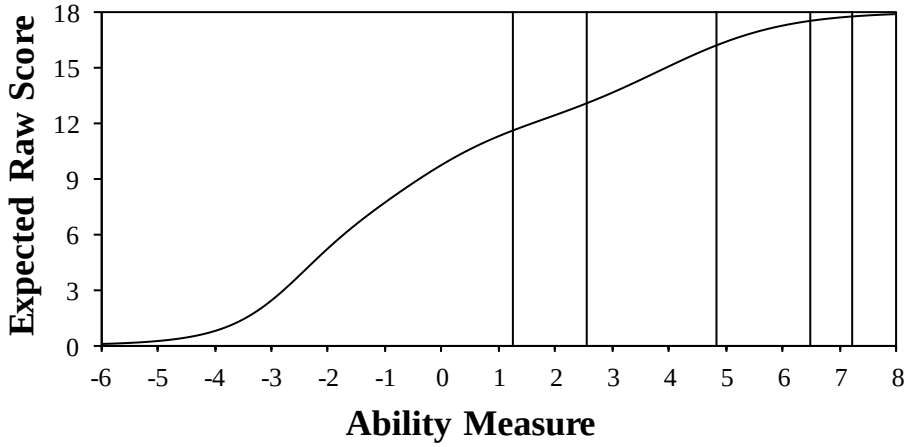


Figure 2.8.4.6.2
Test Characteristic Curve: Spek 9-12 B/C S501 Paper

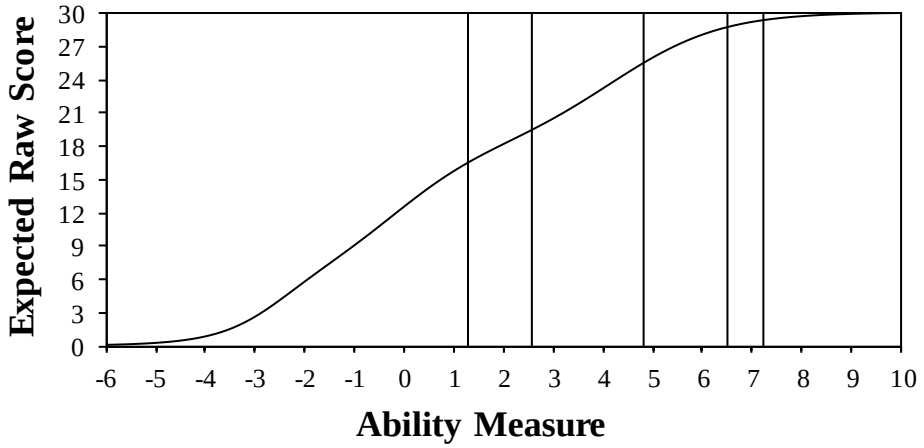
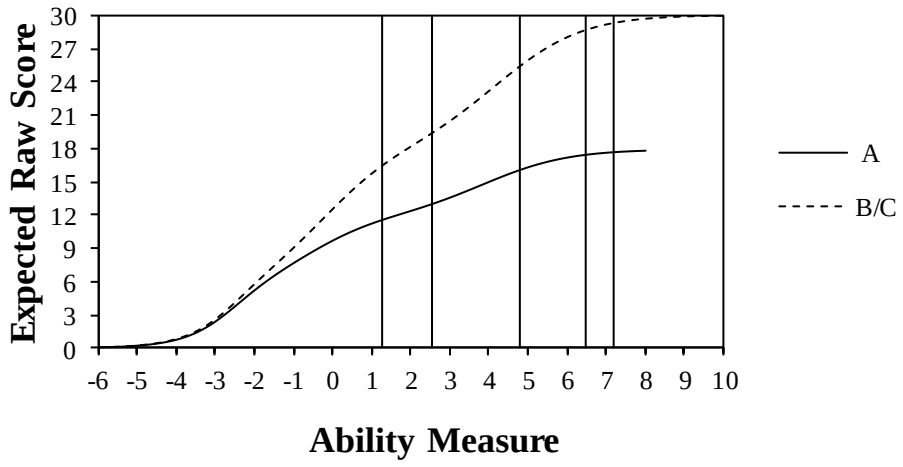


Figure 2.8.4.6.3
Test Characteristic Curve: Spek 9-12 S501 Paper



2.9 Test Information Function

With the Rasch measurement model, as with any measurement model following item response theory, one can use the item information function (Lord, 1980) to model the relationship between the ability measure (in logits) and the accuracy of the ability measure by item. The item information function indicates the amount of information we have about the ability estimate provided by the item, as a function of the ability level. The more information we have about the ability estimate, the more certain or confident we are about the ability estimate. If the amount of information is large, that means the student whose true ability is at that level is estimated with a higher degree of certainty, and all the estimates will be reasonably close to the true values. Conversely, if the amount of information is small, that means the student whose true ability is at that level is estimated with a lower degree of certainty and estimates will be further away from the true values. Mathematically, the amount of item information at a given ability level is the reciprocal of the variance of the ability estimate at the level for the item. In other words, item information value is the inverse squared of the standard errors of measurement of a given ability measure for the item. Therefore item information is also said to provide information about the precision of the ability estimate along the ability continuum provided by the item.

The test information function (TIF) aggregates the item information functions across all the items on the test form or item pool. Since the item information value is the inverse squared of the standard errors of measurement of a given ability measure for the item, the test information value reflects the standard errors of measurement of a given ability level for the test. When the TIF is presented graphically as the test information curve, it shows how well the test is measuring across the continuum of student ability in terms of the amount of information, certainty, or the amount of measurement precision the test provides at each ability level. The higher the curve, the more information the test provides at the ability level.

Since the TIF is the sum of all item characteristic functions on the test form (Lord, 1980), the TIF depends on the item information functions (Lord, 1980) of the items on the test form or in the item pool. The shape of the test information curve depends on several factors, including the number and characteristics of items, the item response theory model used, and the values of the item parameters. With some exceptions, there is a general pattern to the shape of test information curves. Test information curves peak at the area where the test provides higher discrimination and better measurement as compared to other areas where the curve is less peaked, normally at the lower and upper ends of the ability continuum. When the test form consists of multiple-choice items such as on the Listening and Reading domains, the test information is usually unimodal. The shape of test information curves for Writing and Speaking tests, which consist of polytomous tasks, are affected by the values of the item category parameters in addition to the factors mentioned earlier. Since polytomous tasks have more score categories than multiple-choice items and measure a wider range of values on the proficiency scale, adjacent category boundaries are sometimes far apart as a result. In this situation, a test information curve will have a dip in the area between the adjacent category boundaries indicating the loss of information in

this ability range. Therefore the shape of a test information curve for ACCESS Writing and Speaking tests may not be unimodal and instead may have one or more peaks. This is consistent with other tests with polytomous items, such as the National Assessment of Educational Progress Writing assessment (Muraki, 1993).

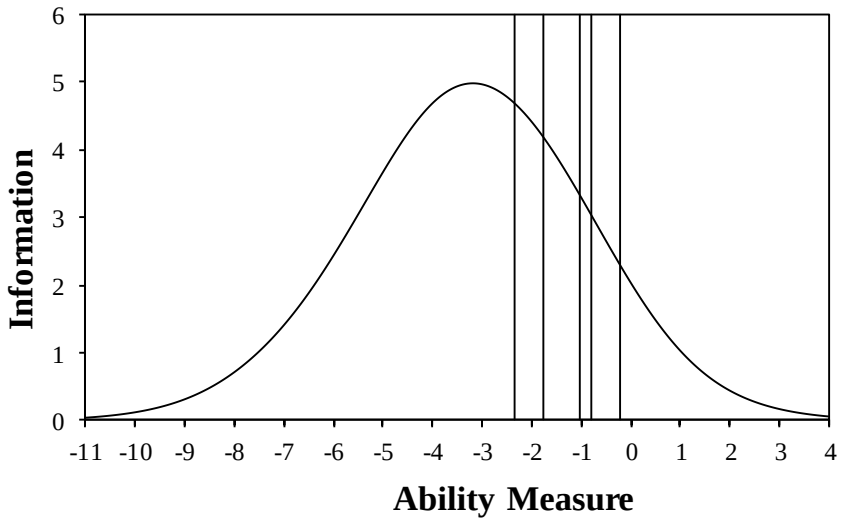
The figures in this section plot the TIF and show graphically the amount of information provided by the test across the continuum of student ability. Five vertical lines in the figure indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1–6) for the domain being tested. The ACCESS cut score lines are presented along with the TIF 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 the Kindergarten and Tier A tests in some domains, it was not possible to place into all six proficiency levels.

In addition to the TIF graphs by tier, we provide plots of the TIFs across tiers, by grade cluster, on the same graph. It is informative to compare the ability ranges where the curves are peaked (where the best measurement information is provided) across tiers. For example, the TIF across tiers for Listening Grade 1 shows that the Listening Grade 1 Tier A form provides the most information right below PL 2, while the Listening Grade 1 Tier B/C form provides the most information at a higher proficiency level (right below PL 3), as expected. In addition, the plot shows that the Listening Grade 1 Tier A form provides more information than the B/C form before the PL 2 cut, while the B/C form provides more information than the Tier A form after the PL 2 cut.

2.9.1 Listening

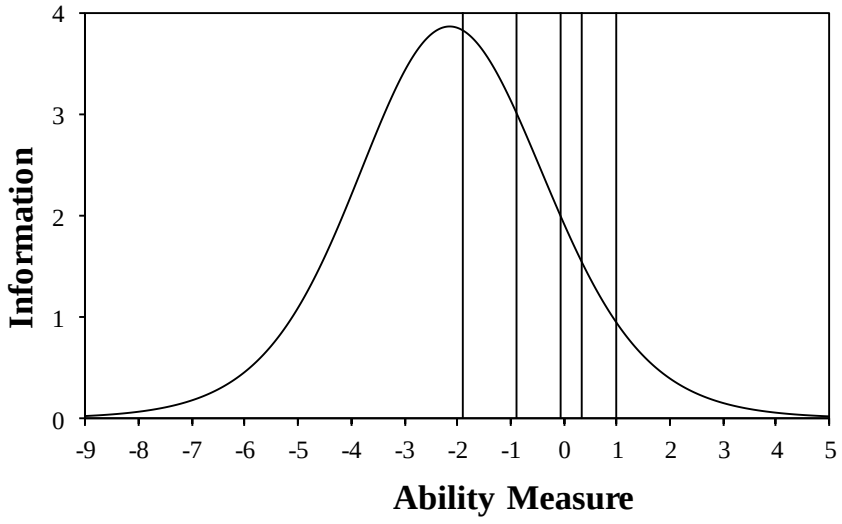
2.9.1.0 Kindergarten

Figure 2.9.1.0
Test Information Function: List K S501 Paper



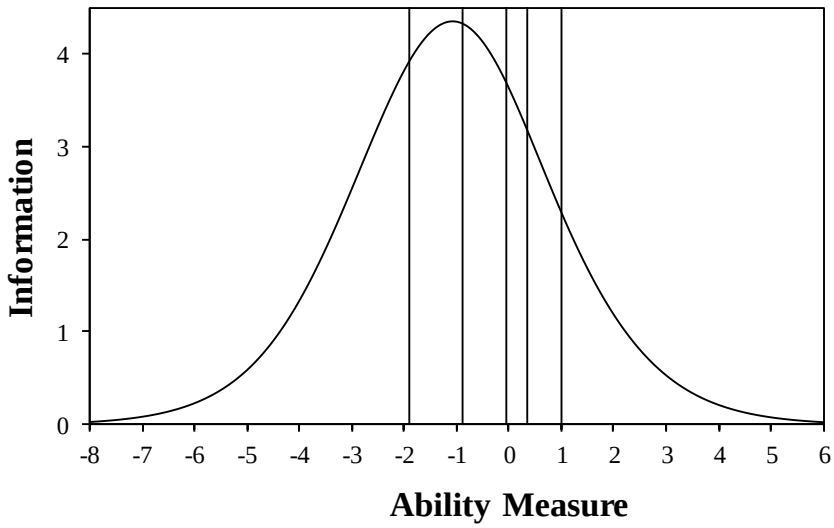
2.9.1.1 Grade 1

Figure 2.9.1.1.1
Test Information Function: List 1A S501 Paper



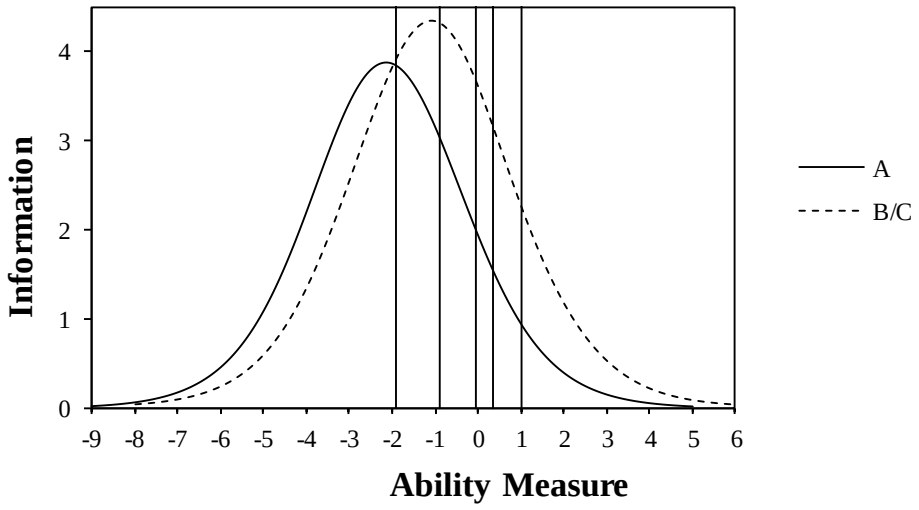
Note: The test form is shared between 1A and 2A.

Figure 2.9.1.1.2
Test Information Function: List 1B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

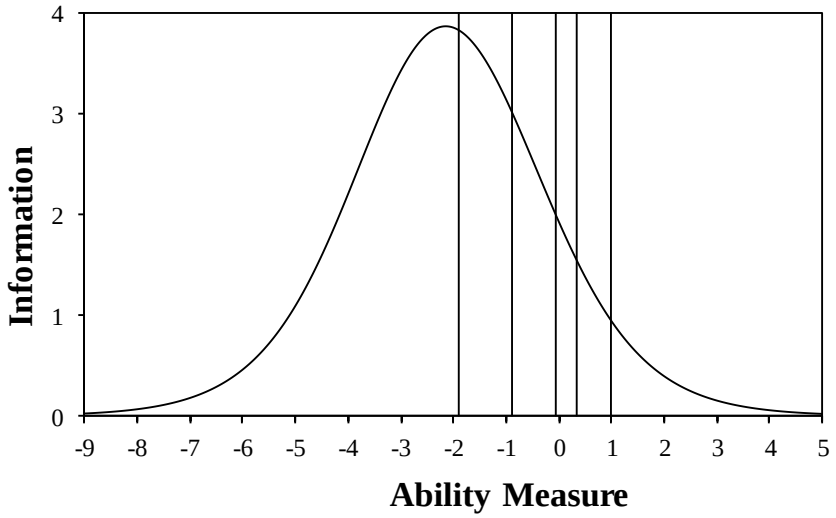
Figure 2.9.1.1.3
Test Information Function: List 1 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

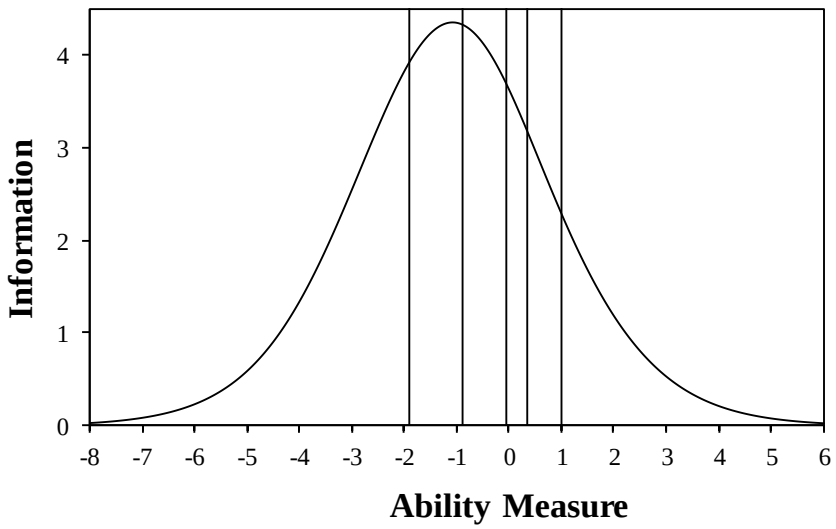
2.9.1.2 Grade 2

Figure 2.9.1.2.1
Test Information Function: List 2A S501 Paper



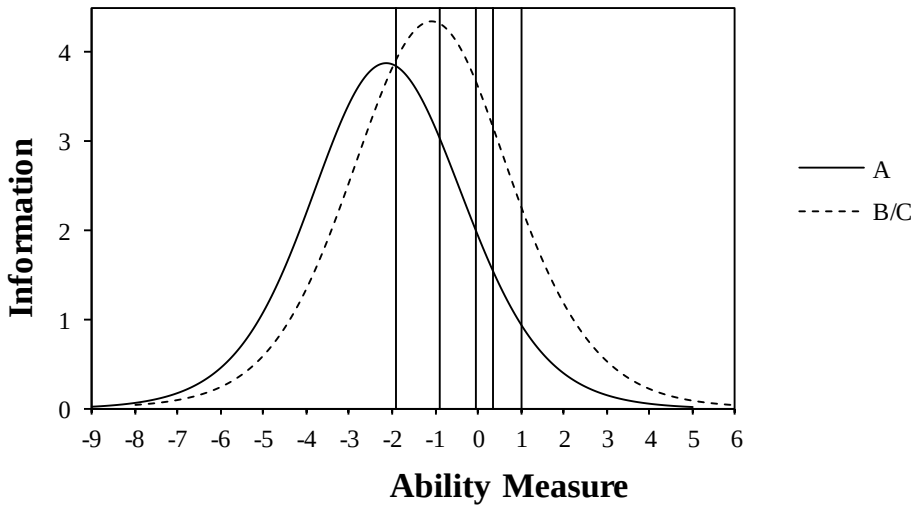
Note: The test form is shared between 1A and 2A.

Figure 2.9.1.2.2
Test Information Function: List 2B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

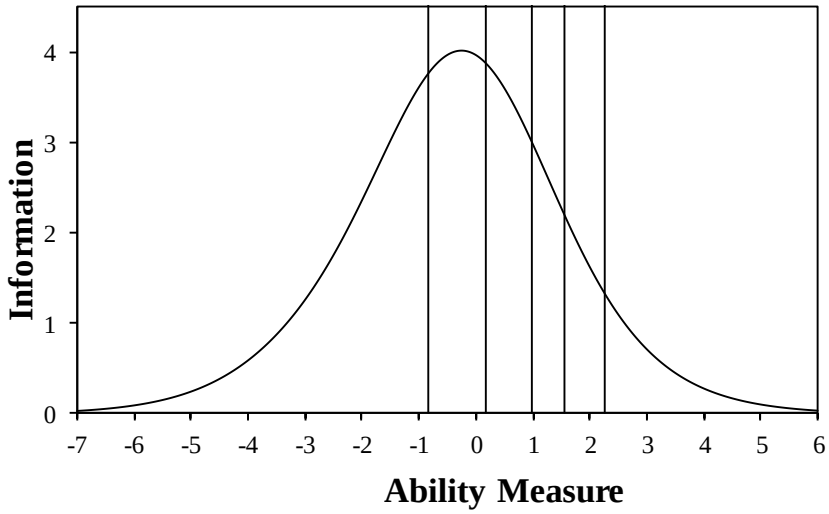
Figure 2.9.1.2.3
Test Information Function: List 2 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

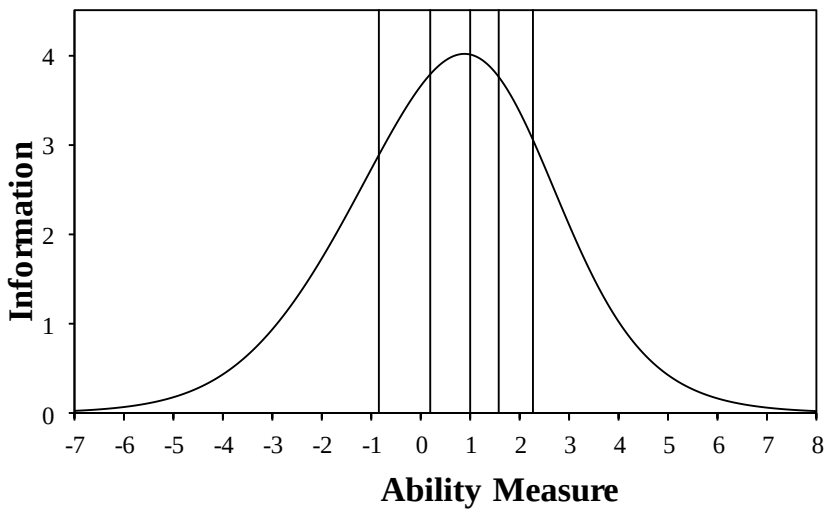
2.9.1.3 Grade 3

Figure 2.9.1.3.1
Test Information Function: List 3A S501 Paper



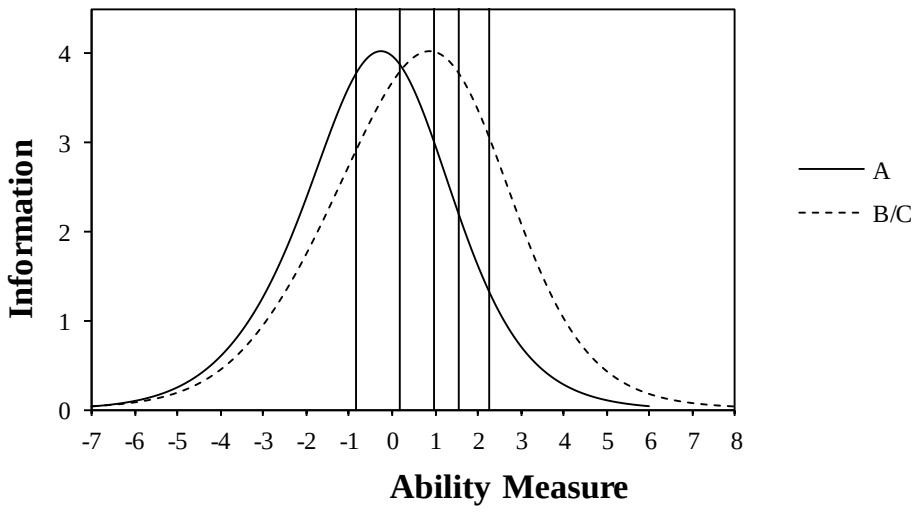
Note: The test form is shared between 3A and 4–5A.

Figure 2.9.1.3.2
Test Information Function: List 3B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

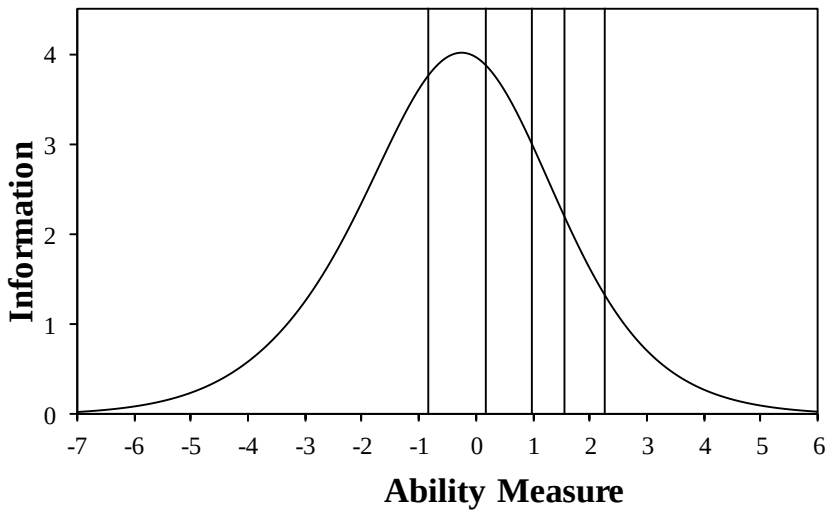
Figure 2.9.1.3.3
Test Information Function: List 3 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

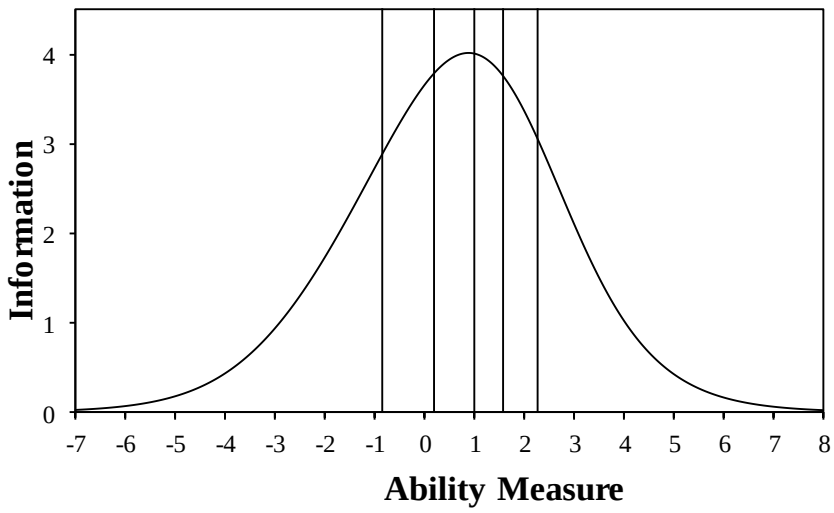
2.9.1.4 Grades 4–5

Figure 2.9.1.4.1
Test Information Function: List 4-5A S501 Paper



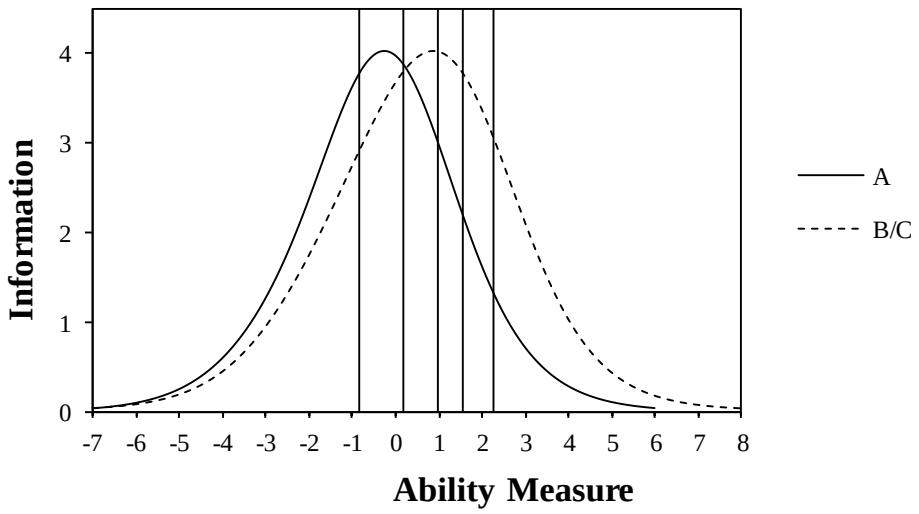
Note: The test form is shared between 3A and 4–5A.

Figure 2.9.1.4.2
Test Information Function: List 4-5B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

Figure 2.9.1.4.3
Test Information Function: List 4-5 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

2.9.1.5 Grades 6–8

Figure 2.9.1.5.1
Test Information Function: List 6-8A S501 Paper

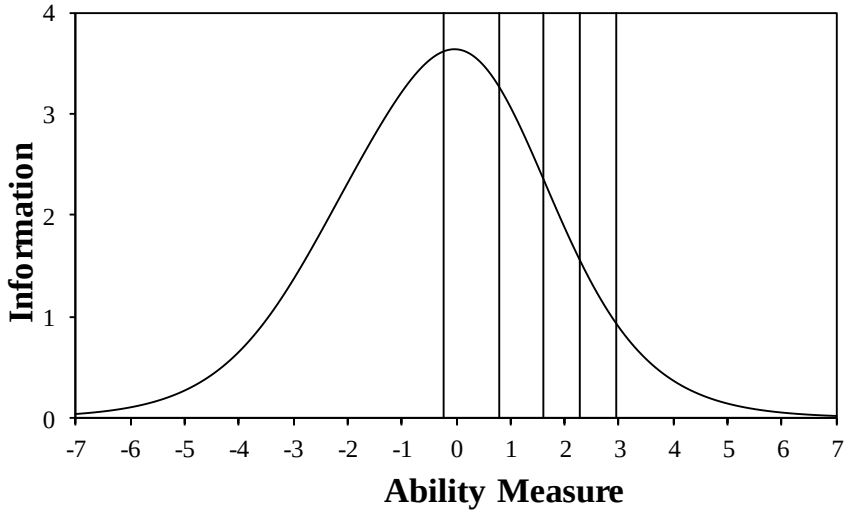


Figure 2.9.1.5.2
Test Information Function: List 6-8B/C S501 Paper

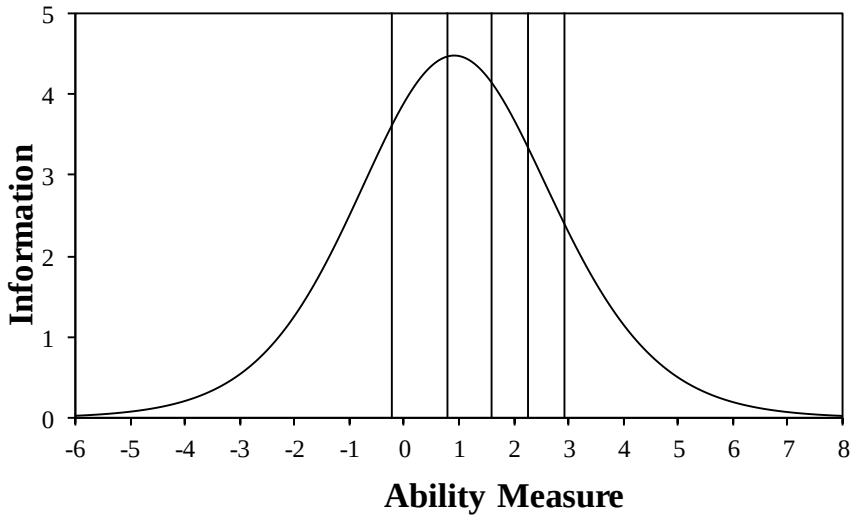
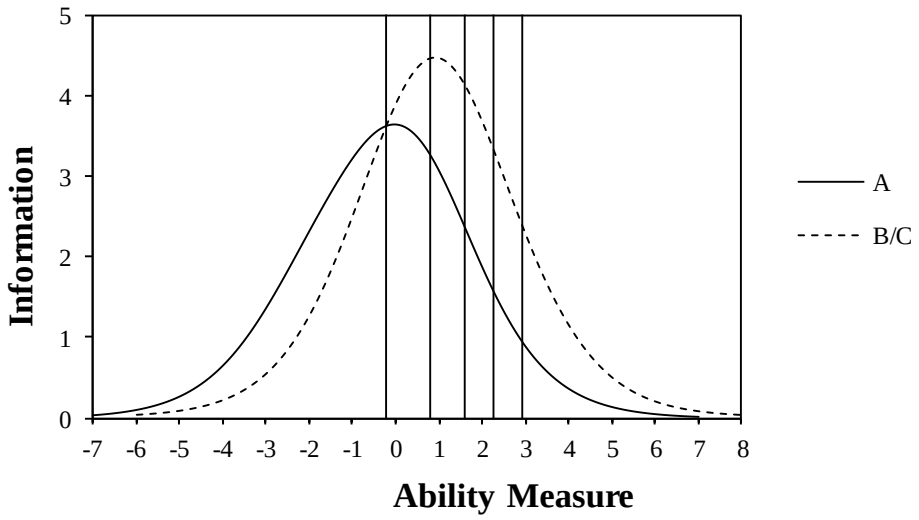


Figure 2.9.1.5.3
Test Information Function: List 6-8 S501 Paper



2.9.1.6 Grades 9–12

Figure 2.9.1.6.1
Test Information Function: List 9-12A S501 Paper

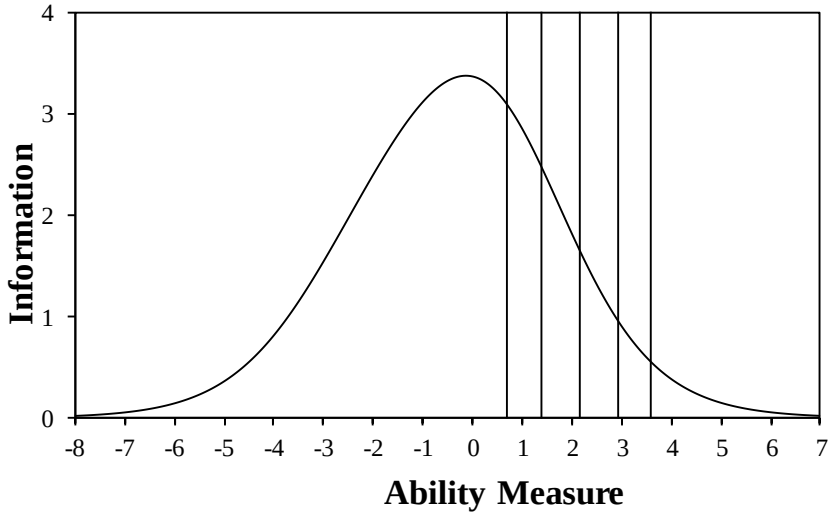


Figure 2.9.1.6.2
Test Information Function: List 9-12B/C S501 Paper

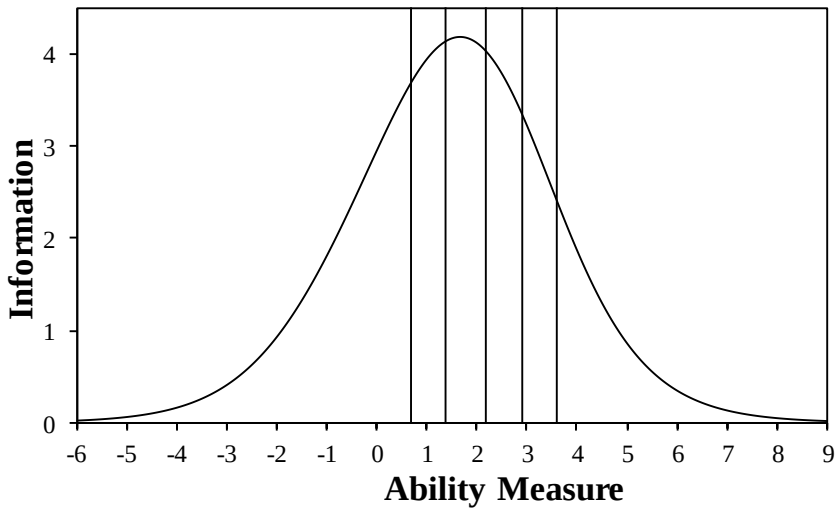
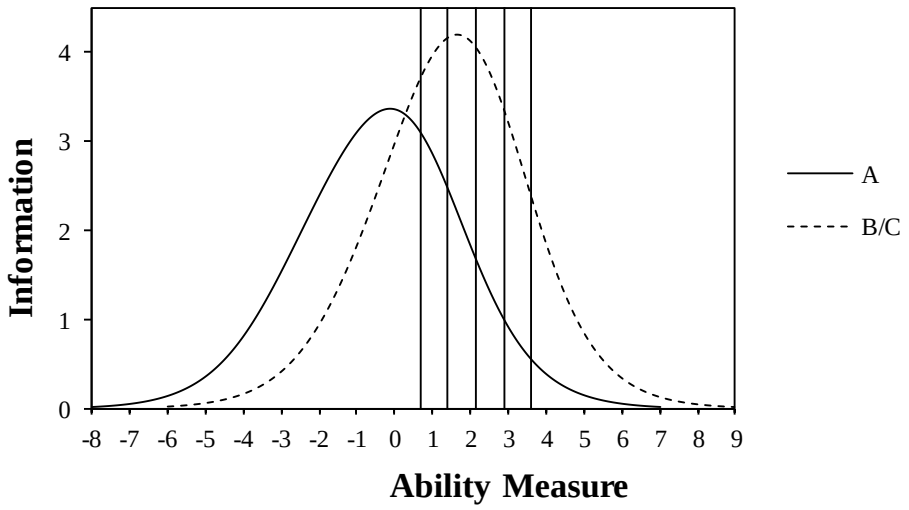


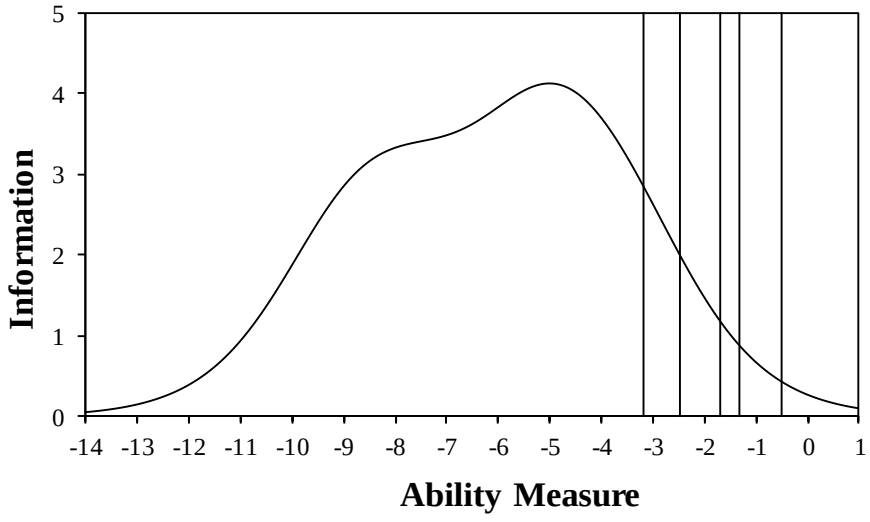
Figure 2.9.1.6.3
Test Information Function: List 9-12 S501 Paper



2.9.2 Reading

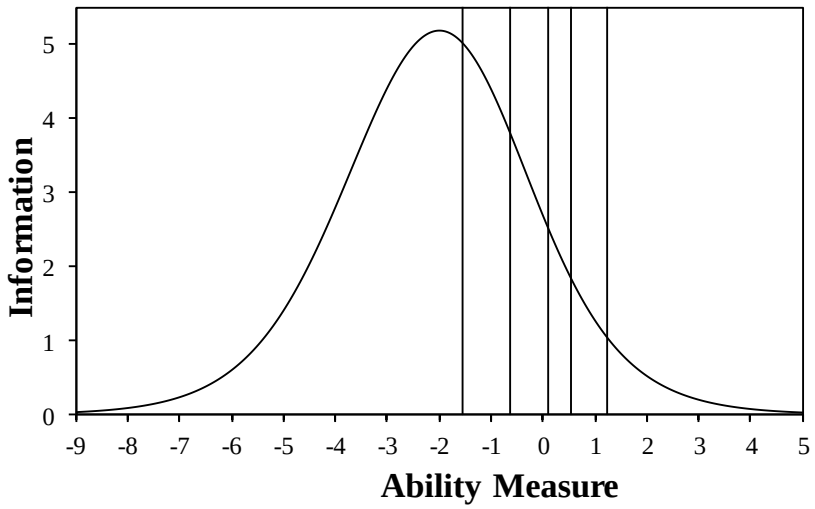
2.9.2.0 Kindergarten

Figure 2.9.2.0
Test Information Function: Read K S501 Paper



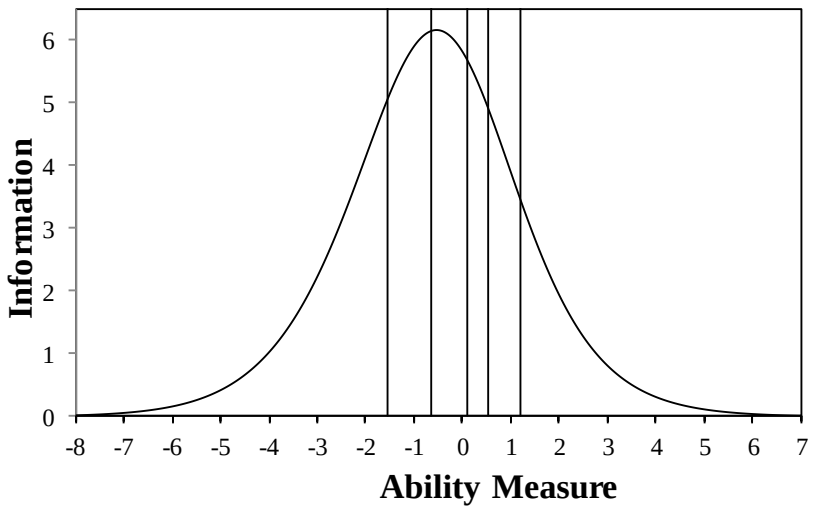
2.9.2.1 Grade 1

Figure 2.9.2.1.1
Test Information Function: Read 1A S501 Paper



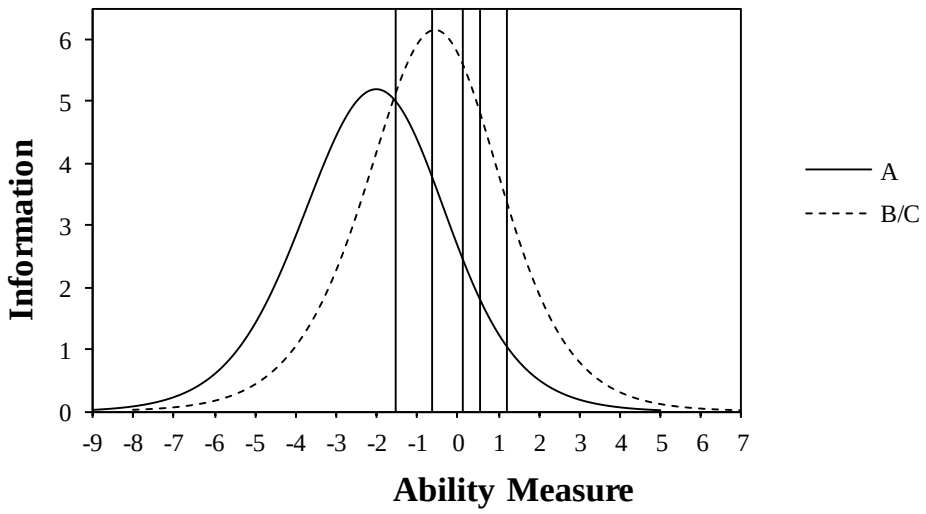
Note: The test form is shared between 1A and 2A.

Figure 2.9.2.1.2
Test Information Function: Read 1B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

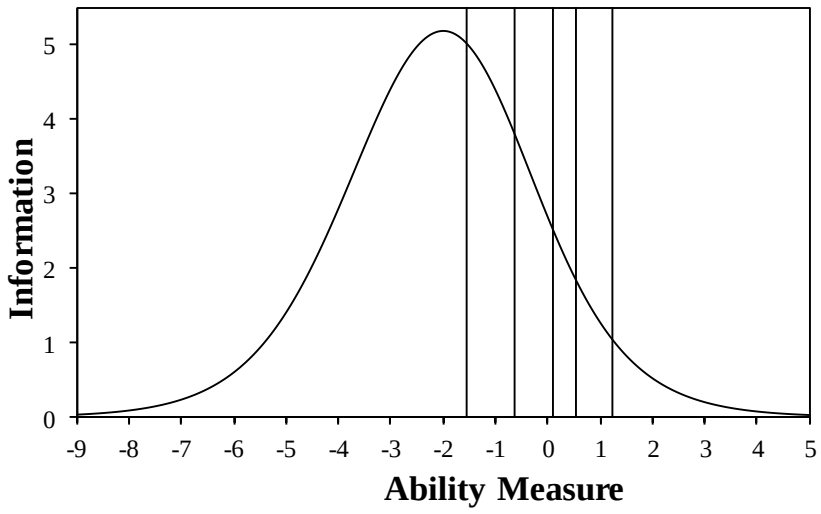
Figure 2.9.2.1.3
Test Information Function: Read 1 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

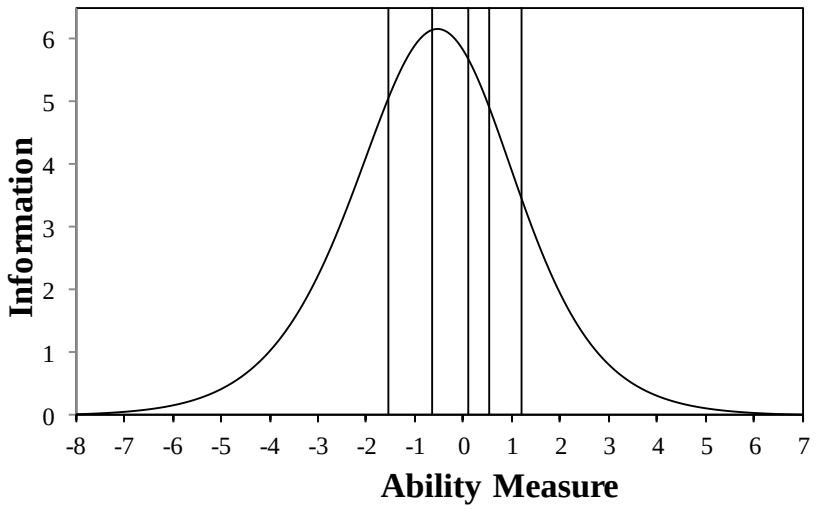
2.9.2.2 Grade 2

Figure 2.9.2.2.1
Test Information Function: Read 2A S501 Paper



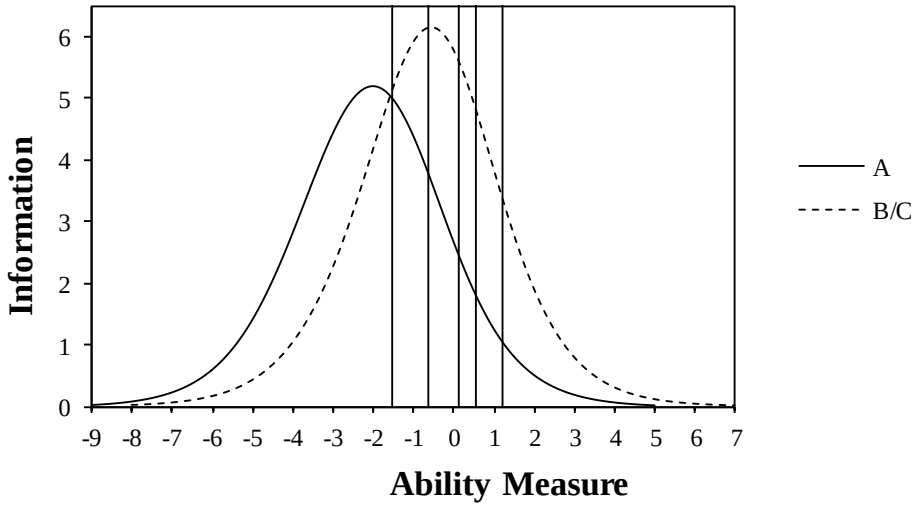
Note: The test form is shared between 1A and 2A.

Figure 2.9.2.2.2
Test Information Function: Read 2B/C S501 Paper



Note: The test form is shared between 1B/C and 2B/C.

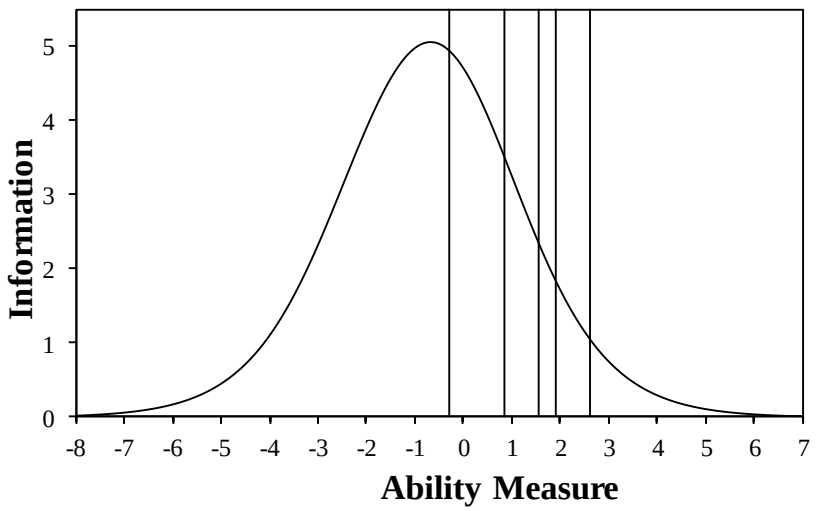
Figure 2.9.2.2.3
Test Information Function: Read 2 S501 Paper



Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.

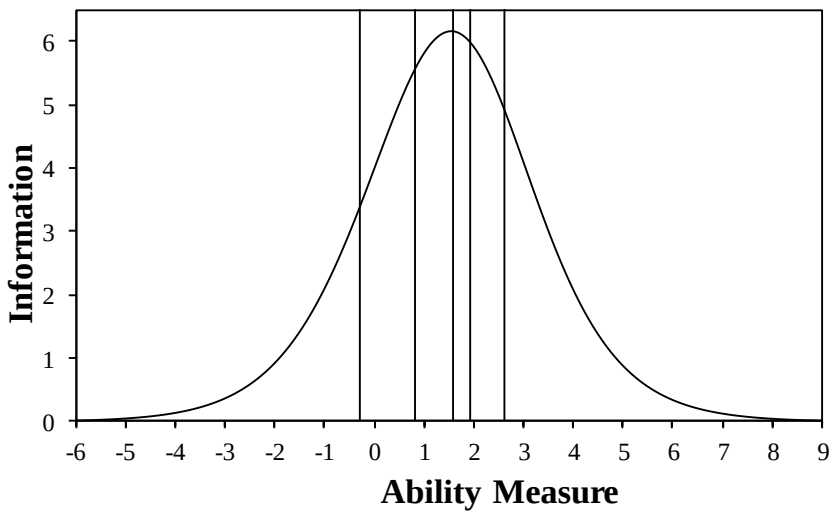
2.9.2.3 Grade 3

Figure 2.9.2.3.1
Test Information Function: Read 3A S501 Paper



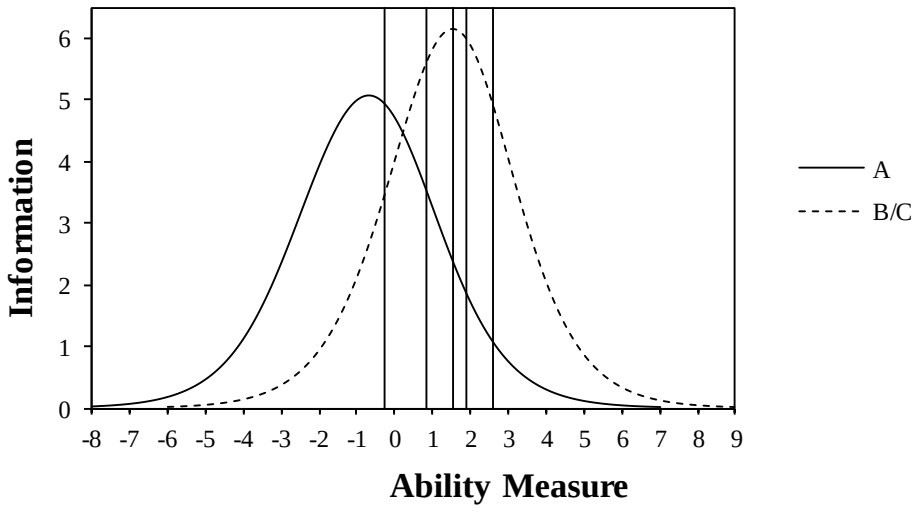
Note: The test form is shared between 3A and 4–5A.

Figure 2.9.2.3.2
Test Information Function: Read 3B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

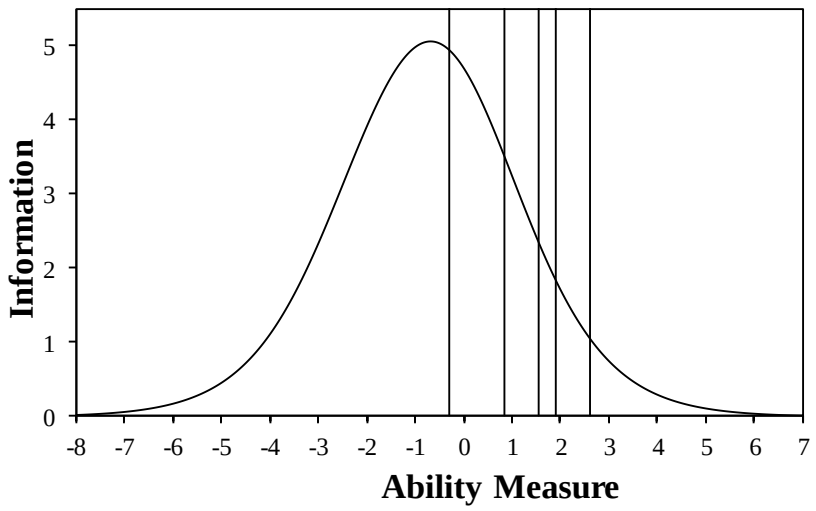
Figure 2.9.2.3.3
Test Information Function: Read 3 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

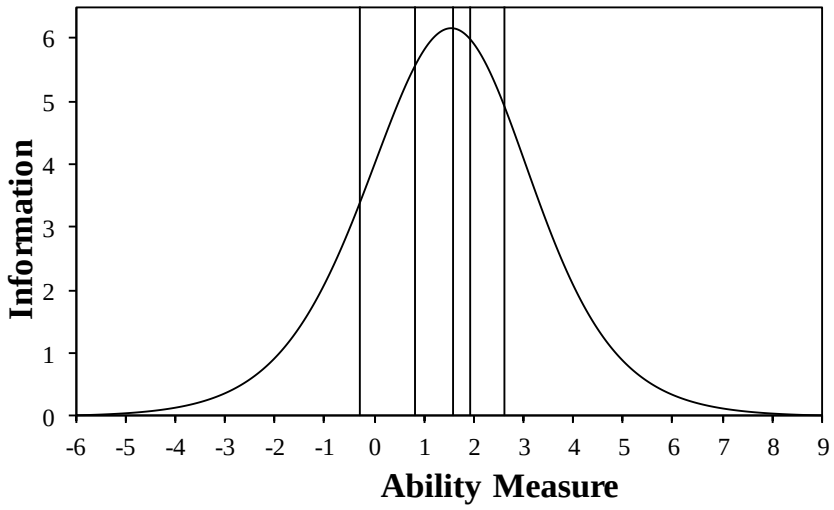
2.9.2.4 Grades 4–5

Figure 2.9.2.4.1
Test Information Function: Read 4-5A S501 Paper



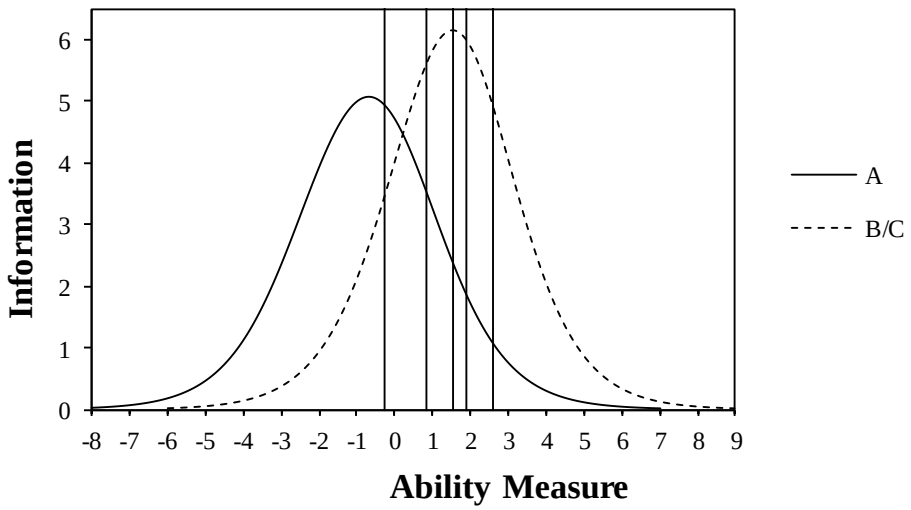
Note: The test form is shared between 3A and 4–5A.

Figure 2.9.2.4.2
Test Information Function: Read 4-5B/C S501 Paper



Note: The test form is shared between 3B/C and 4–5B/C.

Figure 2.9.2.4.3
Test Information Function: Read 4-5 S501 Paper



Note: The test form is shared between 3A and 4–5A, 3B/C and 4–5B/C.

2.9.2.5 Grades 6–8

Figure 2.9.2.5.1
Test Information Function: Read 6-8A S501 Paper

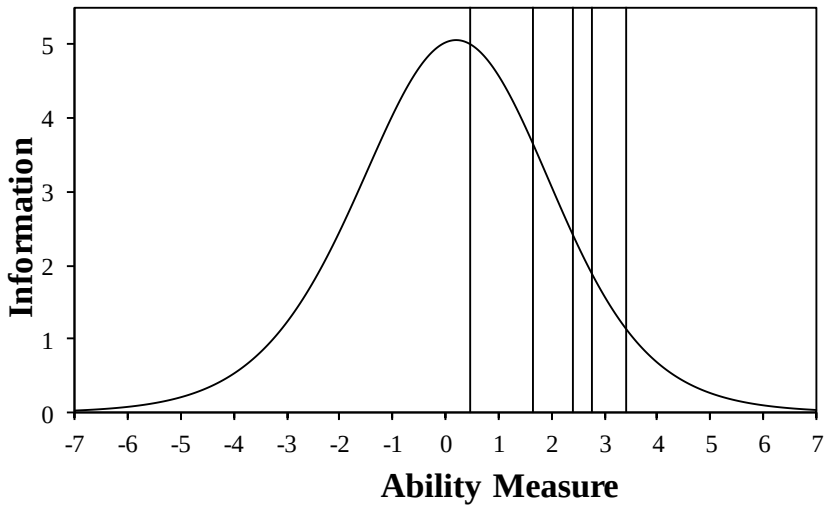


Figure 2.9.2.5.2
Test Information Function: Read 6-8B/C S501 Paper

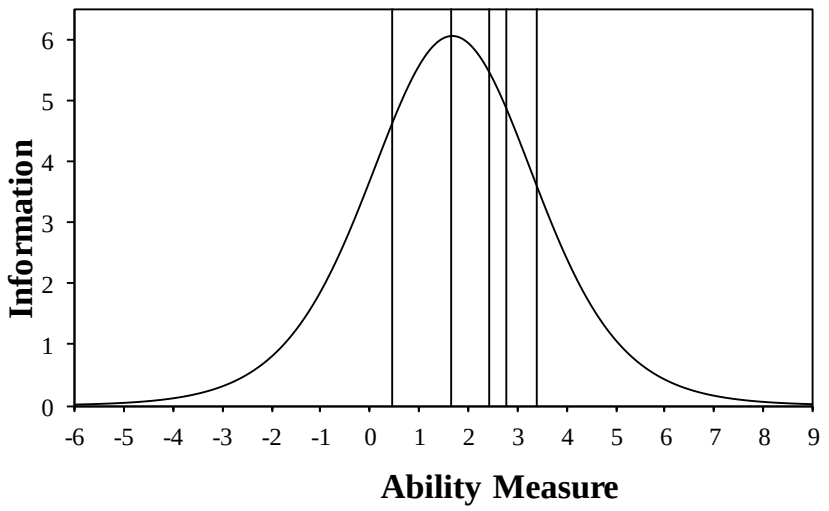
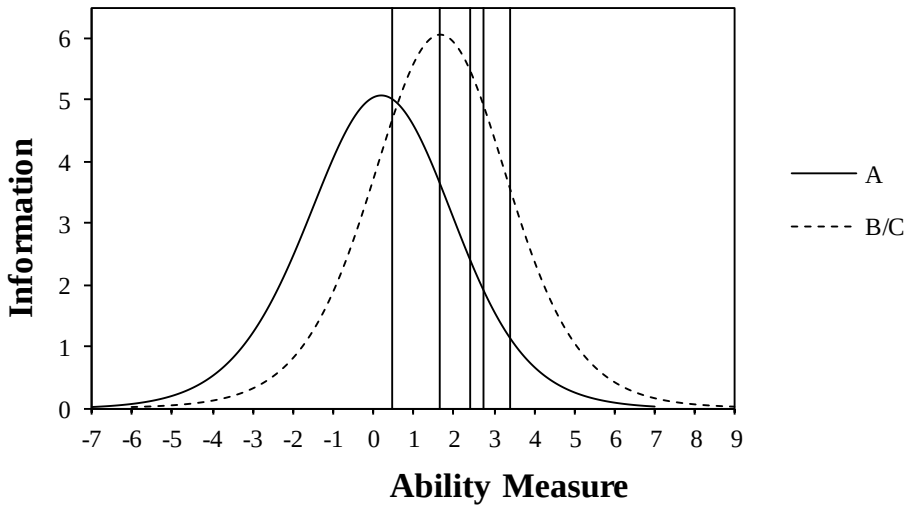


Figure 2.9.2.5.3
Test Information Function: Read 6-8 S501 Paper



2.9.2.6 Grades 9–12

Figure 2.9.2.6.1
Test Information Function: Read 9-12A S501 Paper

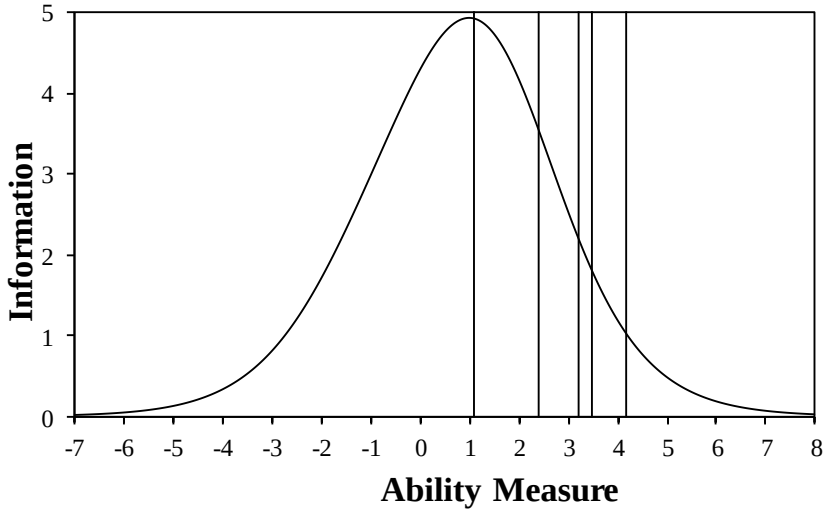


Figure 2.9.2.6.2
Test Information Function: Read 9-12B/C S501 Paper

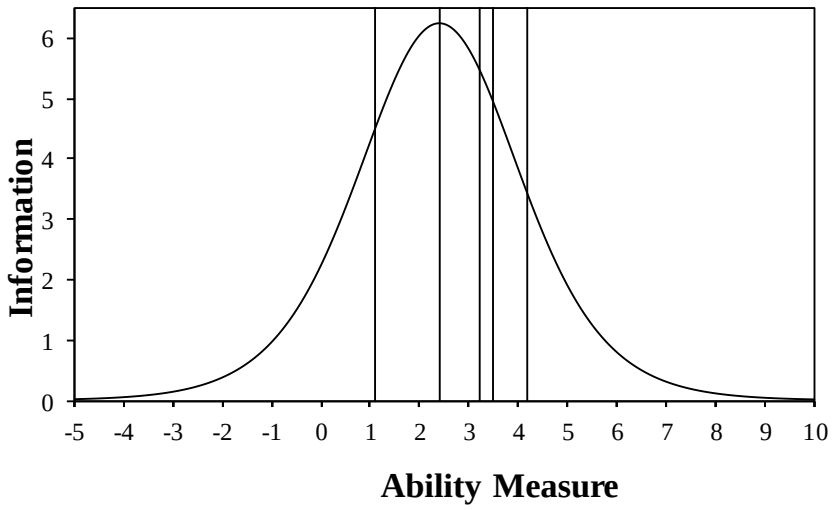
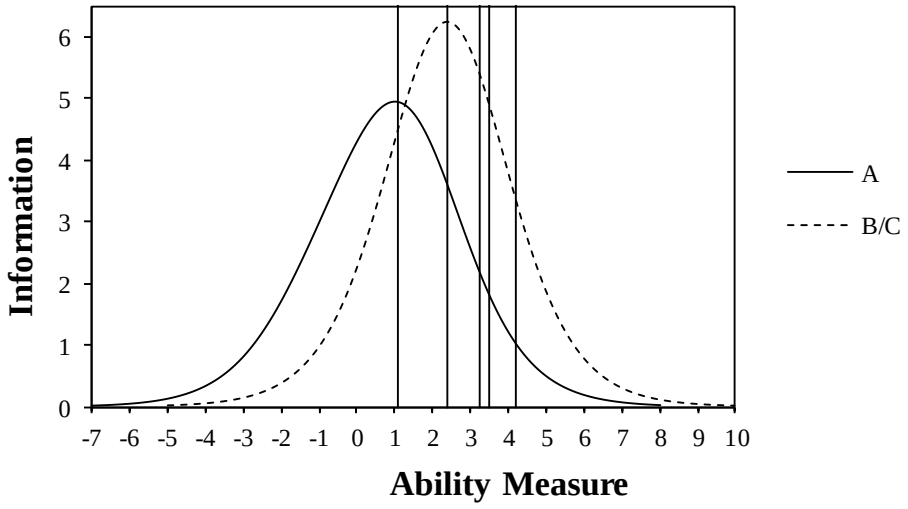


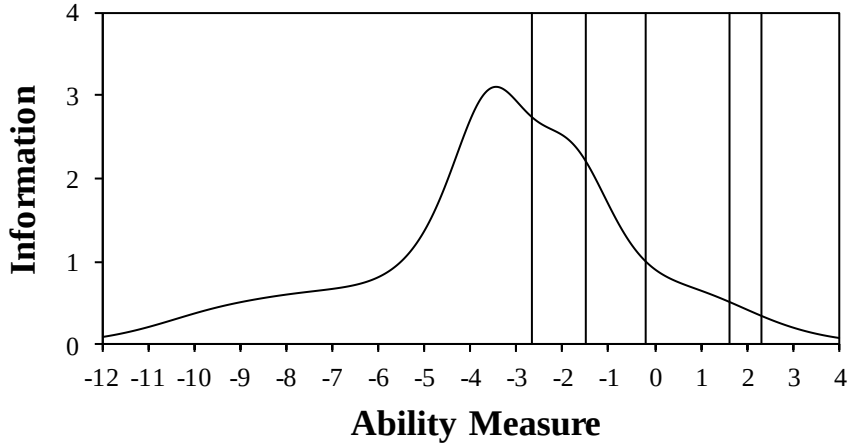
Figure 2.9.2.6.3
Test Information Function: Read 9-12 S501 Paper



2.9.3 Writing

2.9.3.0 Kindergarten

Figure 2.9.3.0
Test Information Function: Writ K S501 Paper



2.9.3.1 Grade 1

Figure 2.9.3.1.1
Test Information Function: Writ 1 A S501 Paper

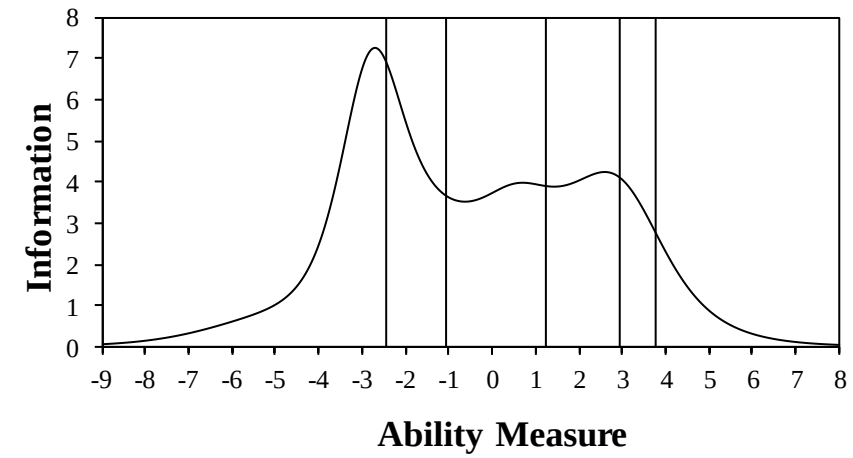


Figure 2.9.3.1.2
Test Information Function: Writ 1 B/C S501 Paper

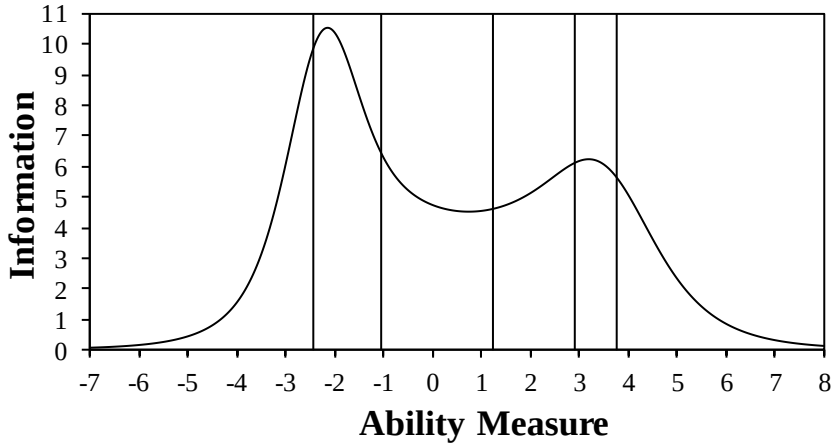
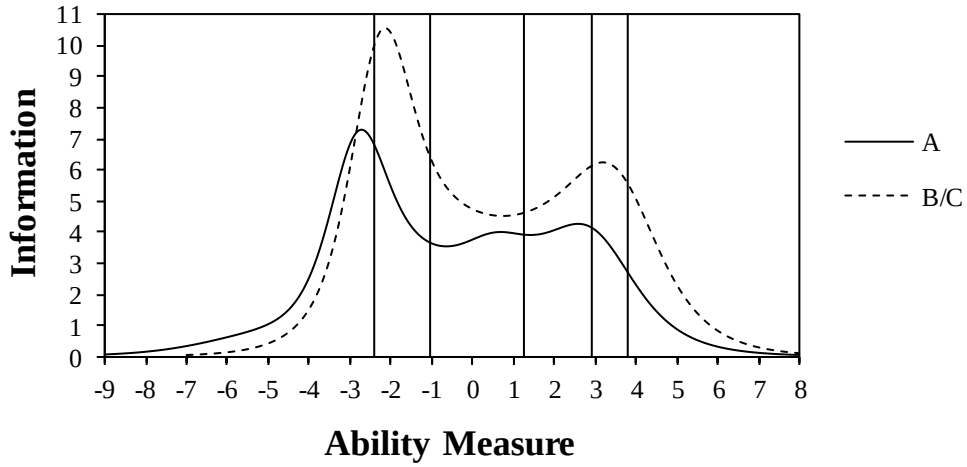
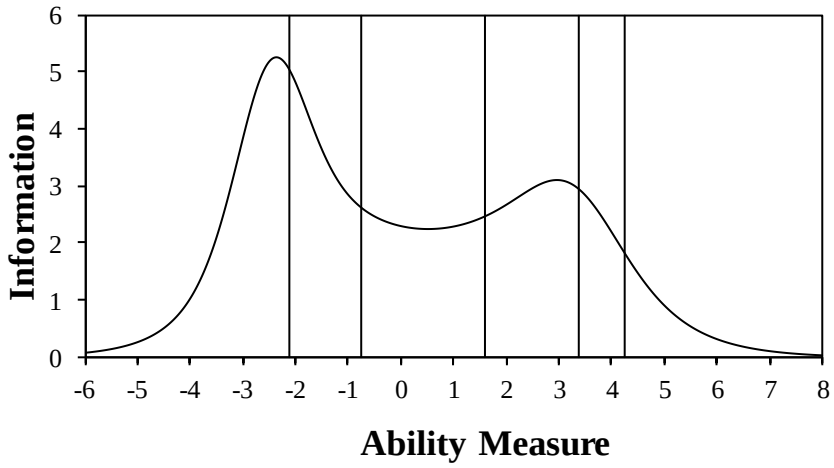


Figure 2.9.3.1.3
Test Information Function: Writ 1 S501 Paper



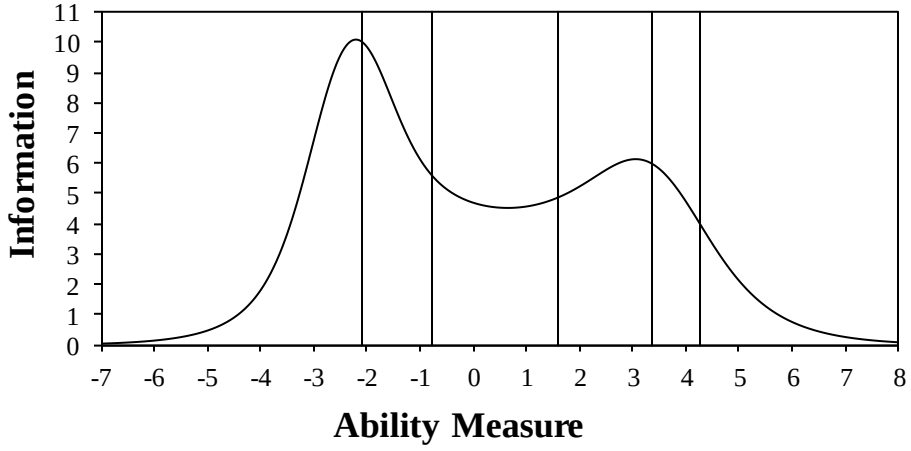
2.9.3.2 Grade 2

Figure 2.9.3.2.1
Test Information Function: Writ 2 A S501 Paper



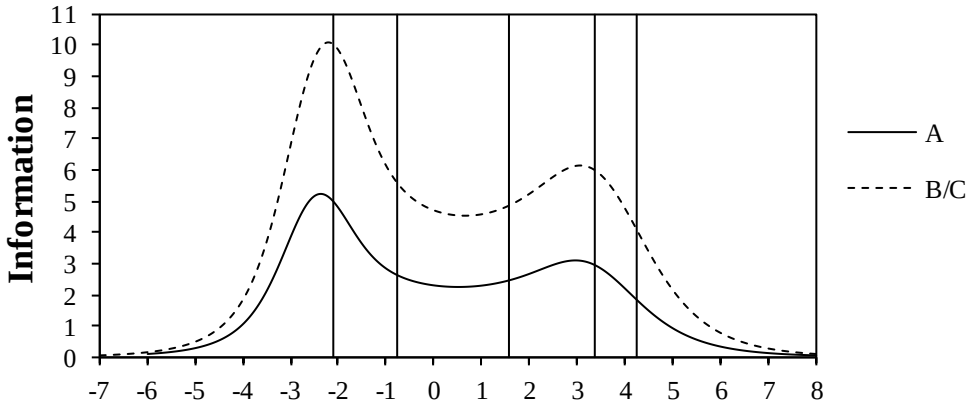
Note: The test form is shared between 2A and 3A.

Figure 2.9.3.2.2
Test Information Function: Writ 2 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

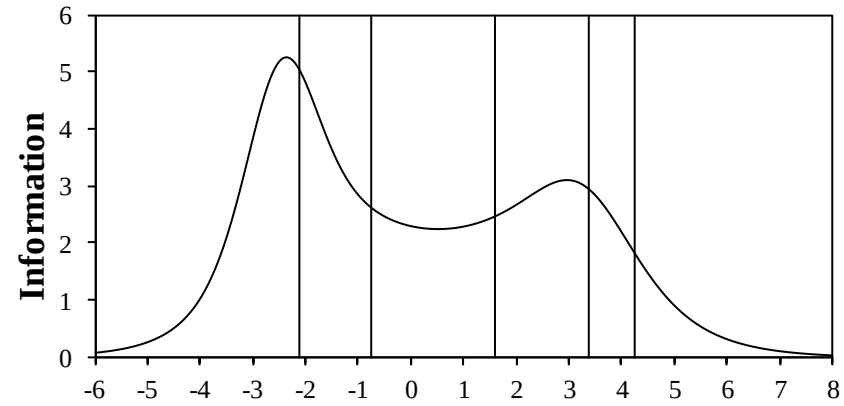
Figure 2.9.3.2.3
Test Information Function: Writ 2 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

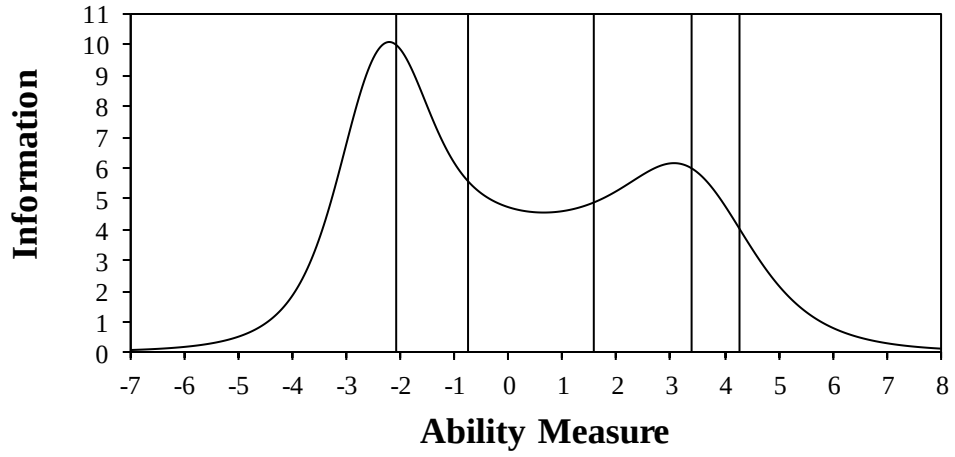
2.9.3.3 Grade 3

Figure 2.9.3.3.1
Test Information Function: Writ 3 A S501 Paper



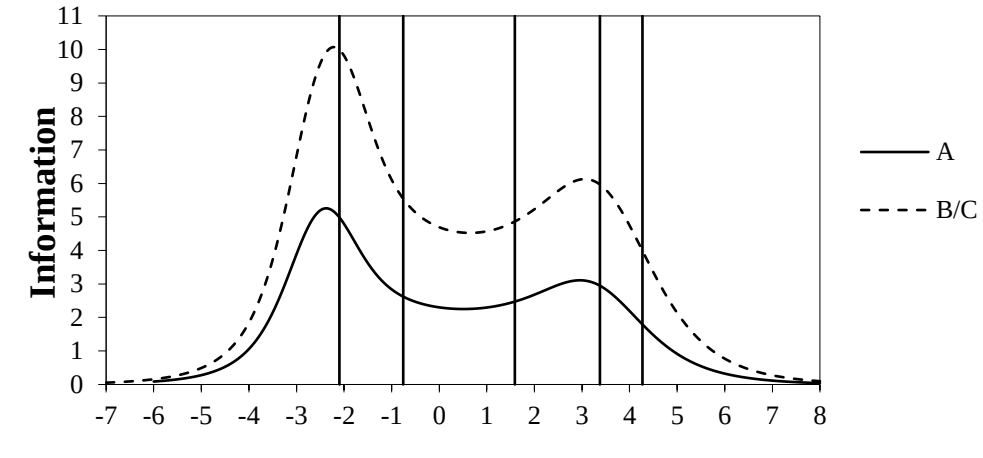
Note: The test form is shared between 2A and 3A.

Figure 2.9.3.3.2
Test Information Function: Writ 3 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

Figure 2.9.3.3.3
Test Information Function: Writ 3 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

2.9.3.4 Grades 4–5

Figure 2.9.3.4.1
Test Information Function: Writ 4-5 A S501 Paper

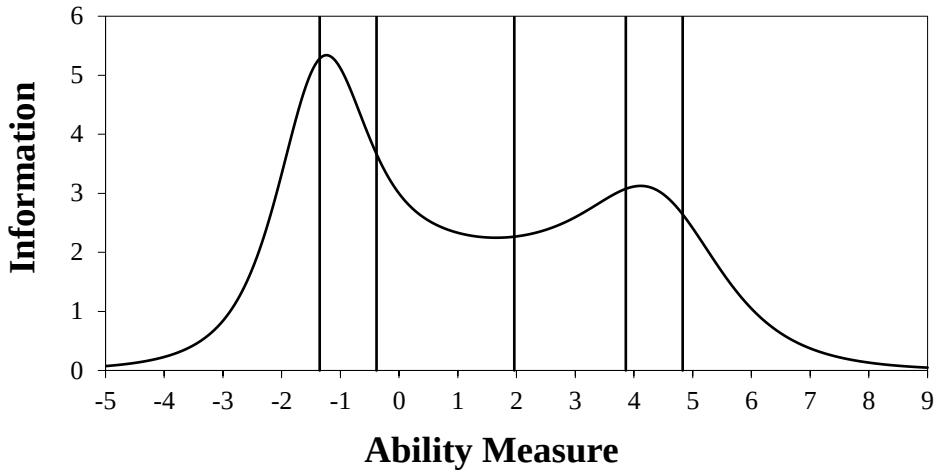


Figure 2.9.3.4.2
Test Information Function: Writ 4-5 B/C S501 Paper

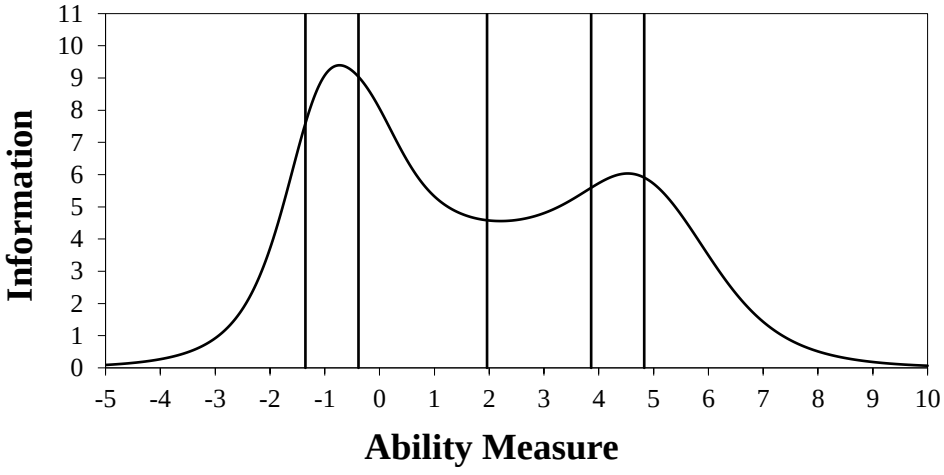
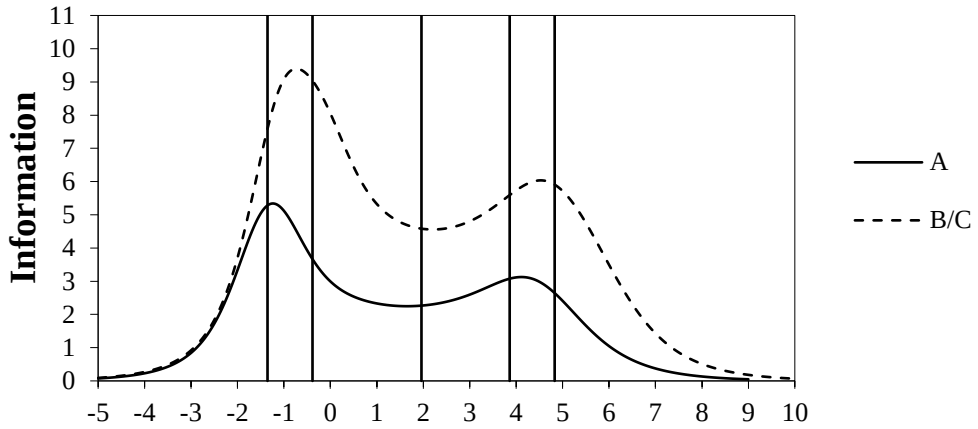


Figure 2.9.3.4.3
Test Information Function: Writ 4-5 S501 Paper



2.9.3.5 Grades 6–8

Figure 2.9.3.5.1
Test Information Function: Writ 6-8 A S501 Paper

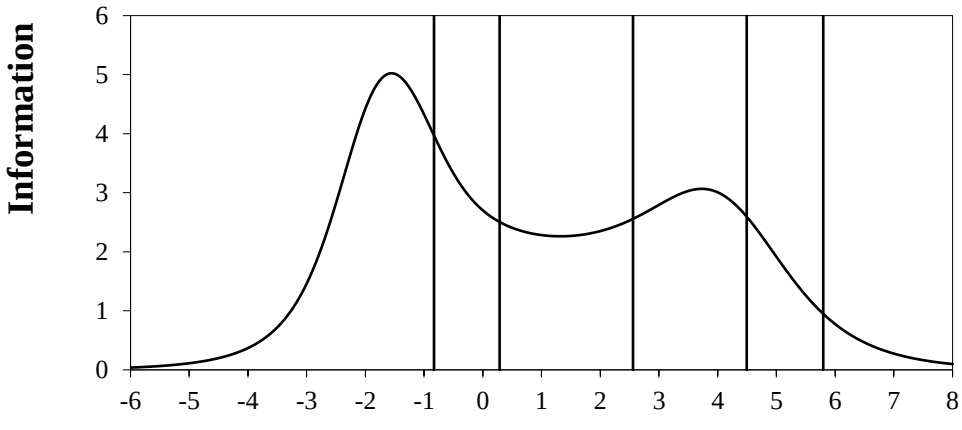


Figure 2.9.3.5.2
Test Information Function: Writ 6-8 B/C S501 Paper

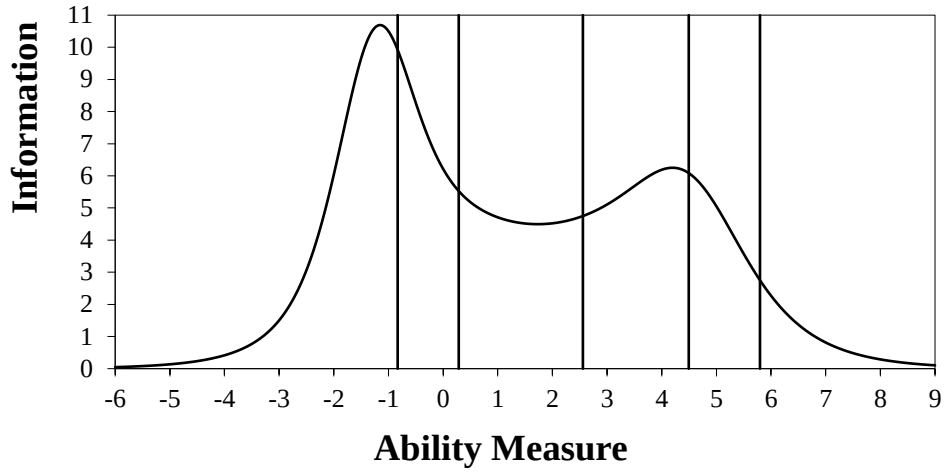
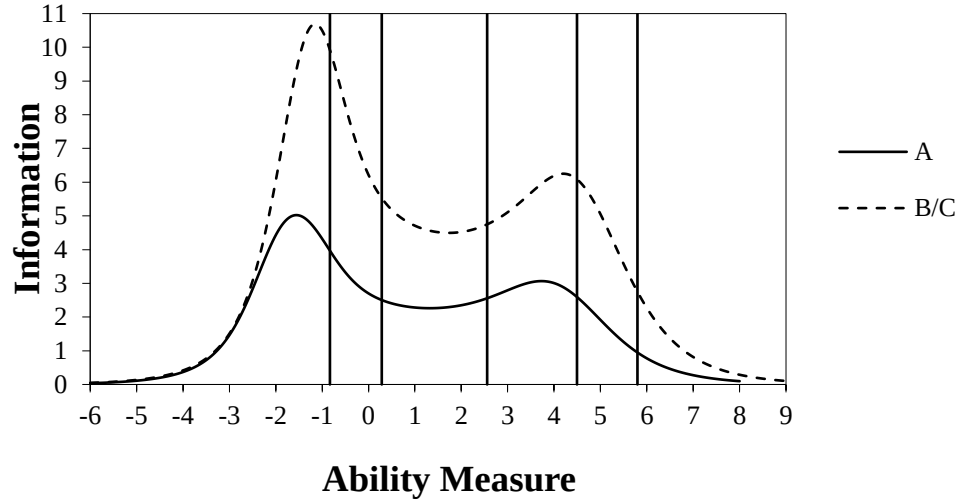


Figure 2.9.3.5.3
Test Information Function: Writ 6-8 S501 Paper



2.9.3.6 Grades 9–12

Figure 2.9.3.6.1
Test Information Function: Writ 9-12 A S501 Paper

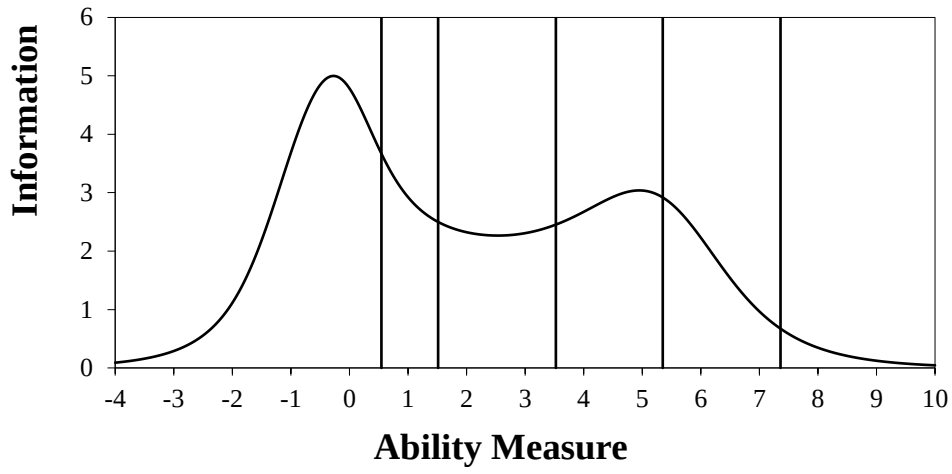


Figure 2.9.3.6.2
Test Information Function: Writ 9-12 B/C S501 Paper

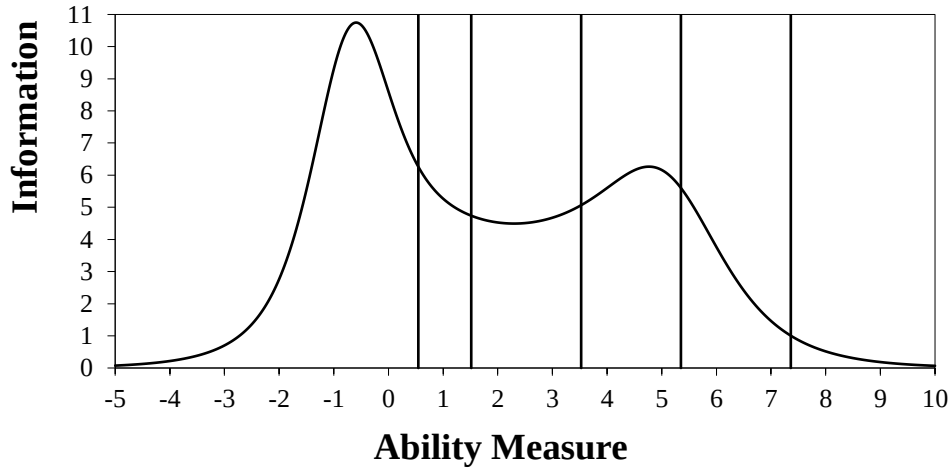
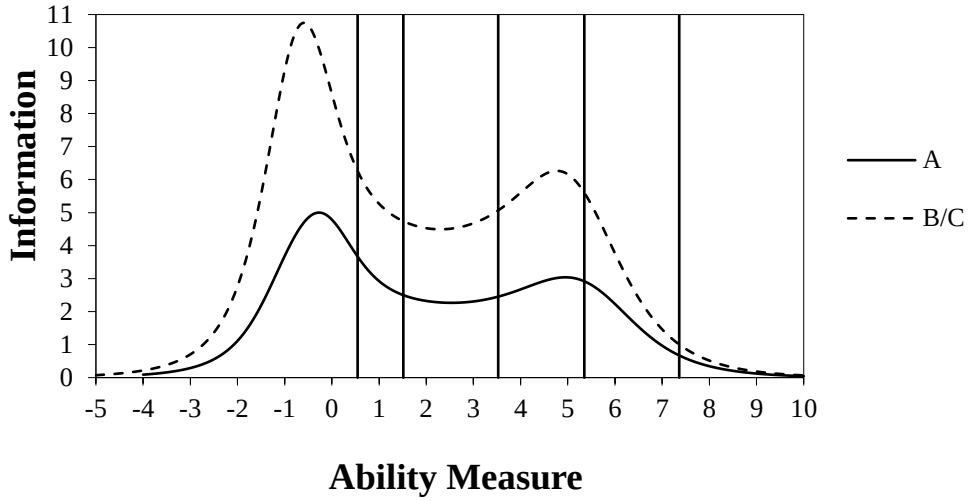


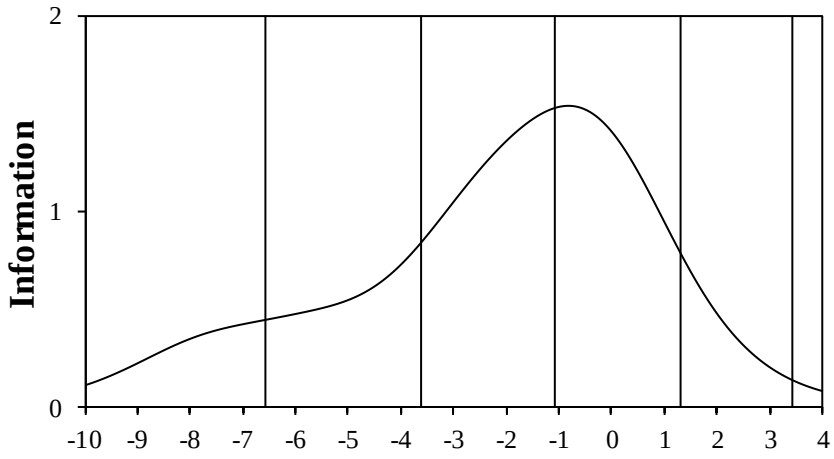
Figure 2.9.3.6.3
Test Information Function: Writ 9-12 S501 Paper



2.9.4 Speaking

2.9.4.0 Kindergarten

Figure 2.9.4.0
Test Information Function: Spek K S501 Paper



2.9.4.1 Grade 1

Figure 2.9.4.1.1
Test Information Function: Spek 1 A S501 Paper

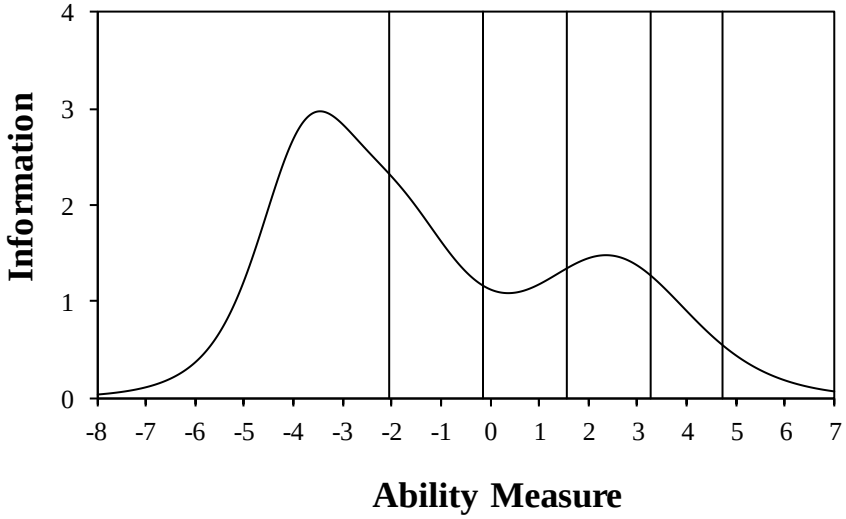


Figure 2.9.4.1.2
Test Information Function: Spek 1 B/C S501 Paper

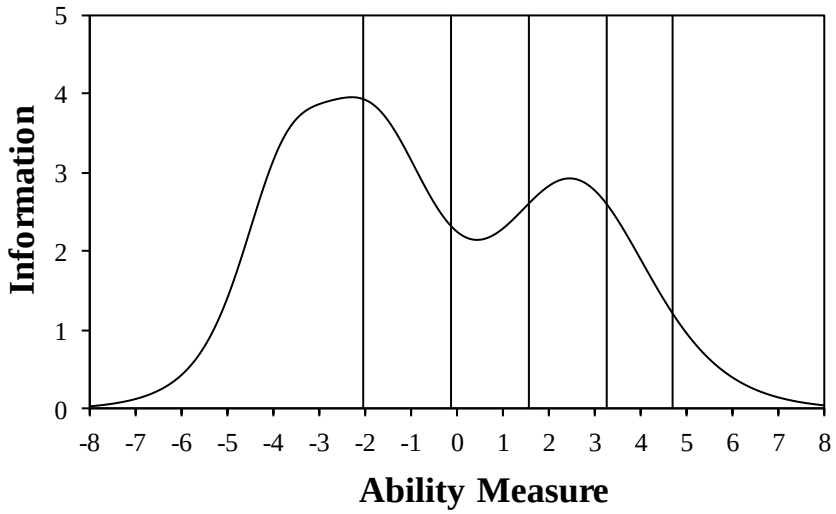
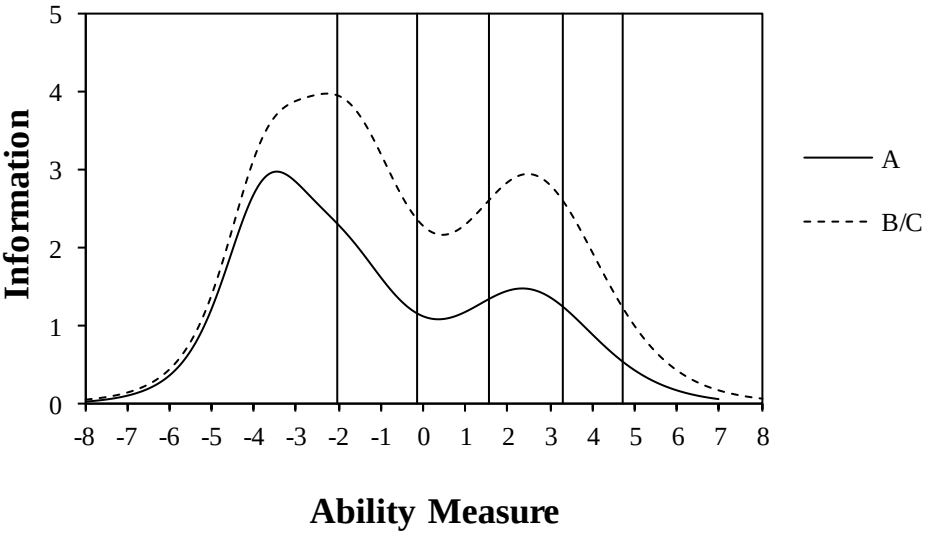
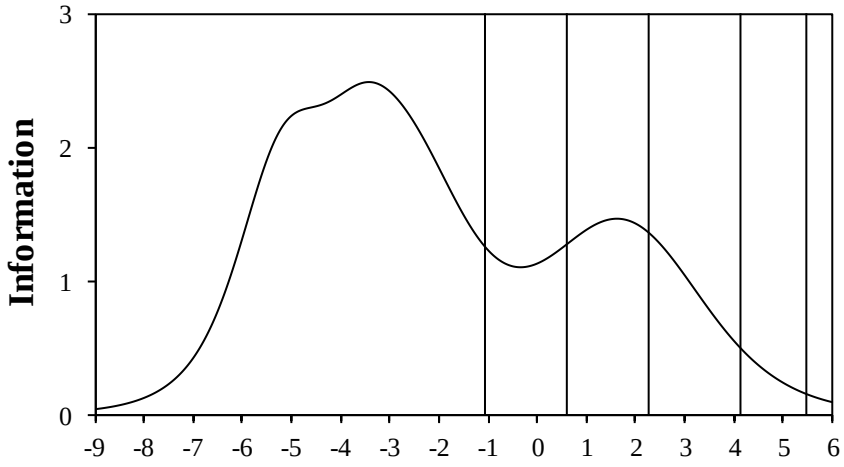


Figure 2.9.4.1.3
Test Information Function: Spek 1 S501 Paper



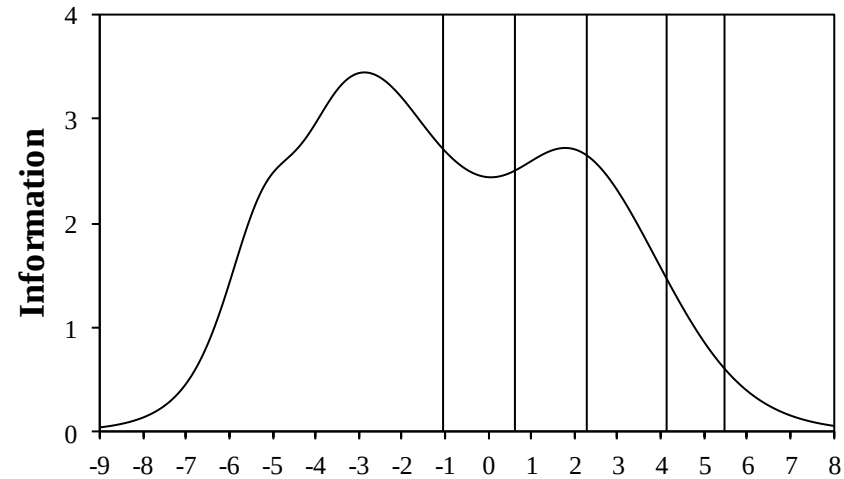
2.9.4.2 Grade 2

Figure 2.9.4.2.1
Test Information Function: Spek 2 A S501 Paper



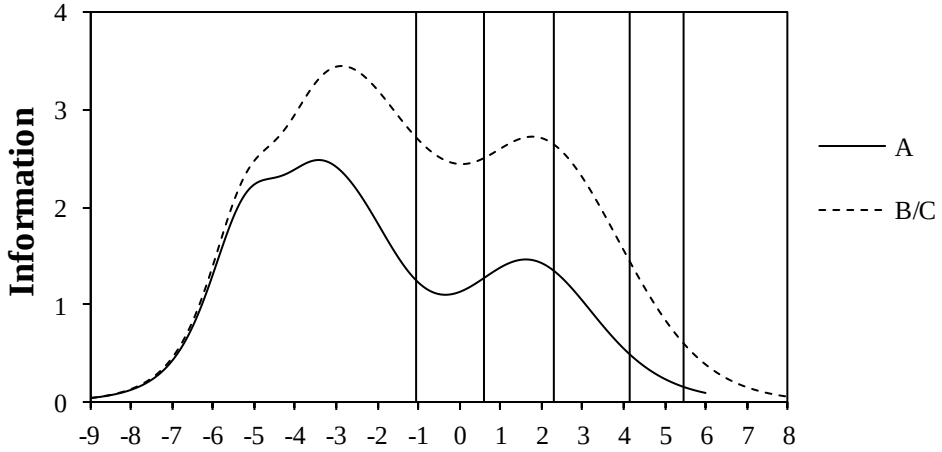
Note: The test form is shared between 2A and 3A.

Figure 2.9.4.2.2
Test Information Function: Spek 2 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

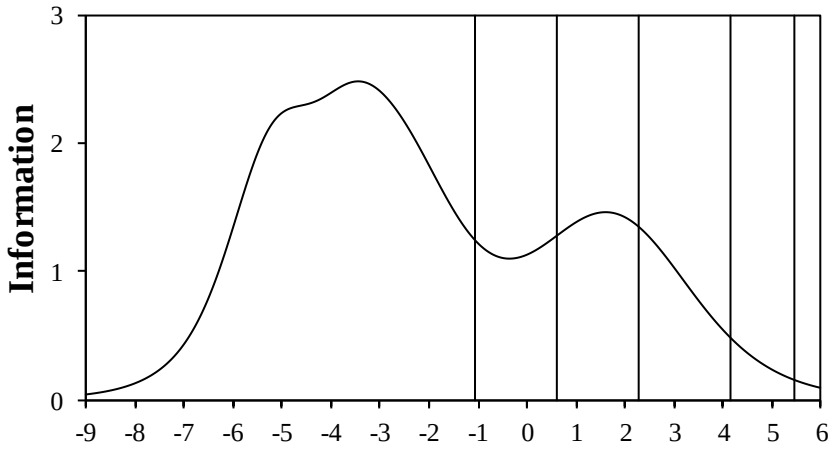
Figure 2.9.4.2.3
Test Information Function: Spek 2 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

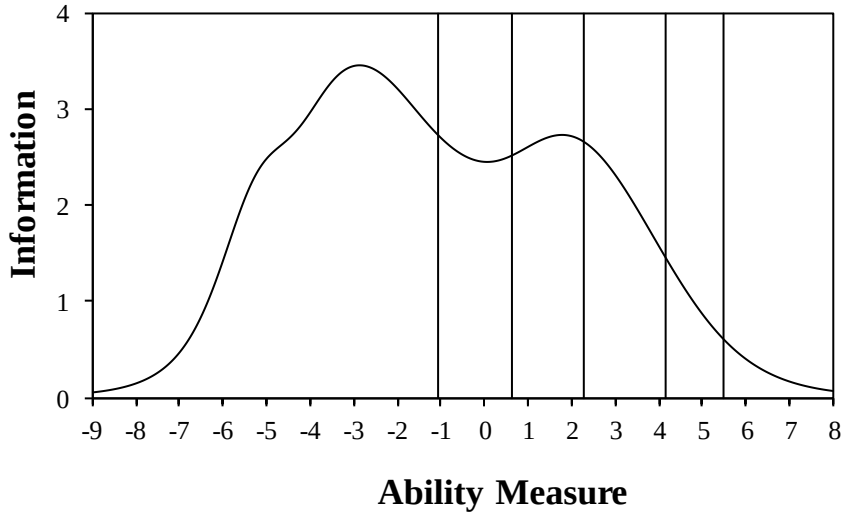
2.9.4.3 Grade 3

Figure 2.9.4.3.1
Test Information Function: Spek 3 A S501 Paper



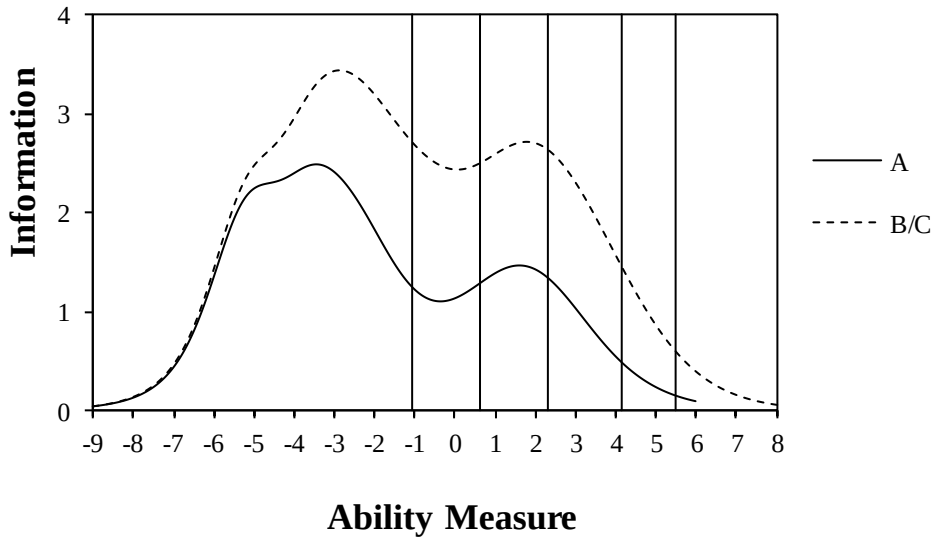
Note: The test form is shared between 2A and 3A.

Figure 2.9.4.3.2
Test Information Function: Spek 3 B/C S501 Paper



Note: The test form is shared between 2B/C and 3B/C.

Figure 2.9.4.3.3
Test Information Function: Spek 3 S501 Paper



Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

2.9.4.4 Grades 4–5

Figure 2.9.4.4.1
Test Information Function: Spek 4-5 A S501 Paper

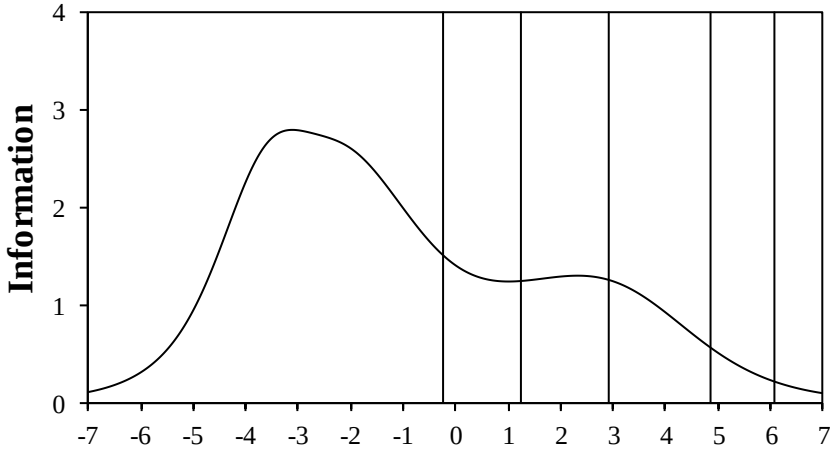


Figure 2.9.4.4.2
Test Information Function: Spek 4-5 B/C S501 Paper

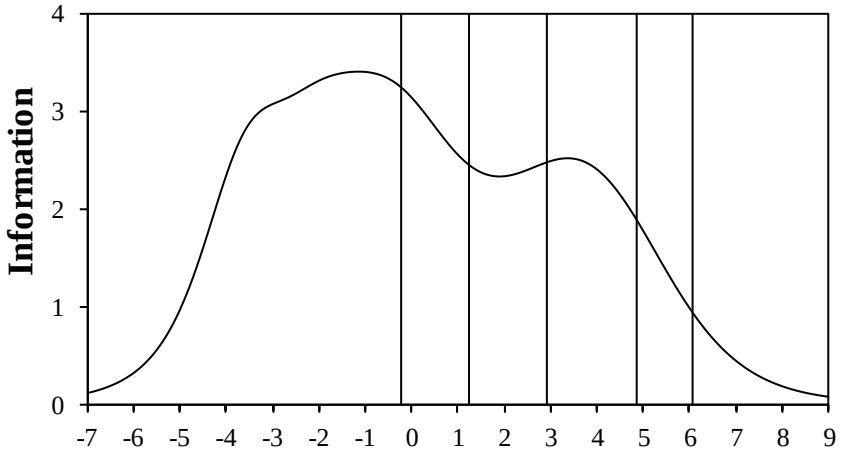
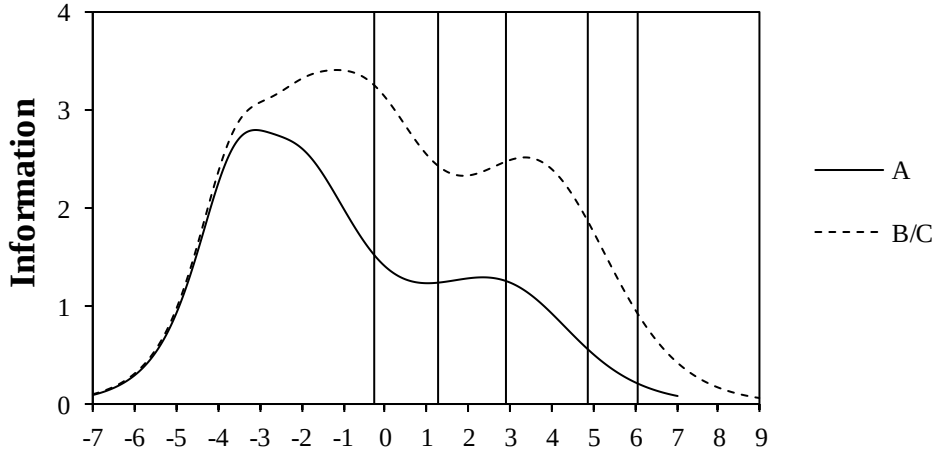


Figure 2.9.4.4.3
Test Information Function: Spek 4-5 S501 Paper



2.9.4.5 Grades 6–8

Figure 2.9.4.5.1
Test Information Function: Spek 6-8 A S501 Paper

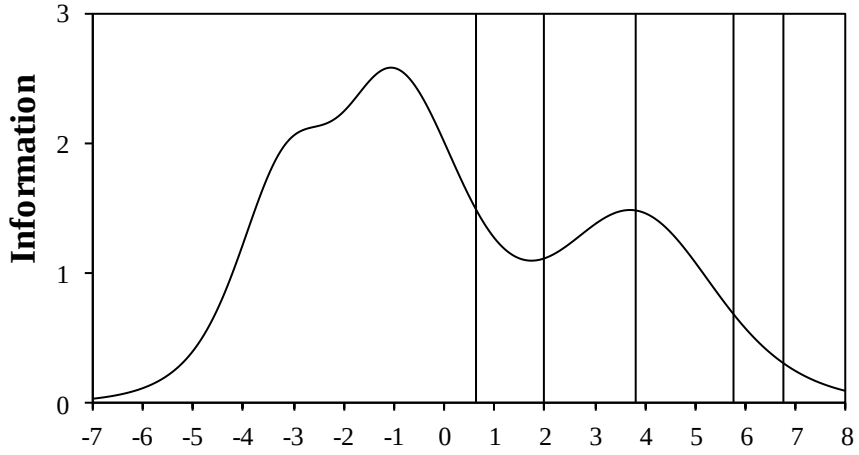


Figure 2.9.4.5.2
Test Information Function: Spek 6-8 B/C S501 Paper

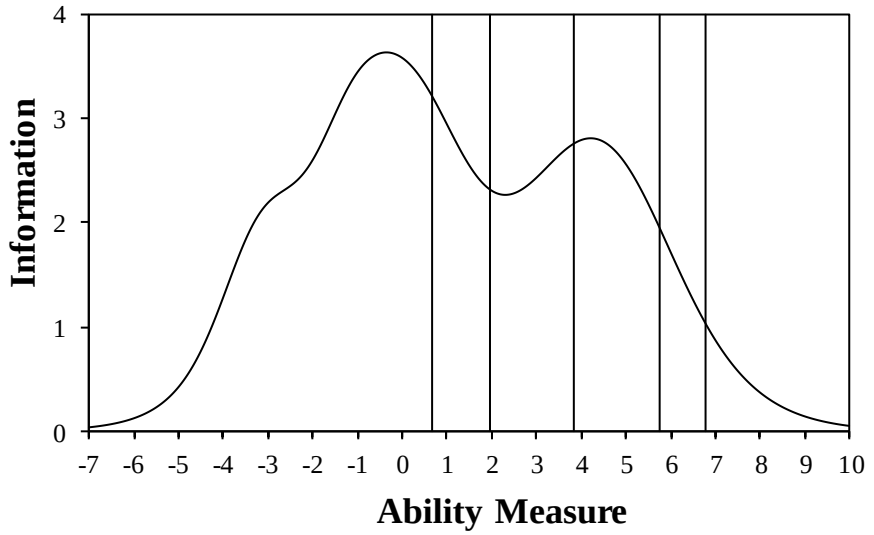
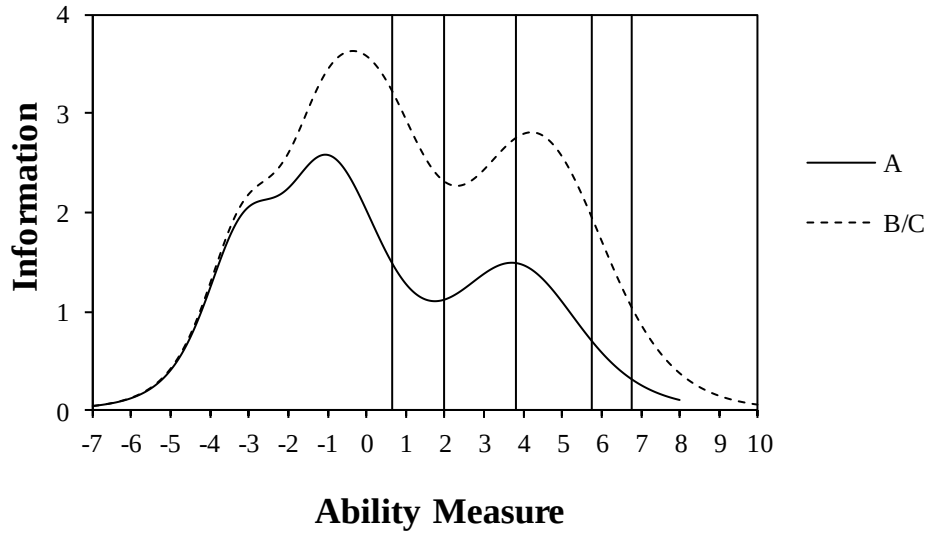


Figure 2.9.4.5.3
Test Information Function: Spek 6-8 S501 Paper



2.9.4.6 Grades 9–1

Figure 2.9.4.6.1
Test Information Function: Spek 9-12 A S501 Paper

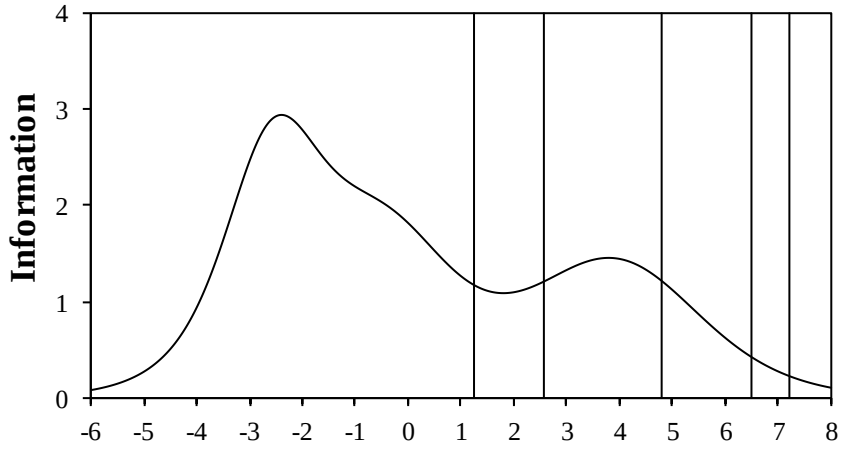


Figure 2.9.4.6.2
Test Information Function: Spek 9-12 B/C S501 Paper

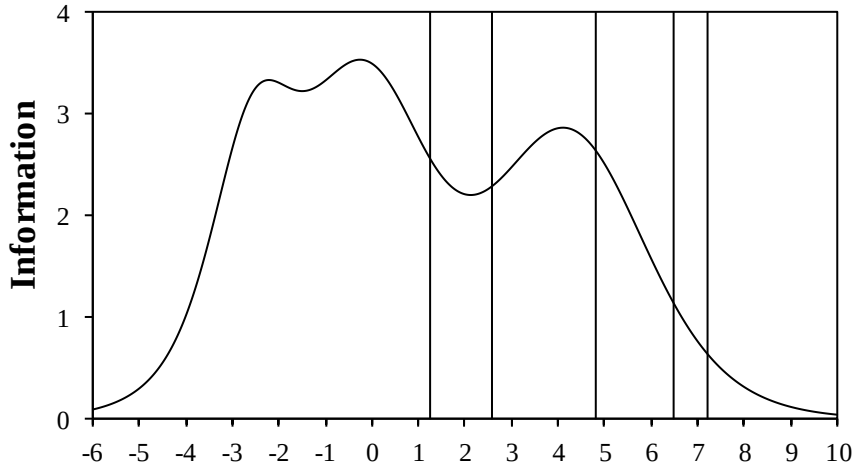
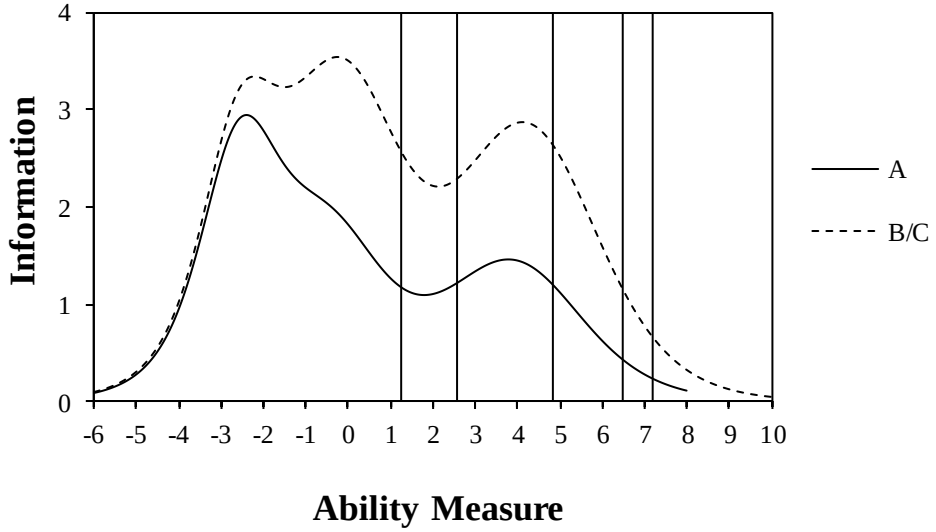


Figure 2.9.4.6.3
Test Information Function: Spek 9-12 S501 Paper



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 score

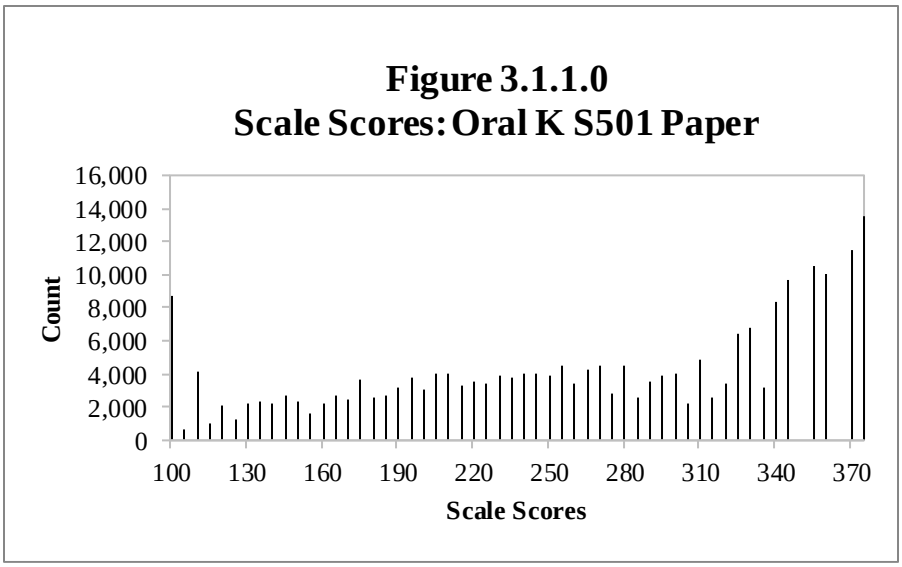
3.1.1 Oral

3.1.1.0 Kindergarten

Table 3.1.1.0

Scale Score Descriptive Statistics: Oral K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,997	100	378	266.84	83.72

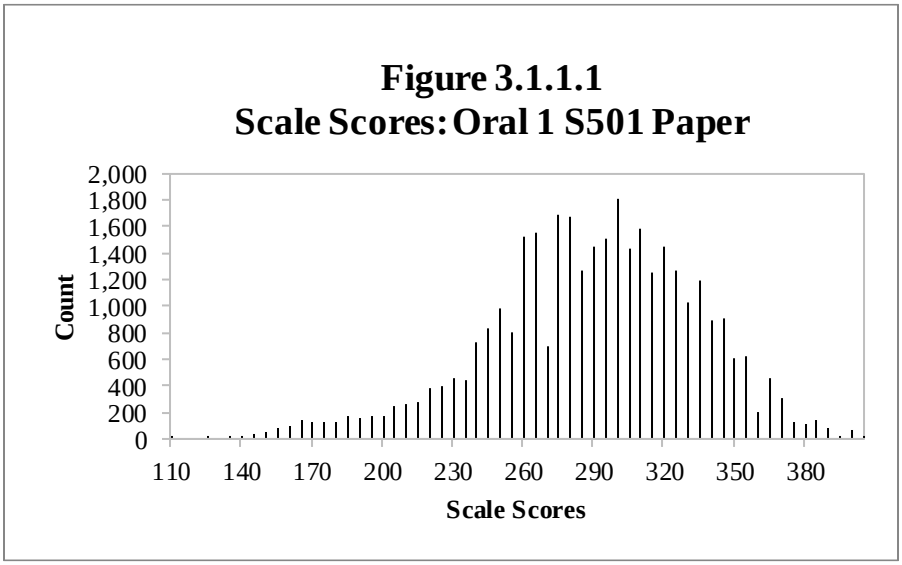


3.1.1.1 Grade 1

Table 3.1.1.1

Scale Score Descriptive Statistics: Oral 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	34,181	114	408	291.39	44.66

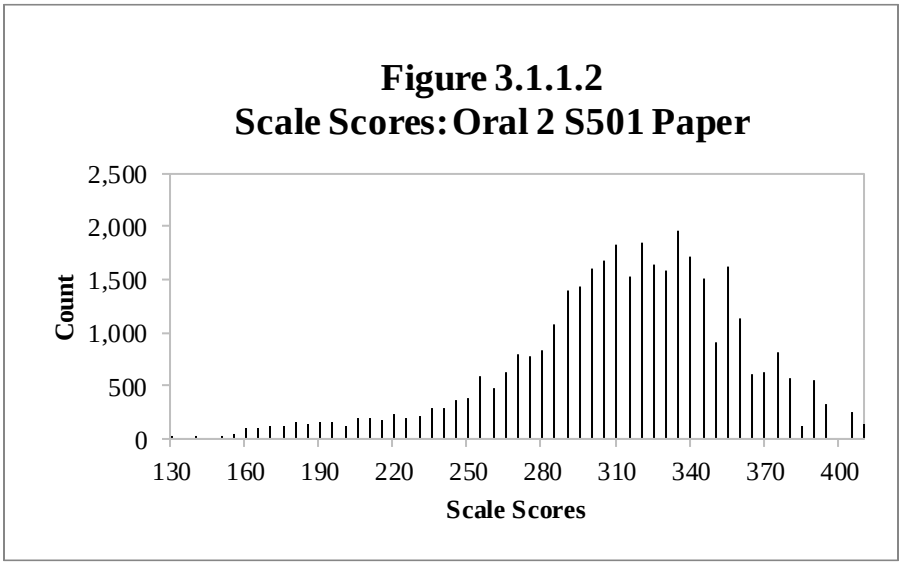


3.1.1.2 Grade 2

Table 3.1.1.2

Scale Score Descriptive Statistics: Oral 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	36,379	134	413	313.48	46.99

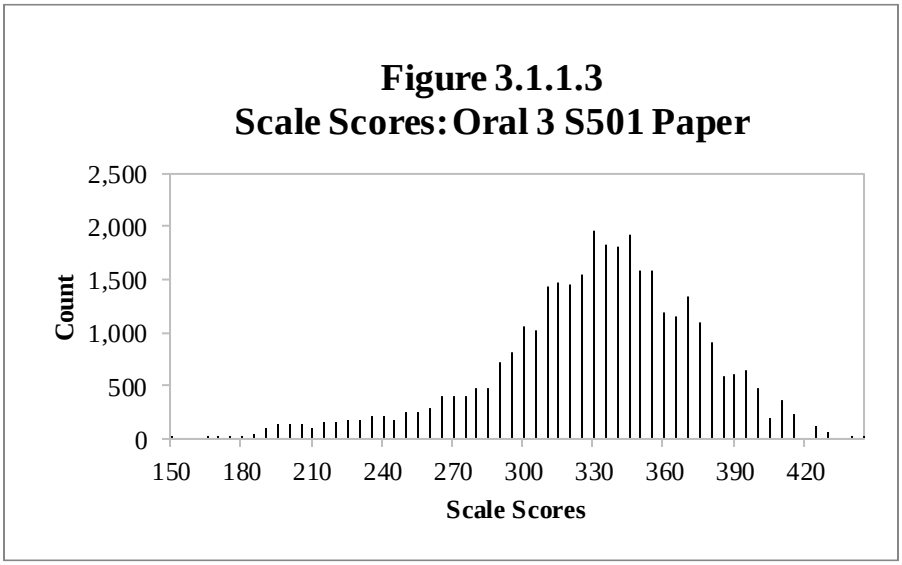


3.1.1.3 Grade 3

Table 3.1.1.3

Scale Score Descriptive Statistics: Oral 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	34,195	151	447	331.82	44.92

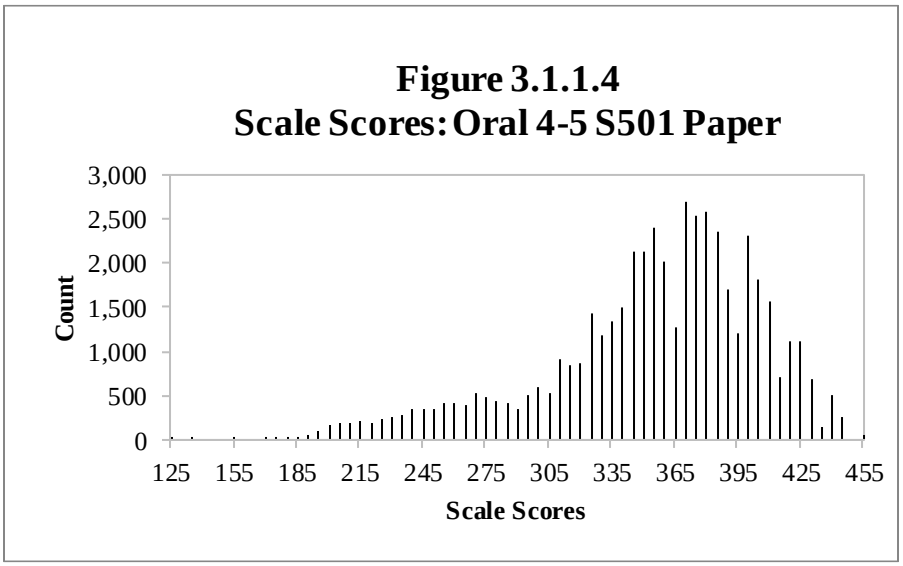


3.1.1.4 Grades 4–5

Table 3.1.1.4

Scale Score Descriptive Statistics: Oral 4-5 S501 Paper

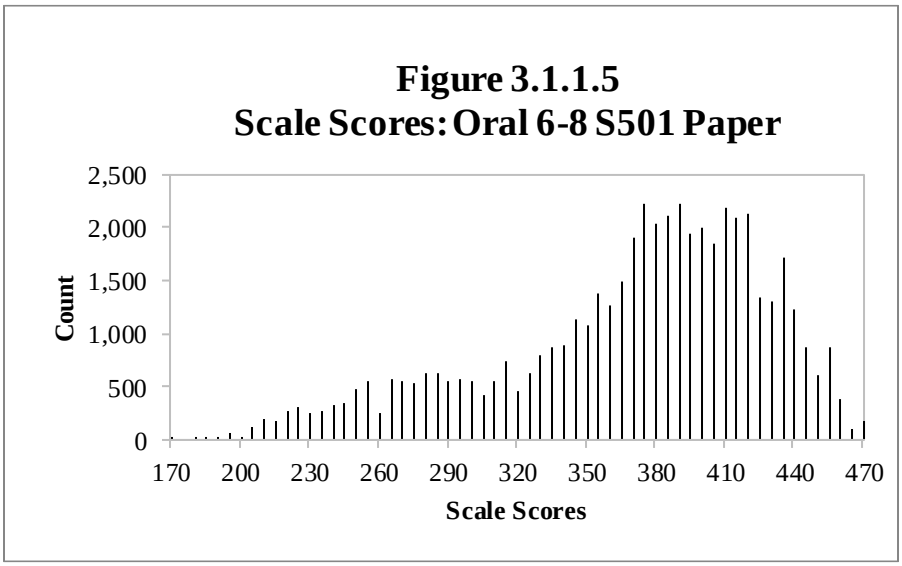
Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	26,143	125	456	352.55	50.86
5	23,170	173	456	359.93	52.83
Total	49,313	125	456	356.02	51.93



3.1.1.5 Grades 6–8

Table 3.1.1.5
Scale Score Descriptive Statistics: Oral 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	19,075	173	472	369.60	55.59
7	16,551	180	472	370.81	59.76
8	15,649	180	472	373.57	61.32
Total	51,275	173	472	371.20	58.76



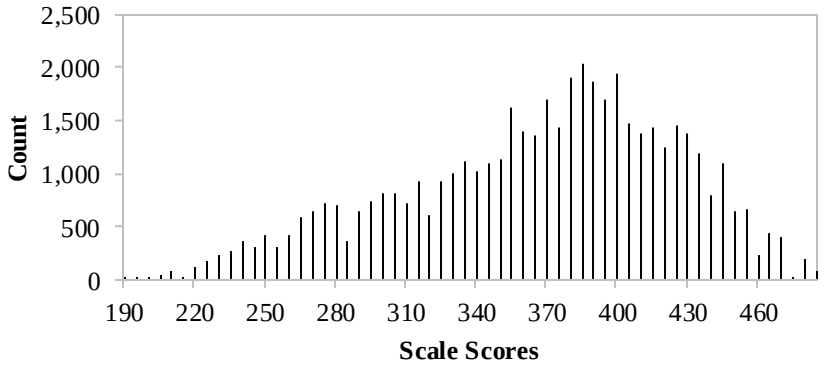
3.1.1.6 Grades 9–12

Table 3.1.1.6

Scale Score Descriptive Statistics: Oral 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	15,259	192	488	360.21	61.66
10	13,599	198	488	367.51	58.03
11	11,779	198	488	374.62	55.59
12	7,836	198	488	376.71	52.36
Total	48,473	192	488	368.43	58.11

Figure 3.1.1.6
Scale Scores: Oral 9-12 S501 Paper



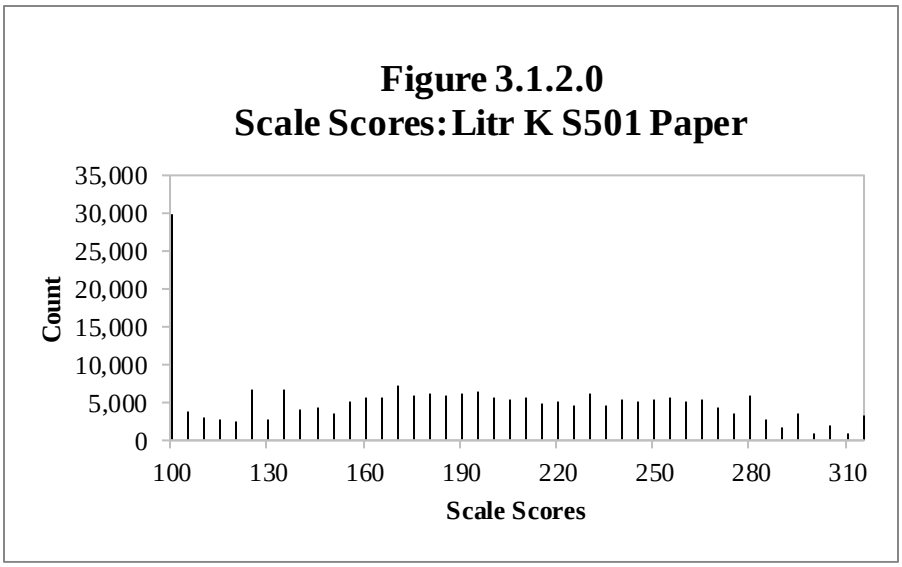
3.1.2 Literacy

3.1.2.0 Kindergarten

Table 3.1.2.0

Scale Score Descriptive Statistics: Litr K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,982	100	315	192.08	62.28

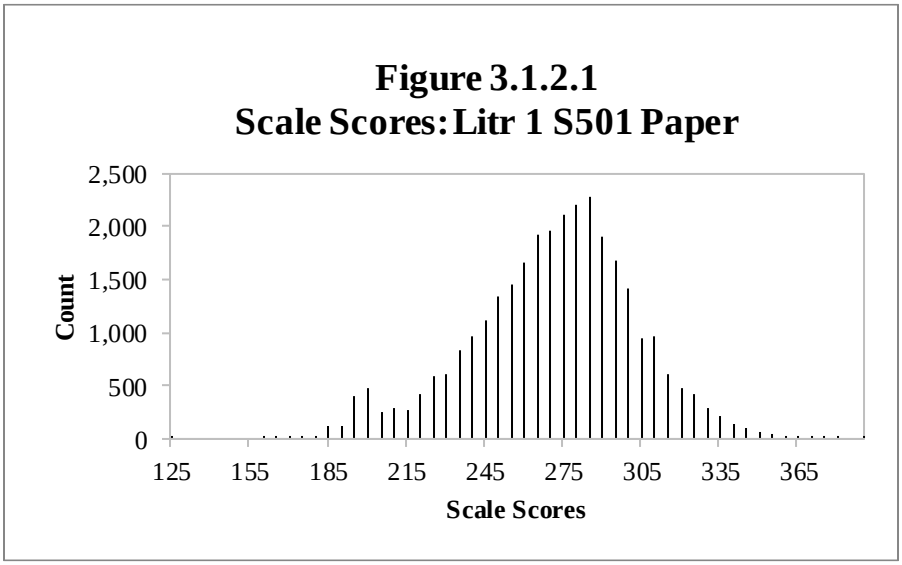


3.1.2.1 Grade 1

Table 3.1.2.1

Scale Score Descriptive Statistics: Litr 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	30,823	126	394	273.02	31.88

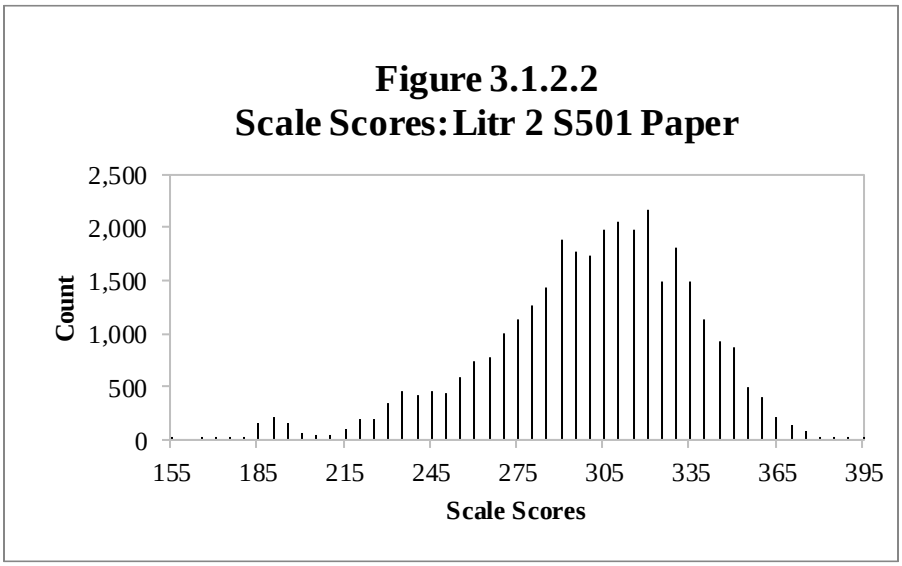


3.1.2.2 Grade 2

Table 3.1.2.2

Scale Score Descriptive Statistics: Litr 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	33,000	157	395	301.98	35.86

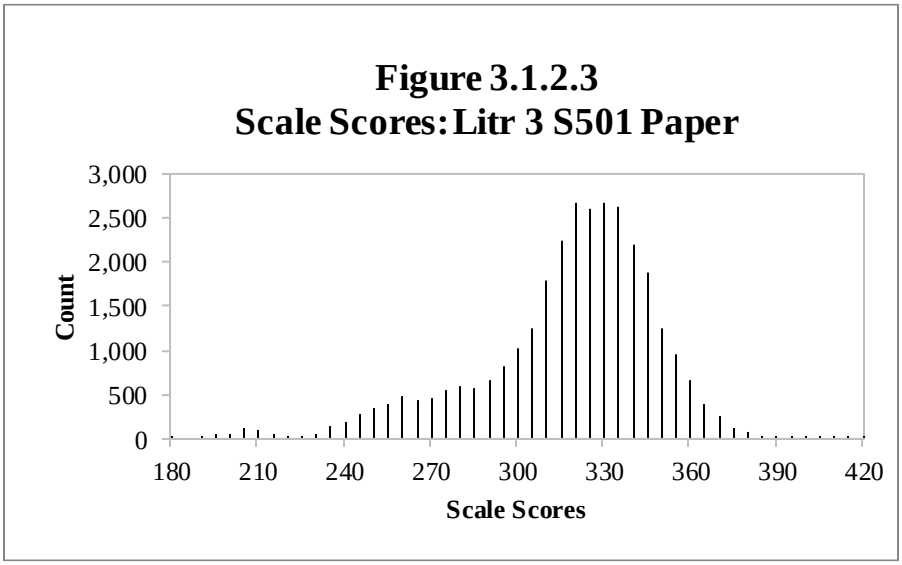


3.1.2.3 Grade 3

Table 3.1.2.3

Scale Score Descriptive Statistics: Litr 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	31,130	184	421	319.48	31.21

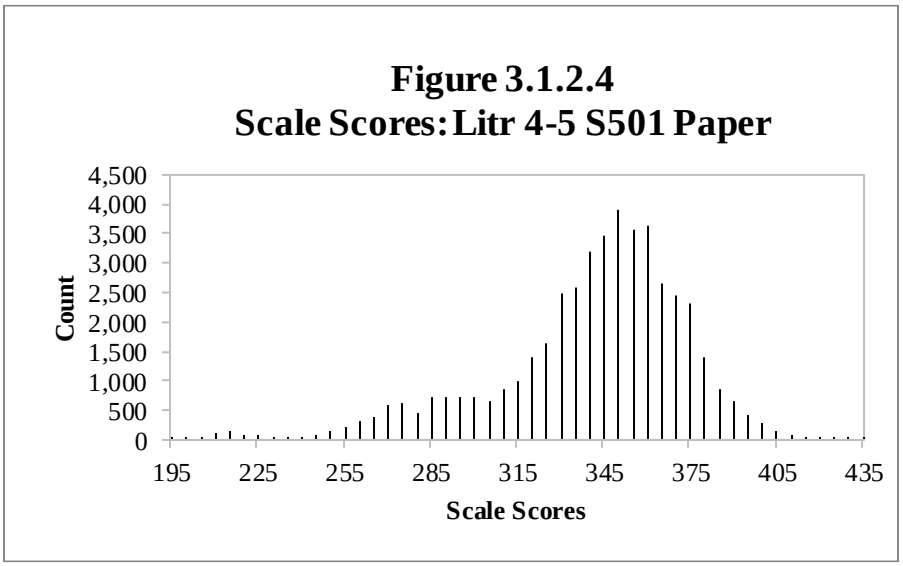


3.1.2.4 Grades 4–5

Table 3.1.2.4

Scale Score Descriptive Statistics: Litr 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	24,128	195	430	338.26	32.61
5	21,881	195	439	347.08	34.32
Total	46,009	195	439	342.45	33.73

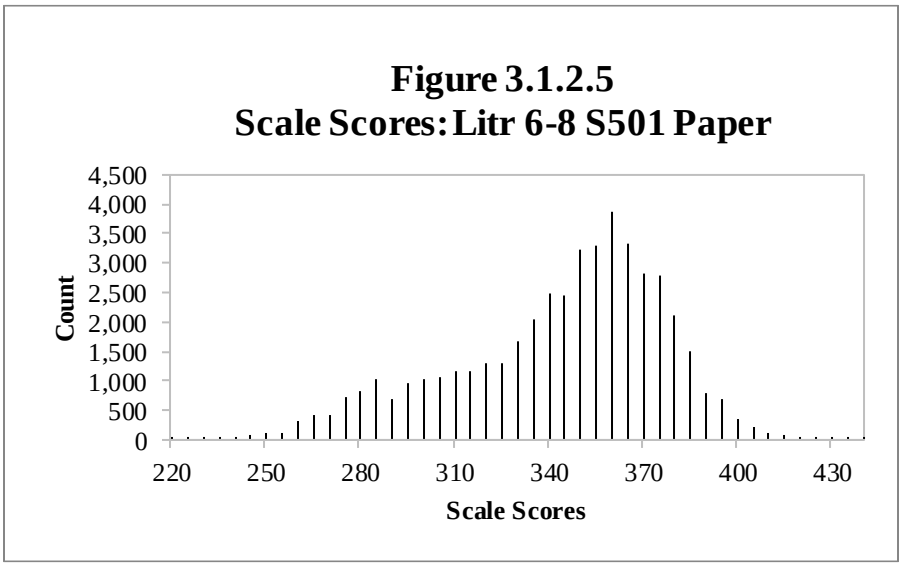


3.1.2.5 Grades 6–8

Table 3.1.2.5

Scale Score Descriptive Statistics: Litr 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,163	222	426	342.07	31.36
7	15,068	234	444	345.31	33.73
8	14,356	238	435	349.34	35.08
Total	46,587	222	444	345.36	33.44

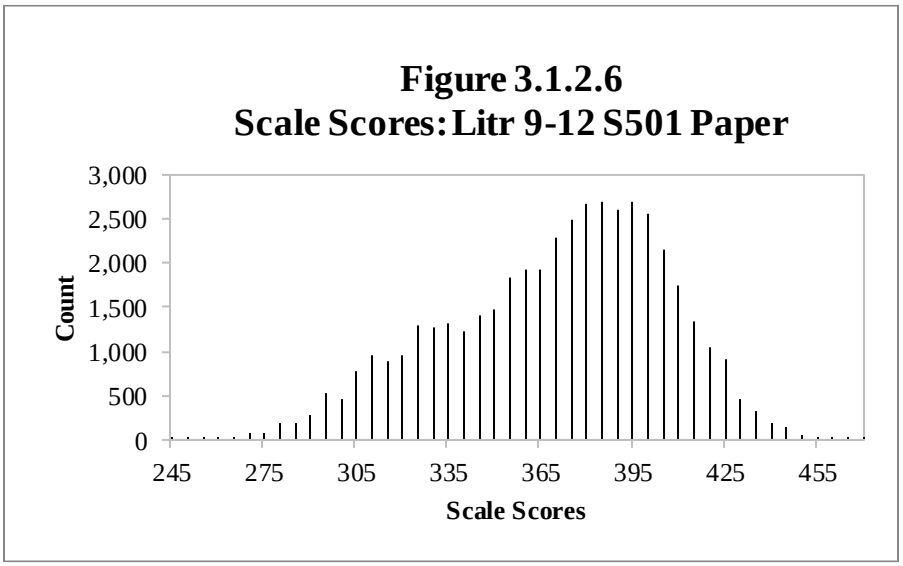


3.1.2.6 Grades 9–12

Table 3.1.2.6

Scale Score Descriptive Statistics: Litr 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	14,183	251	462	364.96	36.90
10	12,697	263	467	371.78	35.38
11	11,121	268	472	377.99	34.27
12	7,454	245	467	376.82	32.61
Total	45,455	245	472	372.00	35.55

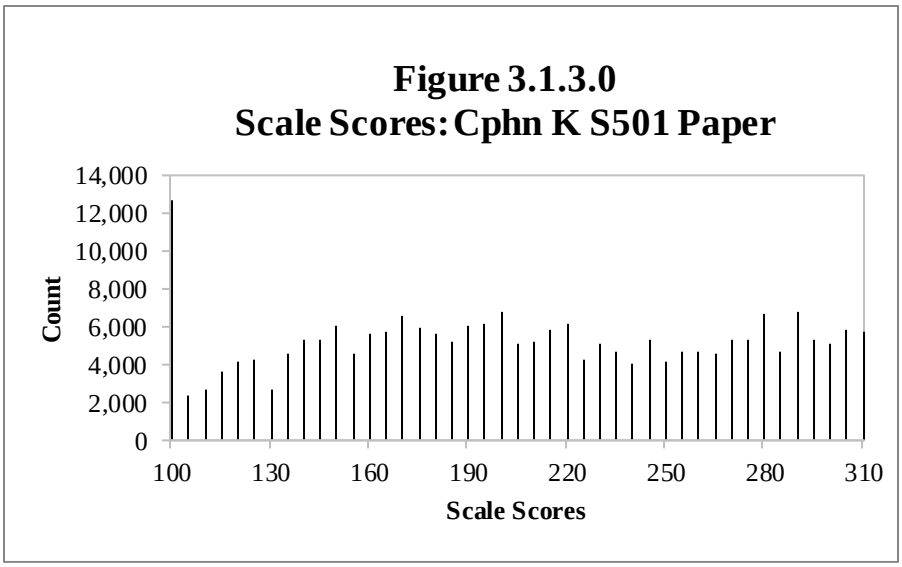


3.1.3 Comprehension

3.1.3.0 Kindergarten

Table 3.1.3.0
Scale Score Descriptive Statistics: Cphn K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,990	100	312	208.03	62.28

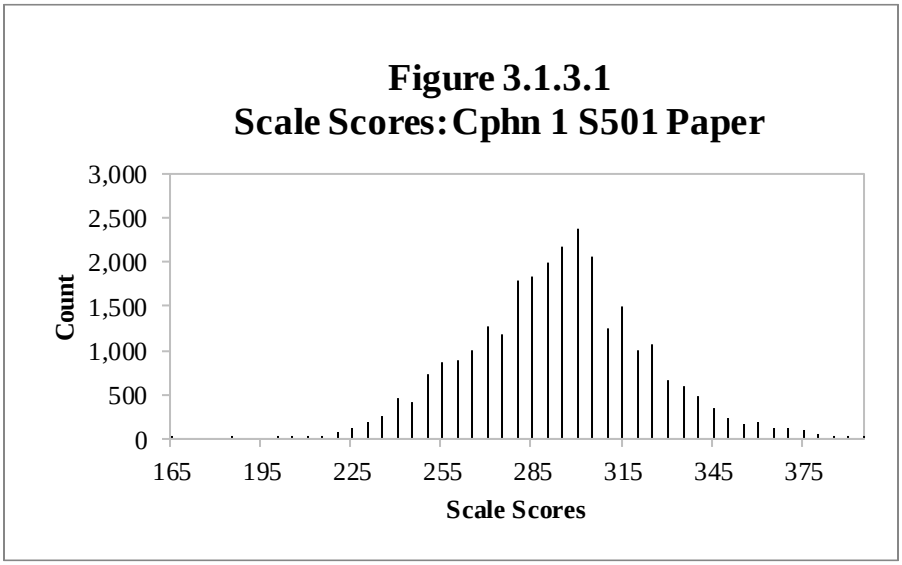


3.1.3.1 Grade 1

Table 3.1.3.1

Scale Score Descriptive Statistics: Cphn 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	27,624	166	395	295.66	28.82

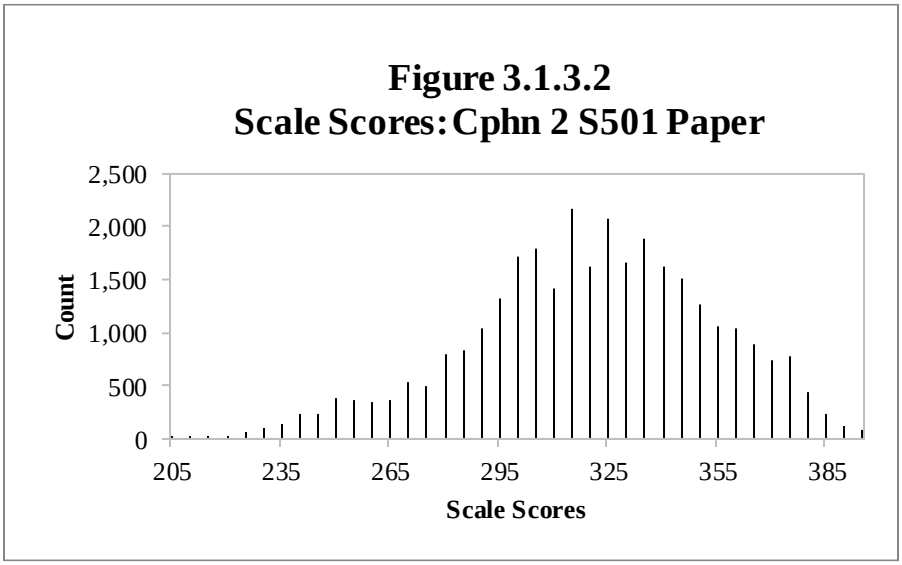


3.1.3.2 Grade 2

Table 3.1.3.2

Scale Score Descriptive Statistics: Cphn 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	31,407	205	395	321.98	33.25

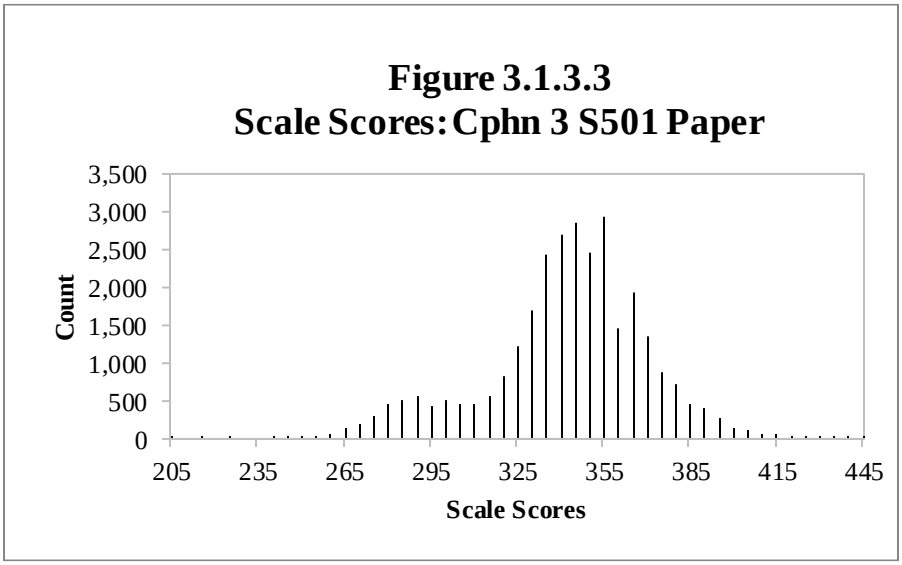


3.1.3.3 Grade 3

Table 3.1.3.3

Scale Score Descriptive Statistics: Cphn 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,735	206	449	344.19	28.15

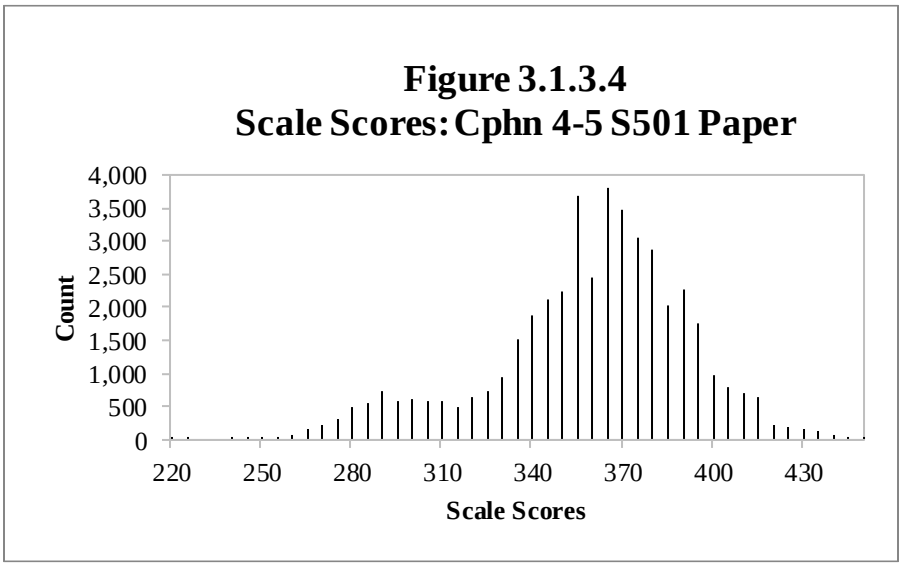


3.1.3.4 Grades 4–5

Table 3.1.3.4

Scale Score Descriptive Statistics: Cphn 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	23,426	221	449	356.83	31.70
5	21,289	243	454	365.17	33.72
Total	44,715	221	454	360.80	32.94

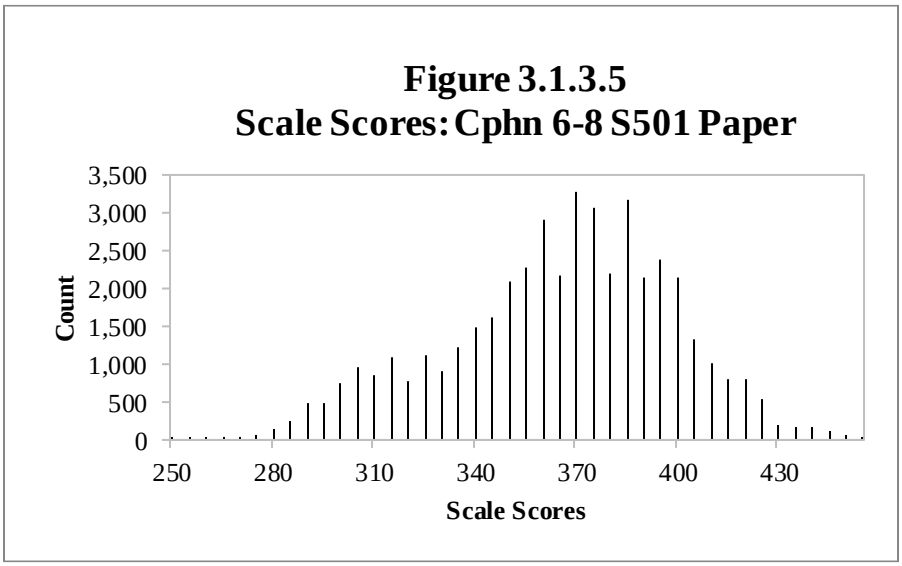


3.1.3.5 Grades 6–8

Table 3.1.3.5

Scale Score Descriptive Statistics: Cphn 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	16,617	265	454	362.72	30.49
7	14,634	254	459	367.03	33.94
8	13,849	258	459	372.09	35.86
Total	45,100	254	459	367.00	33.55

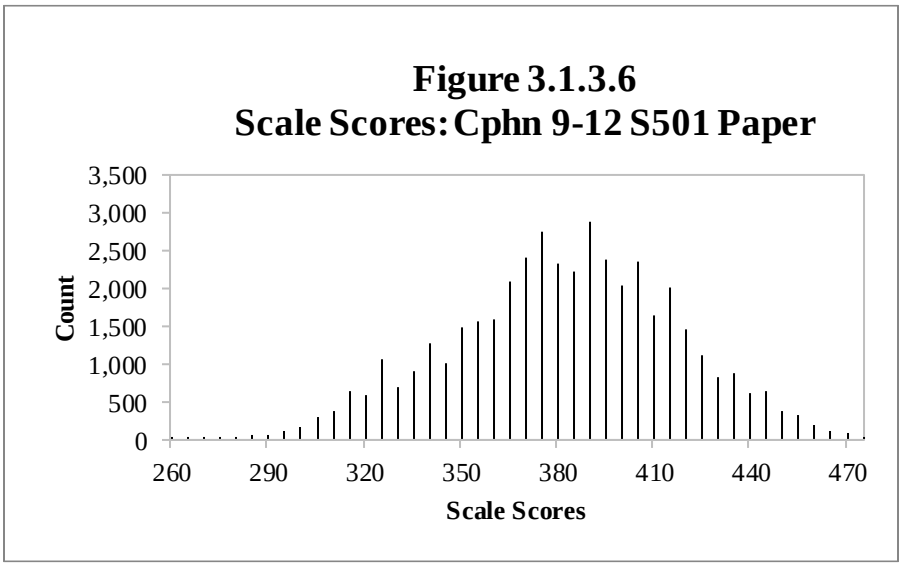


3.1.3.6 Grades 9–12

Table 3.1.3.6

Scale Score Descriptive Statistics: Cphn 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	13,578	262	479	377.39	35.87
10	12,230	273	479	383.96	34.74
11	10,720	283	479	389.39	34.11
12	7,189	274	479	388.50	31.94
Total	43,717	262	479	384.00	34.84

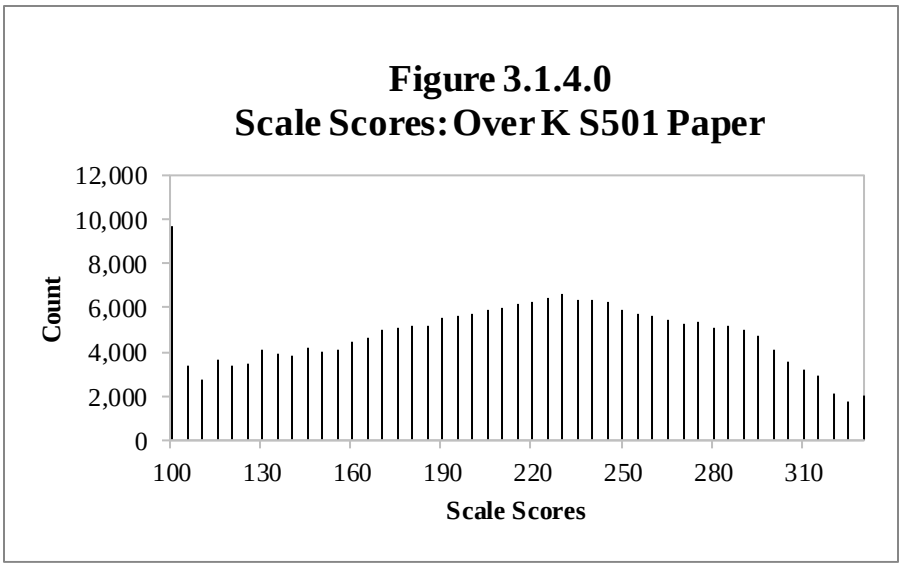


3.1.4 Overall

3.1.4.0 Kindergarten

Table 3.1.4.0
Scale Score Descriptive Statistics: Over K S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	225,978	100	333	214.28	61.98

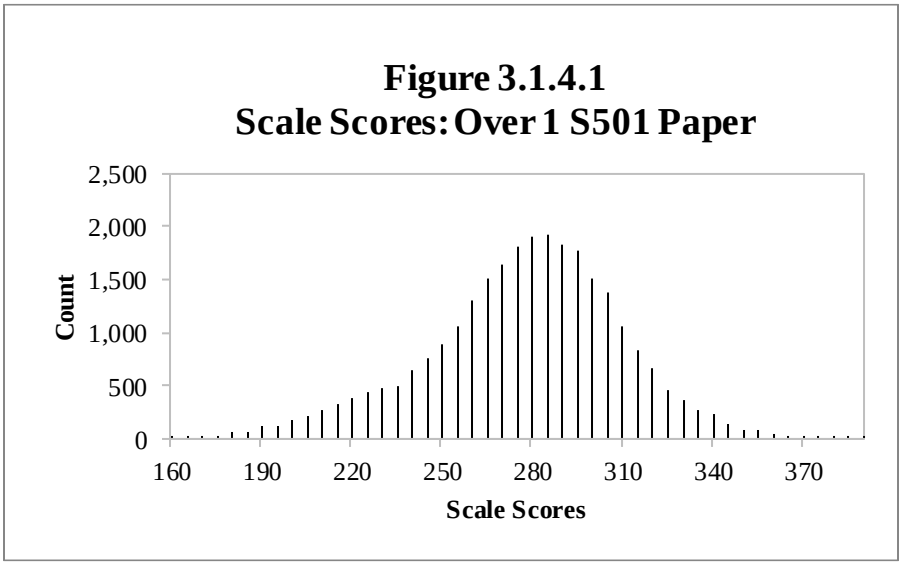


3.1.4.1 Grade 1

Table 3.1.4.1

Scale Score Descriptive Statistics: Over 1 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	27,403	164	394	279.20	32.05

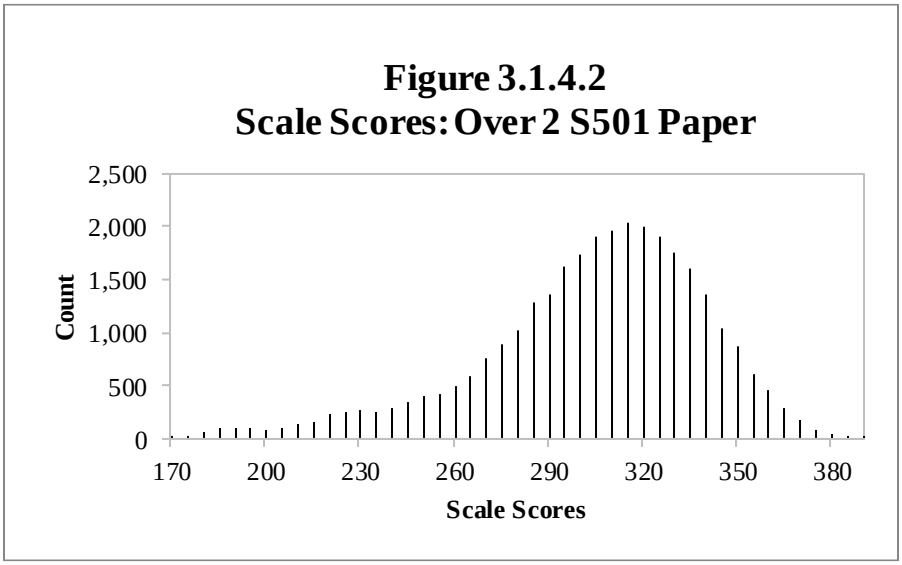


3.1.4.2 Grade 2

Table 3.1.4.2

Scale Score Descriptive Statistics: Over 2 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	31,166	170	394	305.99	35.87

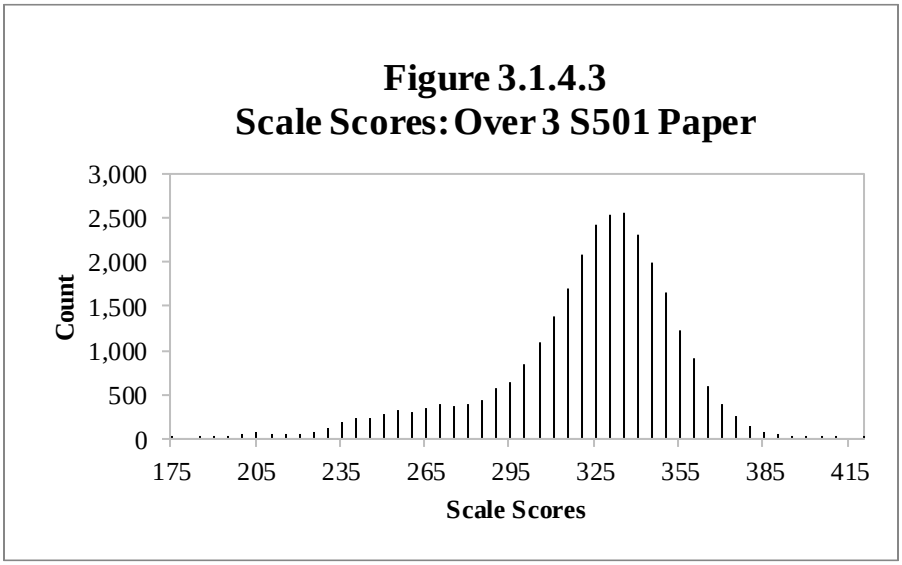


3.1.4.3 Grade 3

Table 3.1.4.3

Scale Score Descriptive Statistics: Over 3 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	29,512	178	424	323.49	32.45

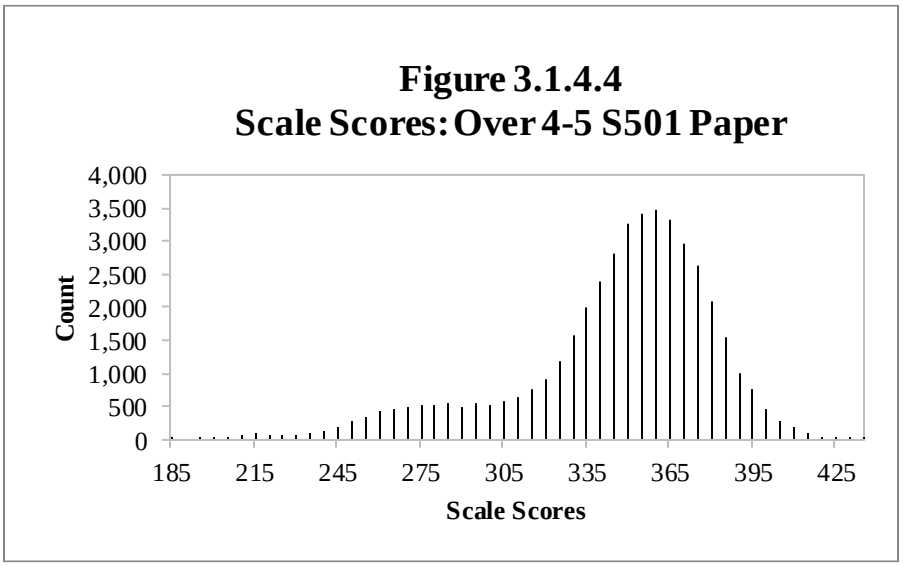


3.1.4.4 Grades 4–5

Table 3.1.4.4

Scale Score Descriptive Statistics: Over 4-5 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	23,250	187	435	342.78	35.67
5	21,121	199	437	351.18	37.45
Total	44,371	187	437	346.78	36.77

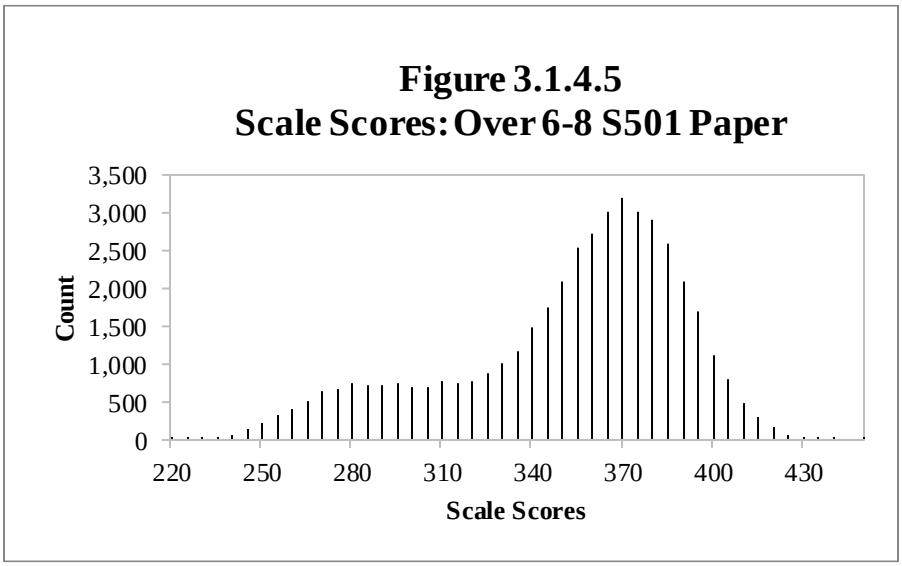


3.1.4.5 Grades 6–8

Table 3.1.4.5

Scale Score Descriptive Statistics: Over 6-8 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	16,459	221	434	350.38	36.55
7	14,502	227	452	352.97	39.74
8	13,724	228	440	356.88	41.12
Total	44,685	221	452	353.22	39.13

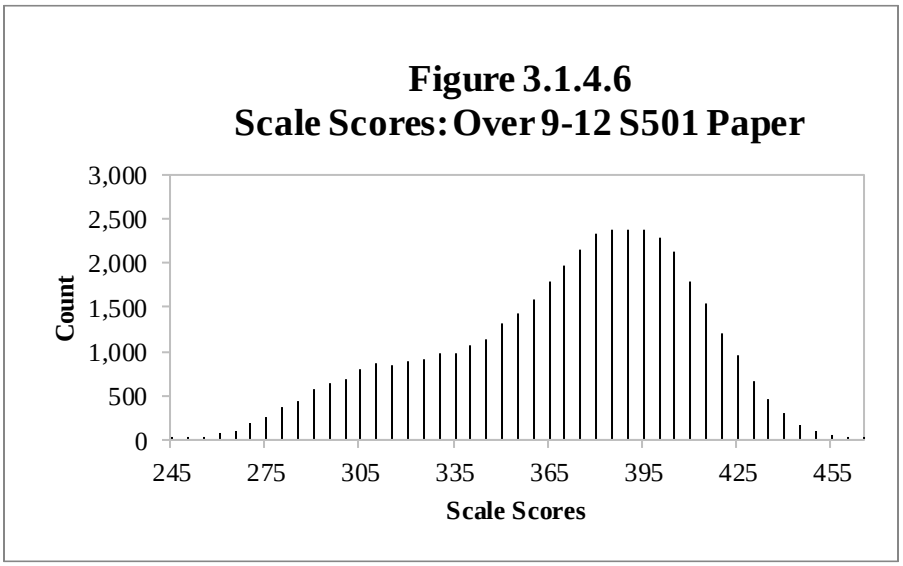


3.1.4.6 Grades 9–12

Table 3.1.4.6

Scale Score Descriptive Statistics: Over 9-12 S501 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	13,415	249	464	363.87	42.38
10	12,072	254	465	370.94	40.08
11	10,605	251	468	377.46	38.37
12	7,099	251	461	377.32	35.70
Total	43,191	249	468	371.39	40.13



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

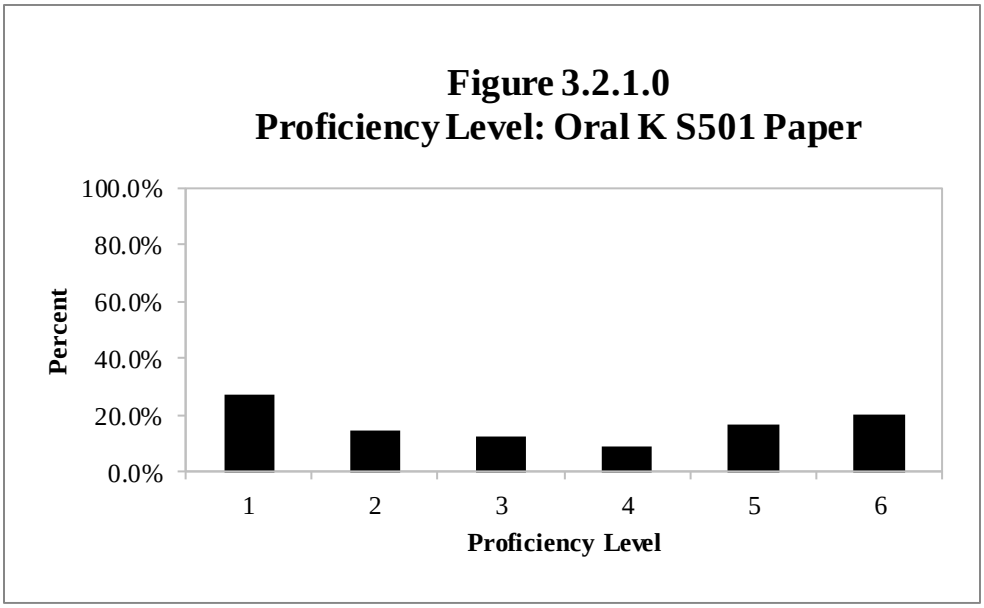
The percentage of students, out of the total number of students taking the form, who were placed into that proficiency level in the domain being tested

3.2.1 Oral

3.2.1.0 Kindergarten

Table 3.2.1.0
Proficiency Level Distribution: Oral K S501 Paper

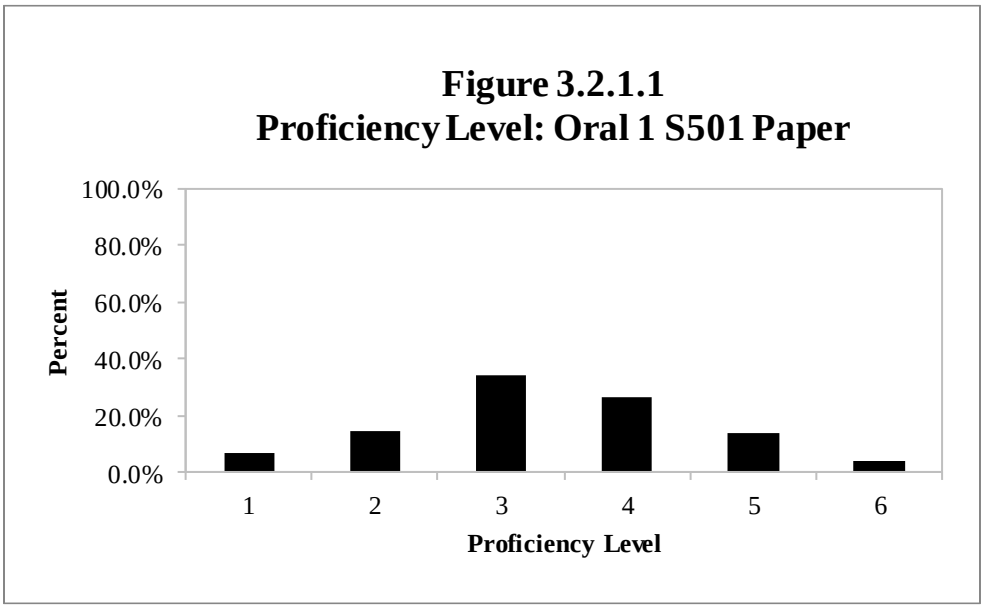
Level	Count	Percent
1	61,439	27.2%
2	32,546	14.4%
3	27,636	12.2%
4	21,007	9.3%
5	37,831	16.7%
6	45,538	20.1%
Total	225,997	100.0%



3.2.1.1 Grade 1

Table 3.2.1.1
Proficiency Level Distribution: Oral 1 S501 Paper

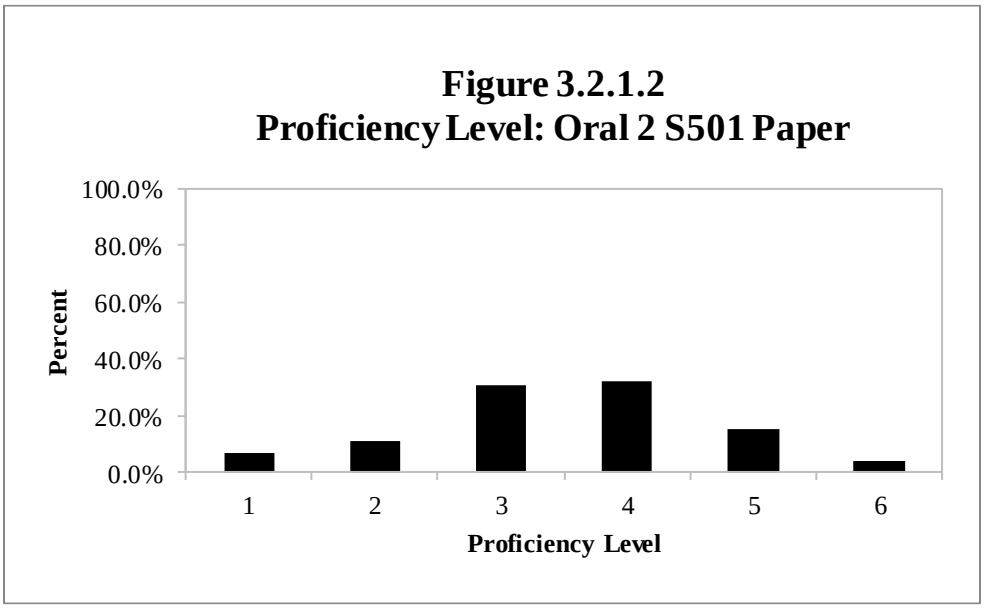
Level	Count	Percent
1	2,306	6.7%
2	4,914	14.4%
3	11,828	34.6%
4	9,040	26.4%
5	4,761	13.9%
6	1,332	3.9%
Total	34,181	100.0%



3.2.1.2 Grade 2

Table 3.2.1.2
Proficiency Level Distribution: Oral 2 S501 Paper

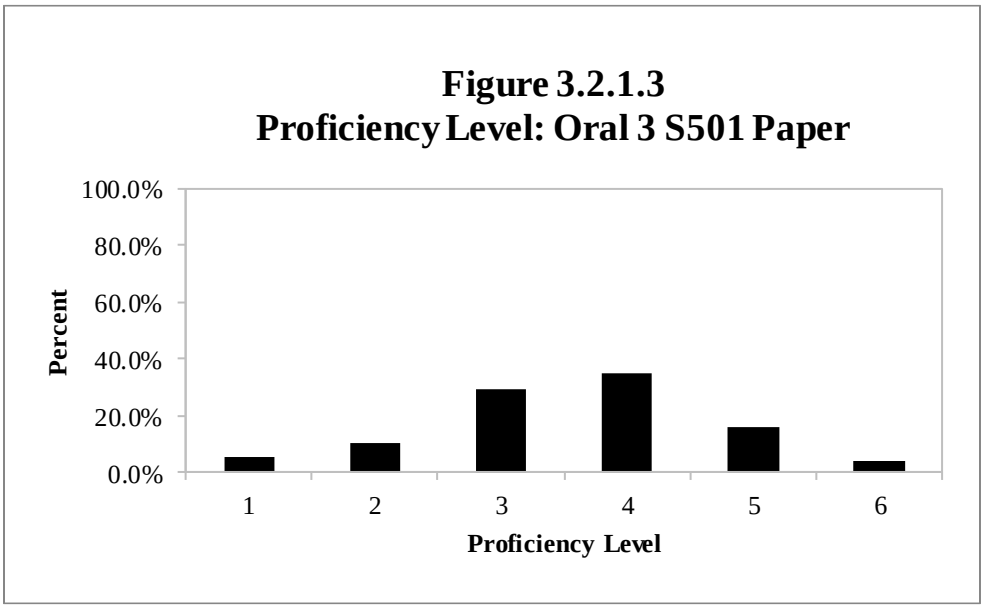
Level	Count	Percent
1	2,437	6.7%
2	4,104	11.3%
3	11,201	30.8%
4	11,589	31.9%
5	5,668	15.6%
6	1,380	3.8%
Total	36,379	100.0%



3.2.1.3 Grade 3

Table 3.2.1.3
Proficiency Level Distribution: Oral 3 S501 Paper

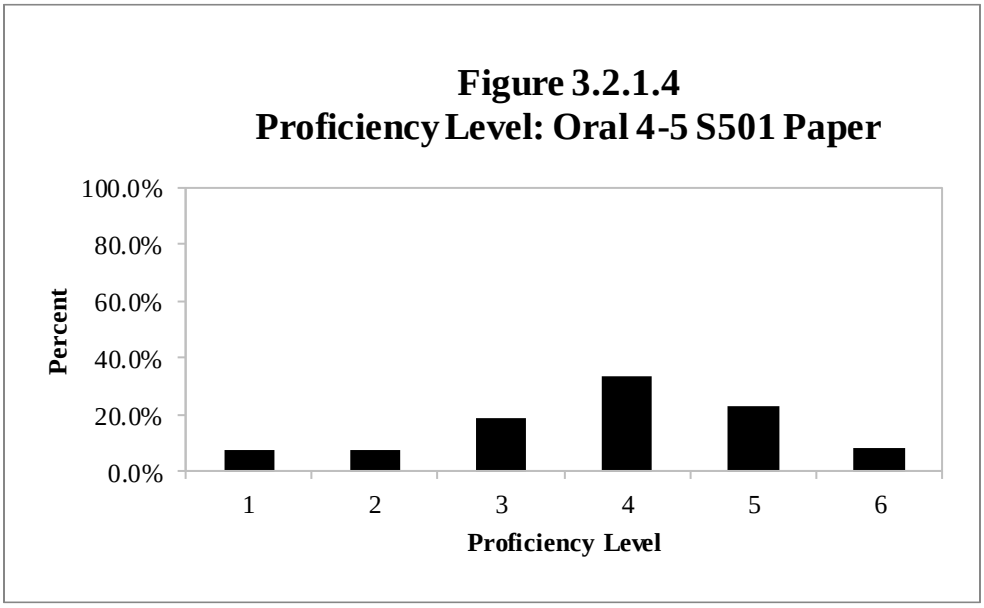
Level	Count	Percent
1	1,939	5.7%
2	3,454	10.1%
3	9,971	29.2%
4	11,932	34.9%
5	5,408	15.8%
6	1,491	4.4%
Total	34,195	100.0%



3.2.1.4 Grades 4–5

Table 3.2.1.4
Proficiency Level Distribution: Oral 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,869	7.1%	1,997	8.6%	3,866	7.8%
2	2,068	7.9%	1,850	8.0%	3,918	7.9%
3	5,010	19.2%	4,459	19.2%	9,469	19.2%
4	8,703	33.3%	7,714	33.3%	16,417	33.3%
5	5,962	22.8%	5,431	23.4%	11,393	23.1%
6	2,531	9.7%	1,719	7.4%	4,250	8.6%
Total	26,143	100.0%	23,170	100.0%	49,313	100.0%

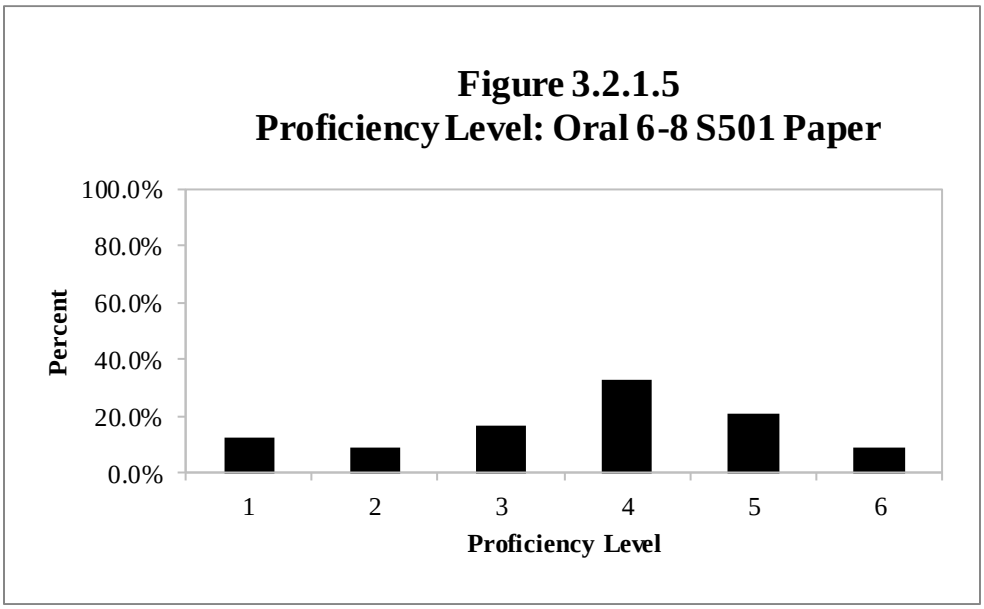


3.2.1.5 Grades 6–8

Table 3.2.1.5

Proficiency Level Distribution: Oral 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,855	9.7%	2,205	13.3%	2,294	14.7%	6,354	12.4%
2	1,593	8.4%	1,429	8.6%	1,471	9.4%	4,493	8.8%
3	3,134	16.4%	2,798	16.9%	2,665	17.0%	8,597	16.8%
4	6,396	33.5%	5,209	31.5%	5,096	32.6%	16,701	32.6%
5	4,292	22.5%	3,494	21.1%	2,824	18.0%	10,610	20.7%
6	1,805	9.5%	1,416	8.6%	1,299	8.3%	4,520	8.8%
Total	19,075	100.0%	16,551	100.0%	15,649	100.0%	51,275	100.0%

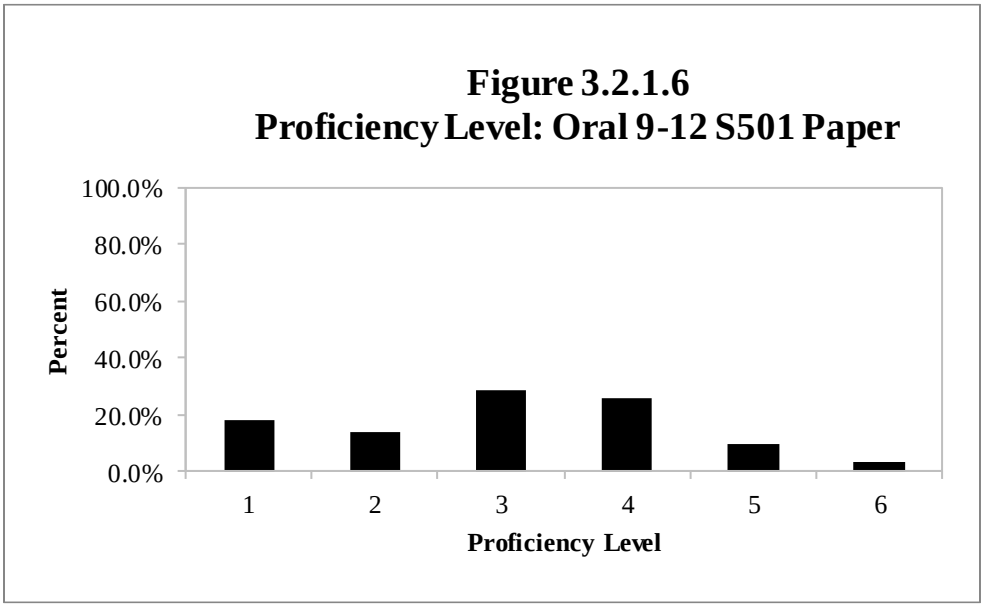


3.2.1.6 Grades 9–12

Table 3.2.1.6

Proficiency Level Distribution: Oral 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,095	20.3%	2,528	18.6%	1,978	16.8%	1,269	16.2%	8,870	18.3%
2	2,087	13.7%	1,804	13.3%	1,636	13.9%	1,117	14.3%	6,644	13.7%
3	3,756	24.6%	3,880	28.5%	3,490	29.6%	2,689	34.3%	13,815	28.5%
4	3,991	26.2%	3,492	25.7%	3,072	26.1%	1,929	24.6%	12,484	25.8%
5	1,648	10.8%	1,411	10.4%	1,190	10.1%	615	7.8%	4,864	10.0%
6	682	4.5%	484	3.6%	413	3.5%	217	2.8%	1,796	3.7%
Total	15,259	100.0%	13,599	100.0%	11,779	100.0%	7,836	100.0%	48,473	100.0%



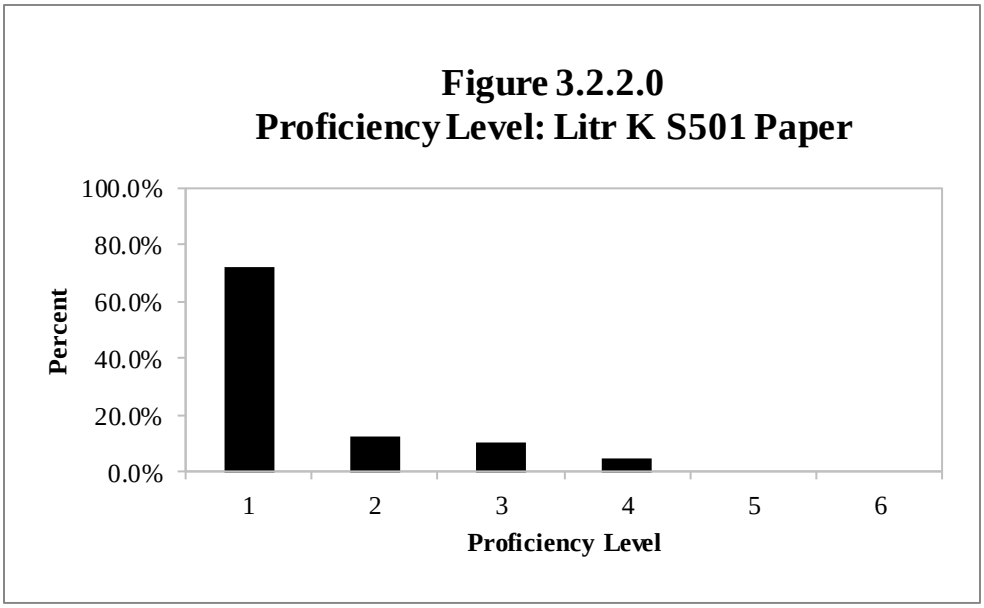
3.2.2 Literacy

3.2.2.0 Kindergarten

Table 3.2.2.0

Proficiency Level Distribution: Litr K S501 Paper

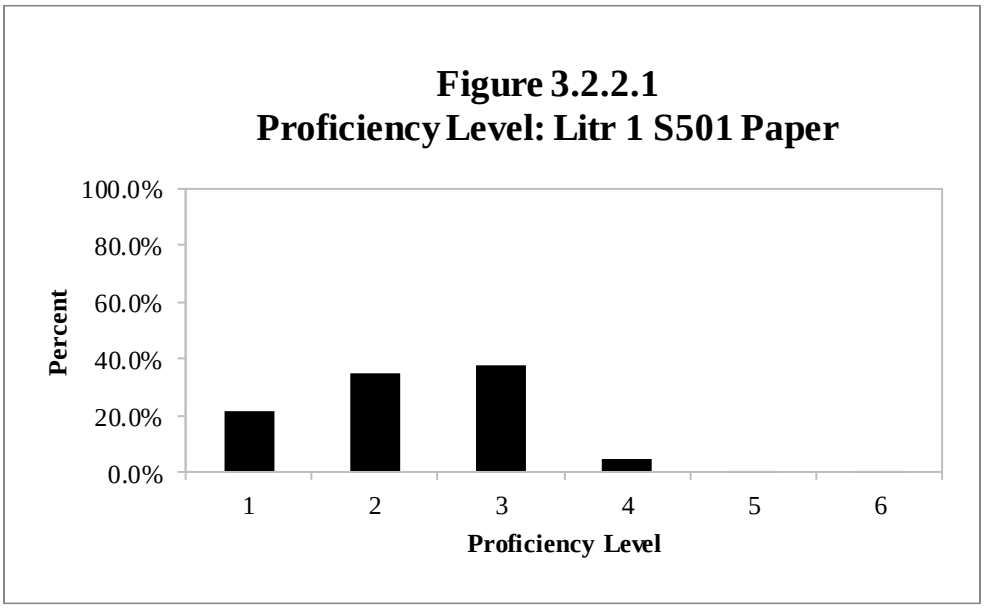
Level	Count	Percent
1	163,715	72.4%
2	28,119	12.4%
3	23,733	10.5%
4	10,415	4.6%
5	0	0.0%
6	0	0.0%
Total	225,982	100.0%



3.2.2.1 Grade 1

Table 3.2.2.1
Proficiency Level Distribution: Litr 1 S501 Paper

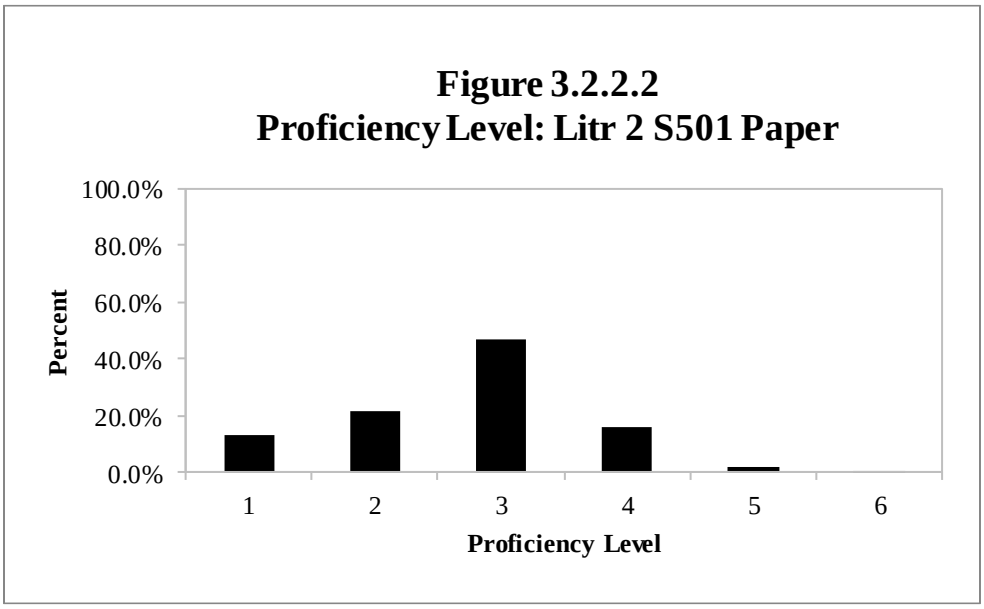
Level	Count	Percent
1	6,764	21.9%
2	10,730	34.8%
3	11,641	37.8%
4	1,499	4.9%
5	170	0.6%
6	19	0.1%
Total	30,823	100.0%



3.2.2.2 Grade 2

Table 3.2.2.2
Proficiency Level Distribution: Litr 2 S501 Paper

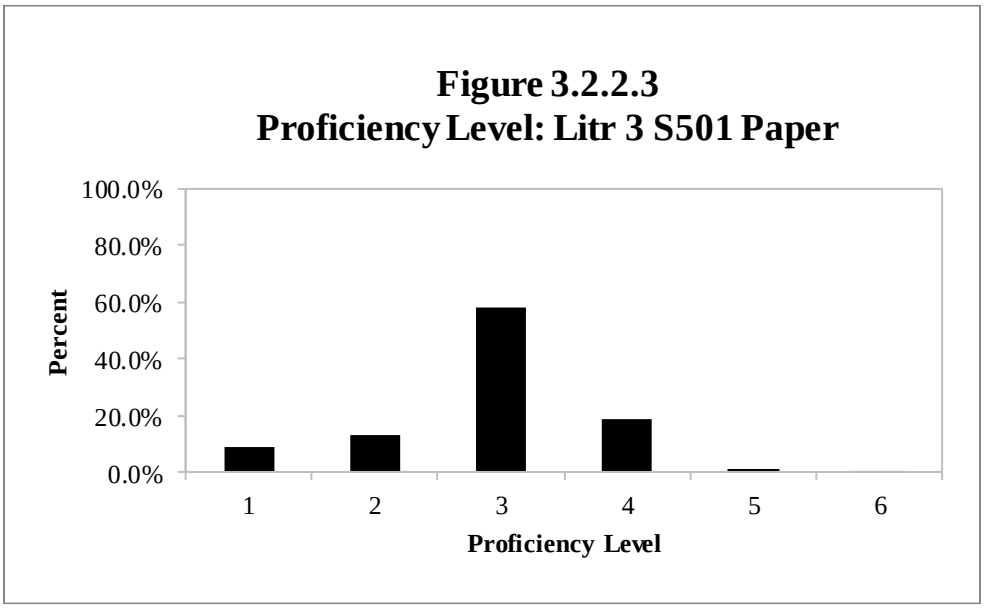
Level	Count	Percent
1	4,422	13.4%
2	7,105	21.5%
3	15,563	47.2%
4	5,263	15.9%
5	621	1.9%
6	26	0.1%
Total	33,000	100.0%



3.2.2.3 Grade 3

Table 3.2.2.3
Proficiency Level Distribution: Litr 3 S501 Paper

Level	Count	Percent
1	2,791	9.0%
2	4,090	13.1%
3	18,055	58.0%
4	5,791	18.6%
5	367	1.2%
6	36	0.1%
Total	31,130	100.0%

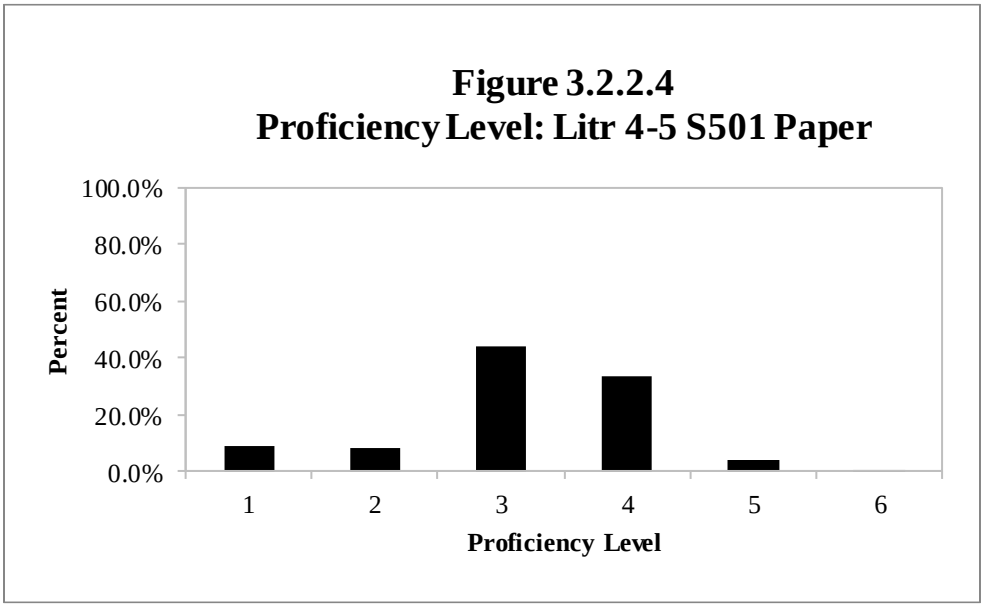


3.2.2.4 Grades 4–5

Table 3.2.2.4

Proficiency Level Distribution: Litr 4-5 S501 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	2,124	8.8%	1,984	9.1%	4,108	8.9%
2	1,952	8.1%	1,798	8.2%	3,750	8.2%
3	11,435	47.4%	8,994	41.1%	20,429	44.4%
4	7,679	31.8%	7,924	36.2%	15,603	33.9%
5	835	3.5%	1,100	5.0%	1,935	4.2%
6	103	0.4%	81	0.4%	184	0.4%
Total	24,128	100.0%	21,881	100.0%	46,009	100.0%

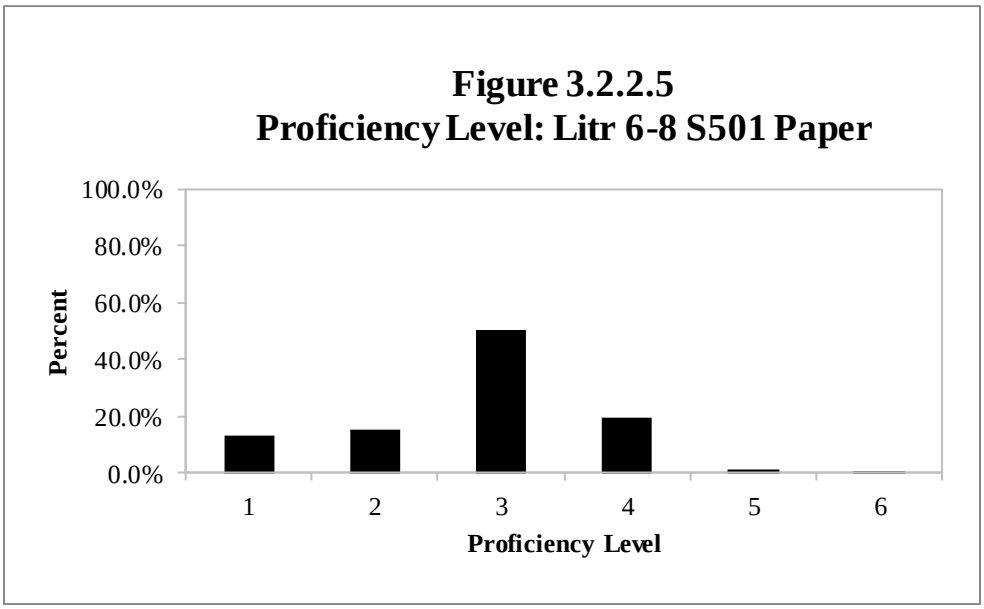


3.2.2.5 Grades 6–8

Table 3.2.2.5

Proficiency Level Distribution: Litr 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,919	11.2%	2,046	13.6%	2,186	15.2%	6,151	13.2%
2	2,467	14.4%	2,397	15.9%	2,404	16.7%	7,268	15.6%
3	9,133	53.2%	7,602	50.5%	6,836	47.6%	23,571	50.6%
4	3,506	20.4%	2,863	19.0%	2,768	19.3%	9,137	19.6%
5	132	0.8%	155	1.0%	162	1.1%	449	1.0%
6	6	0.0%	5	0.0%	0	0.0%	11	0.0%
Total	17,163	100.0%	15,068	100.0%	14,356	100.0%	46,587	100.0%

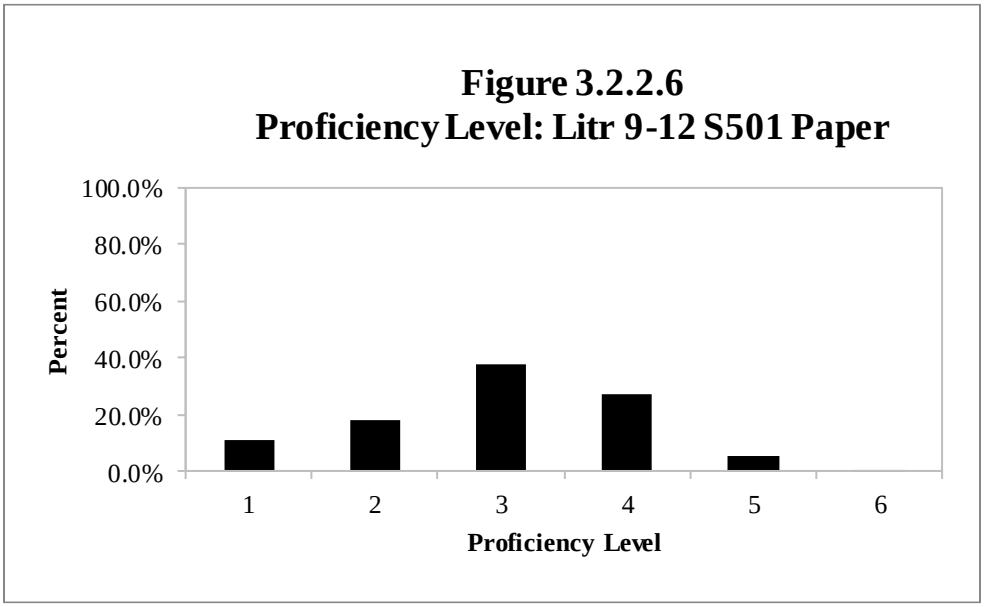


3.2.2.6 Grades 9–12

Table 3.2.2.6

Proficiency Level Distribution: Litr 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,710	12.1%	1,337	10.5%	1,099	9.9%	916	12.3%	5,062	11.1%
2	2,564	18.1%	2,165	17.1%	1,946	17.5%	1,577	21.2%	8,252	18.2%
3	4,930	34.8%	4,838	38.1%	4,320	38.8%	3,258	43.7%	17,346	38.2%
4	4,142	29.2%	3,627	28.6%	3,089	27.8%	1,439	19.3%	12,297	27.1%
5	806	5.7%	717	5.6%	662	6.0%	263	3.5%	2,448	5.4%
6	31	0.2%	13	0.1%	5	0.0%	1	0.0%	50	0.1%
Total	14,183	100.0%	12,697	100.0%	11,121	100.0%	7,454	100.0%	45,455	100.0%

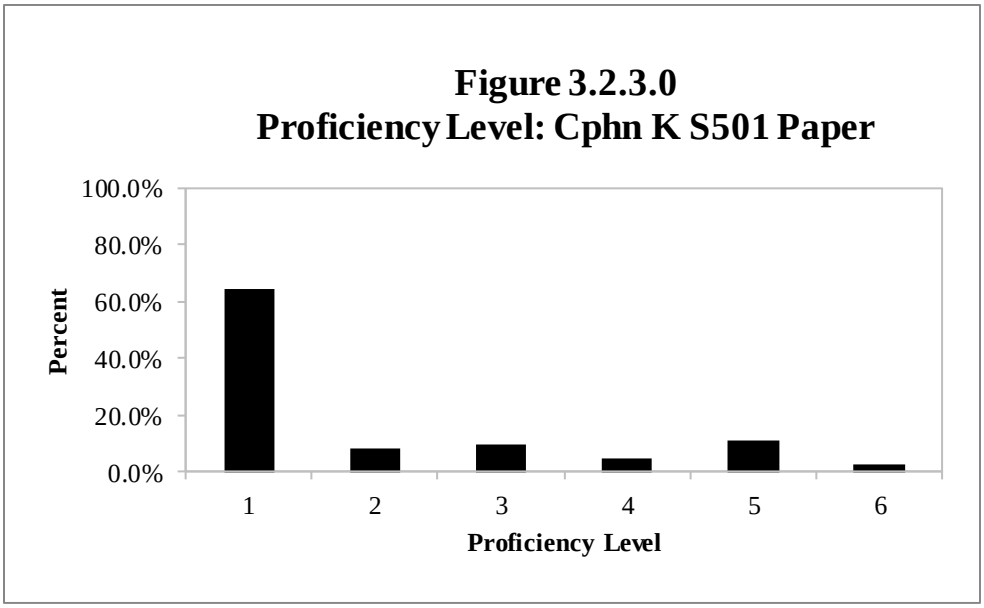


3.2.3 Comprehension

3.2.3.0 Kindergarten

Table 3.2.3.0
Proficiency Level Distribution: Cphn K S501 Paper

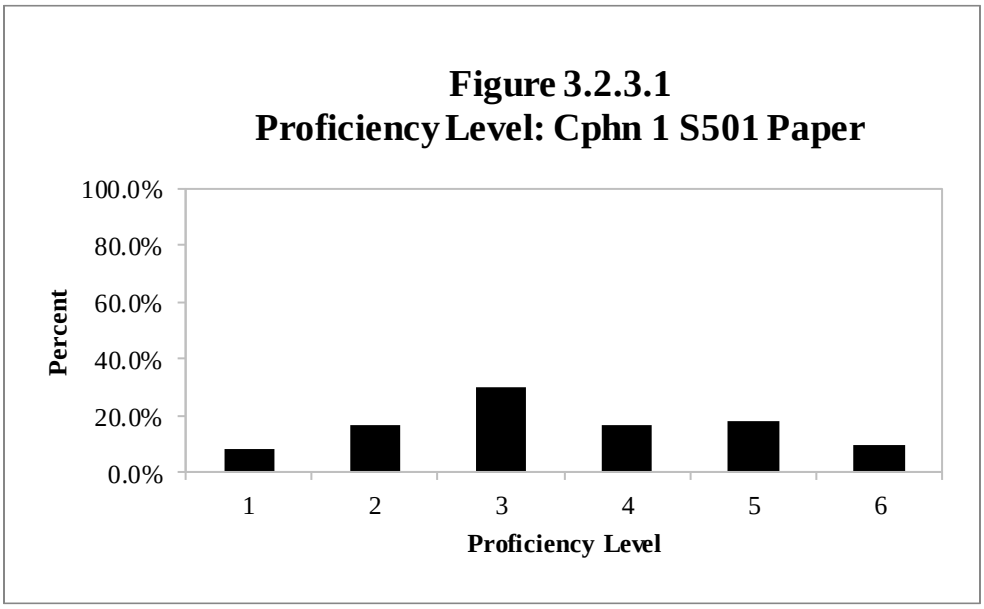
Level	Count	Percent
1	145,534	64.4%
2	18,326	8.1%
3	21,214	9.4%
4	10,761	4.8%
5	24,456	10.8%
6	5,699	2.5%
Total	225,990	100.0%



3.2.3.1 Grade 1

Table 3.2.3.1
Proficiency Level Distribution: Cphn 1 S501 Paper

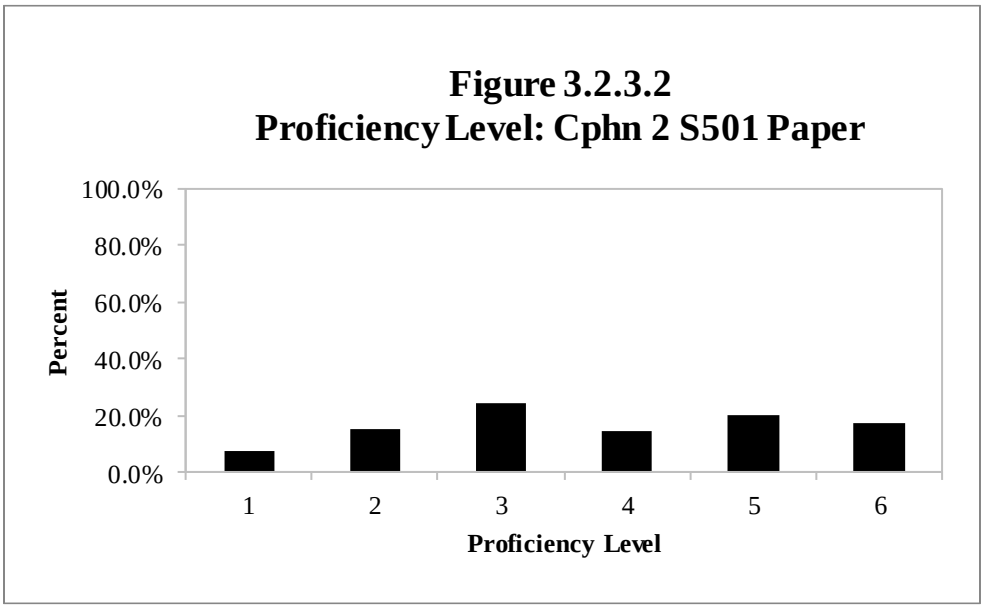
Level	Count	Percent
1	2,391	8.7%
2	4,579	16.6%
3	8,298	30.0%
4	4,628	16.8%
5	4,974	18.0%
6	2,754	10.0%
Total	27,624	100.0%



3.2.3.2 Grade 2

Table 3.2.3.2
Proficiency Level Distribution: Cphn 2 S501 Paper

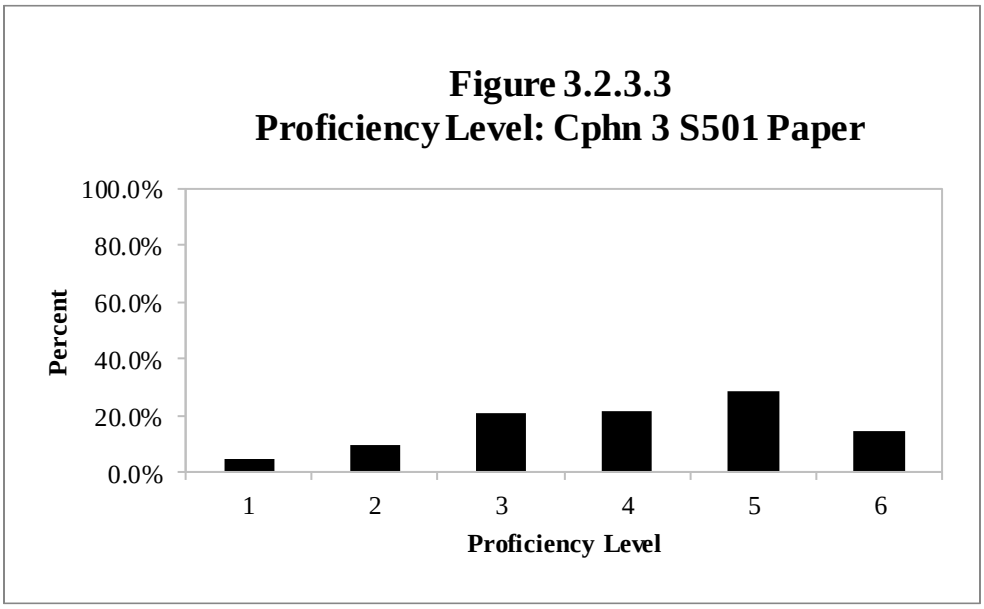
Level	Count	Percent
1	2,474	7.9%
2	4,795	15.3%
3	7,793	24.8%
4	4,692	14.9%
5	6,279	20.0%
6	5,374	17.1%
Total	31,407	100.0%



3.2.3.3 Grade 3

Table 3.2.3.3
Proficiency Level Distribution: Cphn 3 S501 Paper

Level	Count	Percent
1	1,384	4.7%
2	2,854	9.6%
3	6,228	20.9%
4	6,415	21.6%
5	8,517	28.6%
6	4,337	14.6%
Total	29,735	100.0%

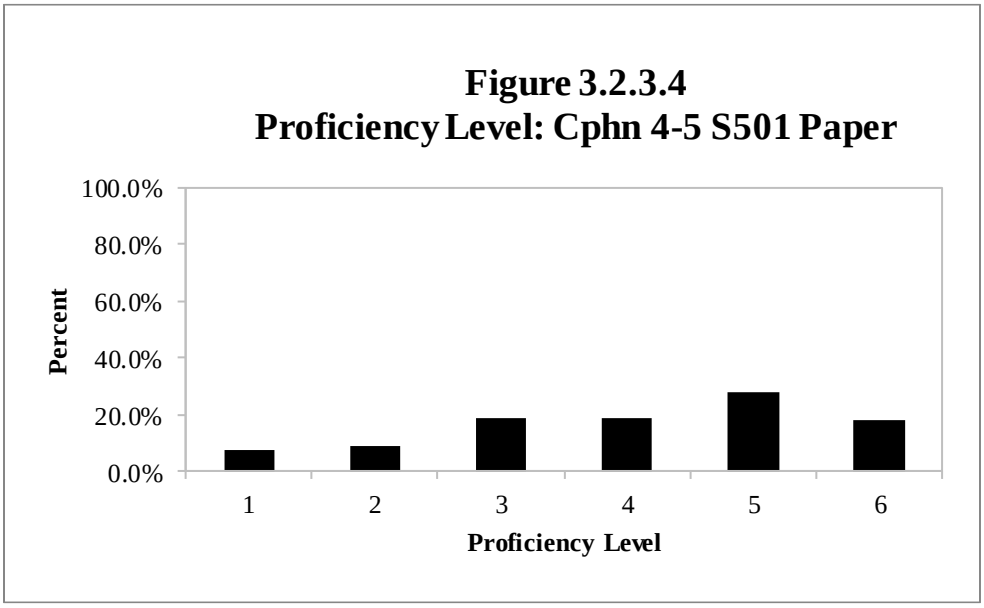


3.2.3.4 Grades 4–5

Table 3.2.3.4

Proficiency Level Distribution: Cphn 4-5 S501 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	1,558	6.7%	1,742	8.2%	3,300	7.4%
2	2,012	8.6%	1,911	9.0%	3,923	8.8%
3	4,439	18.9%	4,108	19.3%	8,547	19.1%
4	4,636	19.8%	3,722	17.5%	8,358	18.7%
5	6,691	28.6%	5,846	27.5%	12,537	28.0%
6	4,090	17.5%	3,960	18.6%	8,050	18.0%
Total	23,426	100.0%	21,289	100.0%	44,715	100.0%

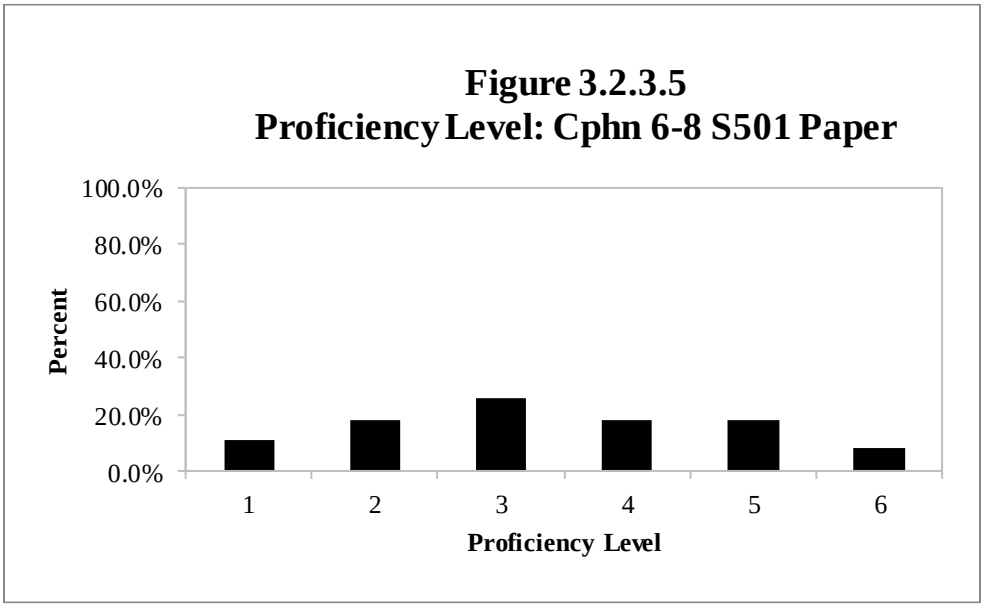


3.2.3.5 Grades 6–8

Table 3.2.3.5

Proficiency Level Distribution: Cphn 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,432	8.6%	1,773	12.1%	1,925	13.9%	5,130	11.4%
2	2,851	17.2%	2,693	18.4%	2,610	18.8%	8,154	18.1%
3	4,662	28.1%	3,850	26.3%	3,258	23.5%	11,770	26.1%
4	3,319	20.0%	2,527	17.3%	2,383	17.2%	8,229	18.2%
5	3,091	18.6%	2,545	17.4%	2,546	18.4%	8,182	18.1%
6	1,262	7.6%	1,246	8.5%	1,127	8.1%	3,635	8.1%
Total	16,617	100.0%	14,634	100.0%	13,849	100.0%	45,100	100.0%

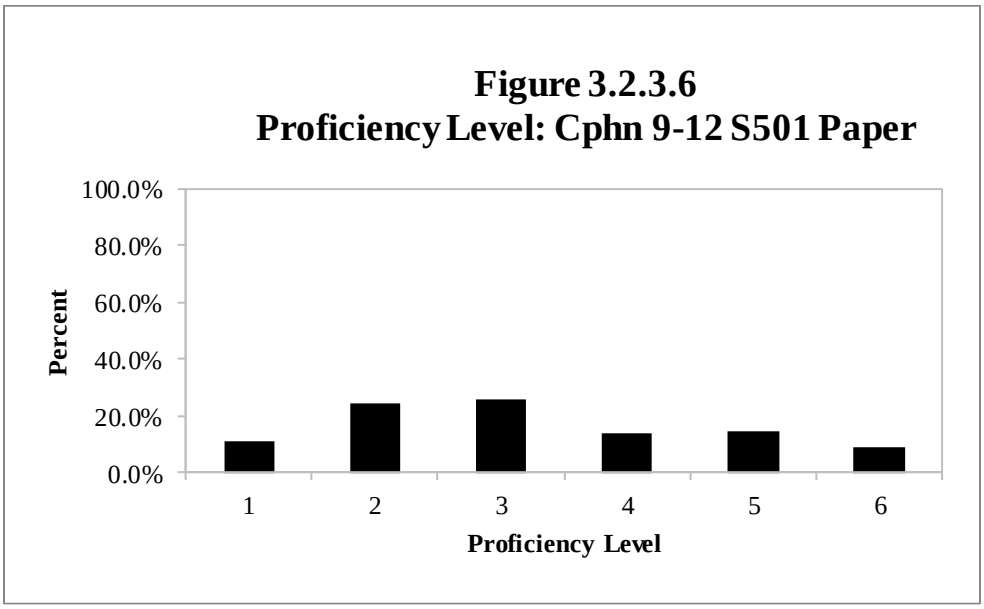


3.2.3.6 Grades 9–12

Table 3.2.3.6

Proficiency Level Distribution: Cphn 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,655	12.2%	1,390	11.4%	1,107	10.3%	724	10.1%	4,876	11.2%
2	3,086	22.7%	2,660	21.7%	2,685	25.0%	2,247	31.3%	10,678	24.4%
3	3,429	25.3%	3,323	27.2%	2,690	25.1%	2,000	27.8%	11,442	26.2%
4	1,989	14.6%	1,819	14.9%	1,524	14.2%	903	12.6%	6,235	14.3%
5	2,116	15.6%	1,857	15.2%	1,651	15.4%	797	11.1%	6,421	14.7%
6	1,303	9.6%	1,181	9.7%	1,063	9.9%	518	7.2%	4,065	9.3%
Total	13,578	100.0%	12,230	100.0%	10,720	100.0%	7,189	100.0%	43,717	100.0%

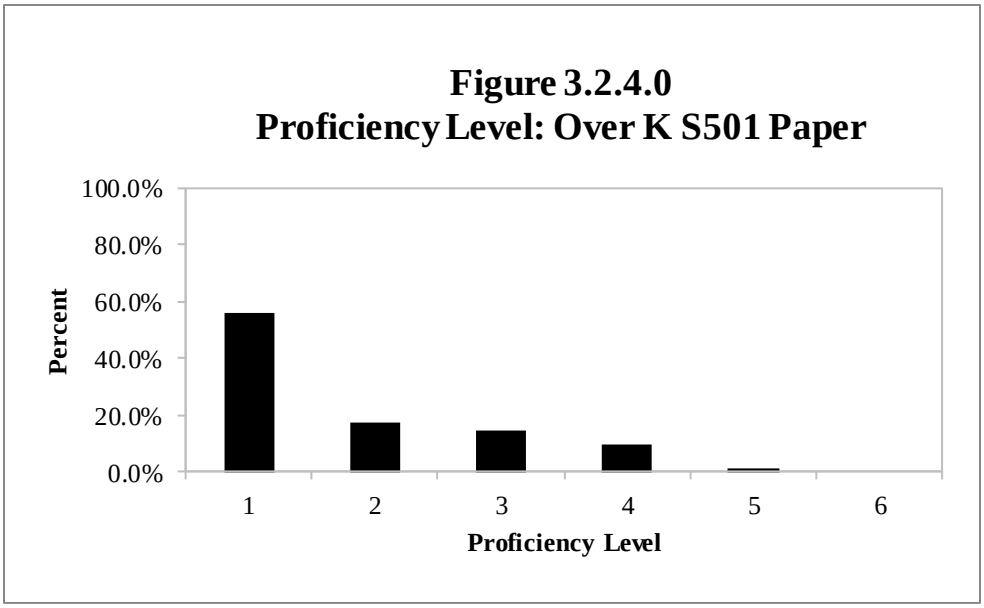


3.2.4 Overall

3.2.4.0 Kindergarten

Table 3.2.4.0
Proficiency Level Distribution: Over K S501 Paper

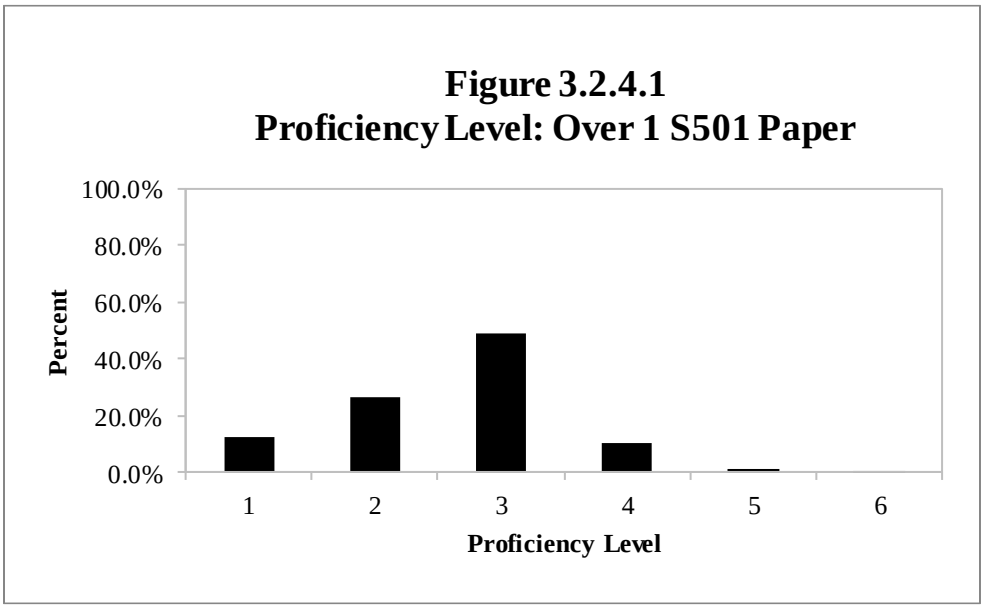
Level	Count	Percent
1	126,488	56.0%
2	39,355	17.4%
3	33,870	15.0%
4	22,566	10.0%
5	3,699	1.6%
6	0	0.0%
Total	225,978	100.0%



3.2.4.1 Grade 1

Table 3.2.4.1
Proficiency Level Distribution: Over 1 S501 Paper

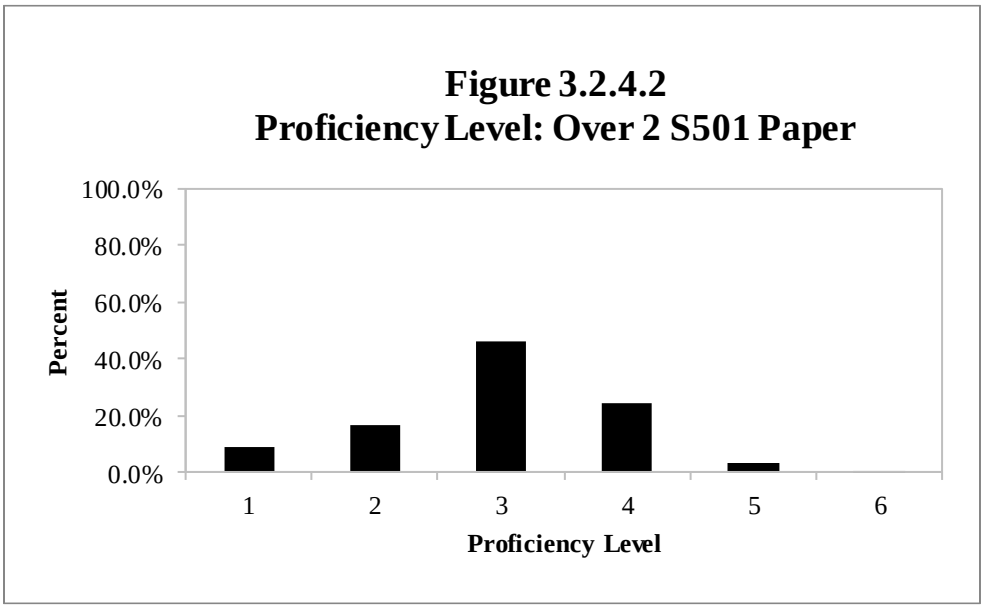
Level	Count	Percent
1	3,400	12.4%
2	7,258	26.5%
3	13,529	49.4%
4	2,795	10.2%
5	392	1.4%
6	29	0.1%
Total	27,403	100.0%



3.2.4.2 Grade 2

Table 3.2.4.2
Proficiency Level Distribution: Over 2 S501 Paper

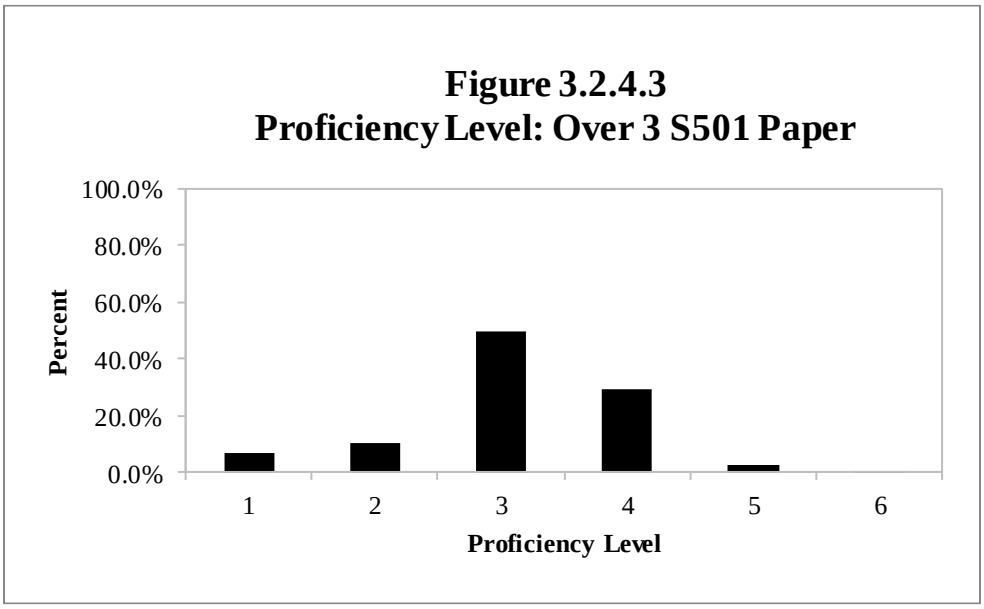
Level	Count	Percent
1	2,794	9.0%
2	5,227	16.8%
3	14,460	46.4%
4	7,545	24.2%
5	1,099	3.5%
6	41	0.1%
Total	31,166	100.0%



3.2.4.3 Grade 3

Table 3.2.4.3
Proficiency Level Distribution: Over 3 S501 Paper

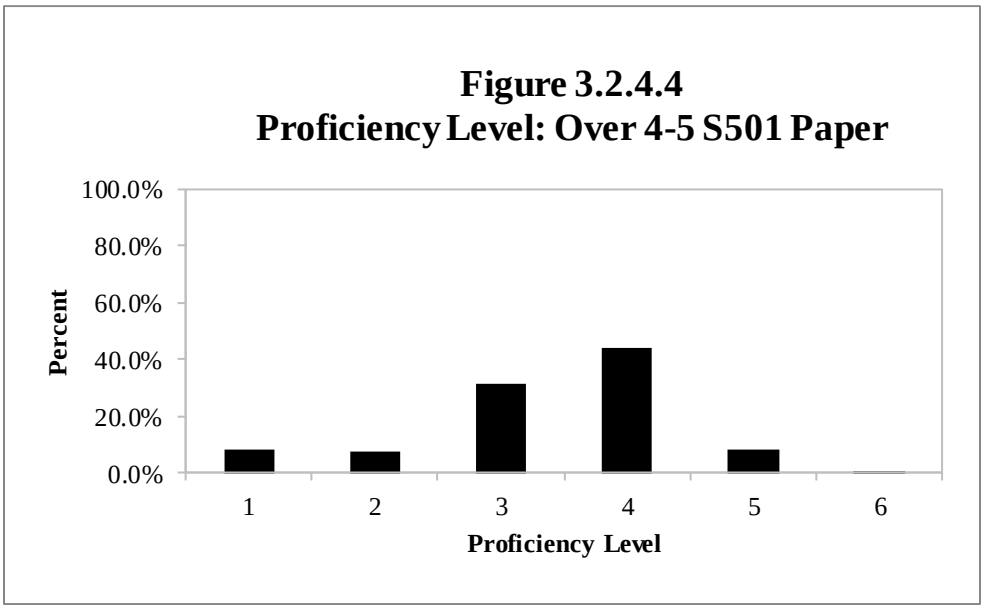
Level	Count	Percent
1	2,055	7.0%
2	3,163	10.7%
3	14,632	49.6%
4	8,758	29.7%
5	852	2.9%
6	52	0.2%
Total	29,512	100.0%



3.2.4.4 Grades 4–5

Table 3.2.4.4
Proficiency Level Distribution: Over 4-5 S501 Paper

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,824	7.8%	1,796	8.5%	3,620	8.2%
2	1,715	7.4%	1,575	7.5%	3,290	7.4%
3	7,677	33.0%	6,329	30.0%	14,006	31.6%
4	10,162	43.7%	9,454	44.8%	19,616	44.2%
5	1,711	7.4%	1,865	8.8%	3,576	8.1%
6	161	0.7%	102	0.5%	263	0.6%
Total	23,250	100.0%	21,121	100.0%	44,371	100.0%

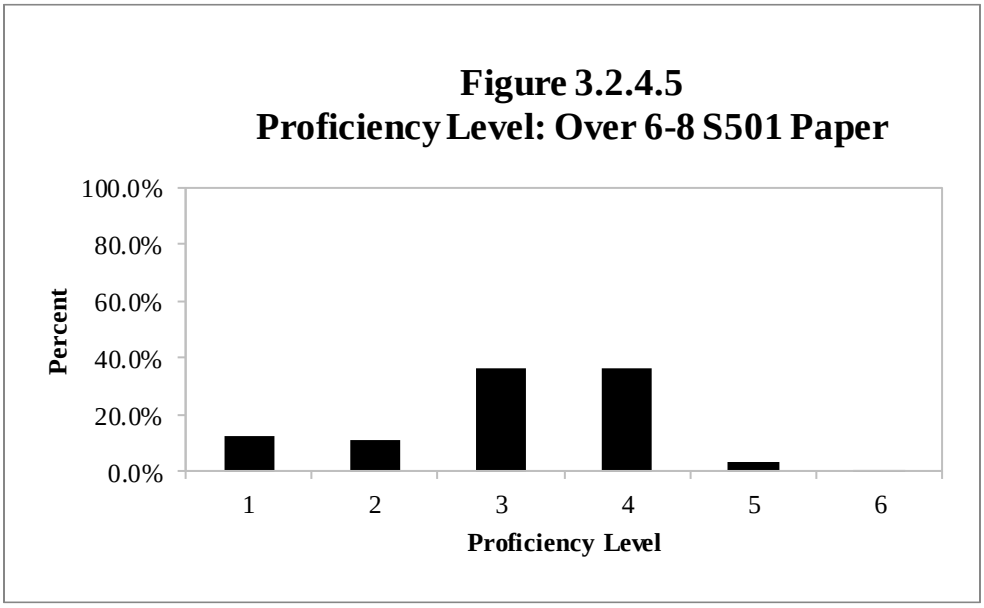


3.2.4.5 Grades 6–8

Table 3.2.4.5

Proficiency Level Distribution: Over 6-8 S501 Paper

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	1,658	10.1%	1,927	13.3%	2,008	14.6%	5,593	12.5%
2	1,701	10.3%	1,619	11.2%	1,661	12.1%	4,981	11.1%
3	6,101	37.1%	5,262	36.3%	4,801	35.0%	16,164	36.2%
4	6,434	39.1%	5,172	35.7%	4,777	34.8%	16,383	36.7%
5	552	3.4%	513	3.5%	473	3.4%	1,538	3.4%
6	13	0.1%	9	0.1%	4	0.0%	26	0.1%
Total	16,459	100.0%	14,502	100.0%	13,724	100.0%	44,685	100.0%

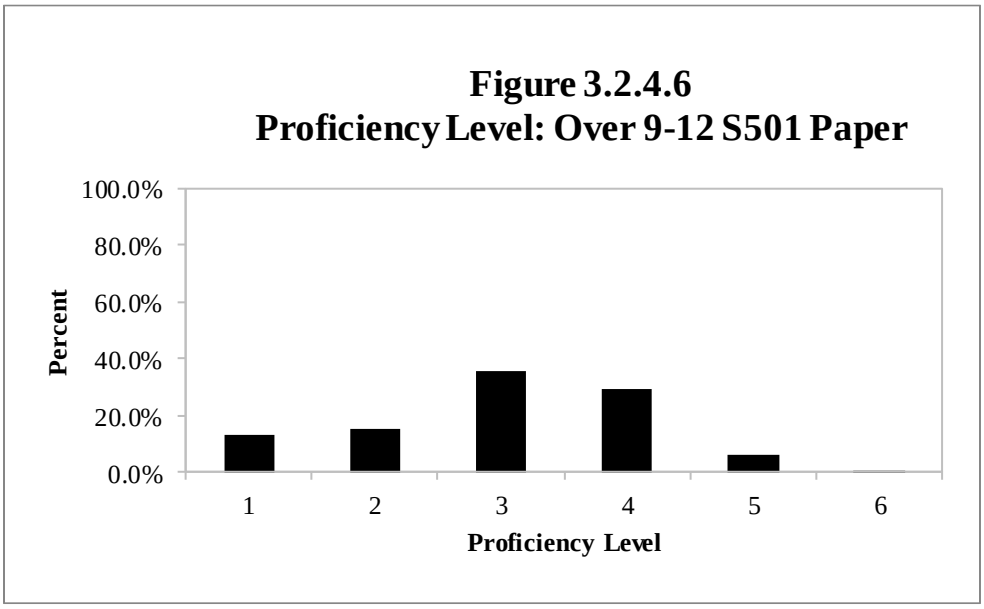


3.2.4.6 Grades 9–12

Table 3.2.4.6

Proficiency Level Distribution: Over 9-12 S501 Paper

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,019	15.1%	1,578	13.1%	1,230	11.6%	841	11.8%	5,668	13.1%
2	2,029	15.1%	1,770	14.7%	1,587	15.0%	1,332	18.8%	6,718	15.6%
3	4,277	31.9%	4,277	35.4%	3,967	37.4%	2,995	42.2%	15,516	35.9%
4	4,102	30.6%	3,691	30.6%	3,145	29.7%	1,652	23.3%	12,590	29.1%
5	938	7.0%	736	6.1%	669	6.3%	279	3.9%	2,622	6.1%
6	50	0.4%	20	0.2%	7	0.1%	0	0.0%	77	0.2%
Total	13,415	100.0%	12,072	100.0%	10,605	100.0%	7,099	100.0%	43,191	100.0%



4 Annual Updates of Validity Evidence

This section presents studies conducted as validity evidence for the WIDA ACCESS assessments. According to the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 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. 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) relation 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 judgment 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 subgroups 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' performance 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. Relation to Other Variables

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 the accuracy with which 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.

Section 4.2.1, English Learner Reclassification Study–Phase 1, addresses the validity of using the ACCESS test to reclassify EL learners for exiting from the supporting programs.

4.2. Annual Validity Studies

4.2.1. English Learner Reclassification Study–Phase 1

Kim, A., Ho, P., Chapman, M., & Cook, H. G. (2020a). *Examination of reclassification decisions made for K–12 English learners: Survey report of Delaware* (WIDA Internal Report). Madison, WI: WIDA at the Wisconsin Center for Education Research.

Kim, A., Ho, P., Chapman, M., & Cook, H. G. (2020b). *Examination of reclassification decisions made for K–12 English learners: Survey report of Pennsylvania* (WIDA Internal Report). Madison, WI: WIDA at the Wisconsin Center for Education Research.

This survey study investigated how English learners (ELs) are reclassified across districts in select WIDA Consortium member states. Despite the high-stakes nature of the reclassification decision, little is known regarding the decision-making process across WIDA states. A pilot survey was distributed across districts in Vermont in spring of 2019; findings were used to update the main survey. The revised survey consisted of five sections: (1) educator background information, (2) reclassification criteria, (3) reclassification procedures and decision-makers, (4) reclassification monitoring, and (5) perceived effectiveness of reclassification.

Two states—Delaware and Pennsylvania—were recruited for the main study (Kim, Ho, Chapman, & Cook, 2020a, 2020b). According to its reclassification policy, Delaware uses only English language proficiency assessment scores, whereas Pennsylvania uses both English language proficiency assessment scores and teacher judgments on students' classroom language proficiency. Online surveys were distributed across districts in September to October 2019.

Collected data were primarily analyzed using descriptive analyses. Open-ended responses were qualitatively analyzed for emerging patterns.

Results from Pennsylvania indicated that EL reclassification criteria varied across districts (Kim, et al., 2020b). The state's policy requires a minimum of two criteria for making reclassification decisions: ELs' scores on an English language proficiency assessment (ACCESS for ELLs) and educator input (standardized language use inventory). Findings indicated that over half of the districts (65%) used three or more criteria for EL reclassification, for example, students' writing samples, performance in content areas, and grade-point average. Such variability in the number and types of criteria could potentially result in ELs qualifying for reclassification in one district but not in others.

Survey findings also indicated that reclassification decisions were either made by a single decision-maker (37%) or through a reclassification meeting (46%) attended by several educators. In either case, district EL/Title III coordinators and EL/Bilingual program directors were often the primary decision-makers for EL reclassification. Although few educators believed that ELs were inappropriately reclassified, students' disability status was considered the main factor leading to inappropriate reclassification. Overall, these results suggest that the majority of Pennsylvania districts and schools exercise local autonomy regarding EL reclassification, creating wide variability in decision-making across districts. Furthermore, these findings from Phase 1 will guide Phase 2 of the study (see Phase 2 under Ongoing Research).

In the survey, educators shared their suggestions for improving EL reclassification. They requested more targeted training from the state. Examples included more training for content teachers, who were not as familiar with English language proficiency terminology and concepts, and more professional development on reclassifying ELs with disabilities. Some educators also believed that ACCESS for ELLs could be enhanced by ensuring that its Speaking domain better reflects students' actual speaking language ability.

5 Reliability

In accordance with the *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014), in interpreting test scores, it is important to evaluate their reliability, as the interpretation of test scores depends on assumptions that students exhibit some degree of consistency in their scores across independent administrations of the same testing procedure. It is expected that students mastering the domain will consistently perform well and those who have not mastered the domain will consistently perform less well, regardless of the particular sample of items and tasks used to assess students. Furthermore, because it is assumed that all items on such a test measure some aspect of the domain of interest, it is expected that students will perform consistently across different items and tasks measuring the same ability within the test. Therefore, it is important to evaluate the degree to which students' test scores are consistent across replications of the same testing condition.

However, different samples of performances from the same student are rarely identical. A student's responses to sets of test questions or tasks vary from one sample of test questions or tasks targeting the domain to another, and from one occasion to another, even under strictly controlled conditions. In addition, different raters may award different scores to the same student performance on a test task. These sources of variation are reflected in the students' scores. Therefore, it is important to evaluate the extent to which differences in students' test scores reflect true differences in the knowledge, skills, or ability being tested, rather than fluctuations due to chance.

The reliability of the test scores depends on how much the scores vary across replications of the testing procedure, and analyses of reliability depend on the types of variability likely to be of concern in the testing procedure as well as how the test scores will be interpreted. There are several ways to collect reliability data and to estimate reliability, many of which depend on the exact nature of the measurement, the intended use of the test scores, the assessment design, and the potential sources of measurement error that might contribute to inconsistency in students' scores across different test administrations.

The reliability information presented in this section is organized to be in compliant with critical element 4.1 of the ESSA Peer Review requirements (U.S. Department of Education, 2018) and follows the guidelines of the *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014). Reliability of domain score is presented first, followed by reliability of composite scores.

ACCESS Listening, Reading, Writing, and Speaking scores are used to determine the English language proficiency of students based on students' test scores in each of the four domains. Therefore, the main concern in interpreting the ACCESS test scores is how consistent the scores of the students would be over replications of the same testing procedure in each domain. We use **internal-consistency reliability statistics** to address this question (Section 5.1).

Additionally, for the Writing and Speaking domains, inconsistency in test scores may be introduced by different raters as a potential source of variation. The **interrater agreement** in scoring Writing tasks is reported in Section 5.2, to examine how consistent the scores of the students would be if their responses were scored by different raters. As noted in Part 1, Section 3.2.4, ACCESS Paper Speaking tasks are scored locally; therefore, interrater agreement data are not available for Speaking. Since an item response theory–based method is used in estimating students' latent scores, we also examine the amount of **measurement error** in students' scores using conditional standard error of measurement (Section 5.3). Lastly, in Section 5.4, we evaluate the reliability of classification into WIDA proficiency levels (the most important interpretation of the test scores) in terms of the **accuracy and consistency** of the classification decisions made based on the students' domain test scores. Detailed descriptions of the methods, data sources, and procedures are presented in each subsection.

ACCESS composite scores are used to describe the English language proficiency of students in the respective composites. Therefore, the most important concern in interpreting the ACCESS composite scores is how consistent the composites scores of these students would be over replications of the same testing procedure. We use internal consistency reliability statistics to address this question, and results are provided in Section 5.5. In addition, we examine conditional standard error of measurement of the composites in Section 5.6. Lastly, we evaluate the reliability of classification in terms of the accuracy and the consistency of the decisions made about students' level of English language proficiency based on their composite scores in Section 5.7. Detailed descriptions of the methods, data sources, and procedures are presented in each section.

Internal Consistency Reliability Statistics

One way to evaluate the consistency of students' test scores across test administrations is to examine how the students would have performed on alternate forms of the same test (parallel test form reliability). Given that the abilities being measured are assumed to be constant for each student over two administrations of alternate forms, the more variation found across the two administrations, the more evidence for lower reliability. In this case, the sources of inconsistency across the two administrations taken together are called "measurement error." Measurement error is considered to be random and to occur by chance. For example, there may be some kinds of knowledge and skills assessed by some items or tasks that affect students' scores, but which are not part of what the test intends to measure.

Unless students take two alternate versions of the same test, test reliability cannot be calculated directly. Thus, it is usually estimated from student responses to a single form of the test. Methods used to estimate reliability using test scores from a single test administration are modeled from classical test theory and are referred to as estimates of ***internal consistency***. Internal consistency reliability statistics are a good estimate of alternate-forms reliability statistics, providing an estimate of the consistency of the performance of students across items within a test. The most

common index of internal consistency reliability is referred to as Cronbach's alpha (Cronbach, 1951), which is a lower bound estimate of test reliability. Conceptually, it may be thought of as the correlation obtained between performances on two halves of the test, if every possible way of dividing the test items in two were attempted. Because Cronbach's alpha is a correlation of all possible pairs of test items, Cronbach's alpha may be low if some items are measuring something other than what most of the other items are measuring (and thus leading to inconsistent student performances). In this way, Cronbach's alpha expresses how well the items and tasks on a test appear to measure the same ability. The Cronbach's alpha coefficient of internal consistency ranges from 0 to 1. If scores are assigned to students by a completely random process (i.e., scores are not correlated or share no covariance), then the reliability estimate is very close to 0. If scores assigned to students are perfectly consistent (i.e., scores have high covariances), then the internal consistency coefficient will approach 1.

Reliability statistics such as the Cronbach alpha coefficient of internal consistency are 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 measuring similar abilities there are on the test, the higher the internal consistency reliability statistics. Additionally, because reliability statistics refer to the consistency of scores for a group of students, they are affected by the distribution of abilities measured by the test within the specific group of students tested. If the students in the group are nearly equal in the abilities measured by the test (i.e., are very homogeneous in the ability distribution), small changes in their scores can easily change their relative positions in the group. Consequently, the internal consistency reliability statistics will be low. In this case, the statistic may be telling us more about the group of examinees tested than the test itself. On the other hand, if the students in the group differ widely in the abilities the test measures (i.e., are very heterogeneous in the ability distribution), small changes in their scores will not affect their relative positions in the group as much, and the internal consistency reliability statistics will be higher. Therefore, it is widely recognized that reliability can be as much a function of the test items and tasks as of the sample of students tested. That is, the exact same test can produce widely disparate reliability indices based on the distribution of the group of students. Therefore, when interpreting estimates of internal consistency, it is wise to keep in mind the specific set of test items and the distribution of ability in the group of students used in the estimation.

Interrater Agreement

A potential source of variance in students' scores on the productive domains of ACCESS (Writing and Speaking) lies in the behavior of raters. ACCESS scoring procedures and steps taken to provide rater training and consistency are described elsewhere in this report (see Part 1, Section 3.2.2). The **interrater agreement** rates in scoring Writing tasks are reported in Section 5.2. (As noted above, these data are not available for Paper ACCESS Speaking.) These values examine how consistent the scores of the students would be if their responses were scored by different groups of raters. Detailed descriptions of the methods, data sources, and procedures are presented in the section.

Measurement Error

In addition to evaluating test reliability in terms of estimates of internal consistency, the amount of measurement error in students' test scores is commonly addressed in two different ways in educational and psychological testing. One way is to hypothesize that there is an error-free measure of students' true ability, skills, or proficiency. In classical test theory, it is referred to as the true score. True score is a theoretical value, so it is not a known quantity. Rather, it is viewed as the hypothetical average score over repeated replications of the same testing condition. Under the assumption of classical test theory, the error of measurement over replication of a testing condition provides an estimate of the amount of variability we would expect from students' true scores. In practical testing contexts, it is generally not possible to replicate a testing condition (i.e., have students take the same test form over and over again), so it is not possible to estimate the standard error of the students' scores using a repeated measure design. Instead, the average error of measurement over the population of students who take the test is estimated and used as an indication of the amount of variation we would expect in any individual student's score. This statistic is referred to as the *standard error of measurement* (SEM). It provides an indication of how much students' scores differ from their true scores, on average, on the raw score metric. Because it is a standard deviation of the distribution of errors of measurement, a confidence interval can be constructed to indicate how the errors of measurement are affecting the scores. Test scores with large SEMs pose a challenge to the interpretation of the reliability of any single test score.

Classification 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 into six WIDA proficiency levels, it is important to know how consistently ACCESS scores do indeed *classify* students into the WIDA proficiency levels (American Educational Research Association et al., 2014). The questions we want to answer are different from the questions answered by the reliability coefficient. Instead of looking at the reliability of a specific student score, we want to know how consistently the classifications are being made about students when placed by their test results into a smaller number of proficiency levels. One way to approach this question is to estimate the degree to which classification decisions we are making on the basis of the students' observed test scores agree with the classification decisions we would make based on students' theoretical true score. This estimate is known as decision accuracy. A second way to approach this question is to estimate the degree to which classification decisions we are making on the basis of the students' test scores agree with the classification decisions we would make based on students' scores on a different edition of the test. This estimate is known as decision consistency.

5.1 Reliability of Domain Scores

Cronbach's coefficient alpha is widely used as an estimate of reliability, particularly of the internal consistency of test items. Conceptually, it may be thought of as the correlation obtained between performances on two halves of the test, if every possible way of dividing the test tasks 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. In this way, Cronbach's alpha expresses how well the items and tasks on a test appear to measure the same ability.

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 Cronbach's alpha for tiered forms that have differential weighting across tasks. This modification is an attempt to take into account the different weighting of tasks when deriving students' ability measures for these tiered forms. For Writing tasks with a weight greater than one, students' responses to the tasks are 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 the Kindergarten Writing domain, a stratified Cronbach's alpha is reported instead of Cronbach's alpha because the dichotomous and polytomous items are heterogeneous and each have different true score variance. It is more appropriate to report stratified alpha (Feldt & Brennan, 1989), as this statistic was derived to measure the consistency in students' scores when the total score consists of heterogeneous parts. Stratified alpha is a weighted average of coefficient alphas for item sets with different maximum score points or "strata." Stratified alpha is a reliability estimate computed by dividing the test into parts (strata), computing Cronbach's alpha separately for each part, and using the results to estimate a reliability coefficient for the total score. (See Section 5.5 for more details regarding stratified Cronbach's alpha.) In computing the stratified Cronbach's alpha for Kindergarten Writing, each part that makes up the total score is treated as a strata. In other words, two strata (dichotomous and polytomous) are entered into the computation. The stratified Cronbach's alpha is interpreted like other traditional

internal consistency statistics such as Cronbach's coefficient alpha. Like Cronbach's alpha, stratified Cronbach's alpha is an estimate of the proportion of the total variance of the observed composite score that can be explained by the variance of the true composite score.

Tables in this section also present the standard error of measurement (SEM), which provides a value for the errors of measurement in students' scores using classical test theory. It is a function of two statistics: the reliability estimate of the test and the (observed) standard deviation (SD) of the test scores in the student population, and it is on the raw score metric. It is calculated as

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

Since the SEM is an estimate of the standard deviation of the distribution of measurement errors, SEM can be used to create a band around a student's observed score. Under the assumption that the error of measurement follows a normal distribution, the student's true score would lie with a certain degree of probability within this band. Statistically speaking, then, there is an expectation that a student'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. Since SEMs are expressed on the raw score metric, it is wise to keep the range of the raw score distribution in mind when interpreting the SEM. Raw score statistics by domains are reported in Section 2.

In the tables below, we provide the number of tasks, Cronbach's alpha, and SEM for all students and for subgroups as required by the ESSA Peer Review so that the reliability estimates of the subgroups can be compared with those computed based on all students. For these domains, the first table provides Cronbach's alpha and the SEM for all students. Each row in the table represents a specific grade cluster and test form. For each form, the numbers of students, numbers of tasks, Cronbach's alpha, and SEM are provided. 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).

Kindergarten: For the Kindergarten Listening test, the reliability for all students was 0.94, and reliability values across subgroups ranged from 0.94 to 0.95. For the Kindergarten Reading test, the reliability for all students was 0.95, and reliability values across subgroups ranged from 0.95 to 0.96. For the Kindergarten Writing test, the reliability for all students was 0.92, and reliability values across subgroups ranged from 0.92 to 0.93. For the Kindergarten Speaking test, the reliability for all students was 0.90, and reliability values across subgroups ranged from 0.89 to 0.91.

Listening Tier A: The Listening Tier A Cronbach's alpha computed for all students ranged from 0.63 to 0.77. The Listening Tier A Cronbach's alpha ranged from 0.65 to 0.78 for male students; 0.61 to 0.77 for female students; 0.62 to 0.77 for Hispanic students; 0.66 to 0.77 for non-Hispanic students; and 0.63 to 0.76 for students with an IEP.

Listening Tier B/C: The Listening Tier B/C Cronbach's alpha computed for all students ranged from 0.63 to 0.68. The Listening Tier B/C Cronbach's alpha ranged from 0.64 to 0.69 for male students; 0.61 to 0.67 for female students; 0.62 to 0.68 for Hispanic students; 0.63 to 0.70 for non-Hispanic students; and 0.61 to 0.72 for students with an IEP.

Reading Tier A: The Reading Tier A Cronbach's alpha computed for all students ranged from 0.78 to 0.83. The Reading Tier A Cronbach's alpha ranged from 0.78 to 0.83 for male students; 0.78 to 0.84 for female students; 0.77 to 0.82 for Hispanic students; 0.80 to 0.85 for non-Hispanic students; and 0.73 to 0.76 for students with an IEP.

Reading Tier B/C: The Reading Tier B/C Cronbach's alpha computed for all students ranged from 0.75 to 0.84. The Reading Tier B/C Cronbach's alpha ranged from 0.76 to 0.84 for male students; 0.73 to 0.84 for female students; 0.74 to 0.83 for Hispanic students; 0.77 to 0.85 for non-Hispanic students; and 0.68 to 0.79 for students with an IEP.

Writing Tier A: The Writing Tier A Cronbach's alpha computed for all students ranged from 0.88 to 0.92. The Writing Tier A Cronbach's alpha ranged from 0.88 to 0.92 for male students; 0.87 to 0.92 for female students; 0.88 to 0.92 for Hispanic students; 0.86 to 0.91 for non-Hispanic students; and 0.87 to 0.92 for students with an IEP.

Writing Tier B/C: The Writing Tier B/C Cronbach's alpha computed for all students ranged from 0.91 to 0.95. The Writing Tier B/C Cronbach's alpha ranged from 0.91 to 0.95 for male students; 0.90 to 0.94 for female students; 0.91 to 0.95 for Hispanic students; 0.91 to 0.95 for non-Hispanic students; and 0.92 to 0.96 for students with an IEP.

Speaking Tier A: The Speaking Tier A Cronbach's alpha computed for all students ranged from 0.88 to 0.91. Cronbach's alpha ranged from 0.88 to 0.91 for male students; 0.88 to 0.91 for female students; 0.88 to 0.92 for Hispanic students; 0.85 to 0.88 for non-Hispanic students; and 0.85 to 0.86 for students with an IEP.

Speaking Tier B/C: The Speaking Tier B/C Cronbach's alpha computed for all students ranged from 0.90 to 0.94. Cronbach's alpha ranged from 0.90 to 0.93 for male students; 0.90 to 0.94 for female students; 0.90 to 0.94 for Hispanic students; 0.90 to 0.93 for non-Hispanic students; and 0.90 to 0.93 for students with an IEP.

5.1.1 Listening

Table 5.1.1.1

Reliability: List S501 Paper

Cluster	Tier	No. of Students	No. of Items	Cronbach's Alpha	SEM
K	-	226,001	30	0.943	1.831
1	A	24,949	18	0.773	1.572
	B/C	47,022	21	0.683	1.719
2	A	24,949	18	0.773	1.572
	B/C	47,022	21	0.683	1.719
3	A	16,595	18	0.743	1.857
	B/C	68,570	21	0.636	1.878
4-5	A	16,595	18	0.743	1.857
	B/C	68,570	21	0.636	1.878
6-8	A	12,504	18	0.716	1.846
	B/C	39,779	21	0.626	1.786
9-12	A	12,490	18	0.632	1.778
	B/C	36,883	21	0.673	1.954

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.1.2

Reliability: List S501 Paper by Gender

			Female			Male		
Cluster	Tier	No. of Items	No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	30	102,782	0.942	1.798	116,888	0.943	1.853
1	A	18	10,833	0.766	1.558	13,076	0.779	1.579
	B/C	21	22,117	0.671	1.700	23,665	0.692	1.732
2	A	18	10,833	0.766	1.558	13,076	0.779	1.579
	B/C	21	22,117	0.671	1.700	23,665	0.692	1.732
3	A	18	7,311	0.739	1.849	8,732	0.747	1.859
	B/C	21	30,464	0.621	1.882	35,889	0.647	1.871
4-5	A	18	7,311	0.739	1.849	8,732	0.747	1.859
	B/C	21	30,464	0.621	1.882	35,889	0.647	1.871
6-8	A	18	5,384	0.721	1.826	6,624	0.714	1.858
	B/C	21	17,385	0.609	1.767	20,652	0.637	1.798
9-12	A	18	5,537	0.614	1.767	6,237	0.649	1.783
	B/C	21	16,289	0.668	1.944	18,910	0.679	1.958

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.1.3

Reliability: List S501 Paper by Ethnicity

Cluster	Tier	No. of Items	Hispanic			Other		
			No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	30	147,166	0.944	1.858	70,832	0.936	1.768
1	A	18	19,443	0.771	1.582	4,863	0.769	1.541
	B/C	21	35,587	0.678	1.719	10,124	0.700	1.729
2	A	18	19,443	0.771	1.582	4,863	0.769	1.541
	B/C	21	35,587	0.678	1.719	10,124	0.700	1.729
3	A	18	13,081	0.733	1.873	2,915	0.758	1.804
	B/C	21	52,552	0.631	1.882	13,950	0.649	1.879
4-5	A	18	13,081	0.733	1.873	2,915	0.758	1.804
	B/C	21	52,552	0.631	1.882	13,950	0.649	1.879
6-8	A	18	10,162	0.700	1.861	1,995	0.727	1.791
	B/C	21	30,446	0.624	1.790	8,084	0.631	1.784
9-12	A	18	9,888	0.622	1.778	2,317	0.662	1.767
	B/C	21	27,576	0.671	1.953	8,503	0.677	1.962

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.1.4

Reliability: List S501 Paper by IEP Status

Cluster	Tier	No. of Students	No. of Items	Cronbach's Alpha	SEM
K	-	19,371	30	0.951	1.902
1	A	3,195	18	0.757	1.616
	B/C	5,480	21	0.718	1.824
2	A	3,195	18	0.757	1.616
	B/C	5,480	21	0.718	1.824
3	A	1,609	18	0.692	1.852
	B/C	12,913	21	0.632	1.951
4-5	A	1,609	18	0.692	1.852
	B/C	12,913	21	0.632	1.951
6-8	A	686	18	0.708	1.841
	B/C	5,645	21	0.620	1.925
9-12	A	530	18	0.632	1.795
	B/C	2,793	21	0.613	2.025

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

5.1.2 Reading

Table 5.1.2.1

Reliability: Read S501 Paper

Cluster	Tier	No. of Students	No. of Items	Cronbach's Alpha	SEM
K	-	225,994	30	0.950	1.756
1	A	23,605	24	0.802	2.165
	B/C	41,032	27	0.837	2.284
2	A	23,605	24	0.802	2.165
	B/C	41,032	27	0.837	2.284
3	A	15,781	24	0.832	2.118
	B/C	62,328	27	0.747	2.378
4-5	A	15,781	24	0.832	2.118
	B/C	62,328	27	0.747	2.378
6-8	A	12,109	24	0.780	2.163
	B/C	34,994	27	0.773	2.345
9-12	A	12,422	24	0.799	2.092
	B/C	33,400	27	0.815	2.349

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.2.2

Reliability: Read S501 Paper by Gender

			Female			Male		
Cluster	Tier	No. of Items	No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	30	102,780	0.947	1.759	116,883	0.953	1.751
1	A	24	10,242	0.800	2.160	12,408	0.807	2.165
	B/C	27	19,282	0.835	2.277	20,703	0.839	2.287
2	A	24	10,242	0.800	2.160	12,408	0.807	2.165
	B/C	27	19,282	0.835	2.277	20,703	0.839	2.287
3	A	24	7,033	0.839	2.091	8,223	0.826	2.136
	B/C	27	27,855	0.731	2.379	32,474	0.761	2.374
4-5	A	24	7,033	0.839	2.091	8,223	0.826	2.136
	B/C	27	27,855	0.731	2.379	32,474	0.761	2.374
6-8	A	24	5,284	0.780	2.146	6,358	0.779	2.170
	B/C	27	15,344	0.762	2.332	18,021	0.781	2.353
9-12	A	24	5,547	0.796	2.072	6,142	0.802	2.101
	B/C	27	14,871	0.807	2.345	17,039	0.821	2.347

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.2.3

Reliability: Read S501 Paper by Ethnicity

Cluster	Tier	No. of Items	Hispanic			Other		
			No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	30	147,164	0.946	1.759	70,828	0.954	1.719
1	A	24	18,331	0.788	2.182	4,655	0.831	2.111
	B/C	27	30,980	0.831	2.296	8,902	0.845	2.258
2	A	24	18,331	0.788	2.182	4,655	0.831	2.111
	B/C	27	30,980	0.831	2.296	8,902	0.845	2.258
3	A	24	12,444	0.822	2.129	2,778	0.850	2.088
	B/C	27	47,643	0.735	2.381	12,739	0.765	2.375
4-5	A	24	12,444	0.822	2.129	2,778	0.850	2.088
	B/C	27	47,643	0.735	2.381	12,739	0.765	2.375
6-8	A	24	9,844	0.772	2.171	1,929	0.796	2.128
	B/C	27	26,839	0.767	2.351	7,017	0.778	2.334
9-12	A	24	9,864	0.794	2.100	2,277	0.808	2.050
	B/C	27	25,121	0.812	2.350	7,536	0.819	2.350

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

Table 5.1.2.4

Reliability: Read S501 Paper by IEP Status

Cluster	Tier	No. of Students	No. of Items	Cronbach's Alpha	SEM
K	-	19,367	30	0.960	1.744
1	A	3,027	24	0.761	2.209
	B/C	4,847	27	0.794	2.371
2	A	3,027	24	0.761	2.209
	B/C	4,847	27	0.794	2.371
3	A	1,537	24	0.762	2.197
	B/C	11,860	27	0.677	2.380
4-5	A	1,537	24	0.762	2.197
	B/C	11,860	27	0.677	2.380
6-8	A	639	24	0.730	2.191
	B/C	4,966	27	0.700	2.383
9-12	A	507	24	0.733	2.159
	B/C	2,535	27	0.750	2.408

Note: The test form is shared between 1A and 2A, 1B/C and 2B/C.
The test form is shared between 3A and 4-5A, 3B/C and 4-5B/C.

5.1.3 Writing

Table 5.1.3.1

Reliability: Writ S501 Paper

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha*	SEM
K	-	225,987	6	0.924	1.176
1	A	19,965	4	0.880	2.031
	B/C	20,779	3	0.947	1.839
2	A	16,588	3	0.922	1.322
	B/C	60,363	3	0.940	1.653
3	A	16,588	3	0.922	1.322
	B/C	60,363	3	0.940	1.653
4-5	A	10,478	3	0.897	1.369
	B/C	41,920	3	0.912	1.722
6-8	A	13,517	3	0.885	1.434
	B/C	41,227	3	0.917	1.691
9-12	A	13,589	3	0.875	1.628
	B/C	38,565	3	0.915	2.181

*Note that for Kindergarten, which includes both dichotomous and polytomous tasks in the Writing test, a stratified Cronbach's alpha is computed.

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.3.2

Reliability: Writ S501 Paper by Gender

			Female			Male		
Cluster	Tier	No. of Tasks	No. of Students	Cronbach's Alpha*	SEM	No. of Students	Cronbach's Alpha*	SEM
K	-	6	102,780	0.922	1.177	116,876	0.924	1.175
1	A	4	8,514	0.877	1.984	10,525	0.882	2.051
	B/C	3	9,679	0.943	1.809	10,487	0.949	1.862
2	A	3	7,089	0.921	1.321	8,949	0.921	1.325
	B/C	3	27,387	0.932	1.629	31,324	0.942	1.671
3	A	3	7,089	0.921	1.321	8,949	0.921	1.325
	B/C	3	27,387	0.932	1.629	31,324	0.942	1.671
4-5	A	3	4,675	0.893	1.374	5,444	0.897	1.363
	B/C	3	18,467	0.898	1.697	21,966	0.914	1.735
6-8	A	3	5,814	0.879	1.438	7,171	0.889	1.429
	B/C	3	17,924	0.900	1.633	21,464	0.919	1.737
9-12	A	3	5,973	0.867	1.659	6,805	0.880	1.601
	B/C	3	16,856	0.910	2.067	19,897	0.915	2.263

*Note that for Kindergarten, which includes both dichotomous and polytomous tasks in the Writing test, a stratified Cronbach's alpha is computed.

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.3.3

Reliability: Writ S501 Paper by Ethnicity

Cluster	Tier	No. of Tasks	Hispanic			Other		
			No. of Students	Cronbach's Alpha*	SEM	No. of Students	Cronbach's Alpha*	SEM
K	-	6	147,160	0.916	1.169	70,827	0.929	1.180
1	A	4	15,631	0.880	2.031	3,935	0.880	2.015
	B/C	3	15,371	0.947	1.830	4,721	0.950	1.869
2	A	3	12,850	0.922	1.318	3,153	0.914	1.338
	B/C	3	46,163	0.940	1.652	12,654	0.941	1.659
3	A	3	12,850	0.922	1.318	3,153	0.914	1.338
	B/C	3	46,163	0.940	1.652	12,654	0.941	1.659
4-5	A	3	8,281	0.895	1.376	1,815	0.890	1.343
	B/C	3	31,991	0.911	1.713	8,562	0.915	1.746
6-8	A	3	10,959	0.883	1.421	2,186	0.881	1.458
	B/C	3	31,470	0.919	1.673	8,441	0.912	1.764
9-12	A	3	10,709	0.879	1.585	2,561	0.859	1.759
	B/C	3	28,765	0.915	2.153	8,963	0.916	2.267

*Note that for Kindergarten, which includes both dichotomous and polytomous tasks in the Writing test, a stratified Cronbach's alpha is computed.

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.3.4

Reliability: Writ S501 Paper by IEP Status

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha*	SEM
K	-	19,363	6	0.925	1.123
1	A	2,511	4	0.884	2.008
	B/C	2,138	3	0.959	1.899
2	A	2,171	3	0.915	1.306
	B/C	9,009	3	0.954	1.700
3	A	2,171	3	0.915	1.306
	B/C	9,009	3	0.954	1.700
4-5	A	818	3	0.886	1.336
	B/C	8,521	3	0.923	1.809
6-8	A	719	3	0.874	1.427
	B/C	5,853	3	0.926	1.820
9-12	A	563	3	0.875	1.638
	B/C	2,923	3	0.922	2.190

*Note that for Kindergarten, which includes both dichotomous and polytomous tasks in the Writing test, a stratified Cronbach's alpha is computed.

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

5.1.4 Speaking

Table 5.1.4.1

Reliability: Spek S501 Paper

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
K	-	226,000	10	0.904	1.031
1	A	19,813	6	0.877	1.363
	B/C	20,613	6	0.905	1.433
2	A	16,424	6	0.910	1.369
	B/C	59,952	6	0.906	1.343
3	A	16,424	6	0.910	1.369
	B/C	59,952	6	0.906	1.343
4-5	A	10,396	6	0.905	1.404
	B/C	41,592	6	0.903	1.346
6-8	A	13,376	6	0.903	1.401
	B/C	40,881	6	0.917	1.349
9-12	A	13,367	6	0.884	1.457
	B/C	38,223	6	0.935	1.327

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.4.2

Reliability: Spek S501 Paper by Gender

			Female			Male		
Cluster	Tier	No. of Tasks	No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	10	102,783	0.907	1.011	116,886	0.901	1.046
1	A	6	8,446	0.876	1.383	10,443	0.877	1.347
	B/C	6	9,594	0.904	1.434	10,410	0.906	1.432
2	A	6	7,020	0.907	1.387	8,859	0.911	1.358
	B/C	6	27,200	0.904	1.347	31,118	0.906	1.336
3	A	6	7,020	0.907	1.387	8,859	0.911	1.358
	B/C	6	27,200	0.904	1.347	31,118	0.906	1.336
4-5	A	6	4,648	0.902	1.421	5,391	0.907	1.386
	B/C	6	18,317	0.902	1.342	21,795	0.903	1.344
6-8	A	6	5,755	0.897	1.420	7,097	0.909	1.382
	B/C	6	17,760	0.915	1.370	21,306	0.920	1.325
9-12	A	6	5,905	0.879	1.482	6,710	0.886	1.439
	B/C	6	16,707	0.935	1.335	19,724	0.934	1.315

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.4.3

Reliability: Spek S501 Paper by Ethnicity

Cluster	Tier	No. of Tasks	Hispanic			Other		
			No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
K	-	10	147,165	0.907	1.038	70,833	0.890	1.019
1	A	6	15,522	0.881	1.352	3,897	0.850	1.392
	B/C	6	15,246	0.907	1.425	4,682	0.900	1.451
2	A	6	12,731	0.916	1.349	3,118	0.880	1.427
	B/C	6	45,854	0.906	1.339	12,566	0.904	1.354
3	A	6	12,731	0.916	1.349	3,118	0.880	1.427
	B/C	6	45,854	0.906	1.339	12,566	0.904	1.354
4-5	A	6	8,217	0.908	1.377	1,799	0.874	1.461
	B/C	6	31,746	0.904	1.339	8,490	0.902	1.360
6-8	A	6	10,852	0.904	1.381	2,157	0.876	1.458
	B/C	6	31,244	0.919	1.349	8,328	0.913	1.357
9-12	A	6	10,531	0.886	1.446	2,525	0.866	1.475
	B/C	6	28,556	0.936	1.322	8,836	0.929	1.347

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

Table 5.1.4.4

Reliability: Spek S501 Paper by IEP Status

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
K	-	19,370	10	0.897	1.059
1	A	2,495	6	0.856	1.286
	B/C	2,124	6	0.915	1.368
2	A	2,148	6	0.864	1.322
	B/C	8,957	6	0.910	1.339
3	A	2,148	6	0.864	1.322
	B/C	8,957	6	0.910	1.339
4-5	A	811	6	0.854	1.391
	B/C	8,453	6	0.899	1.359
6-8	A	710	6	0.859	1.499
	B/C	5,816	6	0.912	1.370
9-12	A	557	6	0.864	1.451
	B/C	2,884	6	0.930	1.371

Note: The test form is shared between 2A and 3A, 2B/C and 3B/C.

5.2 Interrater Agreement

For the Writing tests (except Kindergarten, which is scored by the test administrator), 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 show 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).

As the Speaking test is scored locally, it is not possible to provide interrater agreement data for Speaking. Section 3.2.3 in Part 1 of this report describes training procedures that local raters must complete before being certified to administer and score the Speaking test.

WIDA stipulates a minimum interrater agreement rate of 70%. Tasks with interrater agreement rates between 70% and 74% are regarded as borderline.

For Writing, the lowest value for interrater agreement is 95%.

5.2.1 Listening

Interrater Agreement is not relevant for the domain of Listening, as all items are multiple choice items.

5.2.2 Reading

Interrater Agreement is not relevant for the domain of Listening, as all items are multiple choice items.

5.2.3 Writing

5.2.3.0 Kindergarten

Table 5.2.3.0.1

Interrater Agreement: Writ K S501 Paper

Interrater Agreement	n/a

5.2.3.1 Grade 1

Table 5.2.3.1.1

Interrater Agreement: Writ 1 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	7,868	100	0	0
	2	11,246	100	0	0
	3	12,476	99	1	0
	4	11,622	99	1	0

Table 5.2.3.1.2

Interrater Agreement: Writ 1 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	9,452	99	1	0
	2	11,868	98	2	0
	3	11,092	99	1	0

5.2.3.2 Grade 2

Table 5.2.3.2.1

Interrater Agreement: Writ 2 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	5,880	98	2	0
	2	5,984	99	1	0
	3	6,856	99	1	0

Note: the test form is shared between 2A and 3A.

Table 5.2.3.2.2

Interrater Agreement: Writ 2 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	13,650	96	4	0
	2	13,806	96	4	0
	3	13,704	98	2	0

Note: the test form is shared between 2B/C and 3B/C.

5.2.3.3 Grade 3

Table 5.2.3.3.1

Interrater Agreement: Writ 3 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	4,212	97	3	0
	2	4,378	99	1	0
	3	4,912	99	1	0

Note: the test form is shared between 2A and 3A.

Table 5.2.3.3.2

Interrater Agreement: Writ 3 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	12,730	96	4	0
	2	12,834	97	3	0
	3	12,742	98	2	0

Note: the test form is shared between 2B/C and 3B/C.

5.2.3.4 Grades 4–5

Table 5.2.3.4.1

Interrater Agreement: Writ 4-5 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	6,610	98	2	0
	2	5,798	99	1	0
	3	6,782	99	1	0

Table 5.2.3.4.2

Interrater Agreement: Writ 4-5 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	17,754	98	2	0
	2	17,918	98	2	0
	3	18,098	98	2	0

5.2.3.5 Grades 6–8

Table 5.2.3.5.1

Interrater Agreement: Writ 6-8 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	7,004	97	3	0
	2	7,954	97	3	0
	3	6,474	97	3	0

Table 5.2.3.5.2

Interrater Agreement: Writ 6-8 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	17,448	98	2	0
	2	17,440	97	3	0
	3	17,602	98	2	0

5.2.3.6 Grades 9–12

Table 5.2.3.6.1

Interrater Agreement: Writ 9-12 A S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	7,956	97	3	0
	2	6,366	97	3	0
	3	10,078	98	2	0

Table 5.2.3.6.2

Interrater Agreement: Writ 9-12 B/C S501 Paper

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	17,652	97	3	0
	2	18,032	96	3	0
	3	17,842	95	5	0

5.3 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.

For each domain, the values are presented by tier. From these tables, it is possible to identify how well the different 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 a relatively low CSEM at the 4/5 proficiency level cut. Information from these tables provides 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 below, 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.

As a general rule, lower CSEM values around decision points are desirable. For the ACCESS population, CSEM values for the highest cut points are typically high. Students are exited from the ACCESS population upon gaining English language proficiency, and therefore these students are removed from the ACCESS population, resulting in smaller numbers of students at the highest cut points.

5.3.1 Listening

5.3.1.0 Kindergarten

Table 5.3.1.0
Conditional Standard Error of Measurement at Cut Scores: List K S501 Paper

Proficiency Level Cut Point	Cut Score	SEM
1/2	229	17.28
2/3	251	18.41
3/4	278	20.66
4/5	286	21.42
5/6	308	24.80

5.3.1.1 Grade 1

Table 5.3.1.1
Conditional Standard Error of Measurement at Cut Scores: List 1 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	1	236	19.16	19.54
2/3	1	259	19.54	18.41
3/4	1	291	22.54	18.41
4/5	1	303	24.42	18.79
5/6	1	327	29.31	20.66

5.3.1.2 Grade 2

Table 5.3.1.2
Conditional Standard Error of Measurement at Cut Scores: List 2 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	2	245	19.16	19.01
2/3	2	283	21.42	18.03
3/4	2	314	26.30	19.54
4/5	2	330	30.43	21.04
5/6	2	354	38.32	24.80

5.3.1.3 Grade 3

Table 5.3.1.3
Conditional Standard Error of Measurement at Cut Scores: List 3 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	3	262	21.42	24.80
2/3	3	300	18.79	20.66
3/4	3	331	19.54	19.16
4/5	3	349	21.04	18.79
5/6	3	374	25.17	19.16

5.3.1.4 Grade 4–5

Table 5.3.1.4
Conditional Standard Error of Measurement at Cut Scores: List 4-5 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	275	20.06	23.29
	5	285	19.54	22.17
2/3	4	313	18.79	19.91
	5	323	19.16	19.16
3/4	4	343	20.66	18.79
	5	354	21.79	18.79
4/5	4	363	22.92	18.79
	5	375	25.55	19.35
5/6	4	388	28.55	20.29
	5	401	32.31	21.42

5.3.1.5 Grade 6–8

Table 5.3.1.5
Conditional Standard Error of Measurement at Cut Scores: List 6-8 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	294	20.29	21.42
	7	302	19.91	20.29
	8	308	19.91	19.91
2/3	6	332	19.91	18.03
	7	340	20.40	18.03
	8	347	21.04	17.77
3/4	6	363	22.54	18.03
	7	370	23.29	18.03
	8	377	24.42	18.41
4/5	6	385	25.92	19.16
	7	394	28.18	19.91
	8	402	30.06	20.66
5/6	6	411	33.06	21.79
	7	420	36.07	23.29
	8	427	39.07	24.42

5.3.1.6 Grade 9–12

Table 5.3.1.6
Conditional Standard Error of Measurement at Cut Scores: List 9-12 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	9	314	20.66	22.17
	10	325	20.66	21.04
	11	335	21.04	19.91
	12	342	21.42	19.54
2/3	9	353	22.17	18.79
	10	358	22.54	18.79
	11	364	23.26	18.41
	12	368	23.67	18.41
3/4	9	383	25.92	18.41
	10	389	27.43	18.41
	11	394	28.55	18.41
	12	398	29.31	18.79
4/5	9	409	32.31	19.16
	10	415	34.19	19.54
	11	420	36.07	19.91
	12	426	37.95	20.66
5/6	9	434	41.33	21.42
	10	441	44.71	22.54
	11	447	47.72	23.29
	12	452	50.72	24.42

5.3.2 Reading

5.3.2.0 Kindergarten

Table 5.3.2.0
Conditional Standard Error of Measurement at Cut Scores: Read K S501 Paper

Proficiency Level Cut Point	Cut Score	SEM
1/2	241	15.34
2/3	259	18.46
3/4	279	23.92
4/5	289	27.82
5/6	310	39.26

5.3.2.1 Grade 1

Table 5.3.2.1
Conditional Standard Error of Measurement at Cut Scores: Read 1 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	1	264	11.44	13.78
2/3	1	286	11.70	11.44
3/4	1	304	13.00	10.40
4/5	1	315	14.56	10.40
5/6	1	334	18.46	11.44

5.3.2.2 Grade 2

Table 5.3.2.2
Conditional Standard Error of Measurement at Cut Scores: Read 2 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	2	283	11.70	11.52
2/3	2	307	13.52	10.40
3/4	2	326	16.51	10.92
4/5	2	337	19.24	11.78
5/6	2	355	25.48	14.04

5.3.2.3 Grade 3

Table 5.3.2.3
Conditional Standard Error of Measurement at Cut Scores: Read 3 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	3	297	11.70	17.94
2/3	3	323	11.96	13.00
3/4	3	342	13.52	11.13
4/5	3	352	14.82	10.69
5/6	3	370	18.46	10.66

5.3.2.4 Grade 4–5

Table 5.3.2.4
Conditional Standard Error of Measurement at Cut Scores: Read 4-5 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	307	11.44	15.86
	5	316	11.70	14.04
2/3	4	335	12.74	11.70
	5	345	13.83	10.92
3/4	4	354	15.34	10.66
	5	364	17.16	10.48
4/5	4	364	17.16	10.48
	5	373	19.50	10.61
5/6	4	382	22.10	11.00
	5	391	25.48	11.70

5.3.2.5 Grade 6–8

Table 5.3.2.5
Conditional Standard Error of Measurement at Cut Scores: Read 6-8 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	323	11.60	13.52
	7	329	11.44	12.74
	8	335	11.70	12.22
2/3	6	353	12.48	10.92
	7	360	13.00	10.61
	8	366	13.52	10.66
3/4	6	373	14.56	10.61
	7	380	15.60	10.92
	8	386	16.67	11.18
4/5	6	382	16.12	10.89
	7	389	17.42	11.18
	8	395	18.98	11.70
5/6	6	399	20.02	12.22
	7	406	22.36	13.00
	8	412	24.70	13.78

5.3.2.6 Grade 9–12

Table 5.3.2.6
Conditional Standard Error of Measurement at Cut Scores: Read 9-12 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	9	340	11.83	14.04
	10	344	11.70	13.21
	11	348	11.70	12.74
	12	352	11.70	12.22
2/3	9	372	12.48	10.66
	10	377	12.74	10.40
	11	382	13.26	10.40
	12	386	13.78	10.40
3/4	9	392	14.82	10.45
	10	397	15.60	10.66
	11	402	16.38	10.92
	12	407	17.68	11.18
4/5	9	401	16.38	10.66
	10	406	17.42	11.18
	11	410	18.46	11.31
	12	414	19.50	11.70
5/6	9	418	20.54	12.22
	10	423	22.36	12.74
	11	427	23.66	13.21
	12	432	25.74	14.04

5.3.3 Writing

5.3.3.0 Kindergarten

Table 5.3.3.0
Conditional Standard Error of Measurement at Cut Scores: Writ K S501 Paper

Proficiency Level Cut Point	Cut Score	SEM
1/2	234	18.97
2/3	271	21.15
3/4	311	31.41
4/5	367	43.22
5/6	389	52.55

5.3.3.1 Grade 1

Table 5.3.3.1
Conditional Standard Error of Measurement at Cut Scores: Writ 1 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	1	238	10.31	8.59
2/3	1	275	13.96	10.74
3/4	1	337	13.69	12.51
4/5	1	382	13.16	10.82
5/6	1	405	16.38	11.38

5.3.3.2 Grade 2

Table 5.3.3.2
Conditional Standard Error of Measurement at Cut Scores: Writ 2 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	2	242	11.81	8.46
2/3	2	279	16.27	11.28
3/4	2	341	17.18	12.35
4/5	2	388	15.31	10.87
5/6	2	411	18.26	12.35

5.3.3.3 Grade 3

Table 5.3.3.3
Conditional Standard Error of Measurement at Cut Scores: Writ 3 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	3	247	12.08	8.59
2/3	3	283	16.65	11.44
3/4	3	346	17.18	12.08
4/5	3	394	15.65	11.01
5/6	3	418	20.41	13.69

5.3.3.4 Grade 4–5

Table 5.3.3.4
Conditional Standard Error of Measurement at Cut Scores: Writ 4-5 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	266	11.81	9.93
	5	267	11.55	9.67
2/3	4	288	13.43	8.86
	5	293	14.04	8.86
3/4	4	351	17.99	12.35
	5	356	17.72	12.62
4/5	4	401	15.57	11.55
	5	407	15.31	11.33
5/6	4	425	15.57	11.01
	5	433	16.65	11.01

5.3.3.5 Grade 6–8

Table 5.3.3.5
Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	268	12.35	8.32
	7	273	12.62	8.32
	8	281	13.69	8.59
2/3	6	298	15.84	10.20
	7	305	16.65	11.01
	8	311	16.92	11.55
3/4	6	361	17.45	12.62
	7	367	17.18	12.35
	8	372	16.92	12.35
4/5	6	413	15.57	10.74
	7	419	16.11	10.74
	8	424	16.92	11.01
5/6	6	441	20.68	12.35
	7	450	23.90	14.23
	8	459	27.93	16.38

5.3.3.6 Grade 9–12

Table 5.3.3.6
Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	9	289	12.35	8.22
	10	298	12.08	8.59
	11	308	12.89	9.67
	12	318	14.23	10.77
2/3	9	319	14.23	11.01
	10	326	15.31	11.55
	11	335	16.38	12.08
	12	344	17.02	12.35
3/4	9	378	17.72	12.62
	10	385	17.72	12.35
	11	391	17.45	12.08
	12	398	17.18	11.81
4/5	9	430	15.47	10.74
	10	436	15.31	10.74
	11	441	15.57	11.01
	12	447	15.84	11.55
5/6	9	469	19.33	15.04
	10	479	22.29	17.72
	11	490	27.12	22.02
	12	501	33.03	27.12

5.3.4 Speaking

5.3.4.0 Kindergarten

Table 5.3.4.0
Conditional Standard Error of Measurement at Cut Scores: Spek K S501 Paper

Proficiency Level Cut Point	Cut Score	SEM
1/2	191	28.06
2/3	250	20.92
3/4	301	16.33
4/5	349	22.45
5/6	392	53.57

5.3.4.1 Grade 1

Table 5.3.4.1
Conditional Standard Error of Measurement at Cut Scores: Spek 1 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	1	205	19.30	14.74
2/3	1	261	27.20	19.30
3/4	1	311	25.15	18.13
4/5	1	361	26.32	18.13
5/6	1	403	39.48	26.91

5.3.4.2 Grade 2

Table 5.3.4.2
Conditional Standard Error of Measurement at Cut Scores: Spek 2 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	2	220	23.98	16.96
2/3	2	273	26.91	18.72
3/4	2	322	24.57	17.84
4/5	2	374	35.97	21.94
5/6	2	415	64.05	33.34

5.3.4.3 Grade 3

Table 5.3.4.3

Conditional Standard Error of Measurement at Cut Scores: Spek 3 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	3	234	26.32	17.84
2/3	3	283	25.74	18.43
3/4	3	332	25.15	18.13
4/5	3	386	41.82	24.28
5/6	3	425	74.87	38.02

5.3.4.4 Grade 4–5

Table 5.3.4.4

Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	246	22.23	16.09
	5	258	23.98	16.38
2/3	4	293	26.29	18.25
	5	302	26.32	18.72
3/4	4	342	25.74	18.78
	5	350	26.03	18.72
4/5	4	397	35.10	20.18
	5	407	39.19	21.35
5/6	4	435	57.33	27.79
	5	443	64.64	30.71

5.3.4.5 Grade 6–8

Table 5.3.4.5
Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	268	21.18	15.50
	7	277	22.81	15.79
	8	284	24.28	16.38
2/3	6	310	27.79	18.43
	7	317	27.79	19.01
	8	323	27.79	19.30
3/4	6	360	24.57	18.43
	7	369	23.98	17.84
	8	377	23.98	17.55
4/5	6	417	29.54	18.72
	7	425	32.17	19.89
	8	433	35.10	21.06
5/6	6	451	44.46	24.86
	7	457	48.55	26.62
	8	463	53.23	28.96

5.3.4.6 Grade 9–12

Table 5.3.4.6
Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 S501 Paper

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	9	290	25.45	17.26
	10	295	26.03	17.55
	11	299	26.62	18.13
	12	302	27.20	18.43
2/3	9	328	27.49	19.60
	10	333	27.20	19.60
	11	337	26.76	19.60
	12	340	26.62	19.30
3/4	9	385	24.57	17.26
	10	393	24.86	17.26
	11	400	25.74	17.67
	12	406	26.62	18.13
4/5	9	440	36.85	23.40
	10	446	40.07	24.86
	11	451	42.70	26.32
	12	455	45.04	27.79
5/6	9	468	54.40	32.76
	10	471	57.03	34.22
	11	474	59.67	35.68
	12	476	61.42	36.85

5.4 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 domain scores can classify students into WIDA proficiency categories determined by the 2016 ACCESS standard setting process (Cook & MacGregor, 2017). The accuracy and consistency of these classifications can be useful for test users to judge the utility of this information and to policy makers to make decisions about test design and score reporting (American Educational Research Association et al., 2014). 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 observed test scores would agree with those made on the basis of their true scores (Livingston & Lewis, 1995). True scores are assumed to be measured perfectly but are unknown. Therefore, to provide the best estimation of classification accuracy, we use test data from one test administration to estimate the true scores based on observed scores and the parameters of the model used in estimating the true scores. It is then possible to estimate the percentages of the students who were accurately classified into each proficiency level.

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. 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. However, it is possible to estimate the percentages of the students who would be consistently classified with the assumption that the same test is independently administered twice to the same group of students.

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 student 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 5.1).

Overall Classification Accuracy and Consistency

Overall classification 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 77% of students would be

classified into the correct performance level across all six proficiency levels according to observed and true scores. **Overall classification 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 66% of students would be classified into the same performance level if two parallel forms were administered. Classification consistency values are always lower than the corresponding classification accuracy values, because in classification consistency, both of the classifications are subject to measurement error. In classification accuracy, only one of the classifications is based on a score that contains error.

Marginal Classification Accuracy and Consistency

Overall classification accuracy and consistency indicate the degree to which students are accurately and consistently classified in the same WIDA proficiency levels, but not the degree to which students are accurately or consistently classified into the proficiency levels below or above at the specific cut point (e.g., at the PL 4 or PL 5 cut). The statistics that can address this question are **marginal classification accuracy and consistency** or classification accuracy and consistency indices at the cut score level. These two terms are used interchangeably in this report. From an accountability perspective, the most important information for test users and policy makers to examine is the marginal classification accuracy and consistency.

The **classification accuracy indices at the cut** score examine the percentage of students who are accurately placed above and below the cut score. A classification accuracy index at cut score 4/5 of 0.774 means that 77% of students would be classified in the same way if they were classified according to their observed score and their true score, either into the proficiency levels below the cut score (i.e., PL 1 to PL 4) or into the proficiency levels above the cut score (i.e., PL 5 to PL 6). The **classification consistency indices at the cut** score examine the percentage of students classified consistently above and below the cut score. A classification consistency index at cut score 4/5 of 0.664 means that 66% of students would be classified in the same way if two parallel forms were administered, either into the proficiency levels below the cut score (i.e., PL 1 to PL 4) or into the proficiency levels above the cut score (i. e., PL 5 to PL 6). Note that the accuracy and consistency are generally higher at the cut scores than over the proficiency levels, or the overall classification accuracy and consistency. This is because the accuracy and consistency indices at the cut examine the classification decisions at one cut point at a time while the overall accuracy and consistency statistics examine the classification decisions at all five ACCESS cut scores at the same time.

Classification accuracy and consistency indices are affected by the interaction of the number of proficiency cuts, the magnitude of the test reliability coefficient, measurement accuracy at the cut score, the distance between adjacent cuts, the location of the cut scores on the ability scale, and the proportion of students around a cut score (Lee, Hanson, & Brennan, 2002; Ercikan & Julian, 2002), and these factors are functions of the test design and most importantly the standard setting decisions. The greater the number of proficiency levels, the lower the test reliability, the

higher the measurement accuracy at the cut scores, the closer the two adjacent cut scores, and the greater the proportion of students around a cut score, the lower the indices. Furthermore, the test reliability coefficient is affected by the numbers and types of items. For example, the test reliability estimate for the ACCESS Paper Writing domain would be lower than similar tests with more items or tasks since it is estimated based on only three or four tasks.

For each test domain, we present three tables. The first provides the overall accuracy and the overall consistency, for each grade level. The second provides the classification accuracy at the cut score, for each grade level. The third provides the classification consistency at the cut score, for each grade level.

If the overall and marginal classification accuracy and consistency indices cannot be estimated because there are fewer than 200 students in the proficiency level, we collapsed the affected proficiency level category with the category below it and placed 'N/A' in the table for the affected proficiency level.

There has been very little guidance for the ideal or expected levels of decision consistency and accuracy needed for educational assessments since these statistics are affected by many different factors, as discussed earlier. We summarize the range of overall classification accuracy and consistency of domains across grades, by domains, and highlight the grade level with the lowest classification accuracy and consistency for test users and policy makers. Since the overall accuracy and consistency statistics are a summary of the degree of classification accuracy and consistency across all proficiency level cut points, the marginal classification accuracy and consistency for these grades were further examined to identify the specific source(s) of low classification accuracy and consistency.

For Listening, as shown in Table 5.4.1.1, overall classification accuracy ranged from 0.408 to 0.695 and overall classification consistency ranged from 0.321 to 0.633. The lowest overall classification accuracy and consistency values were found for students in Grade 9.

For Reading, as shown in Table 5.4.2.1, overall classification accuracy ranged from 0.428 to 0.821 and overall classification consistency ranged from 0.331 to 0.794. The lowest overall classification accuracy and consistency values were found for students in Grade 3.

For Writing, as shown in Table 5.4.3.1, overall classification accuracy ranged from 0.719 to 0.844 and overall classification consistency ranged from 0.647 to 0.797. The lowest overall classification accuracy and consistency values were found for students in Grade 4.

For Speaking, as shown in Table 5.4.4.1, overall classification accuracy ranged from 0.466 to 0.683 and overall classification consistency ranged from 0.478 to 0.600. The lowest overall classification accuracy and consistency values were found for students in Kindergarten.

These results suggest that the grade level with the lowest classification accuracy and consistency tends to vary across the four domains.

From an accountability perspective, the most important information for test users and policy makers to examine is the marginal classification accuracy and consistency. We summarize the range of the marginal classification accuracy and consistency of domains across grades, by domain, and highlight the grade level with the lowest marginal classification accuracy and with the lowest consistency, by domain, for test users and policy makers.

For Listening, classification accuracy indices at the cut scores ranged from 0.735 to 0.993 (Table 5.4.1.2) and classification consistency at the cut scores ranged from 0.685 to 0.987 (Table 5.4.1.3). The lowest classification accuracy and consistency values were found for students in Grade 8 at the PL 5/PL 6 cut level. The low marginal classification consistency at the PL 5/PL 6 cut appeared to have contributed to its low overall classification consistency.

For Reading, classification accuracy indices at the cut scores ranged from 0.777 to 0.967 (Table 5.4.2.2) and classification consistency at the cut scores ranged from 0.710 to 0.950 (Table 5.4.2.3). The lowest classification accuracy and consistency values were found for students in Grade 3 at the PL 4/PL 5 cut. Note that Grade 3 was also identified as having the lowest overall classification consistency in the Reading domain. The low marginal classification consistency at the PL 4/PL 5 cut appeared to have contributed to its low overall classification consistency. However it should be noted that the marginal classification accuracy and consistency for Grade 3 Reading are still in the 70's.

For Writing, classification accuracy indices at the cut scores ranged from 0.777 to 0.988 (Table 5.4.3.2) and classification consistency at the cut scores ranged from 0.726 to 0.988 (Table 5.4.3.3). The lowest classification accuracy and consistency values were found for students in Grade 4 at the PL 3/PL 4 cut. The low marginal classification accuracy and consistency at the PL 3/PL 4 cut appeared to have contributed to its low overall classification accuracy and consistency. However it should be noted that the marginal classification accuracy and consistency for Grade 4 Writing are still in the 70's.

For Speaking, classification accuracy indices at the cut scores ranged from 0.723 to 0.990 (Table 5.4.4.2) and classification consistency at the cut scores ranged from 0.795 to 0.990 (Table 5.4.4.3). The lowest classification accuracy and consistency values were found for students in Kindergarten at the PL 5/PL 6 cut. Note that Kindergarten was also identified as having the lowest overall classification accuracy and consistency in the Speaking domain. The low marginal classification accuracy and consistency at the PL 5/PL 6 cut appeared to have contributed to its low overall classification accuracy and consistency. However it should be noted that the marginal classification accuracy and consistency for Kindergarten Speaking are still in the 70's.

The grades with the lowest overall classification accuracy and consistency are the same grades with the lowest marginal classification accuracy and consistency for three domains, Reading (Grade 3), Writing (Grade 4)and Speaking (Kindergarten). In Listening, Grade 9 had the lowest overall classification accuracy and consistency, and Grade 8 had the lowest marginal classification accuracy and consistency.

We observed that the lowest marginal classification accuracy and consistency for two domains (Reading and Writing) occurred at the PL 3/PL 4 and PL 4/PL 5 cut points. This finding is consistent with previous research (Lee et al., 2000; Ercikan & Julian, 2002) in that classification accuracy and consistency at cut points in the middle of the proficiency level range are lower than those in the lower and upper ends.

The higher number of proficiency levels typically results in cut scores that are closer to each other than if a smaller number of proficiency levels are used. Classification accuracy and consistency are expected to vary for different ability levels due to variation in measurement accuracy. The further away the scores are from the cut scores, the smaller the classification errors would be or the more accurate the classification decisions would be. When there is a large number of proficiency levels, more students are near the cut scores than there would be if there were fewer proficiency levels. Therefore, the higher the number of proficiency levels, the higher the probability that students would be misclassified (Ercikan & Julian, 2002). Since ACCESS has six proficiency levels and PL 3 and PL 4 occupy relatively narrow ranges on the ability scale as compared to other proficiency levels, the classification accuracy and consistency for the 3/4 and 4/5 cuts are lower than for other cuts.

The lowest marginal classification accuracy and consistency of the other two domains (Grade 8 Listening and Kindergarten Speaking) occurred at the PL 5/PL 6 cut point, which is the highest cut point on the proficiency scale. Extreme cuts tend to have larger measurement error. Thus, among the many factors mentioned earlier that affect the magnitude of classification accuracy and consistency, a large standard error at the PL 5/PL 6 cut point may have contributed to the lower classification accuracy and consistency at this cut point.

Although there has been very little guidance for the ideal or expected levels of decision consistency and accuracy needed for educational assessments since these statistics are affected by many different factors, as discussed earlier, the range of classification accuracy and consistency statistics for ACCESS domains are very similar to those reported for similar testing programs such as ELPA21 (American Institutes of Research, 2018). Also note that we do not expect the values estimated for ACCESS domains to be exactly the same as those computed in other programs, because testing programs differ in the student population, numbers of proficiency levels, test design, score distributions, and methods used to compute classification accuracy and consistency statistics. For example, ACCESS has a much larger and more diverse population and states, more proficiency levels, and a more complex test design than other similar testing programs. Therefore, it is difficult to make an absolute comparison between the classification accuracy and consistency statistics for ACCESS domains with those from other testing programs.

5.4.1 Listening

Table 5.4.1.1

Overall Accuracy and Consistency of Classification Indices: List S501 Paper

Grade	Accuracy	Consistency
K	0.695	0.633
1	0.512	0.423
2	0.479	0.396
3	0.487	0.405
4	0.501	0.411
5	0.495	0.403
6	0.447	0.362
7	0.416	0.336
8	0.419	0.336
9	0.408	0.321
10	0.416	0.322
11	0.420	0.330
12	0.434	0.344

Table 5.4.1.2

Classification Accuracy Indices at Cut Score Level: List S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.944	0.933	0.922	0.913	0.902
1	0.966	0.922	0.852	0.822	0.825
2	0.973	0.922	0.849	0.804	0.806
3	0.993	0.938	0.853	0.796	0.795
4	0.986	0.936	0.875	0.808	0.782
5	0.982	0.932	0.879	0.815	0.774
6	0.961	0.910	0.860	0.791	0.790
7	0.920	0.892	0.854	0.803	0.781
8	0.917	0.891	0.880	0.826	0.735
9	0.911	0.844	0.806	0.821	0.886
10	0.910	0.850	0.806	0.812	0.909
11	0.902	0.848	0.798	0.834	0.912
12	0.884	0.832	0.813	0.868	0.924

Table 5.4.1.3

Classification Consistency Indices at Cut Score Level: List S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.922	0.907	0.891	0.883	0.863
1	0.950	0.890	0.790	0.756	0.761
2	0.960	0.885	0.790	0.738	0.732
3	0.987	0.910	0.791	0.723	0.722
4	0.977	0.907	0.814	0.737	0.708
5	0.971	0.902	0.818	0.743	0.700
6	0.945	0.871	0.794	0.720	0.717
7	0.900	0.843	0.793	0.732	0.716
8	0.890	0.845	0.818	0.738	0.685
9	0.873	0.782	0.736	0.755	0.835
10	0.872	0.787	0.735	0.751	0.856
11	0.861	0.783	0.730	0.768	0.869
12	0.838	0.765	0.741	0.811	0.897

5.4.2 Reading

Table 5.4.2.1

Overall Accuracy and Consistency of Classification Indices: Read S501 Paper

Grade	Accuracy	Consistency
K	0.821	0.794
1	0.510	0.406
2	0.552	0.451
3	0.428	0.331
4	0.469	0.376
5	0.469	0.375
6	0.565	0.454
7	0.557	0.449
8	0.542	0.437
9	0.552	0.446
10	0.562	0.453
11	0.575	0.469
12	0.606	0.495

Table 5.4.2.2

Classification Accuracy Indices at Cut Score Level: Read S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.958	0.959	0.950	0.926	N/A
1	0.866	0.832	0.871	0.913	0.956
2	0.908	0.871	0.885	0.897	0.928
3	0.937	0.885	0.786	0.777	0.912
4	0.948	0.892	0.819	0.821	0.883
5	0.939	0.878	0.822	0.823	0.885
6	0.928	0.844	0.867	0.900	0.967
7	0.905	0.839	0.873	0.914	0.965
8	0.896	0.846	0.866	0.898	0.959
9	0.920	0.861	0.867	0.896	0.944
10	0.930	0.853	0.872	0.898	0.941
11	0.930	0.861	0.884	0.896	0.930
12	0.901	0.863	0.911	0.925	0.957

Table 5.4.2.3

Classification Consistency Indices at Cut Score Level: Read S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.937	0.940	0.932	0.919	N/A
1	0.813	0.772	0.820	0.871	0.936
2	0.869	0.822	0.837	0.854	0.898
3	0.916	0.820	0.714	0.710	0.853
4	0.928	0.840	0.759	0.756	0.838
5	0.913	0.823	0.760	0.760	0.837
6	0.894	0.784	0.816	0.860	0.950
7	0.866	0.779	0.824	0.873	0.948
8	0.853	0.785	0.818	0.857	0.936
9	0.885	0.804	0.821	0.853	0.917
10	0.898	0.798	0.822	0.853	0.915
11	0.899	0.808	0.833	0.852	0.902
12	0.863	0.810	0.870	0.891	0.940

5.4.3 Writing

Table 5.4.3.1

Overall Accuracy and Consistency of Classification Indices: Writ S501 Paper

Grade	Accuracy	Consistency
K	0.787	0.739
1	0.768	0.693
2	0.844	0.797
3	0.842	0.782
4	0.719	0.647
5	0.773	0.689
6	0.780	0.697
7	0.781	0.699
8	0.773	0.692
9	0.797	0.719
10	0.793	0.714
11	0.794	0.714
12	0.780	0.697

Table 5.4.3.2

Classification Accuracy Indices at Cut Score Level: Writ S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.926	0.889	0.962	N/A	N/A
1	0.941	0.843	0.988	N/A	N/A
2	0.968	0.936	0.940	N/A	N/A
3	0.976	0.949	0.917	N/A	N/A
4	0.979	0.962	0.777	N/A	N/A
5	0.981	0.963	0.828	N/A	N/A
6	0.970	0.949	0.860	N/A	N/A
7	0.965	0.943	0.874	N/A	N/A
8	0.963	0.945	0.864	N/A	N/A
9	0.959	0.938	0.899	N/A	N/A
10	0.961	0.936	0.896	N/A	N/A
11	0.963	0.941	0.889	N/A	N/A
12	0.955	0.934	0.889	N/A	N/A

Table 5.4.3.3

Classification Consistency Indices at Cut Score Level: Writ S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.894	0.850	0.958	N/A	N/A
1	0.913	0.789	0.988	N/A	N/A
2	0.954	0.912	0.932	N/A	N/A
3	0.966	0.928	0.888	N/A	N/A
4	0.969	0.948	0.726	N/A	N/A
5	0.972	0.947	0.766	N/A	N/A
6	0.958	0.928	0.808	N/A	N/A
7	0.950	0.918	0.828	N/A	N/A
8	0.949	0.921	0.819	N/A	N/A
9	0.942	0.915	0.857	N/A	N/A
10	0.944	0.910	0.854	N/A	N/A
11	0.947	0.917	0.844	N/A	N/A
12	0.937	0.905	0.846	N/A	N/A

5.4.4 Speaking

Table 5.4.4.1

Overall Accuracy and Consistency of Classification Indices: Spek S501 Paper

Grade	Accuracy	Consistency
K	0.466	0.478
1	0.683	0.571
2	0.671	0.564
3	0.664	0.558
4	0.641	0.534
5	0.624	0.523
6	0.617	0.524
7	0.639	0.539
8	0.639	0.543
9	0.669	0.583
10	0.674	0.592
11	0.680	0.600
12	0.668	0.593

Table 5.4.4.2

Classification Accuracy Indices at Cut Score Level: Spek S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.887	0.928	0.949	0.952	0.723
1	0.942	0.889	0.908	0.952	0.990
2	0.955	0.902	0.897	0.945	0.966
3	0.953	0.899	0.894	0.936	0.956
4	0.960	0.926	0.891	0.900	0.956
5	0.951	0.918	0.896	0.901	0.940
6	0.956	0.932	0.905	0.904	0.911
7	0.948	0.926	0.909	0.900	0.942
8	0.946	0.924	0.901	0.916	0.936
9	0.933	0.919	0.913	0.950	0.936
10	0.936	0.911	0.917	0.966	0.933
11	0.936	0.918	0.923	0.966	0.929
12	0.936	0.913	0.932	0.965	0.912

Table 5.4.4.3

Classification Consistency Indices at Cut Score Level: Spek S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.850	0.900	0.924	0.926	0.795
1	0.913	0.846	0.868	0.932	0.990
2	0.934	0.863	0.855	0.919	0.966
3	0.931	0.859	0.852	0.920	0.956
4	0.941	0.894	0.850	0.854	0.952
5	0.929	0.886	0.857	0.856	0.938
6	0.936	0.903	0.867	0.856	0.907
7	0.926	0.897	0.873	0.856	0.933
8	0.921	0.892	0.863	0.874	0.933
9	0.904	0.885	0.878	0.920	0.938
10	0.907	0.875	0.881	0.945	0.942
11	0.909	0.884	0.891	0.946	0.938
12	0.908	0.878	0.901	0.948	0.925

5.5 Reliability of Composite Scores

The reliability of ACCESS composites evaluates the consistency of the composite scores of the students over replications of the testing procedure. Because the domains that make up the composites consist of different test items, and because items from different domains may measure different attributes, even though items within the domain are assumed to measure similar attributes, a traditional internal consistency statistic such as Cronbach alpha is not appropriate, as such statistics were developed assuming items in a test measure similar attributes. It is more appropriate to report stratified alpha (Feldt & Brennan, 1989), derived to measure consistency in students' scores when the total score consists of heterogeneous parts. Stratified alpha is a weighted average of coefficient alphas for item sets with different maximum score points or "strata." Stratified alpha is a reliability estimate computed by dividing the test into parts (strata), computing Cronbach's alpha separately for each part, and using the results to estimate a reliability coefficient for the total score. In computing the stratified Cronbach's alpha for ACCESS composites, each domain that makes up a composite is treated as a strata. For example, in computing stratified Cronbach's alpha for Literacy, two strata (Reading and Writing) are entered into the computation. The stratified Cronbach's alpha is interpreted like other traditional internal consistency statistics such as Cronbach's coefficient alpha. Like Cronbach's alpha, stratified Cronbach's alpha is an estimate of the proportion of the total variance of the composite that can be explained by the variance of the true score.

Because of the differential weights applied to the ACCESS domains that contribute to the composites, the stratified Cronbach's alpha coefficient is weighted by the contribution of each domain score into the composite (Rudner, 2001; Kamata, Turhan, & Darandari, 2003; Kane & Case, 2004). 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

σ_j^2 = variance of component j

σ_c^2 = variance of composite

ρ_j = reliability coefficient of component j

The tables below express the stratified Cronbach's alpha for each of the composites. The first table for each composite provides stratified Cronbach's alpha for all students. 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.

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. (Detail on how the composites are computed is provided in the introduction to Chapter 3.)

For each grade-level cluster excluding Kindergarten, we derive a reliability coefficient across tiers for each domain. (The Kindergarten test is not tiered and so this step is not necessary.) To produce this coefficient, values for Cronbach's alpha for each of the tiers in the grade-level cluster (provided in Section 5.1) 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, we provide the variance of the scale score. We also provide the variance of the composite scale score. The variances of domains and composites are computed for students who had valid results in all four domains.

Finally, the table presents the computed stratified Cronbach's alpha value for the composite, 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 5.7).

The stratified Cronbach's alpha of the Oral composite computed for all students ranged from 0.88 to 0.95. The stratified Cronbach's alpha of the Oral composite ranged from 0.88 to 0.95 for male students; from 0.88 to 0.95 for female students; from 0.88 to 0.96 for Hispanic students; from 0.88 to 0.95 for non-Hispanic students; and from 0.85 to 0.96 for students with an IEP.

The stratified Cronbach's alpha of the Literacy composite computed for all students ranged from 0.90 to 0.96. The stratified Cronbach's alpha of the Literacy composite ranged from 0.90 to 0.97 for male students; from 0.90 to 0.96 for female students; from 0.90 to 0.96 for Hispanic students; from 0.91 to 0.97 for non-Hispanic students; and from 0.88 to 0.97 for students with an IEP.

The stratified Cronbach's alpha of the Comprehension composite computed for all students ranged from 0.78 to 0.96. The stratified Cronbach's alpha of the Comprehension composite ranged from 0.79 to 0.97 for male students; from 0.76 to 0.96 for female students; from 0.77 to 0.96 for Hispanic students; from 0.79 to 0.97 for non-Hispanic students; and from 0.69 to 0.97 for students with an IEP.

The stratified Cronbach's alpha of the Overall composite computed for all students ranged from 0.94 to 0.97. The stratified Cronbach's alpha of the Overall composite ranged from 0.94 to 0.98 for male students; from 0.93 to 0.97 for female students; from 0.94 to 0.97 for Hispanic students; from 0.94 to 0.97 for non-Hispanic students; and from 0.92 to 0.98 for students with an IEP.

5.5.1 Oral

Table 5.5.1.1

Reliability of Composite: Oral S501 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.50	5771.451	0.943
	Speaking	0.50	10216.448	0.904
	Oral		7008.004	0.953
1	Listening	0.50	1481.307	0.700
	Speaking	0.50	4011.636	0.891
	Oral		1986.789	0.889
2	Listening	0.50	1664.999	0.678
	Speaking	0.50	4262.137	0.907
	Oral		2219.334	0.895
3	Listening	0.50	1402.071	0.598
	Speaking	0.50	4186.433	0.905
	Oral		2017.455	0.881
4-5	Listening	0.50	1816.602	0.635
	Speaking	0.50	5157.761	0.903
	Oral		2695.605	0.892
6-8	Listening	0.50	2589.510	0.648
	Speaking	0.50	6056.440	0.914
	Oral		3532.991	0.898
9-12	Listening	0.50	2499.567	0.663
	Speaking	0.50	6051.411	0.922
	Oral		3440.068	0.904

Table 5.5.1.2

Reliability of Composite: Oral S501 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.50	5514.549	0.942	5881.517	0.943
	Speaking	0.50	10258.485	0.907	9998.351	0.901
	Oral		6894.582	0.954	6962.548	0.952
1	Listening	0.50	1419.122	0.686	1538.382	0.711
	Speaking	0.50	3984.504	0.891	4004.217	0.892
	Oral		1939.559	0.886	2018.077	0.891
2	Listening	0.50	1601.314	0.664	1700.846	0.687
	Speaking	0.50	4268.836	0.906	4207.864	0.907
	Oral		2193.276	0.893	2208.019	0.895
3	Listening	0.50	1330.386	0.581	1456.863	0.611
	Speaking	0.50	4186.314	0.902	4135.265	0.905
	Oral		1977.078	0.878	2027.597	0.882
4-5	Listening	0.50	1744.322	0.619	1870.069	0.645
	Speaking	0.50	5191.717	0.902	5104.365	0.903
	Oral		2660.243	0.890	2709.858	0.893
6-8	Listening	0.50	2583.343	0.635	2591.180	0.655
	Speaking	0.50	5983.455	0.911	6159.317	0.917
	Oral		3508.467	0.895	3566.988	0.901
9-12	Listening	0.50	2397.569	0.654	2565.732	0.672
	Speaking	0.50	6046.394	0.920	5948.524	0.922
	Oral		3379.588	0.903	3433.337	0.905

Table 5.5.1.3

Reliability of Composite: Oral S501 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.50	5908.101	0.944	5028.816	0.936
	Speaking	0.50	10427.358	0.907	9259.327	0.890
	Oral		7174.774	0.955	6168.033	0.946
1	Listening	0.50	1480.233	0.698	1447.740	0.706
	Speaking	0.50	4057.281	0.894	3602.054	0.878
	Oral		2002.323	0.890	1807.410	0.880
2	Listening	0.50	1698.219	0.673	1551.983	0.690
	Speaking	0.50	4346.171	0.909	3840.019	0.898
	Oral		2272.234	0.895	1975.713	0.890
3	Listening	0.50	1397.926	0.588	1376.735	0.617
	Speaking	0.50	4313.637	0.906	3688.454	0.898
	Oral		2061.955	0.881	1805.767	0.875
4-5	Listening	0.50	1859.857	0.628	1606.061	0.646
	Speaking	0.50	5391.290	0.905	4175.704	0.897
	Oral		2820.376	0.893	2139.580	0.883
6-8	Listening	0.50	2700.274	0.643	2131.216	0.650
	Speaking	0.50	6408.645	0.915	4563.414	0.905
	Oral		3744.462	0.899	2633.966	0.888
9-12	Listening	0.50	2535.262	0.658	2287.028	0.674
	Speaking	0.50	6276.683	0.923	5096.237	0.915
	Oral		3559.070	0.905	2901.382	0.899

Table 5.5.1.4

Reliability of Composite: Oral S501 Paper by IEP Status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.50	6583.408	0.951
	Speaking	0.50	9044.969	0.897
	Oral		6932.713	0.955
1	Listening	0.50	1510.575	0.714
	Speaking	0.50	3297.795	0.883
	Oral		1729.199	0.882
2	Listening	0.50	1563.513	0.715
	Speaking	0.50	3388.686	0.903
	Oral		1776.753	0.891
3	Listening	0.50	1159.801	0.570
	Speaking	0.50	3047.597	0.896
	Oral		1399.389	0.854
4-5	Listening	0.50	1275.239	0.619
	Speaking	0.50	3153.847	0.895
	Oral		1518.823	0.865
6-8	Listening	0.50	1561.073	0.630
	Speaking	0.50	3683.661	0.906
	Oral		1897.791	0.878
9-12	Listening	0.50	1715.407	0.616
	Speaking	0.50	4868.811	0.919
	Oral		2476.498	0.894

5.5.2 Literacy

Table 5.5.2.1

Reliability of Composite: Litr S501 Paper

Cluster	Component	Weight	Variance	Reliability
K	Reading	0.50	4548.221	0.950
	Writing	0.50	4472.087	0.924
	Literacy		3879.382	0.963
1	Reading	0.50	922.950	0.768
	Writing	0.50	1678.607	0.914
	Literacy		991.793	0.910
2	Reading	0.50	1200.405	0.838
	Writing	0.50	1804.048	0.938
	Literacy		1247.466	0.939
3	Reading	0.50	772.754	0.667
	Writing	0.50	1617.750	0.929
	Literacy		944.992	0.901
4-5	Reading	0.50	1024.862	0.769
	Writing	0.50	1638.155	0.909
	Literacy		1112.615	0.913
6-8	Reading	0.50	879.522	0.774
	Writing	0.50	1804.690	0.909
	Literacy		1095.670	0.917
9-12	Reading	0.50	1016.586	0.810
	Writing	0.50	1999.257	0.905
	Literacy		1240.615	0.923

Table 5.5.2.2

Reliability of Composite: Litr S501 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Reading	0.50	4382.288	0.947	4702.879	0.953
	Writing	0.50	4344.313	0.922	4539.795	0.924
	Literacy		3743.959	0.962	3984.841	0.965
1	Reading	0.50	892.374	0.764	955.617	0.773
	Writing	0.50	1508.196	0.912	1794.159	0.916
	Literacy		919.810	0.907	1048.153	0.912
2	Reading	0.50	1150.753	0.834	1238.438	0.842
	Writing	0.50	1668.244	0.931	1827.428	0.939
	Literacy		1172.294	0.935	1274.000	0.940
3	Reading	0.50	726.497	0.650	810.008	0.683
	Writing	0.50	1524.613	0.921	1604.039	0.929
	Literacy		896.335	0.895	957.285	0.903
4-5	Reading	0.50	962.931	0.756	1079.749	0.781
	Writing	0.50	1561.727	0.897	1634.779	0.911
	Literacy		1063.078	0.907	1135.978	0.916
6-8	Reading	0.50	827.408	0.767	919.580	0.780
	Writing	0.50	1752.365	0.895	1792.814	0.912
	Literacy		1048.223	0.910	1110.967	0.919
9-12	Reading	0.50	986.260	0.804	1042.366	0.816
	Writing	0.50	1947.131	0.899	1982.040	0.906
	Literacy		1214.087	0.920	1237.186	0.924

Table 5.5.2.3

Reliability of Composite: Litr S501 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Reading	0.50	4140.677	0.946	4810.608	0.954
	Writing	0.50	4201.218	0.916	4485.604	0.929
	Literacy		3540.548	0.959	4012.926	0.966
1	Reading	0.50	882.597	0.752	983.478	0.797
	Writing	0.50	1658.320	0.913	1700.082	0.918
	Literacy		960.664	0.905	1032.966	0.918
2	Reading	0.50	1185.691	0.832	1199.209	0.849
	Writing	0.50	1848.880	0.938	1637.730	0.936
	Literacy		1258.450	0.938	1179.189	0.939
3	Reading	0.50	763.488	0.647	776.824	0.701
	Writing	0.50	1638.633	0.928	1519.197	0.931
	Literacy		950.537	0.898	901.638	0.906
4-5	Reading	0.50	1036.562	0.758	924.655	0.784
	Writing	0.50	1715.661	0.908	1342.116	0.910
	Literacy		1155.143	0.912	923.063	0.913
6-8	Reading	0.50	876.526	0.768	822.251	0.782
	Writing	0.50	1881.837	0.910	1503.022	0.906
	Literacy		1127.330	0.917	936.059	0.914
9-12	Reading	0.50	1028.307	0.807	935.832	0.817
	Writing	0.50	2033.792	0.905	1833.256	0.904
	Literacy		1263.391	0.923	1120.139	0.922

Table 5.5.2.4

Reliability of Composite: Litr S501 Paper by IEP Status

Cluster	Component	Weight	Variance	Reliability
K	Reading	0.50	4781.983	0.960
	Writing	0.50	4267.931	0.925
	Literacy		3813.961	0.967
1	Reading	0.50	769.068	0.701
	Writing	0.50	2032.875	0.918
	Literacy		986.153	0.900
2	Reading	0.50	975.932	0.808
	Writing	0.50	1870.096	0.948
	Literacy		1125.073	0.937
3	Reading	0.50	609.917	0.547
	Writing	0.50	1444.758	0.939
	Literacy		769.261	0.881
4-5	Reading	0.50	647.252	0.702
	Writing	0.50	1193.529	0.920
	Literacy		701.613	0.897
6-8	Reading	0.50	563.734	0.704
	Writing	0.50	1269.673	0.920
	Literacy		675.246	0.901
9-12	Reading	0.50	684.770	0.747
	Writing	0.50	1686.056	0.914
	Literacy		892.989	0.911

5.5.3 Comprehension

Table 5.5.3.1

Reliability of Composite: Cphn S501 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.30	5771.451	0.943
	Reading	0.70	4548.221	0.950
	Comprehension		3878.875	0.964
1	Listening	0.30	1481.307	0.700
	Reading	0.70	922.950	0.768
	Comprehension		830.803	0.825
2	Listening	0.30	1664.999	0.678
	Reading	0.70	1200.405	0.838
	Comprehension		1104.575	0.870
3	Listening	0.30	1402.071	0.598
	Reading	0.70	772.754	0.667
	Comprehension		790.876	0.777
4-5	Listening	0.30	1816.602	0.635
	Reading	0.70	1024.862	0.769
	Comprehension		1085.800	0.838
6-8	Listening	0.30	2589.510	0.648
	Reading	0.70	879.522	0.774
	Comprehension		1125.295	0.841
9-12	Listening	0.30	2499.567	0.663
	Reading	0.70	1016.586	0.810
	Comprehension		1211.418	0.859

Table 5.5.3.2

Reliability of Composite: Cphn S501 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.30	5514.549	0.942	5881.517	0.943
	Reading	0.70	4382.288	0.947	4702.879	0.953
	Comprehension		3745.433	0.962	3977.697	0.965
1	Listening	0.30	1419.122	0.686	1538.382	0.711
	Reading	0.70	892.374	0.764	955.617	0.773
	Comprehension		805.710	0.822	858.612	0.830
2	Listening	0.30	1601.314	0.664	1700.846	0.687
	Reading	0.70	1150.753	0.834	1238.438	0.842
	Comprehension		1059.901	0.866	1134.084	0.873
3	Listening	0.30	1330.386	0.581	1456.863	0.611
	Reading	0.70	726.497	0.650	810.008	0.683
	Comprehension		741.739	0.764	829.393	0.787
4-5	Listening	0.30	1744.322	0.619	1870.069	0.645
	Reading	0.70	962.931	0.756	1079.749	0.781
	Comprehension		1025.861	0.830	1136.574	0.845
6-8	Listening	0.30	2583.343	0.635	2591.180	0.655
	Reading	0.70	827.408	0.767	919.580	0.780
	Comprehension		1083.118	0.834	1156.210	0.845
9-12	Listening	0.30	2397.569	0.654	2565.732	0.672
	Reading	0.70	986.260	0.804	1042.366	0.816
	Comprehension		1171.377	0.856	1241.830	0.863

Table 5.5.3.3

Reliability of Composite: Cphn S501 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.30	5908.101	0.944	5028.816	0.936
	Reading	0.70	4140.677	0.946	4810.608	0.954
	Comprehension		3622.925	0.961	3881.163	0.965
1	Listening	0.30	1480.233	0.698	1447.740	0.706
	Reading	0.70	882.597	0.752	983.478	0.797
	Comprehension		799.510	0.815	874.410	0.845
2	Listening	0.30	1698.219	0.673	1551.983	0.690
	Reading	0.70	1185.691	0.832	1199.209	0.849
	Comprehension		1101.161	0.866	1078.765	0.878
3	Listening	0.30	1397.926	0.588	1376.735	0.617
	Reading	0.70	763.488	0.647	776.824	0.701
	Comprehension		782.771	0.765	784.551	0.794
4-5	Listening	0.30	1859.857	0.628	1606.061	0.646
	Reading	0.70	1036.562	0.758	924.655	0.784
	Comprehension		1105.670	0.833	954.617	0.844
6-8	Listening	0.30	2700.274	0.643	2131.216	0.650
	Reading	0.70	876.526	0.768	822.251	0.782
	Comprehension		1144.408	0.837	987.530	0.843
9-12	Listening	0.30	2535.262	0.658	2287.028	0.674
	Reading	0.70	1028.307	0.807	935.832	0.817
	Comprehension		1230.367	0.858	1091.408	0.862

Table 5.5.3.4

Reliability of Composite: Cphn S501 Paper by IEP Status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.30	6583.408	0.951
	Reading	0.70	4781.983	0.960
	Comprehension		4004.060	0.969
1	Listening	0.30	1510.575	0.714
	Reading	0.70	769.068	0.701
	Comprehension		696.732	0.783
2	Listening	0.30	1563.513	0.715
	Reading	0.70	975.932	0.808
	Comprehension		878.687	0.850
3	Listening	0.30	1159.801	0.570
	Reading	0.70	609.917	0.547
	Comprehension		586.208	0.692
4-5	Listening	0.30	1275.239	0.619
	Reading	0.70	647.252	0.702
	Comprehension		649.707	0.787
6-8	Listening	0.30	1561.073	0.630
	Reading	0.70	563.734	0.704
	Comprehension		657.218	0.796
9-12	Listening	0.30	1715.407	0.616
	Reading	0.70	684.770	0.747
	Comprehension		784.632	0.816

5.5.4 Overall

Table 5.5.4.1

Reliability of Composite: Over S501 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.15	5771.451	0.943
	Reading	0.35	4548.221	0.950
	Writing	0.35	4472.087	0.924
	Speaking	0.15	10216.448	0.904
	Overall Composite		3841.385	0.974
1	Listening	0.15	1481.307	0.700
	Reading	0.35	922.950	0.768
	Writing	0.35	1678.607	0.914
	Speaking	0.15	4011.636	0.891
	Overall Composite		1027.368	0.938
2	Listening	0.15	1664.999	0.678
	Reading	0.35	1200.405	0.838
	Writing	0.35	1804.048	0.938
	Speaking	0.15	4262.137	0.907
	Overall Composite		1286.324	0.955
3	Listening	0.15	1402.071	0.598
	Reading	0.35	772.754	0.667
	Writing	0.35	1617.750	0.929
	Speaking	0.15	4186.433	0.905
	Overall Composite		1052.924	0.936
4-5	Listening	0.15	1816.602	0.635
	Reading	0.35	1024.862	0.769
	Writing	0.35	1638.155	0.909
	Speaking	0.15	5157.761	0.903
	Overall Composite		1351.735	0.946
6-8	Listening	0.15	2589.510	0.648
	Reading	0.35	879.522	0.774
	Writing	0.35	1804.690	0.909
	Speaking	0.15	6056.440	0.914
	Overall Composite		1530.817	0.950
9-12	Listening	0.15	2499.567	0.663
	Reading	0.35	1016.586	0.810
	Writing	0.35	1999.257	0.905
	Speaking	0.15	6051.411	0.922
	Overall Composite		1610.196	0.952

Table 5.5.4.2

Reliability of Composite: Over S501 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.15	5514.549	0.942	5881.517	0.943
	Reading	0.35	4382.288	0.947	4702.879	0.953
	Writing	0.35	4344.313	0.922	4539.795	0.924
	Speaking	0.15	10258.485	0.907	9998.351	0.901
	Overall Composite		3735.869	0.974	3891.854	0.975
1	Listening	0.15	1419.122	0.686	1538.382	0.711
	Reading	0.35	892.374	0.764	955.617	0.773
	Writing	0.35	1508.196	0.912	1794.159	0.916
	Speaking	0.15	3984.504	0.891	4004.217	0.892
	Overall Composite		970.582	0.936	1071.226	0.939
2	Listening	0.15	1601.314	0.664	1700.846	0.687
	Reading	0.35	1150.753	0.834	1238.438	0.842
	Writing	0.35	1668.244	0.931	1827.428	0.939
	Speaking	0.15	4268.836	0.906	4207.864	0.907
	Overall Composite		1232.198	0.952	1296.797	0.955
3	Listening	0.15	1330.386	0.581	1456.863	0.611
	Reading	0.35	726.497	0.650	810.008	0.683
	Writing	0.35	1524.613	0.921	1604.039	0.929
	Speaking	0.15	4186.314	0.902	4135.265	0.905
	Overall Composite		1011.231	0.933	1062.590	0.937
4-5	Listening	0.15	1744.322	0.619	1870.069	0.645
	Reading	0.35	962.931	0.756	1079.749	0.781
	Writing	0.35	1561.727	0.897	1634.779	0.911
	Speaking	0.15	5191.717	0.902	5104.365	0.903
	Overall Composite		1311.816	0.943	1371.185	0.947
6-8	Listening	0.15	2583.343	0.635	2591.180	0.655
	Reading	0.35	827.408	0.767	919.580	0.780
	Writing	0.35	1752.365	0.895	1792.814	0.912
	Speaking	0.15	5983.455	0.911	6159.317	0.917
	Overall Composite		1496.066	0.947	1548.881	0.951
9-12	Listening	0.15	2397.569	0.654	2565.732	0.672
	Reading	0.35	986.260	0.804	1042.366	0.816
	Writing	0.35	1947.131	0.899	1982.040	0.906
	Speaking	0.15	6046.394	0.920	5948.524	0.922
	Overall Composite		1591.483	0.951	1598.505	0.953

Table 5.5.4.3

Reliability of Composite: Over S501 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.15	5908.101	0.944	5028.816	0.936
	Reading	0.35	4140.677	0.946	4810.608	0.954
	Writing	0.35	4201.218	0.916	4485.604	0.929
	Speaking	0.15	10427.358	0.907	9259.327	0.890
	Overall Composite		3627.948	0.972	3752.801	0.974
1	Listening	0.15	1480.233	0.698	1447.740	0.706
	Reading	0.35	882.597	0.752	983.478	0.797
	Writing	0.35	1658.320	0.913	1700.082	0.918
	Speaking	0.15	4057.281	0.894	3602.054	0.878
	Overall Composite		1006.552	0.936	1013.813	0.940
2	Listening	0.15	1698.219	0.673	1551.983	0.690
	Reading	0.35	1185.691	0.832	1199.209	0.849
	Writing	0.35	1848.880	0.938	1637.730	0.936
	Speaking	0.15	4346.171	0.909	3840.019	0.898
	Overall Composite		1304.323	0.954	1182.139	0.954
3	Listening	0.15	1397.926	0.588	1376.735	0.617
	Reading	0.35	763.488	0.647	776.824	0.701
	Writing	0.35	1638.633	0.928	1519.197	0.931
	Speaking	0.15	4313.637	0.906	3688.454	0.898
	Overall Composite		1067.928	0.935	966.395	0.936
4-5	Listening	0.15	1859.857	0.628	1606.061	0.646
	Reading	0.35	1036.562	0.758	924.655	0.784
	Writing	0.35	1715.661	0.908	1342.116	0.910
	Speaking	0.15	5391.290	0.905	4175.704	0.897
	Overall Composite		1414.561	0.945	1070.522	0.942
6-8	Listening	0.15	2700.274	0.643	2131.216	0.650
	Reading	0.35	876.526	0.768	822.251	0.782
	Writing	0.35	1881.837	0.910	1503.022	0.906
	Speaking	0.15	6408.645	0.915	4563.414	0.905
	Overall Composite		1603.532	0.950	1203.012	0.945
9-12	Listening	0.15	2535.262	0.658	2287.028	0.674
	Reading	0.35	1028.307	0.807	935.832	0.817
	Writing	0.35	2033.792	0.905	1833.256	0.904
	Speaking	0.15	6276.683	0.923	5096.237	0.915
	Overall Composite		1658.397	0.953	1381.308	0.950

Table 5.5.4.4

Reliability of Composite: Over S501 Paper by IEP Status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.15	6583.408	0.951
	Reading	0.35	4781.983	0.960
	Writing	0.35	4267.931	0.925
	Speaking	0.15	9044.969	0.897
	Overall Composite		3712.706	0.976
1	Listening	0.15	1510.575	0.714
	Reading	0.35	769.068	0.701
	Writing	0.35	2032.875	0.918
	Speaking	0.15	3297.795	0.883
	Overall Composite		945.730	0.929
2	Listening	0.15	1563.513	0.715
	Reading	0.35	975.932	0.808
	Writing	0.35	1870.096	0.948
	Speaking	0.15	3388.686	0.903
	Overall Composite		1059.380	0.951
3	Listening	0.15	1159.801	0.570
	Reading	0.35	609.917	0.547
	Writing	0.35	1444.758	0.939
	Speaking	0.15	3047.597	0.896
	Overall Composite		747.318	0.916
4-5	Listening	0.15	1275.239	0.619
	Reading	0.35	647.252	0.702
	Writing	0.35	1193.529	0.920
	Speaking	0.15	3153.847	0.895
	Overall Composite		734.803	0.927
6-8	Listening	0.15	1561.073	0.630
	Reading	0.35	563.734	0.704
	Writing	0.35	1269.673	0.920
	Speaking	0.15	3683.661	0.906
	Overall Composite		806.974	0.934
9-12	Listening	0.15	1715.407	0.616
	Reading	0.35	684.770	0.747
	Writing	0.35	1686.056	0.914
	Speaking	0.15	4868.811	0.919
	Overall Composite		1102.679	0.943

5.6 Conditional Standard Error of Measurement 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 Section 3 for weighting conventions), we estimate the CSEMs 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. We used this method to estimate the CSEM for ACCESS composites by treating the ACCESS domains as subtests.

We use a three-step process to derive the CSEM for ACCESS composites. We conduct the derivation by grade and composite to obtain a unique CSEM for each composite score by grade. Since this procedure replies on empirical student data, which are subject to year-to-year fluctuation, we use all population student data from the previous ACCESS series in the derivation to obtain more stable estimates than using only data from a single series.

Step 1. Since we calibrated ACCESS domains 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, we took an additional step to obtain a unique value for each composite score. Specifically, we estimated the expected value of the CSEM functions for a composite score using a regression approach, and we reported this expected value as the 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. We do not plot values for students who received the lowest possible scores on any ACCESS domains, as it is not possible to compute accurately the conditional measurement errors for these students. For grade-level clusters with multiple grades, we use different colors in the figures to represent students in different grades.

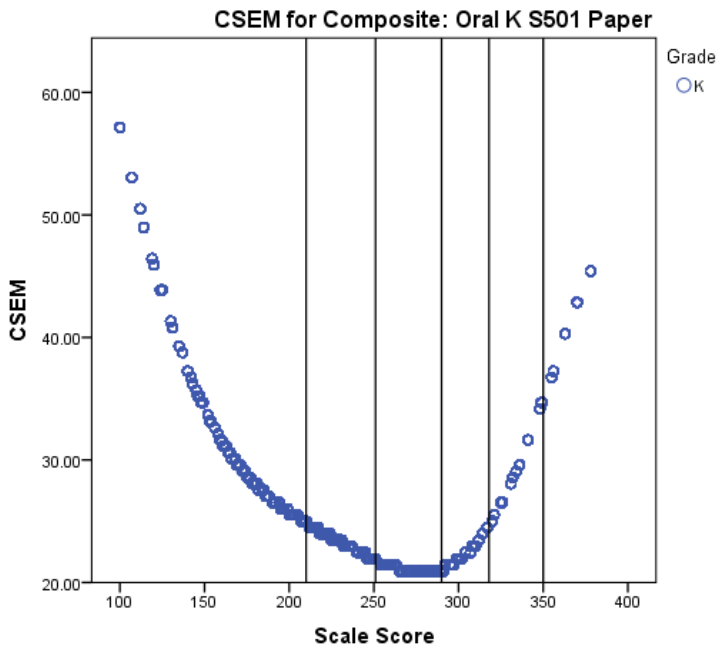
Five vertical lines in the figure indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1–6) for the composites.

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. As noted elsewhere in this report, students are exited from the ACCESS population upon gaining English language proficiency, and therefore these students are removed from the ACCESS population, resulting in smaller numbers of students at the highest cut points.

5.6.1 Oral

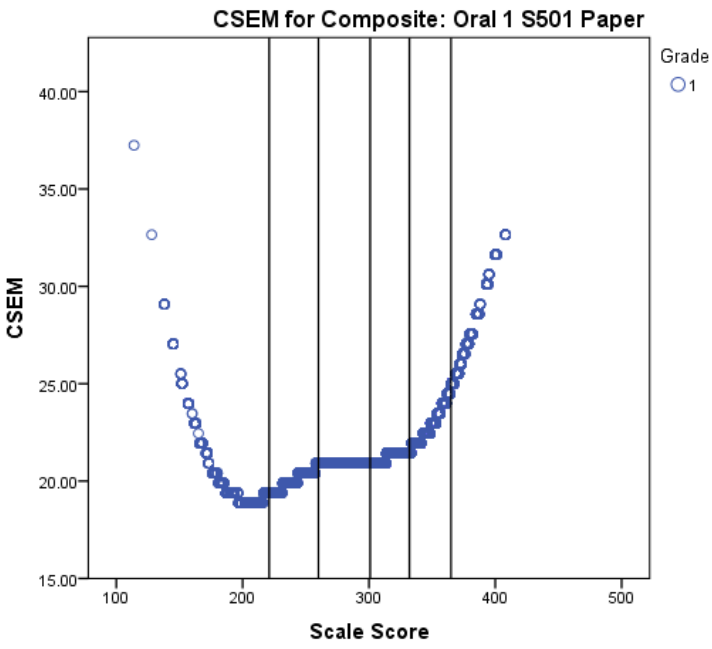
5.6.1.0 Kindergarten

Figure 5.6.1.0



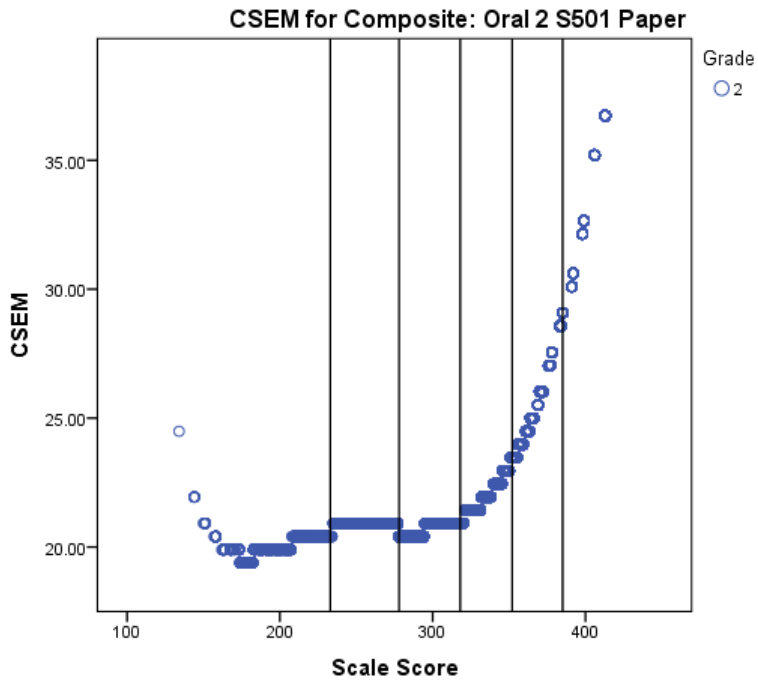
5.6.1.1 Grade 1

Figure 5.6.1.1



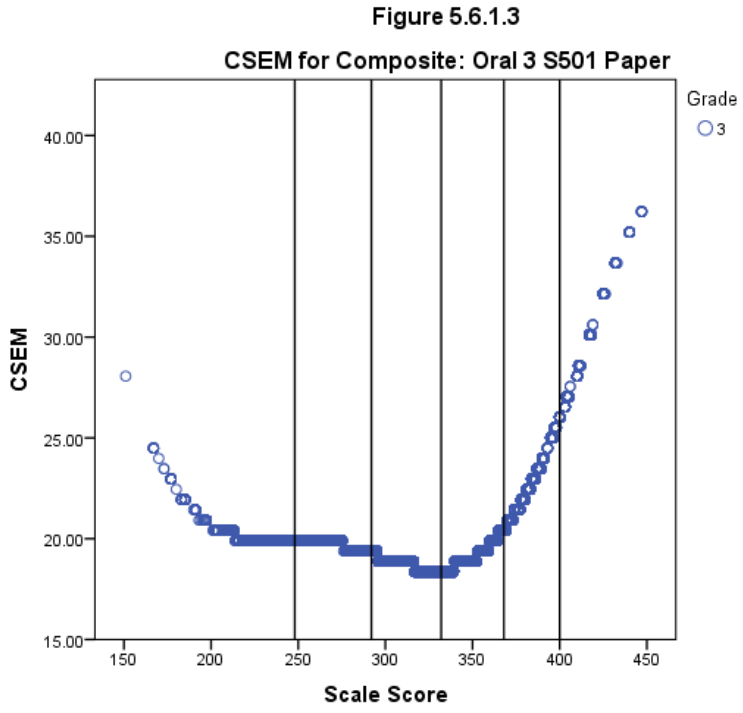
5.6.1.2 Grade 2

Figure 5.6.1.2



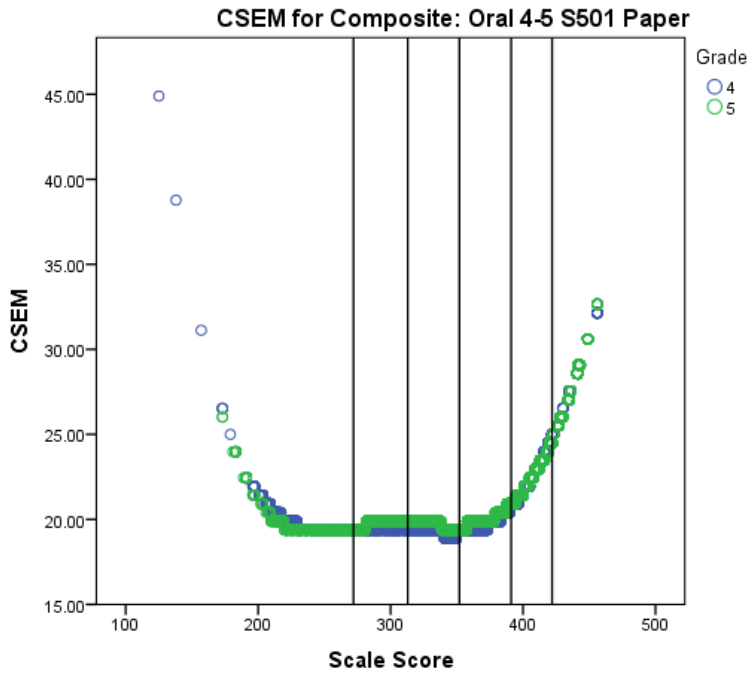
5.6.1.3 Grade 3

Figure 5.6.1.3



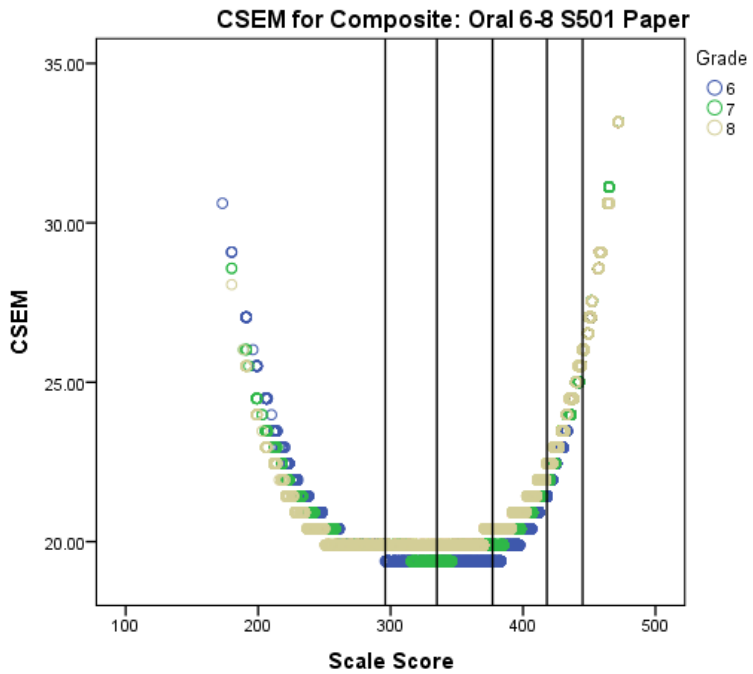
5.6.1.4 Grades 4–5

Figure 5.6.1.4



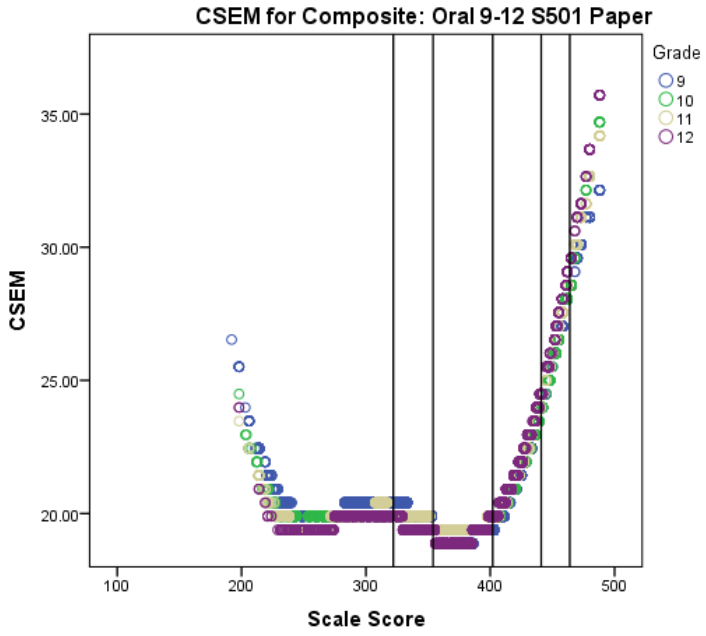
5.6.1.5 Grades 6–8

Figure 5.6.1.5



5.6.1.6 Grades 9–12

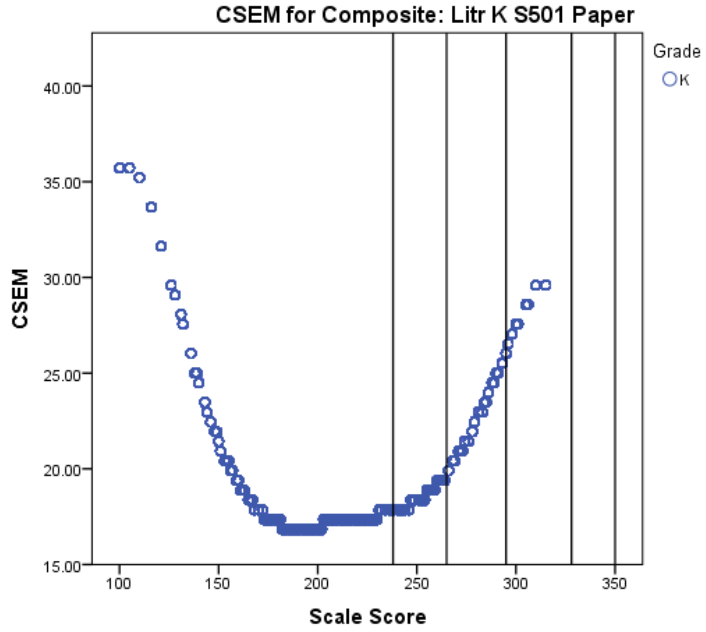
Figure 5.6.1.6



5.6.2 Literacy

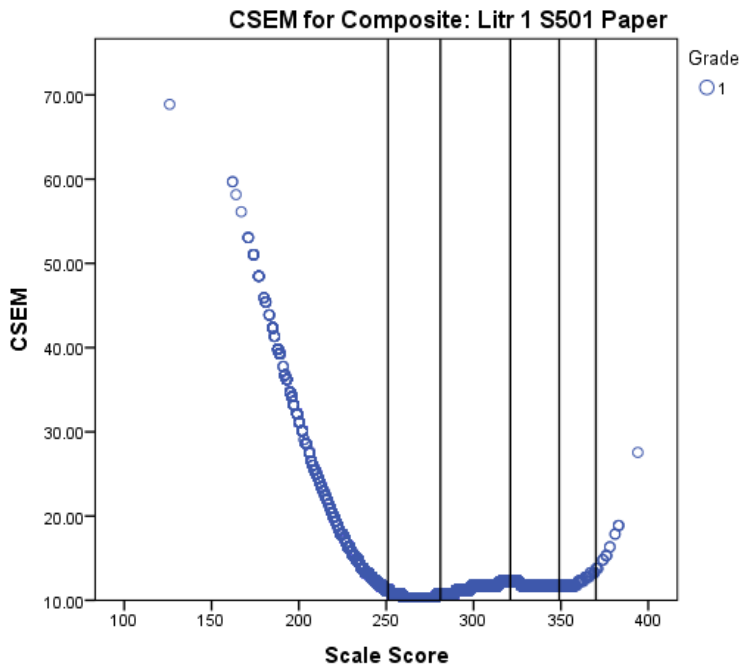
5.6.2.0 Kindergarten

Figure 5.6.2.0



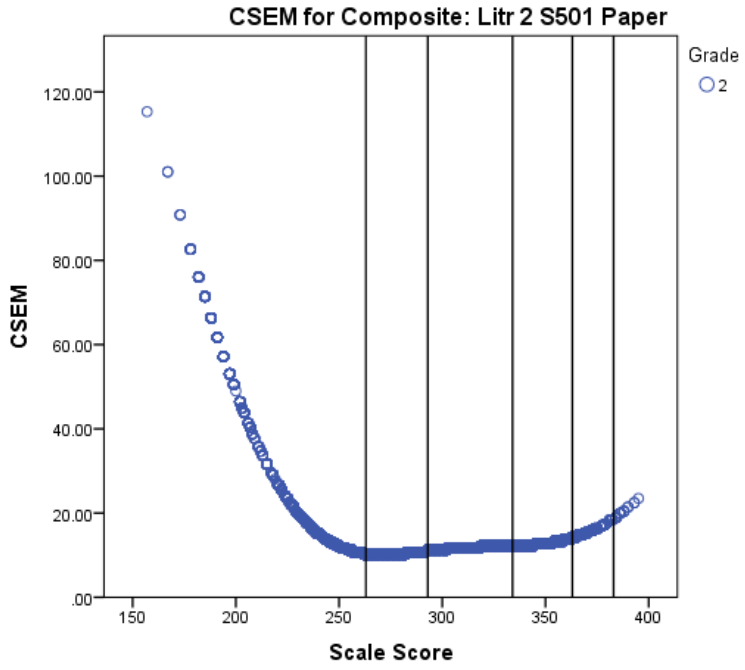
5.6.2.1 Grade 1

Figure 5.6.2.1



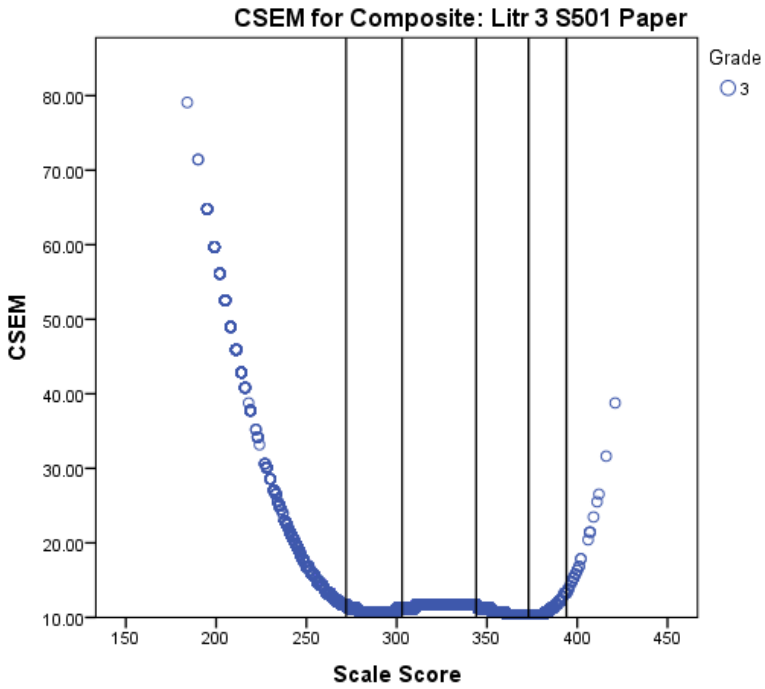
5.6.2.2 Grade 2

Figure 5.6.2.2



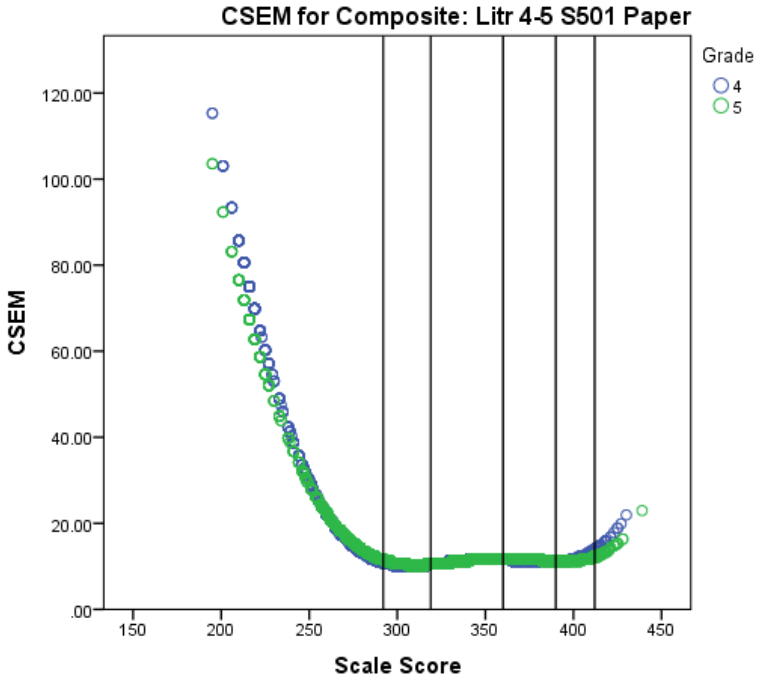
5.6.2.3 Grade 3

Figure 5.6.2.3



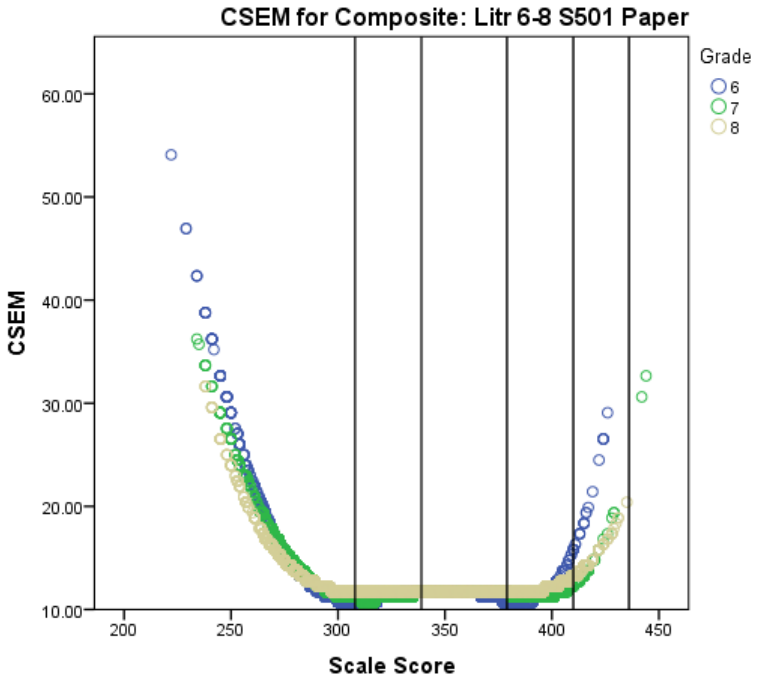
5.6.2.4 Grades 4–5

Figure 5.6.2.4



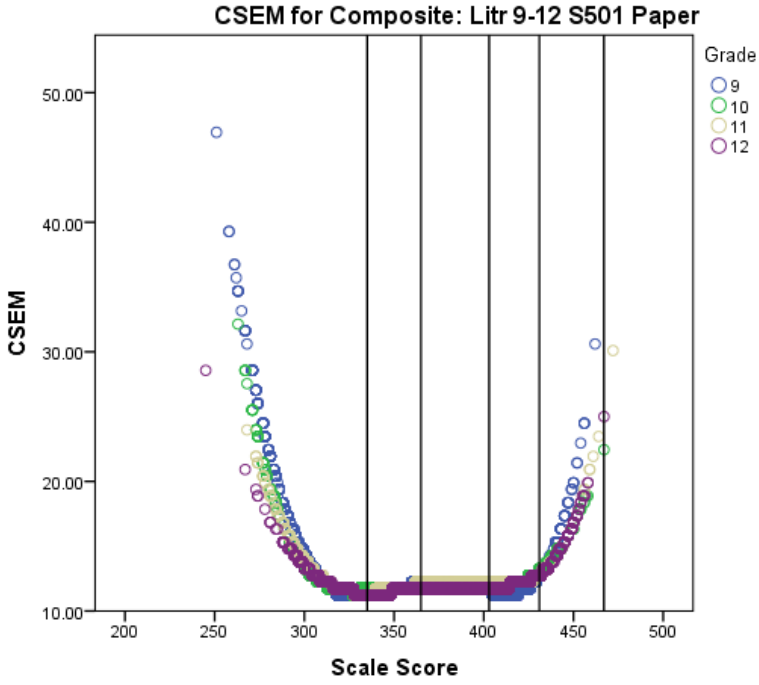
5.6.2.5 Grades 6–8

Figure 5.6.2.5



5.6.2.6 Grades 9–12

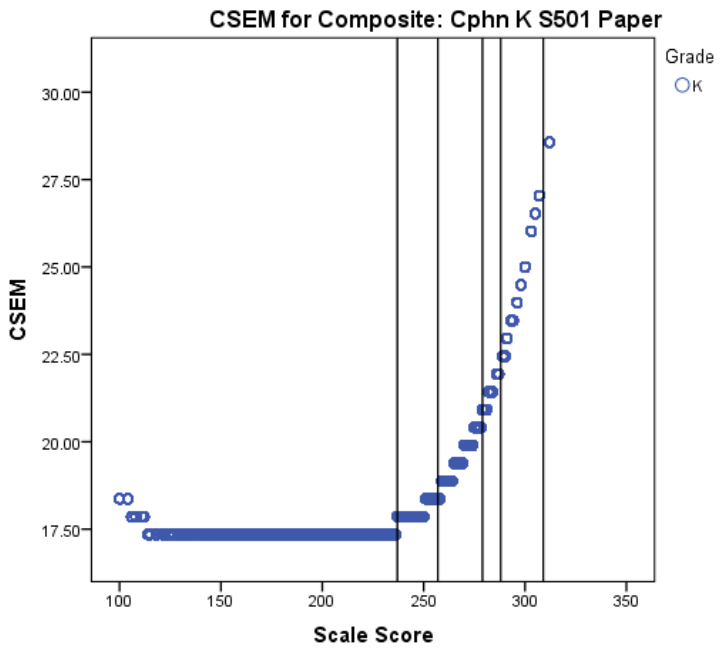
Figure 5.6.2.6



5.6.3 Comprehension

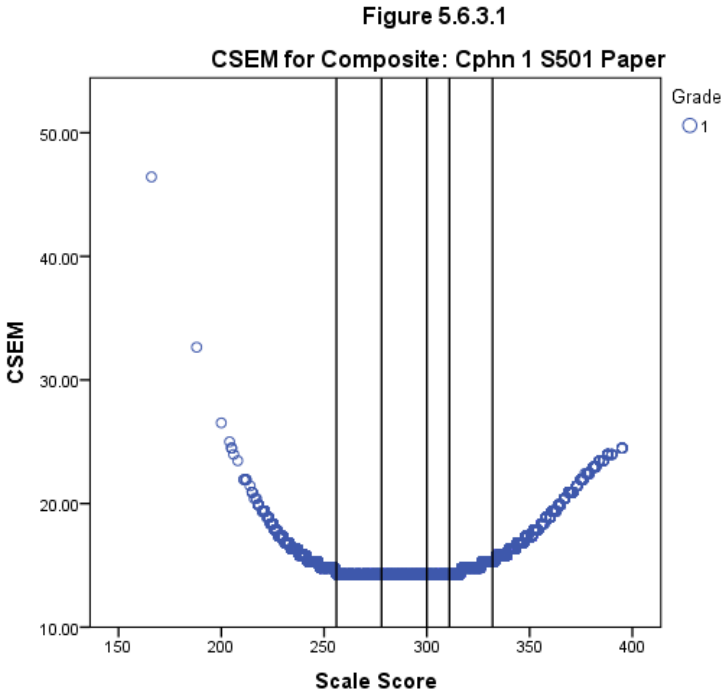
5.6.3.0 Kindergarten

Figure 5.6.3.0



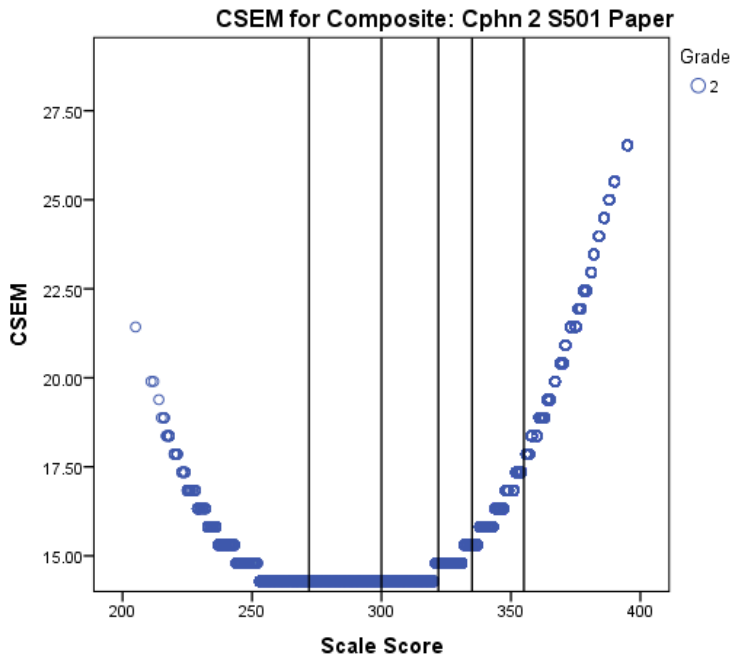
5.6.3.1 Grade 1

Figure 5.6.3.1



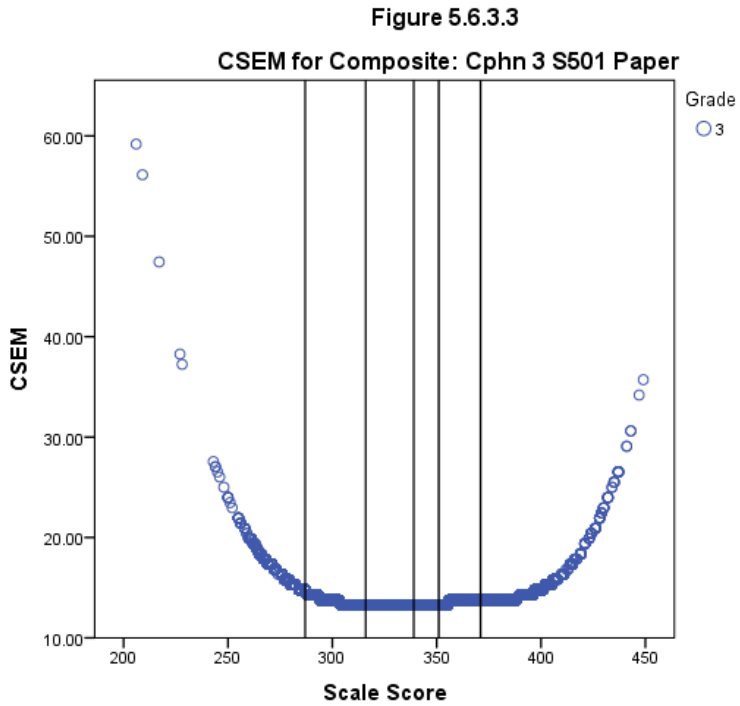
5.6.3.2 Grade 2

Figure 5.6.3.2



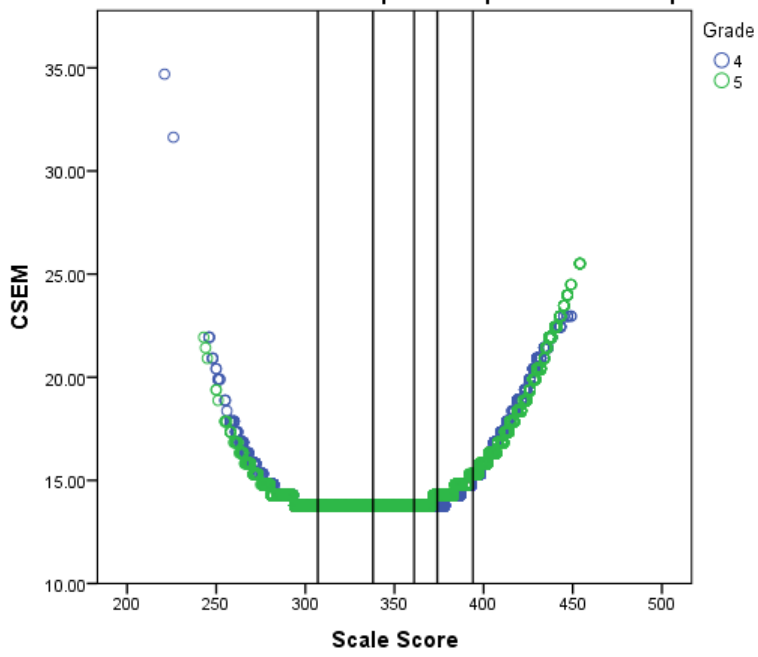
5.6.3.3 Grade 3

Figure 5.6.3.3



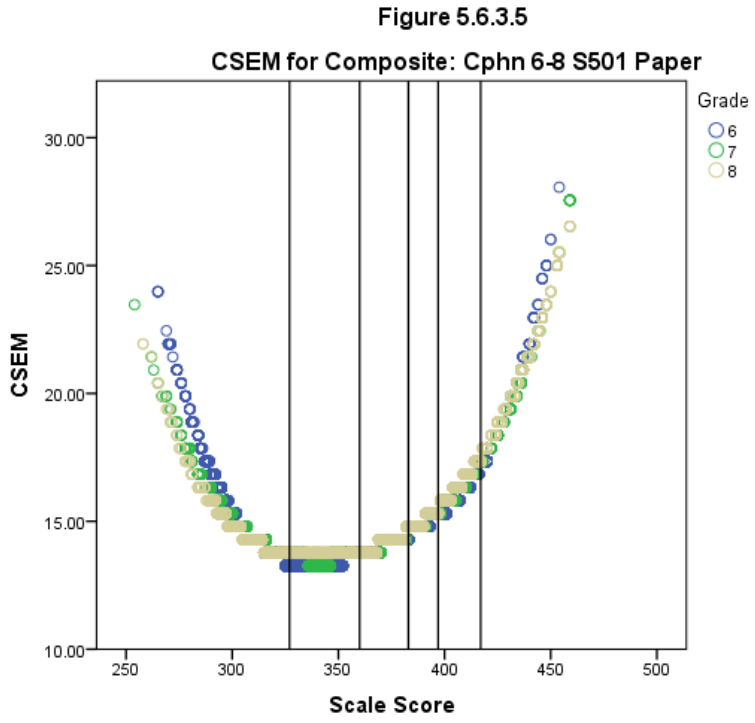
5.6.3.4 Grades 4–5

Figure 5.6.3.4



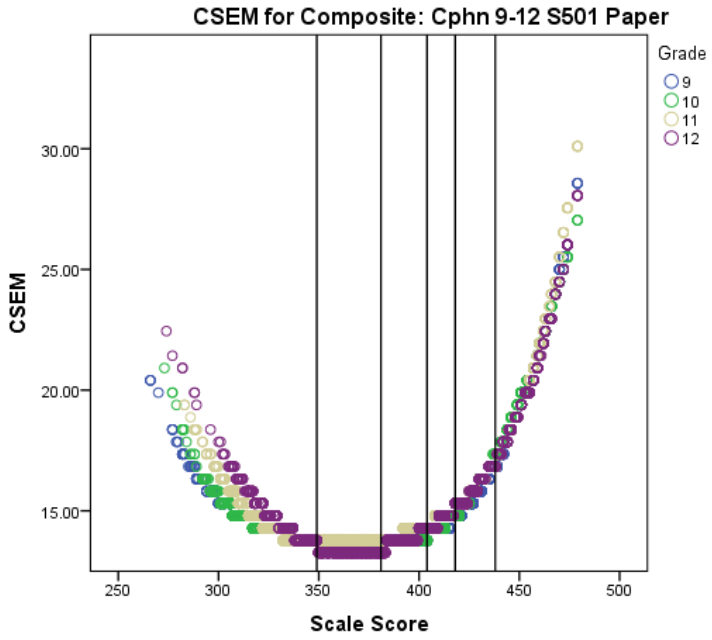
5.6.3.5 Grades 6–8

Figure 5.6.3.5



5.6.3.6 Grades 9–12

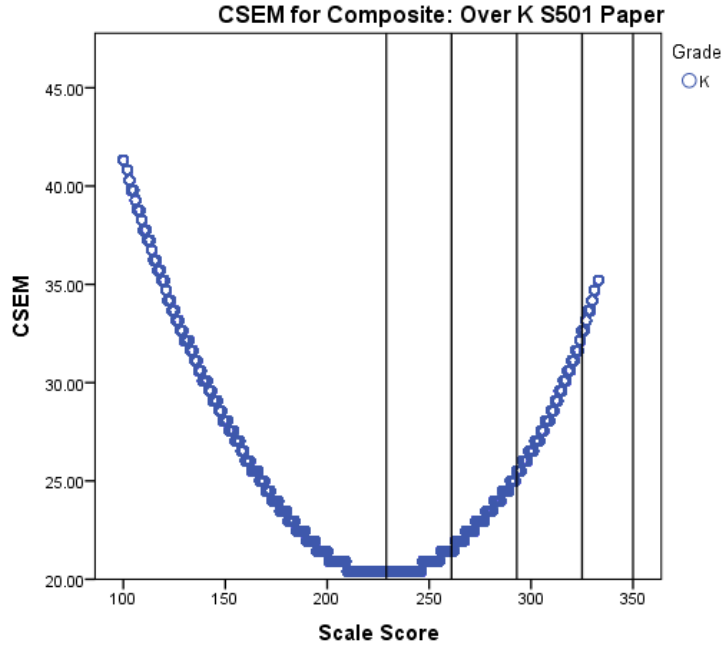
Figure 5.6.3.6



5.6.4 Overall

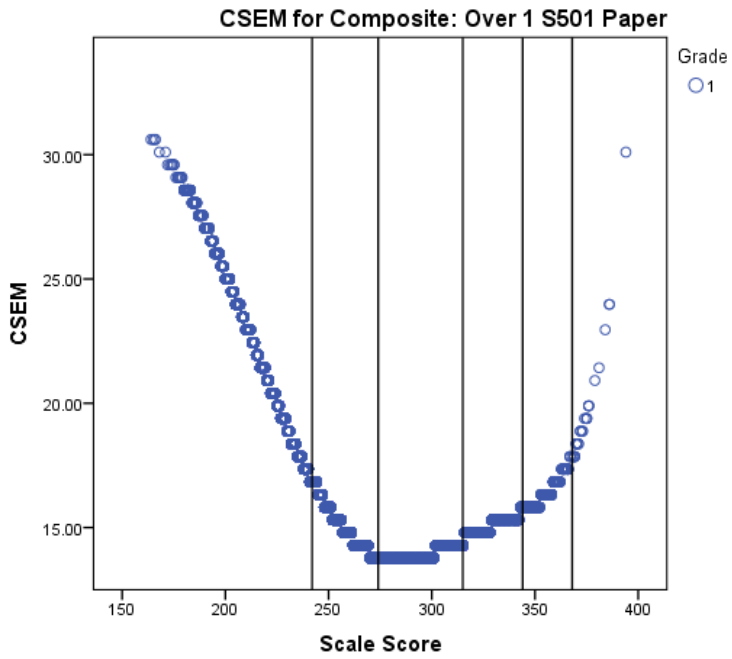
5.6.4.0 Kindergarten

Figure 5.6.4.0



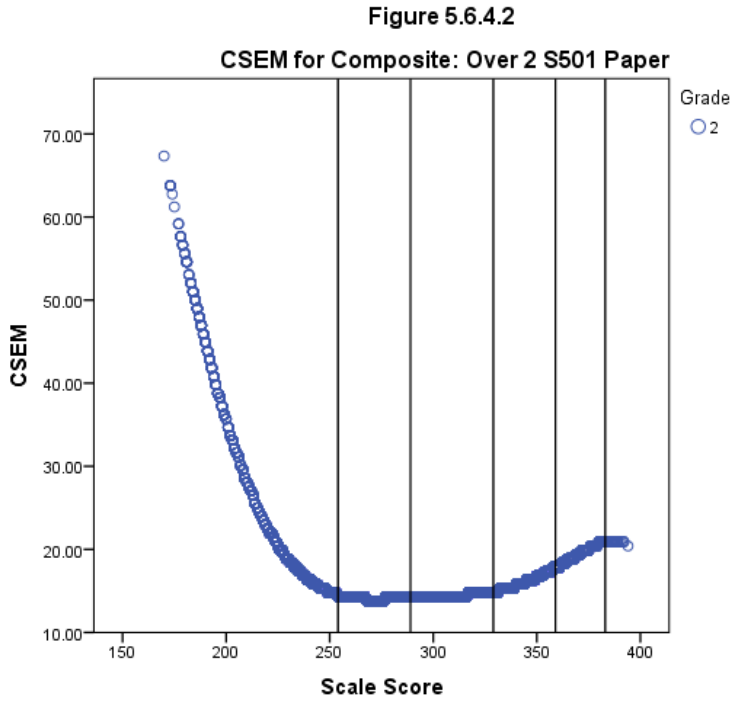
5.6.4.1 Grade 1

Figure 5.6.4.1



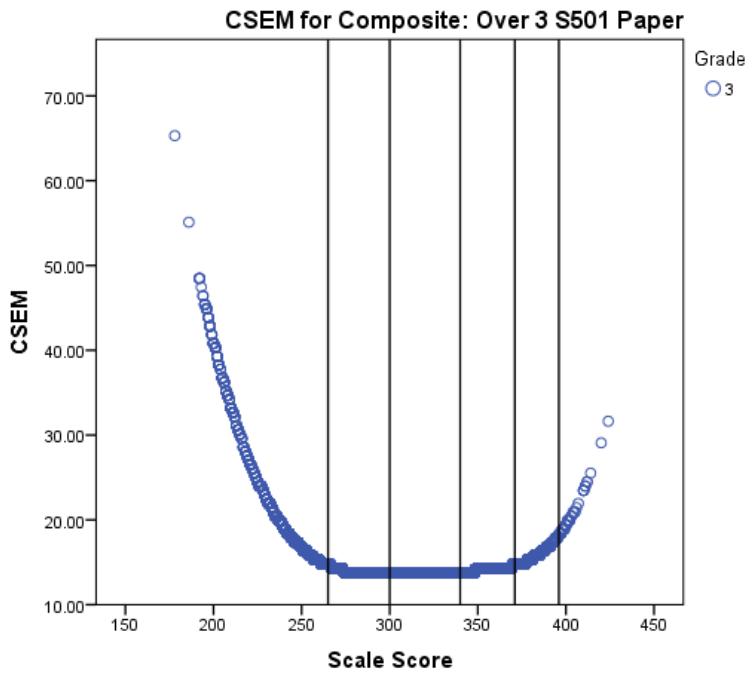
5.6.4.2 Grade 2

Figure 5.6.4.2



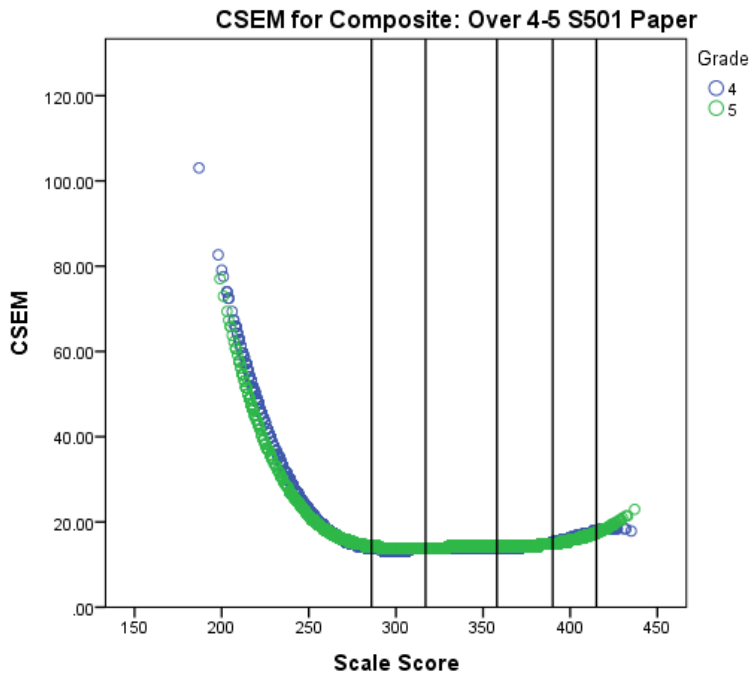
5.6.4.3 Grade 3

Figure 5.6.4.3



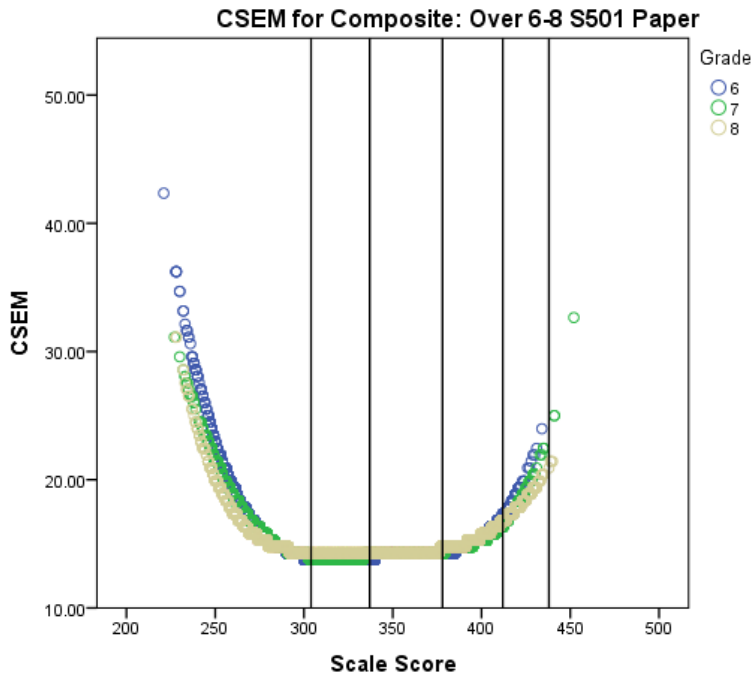
5.6.4.4 Grades 4–5

Figure 5.6.4.4



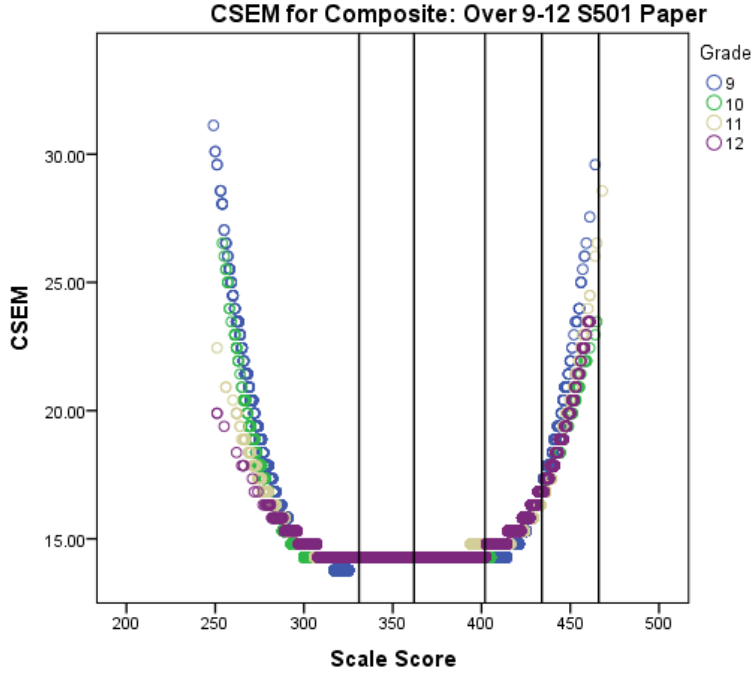
5.6.4.5 Grades 6–8

Figure 5.6.4.5



5.6.4.6 Grades 9–12

Figure 5.6.4.6



5.7 Accuracy and Consistency of Composites

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 composite scores can classify students into WIDA proficiency categories determined by the 2016 ACCESS standard setting process (Cook & MacGregor, 2017). Although states in the WIDA Consortium incorporate one or more of the domains and composite scores in making accountability decisions, all WIDA Consortium states use the **Overall composite** as the primary score in making classification decisions about students. Therefore it is especially important to examine the accuracy and consistency of the classifications based on the Overall composite to help test users and policy makers to judge the utility of this information and to make decisions about score reporting (American Educational Research Association et al., 2014). 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 et al., 2002).

The method and descriptions of the classification accuracy and consistency indices reported in this section appear in detail in Section 5.4. The only substantive methodological difference between the estimation of classification accuracy and consistency of the domains versus composites is that in order to estimate classification accuracy and consistency of the composites, we first estimated the reliability of the composite scores using a stratified Cronbach's alpha coefficient, as described in Section 5.4.

For each test domain, we present three tables. The first provides the overall accuracy and the overall consistency for each grade level. The second provides the classification accuracy at the cut score for each grade level. The third provides the classification consistency at the cut score for each grade level.

If the overall and marginal classification accuracy and consistency indices cannot be estimated because there are fewer than 200 students in the proficiency level, we collapsed the affected proficiency level category with the category below it and placed 'N/A' in the table for the affected proficiency level.

As noted in Section 5.4, there has been very little guidance for the ideal or expected levels of decision consistency and accuracy needed for educational assessments. We summarize the range of overall classification accuracy and consistency of domains across grades, by composite, and highlight the grade level with the lowest classification accuracy and consistency for test users and policy makers.

Since overall accuracy and consistency statistics are a summary of the degree of classification accuracy and consistency across all proficiency level cut points, the marginal classification accuracy and consistency for these grades were further examined to identify the specific source(s) of low classification accuracy and consistency.

For the Oral composite, as shown in Table 5.7.1.1, overall classification accuracy ranged from 0.609 to 0.710 and overall classification consistency ranged from 0.504 to 0.624 across grades. The lowest overall classification accuracy and consistency values were found for students in Grade 7 and Grade 8.

For the Literacy composite, overall classification accuracy ranged from 0.759 to 0.861 and overall classification consistency ranged from 0.670 to 0.826 across grades, as shown in Table 5.7.2.1. The lowest overall classification accuracy and consistency values were found for students in Grade 3 for classification accuracy and Grade 9 for classification consistency.

For the Comprehension composite, as shown in Table 5.7.3.1, overall classification accuracy ranged from 0.518 to 0.824 and overall classification consistency ranged from 0.408 to 0.774 across grades. The lowest overall classification accuracy and consistency values were found for students in Grade 3.

For the Overall composite, as shown in Table 5.7.4.1, overall classification accuracy ranged from 0.787 to 0.858 and overall classification consistency ranged from 0.711 to 0.808 across grades. The lowest overall classification accuracy and consistency values were found for students in Grade 5.

The results suggest that the grade level with the lowest classification accuracy and the lowest classification consistency tends to vary across these two indices and across the four composites.

The range of the marginal classification accuracy and consistency of composites are summarized and compared across grades by domains. In addition, the grade level with the lowest marginal classification accuracy and consistency of the composites is highlighted so that the test users and policy makers can use caution when making classification decisions in these grades at the specific cuts in the composites.

For the Oral composite, classification accuracy indices at the cut ranged from 0.868 to 0.978 (Table 5.7.1.2) and classification consistency at the cut ranged from 0.814 to 0.970 (Table 5.7.1.3). The lowest marginal classification accuracy and consistency values were found for students in Grade 5 at the PL 4/PL 5 cut. Additionally, Grade 5 was identified as having the lowest overall classification accuracy and consistency for the Overall composite. The low marginal classification accuracy and consistency at the PL 4/PL 5 cut appeared to have contributed to its low overall classification accuracy and consistency. However, it should be noted that the marginal classification accuracy and consistency for Grade 5 Oral composite are still in the 80's.

For the Literacy composite, classification accuracy indices at the cut ranged from 0.853 to 0.978 (Table 5.7.2.2) and classification consistency at the cut range from 0.803 to 0.986 (Table 5.7.2.3). The lowest marginal classification accuracy and consistency values were found for students in Grade 3 at the PL 3/PL 4 cut. Note that Grade 3 was also identified as having the lowest overall classification accuracy and second lowest overall classification consistency in the

Literacy composite. The low marginal classification accuracy and consistency at the PL 3/PL 4 cut appeared to have contributed to its low overall classification accuracy and consistency. However, it should be noted that the marginal classification accuracy and consistency for Grade 3 Literacy composite are still in the 80's.

For the Comprehension composite, classification accuracy indices at the cut ranged from 0.820 to 0.983 (Table 5.7.3.2) and classification consistency at the cut ranged from 0.760 to 0.975 (Table 5.7.3.3). The lowest marginal classification accuracy and consistency values were found for students in Grade 3 at the PL 4/PL 5 cut. Note that Grade 3 was also identified as having the lowest overall classification accuracy and consistency in the Comprehension composite. The low marginal classification accuracy and consistency at the PL 4/PL 5 cut appeared to have contributed to its low overall classification accuracy and consistency. However, it should be noted that the marginal classification accuracy and consistency for Grade 3 Comprehension composite are still in the high 70's and low 80's.

For the Overall composite, classification accuracy indices at the cut ranged from 0.895 to 0.985 (Table 5.7.4.2) and classification consistency at the cut ranged from 0.853 to 0.985 (Table 5.7.4.3). The lowest marginal classification accuracy and consistency values were found for students in Grade 3 at the PL 3/PL 4 cut. Note that Grade 3 was also identified as having the lowest overall classification accuracy and consistency in the Comprehension composite. Additionally, Grade 3 was also identified as having the lowest overall classification accuracy and second lowest overall classification consistency in the Literacy composite. The low marginal classification accuracy and consistency at the PL 3/PL 4 cut appeared to have contributed to its low overall classification accuracy and consistency. However, it should be noted that the marginal classification accuracy and consistency for Grade 3 Overall composite are still in the 80's.

Grade 3 had the lowest marginal classification accuracy and consistency in three of the four composites (Literacy, Comprehension, and Overall). Grade 5 had the lowest overall and marginal classification accuracy and consistency in the Oral composite.

In addition, the lowest marginal classification accuracy and consistency of the composites occurred at the PL 3/PL 4 and PL 4/PL 5 cut points. This finding is consistent with previous research (Lee et al., 2000), in that classification accuracy and consistency at cut points in the middle of the proficiency level range are lower than those at the lower and upper ends.

A higher number of proficiency levels typically results in cut scores that are closer to each other than if a smaller number of proficiency levels are used. Classification accuracy and consistency are expected to vary for different ability levels due to variation in measurement accuracy. The further away the scores are from the cut scores, the smaller the classification errors would be or the more accurate the classification decisions would be. When there is a large number of proficiency levels, more students are near the cut scores than there would be if there were fewer proficiency levels. Therefore, the higher the number of proficiency levels, the higher the probability that students would be misclassified (Ercikan & Julian, 2002). Since ACCESS has

six proficiency levels and PL 3 and PL 4 occupy relatively narrow ranges on the ability scale compared with other proficiency levels, the classification accuracy and consistency for the 3/4 and 4/5 cuts are lower than for other cuts.

There has been very little guidance for the ideal or expected levels of decision consistency and accuracy needed for educational assessments that use composite scores. From an accountability perspective, the most important information for test users and policy makers to examine is the marginal classification accuracy and consistency. The marginal classification accuracy and consistency indices were at or above 0.800 for all composites except for the Comprehension composite. The lowest marginal classification consistency for the Comprehension composite was 0.760 for Grade 3. Additionally, the marginal classification accuracy and consistency indices were at or above 0.853 for the Overall composite, where the major accountability decisions are being made.

5.7.1 Oral

Table 5.7.1.1

Overall Accuracy and Consistency of Classification Indices: Oral S501 Paper

Grade	Accuracy	Consistency
K	0.710	0.624
1	0.674	0.563
2	0.668	0.557
3	0.658	0.547
4	0.634	0.526
5	0.626	0.514
6	0.616	0.511
7	0.611	0.504
8	0.609	0.504
9	0.650	0.546
10	0.664	0.558
11	0.670	0.565
12	0.695	0.591

Table 5.7.1.2

Classification Accuracy Indices at Cut Score Level: Oral S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.946	0.939	0.943	0.936	0.934
1	0.971	0.926	0.886	0.923	0.966
2	0.976	0.939	0.887	0.901	0.962
3	0.978	0.944	0.885	0.892	0.956
4	0.975	0.954	0.903	0.884	0.914
5	0.972	0.951	0.904	0.868	0.926
6	0.970	0.953	0.910	0.872	0.905
7	0.961	0.947	0.906	0.874	0.914
8	0.957	0.943	0.905	0.876	0.917
9	0.944	0.923	0.900	0.918	0.955
10	0.946	0.923	0.899	0.926	0.964
11	0.946	0.921	0.900	0.932	0.966
12	0.946	0.919	0.904	0.947	0.975

Table 5.7.1.3

Classification Consistency Indices at Cut Score Level: Oral S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.924	0.915	0.917	0.911	0.908
1	0.958	0.894	0.842	0.888	0.960
2	0.966	0.911	0.844	0.856	0.957
3	0.970	0.918	0.840	0.843	0.949
4	0.966	0.932	0.865	0.837	0.894
5	0.961	0.928	0.867	0.814	0.901
6	0.958	0.932	0.873	0.821	0.881
7	0.946	0.922	0.869	0.823	0.886
8	0.940	0.916	0.867	0.826	0.892
9	0.922	0.890	0.861	0.883	0.943
10	0.923	0.889	0.859	0.894	0.954
11	0.924	0.888	0.860	0.901	0.958
12	0.923	0.884	0.865	0.923	0.970

5.7.2 Literacy

Table 5.7.2.1

Overall Accuracy and Consistency of Classification Indices: Litr S501 Paper

Grade	Accuracy	Consistency
K	0.861	0.826
1	0.799	0.718
2	0.801	0.726
3	0.759	0.677
4	0.769	0.689
5	0.766	0.680
6	0.804	0.726
7	0.796	0.715
8	0.791	0.707
9	0.763	0.670
10	0.769	0.679
11	0.770	0.681
12	0.790	0.706

Table 5.7.2.2

Classification Accuracy Indices at Cut Score Level: Litr S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.950	0.956	0.954	N/A	N/A
1	0.932	0.902	0.966	N/A	N/A
2	0.964	0.931	0.925	0.980	N/A
3	0.976	0.943	0.853	0.987	N/A
4	0.978	0.957	0.872	0.961	N/A
5	0.977	0.958	0.885	0.946	N/A
6	0.970	0.942	0.893	N/A	N/A
7	0.966	0.937	0.893	N/A	N/A
8	0.961	0.931	0.899	N/A	N/A
9	0.962	0.935	0.914	0.952	N/A
10	0.967	0.938	0.912	0.953	N/A
11	0.969	0.937	0.913	0.952	N/A
12	0.960	0.924	0.929	0.978	N/A

Table 5.7.2.3

Classification Consistency Indices at Cut Score Level: Litr S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.930	0.936	0.952	N/A	N/A
1	0.904	0.862	0.951	N/A	N/A
2	0.949	0.902	0.894	0.980	N/A
3	0.966	0.918	0.803	0.986	N/A
4	0.969	0.937	0.822	0.955	N/A
5	0.969	0.938	0.838	0.930	N/A
6	0.959	0.916	0.850	N/A	N/A
7	0.954	0.909	0.851	N/A	N/A
8	0.946	0.902	0.858	N/A	N/A
9	0.946	0.908	0.880	0.933	N/A
10	0.954	0.911	0.876	0.935	N/A
11	0.956	0.911	0.877	0.935	N/A
12	0.944	0.892	0.899	0.970	N/A

5.7.3 Comprehension

Table 5.7.3.1

Overall Accuracy and Consistency of Classification Indices: Cphn S501 Paper

Grade	Accuracy	Consistency
K	0.824	0.774
1	0.554	0.443
2	0.594	0.487
3	0.518	0.408
4	0.555	0.445
5	0.540	0.434
6	0.559	0.453
7	0.553	0.446
8	0.549	0.442
9	0.586	0.479
10	0.589	0.480
11	0.592	0.483
12	0.624	0.510

Table 5.7.3.2

Classification Accuracy Indices at Cut Score Level: Cphn S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.951	0.956	0.958	0.961	0.983
1	0.941	0.886	0.855	0.883	0.948
2	0.961	0.912	0.878	0.888	0.922
3	0.974	0.936	0.852	0.820	0.883
4	0.972	0.946	0.877	0.845	0.877
5	0.969	0.942	0.877	0.849	0.856
6	0.963	0.913	0.857	0.866	0.927
7	0.949	0.900	0.861	0.874	0.930
8	0.942	0.897	0.861	0.872	0.932
9	0.944	0.892	0.876	0.896	0.944
10	0.948	0.895	0.876	0.897	0.942
11	0.948	0.889	0.879	0.900	0.945
12	0.938	0.878	0.897	0.930	0.963

Table 5.7.3.3

Classification Consistency Indices at Cut Score Level: Cphn S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.930	0.938	0.942	0.946	0.975
1	0.918	0.837	0.804	0.836	0.922
2	0.945	0.875	0.832	0.842	0.890
3	0.966	0.902	0.795	0.760	0.832
4	0.964	0.917	0.829	0.791	0.826
5	0.959	0.911	0.831	0.793	0.804
6	0.950	0.873	0.806	0.812	0.902
7	0.930	0.855	0.810	0.825	0.899
8	0.919	0.852	0.811	0.823	0.900
9	0.922	0.847	0.830	0.854	0.918
10	0.927	0.851	0.828	0.855	0.916
11	0.926	0.844	0.832	0.859	0.920
12	0.910	0.830	0.855	0.898	0.949

5.7.4 Overall

Table 5.7.4.1

Overall Accuracy and Consistency of Classification Indices: Over S501 Paper

Grade	Accuracy	Consistency
K	0.858	0.808
1	0.825	0.756
2	0.818	0.757
3	0.808	0.741
4	0.799	0.728
5	0.787	0.711
6	0.832	0.770
7	0.820	0.750
8	0.816	0.746
9	0.797	0.716
10	0.803	0.728
11	0.811	0.736
12	0.829	0.760

Table 5.7.4.2

Classification Accuracy Indices at Cut Score Level: Over S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.953	0.956	0.966	0.984	N/A
1	0.965	0.927	0.948	0.985	N/A
2	0.980	0.952	0.924	0.963	N/A
3	0.985	0.960	0.895	0.969	N/A
4	0.985	0.971	0.925	0.919	N/A
5	0.983	0.970	0.927	0.907	N/A
6	0.981	0.964	0.922	0.966	N/A
7	0.976	0.960	0.920	0.964	N/A
8	0.972	0.956	0.923	0.965	N/A
9	0.969	0.952	0.931	0.946	N/A
10	0.972	0.953	0.930	0.949	N/A
11	0.973	0.952	0.931	0.956	N/A
12	0.970	0.943	0.939	0.978	N/A

Table 5.7.4.3

Classification Consistency Indices at Cut Score Level: Over S501 Paper

Grade	PL 1/2	PL 2/3	PL 3/4	PL 4/5	PL 5/6
K	0.933	0.938	0.951	0.984	N/A
1	0.951	0.897	0.925	0.985	N/A
2	0.971	0.931	0.893	0.962	N/A
3	0.979	0.941	0.853	0.967	N/A
4	0.979	0.957	0.894	0.898	N/A
5	0.977	0.957	0.897	0.879	N/A
6	0.973	0.948	0.890	0.959	N/A
7	0.967	0.942	0.887	0.954	N/A
8	0.962	0.937	0.891	0.956	N/A
9	0.957	0.932	0.902	0.925	N/A
10	0.961	0.932	0.902	0.933	N/A
11	0.962	0.931	0.903	0.941	N/A
12	0.957	0.919	0.913	0.970	N/A

6 Quality Control

6.1. Content Development Quality Control

The Center for Applied Linguistics (CAL) utilizes educators and other consultants at a number of phases throughout the test-development cycle. These educators and consultants are recruited, vetted, and trained by CAL and/or WIDA and make crucial contributions to these phases of the test development cycle. The phases of development in which educators or consultants are involved, as well as the procedures and criteria for recruitment and training, are described below.

Theme Generation

During theme generation, CAL and WIDA recruit educators to generate raw ideas to be used in new item development. Educators with ESL or content-area expertise and two or more years of teaching experience in a WIDA state (in the grade cluster for which they will generate themes) are invited to participate. Recruitment also focuses on a geographical distribution of educators from across the consortium. Upon selection, educators participate in a short training that introduces the theme-generation process, along with how to understand the item specifications that they use to generate themes.

Item Writing

CAL recruits professional item writers to generate raw item/task content based on the ideas from theme generation. To recruit item writers, CAL has a standing announcement on its website asking prospective item writers to submit their resume and fill out a survey describing their past item-writing experience. CAL selects individuals with significant experience in writing items, both in large-scale assessment programs (ESL/EFL or ELA) and in other contexts (e.g., writing items for assessment-programs in university-based ESL programs).

Item writers undergo a 90-minute orientation prior to beginning item writing. This training focuses on the item specifications, the process and procedures, the item writing checklist, the acceptance criteria for the items, and the security protocols. Item writers also receive an item writing handbook, which formalizes the content of the orientation, along with assignment of themes to develop and the associated item specifications. After the orientation, CAL Language Testing Specialists and managers provide feedback to the item writers on the items, focusing on alignment with the item writing checklist and the item specifications. After completion of item writing for a given development cycle, item writers are evaluated by CAL staff for their compliance with the requirements and the quality of their items.

Standards Expert Review

After items have been drafted by item writers, CAL Language Testing Specialists review all of the raw content internally. This review focuses on determining which sets of items will move on to further development and which will be discontinued, based on criteria from an item review checklist. The Language Testing Specialists then do minor editing and formatting to the items to make sure that they are complete, with no stray comments or other editorial notes from previous drafts, and they produce a short questionnaire for each set of items that becomes part of Standards Expert review. The purpose of Standards Expert review is to ensure that the items are appropriate for the grade-level and intended difficulty level in terms of both the content and the language, and the items have not drifted from their intended target between theme generation and item writing. The questionnaires produced by CAL's Language Testing Specialists guide the Standards Experts through the review process, asking questions specific to the purpose of this review.

Educators are recruited jointly by CAL and WIDA to serve as Standards Experts; educators with ESL or content-area expertise and two or more years of teaching experience in a WIDA state are invited to participate. Recruitment also focuses on a geographical distribution of educators from across the consortium. Standards Experts receive written instructions and a questionnaire to complete for each set of items they review.

Bias and Sensitivity and Content Review

After Standards Expert Review has been completed, all items undergo an additional phase of review and revision internal to CAL, leading up to Bias & Sensitivity and Content Review. These are technically two separate reviews, although a single recruitment effort is conducted by WIDA, and the reviews occur consecutively in a single week (generally 3 days for Content review followed by 2 days for Bias & Sensitivity review). As with other reviews, educators for Content review must have at least 2 years of ESL teaching experience (with a preference for content-area experience as well). Recruitment also focuses on selecting educators with a variety of cultural and linguistic backgrounds and obtaining a geographical distribution of educators from across the consortium. Recruitment for Bias & Sensitivity review focuses on selecting 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.

At the beginning of both Bias & Sensitivity and Content review meetings, CAL and WIDA staff conduct an intensive training to orient the reviewers to the specific purpose of the review (Bias & Sensitivity or Content), how to use the review checklist and what to look for in the review, and the procedures and security protocols for the review. Then, the reviews are conducted in breakout groups by grade cluster (or combinations of grade clusters; for example, Bias & Sensitivity review of Grade 1 and Grades 2–3 is often combined). Although Bias & Sensitivity

and Content reviews are generally held in -person, the reviews for the Writing domain occur virtually each year due to timeline constraints. For both the in-person and virtual contexts, CAL and WIDA facilitators are present in each breakout group to guide the educators in their reviews of the materials.

Writing Tryouts

For the Writing domain, all tasks in the Writing domain are subject to tryouts in the field. The Writing tryouts only occur once the tasks have been through a thorough Bias & Sensitivity and Content review and subsequent revision. CAL and WIDA recruit educators who are willing to administer the Writing tasks to their students; these educators are classroom ESL or content teachers who work with ELLs. All students who participate are required to have parent/guardian consent.

Once the students complete the Writing tasks, both the students and educators fill out questionnaires. Student questionnaires focus on whether the students understood the task, their engagement with the task, and their ability to complete the task; educator surveys ask the teachers to evaluate the effectiveness of the task input, the appropriateness of the task, the comparability of the task with other classroom-based writing tasks, and the ability of the students to complete the task.

CAL provides the teachers with a number of documents outlining the procedures for administering the tasks, recording student responses to the tasks, recording student and teacher responses to the questionnaires, and protecting the personally identifiable information of the students. CAL staff are also available throughout the tryouts process to answer any questions the teachers might have. Following the Writing tryouts, CAL specialists review the writing responses both qualitatively and quantitatively, providing WIDA with a report on how the Writing tasks performed.

6.2. Test Administration Quality Control

This section describes how WIDA monitors test administration to ensure standardized test administration procedures are implemented with fidelity across districts and schools. To support standardized administrations, WIDA provides test administrators with a series of resources, such as a Test Administration Manual, a training course, and a Test Administration Script for each assessment.

Qualifications of Test Administrators

Before, during, and after a state's testing window, educators hold various roles to ensure all tasks are carried out for successful test administration. These roles include Test Coordinators at the

district and school level, Test Administrators. The Test Administrator administers and monitors the test. He or she is also responsible for managing student data prior to, during, and after testing.

WIDA has worked directly with each state education agency to develop the ACCESS for ELLs Checklist for the school year. This list highlights all tasks that need to be completed before, during, and after testing within a school or district and outlines which tasks are assigned to Test Coordinators at the district and school level and Test Administrators. It also provides additional guidance that a state expects test administrators to follow as they prepare for and administer the ACCESS for ELLs suite of assessments.

Test administrators are responsible for reviewing each state's checklist in detail prior to completing any training and for working with the district or school Test Coordinator to complete these tasks. The state's checklist can be found in the training course and on each state's WIDA webpage at www.wida.us/membership/states.

The training course within the WIDA Secure Portal (https://www.wida.us/login.aspx) is where educators can access both training to become certified to administer ACCESS for ELLs as well as additional materials and resources to assist administrators and coordinators before, during, and after each state's testing window. WIDA user accounts provide access to the training course and Facilitator Toolkit within the WIDA Secure Portal. Educators must pass an administration quiz at the end of the training with a score of 80% or higher. WIDA recommends taking the quiz immediately after completing the training. There is no limit to the number of times educators can attempt the quiz. Once individuals pass an administration quiz, training certificates within the WIDA Secure Portal are updated to reflect their status as a certified test administrator for that component of the assessment suite.

Paper Testing (for Writing Grades 1–3)

Depending on state, district, and school policy, not all Test Administrators will be responsible for initially labeling and/or bubbling booklets. However, it is the responsibility of all Test Administrators and Test Coordinators to ensure that correct and complete information is either labeled or bubbled in each student booklet. Each state's ACCESS for ELLs checklist has more information on who is responsible for each task related to materials management in the state.

To ensure all booklets have the detailed and necessary information needed to score, all Test Administrators must adhere to the following:

- Prior to administration
    - Review labels and/or bubbled information to ensure all student information is accurate.
    - Complete labeling or bubbling if needed.

- During administration
    - Distribute the test booklets, as applicable, to the correct students.
    - Verify that students have been given their assigned booklet.
- Immediately following administration
    - Collect all material from all students.
    - Review student test booklets once more for any errors or discrepancies in student information.
    - Confirm all necessary fields are completed and all necessary labels are correctly adhered to student test booklets.
    - Ensure all booklets are in proper condition to be returned, with no loose or damaged pages.
    - Return test materials to a Test Coordinator, or store the booklets in a secure area until they can be handed over to a Test Coordinator.

Failure to address incorrect, missing, or incomplete booklet information and labels may result in late reporting or no student score. In addition, the WIDA Consortium's national research agenda relies on complete and accurate student demographic data to inform the field and benefit English language learners.

When preparing test materials for return to DRC, test administrators need to confirm that any booklet that contains student response information has either a Pre-ID Label or a District/School Label with bubbled student information. If a booklet is unused, there is no need to place any labels on the booklet. Placing a label on a booklet will cause it to be processed (and either scored, if the label is a Pre-ID or School/District label, or not scored, if it is a Do Not Process label).

6.3. Rater Quality Control

Rater Training

Students who take the ACCESS for ELLs Paper Speaking test have their spoken responses scored by the Test Administrator who administered the Speaking test. Another term for this Test Administrator is *rater*. Raters must be trained and certified, so we can be confident that they interpret students' spoken language consistently and fairly, and that the scores are reported according to the WIDA English language proficiency standards. WIDA provides several different types of resources to support raters' training and reliability.

Students who take ACCESS for ELLs Online have their spoken responses digitally recorded and then scored centrally by DRC's trained raters. Students who take ACCESS for ELLs Paper have their spoken responses scored in real time by the Test Administrator who administers the

Speaking test. In both cases, it is important that the individual who scores the spoken responses is trained and certified.

WIDA provides a series of training modules in the Secure Portal on the WIDA website. ACCESS for ELLs Speaking test raters should complete three core modules:

1. Overview and Test Structure
2. Speaking Assessment Scoring Practice
3. Speaking Assessment Recommended Practice

WIDA strongly recommends that all new raters complete all three of these modules. These modules provide a comprehensive introduction to the ACCESS for ELLs Speaking test and the opportunity to learn how to score students' spoken English reliably using the ACCESS for ELLs 2.0 Speaking Scoring Scale.

In addition to the modules described above, WIDA also releases supplemental training materials each year to refamiliarize experienced raters with the Speaking Scoring Scale and introduce new Speaking tasks and sample responses for the coming year. These materials, called Supplemental Training for the Speaking Assessment, reflect the Speaking tasks that will appear on the test in the current year. WIDA recommends that all raters (new and experienced) engage with these supplementary materials at the start of each scoring season. Reading and reviewing these materials will help raters maintain their reliability from year to year and contribute to the fairness of test scores awarded to all students.

Rater Certification

After completing the training modules described in the section above, new raters should take the relevant certification quiz. WIDA provides two quizzes: one for raters who will evaluate students in Grades 1–5 and another for raters who will evaluate students in Grades 6–12. Raters should take the appropriate quiz.

The purpose of the quiz is to ensure that raters have internalized the Speaking Scoring Scale and can apply it consistently. Only raters who pass the quiz(zes) should administer and score the ACCESS for ELLs 2.0 Paper Speaking test.

Checklist for Rater Training, Monitoring, and Recertification

- ✓ New raters complete all Speaking assessment training
- ✓ New raters take and pass the appropriate certification quizzes
- ✓ All raters recertify at the start of each testing season (review new materials, retake quiz)
- ✓ Only certified raters administer and score the ACCESS for ELLs 2.0 Speaking test

- ✓ Raters do not evaluate their own students, if at all possible
- ✓ Rater reliability and/or score point distributions are monitored regularly

6.4. Score Reporting Quality Control

WIDA conducts an annual score reporting quality control process to (1) verify the accuracy of paper-based test scores (i.e., ACCESS for ELLs Paper, Kindergarten ACCESS for ELLs, and Alternate ACCESS) and (2) verify the accuracy of all score reports (the Individual Student Report, the Student Roster Report, the School Frequency Report, the District Frequency Report, and the State Frequency Report) for both ACCESS (Online, Paper, and Kindergarten) and Alternate ACCESS.

The Score Reporting quality control is conducted at DRC's offices in Maple Grove, Minnesota. The team generally includes five state education agency representatives, one CAL employee, and four WIDA employees.[3] This team examines data from three districts: a primary district, for quality control of all score reports; a secondary district, for quality control of State Frequency Reports only; and a tertiary district for quality control of paper-based tests only.

After an introductory presentation, which includes details of the quality control processes undertaken by DRC and WIDA and instructions on using the data entry tools, panelists begin by confirming the scoring of ACCESS Paper. Using the information in the State Student Response file, panelists enter the grade level, grade level cluster, tier, the Listening and Reading responses, and the Speaking and Writing scores into the data entry tool. The tool then calculates the student's raw scores and, using a series of look-ups, the student's scale score, proficiency level score, and confidence bands for all domains and composites. Panelists check student scores on the Individual Student Reports against those calculations. Any discrepancies are brought to the attention of the WIDA facilitator who investigates and, if there seems to be an issue with the report (rather than the data entry or data entry tool), discusses the issue further with DRC.

The panelists follow a similar process with the Kindergarten ACCESS tests, but with the raw scores for these tests copied directly from the response booklets.

After checking the paper-based tests, panelists turn their attention to the score reports. Panelists first check both the demographic information and the student scores in the Individual Student Reports against the information in the Student Roster Reports. Again, any discrepancies are brought to the attention of the facilitator, who investigates and discusses the issue with DRC if necessary. Panelists use the verified Individual Student Reports to check the Student Roster Report. Once the Student Roster Report is verified, panelists use it to check the State Frequency

[3] Due to the COVID-19 pandemic, the 2020 Score Reporting quality control was conducted online, with only WIDA and DRC employees participating.

Report; they then use the verified State Frequency Reports to check the District Frequency Report. Finally, panelists check the State Frequency Reports against verified District Frequency Reports from the primary district along with District Frequency Reports from the secondary district.

6.5. Data Forensic Quality Control

WIDA hired Caveon to perform data forensic analysis during the 2019–2020 test administration cycle to examine whether ACCESS data has been compromised or has evidence of item exposure.

Caveon security statistics are based on mathematical models, where the test response data are used to create a baseline model of normal or "typical" test taking among that population. Individuals or groups are then compared to the baseline, and observations that are significantly different from the baseline are flagged as anomalous. Caveon's statistics are designed to be robust but also conservative regarding which and how many individuals or groups are flagged as anomalous, thereby reducing the chances of false-positive detections.

Data forensics analysis was performed after the administration window for the following administrations:

- December 2019 through Spring 2020 online multistage adaptive test administrations, Listening and Reading domains
- December 2019 through Spring 2020 paper fixed-form administrations, Listening and Reading domains

The analysis utilized several of Caveon's security statistics to detect evidence of whether the assessment instrument has been compromised through disclosure of the content. This analysis attempted to understand where and when disclosure of the test content may have occurred and what items and forms may have been affected. Results of this analysis might enable WIDA to take specific actions to limit the impact of disclosed content. Such actions may include

- Republishing or reworking items or forms
- Rotating disclosed items to limit their exposure
- Designing a republication or rotation strategy for future items and forms

Caveon security statistics were computed for each individual test instance. These data were aggregated or summarized at the group level. The aggregated statistics were compared against the population model.

Analysis of Tests

Caveon aggregated the data according to individual test forms using the security statistics to determine whether rates of detections by the security statistics were higher for certain test forms. For fixed-form paper tests, two forms—A and B/C—were analyzed. For the multistage adaptive

test, there is a finite number of ways a student could progress through the test. Caveon analyzed each pathway as a separate form. Higher rates of security detections for a specific form of the test suggest that compromise of the form may have occurred.

Analysis of Items

Item security: In this portion of the analysis, the security of the items was evaluated using aberrance statistics. Aberrance statistics detect test-taking behaviors such as answering difficult items correctly but answering easy items incorrectly, or unusual patterns in the time taken to answer test items. In the absence of security issues, aberrant test taking is expected to be the result of poor or uneven test preparation, illness or other physical malady, mental and emotional distractions, and so forth. These factors usually result in lower levels of test performance. When aberrance is associated with higher performance, however, test fraud may have occurred, such as preknowledge of test content. By applying aberrance measures and comparing the performance between aberrant and nonaberrant test instances on individual items, inferences can be made about item security.

Item performance changes: Analysis of item performance changes tracks individual item performance rates over time. The item performance shifts are measured within the context of the item response theory model and adjusted for varying test-taker performance levels. This means that detected performance shifts are invariant to fluctuations in the test-taker population. When performance shifts indicate the item has become significantly easier, the item may have been disclosed. Items with significant performance shifts become candidates for revision or replacement. Item performance shifts were detected with a granularity of 1 week, where Monday to Sunday represents 1 week.

Analysis of Groups

Analysis by week: This analysis aggregates the data according to the week in which the test was taken to identify whether security threats and pass rates appeared to be more prevalent at certain times during the testing window. Increases in scores or security detections during certain periods of time suggest the content may have been disclosed at some point prior to that time. This analysis also includes a form-date grouping to determine if increasing security threats are associated with a particular form of the test. This analysis is performed for online and paper tests, where relevant test date data are provided.

Analysis of WIDA jurisdictions: Caveon analyzed WIDA member jurisdictions (states and districts) to determine whether rates of detections by the security statistics were higher for certain jurisdictions. This analysis is intended to detect whether compromise at the state or member jurisdiction level potentially occurred. This analysis is performed for online and paper tests.

Analysis of administration mode: Caveon aggregates the data according to administration mode (i.e., online versus paper) to determine if security threats are associated with the mode of testing.

Other Analyses

Analysis of mean score over time was used to identify whether mean scores increased over time during the testing window. Increases in scores over time suggest the content may have been disclosed during the testing window.

Findings of Data Forensic Analyses

Generally, no major data forensic anomalies were observed across WIDA states. There were some general findings and a few minor localized anomalies:

1. High rates of similar tests with associated score gains and a high rate of tests in large clusters suggest the presence of possible security violations in a district.
2. High rates of identical and/or perfect tests in two states suggest potential item compromise in these states.
3. For lower grades of the Reading Exam, examinees with better performance on old items than new items tended to have higher scores than those who did not exhibit a performance difference.
4. Paper-and-pencil exams had higher rates of identical and perfect tests than online exams. Within paper-and-pencil administrations, the Listening exam generally had higher rates of identical and perfect tests than the Reading exam.
5. Analysis of items suggested that some items may have been disclosed or become well known. This was especially prevalent among the online exams. However, if true, the disclosure appears to have occurred only among a low proportion of the examinees.
6. Analysis of test forms, test formats (i.e., administration mode), and test weeks did not find evidence of widespread item compromise or security violations. Mean scores were generally stable over the testing window.

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.

American Institutes of Research. (2018). *ELPA21 technical report, part I—summative assessment*. Washington DC: Author.

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

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.

Center for Applied Linguistics. (2018). *Annual technical report for ACCESS for ELLs® 2.0 Online English Language Proficiency Test, Series 401, 2016–2017 administration*. Washington, DC: Author.

Center for Applied Linguistics. (2019). *Annual technical report for ACCESS for ELLs® English Language Proficiency Test, Series 402 Paper, 2017–2018 administration* (WIDA Consortium Annual Technical Report No. 14B). 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.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Elementary and Secondary Education Act of 1965, amended 2015. 20 USC §6301-8961.

Ercikan, K, & Julian, M. (2002). Classification accuracy of assigning student performance to proficiency levels: Guidelines for assessment design. *Applied Measurement in Education, 15*(3), 269–294.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105–146). New York, NY: Macmillan.

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.

Kim, A., Ho, P., Chapman, M., & Cook, H. G. (2020a). *Examination of reclassification decisions made for K–12 English learners: Survey report of Delaware* (WIDA Internal Report). Madison, WI: WIDA at the Wisconsin Center for Education Research.

Kim, A., Ho, P., Chapman, M., & Cook, H. G. (2020b). *Examination of reclassification decisions made for K–12 English learners: Survey report of Pennsylvania* (WIDA Internal Report). Madison, WI: WIDA at the Wisconsin Center for Education Research.

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. (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. (2004). Optimizing rating scale category effectiveness. In E. V. Smith Jr. & R. M. Smith (Eds.), *Introduction to Rasch measurement* (pp. 258–278). Maple Grove, MN: JAM Press.

Linacre, J. M. (2006). Winsteps Rasch analysis (Version 3.60.1) [Computer software]. Retrieved from http://www.winsteps.com

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

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacGregor, D., Kenyon, D. M., Gibson, S., & Evans, E. (2009). *Development and field test of Kindergarten ACCESS for ELLs®*. Madison, WI: WIDA Consortium.

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/

Muraki, E. (1993). Information functions of the generalized partial credit model. *Applied Psychological Measurement, 17*(4), 151–363.

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.

U.S. Department of Education. (2018). *A state's guide to the U.S. Department of Education's assessment peer review process.* Retrieved from https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf?utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=

Waller, N. G. (n.d.). *EZDIF: A computer program for detecting uniform and nonuniform differential item functioning with the Mantel-Haenszel and logistic regression procedures* [Computer software]. Davis, CA: University of California – Davis.

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 & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ: Erlbaum.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30,* 233–251.

Acknowledgments

We would like to extend our appreciation to the many CAL and WIDA staff members who have supported this work.

From CAL:

Keira Ballantyne, Ph.D.
Tanya Bitterman, M.A.
Caitlin Gdowski, M.A.
Yage (Leah) Guo, Ph.D.
Michele Kawood, M.S.Ed.
Justin Kelly, Ph.D.
Dorry M. Kenyon, Ph.D.
Nicholas Luzio Jr., B.S.
Erin Shaw-Meadow, M.Sc.
Samantha Musser, M.A.
Rachel Myers, M.S.
Yoon Ah Song, Ph.D.
Alice Tsai, M.S.
Shu Jing Yen, Ph.D.
Xin Yu, M.A.

From WIDA:

Mohammad Akanda, Ph.D.
Kyoungwon Bishop, Ph.D.
Sakine Göçer Sahin, Ph.D.