World-Class Instructional Design and Assessment

# WIDA™
## CONSORTIUM

# Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 302, 2013-2014 Administration

**Annual Technical Report No. 10**
**Volume 1 of 3: Description, Validity, and Student Results**

Prepared by:

Center for Applied Linguistics

CAL/WIDA Partnership Activities
Psychometrics/Research Team

May 30, 2015

**The WIDA ACCESS for ELLs Technical Advisory Committee**

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

- Jamal Abedi, Ph.D., Professor at the Graduate School of Education at the University of California, 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
- Akihito Kamata, Ph.D., Professor, Department of Education Policy and Leadership, Department of Psychology, Southern Methodist University.
- Timothy Kurtz, Hanover High School, Hanover, New Hampshire
- Carol Myford, Ph.D., Associate Professor, Educational Psychology at the University of Illinois at Chicago.
- Elizabeth Peña, Ph.D., Professor, Department of Communication Sciences and Disorders, University of Texas at Austin.

More information on the TAC members can be found at the WIDA website (www.wida.us/assessment/access/TAC/index.aspx).

# Executive Summary

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

ACCESS for ELLs serves two purposes: 1) To assess reliably and validly the English language development (ELD) of English language learners (ELLs) in Grades K–12 according to WIDA *2012 Amplification of the English Language Development Standards Kindergarten–Grade 12* (WIDA,2012b) To place students appropriately into proficiency levels described by the ELD Standards. 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.

This report provides detailed information from the analysis of the tenth series of the test, Series 302. Series 302 was administered during the academic year 2013–2014 in 33 WIDA Consortium states. Because the main focus of this report is on the technical quality of the test forms and not on the performance of students, analyses in this report are aggregated across all participating states.

Beginning with ACCESS Series 302 (operational year 2013-2014), the ACCESS Listening test transitioned from a traditional test administrator-read script to a media-delivered format, played either from CD or from streaming audio available online, for all grade level clusters except for Kindergarten. For more information, please see the *ACCESS for ELLs Series 302 Media-Based Listening Field Test Technical Brief* (Center for Applied Linguistics, forthcoming).


As in the previous annual technical reports, this report provides background to the test (Chapter 1). The current report has been modified for Series 302 to introduce an argument-based validation framework to support the use of ACCESS for ELLs and to contextualize the data so that its interpretation and use are more transparent to stakeholders (Chapter 2). The rest of the report consists of paired chapters. The first chapter within each pair contains text that explains the data tables that follow in the second chapter. Information on the students who participated in the operational administration is presented (Chapters 3 and 4), followed by an explanation of the technical analyses conducted on each of the 44 test forms that constitute ACCESS for ELLs (Chapter 5) and the tables and figures of results (Chapter 6). The final chapters explain (Chapter 7) and present (Chapter 8) technical analyses based on the domain scores and composite scores by grade-level cluster. Note that Chapters 1–4 are in Volume 1, Chapters 5–6 are in Volume 2, and Chapters 7–8 are in Volume 3.

## *Summary Highlights*

This report presents a wealth of data documenting the technical properties of the 44 test forms of ACCESS for ELLs Series 302, which is impossible to summarize here. In addition to information on validity, the report presents information on reliability of test scores and the accuracy and consistency of proficiency level classifications, including information on

conditional standard errors of measurement for all scores and a separate table highlighting conditional standard errors around the cut scores. The report also provides details on scaling and the equating of the Series 302 test forms to those of Series 301. Item-level analyses include item difficulty levels, fit of the items to the Rasch measurement model, and differential item functioning (DIF) analyses for each item or assessment task. The annual analyses of the technical properties of ACCESS for ELLs test forms are used in the continual refinement and improvement of ACCESS for ELLs.

Here we would like to highlight the following results of this report.

### *Argument-based validation framework for ACCESS for ELLs*

Starting with Series 301, Chapter 2 of the ACCESS for ELLs Annual Technical Report consists of an argument-based framework for supporting the validity of ACCESS for ELLs. This framework structures the information contained in this Annual Technical Report to support assertions about data collected via the assessment (i.e., *Assessment Records*). Specifically, tables and figures from this report are explicitly linked to claims related to *Assessment Records* through an Assessment Use Argument (AUA), which allows stakeholders to better interpret and use ACCESS for ELLs. A larger, forthcoming (as of 2015), validation framework for the complete assessment from its inception to its consequences is currently under development by WIDA.

### *Demographic data*

The Series 302 data set for analyses included the results of 1,372,806 students. The largest grade was Kindergarten with 204,828 students, while the smallest was Grade 12 with 31,299 students. Of the participating WIDA states, the largest was Illinois with 176,389 students, while the smallest was Vermont with 1,533 students. Technical analyses in this report are based on the performance of all students who were administered Series 302 of ACCESS for ELLs.

### *Reliability and accuracy data*

For most test users, the Overall Composite proficiency score, based on performances in Listening, Speaking, Reading, and Writing, is the major score used for making decisions about gains in student proficiency and exiting from language support services, and for Annual Measureable Achievement Objectives (AMAOs). As explained by Keng, Miller, O'Malley, and Turhan (2008), "the use of composite scores has become more widespread with federal testing requirements under Title III of No Child Left Behind now calling for states to assess students with limited English proficiency (LEP) annually from Kindergarten through 12th grade in the four language domains of listening, speaking, reading and writing. A composite of the student's performance on each of these domains is calculated to represent the student's overall English language proficiency." Results indicate that the reliability of the Overall Composite score for Series 302, presented in Chapter 8 Table D, is very high across all grade-level clusters. For Kindergarten it was .973; for Grades 1–2, .943; for Grades 3–5, .937; for Grades 6–8, .930; and for Grades 9–12, .945. Likewise, as Table 0.1 shows, the accuracy of decisions about student placement using the Overall Composite score around the proficiency level cut scores is very high across the grade and proficiency levels. Because many WIDA Consortium states use the proficiency level score of 5.0 as a criterion for exiting students from language support services, the column headed 4/5 Cut (the proficiency level score of 5.0) is of particular interest.

Table 0.1

Accuracy of Overall Score at Cut Points (Proficiency Level Score)

| Grade | 1/2 Cut (2.0) | 2/3 Cut (3.0) | 3/4 Cut (4.0) | 4/5 Cut (5.0) | 5/6 Cut (6.0) |
|---|---|---|---|---|---|
| K (instructional) | 0.976 | 0.959 | 0.950 | 0.953 | 0.950 |
| K (accountability) | 0.953 | 0.951 | 0.955 | 0.953 | 0.991 |
| 1 | 0.985 | 0.938 | 0.930 | 0.970 | 0.991 |
| 2 | 0.985 | 0.956 | 0.911 | 0.936 | 0.985 |
| 3 | 0.996 | 0.982 | 0.938 | 0.910 | 0.930 |
| 4 | 0.994 | 0.979 | 0.931 | 0.900 | 0.918 |
| 5 | 0.992 | 0.974 | 0.925 | 0.897 | 0.931 |
| 6 | 0.988 | 0.962 | 0.908 | 0.889 | 0.983 |
| 7 | 0.985 | 0.958 | 0.903 | 0.886 | 0.984 |
| 8 | 0.981 | 0.951 | 0.900 | 0.908 | 0.985 |
| 9 | 0.980 | 0.960 | 0.931 | 0.906 | 0.901 |
| 10 | 0.984 | 0.956 | 0.925 | 0.904 | 0.942 |
| 11 | 0.985 | 0.956 | 0.923 | 0.896 | 0.941 |
| 12 | 0.984 | 0.954 | 0.915 | 0.868 | 0.956 |

*Overview of the Annual Technical Report*

The multistate WIDA Consortium's ACCESS for ELLs was first operationally administered in 2005 in three states: Alabama, Maine, and Vermont. Results of that administration were reported in Annual Technical Report 1 (Series 100, 2004–2005).

Each year, the Center for Applied Linguistics refreshes a third to a half of all ACCESS for ELLs test items. In academic year 2013–2014, Series 302 was administered in 33 WIDA Consortium states. The 44 test forms in Listening, Reading, Writing, and Speaking across the grades from Kindergarten to 12 constitute a test series. These test forms are grouped into five grade-level clusters: Kindergarten, Grades 1–2, Grades 3–5, Grades 6–8, and Grades 9–12. Within each grade-level cluster except Kindergarten, there are three overlapping tiers of test forms for Listening, Reading, and Writing: A, B, and C. This report presents the results of research into the technical properties of the 44 test forms (e.g., Grades 3–5, Reading, Tier C) that constitute Series 302. Data come from the 1,372,806 students who were administered the test operationally in the winter and spring of 2014.

Because of the size of the complete report, it is presented in three volumes.

*Volume I* contains an executive summary, this overview, an annotated bibliography, and Chapters 1 to 4. Chapter 1 provides background to the test. Readers unfamiliar with ACCESS for ELLs should pay particular attention to this chapter. Chapter 2 presents an argument-based approach for structuring the data contained in this report so that its interpretation and use are more transparent to stakeholders. Chapters 3 and 4 present information on the students who participated in the Series 302 (2013–2014) operational administration, including overall results.

*Volume II* contains Chapters 5 and 6. Chapter 5 presents background on the technical analyses conducted on each of the test forms and explains how to understand the tables and figures of results. Chapter 6 presents the results organized by

- Grade-level cluster (K, 1–2, 3–5, 6–8, 9–12)

- Domain (Listening, Reading, Writing, and Speaking, abbreviated List, Read, Writ, and Spek, respectively)

- Tier (A, B, C)

Thus, all of the results for Kindergarten are presented before the results for Grades 1–2, and all of the results for Listening (i.e., for tiers A, B, and C where applicable) are presented before results for Reading.

*Volume III* contains Chapters 7 and 8. These chapters focus on results across tiers within grade-level clusters, including the four composite scores (Oral Language, Literacy, Comprehension, and Overall). Chapter 7 presents background on the technical analyses and explains how to understand the tables and figures of results. Chapter 8 presents the results organized by

- Grade-level cluster (K, 1–2, 3–5, 6–8, 9–12)

- Score (Listening, Reading, Writing, Speaking, Oral Language Composite, Literacy Composite, Comprehension Composite, and Overall Composite, abbreviated List, Read, Writ, Spek, Oral, Litr, Cphn, and Over, respectively)

# Annotated Bibliography: 2013-2014

## *Technical Reports*

This is a list of reports that describe the development of ACCESS for ELLs.

Center for Applied Linguistics (forthcoming). *ACCESS for ELLs Series 302 Media-Based Listening Field Test Technical Brief.* (WIDA Consortium).

> This report (forthcoming) provides detailed information on the conceptualization, development, and field testing of ACCESS for ELLs Media-Based Listening Test.

Gottlieb, M., & Boals, T. (2005). Considerations in Reconfiguring Cohorts and Resetting Annual Measurable Achievement Objectives (AMAOs) based on ACCESS for ELLs Data (WIDA Consortium Technical Report No. 3).

> This report is intended to assist states with the transition to a standards-based test and determining their AMAOs using ACCESS for ELLs.

Gottlieb, M. & Kenyon, D.M. (2006). The Bridge Study between Tests of English Language Proficiency and ACCESS for ELLs (WIDA Consortium Technical Report No. 2).

> This report provides the background, procedures, and results of a study intended to establish estimates of comparability between ACCESS for ELLs and four other English language tests used by Consortium member states. Students in Illinois and Rhode Island were administered ACCESS for ELLs along with one of the other four tests, and results on the four tests were compared with results on ACCESS for ELLs. Results allow states, districts, and schools to understand and report ACCESS for ELLs scores and to establish continuity between previous tests and ACCESS for ELLs.

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

> This report provides detailed information on the conceptualization, development, and field testing of ACCESS for ELLs. It also provides technical data on equating and scaling procedures, standard setting and operational score reporting, analyses of reliability and errors of measurement, and two initial validity studies.

Kenyon, D. M., Ryu, J.R. (Willow), & MacGregor, D. (2013). *Setting Grade Level Cut Scores for ACCESS for ELLs.* (WIDA Consortium Technical Report No. 4).

> This report describes the technical procedures and outcomes of the process to move from grade-level-cluster cut scores to grade-level cut scores. Proposed cut scores were determined mathematically and then reviewed and revised in a standard setting process involving 75 teachers from 14 WIDA Consortium states.

MacGregor, D., Kenyon, D. M., Gibson, S., & Evans, E. (2009). *Development and Field Test of Kindergarten ACCESS for ELLs*. (WIDA Consortium).

> This report provides detailed information on the conceptualization, development, and field testing of Kindergarten ACCESS for ELLs. It also provides technical data on equating and scaling procedures, standard setting and operational score reporting, and analyses of reliability and errors of measurement.

## Annual Technical Reports for ACCESS for ELLs

Below is a list of annual technical reports for ACCESS for ACCESS for ELLs, listed by year of publication. These reports provide extensive analysis of the results from the operational administration of ACCESS for ELLs. They provide detailed information on student results broken down by grade-level cluster, grade, and tier. They also provide detailed information on test and item characteristics.

Kenyon, D. M., MacGregor, D., Ryu, J.R. (Willow), Cho, B., and Louguit, M. (2006). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 100, 2004-2005 Administration.* (WIDA Consortium Annual Technical Report No. 1).

Kenyon, D. M., MacGregor, D., Louguit, M. Cho, B., and Ryu, J.R. (Willow). (2007). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 101, 2005-2006 Administration.* (WIDA Consortium Annual Technical Report No. 2).

MacGregor, D., Louguit, M., Ryu, J.R. (Willow), Kenyon, D.M., and Li, D. (2008). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 102, 2006-2007 Administration.* (WIDA Consortium Annual Technical Report No. 3).

MacGregor, D., Louguit, M., Huang, X., and Kenyon, D.M. (2009). Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 103, 2007-2008 Administration. (WIDA Consortium Annual Technical Report No. 4).

MacGregor, D., Louguit, M., Yanosky, T., Fidelman, C. G., Pan, M., Huang, X., and Kenyon, D.M. (2010). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 200, 2008-2009 Administration.* (WIDA Consortium Annual Technical Report No. 5).

Yanosky, T., Yen, S., Louguit, M., MacGregor, D., Zhang, Y., and Kenyon, D.M. (2011). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 201, 2009-2010 Administration.* (WIDA Consortium Annual Technical Report No. 6).

Yanosky, T., Chong, A., Louguit, M., Olson, E., Choi, Y., MacGregor, D., Yen, S., Cameron, C., and Kenyon, D.M. (2012). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 202, 2010-2011 Administration.* (WIDA Consortium Annual Technical Report No. 7).

Yanosky, T., Amos, M., Louguit, M., Olson, Cameron, C., Louguit, M., MacGregor, D., Yen, S., and Kenyon, D.M. (2013). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 203, 2011-2012 Administration.* (WIDA Consortium Annual Technical Report No. 8).

Center for Applied Linguistics (2014). Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 301, 2012-2013 Administration. (WIDA Consortium Annual Technical Report No. 9).

## *Other Documentation*

Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly, 2(1),* 1–34.

This article describes how an argument for test use might be structured so as to provide a clear linkage from test performance to interpretations and from interpretations to uses.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice.* Oxford: Oxford University Press.

This book presents the Assessment Use Argument, which provides a framework for justifying the intended uses of an assessment, as well as a guide for the design and development of the assessment itself.

Bauman, J., Boals, T., Cranley, E., Gottlieb, M., and Kenyon, D.M. (2007). The Newly Developed English Language Tests (World-Class Instructional Design and Assessment – WIDA). In Abedi, Jamal (Ed.), *English Language Proficiency Assessment in the Nation: Current Status and Future Practice.* Davis: University of California.

In this book chapter, the authors describe the test development process, from the development of standards through the development of items, field testing, and operationalization. They also report on validation of the test, accommodations, the test administration and technical manuals, and score reporting.

Chapelle, C. A., Enright, M. E., & Jamieson, J. (2010). Does an argument-based approach to validity make a difference? *Educational Measurement: Issues and Practice, 29(1),* 3–13.

Drawing on experience between 2000 and 2007 in developing a validity argument for the high-stakes Test of English as a Foreign Language™, this paper evaluates the differences between the argument-based approach to validity as presented by Kane (2006) and that described in the 1999 AERA/APA/NCME Standards for Educational and Psychological Testing.

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

This book uses the Test of English as a Foreign Language™ as a case study for validating test design. It attempts to meet the standards of educational measurement while also drawing on theory related to English language proficiency.

Cook, H. G. (2007). Alignment Study Report: The WIDA Consortium's English Language Proficiency Standards for English Language Learners in Kindergarten through Grade 12 to ACCESS for ELLs® Assessment. Madison, WI: WIDA Consortium.

In this report, the author describes a study to align the WIDA Standards to the ACCESS for ELLs test. The study was designed to address two questions: how well the test measures the proficiency levels described in the Standards, and how well the different domains of each standard are addressed by the domains of the test. The author concludes that overall ACCESS for ELLs is adequately aligned to the Standards.

Cook, H. G., Boals, T., Wilmes, C., and Santos, M. (2007). Issues in the Development of Annual Measurable Achievement Objectives (AMAOs) for WIDA Consortium States. Madison, WI: WIDA Consortium.

In this paper, the authors offer guidance to states in formulating Annual Measurable Achievement Objectives for English language learners.

Fox, J. (2011). Test review: ACCESS for ELLs[®]. *Language Testing 28* (3): 425-431.

The author provides a thorough review of ACCESS for ELLs, using the eight criteria enumerated in Fairbairn and Fox (2009).

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.

These documents contain the WIDA Standards and describe the rationale behind and development of the frameworks for large-scale state and classroom assessments. These frameworks comprise English Language Development standards, language domains, grade-level clusters, language proficiency levels and the model performance indicators upon which ACCESS for ELLs is based. They are meant to guide curriculum development, instruction, and assessment of English language learners.

Kane, M. (2006). Validation. In R. Brennan, (Ed.), *Educational Measurement (4th Edition)*, pp. 18-64. Westport, CT: Greenwood Publishing.

This book chapter presents a conceptualization of test validity where evidence and logical argument are brought together to evaluate claims and propositions about the proposed uses and interpretations of test results.

Kenyon, D.M., MacGregor, D., Li, D., and Cook, H. G. (2011). Issues in vertical scaling of a K-12 English language proficiency test. *Language Testing 28* (3): 383-400.

In this article, the authors describe the procedure used to place ACCESS for ELLs results on a vertical scale, and they discuss studies conducted to test the effectiveness of that scale.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). A Brief Introduction to Evidence-Centered Design (CSE Report 632). CA: Center for Research on Evaluation, Standards, and Student Testing.

This paper provides an introduction to the basic ideas of Evidence Centered Design, an approach to constructing educational assessments in terms of evidentiary arguments. It includes some of the terminology and models that have been developed to implement the approach.

National Research Council. (2011). Allocating federal funds for state programs for English language learners. Panel to Review Alternative Data Sources for the Limited-English Proficiency Allocation Formula under Title III, Part A, Elementary and Secondary Education Act, Committee on National Statistics and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

This report includes detailed descriptions of six English language proficiency tests, including ACCESS for ELLs, along with information about the reliability and validity of the tests.

Parker, C. E., Louie, J., and O'Dwyer, L. (2009). New measures of English language proficiency and their relationship to performance on large-scale content assessments (Issues & Answers Report, REL 2009–No. 066). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs, January 29, 2009.

This report describes a study investigating how well the domain tests on ACCESS for ELLs predict performance on a content test. Results indicate that the Reading and Writing tests are the strongest predictors.

Römhild, A., Kenyon, D. M., and MacGregor, D. (2011). Exploring domain-general and domain-specific linguistic knowledge in the assessment of academic English language proficiency. *Language Assessment Quarterly*, 8:213-228.

This article reports on a confirmatory factor analysis study conducted to model domain-specific and domain-general variance on ACCESS for ELLs. The authors found that, while domain-general linguistic knowledge represents the primary dimension across almost all test forms, domain-specific knowledge becomes increasingly salient as proficiency level increases.

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

This document presents the second edition of the WIDA English Language Development Standards, which were released in 2007. The second edition included the addition of formative and summative frameworks for assessment and instruction, the separation of Kindergarten into its own grade-level cluster, and the addition of the sixth proficiency level, "Reaching".

WIDA Consortium. (2012a). *2012 Amplification of the English Language Development Standards Kindergarten–Grade 12*. Madison, Wisconsin: Board of Regents of the University of Wisconsin System.

This document describes the amplified Strands of Model Performance Indicators that represent the WIDA English Language Development Standards. The amplification reflects states' content standards and the fluid and ongoing process of language development.

WIDA Consortium (2012b). *WIDA ACCESS for ELLs® Test Administration Manual*. Retrieved from www.wida.us/assessment/ACCESS/#about.

This document details the test administration procedures for ACCESS for ELLs.

WIDA Consortium. (2013). *Interpretive Guide for Score Reports Spring 2013* (WIDA Consortium). Madison, WI: The Board of Regents of the University of Wisconsin System.

This report provides an overview on how ACCESS for ELLs is scored and how those scores are reported. Part 1 gives a description of scores for 2014. Part 2 gives suggestions on how states can use scores, as well as examples of score reports to various stakeholders. Part 3 provides guidance on interpreting the reports.

Wolf, M., Kao, J., Griffin, N., Herman, J., Bachman, P., Chang, S., and Farnsworth, T. (2008). Issues in assessing English language learners: English language proficiency measures and accommodation uses—Practice review. Retrieved from the University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing Web site: http://www.cse.ucla.edu/products/rsearch.asp.

This paper describes the English language proficiency tests in use in school year 2006, including ACCESS for ELLs, and provides a summary of validity evidence for the tests.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P.Holland & H. Wainer (EMS.), *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

This book chapter describes procedures for conducting DIF analysis.

## Table of Contents
### Volume 1

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

## 1.1    Purpose of ACCESS for ELLs

The overarching purpose of Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs) is to assess the developing English language proficiency of English language learners in Grades K–12 in the United States following the English Language Development Standards (2012) of the multi-state WIDA Consortium. The WIDA English Language Proficiency (ELP) Standards (2004, 2007) were amplified in 2012 to become English Language Development (ELD) Standards, hereafter referred to as ELD Standards. The WIDA ELD Standards, aligned to state academic content standards, form the core of the WIDA Consortium's approach to instructing and testing English language learners and describe six levels of developing English language proficiency. ACCESS for ELLs may thus be described as a standards-based English language proficiency test designed to measure English language learners' social and academic language proficiency in English. It assesses social and instructional English as well as the language associated with language arts, mathematics, science, and social studies within the school context across the four language domains (Listening, Reading, Writing, and Speaking).

Other major purposes of ACCESS for ELLs 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 ESL/bilingual programs and determine staffing requirements,

- Providing data for meeting federal and state statutory requirements with respect to student assessment, and

- Providing information that enhances instruction and learning in programs for English language learners.

## 1.2    Format of ACCESS for ELLs

### 1.2.1    Integration with the Standards

The design of ACCESS for ELLs, from the structure of the assessment system to the content of each test booklet and item, is built upon the five foundational WIDA ELD Standards:

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

Standard 2 - English language learners communicate information, ideas, and concepts necessary for academic success in the content area of **Language Arts.**

Standard 3 - English language learners communicate information, ideas, and concepts necessary for academic success in the content area of **Mathematics.**

Standard 4 - English language learners communicate information, ideas, and concepts necessary for academic success in the content area of **Science.**

Standard 5 - English language learners 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: SI
- Language of English Language Arts: LA
- Language of Math: MA
- Language of Science: SC
- Language of Social Studies: SS

Every selected response item and every performance-based task on ACCESS for ELLs targets at least one of these five Standards.

## 1.2.2    Grade-level Clusters

The WIDA ELD Standards describe developing English language proficiency by five grade-level clusters. These are PreK-K, 1–2, 3–5, 6–8, and 9–12. Test booklets follow this grade-level clustering.

## 1.2.3    Language Domains

The WIDA ELD Standards describe developing English language proficiency for each of the four language domains: Listening, Speaking, Reading, and Writing. Thus, there is a section of the ACCESS for ELLs test assessing each of these four language domains.

## 1.2.4    Language Proficiency Levels

The WIDA ELD Standards describe the continuum of language development with five language proficiency levels that are fully delineated in the Standards document. These levels are "Entering," "Emerging," "Developing," "Expanding," and "Bridging." There is also a final exit stage known as Level 6, "Reaching," that describes students who have progressed across the entire WIDA English language proficiency continuum. These levels are shown graphically in Figure 1.2.4A.
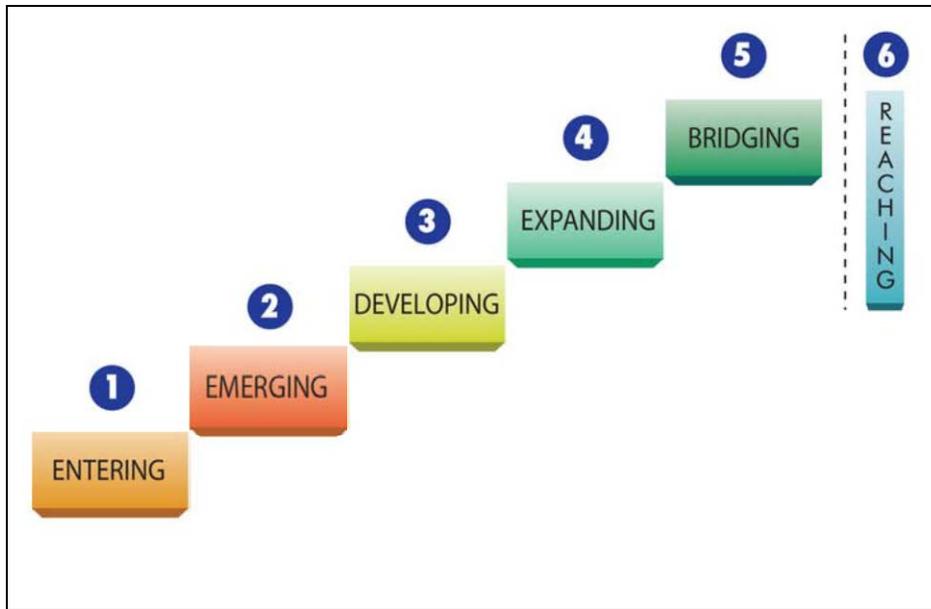
Figure 1.2.4A. The language proficiency levels of the WIDA ELD Standards

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

First, they appear in the **performance definitions**. According to the WIDA ELD Standards, the performance definitions provide a global overview of the stages of the language acquisition process. As such, they complement the model **performance indicators** (PIs, see below) for each language proficiency level. Being general definitions applicable across the PIs, the performance definitions are not explicitly replicated within the PIs. The performance definitions are based on three criteria. The first is students' increasing comprehension and production of the technical language required for success in the academic content areas. The second criterion is students' demonstration of oral interaction or writing of increasing linguistic complexity. The final criterion is the increasing development of phonological, syntactic, and semantic understanding in receptive skills or control in usage in productive language skills.

Second, the language proficiency levels of the WIDA ELD Standards are fully embedded in the accompanying PIs, which exemplify the Standards. The PIs describe the expectations for ELL students for each of the five **Standards,** at five different **grade-level clusters**, across four **language domains**, and at the five **language proficiency levels**. That is, within each combination of standard, grade-level cluster, and language domain is a PI at each of the five language proficiency levels. Proficiency Level 6, Reaching, represents the end of the continuum rather than another level of language proficiency. The sequence of these five PIs together describes a logical progression and accumulation of skills on the path from the lowest level of English language proficiency to full English language proficiency for academic success. These groupings of five PIs in logical progression are called a "strand."

ACCESS for ELLs is based on the 80 strands, containing 400 individual PIs, within the WIDA ELD Standards. (The Standards and the accompanying model PIs are available at the WIDA web site, (www.wida.us.) 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 PI. (See the sample items at the WIDA web site for examples.)

## 1.2.5    Tiers

Obviously, test items and tasks suitable for allowing Entering (Level 1) or Emerging (Level 2) students to demonstrate accomplishment of the PIs at their level of language proficiency (i.e., that allow them to demonstrate what they can do) will not allow Expanding (Level 4) or Bridging (Level 5) students to demonstrate the full extent of their language proficiency. Likewise, items and tasks developed to allow Expanding (Level 4) and Bridging (Level 5) students to demonstrate accomplishment of the PIs at their level would be far too challenging for Entering (Level 1) or Emerging (Level 2) students. Items that are far too easy for test takers may be boring and lead to inattentiveness on the part of students. Likewise, items that are far too difficult for test takers may be frustrating, discouraging them from giving their best performance. But more importantly, a test is a measure, and 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 proficiency. Tests need to be at the right difficulty level for individual test takers.

The solution to making ACCESS for ELLs appropriate to the proficiency level of individual students across the wide range of proficiencies described in the WIDA ELD Standards is to present the test items in three overlapping **tiers** for each grade-level cluster: A, B, and C. Figure 1.2.5A shows how the different tiers map to the language proficiency levels.
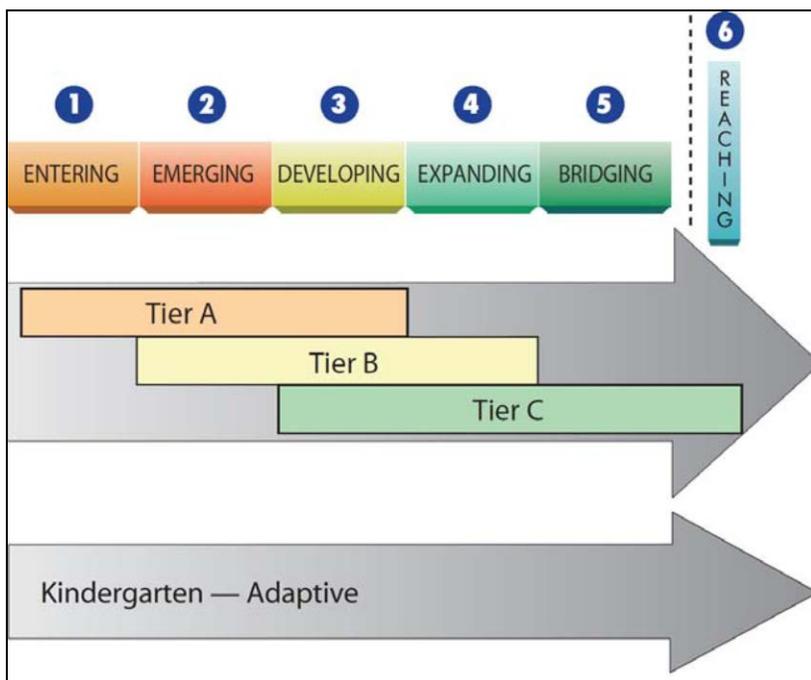


Figure 1.2.5A. Tier structure of ACCESS for ELLs

Thus, Tier A has items and tasks designed to allow students at the lowest language proficiency levels (Levels 1 and 2) to demonstrate meeting the WIDA ELD Standards at their language proficiency levels, and it includes some items targeted to Language Proficiency Level 3.

Likewise, Tier C has items and tasks designed to allow students at the highest language proficiency levels (Levels 4 and 5) to demonstrate meeting the WIDA ELD Standards at their language proficiency levels, while also containing some items targeted to Language Proficiency Level 3. In this test design, the tiers overlap: while Tier A and Tier C would share little in common, Tier B is composed of tasks from both Tiers A (Level 2) and C (Level 4), as well as tasks from Level 3. This overlap of tiers ensures that all of the PIs from the WIDA ELD Standards appear on the assessment; however, each test booklet does not need to contain an unduly large number of test items. The overlap also ensures that the entire language proficiency range is covered. Finally, the overlap ensures that the assessment is *horizontally equated*; that is, common items and tasks across tiers ensure that each tier is measuring to a common language proficiency scale. Thus, a test booklet at any given tier is primarily composed of items and tasks that span three targeted language proficiency levels. (Note that in order to assure that students are accurately measured to Level 6, Tier C also includes some items that are slightly more difficult than Language Proficiency Level 5. The Tier structure only applies to the Grade 1-12 Listening, Reading, and Writing portions of ACCESS for ELLs. Kindergarten (PreK-K) does not have tiers because it is an adaptive assessment.

The individually administered Speaking portion of the assessment for each grade-level cluster is designed as an adaptive measure. In each of its three parts, the test administrator begins with a task that allows students to demonstrate meeting the performance level expectations of the PIs at Level 1 and then presents continually more challenging tasks (tasks at Level 2, then Level 3, and so on). Within each part, the administrator stops presenting additional tasks when the student can no longer demonstrate meeting the expectations of the tasks. Table 1.2.5A summarizes the main points in the above discussion and illustrates the number of unique components in ACCESS for ELLs.

**Table 1.2.5A**
Unique Components in ACCESS for ELLs

| Grade-Level Clusters | List/Read/Write | | | Speak (adaptive) |
|:---:|:---:|:---:|:---:|:---:|
| | Tier A | Tier B | Tier C | |
| 9-12 | x | x | x | x |
| 6-8 | x | x | x | x |
| 3-5 | x | x | x | x |
| 1-2 | x | x | x | x |
| K | | x   (adaptive) | | |

## *1.3    Test Development*

### 1.3.1    Field Test

In 2004 the field test of ACCESS for ELLs was conducted. The purpose of the field test was to collect extensive data on items and forms in order to equate forms both horizontally (i.e., across tiers within the same grade-level clusters) and vertically (i.e., across grade-level clusters), as well as to judge the strength of individual items. The item pool for the field test consisted of 376 Listening items, 355 Reading items, and 51 Writing tasks. Two forms were used for each tier in each grade-level cluster. For equating purposes, common items were used across tiers, as well as

across forms, within grade-level clusters for the Listening, Reading, and Writing tests. In addition, common items were used across grade-level clusters for the Listening and Reading tests.

Table 1.3.1A shows the number of students who participated in the field test by grade-level cluster. 72.3% of the sample came from two states, Illinois and Wisconsin. Over half of the students (61.8%) had Spanish as their native language. The only other sizable language group was Hmong (13.8%). Indeed, of the 96 languages represented, only four languages (Spanish, Hmong, English, and French) had more than 100 students in the field test sample.

**Table 1.3.1A**
Field Test for Listening, Reading and Writing: Students per Grade-level Cluster

| Grade-Level Cluster | Students |
|---|---|
| 1-2 | 1,647 |
| 3-5 | 1,850 |
| 6-8 | 1,449 |
| 9-12 | 1,716 |
| Total | 6,662 |

A separate, individually administered field test was conducted for Speaking. One form was developed for each grade-level cluster, using the adaptive design described in 1.2.5, for a total of 52 tasks. Field testing for Speaking was conducted in Wisconsin and the District of Columbia. Table 1.3.1.B shows the number of students who participated in the Speaking field test by grade-level cluster.

**Table 1.3.1B**
Speaking Field Test: Students per Grade-level Cluster

| Grade-Level Cluster | Students |
|---|---|
| 1-2 | 103 |
| 3-5 | 159 |
| 6-8 | 136 |
| 9-12 | 125 |
| Total | 523 |

In addition, a separate field test was conducted in DC for the Kindergarten test. The final version of the adaptive Kindergarten assessment was produced by first choosing the Listening and Reading folders (i.e., sets of thematically related items) that contained items that were empirically the easiest for first graders based on the data collected from the field test. These folders were placed in the Kindergarten assessment in order from easiest to hardest. The Speaking portion of the Kindergarten assessment was the same as that for the 1–2 grade-level cluster, except it included only the SI and LA/SS folders, in order to reduce testing time. Special, very simple writing tasks were adapted from the 1–2 grade-level cluster Tier A SI writing folder.

The adaptive administration of the Kindergarten assessment is similar to that of the Speaking test. Thus, in any domain, if a student does not get at least two items in any folder (part) correct, the administrator stops testing in that domain and moves on to the next domain. (The exception is Speaking, which operates exactly as the standard ACCESS Speaking assessment.)

A total of 154 students participated in the Kindergarten field test. Of those, 55% were boys (84 students) and 45% were girls (70 students). 90.2% (139) of the students were Spanish speakers; the only other language with more than one student was Vietnamese (3).

## 1.3.2    Equating and Scaling

If test results are to be meaningful, they need to be reported on a standard scale that is familiar to test users and that keeps the same meaning whenever it is used. Scaling is the process of developing such a scale. Equating, in the present context, is the process of putting all of the tests onto the same scale, such that results mean the same regardless of which test items the test taker takes.

Of particular challenge for ACCESS for ELLs and similar tests is the need to have a vertically equated scale (i.e., one that can measure progress across the grade levels from K to 12), in addition to the horizontal equating needed across the three tiers of ACCESS for ELLs within each grade-level cluster.

For ACCESS for ELLs, a three-digit scale score was chosen for reporting purposes. The reporting scale would have an interpretive center point across domains and composites. The centering value was chosen as 350, which would represent the cut score between language proficiency Levels 3 and 4 for the 3–5 grade-level cluster. As an additional defining characteristic, the scale would have a lower bound of 100 (i.e., 250 points lower than the center of 350) and an upper bound of 600 (i.e., 250 points higher than 350). In other words, conceptually, students from grades K–2 with the lowest language proficiency in any domain could go no lower than a scale score of 100, thus making 100 a lower bound. Conceptually, students from the 9–12 grade-level cluster with the highest language proficiency in any domain could go no higher than 600, thus making 600 a higher bound. Observed scores on all tests should fall between these extremes.

It should be kept in mind that a scale score is an interpretation of a latent ability measure and not a record of "points" earned on the test. In other words, 100 does not necessarily represent a score of 0 at all grade-level clusters, nor does 600 represent a perfect score. In fact, due to the technical nature of a vertical scale—and one of the criticisms of it—as one goes up the scale from grade level to grade level (or grade-level cluster to grade-level cluster in the case of ACCESS for ELLs), the scales adjust for the developmental growth. Thus, even if a student consistently gets a score of 0 while moving from grade-level cluster to grade-level cluster, the student's scale score on a vertical scale would show an increase, even if very slight.

Thus, to interpret appropriately what the scale scores mean, a standard-setting study was conducted (see Section 1.3.3). However, in this section, we focus on the creation of the ACCESS for ELLs scale score.

The procedure for developing the scale was complex but involved a number of basic steps. These steps were carried out separately for each of the four domains until the last stage, when the separate domain scales were combined to form the composite scores. These steps, as conducted following the field test administration, are briefly summarized here. They are explained more fully in ACCESS for ELLs Technical Report 1, *Development and Field Test of ACCESS for ELLs*, as well as in Kenyon et al. (2011).

**Equating Design:** As previously described in Section 1.3.1, within each grade-level cluster, the Listening, Reading, and Writing test booklets were presented in three tiers (A, B, and C) and two

series (100 and 999), such that within each grade-level cluster, a different sample of test takers took different test booklets. However, the booklets had common folders of items (in the case of Listening and Reading) or common tasks on one or more test booklets for horizontal equating. In addition, there were common folders that went across grade-level clusters for vertical equating. Because of the adaptive design of the Speaking test, described in Section 1.3.1, there was only one form per grade and thus no need for horizontal equating.

For both Writing and Speaking, there were no common items in the equating design that linked the test booklets across different grade-level clusters. This was done intentionally as each task on these performance-based assessments was more complex, involved, and time-consuming. In addition, because these tasks targeted the WIDA Standards so closely, it would have been developmentally inappropriate to ask students to perform on tasks outside of their grade-level cluster. Thus, student performances on the Reading items were used as a scaling test for the Writing tasks, and performances on the Listening items were used as a scaling test for the Speaking tasks.

**Creating the Data Matrix:** The tests were scored, and the matrix of responses—every student's response to every Listening or Reading item or Writing or Speaking task—was the raw input into the scaling procedure.

**Developing the Logit Scale:** A calibration of the ability of the students and items using Rasch procedures was then applied to these data matrices, putting the difficulty of the items or tasks and the ability of the students onto one common interval linear scale. As described in ACCESS for ELLs Technical Report 1, *Development and Field Test of the ACCESS for ELLs®*, the steps of the common rating scale used to score the Writing items were also calibrated. The units of this scale are called logits, and by default the scale is usually centered at 0 (representing the average item difficulty for the ACCESS for ELLs items being calibrated). Theoretically, the logit scale runs from minus infinity to plus infinity, although in practice most tests run from about -4 logits to +4 logits.

**Transforming the Logit Scale to the Reporting Scale:** The logit scale has both negative numbers and decimals, which makes it confusing for many users. Therefore, scores on the logit scale were then transformed onto a reporting scale by means of a linear transformation of the logit scores—in this case, the ACCESS score scale. There is a separate scale for each of the four domains: Listening, Reading, Writing, and Speaking.

**Creating the Composite Scores**: The scores on the four reporting scales were then combined, in predetermined proportions, to create four composite scores:

- Oral Language Composite (50% Listening + 50% Speaking)
- Literacy Composite (50% Reading + 50% Writing)
- Comprehension Composite (30% Listening + 70% Reading)
- Overall Composite (15% Listening + 15% Speaking + 35% Reading + 35% Writing).

The Comprehension Composite score (based on performances in Listening and Reading) and the Overall Composite score (based on performances in all four domains) were created with Series 100. Beginning with Series 101, the Oral Language Composite score (based on performances in Listening and Speaking) and Literacy Composite score (based on performances in Reading and Writing) were added.

## 1.3.3    Standard Setting

In order to interpret appropriately what the scale scores mean, a standard-setting study was conducted. The standard-setting study was held in Madison, WI between April 20 and 27, 2005. The purpose of the study was not to set new standards on WIDA ACCESS for ELLs per se. Rather, the purpose was to use the WIDA ELD [1]Standards together with empirical information from the field test data to conduct a defensible and replicable approach to determine the relationship between student performances on the four domains of the ACCESS for ELLs and the language proficiency levels defined by the WIDA ELD Standards. The following is a brief summary of the Standard Setting study. For a fuller description, see ACCESS for ELLs Technical Report 1, *Development and Field Test of ACCESS for ELLs®*.

Four panels were convened, one for each major grade-level cluster: 1–2, 3–5, 6–8, and 9–12. On each panel were 20–22 teachers or administrators who were deemed qualified to participate in the study by the WIDA office, then located at the Wisconsin Department of Public Instruction. For Listening and Reading, a bookmarking procedure was used. Panelists were given books with all items within their grade-level cluster arranged by empirical difficulty, from least difficult to most difficult. After discussing the model performance indicators and the performance level descriptions from the WIDA ELD Standards, panelists were asked to read through the items and place a bookmark at the item that they determined a student at Language Proficiency Level 1 would have a 50% chance of answering a question correctly. They were then asked to repeat this procedure for all levels up to Level 5. During this procedure, panelists were encouraged to work independently.

After the initial round of bookmarking, the results were compiled and discussed with the panelists as a group. The panelists then were given the opportunity to reconsider and adjust their bookmarking, if they so chose. These results were compiled and presented to the WIDA management team, who used this data to help determine the final cut scores.

For Writing and Speaking, a modified body of work method was used. For Writing, the panelists were presented a book of portfolios from their grade-level cluster. Each portfolio consisted of all of the writings from the test of one student. The portfolios were chosen from students from each tier, and an attempt was made to choose students whose performances did not vary widely from one task to another. Within each grade-level cluster, portfolios were presented in ascending order; that is, the first portfolio represented a student's work that had received the lowest total raw score across the four pieces of writing, and the last portfolio presented was that of a student with a very high total raw score on the four pieces of writing.

After discussing the model performance indicators and the performance level descriptions as a group, the panelists were asked to read the portfolios and, working independently, make a judgment as to the probability that the work represents the work of a student at a given WIDA language proficiency level. For example, if they felt the portfolio represented the work of a student at Language Proficiency Level 3, they would write 100% under the column "3" on their paper. If they felt that it was a borderline performance between Levels 2 and 3, they would write 50% under "2" and 50% under "3". They were allowed to indicate up to two language

---

[1] Note: The 2005 ACCESS for ELLs field test and standard setting were based on the 2004 WIDA ELP standards. The WIDA English Language Proficiency (ELP) Standards (2004, 2007) were amplified in 2012 to become English Language Development (ELD) Standards (WIDA, 2012). In this section, the standards are referred to as ELD standards for consistency.

proficiency levels and a range in 10-point increments (i.e., 50/50, 60/40, 70/30, 80/20, or 90/10), or to indicate 100 under one language proficiency level. The results were compiled and discussed with the panelists as a group. The panelists then were given the opportunity to reconsider and adjust their bookmarking, if they so chose. The final results were analyzed by CAL using a logistic regression procedure to determine the points along the underlying writing proficiency continuum at which at least 50% of the panelists would be expected to agree that the writing represents the work of the next higher proficiency level than the current proficiency level. The results from this analysis were used to set the cut scores for the language proficiency levels.

The procedure for Speaking was similar, with the panelists listening to portfolios and recording their judgments.

## *1.4      Ongoing Item Development*

To keep ACCESS for ELLs secure, as well as to incorporate a program of continual refinement to the assessment (e.g., using colored illustrations), new items are being developed and field tested every year. In fact, one third to one half of the items is replaced yearly. The intent of the ongoing item development is to replace completely all items or tasks in each test form over a three-year period.

The schedule for refreshing items in the ACCESS for ELLs is illustrated in Table 1.4A. This table applies to all grade-level clusters except K, which was redesigned for Series 200 and is not refreshed annually. As can be seen from this table, for the Listening and Reading tests, all of the LA and MA items are replaced in alternating years, while the SC, SS, and SI items are replaced in a three-year cycle. For Series 302, all of the items on the Listening were refreshed, so Reading was not refreshed.

Beginning with Series 302, the Listening test transitioned from a traditional test administrator-read script to a media-delivered format, played either from CD or from streaming audio available online, for all grade level clusters except for Kindergarten. For more information, please see the *ACCESS for ELLs Series 302 Media-Based Listening Field Test Technical Brief* (Center for Applied Linguistics, forthcoming).

For Speaking, the SI task is replaced yearly, while the MS and LS tasks are replaced in alternating years. New items are field tested on separate forms during the operational administration of ACCESS for ELLs.

Table 1.4A also reflects a change in the Writing test that took effect starting with Series 201. In that series, the separate Math and Science folders were replaced with a combined Language of Math/Language of Science folder. Starting with that series, while the IT task will continue to be replaced yearly, the MS and LA tasks will be replaced in a two-year cycle for Tier A, and the MS and SI tasks will be replaced in a two-year cycle for Tiers B and C.

From Table 1.4A, we see that between Series 101 and Series 100, the IT Writing task and the MA/SC Speaking task were replaced. In the Listening and Reading portion of the test, various item folders were replaced following analysis of the field test and operational Series 100. Because ACCESS for ELLs was so new, it was decided that it was most important to be able to improve and/or replace weaker items across all five Standards than to choose only two Standards to be replaced.

**Table 1.4A**
Schedule for Refreshing ACCESS for ELLs Items

| Test Series | Year | Listening | | Reading | | Writing Tier A | Writing Tiers B &C | | Speaking | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | LA/MA | SC/SS/SI | LA/MA | SC/SS/SI | SI/LA/MA/SC | SI/MA/SC | IT | SI | LS/MS |
| 100 | 04-05 | | | | | | | | | |
| 101 | 05-06 | Various | Various | Various | Various | - | | IT | | MS |
| 102 | 06-07 | MA | SI | LA | SC | - | MA | IT | SI | LS |
| 103 | 07-08 | LA | SC | MA | SS | - | SC | IT | SI | MS |
| 200 | 08-09 | Some MA | Some SS | Some LA | Some SI | - | SI | IT | SI | LS |
| 201 | 09-10 | Various + Remaining MA | Various + Remaining SS | Various + Remaining LA | Various + Remaining SI | - | - | IT | SI | MS |
| 202 | 10-11 | LA | SI | MA | SC | MS | MS | IT | SI | LS |
| 203 | 11-12 | MA | SC | LA | SS | LA | SI | IT | SI | MS |
| 301 | 12-13 | LA | SS | MA | SI | MS | MS | IT | SI | LS |
| 302 | 13-14 | MA | SI | - | - | - | - | IT | SI | MS |
| 303 | 14-15 | MA | SC | LA* | SC* | LA | SI | IT | SI | LS |

Social and Instructional language (SI); Language of English Language Arts (LA); Language of Math (MA); Language of Science (SC); Language of Social Studies (SS); Integrated Language of Science, Language of Language Arts, and Language of Social Studies (IT); Language of Math and Language of Science (MS); Language of English Language Arts and Language of Social Studies (LS)

*Reading not refreshed for 302 because of full refreshment of 302 Listening. New specs for 303 are LA & SC.

The following paragraphs describe annual procedures currently in place that influence the development of future items.

## 1.4.1     Item Writing and Editing

The initial item writing is done by participants in an online item writing course conducted by CAL. An internal review of the items generated by that course is conducted, and items are chosen for further development based on how well they fit the Standards and PIs, and how different they are in terms of content from the previous year's items. Those items chosen are refined within CAL before undergoing an item content and bias and sensitivity reviews (see Section 1.4.2). Afterward, some items require further revision at CAL before being sent to MetriTech and WIDA central office for final review. Once returned to CAL, they are prepared for the field test.

### 1.4.2   Item Content and Bias and Sensitivity Reviews

After items are internally refined, they are reviewed by two panels: a content review panel and a bias and sensitivity review panel. The panels consist of educators from the WIDA Consortium states. Items are first submitted to the content review panel to assure that the content is accessible and relevant to students in the grade-level cluster, and that each item or assessment task matches the model performance indicators from the WIDA ELD Standards that it is intended to assess. After the items are revised based on the comments from the panel members, they are submitted to the bias and sensitivity review panel, which inspects the items for potential bias. For the bias and sensitivity review panel, panelists represent a wide variety of language backgrounds and ethnicities. Based on their recommendations, the items are revised as necessary.

### 1.4.3   Item Field Testing

All new items are field tested in conjunction with the current year's operational administration. Larger districts from across the WIDA Consortium states are invited to participate on a rotating schedule, and only districts that accept the invitation actually participate in the field test. Field testing occurs in WIDA states across the country immediately after the operational test is administered. Each participating student is administered items in only one domain. The field test is designed to take no more than 15 minutes on the part of any student participant.

For Listening and Reading, several forms of new items are prepared for each grade-level cluster, each containing two folders of new items and one folder of anchor items, in order to understand the difficulty of the new items in relation to the ACCESS for ELLs score scale. Thus, there are a total of three folders (9 items) per form. Within each form, an effort is made to alternate Standards. Thus, one form of the Listening field test might have two MA and one SI folders, while the other form has one MA and two SI folders. For Writing, four tasks are prepared per grade-level cluster: one task at each tier for the year's standard, and one IT  task. Students are presented with just one task, when possible at the appropriate tier. For Speaking, two folders of tasks are prepared per grade-level cluster, and each student is presented with both folders.

### 1.4.4   Item Calibration and Analysis

After the items are field tested, the results are analyzed using a Rasch model to determine their difficulty measure on the ACCESS for ELLs score scale. Items are also analyzed as to all aspects of their functioning (e.g., fit statistics) to determine whether they may be included in the next year's operational form. Only folders of items meeting all technical requirements are placed into the operational form.

### 1.4.5   DIF Items

Starting with Series 203, two phases of analysis (Phase I and Phase II) for differential item functioning (DIF) are conducted on the operational form while operational testing is still ongoing, in addition to the DIF analysis conducted for the Annual Technical Report. Each item is categorized into three levels of DIF: A, B, or C (Zieky, 1993). An item exhibiting A level DIF shows little or no bias toward a particular group, and an item exhibiting C level DIF is considered to display bias and should be closely examined by test developers.

Phase I is conducted at the same time as equating (see Section 1.3.2) using two sources of data: one, all student data available a week before the equating sample is pulled, called Early Return;

two, the equating sample, called Equating Sample. During Phase I analysis, only ethnicity DIF (Hispanic vs. Non-Hispanic) is investigated. In this phase, items that show high levels of DIF in both data sets are investigated by a team of content experts to determine if any construct-irrelevant factors can be identified that may contribute to DIF. Items which are identified as having construct-irrelevant sources of DIF will not be scored operationally. Two items were identified as having a C-level ethnicity DIF favoring Hispanics in the Early Return data but a A-level DIF favoring Hispanics for the Equating Sample; therefore, no further action was required. For Series 302, no items were unscored because of DIF in Phase I.

Phase II is conducted using all student data available in early May. During Phase II analysis, ethnicity and gender DIF were investigated. As with Phase I, items that show high levels of DIF are investigated by a team of content experts to determine if any construct-irrelevant factors can be identified that may contribute to DIF. Items which are identified as having construct-irrelevant sources of DIF will be removed from the test in the next operational year. For Series 302, one listening item was identified as having C-level ethnicity based DIF, favoring Hispanics; one reading item was identified as having C-level ethnicity based DIF, favoring Non-Hispanics.

For the Annual Technical Report, an ethnicity and gender DIF analysis is conducted using all student data. For Series 302, five items showed DIF. Out of 270 Listening items, two (0.7%) showed C-level DIF based on ethnicity, favoring Hispanics. Out of 342 Reading items, one (0.3%) showed C-level DIF based on ethnicity, favoring Non-Hispanics. Out of 43 Writing tasks, one (2.3%) showed C-level DIF based on ethnicity, favoring Non-Hispanics. Out of 62 Speaking items, one (1.6%) showed C-level DIF based on ethnicity, favoring Hispanics. These items are thoroughly analyzed by the Psychometrics/Research team at CAL to determine the potential sources of DIF. In terms of DIF by ethnicity (Hispanics versus Non-Hispanics), special attention is paid to the presence of Spanish-English cognates or false cognates that may affect student performance. That information is provided to the test development team, which makes necessary revisions to continuing items and keeps a record of such cognates for future reference. The test development team uses this information to guide the item development and review process for future items.

For information on the procedures used to calculate DIF, see Section 5.1.4.

## 1.5    Reporting of Results

### 1.5.1    Scale Scores

ACCESS for ELLs scores are reported as both scale scores and proficiency level scores. Scores are given for all four language domains. In addition, four composite scores are given: Oral Language Composite (based on performances in Listening and Speaking), Literacy Composite (based on performances in Reading and Writing), Comprehension Composite (based on performances in Listening and Reading), and Overall Composite (based on performances in all four domains).

Raw scores are converted to scale scores through processes called equating and scaling (see section 1.3.2 for details). These processes allow us to report scores on a standard scale that is familiar to test users and that remains constant across test forms and grade-level clusters. Scale scores range from 100 to 600. Beginning with Series 102, the center point of the scale, 350,

which formerly represented the cut score between Language Proficiency Levels 3 and 4 for the 3–5 grade-level cluster, represents the same cut score for Grade 5 only.

The scores for the four composite scores are calculated using the following weights:

- Oral Language Composite (50% Listening + 50% Speaking)
- Literacy Composite (50% Reading + 50% Writing)
- Comprehension Composite (30% Listening + 70% Reading)
- Overall Composite (15% Listening + 15% Speaking + 35% Reading + 35% Writing).

Figure 1.5.1A depicts the weighting for each of the composite scores. As shown, the Overall Composite is computed using scores from all four domains. Each of the other three composites is shown with the weighting of domains, in terms of the weighting used for the Overall Composite. As the diagram shows, more weighting is given to the literacy skills than to the oral skills for the Overall Composite
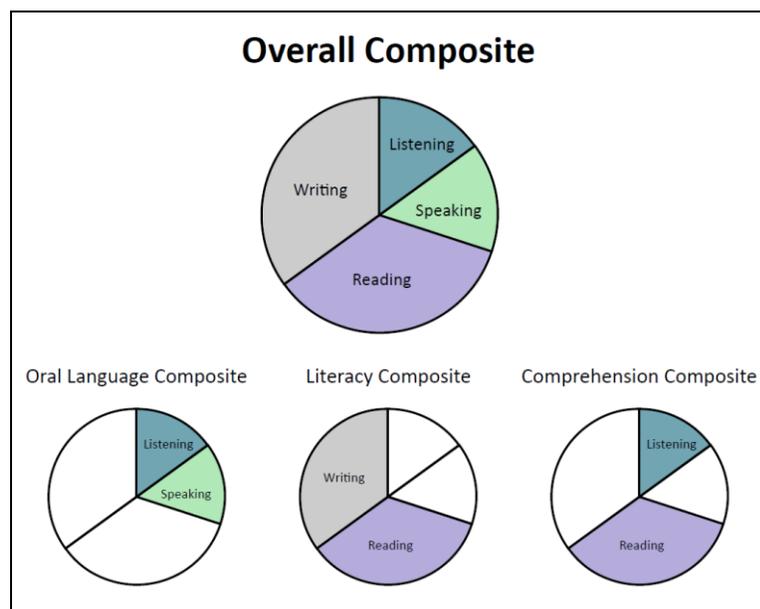


Figure 1.5.1A. Domain Composites

## 1.5.2    Language Proficiency Level Scores

In addition to the ACCESS scale scores, users of ACCESS also receive proficiency level scores. These scores are *interpretive*; that is, they interpret a student's scale score in terms of the results of the standard setting study. The cut scores between proficiency levels are presented in Tables 1.5.2A–H and reflect the adoption of the grade-level cut scores for Series 102 and beyond, as well as the Instructional and Accountability cut scores adapted for Kindergarten for Series 200 and beyond.

**Table 1.5.2A**
Cut Scores (Listening)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | List | 175 | 204 | 240 | 279 | 322 |
| K (Accountability) | List | 229 | 251 | 278 | 286 | 308 |
| 1 | List | 238 | 267 | 295 | 305 | 330 |
| 2 | List | 247 | 281 | 311 | 324 | 350 |
| 3 | List | 255 | 295 | 325 | 340 | 367 |
| 4 | List | 264 | 307 | 338 | 355 | 383 |
| 5 | List | 274 | 318 | 350 | 368 | 397 |
| 6 | List | 283 | 328 | 359 | 380 | 409 |
| 7 | List | 293 | 337 | 368 | 390 | 418 |
| 8 | List | 302 | 345 | 375 | 399 | 426 |
| 9 | List | 312 | 352 | 381 | 406 | 432 |
| 10 | List | 322 | 358 | 386 | 412 | 436 |
| 11 | List | 332 | 363 | 389 | 416 | 438 |
| 12 | List | 343 | 366 | 391 | 418 | 439 |

**Table 1.5.2B**
Cut Scores (Reading)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | Read | 121 | 159 | 204 | 228 | 255 |
| K (Accountability) | Read | 238 | 251 | 261 | 274 | 295 |
| 1 | Read | 253 | 269 | 283 | 294 | 314 |
| 2 | Read | 267 | 286 | 303 | 312 | 331 |
| 3 | Read | 279 | 302 | 320 | 328 | 347 |
| 4 | Read | 291 | 316 | 336 | 343 | 360 |
| 5 | Read | 302 | 328 | 350 | 355 | 372 |
| 6 | Read | 312 | 340 | 360 | 366 | 382 |
| 7 | Read | 321 | 349 | 369 | 375 | 391 |
| 8 | Read | 329 | 358 | 376 | 382 | 398 |
| 9 | Read | 336 | 364 | 381 | 387 | 402 |
| 10 | Read | 341 | 370 | 383 | 390 | 406 |
| 11 | Read | 346 | 374 | 384 | 392 | 407 |
| 12 | Read | 350 | 376 | 385 | 393 | 408 |

**Table 1.5.2C**
Cut Scores (Writing)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | Writ | 145 | 218 | 244 | 269 | 326 |
| K (Accountability) | Writ | 225 | 259 | 295 | 323 | 350 |
| 1 | Writ | 238 | 272 | 308 | 336 | 362 |
| 2 | Writ | 251 | 285 | 320 | 348 | 373 |
| 3 | Writ | 264 | 297 | 330 | 360 | 384 |
| 4 | Writ | 275 | 308 | 340 | 371 | 394 |
| 5 | Writ | 287 | 319 | 350 | 381 | 403 |
| 6 | Writ | 298 | 329 | 361 | 391 | 412 |
| 7 | Writ | 308 | 339 | 371 | 399 | 420 |
| 8 | Writ | 318 | 348 | 381 | 408 | 428 |
| 9 | Writ | 327 | 356 | 389 | 415 | 435 |
| 10 | Writ | 336 | 363 | 397 | 422 | 441 |
| 11 | Writ | 344 | 370 | 404 | 428 | 447 |
| 12 | Writ | 352 | 377 | 410 | 434 | 452 |

**Table 1.5.2D**
Cut Scores (Speaking)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | Spek | 256 | 285 | 308 | 342 | 365 |
| K (Accountability) | Spek | 269 | 314 | 343 | 366 | 383 |
| 1 | Spek | 278 | 318 | 344 | 367 | 385 |
| 2 | Spek | 286 | 322 | 345 | 368 | 386 |
| 3 | Spek | 293 | 326 | 346 | 369 | 389 |
| 4 | Spek | 299 | 329 | 348 | 371 | 391 |
| 5 | Spek | 305 | 333 | 350 | 374 | 394 |
| 6 | Spek | 310 | 337 | 353 | 377 | 397 |
| 7 | Spek | 314 | 340 | 358 | 380 | 400 |
| 8 | Spek | 317 | 344 | 361 | 384 | 404 |
| 9 | Spek | 319 | 347 | 366 | 388 | 407 |
| 10 | Spek | 321 | 351 | 371 | 393 | 412 |
| 11 | Spek | 322 | 354 | 377 | 399 | 416 |
| 12 | Spek | 32~~32~~ | 357 | 384 | 405 | 421 |

**Table 1.5.2E**
Cut Scores (Oral Language Composite)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | Oral | 216 | 245 | 274 | 311 | 344 |
| K (Accountability) | Oral | 249 | 283 | 311 | 326 | 346 |
| 1 | Oral | 258 | 293 | 320 | 336 | 358 |
| 2 | Oral | 267 | 302 | 328 | 346 | 368 |
| 3 | Oral | 274 | 311 | 336 | 355 | 378 |
| 4 | Oral | 282 | 318 | 343 | 363 | 387 |
| 5 | Oral | 290 | 326 | 350 | 371 | 396 |
| 6 | Oral | 297 | 333 | 356 | 379 | 403 |
| 7 | Oral | 304 | 339 | 363 | 385 | 409 |
| 8 | Oral | 310 | 345 | 368 | 392 | 415 |
| 9 | Oral | 316 | 350 | 374 | 397 | 420 |
| 10 | Oral | 322 | 355 | 379 | 403 | 424 |
| 11 | Oral | 327 | 359 | 383 | 408 | 427 |
| 12 | Oral | 333 | 362 | 388 | 412 | 430 |

**Table 1.5.2F**
Cut Scores (Literacy Composite)

| Grades | Domain | Cut | | | | |
|---|---|---|---|---|---|---|
| | | 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
| K (Instructional) | Litr | 133 | 189 | 224 | 249 | 291 |
| K (Accountability) | Litr | 232 | 255 | 278 | 299 | 323 |
| 1 | Litr | 246 | 271 | 296 | 315 | 338 |
| 2 | Litr | 259 | 286 | 312 | 330 | 352 |
| 3 | Litr | 272 | 300 | 325 | 344 | 366 |
| 4 | Litr | 283 | 312 | 338 | 357 | 377 |
| 5 | Litr | 295 | 324 | 350 | 368 | 388 |
| 6 | Litr | 305 | 335 | 361 | 379 | 397 |
| 7 | Litr | 315 | 344 | 370 | 387 | 406 |
| 8 | Litr | 324 | 353 | 379 | 395 | 413 |
| 9 | Litr | 332 | 360 | 385 | 401 | 419 |
| 10 | Litr | 339 | 367 | 390 | 406 | 424 |
| 11 | Litr | 345 | 372 | 394 | 410 | 427 |
| 12 | Litr | 351 | 377 | 398 | 414 | 430 |

**Table 1.5.2G**
Cut Scores (Comprehension Composite)

| Grades | Domain | Cut 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
|---|---|---|---|---|---|---|
| K (Instructional) | Cphn | 138 | 173 | 215 | 244 | 276 |
| K (Accountability) | Cphn | 235 | 251 | 266 | 278 | 299 |
| 1 | Cphn | 249 | 268 | 287 | 297 | 319 |
| 2 | Cphn | 261 | 285 | 305 | 316 | 337 |
| 3 | Cphn | 272 | 300 | 322 | 332 | 353 |
| 4 | Cphn | 283 | 313 | 337 | 347 | 367 |
| 5 | Cphn | 294 | 325 | 350 | 359 | 380 |
| 6 | Cphn | 303 | 336 | 360 | 370 | 390 |
| 7 | Cphn | 313 | 345 | 369 | 380 | 399 |
| 8 | Cphn | 321 | 354 | 376 | 387 | 406 |
| 9 | Cphn | 329 | 360 | 381 | 393 | 411 |
| 10 | Cphn | 335 | 366 | 384 | 397 | 415 |
| 11 | Cphn | 342 | 371 | 386 | 399 | 416 |
| 12 | Cphn | 348 | 373 | 387 | 401 | 417 |

**Table 1.5.2H**
Cut Scores (Overall Composite)

| Grades | Domain | Cut 1/2 | 2/3 | 3/4 | 4/5 | 5/6 |
|---|---|---|---|---|---|---|
| K (Instructional) | Over | 158 | 206 | 239 | 268 | 307 |
| K (Accountability) | Over | 237 | 263 | 288 | 307 | 329 |
| 1 | Over | 249 | 277 | 303 | 321 | 344 |
| 2 | Over | 261 | 290 | 316 | 335 | 357 |
| 3 | Over | 272 | 303 | 328 | 347 | 369 |
| 4 | Over | 283 | 314 | 340 | 359 | 380 |
| 5 | Over | 293 | 324 | 350 | 369 | 390 |
| 6 | Over | 302 | 334 | 359 | 379 | 399 |
| 7 | Over | 311 | 342 | 368 | 386 | 407 |
| 8 | Over | 319 | 350 | 375 | 394 | 414 |
| 9 | Over | 327 | 357 | 382 | 400 | 419 |
| 10 | Over | 333 | 363 | 387 | 405 | 424 |
| 11 | Over | 340 | 368 | 391 | 409 | 427 |
| 12 | Over | 34~~65~~ | 372 | 395 | 413 | 430 |

A proficiency level score consists of a two-digit decimal number (e.g., 4.5). The first digit represents the student's overall language proficiency level range based on the student's scale score. A score of 4.5 indicates that the student is in language proficiency Level 4. 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 tells us that the student's scale score is halfway between the cut scores for Levels 4 and 5.

Unlike ACCESS scale scores, which form an interval scale and are continuous across the grades from K to 12, ACCESS proficiency level scores are, of course, dependent upon which grade a student was in when ACCESS for ELLs was taken. See, for example, the Listening cut scores in Table 1.5.2A. If a child is in Grade 2 and receives a 350 in Listening, that would be a proficiency level score of 6.0; if the child is in Grade 5 and receives a 350 in Listening, that would be 4.0; if the child is in Grade 8 and receives a 350 in Listening, that would be a 3.2; and if a child is in Grade 12 and receives a 350 in Listening, that would be a 2.3. (Note that grade-level-cluster cut scores were used to interpret performances on ACCESS for ELLs for Series 100 and 101. Beginning with Series 102, grade-level cut scores were used.)

Note that because the width between cut scores varies, proficiency level scores should not be considered as forming an interval scale. That is, it cannot be assumed to be the same distance between proficiency level scores of 1.5 and 2.5 as between 2.5 and 3.5. Only scale scores should be used as interval measures. Proficiency level scores are interval within a grade and level (e.g., it is the same distance in grade 3 between 3.1 and 3.2 as between 3.7 and 3.8), but they do not form an interval scale across language proficiency levels.

## 1.5.3 Results by English Language Development Standards

To provide a more complete picture of a student's performance, raw scores are reported by ELD Standards.

For Comprehension (combined Listening and Reading), the five ELD  scores (Social and Instructional language, Language Arts, Mathematics, Science, and Social Studies) are reported as number correct out of maximum possible (e.g., 3 of 8). It should be noted that the absolute number of items that a student sees in any given language proficiency area varies by tier.

For Speaking, ELD scores are reported as raw numbers based on the number of tasks that the student met or exceeded in that standard. The maximum score for Social and Instructional language is 3; the maximum for Language Arts/Social Studies and for Mathematics/Science is 5.

For Writing tasks, three ELD ratings are reported for each of the three or four tasks on the form. The three ratings are for Linguistic Complexity, Vocabulary Usage, and Language Control. Each of these scores can range from 0 to 6.

## *1.6 Test Administration*

## 1.6.1 Test Administrator Training

To prepare individuals to serve as test administrators, test administrator training for Series 302 was conducted through an online course hosted at www.wida.us. Three certifications were offered to participants: a group test administration certification pertaining to the Listening, Reading, and Writing portions of ACCESS for ELLs; a certification for the Speaking test; and a certification for the Kindergarten test. In order to receive a certification, participants had to pass a quiz.

## 1.6.2 Test Security

Every effort is made to keep the test secure at all levels of development and administration. CAL and MetriTech follow policies and procedures regarding the security of the test, and every

individual involved in the administration of the test from the district to the classroom level is trained in issues of test security.

### 1.6.3   Test Accommodations

As a test of developing English language proficiency designed for English language learners, there are no special test accommodations for this group of students. However, if a student also has an IEP, to the extent possible and practical, the recommendations in the student's IEP are to be followed. The extent to which this was successfully accomplished was a local decision during the administration of Series 302.

The WIDA Consortium for the first time ever during the 2011–2012 testing cycle has made available an alternate assessment for ACCESS for ELLs: Alternate ACCESS for ELLs (Alternate ACCESS). Alternate ACCESS is intended only for English language learners who have cognitive disabilities that are severe enough to prevent meaningful participation in the ACCESS for ELLs with accommodations. The results of the Alternate ACCESS operational administration will appear in a separate technical report (forthcoming).

The recommendations regarding physical disabilities, such as deafness or blindness, are available on the WIDA website (http://www.wida.us/get.aspx?id=289) but are being clarified for more standardization.

## *1.7   Scoring*

Test booklets are returned to MetriTech, where they are electronically scanned in preparation for scoring. Listening, Reading, and Writing are scored by Metritech. Speaking is locally scored by the test administrator. Details of the scoring methods are described below.

### 1.7.1   Listening and Reading

In the case of the Listening and Reading tests, all items are selected-response and thus are dichotomously scored as correct or incorrect. Students have entered their answers directly into the test booklets, so each page is scanned into an electronic database.

### 1.7.2   Writing

Students' responses to the Writing tasks are centrally scored at MetriTech by raters who are trained to follow the WIDA Consortium's Writing Rubric (see 1.7.2.1). The rubric reflects the Performance Level Descriptions of the WIDA ELD Standards and is presented in Table 1.7.2A.

Table 1.7.2A

Performance Level Descriptions of the WIDA ELD Standards

At the given level of English language proficiency, English language learners will process, understand, produce or use:

| Level | Description |
|---|---|
| **6 - Reaching** | • specialized or technical language reflective of the content area at grade level<br>• a variety of sentence lengths of varying linguistic complexity in extended oral or written discourse as required by the specified grade level<br>• oral or written communication in English comparable to proficient English peers |
| **5 - Bridging** | • specialized or technical language of the content areas<br>• a variety of sentence lengths of varying linguistic complexity in extended oral or written discourse, including stories, essays, or reports<br>• oral or written language approaching comparability to that of English-proficient peers when presented with grade-level material |
| **4 - Expanding** | • specific and some technical language of the content areas<br>• a variety of sentence lengths of varying linguistic complexity in oral discourse or multiple, related sentences or paragraphs<br>• oral or written language with minimal phonological, syntactic, or semantic errors that do not impede the overall meaning of the communication when presented with oral or written connected discourse with sensory, graphic, or interactive support |
| **3 - Developing** | • general and some specific language of the content areas<br>• expanded sentences in oral interaction or written paragraphs<br>• oral or written language with phonological, syntactic, or semantic errors that may impede the communication, but retain much of its meaning, when presented with oral or written, narrative, or expository descriptions with sensory, graphic, or interactive support |
| **2 - Emerging** | • general language related to the content areas<br>• phrases or short sentences<br>• oral or written language with phonological, syntactic, or semantic errors that often impede the meaning of the communication when presented with one to multiple-step commands, directions, questions, or a series of statements with sensory, graphic, or interactive support |
| **1 - Entering** | • pictorial or graphic representation of the language of the content areas<br>• words, phrases, or chunks of language when presented with one-step commands, directions, WH-, choice, or yes/no questions, or statements with sensory, graphic, or interactive support<br>• oral language with phonological, syntactic, or semantic errors that often impede meaning when presented with basic oral commands, direct questions, or simple statement with sensory, graphic or interactive support |

The Writing rubric contains expectations for three aspects of Writing that play an important role in determining proficiency level: Linguistic Complexity, Vocabulary Usage, and Language Control. Table 1.7.2B presents the WIDA Consortium's Writing Rubric.

Table 1.7.2B

WIDA Consortium's Writing Rubric for Grades 1-12

| Level | Linguistic Complexity | Vocabulary Usage | Language Control |
|---|---|---|---|
| 6 Reaching | A variety of sentence lengths of varying linguistic complexity in a single tightly organized paragraph or in well-organized extended text; tight cohesion and organization | Consistent use of just the right word in just the right place; precise Vocabulary Usage in general, specific, or technical language | Has reached comparability to that of English proficient peers functioning at the "proficient" level in state-wide assessments |
| 5 Bridging | A variety of sentence lengths of varying linguistic complexity in a single organized paragraph or in extended text; cohesion and organization | Usage of technical language related to the content area; evident facility with needed vocabulary | Approaching comparability to that of English proficient peers; errors don't impede comprehensibility |
| 4 Expanding | A variety of sentence lengths of varying linguistic complexity; emerging cohesion used to provide detail and clarity | Usage of specific and some technical language related to the content area; lack of needed vocabulary may be occasionally evident | Generally comprehensible at all times, errors don't impede the overall meaning; such errors may reflect first language interference |
| 3 Developing | Simple and expanded sentences that show emerging complexity used to provide detail | Usage of general and some specific language related to the content area; lack of needed vocabulary may be evident | Generally comprehensible when writing in sentences; comprehensibility may from time to time be impeded by errors when attempting to produce more complex text |
| 2 Emerging | Phrases and short sentences; varying amount of text may be copied or adapted; some attempt at organization may be evidenced | Usage of general language related to the content area; lack of vocabulary may be evident | Generally comprehensible when text is adapted from model or source text, or when original text is limited to simple text; comprehensibility may be often impeded by errors |
| 1 Entering | Single words, set phrases, or chunks of simple language; varying amounts of text may be copied or adapted; adapted text contains original language | Usage of highest frequency vocabulary from school setting and content areas | Generally comprehensible when text is copied or adapted from model or source text; comprehensibility may be significantly impeded in original text |

In addition to training in the generic rubric, training is provided to scorers as to expectations for each grade level and for each Writing task. For example, exceptional vocabulary usage in the 1–2 grade-level cluster would not be so exceptional at the 9–12 grade-level cluster. The amount of writing and sophistication of thought at each performance level generally increases with moving up the grade-level clusters. Thus, a single generic rubric rooted in the WIDA ELD Standards lies at the core of the scoring of Writing, but developmental differences between grade-level clusters are part of the additional training that each rater receives.

Scorers are provided anchor papers for each task. Training sets are also created, as well as calibration sets with which scorers are tested during the operational training session. Raters failing to meet standards on the calibration sets are removed from scoring.

In applying the rubric, the following method of scoring Writing is used. First, the Metritech rater determines the language proficiency level that best characterizes the Writing sample (e.g., Level 3). Then, the rater considers whether in any category the Writing displayed any particular weakness (i.e., was lower in one of the three) or displayed any particular strength (i.e., was higher in one of the three categories). Finally, the rater awards three scores, one for each category: a 3-3-3 represents a solid Level 3 writing sample; a 3-3-2 is a low Level 3 writing sample that is a little weaker than expected in its language control; while a 3-4-3 is a high Level 3 writing sample that is a little stronger than expected in its vocabulary usage. The final score is the sum of the three scores; i.e., 9 for a solid Level 3 paper, 8 for a low Level 3 paper, and 10 for a high Level 3 paper.

In calculating an Overall Composite raw score for Writing, results from the different tasks are given different weights. These weights are intended to reflect the amount of writing that each task may be expected to produce. The weightings for the different tasks are as follows:

- Kindergarten: 1-1-1-1-2-1
- Grades 1–2 Tier A form: 1-1-1-3
- Grades 1–12 Tier B and C forms: 1-2-3
- Grades 3–12 Tier A forms: 1-1-1

For example, for all grades on Tier B and C tests the three tasks are given weights of 1, 2, and 3. Thus, a student who receives scores of 6, 5, and 4 on the three Writing tasks for that test would have an overall writing raw score of 28 ((6*1) +(5*2) + (4*3)).

### 1.7.2.1 Scoring Procedures for Writing

Scoring of ACCESS for ELLs is handled at the MetriTech scoring facilities in Illinois.

All constructed-response scoring for ACCESS for ELLs is performed utilizing a proprietary on-line scoring system (*MTscore*). As with all aspects of ACCESS for ELLs, MetriTech's top concern is security of student data and the items and forms eliciting student responses. Some of the strict security measures implemented as part of *MTscore* include:

- All students' identifying or biographical data (including name, ID number, gender, etc.) will be stripped from scorer images and will not be included in data transferred into *MTscore*

- Students' constructed responses will have an untraceable, non-identifying index number

- Item and student response images will be available only through *MTscore* and cannot be accessed by any outside network or saved on any media

- No image or portion of student response image can be printed, with the exception of Master Scorers only needing to print student responses in cases of alert papers.

- Score session access restrictions, requiring scorer login during predetermined times and dates only

With scoring centers located near several universities, MetriTech has a large pool of qualified scoring applicants from which to choose. Applicants must possess a minimum of a bachelor's degree and pass proprietary pre-employment tests found to predict performance. Many have backgrounds in education and are active or retired teachers.

Applicants are required to attend a pre-employment testing session where they review their already completed online application, answer additional questions specific to the project for which they are applying, and complete a series of proprietary pre-employment screening tests that reliably predict scorer performance. Hiring criteria include, but are not limited to, completion of at least a bachelor's degree from an accredited college or university; work experience, particularly teaching or education-related experience; and test scores.

Lead scoring staff members include master scorers and trainers (already on staff at the start of each project), as well as table leaders and senior scorers for each content area (usually assigned to specific projects based on their content qualifications and background). Before scorer training begins, group leaders and senior scorers are trained by master scoring staff so that they are familiar with the rubrics, annotated anchor papers (originally provided by CAL and augmented by senior MetriTech staff each year), training sets, calibration sets, and scoring procedures. MetriTech master scoring staff has been working with scoring protocols for various programs and states for an average of seven years. This core group works closely with the CAL development staff, augmenting originally supplied training materials for each year of ACCESS scoring.

Each potential reader has been selected to train on a particular grade-span. The training process starts with an on-line training session, where each reader will review the rubrics, the elements of analytic scoring, and anchor papers. Each score point on each rubric is defined, and approved examples of student work that meet the criteria for each score point are presented and analyzed. Following this presentation, the readers work through selected modules of sample papers. Each paper in a training module has already been reviewed and scored by the master reader. As the readers finish the training module, their recorded scores and rationales for their scoring are systematically scored. Discrepancies are noted and feedback and additional modules are presented to the reader to provide further training. Finally, each reader is given a post-test module containing sample student responses. The readers score these modules independently, and the final scores that they assign are compared with those assigned by the master reader. Readers need to reach the criterion of 70% exact agreement with the master reader's score to complete training and to be approved to score live test material. This process is repeated for each scorer selected for training. Training sessions utilizing one-on-one on-line interactive modular trainings supported by printed training manuals and master trainer Q&A provide readers by grade-cluster and typically include eight hours of material.

*How scorers are supervised during the scoring process.*

Group Leaders

- Prioritize work assignments for the scorers in their group for each shift

- Assign scorers work for each shift

- Review completed scoring for their group

- Track scorer attendance

- Monitor decorum within their group

Room Leaders

- Coordinate all Group Leaders on a shift

- Prioritize work assignments for the room for each shift

- Track scorer productivity on each shift

- Monitor decorum for the room

Master Scorers and Trainers

- Complete quality control/scoring checks on all employees on a daily basis, at pre-determined rates

- Provide written as well as verbal one-on-one feedback to scorers on a daily basis

- Provide retraining as needed

- Recommend scorer reassignment as needed

For the ACCESS for ELLs constructed-response scoring, papers from each scorer are randomly directed to the group leader for re-checking. If a group leader finds that a scorer's rates fall below the expected standard, the scorer is directed to retraining.

To monitor that the scoring rubric is being applied consistently across scoring sessions, specially prepared calibration sets are routed to each scorer daily. To the scorer, the calibration student images look like regular student responses. However, master scorers have already reviewed each response in these sets, and the master scorer has created a key of expected scores for each sample. Once the scorer completes the set, the scores that he or she assigned are immediately checked against the master key by the system. This approach allows for the immediate detection and correction of scorer drift. Exact agreement levels between the active and master scorer must exceed the standards established for the project (80% exact agreement) or the scorer is locked out of the system until they have successfully completed a retraining with the master scorer.

Twenty percent of all constructed-response items are blindly re-scored by another reader to provide overall inter-rater reliability. This information is kept for future analysis, reporting in the technical report, and reporting to the master scorer, allowing another avenue of feedback to the individual scorers.

### 1.7.3    Speaking

The Speaking test is administered individually to each test taker. Each task is immediately scored by the administrator while the test is being given. The administration and scoring procedure were designed together to be quite simple to implement. As described previously, the Speaking tasks are designed around the PIs to allow students to demonstrate mastery of the performance level for which the task is designed. After administering each task and listening to the student's responses, the administrator decides whether the student's performance exceeds, meets, or approaches task-level expectations. Specifically, the possible ratings are defined as follows:

**Exceeds:** The student's performance exceeds task-level expectations in quantity and/or quality.

**Meets:** The student's performance meets task-level expectations in quantity and quality.

**Approaches:** The student's performance approaches task-level expectations, but falls short in quantity and/or quality.

**No Response:** The student's performance is quite inadequate: there is no response, the response is incomprehensible or in a language other than English, or the student is unable to understand the task directions.

Operationally, a score of 1 is given for every task that either meets or exceeds expectations, and a 0 is given for any task that is rated as approaches or no response. The sum of those scores is the total Speaking raw score for that student.

Table 1.7.3A presents the WIDA Consortium's Speaking Rubric, which summarizes the expectations for each task level on the Speaking assessment. These expectations are drawn from the performance level descriptions of the WIDA ELD Standards and are divided into three components (Linguistic Complexity, Vocabulary Usage, and Language Control). The training for test administrators consists of familiarizing them with the tasks at each level and listening to responses to those tasks, determining whether they meet the task-level expectations or not.

### 1.7.3.1 Training Procedures for Scoring Speaking

The Speaking Test is the only portion of ACCESS for ELLs that the test administrator scores. Test administrators must complete the Speaking Test module of the online ACCESS for ELLs Test Administrator Training and the accompanying quiz. The training focuses on developing the test administrator's ability to conduct the test using standardized testing procedures and to score the test reliably. Test administrators are provided training on test administration procedures such as navigating the test, scores and ratings. To reliably score the test, test administrators are then trained on the Speaking Rubric of the WIDA Consortium (see Table 1.7.3A). Test administrators must study the rubric thoroughly to understand each of the requirements for speech, demonstrating proficiency at each of the different levels.  Speaking Rubric training is accomplished by listening to online ACCESS for ELLs Test Administrator Training speech samples. Each sample presents a task targeted at a particular proficiency level to allow test administrators to evaluate the responses against the three criteria described in the rubric for the task. Scores and rationales that are provided for each sample demonstrate how and why a particular score is assigned. To be considered certified to administer the ACCESS for ELLs Grades 1–12 Speaking test, test administrators will then need to take the Speaking test quiz that accompanies the training test module.

**Table 1.7.3A**
WIDA Consortium's Speaking Rubric

| Task Level | Linguistic Complexity | Vocabulary Usage | Language Control |
|---|---|---|---|
| **1 Entering** | Single words, set phrases, or chunks of memorized oral language | Highest frequency vocabulary from school setting and content areas | Generally comprehensible and fluent when using memorized language; communication may be significantly impeded when going beyond the highly familiar |
| **2 Emerging** | Phrases, short oral sentences | General language related to the content area; groping for vocabulary when going beyond the highly familiar is evident | Generally comprehensible and fluent when using simple discourse; communication may be impeded by groping for language structures or by phonological, syntactic, or semantic errors when going beyond phrases and short, simple sentences |
| **3 Developing** | Simple and expanded oral sentences; responses show emerging complexity used to add detail | General and some specific language related to the content area; may grope for needed vocabulary at times | Generally comprehensible and fluent when communicating in sentences; communication may from time to time be impeded by groping for language structures or by phonological, syntactic, or semantic errors, especially when attempting more complex oral discourse |
| **4 Expanding** | A variety of oral sentence lengths of varying linguistic complexity; responses show emerging cohesion used to provide detail and clarity | Specific and some technical language related to the content area; groping for needed vocabulary may be occasionally evident | Generally comprehensible and fluent at all times, though phonological, syntactic, or semantic errors that don't impede the overall meaning of the communication may appear at times; such errors may reflect first language interference |
| **5 Bridging** | A variety of sentence lengths of varying linguistic complexity in extended oral discourse; responses show cohesion and organization used to support main ideas | Technical language related to the content area; facility with needed vocabulary is evident | Approaching comparability to that of English proficient peers; errors don't impede communication and may be typical of those an English proficient peer may make |

# 2. An Assessment Use Argument for ACCESS for ELLs: Focus on Assessment Records

Validity is "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014, p. 11). Evaluations of test validity assess whether there is evidence that supports the appropriateness and adequacy of the interpretations and decisions made about test takers on the basis of their performance on a test. This chapter contextualizes the information presented in this Annual Technical Report within an argument-based approach to addressing validity (Bachman & Palmer, 2010; Chapelle, Enright, & Jamieson, 2008; Kane, 2002, 2013; Mislevy, Almond, & Lukas, 2004) for ACCESS for ELLs.

A fully developed validation framework, including an Assessment Use Argument (AUA) (Bachman & Palmer, 2010), consists of several steps (described in Section 2.1 below) that connect test design and administration to intended and actual score interpretation and consequences. This chapter begins the process of developing a complete validation framework for ACCESS for ELLs. This argument-based structure organizes the information in this Annual Technical Report to support claims about Assessment Records (i.e., test scores and proficiency level descriptions collected via ACCESS for ELLs). Specifically, tables and figures from this report are explicitly linked to questions related assessment data. Chapelle, Enright, & Jamieson (2010) support using such a structure to present information to assessment users because "based on an analysis of four points of comparison—framing the intended score interpretation, outlining the essential research, structuring research results into a validity argument, and challenging the validity argument—we conclude that an argument-based approach to validity introduces some new and useful concepts and practices" (p.3). A larger, though yet undocumented (as of 2014), validity argument for the complete assessment from its inception to its consequences is currently under development by WIDA.

The complete validity argument that will be employed to support the use of ACCESS for ELLs will show the path from test design to test taker performance to the uses and interpretations of test scores and the subsequent consequences of test use. This framework is structured around assertions, or claims, about the assessment. The claims are presented as a series of statements that connect some aspect of the assessment process to the intended purposes of the assessment. Evidence for each claim is then organized by the action that is used to ensure each claim, and it includes results from analyses of test data, outside documentation, and other resources. In the complete validation argument, this process of identifying evidence to support claims will encompass the entire testing process, from the commencement of the test design to the consequences of test use (Bachman & Palmer, 2010; Llosa, 2008); Figure 2A shows the process by which evidence supports validation actions, which are used to establish larger claims about ACCESS for ELLs.
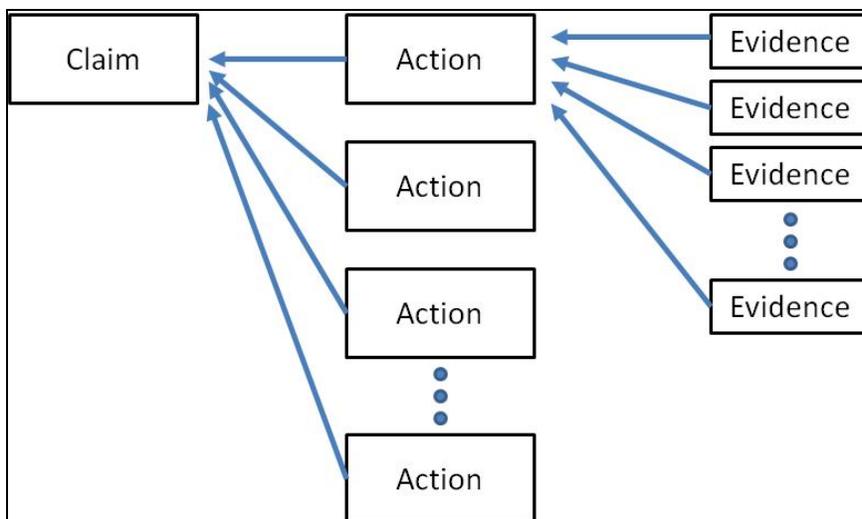
Figure 2A: General Argument Structure for Assessment Validation (simplified from Toulmin, 2003)

## *2.1     The Generic Validation Framework for ACCESS*

The generic validation framework that will be applied to the entire ACCESS for ELLs testing process was developed at the Center for Applied Linguistics (CAL) and is hereafter referred to as CAL's Validation Framework. CAL's Validation Framework, shown in Figure 2.1A, combines models for both test development (i.e., Evidence-Centered Design [Mislevy, Almond, & Lukas, 2004]) and assessment validation (i.e., Bachman and Palmer's (2010) AUA) to cover the assessment development and implementation process from initial conceptualization to the score interpretations and consequences of using the assessment. This framework constantly looks both forward and backward; for example, during the initial *Plan* step (Step 7), test developers state the anticipated decisions and consequences of implementing the assessment program, which are investigated in the *Decisions* step (Step 2) and *Consequences* step (Step 1). Because each subsequent step depends upon the strength of the step below it, the steps are numbered from 7 to 1, with Consequences being the culmination of the previous steps. This structure highlights the fact that any weakness in a lower step affects the steps above it.
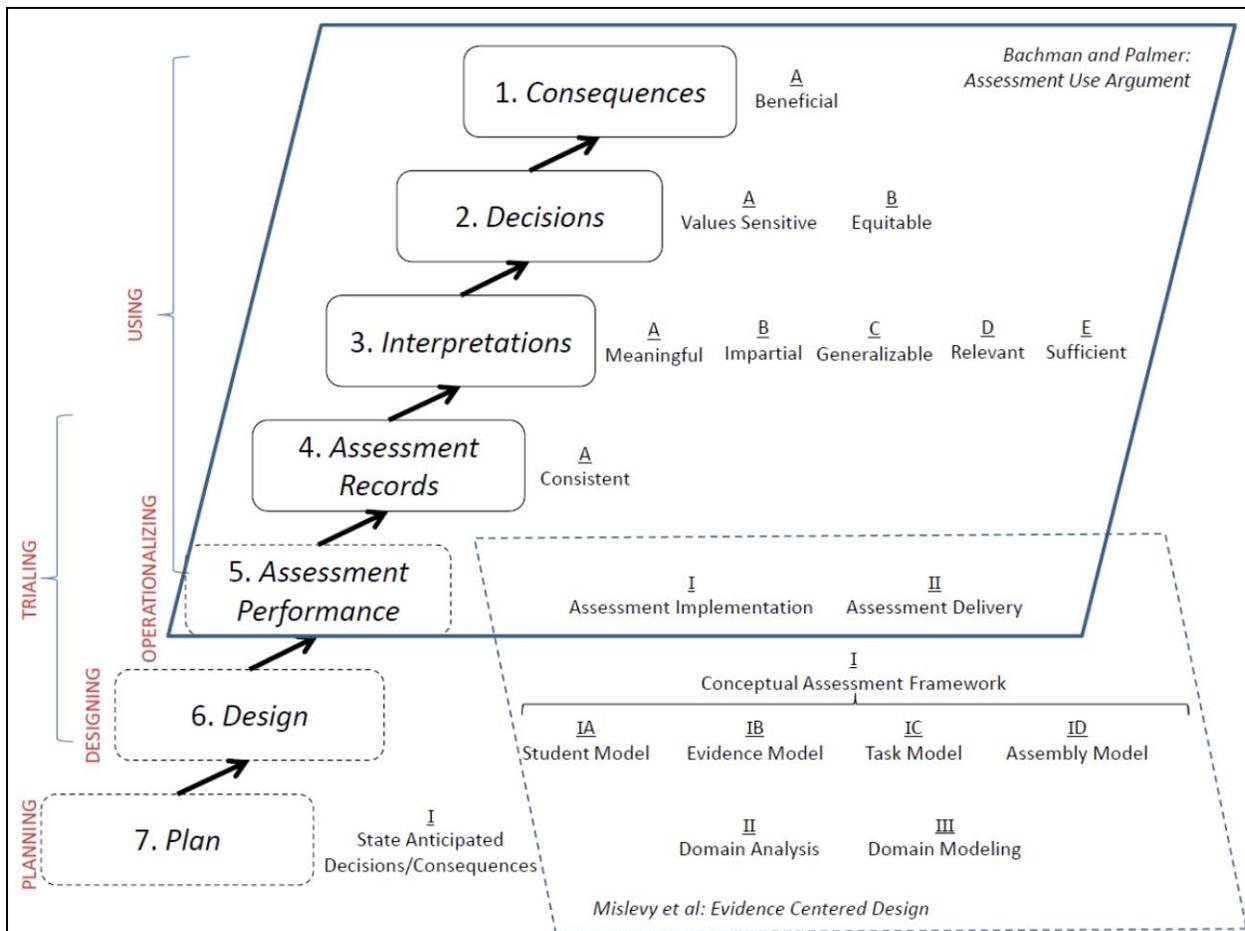
Figure 2.1A: CAL's Validation Framework (based on Bachman & Palmer, 2010; Mislevy, Almond, & Lukas, 2004)

In CAL's Validation Framework, the *Plan* step involves an examination of possible decisions states might make and consequences that might result from the assessment. This leads to the consideration of several models during the *Design* step, where specifications that answer such critical questions as "What are we measuring?" and "How do we measure it?" are developed (Mislevy, Almond, & Lukas, 2004). The subsequent steps of the validation framework highlight the trialing, implementation, and use of the assessment results, beginning with test takers' performance on the assessment (*Assessment Performance*) and continuing through the collection of test scores (*Assessment Records*), interpretations of those test scores (*Interpretations*), decisions made based on the test scores (*Decisions*), and the consequences of test use (*Consequences*).

The WIDA Consortium is using CAL's Validation Framework to present a complete validity argument, which will be updated as needed, for ACCESS for ELLs. To date, information related to Step 4, *Assessment Records*, has been explored and is found in this chapter.

## 2.2     *Focus on Assessment Records*

Although the complete validation framework for ACCESS for ELLs contains seven steps (see Figure 2.1A), the data presented in this document cover the Assessment Records step, which is part of Bachman and Palmer's (2010) AUA. By focusing on Assessment Records (i.e., test scores and proficiency level descriptions), the information in the Annual Technical Report will

be used to support claims related to the quality and consistency of the assessment data gathered and analyzed using ACCESS for ELLs. The claims in this step of the AUA all pertain to the general question "How do we know that the reported language domain scores and composite scores on ACCESS for ELLs are consistent and dependable?" Other questions about the development, administration, and outcomes of ACCESS for ELLs will be evaluated in a forthcoming document, currently in development by WIDA.

The diagram in Figure 2.2A shows a visual representation of an argument-based approach for supporting claims related to Assessment Records. The figure shows how the Assessment Records step, Step 4 of the complete validation framework, will fit in the generic validation framework and be expanded into a series of claims and corresponding actions in this chapter of the Annual Technical Report. Evidence in the form of data from this report or other sources will be presented to support these claims as they relate to ACCESS for ELLs.
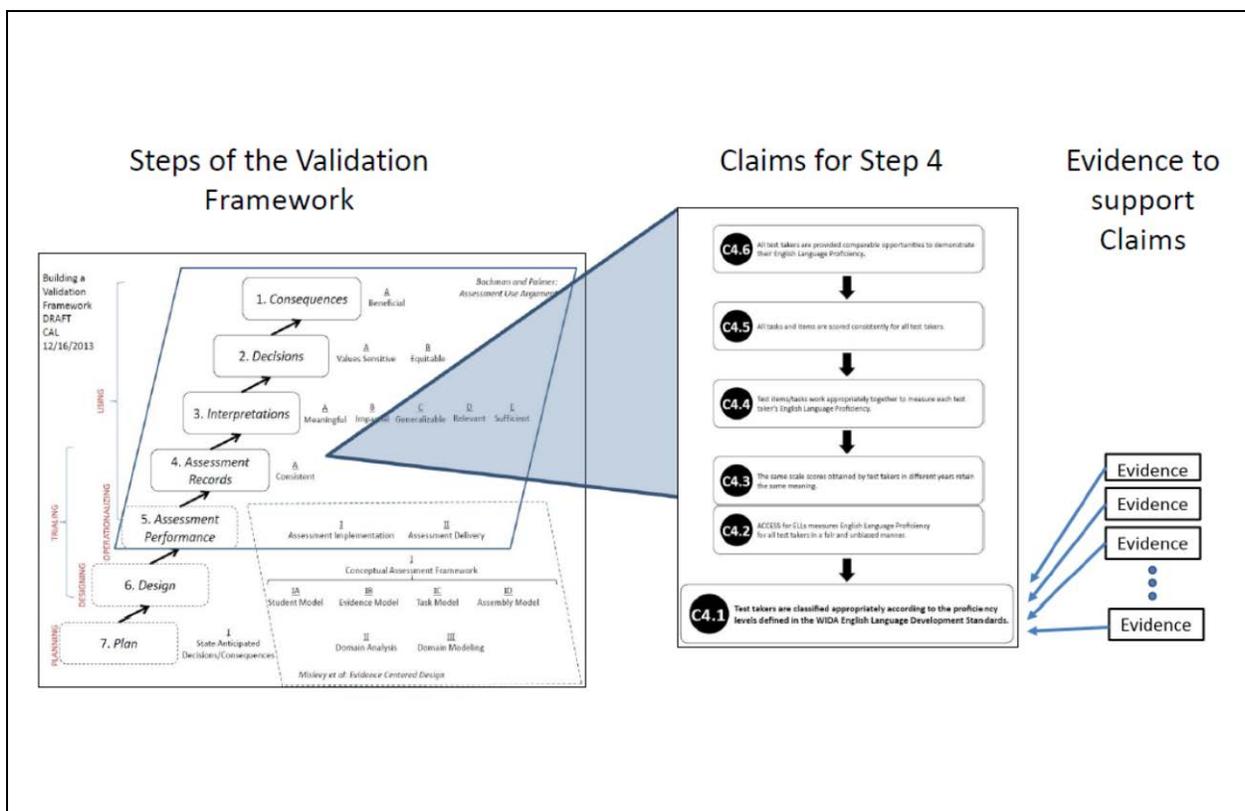


Figure 2.2A: Structure of the Argument-Based Approach Supporting Step 4 Contained in this Chapter

## 2.2.1    Breakdown of Claims for the *Assessment Records* Produced in the ACCESS for ELLs Assessment Program

The general *Assessment Records* step, Step 4 of the complete ACCESS for ELLs validation framework, is broken down into the following six claims:

C4.6.  All test takers are provided comparable opportunities to demonstrate their English Language Proficiency.
C4.5.  All tasks and items are scored consistently for all test takers.

C4.4. Test items/tasks work appropriately together to measure each test taker's English Language Proficiency.

C4.3. The same scale scores obtained by test takers in different years retain the same meaning.

C4.2. ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner.

C4.1. Test takers are classified appropriately according to the proficiency levels defined in the WIDA English Language Development Standards.

As shown in Figure 2.2.1A, these claims depend upon each other, again moving from (C4.6) up to (C4.1). Within this organizational structure, each successive claim builds upon the previous one(s) (e.g., ratings are only useful to test developers and stakeholders if all test takers are provided comparable opportunities to demonstrate their proficiency). In the next section, these claims are broken down even further into actions that are taken to ensure the consistency and reliability of the assessment records.



**C4.6** All test takers are provided comparable opportunities to demonstrate their English Language Proficiency.

**C4.5** All tasks and items are scored consistently for all test takers.

**C4.4** Test items/tasks work appropriately together to measure each test taker's English Language Proficiency.

**C4.3** The same scale scores obtained by test takers in different years retain the same meaning.

**C4.2** ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner.

**C4.1** Test takers are classified appropriately according to the proficiency levels defined in the WIDA English Language Development Standards.

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

## 2.3      *Evidence for Assessment Records Claims of ACCESS for ELLs*

In this section, evidence in the form of data or other sources (e.g., Test Administration Manuals, other information within this report, etc.) is connected to each of the *Assessment Records* claims via the actions taken to ensure those claims. This section denotes the tables, figures, and external sources that provide evidence related to each action. A summary table of the information presented in this section, including hyperlinks to the detailed description of each table or figure in Chapters 5 and 7 of this Annual Technical Report, is contained in Section 2.4. Information on how to navigate the tables and figures throughout this report is presented in Section 2.5.

Because these claims relate to Step 4 of the overall validation framework, their numbering begins with 4. The second number (after the decimal) denotes the level of the claim within Step 4. This numbering system is used in anticipation of the development of more complete documentation of a validity argument for ACCESS for ELLs, which will be completed by WIDA. Individual actions to ensure each claim are denoted by the final letter (a, b, c, and so on).

### Claim 4.6 - All test takers are provided comparable opportunities to demonstrate their English Language Proficiency.

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

*Evidence*: Procedures for administering the test and producing reported scores are documented in the ACCESS for ELLs Test Administration Manual (WIDA, 2012a).


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

*Evidence*: Test administration procedures are documented in the ACCESS for ELLs Test Administration Manual (WIDA, 2012a).

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

*Action 4.5a*: Raters of performance-based tasks undergo thorough training so that they know how to score appropriately.

*Evidence*: Section 1.7 of this report specifies the scoring procedure for ACCESS for ELLs, with Section 1.7.2 providing information on the Writing domain and Section 1.7.3 explicating the procedure for Speaking tasks. Raters of Writing tasks are trained by MetriTech to follow the Writing rubric (see Table 1.7.2B). Since Speaking tasks are scored locally, raters are trained through an online program on the WIDA website to follow the Speaking rubric (see Table 1.7.3A).


*Action 4.5b*: Listening and Reading items are scored electronically using a carefully checked key.

*Evidence*: Section 1.7 of this report specifies the scoring procedure for ACCESS for ELLs. Listening and Reading items are dichotomous and are scored electronically by MetriTech (see Section 1.7.1).

*Action 4.5c*: Raters of performance-based tasks are certified, demonstrating that they can score appropriately.

*Evidence*: Section 1.7 of this report specifies the scoring procedure for ACCESS for ELLs. Writing tasks are centrally scored at MetriTech, and all raters are pre-screened and subsequently trained (see Section 1.7.2). Speaking is scored by the test administrator after the completion of training on test administration and on the Speaking rubric (see Section 1.7.3).

*Action 4.5d*: Raters of Writing tasks are monitored daily to ensure that they are scoring appropriately.

*Evidence*: MetriTech provides Raters of Writing tasks with specially prepared calibration sets each day to monitor that the scoring rubric is being applied consistently across scoring sessions (see Section 1.7.2.1).

*Action 4.5e*: Scoring data for Writing tasks are analyzed for rater agreement to understand how closely raters agree.

*Evidence*: Interrater reliability is calculated for each of the three or four Writing tasks. The percentage of agreement between two raters is calculated in terms of three features (i.e., Linguistic Complexity, Vocabulary Usage, and Language Control). When the two raters agree on a score, this is counted as exact agreement. If the two raters provide feature scores that differ by one point, this is counted as adjacent agreement (see Table 6F for percentages of exact and adjacent agreement).

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

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

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

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

*Evidence*: A single reliability estimate, a stratified Cronbach's alpha (Cronbach, Schonemann, & McKie, 1965), is calculated across the three tiers for each domain. Cronbach's alpha indicates the extent to which items work together to measure the same construct. The stratified Cronbach's alpha is an average reliability, and it is used when test takers are administered several related subtests but are then evaluated based on a composite of those subtest scores. Table 8D presents the data used to calculate an estimate of the reliability of the composite scores using a stratified Cronbach's alpha.

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

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

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

*Action 4.3a*: A sufficient number of items and tasks are used as anchor items across adjacent years to maintain a consistent scale from year to year.

*Evidence*: Each year, while a certain percentage of items on each ACCESS for ELLs test form is refreshed, a number of items and tasks are retained from the previous year's assessment. These retained "anchor items" ensure that performances on the newer form may be interpreted in the same frame of reference as the previous year. For Listening and Reading, a majority of test items are anchor items, while one of three Writing tasks and one of three Speaking folders are retained annually as anchor tasks. Table 6E displays information on the anchor items for each test form.

*Action 4.3b*: New items and tasks are calibrated with anchor items to ensure that their difficulty measures are on the same consistent scale that is used from year to year.

*Evidence*: Both new and previously used items and tasks (i.e., anchor items) are included on each test form (see Table 6H for a list of new and anchored test items/tasks).

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

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

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

*Action 4.2a*: Differential item functioning (DIF) analyses are conducted to determine whether any items or tasks may be biased against certain subgroups.

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

The Complete Items Properties table includes more detailed information on the DIF analyses, showing the degree of measurement bias for each item and which group is favored (ATR Table 6H). Each item is categorized into three levels of DIF: A, B, or C (Zieky, 1993). An item exhibiting A level DIF shows little or no evidence of bias toward a particular group, an item exhibiting B level DIF is displays a moderate amount of bias, and an item exhibiting C level DIF is considered to display considerable evidence for potential bias and should be closely examined by test developers to identify any construct irrelevant factors that may contribute to DIF.

*Action 4.2b*: Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not used for scoring and are removed from future test forms.

*Evidence*: As described in Chapter 1.4.5 (*DIF Items*), ethnicity and gender DIF analyses are conducted using all test taker data. Information on DIF is gathered at different points in the testing cycle and is provided to the test development team. The test development team uses this information to guide the item development and review process for future items.

## Claim 4.1 - Test takers are classified appropriately according to the proficiency levels defined in the WIDA English Language Development Standards.

*Action 4.1a*: Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that ACCESS for ELLs effectively measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA English Language Development (ELD) Standards.

*Evidence*: The distribution of test takers' raw scores on ACCESS for ELLs, organized by individual test form (e.g., Reading 3–5B), shows the extent to which ACCESS for ELLs effectively measures the performance of test takers across the range of ELD abilities that each form was designed to assess (see Table 6A; see Figure 6A).

The distribution of test takers' scale scores on ACCESS for ELLs, organized by test form (e.g., Reading 3–5B), shows that ACCESS for ELLs effectively measures the performance of test takers across the range of ELD abilities that each form was designed to assess (see Table 6B; see Figure 6B).

The proficiency level distribution of test takers' scores on ACCESS for ELLs, organized by individual test form (e.g., Reading 3–5B), shows that ACCESS for ELLs effectively measures the performance of test takers across the range of proficiency levels that each form was designed to assess (see Table 6C; see Figure 6C).

The Raw Score to Proficiency Level Score table shows the interpretive proficiency level score associated with each raw score (see Table 6J). This distribution of scores shows that ACCESS for ELLs effectively measures the performance of test takers across the range of proficiency levels that each form was designed to assess.

The Test Characteristic Curve for each test form graphically shows the relationship between test takers' ability measure (which is calculated based on test performance using Rasch modeling) on the horizontal axis and the expected raw scores on the vertical axis (see Figure 6D). Five vertical lines indicate the five cut scores for the highest grade in the cluster, dividing the figure into six sections for each of the six WIDA proficiency levels. The curve shows that higher expected raw scores are required to be placed into higher language proficiency levels.


*Action 4.1b*: Distributions of scale scores and proficiency levels, organized by grade-level cluster, are analyzed to confirm that ACCESS for ELLs effectively measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA English Language Development (ELD) Standards.

*Evidence*: The distribution of test takers' scale scores on ACCESS for ELLs, organized by grade-level cluster, shows that ACCESS for ELLs effectively measures the performance of test takers across the range of ELD abilities as described by the WIDA ELD Standards (see Table 8A; see Figure 8A).

The proficiency level distribution of test takers' scores on ACCESS for ELLs, organized by grade-level cluster, shows that ACCESS for ELLs effectively measures the performance of test takers across the range of proficiency levels as defined by the WIDA ELD Standards (see Table 8B; see Figure 8B).

The Test Characteristic Curve reflects test takers' mean raw scores by domain on ACCESS for ELLs across the entire test for Kindergarten and across the three tiers for the other grade-level clusters (see Figure 8C). It also graphically illustrates how the tiers differ in difficulty, showing that ACCESS for ELLs effectively captures a range of ELD ability levels. Tier A is represented by a dotted curve, Tier B by a light solid curve, and Tier C by a dark solid curve. As shown, Tier B is more difficult than Tier A, and Tier C is more difficult than Tier B.

*Action 4.1c*: For each test form, analyses are run to confirm that English Language Proficiency is measured with high precision at the cut points pertinent to each tier.

*Evidence*: The Test Information Function graphically shows how well the test is measuring across the ability measure spectrum, which is calculated based on test performance using Rasch modeling (see Figure 6E). High values indicate more accuracy in measurement. Test forms for different tiers are designed to measure most accurately at certain proficiency levels (i.e., PL1 through PL3 for Tier A, PL2 through PL4 for Tier B, and PL3 and up for Tier C), and the expected peak of the distribution occurs within the desired range of the cut scores.

In the Raw Score to Scale Score Conversion Chart, the proficiency level associated with each raw score shows the distribution of proficiency level scores associated with each raw score/scale score for each grade in the cluster, along with the percentage of test takers in that grade who scored at that raw score/scale score/proficiency level score (see Table 6I). Additionally, this table presents the conditional standard error for each scale score, along with the upper and lower bound of the scale scores within this standard error of measurement. This value indicates how accurately or precisely the test is measuring test takers at a particular ability level by estimating the error measurement at each score point. Because there is usually more information about test takers with scores in the middle of the score distribution on each form, the conditional standard error values are usually smallest and scores are more reliable in that region of the score distribution.

*Action 4.1d*: Across domains, analyses are run to confirm that English Language Proficiency is measured with high precision at the cut points pertinent to each tier.

*Evidence*: The conditional standard error of measurement provides information on how precisely test takers' performances on ACCESS for ELLs are measured at the cut points between language proficiency levels. These cut points are critical because they are the points at which decisions are made about test taker placements. Because the cut points depend on the grade level, information for each domain is provided for each grade level within the cluster. From Table 8C, it is possible to examine how well the different tiers measure the English Language Proficiency of test takers at the appropriate proficiency level cut scores (i.e., PL1 through PL3 for Tier A, PL2 through PL4 for Tier B, and PL3 and up for Tier C).

The Test Information Function reflects the precision of measurement by graphically presenting the standard error of measurement across tiers for grade-level clusters (see Figure 8D). Tier A is represented by a dotted curve, Tier B by a light solid curve, and Tier C by a dark solid curve. As shown, Tier B is more difficult than Tier A, and Tier C is more difficult than Tier B. As in Figure C, the cut scores at the highest grade in each cluster are indicated by vertical lines. These lines make it easy to see that the test forms for different tiers measure most accurately at the proficiency levels they are meant to capture.

*Action 4.1e*: Classification and accuracy analyses are conducted by grade level to confirm that proficiency level classifications are reliable for all domain and composite scores.

*Evidence*: Information related to the accuracy of test takers' proficiency-level classifications is presented in multiple ways (see Table 8E). A separate table is provided for each grade level in a cluster. The table provides overall indices related to the accuracy and consistency of

classification. These indices indicate the percent of all test takers who would be classified into the same language proficiency level by both the administered test and either the true score distribution (accuracy) or a parallel test (consistency). Cohen's kappa, which is a statistical measure of interrater agreement between two raters that takes chance agreement between raters into account, is also presented. A kappa value of 1 indicates complete agreement between the two raters, while a kappa value of 0 indicates no agreement other than what would be expected by chance. Table 8E also shows accuracy and consistency information conditional on level and provides indices of classification accuracy and consistency at the cut points.

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

**Table 2.4A**

Summary of *Assessment Records* Claims, Actions, and Evidence

| Claim | Actions | Evidence |
|---|---|---|
| 6. All test takers are provided comparable opportunities to demonstrate their English Language Proficiency. | a. Well-specified procedures were developed for test administrators so that they are able to administer the test consistently.<br><br>b. Test administrators document and report any irregularities that may occur so that appropriate action may be taken. | a. Test Administration Manual<br><br><br>b. Test Administration Manual |
| 5. All items and tasks are scored consistently for all test takers. | a. Raters of performance-based tasks undergo thorough training so that they know how to score appropriately.<br><br>b. Listening and Reading items are scored electronically onsite at MetriTech.<br><br>c. Raters are of performance-based tasks are certified, demonstrating that they can score appropriately.<br><br>d. Raters of Writing tasks are monitored daily to ensure that they are scoring appropriately.<br><br>e. Scoring data for Writing tasks are analyzed for rater agreement to understand how closely raters agree. | a. Chapter 1.7.2 (*Scoring - Writing*); Chapter 1.7.3 (*Scoring - Speaking*)<br><br>b. Chapter 1.7.1 (*Scoring - Listening and Reading*)<br><br>c. Chapter 1.7.2 (*Scoring - Writing*); Chapter 1.7.3 (*Scoring - Speaking*)<br><br>d. Chapter 1.7.2.1 (*Scoring Procedures for Writing*)<br><br>e. Table 6F (*Reliability*) |
| 4. Test items/tasks work appropriately together to measure each test taker's English Language Proficiency. | a. For each test form (e.g., Reading 6-8B), item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.<br><br>b. For each domain and composite score across tiers, item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.<br><br>c. Analyses of Rasch model fit statistics are conducted to show that individual tasks perform appropriately | a. Table 6F (*Reliability*)<br><br><br><br><br>b. Table 8D (*Reliability*)<br><br><br><br><br><br>c. Table 6H (*Complete Item Analysis*) |

| 3. The same scale scores obtained by test takers in different years retain the same meaning. | a. A sufficient number of items and tasks are used as anchor items across adjacent years to maintain a consistent scale from year to year. | a. Table 6E (*Equating Summary*) |
| | b. New items and tasks are calibrated with anchor items to ensure that their difficulty measures are on the same consistent scale that is used from year to year. | b. Table 6D (*Scaling Equation*) |
| | c. The same scaling equation is applied from year to year to ensure that scale scores are obtained consistently over time. | c. Table 6H (*Complete Item Analysis*) |
| 2. ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner. | a. Differential item functioning (DIF) analyses are conducted to determine whether any items or tasks are biased against certain subgroups. | a. Table 6H (*Complete Item Analysis*); Table 6G (*Item/Task Analysis Summary*) |
| | b. Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not used for scoring and are removed from future test forms. | b. Chapter 1.4.5 (*DIF Items*) |
| 1. Test takers are classified appropriately according to the proficiency levels defined in the WIDA English Language Development Standards. | a. Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that ACCESS for ELLs effectively measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA English Language Development Standards. | a. Figure 6A (*Raw Scores*) & Table 6A (*Raw Score Descriptive Statistics*); Figure 6B (*Scale Scores*) & Table 6B (*Scale Score Descriptive Statistics*); Figure 6C (*Proficiency Level*) & Table 6C (*Proficiency Level Distribution*); Table 6J (*Raw Score to Proficiency Level Score Conversion Chart*); Figure 6D (*Test Characteristic Curve*) |
| | b. Distributions of scale scores and proficiency levels, organized by grade-level cluster, are analyzed to confirm that ACCESS for ELLs effectively measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA English Language Development Standards. | b. Figure 8A (*Scale Scores*) & Table 8A (*Scale Score Descriptive Statistics*); Figure 8B (*Proficiency Level*) & Table 8B (*Proficiency Level Distribution*); Figure 8C (*Test Characteristic Curve*) |
| | c. For each test form, analyses are run to confirm that English Language Proficiency is measured with high precision at the cut points pertinent to each tier. | c. Figure 6E (*Test Information Function*); Table 6I (*Raw Score to Scale Score Conversion Chart*) |
| | d. Across domains, analyses are run to confirm that English Language Proficiency is measured with high precision at the cut points pertinent to each tier. | d. Table 8C (*Conditional Standard Error of Measurement*) & Figure 8D (*Test Information Function*) |
| | e. Classification and accuracy analyses are conducted by grade-level to confirm that proficiency level classifications are reliable for all domain and composite scores. | e. Table 8E (*Accuracy and Consistency of Classification Indices*) |

## 2.5    Visual Guide to Tables and Figures

This section provides navigational support for the tables and figures contained in the ACCESS for ELLs Annual Technical Report. The Visual Guide to Tables and Figures, shown in Figures 2.5.1 through 2.5.3, serves as a resource to quickly identify which table and/or figure to look for when seeking specific information based on grade, grade-level cluster, tier, and demographic characteristics, such as state, gender, and ethnicity and race, as well as domains and domain composites.

To use the Visual Guide to Tables and Figures as a navigational tool, click on the links in Figures 2.5.1 through 2.5.3 to navigate to the selected tables and figures in the Annual Technical Report.  A link is provided at the end of each section in Chapters 4, 6, and 8. Detailed descriptions of the information in each of the tables and figures is included in the preceding chapters (i.e., Chapter 5 contains information on tables and figures in Chapter 6, and Chapter 7 contains information on tables and figures in Chapter 8). These descriptions may be accessed through links in Table 2.4A *Summary of Assessment Records Claims, Actions, and Evidence*.

Figure 2.5.1 displays the tables in Chapter 4 that provide information on participation, scale score, and proficiency level results, as well as results by standard.  The key in the upper left corner of the figure describes the tables contained in each section of the chapter. For example, tables in Section 4.1 contain information about participation. To find specific information in Chapter 4, select the Grade or Grade Cluster tab, Domain or Tier tab, and then choose from three categories: Demographic Characteristics, Domain Composites, or Domains. Within each of these categories, several additional options organize information so that individual tables can be accessed. For example, to find a table that displays information on the number of female Grade 2 students who completed the Speaking section, refer to Figure 2.5.1 and complete the following steps: one, select Grade; two, select Domains; three, select Demographic Characteristics; four, select Gender. The information is found in Table 4.2.2.2. Click on 4.2.2.2 to go to the appropriate table in Chapter 4.

 Figure 2.5.2 displays the sections in Chapter 6 that contains analyses for each ACCESS for ELLs test form by grade-level cluster, tier, and domain. The key above the figure describes specific information in each table and figure. For example, to find the Reliability table for Grade-level Cluster 9–12C in the Reading domain, refer to Figure 2.5.2 and complete the following steps: one, select Grade Cluster 9–12; two, select  Tier C; three, select Reading under Domains. Information for 9–12C Reading is shown in section 6.5.2.3. Finally, look at the key that explains that reliability information is located in table F. The result is Table 6.5.2.3F.  Click on 6.5.2.3 to go to the appropriate section, and then locate Table F.

Figure 2.5.3 displays the sections in Chapter 8 that contain analyses across tiers, organized by grade-level cluster, domain composites, and domains. The key above the figure describes the specific information in each table and figure. For example, to find the Conditional Standard Error of Measurement table for Grade-level Cluster 6–8 in the Writing domain, refer to Figure 2.5.3 and complete the following steps: one, select Grade Cluster 6–8; two, select Domain; three, select Writing. Information for 6–8 Writing is shown in section 8.5.3. Finally, look at the key and find Conditional Standard of Error Measurement table. The result is 8.5.3C.  Click on 8.5.3 to go to the appropriate section, and then locate Table C.

## 2.5.1 Chapter 4 Visual Guide to Tables and Figures

| 4.1 | Participation |
|-----|---------------|
| 4.2 | Scale Score Results |
| 4.3 | Proficiency Level Results |
| 4.4 | Results by Standard |

### Test Form Characteristics

| | | Grade | | Grade-Level Cluster | |
|---|---|---|---|---|---|
| | | Tier | Domain | Tier | Domain |
| **Demographic Characteristics** | State | | 4.1.2.1 | | 4.1.1.1 |
| | Gender | 4.2.2.2 | 4.1.2.2 | 4.1.3.3 | 4.2.1.2 · 4.1.1.2 |
| | Ethnicity and Race | 4.2.2.3 | 4.1.2.3 | 4.1.3.4 | 4.2.1.3 · 4.1.1.3 |
| **Domain Composites** | Overall | 4.3.8.2 | 4.3.8.3 | 4.3.8.1 | |
| | Oral Language | 4.3.5.2 | 4.3.5.3 | 4.3.5.1 | |
| | Literacy | 4.3.6.2 | 4.3.6.3 | 4.3.6.1 | |
| | Comprehension | 4.3.7.2 | 4.3.7.3 | 4.3.7.1 | 4.4.1.1 |
| **Domains** | Across All Domains | 4.1.3.2 | 4.2.2.1 | 4.1.3.1 | 4.2.1.1 4.2.3 |
| | Listening | 4.3.1.2 | 4.3.1.3 | 4.3.1.1 | |
| | Reading | 4.3.2.2 | 4.3.2.3 | 4.3.2.1 | |
| | Writing | 4.3.3.2 | 4.3.3.3 4.4.2.2 | 4.3.3.1 | 4.4.2.1 |
| | Speaking | 4.3.4.2 | 4.3.4.3 4.4.3.2 | 4.3.4.1 | 4.4.3.1 |

Figure 2.5.1 Chapter 4 Visual Guide to Tables and Figures

## 2.5.2 Chapter 6 Visual Guide to Tables and Figures

| | |
|---|---|
| **Table A and Figure A** | *Raw Score Information* |
| **Table B and Figure B** | *Scale Score Information* |
| **Table C and Figure C** | *Proficiency and Distribution* |
| **Table D** | *Scaling Equation Table* |
| **Table E** | *Equating Summary* |
| **Figure D** | *Test Characteristic Curve* |
| **Figure E** | *Test Information Function* |
| **Table F** | *Reliability* |
| **Table G** | *Item/Task Analysis Summary* |
| **Table H** | *Complete Item Analysis Table* |
| **Table I** | *Complete Raw Score to Scale Score Conversion Table* |
| **Table J** | *Raw Score to Proficiency Level Score Conversion* |

### Domains

| Grade-Level Cluster and Tier | | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **K** | | 6.1.1 | 6.1.2 | 6.1.3 | 6.1.4 |
| **1-2** | A | 6.2.1.1 | 6.2.2.1 | 6.2.3.1 | 6.2.4 |
| | B | 6.2.1.2 | 6.2.2.2 | 6.2.3.2 | |
| | C | 6.2.1.3 | 6.2.2.3 | 6.2.3.3 | |
| **3-5** | A | 6.3.1.1 | 6.3.2.1 | 6.3.3.1 | 6.3.4 |
| | B | 6.3.1.2 | 6.3.2.2 | 6.3.3.2 | |
| | C | 6.3.1.3 | 6.3.2.3 | 6.3.3.3 | |
| **6-8** | A | 6.4.1.1 | 6.4.2.1 | 6.4.3.1 | 6.4.4 |
| | B | 6.4.1.2 | 6.4.2.2 | 6.4.3.2 | |
| | C | 6.4.1.3 | 6.4.2.3 | 6.4.3.3 | |
| **9-12** | A | 6.5.1.1 | 6.5.2.1 | 6.5.3.1 | 6.5.4 |
| | B | 6.5.1.2 | 6.5.2.2 | 6.5.3.2 | |
| | C | 6.5.1.3 | 6.5.2.3 | 6.5.3.3 | |

Figure 2.5.2 Chapter 6 Visual Guide to Tables and Figures

## 2.5.3 Chapter 8 Visual Guide to Tables and Figures

| | |
|---|---|
| **Table A and Figure A** | *Scale Score Information* |
| **Table B and Figure B** | *Proficiency Level Information* |
| **Table C and Figures C and D** | *Conditional Standard Error Measurement* |
| **Table D** | *Reliability Information* |
| **Table E** | *Accuracy and Consistency of Classification* |

### Grade-Level Cluster

| | | K | 1-2 | 3-5 | 6-8 | 9-12 |
|---|---|---|---|---|---|---|
| **Domain Composites** | Overall | 8.1.8 | 8.2.8 | 8.3.8 | 8.4.8 | 8.5.8 |
| | Oral Language | 8.1.5 | 8.2.5 | 8.3.5 | 8.4.5 | 8.5.5 |
| | Literacy | 8.1.6 | 8.2.6 | 8.3.6 | 8.4.6 | 8.5.6 |
| | Comprehension | 8.1.7 | 8.2.7 | 8.3.7 | 8.4.7 | 8.5.7 |
| **Domains** | Listening | 8.1.1 | 8.2.1 | 8.3.1 | 8.4.1 | 8.5.1 |
| | Reading | 8.1.2 | 8.2.2 | 8.3.2 | 8.4.2 | 8.5.2 |
| | Writing | 8.1.3 | 8.2.3 | 8.3.3 | 8.4.3 | 8.5.3 |
| | Speaking | 8.1.4 | 8.2.4 | 8.3.4 | 8.4.4 | 8.5.4 |

# 3. Descriptions of Student Results

Chapter 3 provides a description of the tables that appear in Chapter 4.

## 3.1 Participation

Participation in ACCESS for ELLs is shown in three ways: grade-level cluster; grade, and tier.

### 3.1.1 Grade-Level Cluster

Section 4.1.1 gives information on participation by **grade-level cluster**.

Table 4.1.1.1 shows participation across the 33 WIDA states that participated in the operational testing program in 2013–2014. The first row shows the grade-level cluster, the next 33 rows show the number of students in that grade-level cluster who took the test, by state, and the final row shows the total number of participants across all 33 states.

Table 4.1.1.2 shows participation by cluster by gender across all 33 states combined, while Table 4.1.1.3 shows participation by cluster by ethnicity across all 33 states.

### 3.1.2 Grade

Section 4.1.2 gives similar data as in the previous section, but broken out by **grade** rather than by grade-level cluster.

### 3.1.3 Tier

Finally, Section 4.1.3 gives participation by **tier**.

Table 4.1.3.1 shows this information by cluster, tier, and domain. Because, for example, Listening in the 1–2 grade-level cluster for Tier A represents a specific test form, this table indicates how many students took each test form. Note that because Speaking is not administered by tiers, the total number shows how many took that cluster's Speaking test.

Table 4.1.3.2 shows the same information, but by grade rather than by grade-level cluster.

Table 4.1.3.3 shows the breakdown by grade-level cluster and tier for gender. When reviewing data on DIF in Chapter 6, it may be useful to refer to these tables to understand the size of the comparison groups on each form.

Table 4.1.3.4 shows the same information for ethnicity (Hispanic vs. Non-Hispanic). Consortium member states use the Census Bureau categories for student ethnicity. Again, this data may be useful when reviewing analyses of DIF in tables G and H in Chapter 6.

Note that in some circumstances there was a mismatch between a student's reported grade and the reported cluster of the test the student took (for example, a student who was reported to be in Kindergarten but who was administered a test in the 1–2 grade-level cluster). In all, 334 students were administered a test form from a cluster other than the grade in which they were reported to be. Table 3.1 below shows the number of students in each grade who were administered out-of-grade-level tests, and the test form that they were administered. The data for these students was eliminated from all analyses in this report.

**Table 3.1**

Students Excluded from Analysis due to Grade/Cluster Mismatch

| Grade/Cluster Tier | K | 1-2A | 1-2B | 1-2C | 3-5A | 3-5B | 3-5C | 6-8A | 6-8B | 6-8C | 9-12A | 9-12B | 9-12C | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| K |  | 40 | 23 | 13 | 4 | 3 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 86 |
| 1 | 0 |  |  |  | 6 | 6 | 4 | 1 | 0 | 0 | 0 | 0 | 1 | 18 |
| 2 | 0 |  |  |  | 8 | 18 | 8 | 0 | 0 | 0 | 0 | 1 | 0 | 35 |
| 3 | 0 | 20 | 37 | 18 |  |  |  | 0 | 1 | 0 | 0 | 0 | 0 | 76 |
| 4 | 0 | 3 | 3 | 2 |  |  |  | 0 | 0 | 1 | 1 | 0 | 0 | 10 |
| 5 | 0 | 3 | 2 | 1 |  |  |  | 3 | 22 | 8 | 0 | 0 | 0 | 39 |
| 6 | 0 | 0 | 1 | 0 | 13 | 12 | 8 |  |  |  | 0 | 1 | 0 | 35 |
| 7 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |  |  |  | 1 | 0 | 0 | 4 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |  |  | 2 | 9 | 3 | 14 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 9 | 3 |  |  |  | 15 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |  |  | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |  |  | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |  |  |  | 2 |
| Total | 0 | 66 | 66 | 34 | 31 | 42 | 21 | 7 | 33 | 13 | 6 | 11 | 4 | 334 |

(Note that the apparent number of Kindergarten students reported as taking tests in the higher grade-level clusters is at least in part spurious. In some states, when a grade level has not been defined for a student before the identification labels for the operational test are sent out, the "Grade" field is filled in with a placeholder of 0, the same code that is used for Kindergarten. If that information is never updated, the grade for the operational data is recorded as Kindergarten. Thus, many of the students who are reported here as Kindergarten students taking tests from higher grade-level clusters may in fact be students for whom the grade level was never defined.)

## *3.2* *Scale Score Results*

### 3.2.1 **Mean Scale Scores Across Domain and Composite Scores Section**

4.2.1 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). In this section, under each average, the number of students in each group is also given.

Table 4.2.1.1 shows mean scale scores by grade-level cluster, while Table 4.2.1.2 shows the same information broken down by gender, and Table 4.2.1.3 shows the same information broken down by ethnicity and race. In 2010, the Census Bureau introduced a new approach to reporting race and ethnicity. Previously, race and ethnicity had been a single category with six values (Hispanic, Asian/Pacific Islander/Hawaiian, Black/African American, American Indian/Alaskan Native, White - Non Hispanic, and Multi-racial/Other). Under the new approach, ethnicity has become a binary category (Hispanic or Non-Hispanic), with five categories for race (American Indian/Alaskan Native, Asian, Black/African American, Pacific Islander/Hawaiian, and White) that are not mutually exclusive. Thus, for example, Student A may be labeled as Hispanic for ethnicity and Asian for race, while Student B may be labeled as Non-Hispanic for ethnicity and both American Indian/Alaskan Native and Black/African American for race. 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 that 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 Multi-racial.

Section 4.2.2 shows the mean scale scores broken down by *grade* rather than by grade-level cluster. Table 4.2.2.1 shows mean scale scores by grade, while Table 4.2.2.2 shows the same information broken down by gender, and Table 4.2.2.3 shows the same information broken down by ethnicity and race.

### 3.2.2 **Correlations**

Tables 4.2.3A through 4.2.3E show 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. Table 4.2.3A shows the results for Kindergarten, Table 4.2.3B shows the results for the 1–2 grade-level cluster, Table 4.2.3C shows the results for the 3–5 grade-level cluster, Table 4.2.3D shows the results for the 6–8 grade-level cluster, and Table 4.2.3E shows the results for the 9–12 grade-level cluster. Beginning with Series 101, caps were placed on students taking Tier A and Tier B test forms in Listening and Reading. This capping of scores may raise the correlation between those two scores, while decreasing the correlation of those two scores with Speaking

and Writing. Note, all correlations in Tables 4.2.3A through 4.2.3E are significant at the 0.01 level (2-tailed).

## *3.3  Proficiency Level Results*

Proficiency level 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 by count and percentage:

4.3.1 – Listening

4.3.2 – Reading

4.3.3 – Writing

4.3.4 – Speaking

4.3.5 – Oral Language Composite

4.3.6 – Literacy Composite

4.3.7 – Comprehension Composite

4.3.8 – Overall Composite

Within each section, results are first presented by *grade-level cluster* and tier in Section 4.3.*.1 (note that the * indicates a subsection variable). Tables 4.3.*.1A shows the number of students who were classified into each language proficiency level, while Table 4.3.*.1B shows the percent of students (within each row) classified into each language proficiency category. These tables clearly show the effect of the capping of scores on Tier A and Tier B for Listening and Reading.

Following the presentation by tier and cluster, results are presented by *grade* and tier in Section 4.3.*.2. Again, the first table in this section shows the number of students classified into each language proficiency level, while the second table shows the results in terms of percentages within each row.

Finally, in Section 4.3.*.3, results are presented by *grade alone*, that is, without the tiers. Again, the first table shows the number of students classified into each language proficiency level, while the second table shows the results in terms of percentages within each row.

## *3.4  Mean Raw Score Results by Standards*

The tables in this section show information on mean raw score results by the five WIDA ELD Standards. These results are in terms of raw scores (i.e., the number of correct responses in Listening/Reading or Speaking or the points on the Writing rubric). Note that scores for Kindergarten students were not categorized by Standard; therefore, these tables include information only for grades 1–12.

### 3.4.1 Comprehension Composite

Section 4.4.1 shows the results for Comprehension (combined Listening and Reading items). The first section (4.4.1.1) shows results by *grade-level cluster,* while the second section (4.4.1.2) shows the results by *grade*. Within each table, the third column shows the Standard (Social and Instructional Language, Language of Language Arts, Language of Math, Language of Science,

and Language of Social Studies). The fourth column shows the maximum possible raw score by Standard, the fifth column shows the mean raw score, and the sixth column shows the mean raw score as a percentage of the maximum.

## 3.4.2 Writing

Section 4.4.2 shows the results for Writing. Again, the first section (4.4.2.1) shows results by *grade-level cluster,* while the second section (4.4.2.2) shows the results by *grade*. Within each table, the third column shows the Standard (Social and Instructional Language, Language of Language Arts/Social Studies, and Language of Mathematics/Science). The next three columns show the mean raw scores (out of a maximum of 6) of the three sub scores for the Writing test: Linguistic Complexity, Vocabulary Usage, and Language Control. The seventh column shows the total mean raw score for each Standard (out of a maximum of 18). The final column shows the mean raw score as a percentage of the maximum possible score.

## 3.4.3 Speaking

Finally, Section 4.4.3 presents the results for Speaking. As in the previous sections, the first section (4.4.3.1) shows results by *grade-level cluster*, while the second section (4.4.3.2) shows the results by *grade*. Note that the Speaking assessment itself is adaptive but not tiered. Student results are categorized here by tier according to the tier of the group-administered assessment that they took. Within each table, the third column shows the Standard (Social and Instructional Language, Language of Language Arts/Social Studies, and Language of Mathematics/Science). The fourth column shows the maximum possible score, the fifth columns shows the mean raw score, and the sixth column shows the mean raw score as a percentage of the maximum possible score.

# 4. Student Results

## 4.1 Participation

### 4.1.1 Participation by Grade-level Cluster

#### 4.1.1.1 By State

**Table 4.1.1.1**

|Participation by Cluster by State| S302

| State | Cluster | | | | | |
|---|---|---|---|---|---|---|
| | K | 1-2 | 3-5 | 6-8 | 9-12 | Total |
| AK | 1,569 | 3,560 | 4,063 | 2,924 | 2,884 | 15,000 |
| AL | 3,762 | 6,298 | 3,651 | 2,105 | 1,972 | 17,788 |
| CO | 12,525 | 25,972 | 30,932 | 21,189 | 16,576 | 107,194 |
| DC | 1,149 | 1,793 | 1,194 | 852 | 1,024 | 6,012 |
| DE | 1,792 | 2,915 | 1,854 | 888 | 923 | 8,372 |
| GA | 17,917 | 33,137 | 23,045 | 10,822 | 8,223 | 93,144 |
| HI | 2,456 | 3,947 | 3,777 | 2,985 | 3,477 | 16,642 |
| IL | 30,799 | 60,513 | 45,318 | 22,284 | 17,475 | 176,389 |
| KY | 3,280 | 6,244 | 5,038 | 2,898 | 2,712 | 20,172 |
| MA | 10,003 | 18,876 | 19,031 | 12,608 | 13,678 | 74,196 |
| MD | 10,110 | 17,921 | 13,922 | 7,573 | 7,919 | 57,445 |
| ME | 522 | 1,021 | 1,273 | 1,161 | 1,280 | 5,257 |
| MI | 10,503 | 19,024 | 21,483 | 14,936 | 15,522 | 81,468 |
| MN | 8,608 | 16,237 | 18,901 | 13,223 | 11,123 | 68,092 |
| MO | 4,618 | 7,788 | 7,423 | 4,399 | 3,556 | 27,784 |
| MP | 63 | 255 | 543 | 534 | 222 | 1,617 |
| MS | 1,312 | 2,391 | 2,376 | 1,357 | 1,034 | 8,470 |
| MT | 321 | 790 | 1,135 | 697 | 478 | 3,421 |
| NC | 14,123 | 27,576 | 25,140 | 15,990 | 13,632 | 96,461 |
| ND | 418 | 718 | 830 | 684 | 764 | 3,414 |
| NH | 407 | 1,085 | 1,092 | 715 | 897 | 4,196 |
| NJ | 11,768 | 19,239 | 12,843 | 7,940 | 10,816 | 62,606 |
| NM | 6,433 | 13,595 | 15,266 | 11,187 | 8,919 | 55,400 |
| NV | 9,663 | 20,134 | 20,690 | 12,804 | 7,866 | 71,157 |
| OK | 7,291 | 13,057 | 11,260 | 6,797 | 5,275 | 43,680 |
| PA | 4,754 | 10,926 | 12,025 | 9,989 | 11,995 | 49,689 |
| RI | 1,228 | 2,464 | 2,589 | 1,646 | 1,869 | 9,796 |
| SD | 662 | 1,107 | 1,157 | 826 | 853 | 4,605 |
| UT | 5,559 | 10,916 | 8,207 | 5,718 | 5,526 | 35,926 |
| VA | 14,803 | 27,739 | 23,153 | 13,731 | 15,813 | 95,239 |
| VT | 196 | 366 | 392 | 227 | 352 | 1,533 |
| WI | 5,792 | 11,932 | 13,497 | 9,005 | 7,616 | 47,842 |
| WY | 422 | 726 | 770 | 484 | 397 | 2,799 |
| Total | 204,828 | 390,262 | 353,870 | 221,178 | 202,668 | 1,372,806 |

## 4.1.1.2    By Gender

**Table 4.1.1.2**

|Participation by Cluster by Gender S302

| Cluster | | Gender | | | Total |
|---------|---|---|---|---|---|
| | | **F** | **M** | **Missing** | **Total** |
| K | Count | 96,381 | 107,416 | 1,031 | 204,828 |
| | % within Cluster | 47.1% | 52.4% | 0.5% | 100.0% |
| 1-2 | Count | 183,929 | 203,920 | 2,413 | 390,262 |
| | % within Cluster | 47.1% | 52.3% | 0.6% | 100.0% |
| 3-5 | Count | 161,237 | 188,972 | 3,661 | 353,870 |
| | % within Cluster | 45.6% | 53.4% | 1.0% | 100.0% |
| 6-8 | Count | 97,779 | 121,107 | 2,292 | 221,178 |
| | % within Cluster | 44.2% | 54.8% | 1.0% | 100.0% |
| 9-12 | Count | 89,173 | 111,305 | 2,190 | 202,668 |
| | % within Cluster | 44.0% | 54.9% | 1.1% | 100.0% |
| Total | Count | 628,499 | 732,720 | 11,587 | 1,372,806 |
| | % within Cluster | 45.8% | 53.4% | 0.8% | 100.0% |

## 4.1.1.3    By Ethnicity

**Table 4.1.1.3**

|Participation by Cluster by Ethnicity S302

| Cluster | | Hispanic/Non-Hispanic | | | Total |
|---------|---|---|---|---|---|
| | | **Hispanic** | **Other** | **Missing** | **Total** |
| K | Count | 137,915 | 63,871 | 3,042 | 204,828 |
| | % within Cluster | 67.3% | 31.2% | 1.5% | 100.0% |
| 1-2 | Count | 269,735 | 115,964 | 4,563 | 390,262 |
| | % within Cluster | 69.1% | 29.7% | 1.2% | 100.0% |
| 3-5 | Count | 242,064 | 107,037 | 4,769 | 353,870 |
| | % within Cluster | 68.4% | 30.2% | 1.3% | 100.0% |
| 6-8 | Count | 146,739 | 70,573 | 3,866 | 221,178 |
| | % within Cluster | 66.3% | 31.9% | 1.7% | 100.0% |
| 9-12 | Count | 121,693 | 76,619 | 4,356 | 202,668 |
| | % within Cluster | 60.0% | 37.8% | 2.1% | 100.0% |
| Total | Count | 918,146 | 434,064 | 20,596 | 1,372,806 |
| | % within Cluster | 66.9% | 31.6% | 1.5% | 100.0% |

## 4.1.2    Participation by Grade

### 4.1.2.1    By State

**Table 4.1.2.1**

Participation by Grade by State| S302

| State | K | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | **Grade** | | | | | | |
| **AK** | 1,569 | 1,797 | 1,763 | 1,614 | 1,374 | 1,075 | 1,005 | 1,032 | 887 | 1,012 | 720 | 573 | 579 | 15,000 |
| **AL** | 3,762 | 3,434 | 2,864 | 2,075 | 924 | 652 | 614 | 706 | 785 | 944 | 467 | 352 | 209 | 17,788 |
| **CO** | 12,525 | 12,914 | 13,058 | 12,212 | 9,801 | 8,919 | 7,658 | 7,246 | 6,285 | 5,529 | 4,152 | 3,391 | 3,504 | 107,194 |
| **DC** | 1,149 | 994 | 799 | 559 | 356 | 279 | 292 | 253 | 307 | 560 | 197 | 123 | 144 | 6,012 |
| **DE** | 1,792 | 1,624 | 1,291 | 1,018 | 484 | 352 | 299 | 286 | 303 | 465 | 225 | 135 | 98 | 8,372 |
| **GA** | 17,917 | 17,589 | 15,548 | 12,871 | 5,947 | 4,227 | 3,540 | 3,781 | 3,501 | 4,418 | 1,841 | 1,120 | 844 | 93,144 |
| **HI** | 2,456 | 2,074 | 1,873 | 1,895 | 997 | 885 | 819 | 1,026 | 1,140 | 1,602 | 799 | 560 | 516 | 16,642 |
| **IL** | 30,799 | 31,002 | 29,511 | 25,508 | 11,678 | 8,132 | 6,917 | 7,782 | 7,585 | 8,053 | 4,111 | 3,116 | 2,195 | 176,389 |
| **KY** | 3,280 | 3,298 | 2,946 | 2,475 | 1,479 | 1,084 | 936 | 1,033 | 929 | 1,085 | 738 | 504 | 385 | 20,172 |
| **MA** | 10,003 | 9,742 | 9,134 | 7,895 | 5,946 | 5,190 | 4,379 | 4,316 | 3,913 | 5,018 | 3,435 | 3,005 | 2,220 | 74,196 |
| **MD** | 10,110 | 9,430 | 8,491 | 7,266 | 3,763 | 2,893 | 2,554 | 2,682 | 2,337 | 3,991 | 2,022 | 999 | 907 | 57,445 |
| **ME** | 522 | 544 | 477 | 505 | 423 | 345 | 347 | 414 | 400 | 394 | 331 | 288 | 267 | 5,257 |
| **MI** | 10,503 | 9,732 | 9,292 | 8,191 | 7,150 | 6,142 | 5,577 | 4,941 | 4,418 | 4,646 | 4,480 | 3,325 | 3,071 | 81,468 |
| **MN** | 8,608 | 8,265 | 7,972 | 7,470 | 6,259 | 5,172 | 4,660 | 4,530 | 4,033 | 3,969 | 2,676 | 2,464 | 2,014 | 68,092 |
| **MO** | 4,618 | 3,992 | 3,796 | 3,315 | 2,276 | 1,832 | 1,528 | 1,481 | 1,390 | 1,336 | 972 | 705 | 543 | 27,784 |
| **MP** | 63 | 115 | 140 | 276 | 134 | 133 | 206 | 165 | 163 | 151 | 33 | 18 | 20 | 1,617 |
| **MS** | 1,312 | 1,282 | 1,109 | 1,040 | 766 | 570 | 516 | 457 | 384 | 462 | 300 | 168 | 104 | 8,470 |
| **MT** | 321 | 362 | 428 | 407 | 410 | 318 | 262 | 229 | 206 | 187 | 130 | 94 | 67 | 3,421 |
| **NC** | 14,123 | 13,968 | 13,608 | 12,951 | 6,782 | 5,407 | 5,222 | 5,383 | 5,385 | 6,728 | 3,284 | 2,071 | 1,549 | 96,461 |
| **ND** | 418 | 389 | 329 | 373 | 246 | 211 | 246 | 236 | 202 | 263 | 189 | 175 | 137 | 3,414 |
| **NH** | 407 | 511 | 574 | 593 | 276 | 223 | 228 | 234 | 253 | 349 | 227 | 179 | 142 | 4,196 |
| **NJ** | 11,768 | 10,475 | 8,764 | 6,459 | 3,679 | 2,705 | 2,513 | 2,650 | 2,777 | 3,615 | 3,120 | 2,421 | 1,660 | 62,606 |
| **NM** | 6,433 | 6,967 | 6,628 | 6,253 | 5,139 | 3,874 | 3,681 | 3,769 | 3,737 | 3,885 | 2,245 | 1,573 | 1,216 | 55,400 |
| **NV** | 9,663 | 10,261 | 9,873 | 9,172 | 6,414 | 5,104 | 4,676 | 4,516 | 3,612 | 2,824 | 1,876 | 1,648 | 1,518 | 71,157 |
| **O K** | 7,291 | 6,997 | 6,060 | 5,355 | 3,327 | 2,578 | 2,353 | 2,229 | 2,215 | 2,409 | 1,261 | 917 | 688 | 43,680 |
| **PA** | 4,754 | 5,396 | 5,530 | 4,834 | 3,821 | 3,370 | 3,310 | 3,404 | 3,275 | 3,907 | 3,246 | 2,561 | 2,281 | 49,689 |
| **RI** | 1,228 | 1,338 | 1,126 | 1,098 | 839 | 652 | 555 | 565 | 526 | 637 | 534 | 391 | 307 | 9,796 |
| **SD** | 662 | 580 | 527 | 539 | 329 | 289 | 276 | 268 | 282 | 334 | 233 | 163 | 123 | 4,605 |
| **UT** | 5,559 | 5,918 | 4,998 | 2,869 | 2,817 | 2,521 | 2,100 | 1,834 | 1,784 | 1,543 | 1,615 | 1,363 | 1,005 | 35,926 |
| **VA** | 14,803 | 14,438 | 13,301 | 11,730 | 6,747 | 4,676 | 4,220 | 4,596 | 4,915 | 7,228 | 3,836 | 3,142 | 1,607 | 95,239 |
| **VT** | 196 | 193 | 173 | 202 | 100 | 90 | 70 | 77 | 80 | 108 | 89 | 92 | 63 | 1,533 |
| **WI** | 5,792 | 6,027 | 5,905 | 5,416 | 4,750 | 3,331 | 2,750 | 3,175 | 3,080 | 3,184 | 1,706 | 1,476 | 1,250 | 47,842 |
| **WY** | 422 | 368 | 358 | 402 | 221 | 147 | 158 | 168 | 158 | 148 | 115 | 68 | 66 | 2,799 |
| **Total** | 204,828 | 202,016 | 188,246 | 164,838 | 105,654 | 83,378 | 74,467 | 75,464 | 71,247 | 80,984 | 51,205 | 39,180 | 31,299 | 1,372,806 |

Return to Summary of Assessment Records Claims, Actions, and Evidence

### 4.1.2.2    By Gender

**Table 4.1.2.2**

Participation by Grade by Gender S302

| Grade | | Gender | | | Total |
|---|---|---|---|---|---|
| | | **F** | **M** | **Missing** | **Total** |
| K | Count | 96,381 | 107,416 | 1,031 | 204,828 |
| | % within Grade | 47.1% | 52.4% | 0.5% | 100.0% |
| 1 | Count | 95,146 | 105,468 | 1,402 | 202,016 |
| | % within Grade | 47.1% | 52.2% | 0.7% | 100.0% |
| 2 | Count | 88,783 | 98,452 | 1,011 | 188,246 |
| | % within Grade | 47.2% | 52.3% | 0.5% | 100.0% |
| 3 | Count | 76,740 | 86,591 | 1,507 | 164,838 |
| | % within Grade | 46.6% | 52.5% | 0.9% | 100.0% |
| 4 | Count | 47,511 | 57,027 | 1,116 | 105,654 |
| | % within Grade | 45.0% | 54.0% | 1.1% | 100.0% |
| 5 | Count | 36,986 | 45,354 | 1,038 | 83,378 |
| | % within Grade | 44.4% | 54.4% | 1.2% | 100.0% |
| 6 | Count | 32,757 | 40,644 | 1,066 | 74,467 |
| | % within Grade | 44.0% | 54.6% | 1.4% | 100.0% |
| 7 | Count | 33,339 | 41,475 | 650 | 75,464 |
| | % within Grade | 44.2% | 55.0% | 0.9% | 100.0% |
| 8 | Count | 31,683 | 38,988 | 576 | 71,247 |
| | % within Grade | 44.5% | 54.7% | 0.8% | 100.0% |
| 9 | Count | 34,392 | 45,683 | 909 | 80,984 |
| | % within Grade | 42.5% | 56.4% | 1.1% | 100.0% |
| 10 | Count | 22,430 | 28,241 | 534 | 51,205 |
| | % within Grade | 43.8% | 55.2% | 1.0% | 100.0% |
| 11 | Count | 17,700 | 21,075 | 405 | 39,180 |
| | % within Grade | 45.2% | 53.8% | 1.0% | 100.0% |
| 12 | Count | 14,651 | 16,306 | 342 | 31,299 |
| | % within Grade | 46.8% | 52.1% | 1.1% | 100.0% |
| Total | Count | 628,499 | 732,720 | 11,587 | 1,372,806 |
| | % within Grade | 45.8% | 53.4% | 0.8% | 100.0% |

### 4.1.2.3 By Ethnicity

**Table 4.1.2.3**

Participation by Grade by Ethnicity S302

| Grade | | Hispanic/Non-Hispanic | | | Total |
|---|---|---|---|---|---|
| | | **Hispanic** | **Other** | **Missing** | |
| K | Count | 137,915 | 63,871 | 3,042 | 204,828 |
| | % within Grade | 67.3% | 31.2% | 1.5% | 100.0% |
| 1 | Count | 139,362 | 60,097 | 2,557 | 202,016 |
| | % within Grade | 69.0% | 29.7% | 1.3% | 100.0% |
| 2 | Count | 130,373 | 55,867 | 2,006 | 188,246 |
| | % within Grade | 69.3% | 29.7% | 1.1% | 100.0% |
| 3 | Count | 114,418 | 48,304 | 2,116 | 164,838 |
| | % within Grade | 69.4% | 29.3% | 1.3% | 100.0% |
| 4 | Count | 71,817 | 32,386 | 1,451 | 105,654 |
| | % within Grade | 68.0% | 30.7% | 1.4% | 100.0% |
| 5 | Count | 55,829 | 26,347 | 1,202 | 83,378 |
| | % within Grade | 67.0% | 31.6% | 1.4% | 100.0% |
| 6 | Count | 49,476 | 23,694 | 1,297 | 74,467 |
| | % within Grade | 66.4% | 31.8% | 1.7% | 100.0% |
| 7 | Count | 50,397 | 23,798 | 1,269 | 75,464 |
| | % within Grade | 66.8% | 31.5% | 1.7% | 100.0% |
| 8 | Count | 46,866 | 23,081 | 1,300 | 71,247 |
| | % within Grade | 65.8% | 32.4% | 1.8% | 100.0% |
| 9 | Count | 52,380 | 26,592 | 2,012 | 80,984 |
| | % within Grade | 64.7% | 32.8% | 2.5% | 100.0% |
| 10 | Count | 30,577 | 19,653 | 975 | 51,205 |
| | % within Grade | 59.7% | 38.4% | 1.9% | 100.0% |
| 11 | Count | 21,793 | 16,603 | 784 | 39,180 |
| | % within Grade | 55.6% | 42.4% | 2.0% | 100.0% |
| 12 | Count | 16,943 | 13,771 | 585 | 31,299 |
| | % within Grade | 54.1% | 44.0% | 1.9% | 100.0% |
| Total | Count | 918,146 | 434,064 | 20,596 | 1,372,806 |
| | % within Grade | 66.9% | 31.6% | 1.5% | 100.0% |

## 4.1.3 Participation by Tier

### 4.1.3.1 By Cluster by Domain (Test Form)

**Table 4.1.3.1**

Participation by Cluster by Tier by Domain S302

| Cluster | | | Domain | | | |
|---------|------|---|-----------|---------|---------|----------|
| | | | Listening | Reading | Writing | Speaking |
| K | Tier | - | 204,694 | 204,696 | 204,695 | 204,695 |
| 1-2 | Tier | A | 85,814 | 85,814 | 85,796 | 85,767 |
| | | B | 194,463 | 194,455 | 194,419 | 194,369 |
| | | C | 109,741 | 109,727 | 109,703 | 109,720 |
| | Total | | 390,018 | 389,996 | 389,918 | 389,856 |
| 3-5 | Tier | A | 29,801 | 29,798 | 29,776 | 29,783 |
| | | B | 149,400 | 149,393 | 149,366 | 149,346 |
| | | C | 174,440 | 174,438 | 174,398 | 174,378 |
| | Total | | 353,641 | 353,629 | 353,540 | 353,507 |
| 6-8 | Tier | A | 24,231 | 24,234 | 24,227 | 24,211 |
| | | B | 82,801 | 82,798 | 82,761 | 82,710 |
| | | C | 113,712 | 113,714 | 113,682 | 113,611 |
| | Total | | 220,744 | 220,746 | 220,670 | 220,532 |
| 9-12 | Tier | A | 29,646 | 29,662 | 29,636 | 29,613 |
| | | B | 76,791 | 76,841 | 76,740 | 76,687 |
| | | C | 93,536 | 93,623 | 93,462 | 93,366 |
| | Total | | 199,973 | 200,126 | 199,838 | 199,666 |

## 4.1.3.2 By Grade by Domain (Test Form)

**Table 4.1.3.2**

Participation by Grade by Tier by Domain S302

| Grade | | | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|---|
| | | | **Domain** | | | |
| K | Tier | - | 204,694 | 204,696 | 204,695 | 204,695 |
| 1 | Tier | A | 67,720 | 67,723 | 67,709 | 67,677 |
| | | B | 91,390 | 91,381 | 91,365 | 91,336 |
| | | C | 42,772 | 42,760 | 42,756 | 42,764 |
| | Total | | 201,882 | 201,864 | 201,830 | 201,777 |
| 2 | Tier | A | 18,094 | 18,091 | 18,087 | 18,090 |
| | | B | 103,073 | 103,074 | 103,054 | 103,033 |
| | | C | 66,969 | 66,967 | 66,947 | 66,956 |
| | Total | | 188,136 | 188,132 | 188,088 | 188,079 |
| 3 | Tier | A | 13,017 | 13,012 | 12,999 | 13,006 |
| | | B | 75,144 | 75,143 | 75,128 | 75,123 |
| | | C | 76,578 | 76,575 | 76,564 | 76,555 |
| | Total | | 164,739 | 164,730 | 164,691 | 164,684 |
| 4 | Tier | A | 8,979 | 8,979 | 8,974 | 8,970 |
| | | B | 42,670 | 42,666 | 42,663 | 42,656 |
| | | C | 53,939 | 53,941 | 53,931 | 53,929 |
| | Total | | 105,588 | 105,586 | 105,568 | 105,555 |
| 5 | Tier | A | 7,805 | 7,807 | 7,803 | 7,807 |
| | | B | 31,586 | 31,584 | 31,575 | 31,567 |
| | | C | 43,923 | 43,922 | 43,903 | 43,894 |
| | Total | | 83,314 | 83,313 | 83,281 | 83,268 |
| 6 | Tier | A | 8,119 | 8,119 | 8,112 | 8,107 |
| | | B | 28,182 | 28,181 | 28,170 | 28,154 |
| | | C | 38,065 | 38,068 | 38,054 | 38,033 |
| | Total | | 74,366 | 74,368 | 74,336 | 74,294 |
| 7 | Tier | A | 8,020 | 8,021 | 8,021 | 8,021 |
| | | B | 28,312 | 28,314 | 28,296 | 28,285 |
| | | C | 38,976 | 38,983 | 38,971 | 38,951 |
| | Total | | 75,308 | 75,318 | 75,288 | 75,257 |
| 8 | Tier | A | 8,092 | 8,094 | 8,094 | 8,083 |
| | | B | 26,307 | 26,303 | 26,295 | 26,271 |
| | | C | 36,671 | 36,663 | 36,657 | 36,627 |
| | Total | | 71,070 | 71,060 | 71,046 | 70,981 |
| 9 | Tier | A | 15,914 | 15,918 | 15,906 | 15,894 |
| | | B | 28,723 | 28,755 | 28,715 | 28,700 |
| | | C | 35,485 | 35,508 | 35,465 | 35,419 |
| | Total | | 80,122 | 80,181 | 80,086 | 80,013 |
| 10 | Tier | A | 7,297 | 7,297 | 7,290 | 7,293 |
| | | B | 20,068 | 20,075 | 20,059 | 20,034 |
| | | C | 23,285 | 23,298 | 23,274 | 23,233 |
| | Total | | 50,650 | 50,670 | 50,623 | 50,560 |
| 11 | Tier | A | 4,103 | 4,111 | 4,110 | 4,107 |
| | | B | 15,695 | 15,702 | 15,686 | 15,679 |
| | | C | 18,846 | 18,870 | 18,826 | 18,810 |
| | Total | | 38,644 | 38,683 | 38,622 | 38,596 |
| 12 | Tier | A | 2,332 | 2,336 | 2,330 | 2,319 |
| | | B | 12,305 | 12,309 | 12,280 | 12,274 |
| | | C | 15,920 | 15,947 | 15,897 | 15,904 |
| | Total | | 30,557 | 30,592 | 30,507 | 30,497 |

Return to Summary of Assessment Records Claims, Actions, and Evidence

### 4.1.3.3    By Cluster by Gender

**Table 4.1.3.3**

|Participation by Cluster by Tier by Gender S302

| Cluster | Tier | | Gender | | | Total |
|---|---|---|---|---|---|---|
| | | | **F** | **M** | **Missing** | |
| K | - | Count | 96,381 | 107,416 | 1,031 | 204,828 |
| | | % within Tier | 47.1% | 52.4% | 0.5% | 100.0% |
| 1-2 | A | Count | 38,715 | 46,380 | 793 | 85,888 |
| | | % within Tier | 45.1% | 54.0% | 0.9% | 100.0% |
| | B | Count | 91,147 | 102,564 | 868 | 194,579 |
| | | % within Tier | 46.8% | 52.7% | 0.4% | 100.0% |
| | C | Count | 54,067 | 54,976 | 752 | 109,795 |
| | | % within Tier | 49.2% | 50.1% | 0.7% | 100.0% |
| 3-5 | A | Count | 13,280 | 15,948 | 611 | 29,839 |
| | | % within Tier | 44.5% | 53.4% | 2.0% | 100.0% |
| | B | Count | 66,075 | 82,278 | 1,148 | 149,501 |
| | | % within Tier | 44.2% | 55.0% | 0.8% | 100.0% |
| | C | Count | 81,882 | 90,746 | 1,902 | 174,530 |
| | | % within Tier | 46.9% | 52.0% | 1.1% | 100.0% |
| 6-8 | A | Count | 10,783 | 12,947 | 541 | 24,271 |
| | | % within Tier | 44.4% | 53.3% | 2.2% | 100.0% |
| | B | Count | 35,626 | 46,894 | 492 | 83,012 |
| | | % within Tier | 42.9% | 56.5% | 0.6% | 100.0% |
| | C | Count | 51,370 | 61,266 | 1,259 | 113,895 |
| | | % within Tier | 45.1% | 53.8% | 1.1% | 100.0% |
| 9-12 | A | Count | 12,368 | 17,012 | 559 | 29,939 |
| | | % within Tier | 41.3% | 56.8% | 1.9% | 100.0% |
| | B | Count | 34,152 | 43,080 | 607 | 77,839 |
| | | % within Tier | 43.9% | 55.3% | 0.8% | 100.0% |
| | C | Count | 42,653 | 51,213 | 1,024 | 94,890 |
| | | % within Tier | 44.9% | 54.0% | 1.1% | 100.0% |

### 4.1.3.4    By Cluster by Ethnicity

**Table 4.1.3.4**

Participation by Cluster by Tier by Ethnicity S302

| Cluster | Tier | | Hispanic/Non-Hispanic | | | Total |
| | | | Hispanic | Other | Missing | |
|---|---|---|---|---|---|---|
| K | - | Count | 137,915 | 63,871 | 3,042 | 204,828 |
| | | % within Tier | 67.3% | 31.2% | 1.5% | 100.0% |
| 1-2 | A | Count | 62,323 | 22,281 | 1,284 | 85,888 |
| | | % within Tier | 72.6% | 25.9% | 1.5% | 100.0% |
| | B | Count | 139,218 | 53,049 | 2,312 | 194,579 |
| | | % within Tier | 71.5% | 27.3% | 1.2% | 100.0% |
| | C | Count | 68,194 | 40,634 | 967 | 109,795 |
| | | % within Tier | 62.1% | 37.0% | 0.9% | 100.0% |
| 3-5 | A | Count | 17,358 | 11,394 | 1,087 | 29,839 |
| | | % within Tier | 58.2% | 38.2% | 3.6% | 100.0% |
| | B | Count | 103,927 | 43,607 | 1,967 | 149,501 |
| | | % within Tier | 69.5% | 29.2% | 1.3% | 100.0% |
| | C | Count | 120,779 | 52,036 | 1,715 | 174,530 |
| | | % within Tier | 69.2% | 29.8% | 1.0% | 100.0% |
| 6-8 | A | Count | 14,277 | 8,886 | 1,108 | 24,271 |
| | | % within Tier | 58.8% | 36.6% | 4.6% | 100.0% |
| | B | Count | 52,345 | 29,285 | 1,382 | 83,012 |
| | | % within Tier | 63.1% | 35.3% | 1.7% | 100.0% |
| | C | Count | 80,117 | 32,402 | 1,376 | 113,895 |
| | | % within Tier | 70.3% | 28.4% | 1.2% | 100.0% |
| 9-12 | A | Count | 18,745 | 9,953 | 1,241 | 29,939 |
| | | % within Tier | 62.6% | 33.2% | 4.1% | 100.0% |
| | B | Count | 43,832 | 32,425 | 1,582 | 77,839 |
| | | % within Tier | 56.3% | 41.7% | 2.0% | 100.0% |
| | C | Count | 59,116 | 34,241 | 1,533 | 94,890 |
| | | % within Tier | 62.3% | 36.1% | 1.6% | 100.0% |

## 4.2 Scale Score Results

## 4.2.1 Mean Scale Scores by Grade-level Cluster Across Domain and Composite Scores

### 4.2.1.1 By Cluster

**Table 4.2.1.1**

Mean Scale Scores by Cluster S302

| Cluster | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---------|------|---------|---------|---------|---------|---------|---------|---------|---------|
| K | Mean | 269.99 | 192.18 | 210.63 | 302.51 | 286.48 | 201.65 | 215.52 | 226.90 |
| | N | 203,841 | 203,853 | 203,840 | 203,827 | 203,823 | 203,837 | 203,837 | 203,809 |
| 1-2 | Mean | 311.79 | 296.51 | 276.02 | 346.38 | 329.35 | 286.56 | 301.17 | 299.16 |
| | N | 388,944 | 388,669 | 388,786 | 388,802 | 388,584 | 388,428 | 388,557 | 388,056 |
| 3-5 | Mean | 357.86 | 338.60 | 345.45 | 357.86 | 358.12 | 342.28 | 344.52 | 346.83 |
| | N | 352,289 | 351,942 | 352,025 | 352,105 | 351,940 | 351,617 | 351,832 | 351,277 |
| 6-8 | Mean | 384.46 | 358.42 | 354.70 | 372.24 | 378.64 | 356.82 | 366.31 | 363.16 |
| | N | 219,478 | 219,320 | 219,327 | 219,219 | 218,979 | 219,028 | 219,163 | 218,525 |
| 9-12 | Mean | 385.46 | 375.73 | 397.54 | 380.05 | 383.05 | 386.92 | 378.73 | 385.57 |
| | N | 197,907 | 198,037 | 197,712 | 197,615 | 196,703 | 197,316 | 197,563 | 196,059 |

#### 4.2.1.2    By Cluster by Gender

**Table 4.2.1.2**

Mean Scale Scores by Cluster by Gender S302

| Cluster | Gender | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| K | F | Mean | 273.79 | 195.70 | 216.67 | 306.29 | 290.26 | 206.43 | 219.11 | 231.37 |
| | | N | 95,911 | 95,914 | 95,908 | 95,907 | 95,905 | 95,908 | 95,909 | 95,899 |
| | M | Mean | 266.61 | 189.15 | 205.37 | 299.21 | 283.14 | 197.51 | 212.39 | 223.00 |
| | | N | 106,900 | 106,908 | 106,901 | 106,890 | 106,888 | 106,898 | 106,898 | 106,880 |
| | Missing | Mean | 266.77 | 179.18 | 194.14 | 293.99 | 280.58 | 186.87 | 205.50 | 214.87 |
| | | N | 1,030 | 1,031 | 1,031 | 1,030 | 1,030 | 1,031 | 1,030 | 1,030 |
| 1-2 | F | Mean | 312.83 | 297.86 | 278.65 | 347.66 | 330.50 | 288.54 | 302.43 | 300.89 |
| | | N | 183,366 | 183,264 | 183,316 | 183,309 | 183,210 | 183,162 | 183,214 | 182,999 |
| | M | Mean | 310.86 | 295.32 | 273.65 | 345.29 | 328.34 | 284.78 | 300.05 | 297.61 |
| | | N | 203,321 | 203,147 | 203,213 | 203,231 | 203,117 | 203,010 | 203,086 | 202,802 |
| | Missing | Mean | 310.75 | 293.74 | 275.94 | 341.41 | 326.30 | 285.17 | 298.90 | 297.30 |
| | | N | 2,257 | 2,258 | 2,257 | 2,262 | 2,257 | 2,256 | 2,257 | 2,255 |
| 3-5 | F | Mean | 358.23 | 339.55 | 348.75 | 357.95 | 358.35 | 344.40 | 345.30 | 348.38 |
| | | N | 160,656 | 160,531 | 160,559 | 160,575 | 160,502 | 160,405 | 160,477 | 160,255 |
| | M | Mean | 357.38 | 337.69 | 342.54 | 357.75 | 357.83 | 340.37 | 343.73 | 345.41 |
| | | N | 188,280 | 188,057 | 188,112 | 188,174 | 188,085 | 187,858 | 188,002 | 187,669 |
| | Missing | Mean | 366.94 | 344.45 | 350.64 | 358.95 | 363.17 | 347.80 | 351.33 | 352.19 |
| | | N | 3,353 | 3,354 | 3,354 | 3,356 | 3,353 | 3,354 | 3,353 | 3,353 |
| 6-8 | F | Mean | 385.45 | 359.81 | 358.61 | 370.65 | 378.33 | 359.47 | 367.57 | 364.91 |
| | | N | 97,139 | 97,090 | 97,087 | 97,028 | 96,933 | 96,982 | 97,021 | 96,773 |
| | M | Mean | 383.62 | 357.29 | 351.56 | 373.52 | 378.86 | 354.69 | 365.27 | 361.74 |
| | | N | 120,355 | 120,248 | 120,262 | 120,223 | 120,078 | 120,070 | 120,161 | 119,785 |
| | Missing | Mean | 386.54 | 358.95 | 353.62 | 372.89 | 380.05 | 356.57 | 367.29 | 363.40 |
| | | N | 1,984 | 1,982 | 1,978 | 1,968 | 1,968 | 1,976 | 1,981 | 1,967 |
| 9-12 | F | Mean | 385.68 | 377.49 | 401.52 | 378.83 | 382.54 | 389.77 | 380.03 | 387.41 |
| | | N | 87,303 | 87,377 | 87,260 | 87,128 | 86,770 | 87,110 | 87,183 | 86,543 |
| | M | Mean | 385.45 | 374.31 | 394.41 | 381.19 | 383.62 | 384.67 | 377.75 | 384.17 |
| | | N | 108,752 | 108,806 | 108,604 | 108,639 | 108,100 | 108,358 | 108,530 | 107,686 |
| | Missing | Mean | 376.29 | 375.38 | 393.34 | 370.58 | 373.78 | 384.63 | 375.71 | 381.11 |
| | | N | 1,852 | 1,854 | 1,848 | 1,848 | 1,833 | 1,848 | 1,850 | 1,830 |

### 4.2.1.3 By Cluster by Ethnicity

**Table 4.2.1.3**

Mean Scale Scores by Cluster by Ethnicity S302

| Cluster | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| K | Non-Hispanic Asian | Mean | 280.02 | 219.99 | 233.95 | 309.54 | 295.01 | 227.24 | 237.98 | 247.36 |
| | | N | 27,047 | 27,046 | 27,044 | 27,047 | 27,046 | 27,042 | 27,045 | 27,042 |
| | Non-Hispanic Pacific Islander | Mean | 260.69 | 182.03 | 202.85 | 303.97 | 282.56 | 192.68 | 205.66 | 219.45 |
| | | N | 1,891 | 1,892 | 1,892 | 1,891 | 1,891 | 1,892 | 1,891 | 1,891 |
| | Non-Hispanic Black | Mean | 277.75 | 205.70 | 223.69 | 317.81 | 298.02 | 214.95 | 227.31 | 239.67 |
| | | N | 8,944 | 8,943 | 8,943 | 8,942 | 8,942 | 8,943 | 8,943 | 8,941 |
| | Hispanic (Of Any Race) | Mean | 265.40 | 184.23 | 203.60 | 298.12 | 281.99 | 194.16 | 208.57 | 220.31 |
| | | N | 137,415 | 137,423 | 137,413 | 137,404 | 137,402 | 137,413 | 137,414 | 137,395 |
| | Non-Hispanic American Indian | Mean | 269.89 | 181.82 | 192.49 | 300.37 | 285.35 | 187.39 | 208.23 | 216.58 |
| | | N | 3,539 | 3,540 | 3,540 | 3,539 | 3,538 | 3,540 | 3,539 | 3,538 |
| | Non-Hispanic Multi-racial | Mean | 304.73 | 223.73 | 238.55 | 336.70 | 320.95 | 231.40 | 248.11 | 258.14 |
| | | N | 967 | 969 | 969 | 966 | 966 | 969 | 967 | 966 |
| | Non-Hispanic White | Mean | 284.48 | 204.71 | 225.05 | 315.90 | 300.41 | 215.14 | 228.63 | 240.51 |
| | | N | 20,220 | 20,220 | 20,219 | 20,220 | 20,220 | 20,218 | 20,220 | 20,218 |
| | Missing | Mean | 264.83 | 190.13 | 204.94 | 296.52 | 280.88 | 197.77 | 212.55 | 222.53 |
| | | N | 3,818 | 3,820 | 3,820 | 3,818 | 3,818 | 3,820 | 3,818 | 3,818 |
| 1-2 | Non-Hispanic Asian | Mean | 316.21 | 304.57 | 284.22 | 349.09 | 332.91 | 294.69 | 308.13 | 305.91 |
| | | N | 48,036 | 48,015 | 48,009 | 48,000 | 47,982 | 47,984 | 48,004 | 47,940 |
| | Non-Hispanic Pacific Islander | Mean | 305.25 | 292.39 | 277.01 | 341.87 | 323.86 | 284.97 | 296.36 | 296.42 |
| | | N | 3,467 | 3,460 | 3,464 | 3,459 | 3,458 | 3,457 | 3,458 | 3,449 |
| | Non-Hispanic Black | Mean | 312.16 | 297.97 | 277.24 | 351.51 | 332.10 | 287.90 | 302.27 | 300.92 |
| | | N | 17,362 | 17,349 | 17,354 | 17,358 | 17,347 | 17,336 | 17,343 | 17,321 |
| | Hispanic (Of Any Race) | Mean | 310.61 | 294.56 | 274.06 | 344.70 | 327.92 | 284.60 | 299.45 | 297.36 |
| | | N | 269,122 | 268,915 | 269,035 | 269,048 | 268,896 | 268,763 | 268,840 | 268,519 |
| | Non-Hispanic American Indian | Mean | 308.16 | 292.64 | 271.65 | 340.18 | 324.46 | 282.45 | 297.36 | 294.84 |
| | | N | 7,091 | 7,085 | 7,073 | 7,089 | 7,079 | 7,066 | 7,079 | 7,050 |
| | Non-Hispanic Multi-racial | Mean | 322.42 | 305.90 | 282.07 | 363.99 | 343.49 | 294.27 | 310.90 | 308.83 |
| | | N | 1,675 | 1,676 | 1,676 | 1,674 | 1,672 | 1,676 | 1,674 | 1,671 |
| | Non-Hispanic White | Mean | 316.37 | 301.00 | 279.99 | 354.45 | 335.66 | 290.79 | 305.67 | 304.02 |
| | | N | 36,073 | 36,055 | 36,065 | 36,060 | 36,047 | 36,041 | 36,046 | 36,013 |
| | Missing | Mean | 306.01 | 292.61 | 273.84 | 341.87 | 324.22 | 283.53 | 296.68 | 295.50 |
| | | N | 6,118 | 6,114 | 6,110 | 6,114 | 6,103 | 6,105 | 6,113 | 6,093 |

| Cluster | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 3-5 | Non-Hispanic Asian | Mean | 363.03 | 344.70 | 350.87 | 355.15 | 359.34 | 348.03 | 350.35 | 351.23 |
| | | N | 42,185 | 42,169 | 42,157 | 42,168 | 42,156 | 42,139 | 42,159 | 42,113 |
| | Non-Hispanic Pacific Islander | Mean | 352.54 | 335.61 | 346.15 | 352.08 | 352.54 | 341.14 | 340.85 | 344.39 |
| | | N | 3,453 | 3,455 | 3,448 | 3,452 | 3,450 | 3,448 | 3,450 | 3,443 |
| | Non-Hispanic Black | Mean | 358.28 | 337.81 | 344.21 | 359.38 | 359.09 | 341.27 | 344.08 | 346.43 |
| | | N | 18,777 | 18,735 | 18,753 | 18,768 | 18,757 | 18,715 | 18,728 | 18,700 |
| | Hispanic (Of Any Race) | Mean | 356.72 | 337.43 | 344.56 | 357.99 | 357.61 | 341.25 | 343.35 | 345.96 |
| | | N | 241,418 | 241,163 | 241,268 | 241,288 | 241,168 | 240,955 | 241,089 | 240,697 |
| | Non-Hispanic American Indian | Mean | 350.03 | 332.90 | 338.50 | 350.71 | 350.65 | 336.00 | 338.18 | 340.19 |
| | | N | 8,534 | 8,533 | 8,514 | 8,525 | 8,517 | 8,508 | 8,528 | 8,495 |
| | Non-Hispanic Multi-racial | Mean | 370.69 | 345.57 | 350.27 | 371.96 | 371.59 | 348.15 | 353.25 | 354.98 |
| | | N | 1,248 | 1,249 | 1,249 | 1,248 | 1,248 | 1,249 | 1,248 | 1,248 |
| | Non-Hispanic White | Mean | 362.43 | 342.22 | 348.31 | 363.27 | 363.10 | 345.52 | 348.42 | 350.59 |
| | | N | 29,727 | 29,694 | 29,704 | 29,713 | 29,706 | 29,673 | 29,689 | 29,659 |
| | Missing | Mean | 355.38 | 336.32 | 341.72 | 351.58 | 353.75 | 339.28 | 342.19 | 343.44 |
| | | N | 6,947 | 6,944 | 6,932 | 6,943 | 6,938 | 6,930 | 6,941 | 6,922 |
| 6-8 | Non-Hispanic Asian | Mean | 386.68 | 364.63 | 359.36 | 367.24 | 377.23 | 362.26 | 371.31 | 366.54 |
| | | N | 27,373 | 27,375 | 27,360 | 27,362 | 27,348 | 27,347 | 27,360 | 27,315 |
| | Non-Hispanic Pacific Islander | Mean | 377.38 | 353.63 | 352.95 | 369.36 | 373.67 | 353.56 | 360.88 | 359.43 |
| | | N | 2,453 | 2,455 | 2,449 | 2,459 | 2,449 | 2,446 | 2,451 | 2,441 |
| | Non-Hispanic Black | Mean | 383.88 | 358.59 | 353.26 | 373.13 | 378.82 | 356.19 | 366.24 | 362.77 |
| | | N | 14,241 | 14,210 | 14,235 | 14,235 | 14,216 | 14,198 | 14,200 | 14,172 |
| | Hispanic (Of Any Race) | Mean | 384.71 | 357.20 | 353.98 | 373.33 | 379.30 | 355.85 | 365.52 | 362.67 |
| | | N | 145,998 | 145,874 | 145,893 | 145,778 | 145,623 | 145,682 | 145,773 | 145,319 |
| | Non-Hispanic American Indian | Mean | 376.67 | 353.72 | 351.37 | 370.17 | 373.70 | 352.83 | 360.67 | 358.88 |
| | | N | 6,619 | 6,615 | 6,600 | 6,600 | 6,583 | 6,591 | 6,608 | 6,552 |
| | Non-Hispanic Multi-racial | Mean | 393.59 | 366.17 | 358.96 | 380.69 | 387.34 | 362.84 | 374.45 | 369.98 |
| | | N | 647 | 647 | 647 | 645 | 644 | 645 | 647 | 642 |
| | Non-Hispanic White | Mean | 386.38 | 362.38 | 357.45 | 375.50 | 381.22 | 360.17 | 369.66 | 366.27 |
| | | N | 17,472 | 17,466 | 17,465 | 17,466 | 17,454 | 17,451 | 17,455 | 17,431 |
| | Missing | Mean | 371.73 | 353.11 | 349.05 | 356.19 | 364.31 | 351.35 | 358.81 | 355.07 |
| | | N | 4,675 | 4,678 | 4,678 | 4,674 | 4,662 | 4,668 | 4,669 | 4,653 |

| Cluster | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 9-12 | Non-Hispanic Asian | Mean | 388.72 | 381.22 | 401.80 | 376.92 | 383.11 | 391.78 | 383.57 | 389.00 |
| | | N | 31,283 | 31,292 | 31,289 | 31,250 | 31,177 | 31,249 | 31,243 | 31,112 |
| | Non-Hispanic Pacific Islander | Mean | 383.89 | 372.63 | 398.52 | 386.64 | 385.67 | 385.94 | 376.20 | 385.78 |
| | | N | 2,384 | 2,399 | 2,381 | 2,385 | 2,354 | 2,377 | 2,377 | 2,338 |
| | Non-Hispanic Black | Mean | 382.64 | 375.85 | 396.96 | 379.04 | 381.12 | 386.68 | 377.96 | 384.81 |
| | | N | 17,062 | 17,082 | 17,055 | 17,047 | 16,985 | 17,026 | 17,027 | 16,922 |
| | Hispanic (Of Any Race) | Mean | 384.07 | 373.93 | 396.34 | 379.90 | 382.28 | 385.43 | 377.06 | 384.29 |
| | | N | 118,990 | 119,066 | 118,831 | 118,812 | 118,170 | 118,551 | 118,764 | 117,734 |
| | Non-Hispanic American Indian | Mean | 390.97 | 372.10 | 398.23 | 384.52 | 388.03 | 385.49 | 377.83 | 386.04 |
| | | N | 5,021 | 5,014 | 4,995 | 4,971 | 4,947 | 4,985 | 5,009 | 4,926 |
| | Non-Hispanic Multi-racial | Mean | 396.45 | 383.40 | 403.18 | 391.39 | 394.19 | 393.79 | 387.31 | 393.65 |
| | | N | 605 | 606 | 606 | 606 | 604 | 604 | 605 | 602 |
| | Non-Hispanic White | Mean | 394.43 | 380.87 | 400.85 | 390.81 | 392.88 | 391.12 | 385.03 | 391.45 |
| | | N | 17,496 | 17,499 | 17,501 | 17,488 | 17,446 | 17,475 | 17,476 | 17,421 |
| | Missing | Mean | 370.57 | 369.91 | 388.00 | 360.18 | 365.76 | 379.27 | 370.18 | 375.04 |
| | | N | 5,066 | 5,079 | 5,054 | 5,056 | 5,020 | 5,049 | 5,062 | 5,004 |

## 4.2.2    Mean Scale Scores by Grade Across Domain and Composite Scores

### 4.2.2.1    By Grade

**Table 4.2.2.1**
Mean Scale Scores by Grade S302

| Grade | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|
| K | Mean | 269.99 | 192.18 | 210.63 | 302.51 | 286.48 | 201.65 | 215.52 | 226.90 |
| | N | 203,841 | 203,853 | 203,840 | 203,827 | 203,823 | 203,837 | 203,837 | 203,809 |
| 1 | Mean | 299.11 | 283.48 | 266.82 | 337.37 | 318.47 | 275.44 | 288.17 | 288.11 |
| | N | 201,336 | 201,177 | 201,249 | 201,244 | 201,132 | 201,059 | 201,119 | 200,853 |
| 2 | Mean | 325.40 | 310.49 | 285.90 | 356.05 | 341.02 | 298.49 | 315.11 | 311.00 |
| | N | 187,608 | 187,492 | 187,537 | 187,558 | 187,452 | 187,369 | 187,438 | 187,203 |
| 3 | Mean | 349.43 | 331.68 | 340.82 | 355.70 | 352.84 | 336.47 | 337.18 | 341.21 |
| | N | 164,211 | 164,072 | 164,071 | 164,141 | 164,066 | 163,916 | 164,030 | 163,789 |
| 4 | Mean | 360.74 | 340.84 | 346.85 | 358.06 | 359.63 | 344.14 | 346.95 | 348.57 |
| | N | 105,153 | 105,032 | 105,084 | 105,094 | 105,043 | 104,940 | 104,998 | 104,822 |
| 5 | Mean | 370.90 | 349.48 | 352.81 | 361.87 | 366.65 | 351.43 | 355.99 | 355.77 |
| | N | 82,925 | 82,838 | 82,870 | 82,870 | 82,831 | 82,761 | 82,804 | 82,666 |
| 6 | Mean | 377.12 | 351.16 | 349.34 | 369.61 | 373.63 | 350.51 | 359.00 | 357.25 |
| | N | 73,891 | 73,817 | 73,833 | 73,805 | 73,741 | 73,722 | 73,777 | 73,582 |
| 7 | Mean | 385.15 | 358.84 | 355.17 | 372.59 | 379.15 | 357.27 | 366.83 | 363.63 |
| | N | 74,881 | 74,844 | 74,840 | 74,819 | 74,722 | 74,744 | 74,773 | 74,562 |
| 8 | Mean | 391.41 | 365.56 | 359.80 | 374.63 | 383.33 | 362.95 | 373.39 | 368.85 |
| | N | 70,706 | 70,659 | 70,654 | 70,595 | 70,516 | 70,562 | 70,613 | 70,381 |
| 9 | Mean | 381.43 | 372.59 | 394.61 | 376.09 | 379.05 | 383.90 | 375.34 | 382.26 |
| | N | 79,568 | 79,602 | 79,499 | 79,452 | 79,128 | 79,349 | 79,449 | 78,900 |
| 10 | Mean | 385.08 | 374.91 | 396.88 | 379.28 | 382.45 | 386.19 | 378.04 | 384.87 |
| | N | 50,140 | 50,148 | 50,096 | 50,051 | 49,875 | 49,998 | 50,057 | 49,719 |
| 11 | Mean | 389.92 | 379.46 | 400.95 | 383.57 | 387.04 | 390.50 | 382.67 | 389.27 |
| | N | 38,200 | 38,245 | 38,170 | 38,164 | 37,978 | 38,092 | 38,136 | 37,857 |
| 12 | Mean | 391.12 | 380.63 | 402.05 | 387.37 | 389.60 | 391.62 | 383.88 | 390.85 |
| | N | 29,999 | 30,042 | 29,947 | 29,948 | 29,722 | 29,877 | 29,921 | 29,583 |

### 4.2.2.2    By Grade by Gender

**Table 4.2.2.2**

Mean Scale Scores by Grade by Gender S302

| Scale | Gender | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| K | F | Mean | 273.79 | 195.70 | 216.67 | 306.29 | 290.26 | 206.43 | 219.11 | 231.37 |
| | | N | 95,911 | 95,914 | 95,908 | 95,907 | 95,905 | 95,908 | 95,909 | 95,899 |
| | M | Mean | 266.61 | 189.15 | 205.37 | 299.21 | 283.14 | 197.51 | 212.39 | 223.00 |
| | | N | 106,900 | 106,908 | 106,901 | 106,890 | 106,888 | 106,898 | 106,898 | 106,880 |
| | Missing | Mean | 266.77 | 179.18 | 194.14 | 293.99 | 280.58 | 186.87 | 205.50 | 214.87 |
| | | N | 1,030 | 1,031 | 1,031 | 1,030 | 1,030 | 1,031 | 1,030 | 1,030 |
| 1 | F | Mean | 300.26 | 284.74 | 269.20 | 338.87 | 319.78 | 277.26 | 289.40 | 289.78 |
| | | N | 94,848 | 94,786 | 94,822 | 94,804 | 94,761 | 94,740 | 94,763 | 94,651 |
| | M | Mean | 298.05 | 282.36 | 264.64 | 336.08 | 317.30 | 273.79 | 287.07 | 286.61 |
| | | N | 105,143 | 105,045 | 105,082 | 105,095 | 105,026 | 104,974 | 105,011 | 104,858 |
| | Missing | Mean | 300.17 | 282.78 | 268.96 | 333.35 | 316.97 | 276.20 | 288.02 | 288.22 |
| | | N | 1,345 | 1,346 | 1,345 | 1,345 | 1,345 | 1,345 | 1,345 | 1,344 |
| 2 | F | Mean | 326.30 | 311.92 | 288.78 | 357.07 | 341.98 | 300.63 | 316.38 | 312.80 |
| | | N | 88,518 | 88,478 | 88,494 | 88,505 | 88,449 | 88,422 | 88,451 | 88,348 |
| | M | Mean | 324.59 | 309.20 | 283.31 | 355.15 | 340.17 | 296.55 | 313.95 | 309.39 |
| | | N | 98,178 | 98,102 | 98,131 | 98,136 | 98,091 | 98,036 | 98,075 | 97,944 |
| | Missing | Mean | 326.34 | 309.90 | 286.25 | 353.23 | 340.06 | 298.40 | 314.94 | 310.69 |
| | | N | 912 | 912 | 912 | 917 | 912 | 911 | 912 | 911 |
| 3 | F | Mean | 350.02 | 332.69 | 344.19 | 356.44 | 353.50 | 338.65 | 338.06 | 342.92 |
| | | N | 76,520 | 76,477 | 76,474 | 76,506 | 76,471 | 76,426 | 76,457 | 76,382 |
| | M | Mean | 348.72 | 330.67 | 337.74 | 355.03 | 352.15 | 334.43 | 336.26 | 339.58 |
| | | N | 86,303 | 86,207 | 86,209 | 86,247 | 86,207 | 86,102 | 86,185 | 86,019 |
| | Missing | Mean | 361.38 | 338.73 | 346.53 | 356.50 | 359.20 | 342.85 | 345.67 | 347.56 |
| | | N | 1,388 | 1,388 | 1,388 | 1,388 | 1,388 | 1,388 | 1,388 | 1,388 |
| 4 | F | Mean | 361.36 | 341.95 | 350.31 | 357.85 | 359.83 | 346.42 | 347.92 | 350.24 |
| | | N | 47,331 | 47,277 | 47,303 | 47,297 | 47,277 | 47,237 | 47,260 | 47,181 |
| | M | Mean | 360.17 | 339.86 | 343.90 | 358.21 | 359.43 | 342.17 | 346.09 | 347.13 |
| | | N | 56,806 | 56,739 | 56,765 | 56,779 | 56,750 | 56,687 | 56,722 | 56,625 |
| | Missing | Mean | 363.37 | 343.82 | 350.89 | 359.94 | 361.85 | 347.64 | 349.86 | 351.66 |
| | | N | 1,016 | 1,016 | 1,016 | 1,018 | 1,016 | 1,016 | 1,016 | 1,016 |
| 5 | F | Mean | 371.29 | 350.74 | 356.20 | 361.25 | 366.54 | 353.75 | 356.99 | 357.37 |
| | | N | 36,805 | 36,777 | 36,782 | 36,772 | 36,754 | 36,742 | 36,760 | 36,692 |
| | M | Mean | 370.42 | 348.36 | 349.98 | 362.38 | 366.66 | 349.46 | 355.06 | 354.40 |
| | | N | 45,171 | 45,111 | 45,138 | 45,148 | 45,128 | 45,069 | 45,095 | 45,025 |
| | Missing | Mean | 378.88 | 353.49 | 356.38 | 361.46 | 370.41 | 355.21 | 361.19 | 359.52 |
| | | N | 949 | 950 | 950 | 950 | 949 | 950 | 949 | 949 |

| Scale | Gender | | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | F | Mean | | 378.22 | 352.40 | 353.39 | 368.13 | 373.43 | 353.14 | 360.20 | 359.02 |
| | | N | | 32,547 | 32,517 | 32,525 | 32,509 | 32,488 | 32,484 | 32,500 | 32,430 |
| | M | Mean | | 376.10 | 350.12 | 346.04 | 370.79 | 373.72 | 348.34 | 357.97 | 355.76 |
| | | N | | 40,406 | 40,362 | 40,374 | 40,368 | 40,325 | 40,305 | 40,340 | 40,224 |
| | Missing | Mean | | 382.36 | 353.33 | 350.79 | 369.65 | 376.45 | 352.38 | 362.11 | 359.42 |
| | | N | | 938 | 938 | 934 | 928 | 928 | 933 | 937 | 928 |
| 7 | F | Mean | | 386.23 | 360.43 | 359.25 | 370.85 | 378.82 | 360.10 | 368.27 | 365.50 |
| | | N | | 33,119 | 33,117 | 33,109 | 33,099 | 33,056 | 33,078 | 33,083 | 33,001 |
| | M | Mean | | 384.22 | 357.52 | 351.88 | 373.95 | 379.37 | 354.97 | 365.62 | 362.08 |
| | | N | | 41,209 | 41,175 | 41,179 | 41,171 | 41,117 | 41,114 | 41,138 | 41,012 |
| | Missing | Mean | | 389.13 | 362.39 | 356.24 | 376.44 | 382.95 | 359.56 | 370.49 | 366.32 |
| | | N | | 553 | 552 | 552 | 549 | 549 | 552 | 552 | 549 |
| 8 | F | Mean | | 392.11 | 366.82 | 363.35 | 373.05 | 382.88 | 365.35 | 374.47 | 370.38 |
| | | N | | 31,473 | 31,456 | 31,453 | 31,420 | 31,389 | 31,420 | 31,438 | 31,342 |
| | M | Mean | | 390.84 | 364.54 | 356.96 | 375.91 | 383.68 | 361.02 | 372.51 | 367.61 |
| | | N | | 38,740 | 38,711 | 38,709 | 38,684 | 38,636 | 38,651 | 38,683 | 38,549 |
| | Missing | Mean | | 391.56 | 365.82 | 356.05 | 375.04 | 383.62 | 361.17 | 373.58 | 367.67 |
| | | N | | 493 | 492 | 492 | 491 | 491 | 491 | 492 | 490 |
| 9 | F | Mean | | 382.87 | 374.90 | 399.50 | 376.52 | 379.97 | 387.47 | 377.38 | 385.04 |
| | | N | | 33,845 | 33,864 | 33,827 | 33,768 | 33,642 | 33,777 | 33,802 | 33,559 |
| | M | Mean | | 380.60 | 370.90 | 391.05 | 376.02 | 378.60 | 381.29 | 373.91 | 380.31 |
| | | N | | 44,902 | 44,916 | 44,853 | 44,866 | 44,675 | 44,753 | 44,827 | 44,530 |
| | Missing | Mean | | 367.80 | 370.44 | 387.62 | 362.12 | 365.32 | 379.32 | 369.70 | 374.84 |
| | | N | | 821 | 822 | 819 | 818 | 811 | 819 | 820 | 811 |
| 10 | F | Mean | | 384.26 | 376.33 | 400.52 | 377.13 | 380.96 | 388.70 | 378.78 | 386.16 |
| | | N | | 22,040 | 22,051 | 22,028 | 21,987 | 21,933 | 21,997 | 22,017 | 21,887 |
| | M | Mean | | 385.79 | 373.73 | 394.00 | 381.07 | 383.71 | 384.17 | 377.44 | 383.84 |
| | | N | | 27,671 | 27,668 | 27,639 | 27,635 | 27,516 | 27,572 | 27,612 | 27,407 |
| | Missing | Mean | | 381.71 | 378.02 | 396.34 | 374.10 | 378.14 | 387.44 | 379.20 | 384.49 |
| | | N | | 429 | 429 | 429 | 429 | 426 | 429 | 428 | 425 |
| 11 | F | Mean | | 389.25 | 380.68 | 404.05 | 381.08 | 385.44 | 392.64 | 383.32 | 390.28 |
| | | N | | 17,308 | 17,331 | 17,310 | 17,300 | 17,215 | 17,273 | 17,285 | 17,169 |
| | M | Mean | | 390.61 | 378.44 | 398.37 | 385.76 | 388.50 | 388.71 | 382.16 | 388.47 |
| | | N | | 20,566 | 20,587 | 20,536 | 20,536 | 20,440 | 20,495 | 20,525 | 20,367 |
| | Missing | Mean | | 382.32 | 379.29 | 398.59 | 377.27 | 380.24 | 389.18 | 380.22 | 386.14 |
| | | N | | 326 | 327 | 324 | 328 | 323 | 324 | 326 | 321 |
| 12 | F | Mean | | 390.21 | 381.60 | 404.80 | 384.26 | 387.61 | 393.47 | 384.28 | 391.53 |
| | | N | | 14,110 | 14,131 | 14,095 | 14,073 | 13,980 | 14,063 | 14,079 | 13,928 |
| | M | Mean | | 392.03 | 379.74 | 399.60 | 390.26 | 391.50 | 389.96 | 383.54 | 390.27 |
| | | N | | 15,613 | 15,635 | 15,576 | 15,602 | 15,469 | 15,538 | 15,566 | 15,382 |
| | Missing | Mean | | 386.00 | 381.37 | 399.52 | 382.33 | 384.47 | 390.69 | 382.83 | 388.58 |
| | | N | | 276 | 276 | 276 | 273 | 273 | 276 | 276 | 273 |

## 4.2.2.3  By Grade by Ethnicity

**Table 4.2.2.3**
Mean Scale Scores by Grade by Ethnicity S302

| Grade | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| K | Non-Hispanic Asian | Mean | 280.02 | 219.99 | 233.95 | 309.54 | 295.01 | 227.24 | 237.98 | 247.36 |
| | | N | 27,047 | 27,046 | 27,044 | 27,047 | 27,046 | 27,042 | 27,045 | 27,042 |
| | Non-Hispanic Pacific Islander | Mean | 260.69 | 182.03 | 202.85 | 303.97 | 282.56 | 192.68 | 205.66 | 219.45 |
| | | N | 1,891 | 1,892 | 1,892 | 1,891 | 1,891 | 1,892 | 1,891 | 1,891 |
| | Non-Hispanic Black | Mean | 277.75 | 205.70 | 223.69 | 317.81 | 298.02 | 214.95 | 227.31 | 239.67 |
| | | N | 8,944 | 8,943 | 8,943 | 8,942 | 8,942 | 8,943 | 8,943 | 8,941 |
| | Hispanic (Of Any Race) | Mean | 265.40 | 184.23 | 203.60 | 298.12 | 281.99 | 194.16 | 208.57 | 220.31 |
| | | N | 137,415 | 137,423 | 137,413 | 137,404 | 137,402 | 137,413 | 137,414 | 137,395 |
| | Non-Hispanic American Indian | Mean | 269.89 | 181.82 | 192.49 | 300.37 | 285.35 | 187.39 | 208.23 | 216.58 |
| | | N | 3,539 | 3,540 | 3,540 | 3,539 | 3,538 | 3,540 | 3,539 | 3,538 |
| | Non-Hispanic Multi-racial | Mean | 304.73 | 223.73 | 238.55 | 336.70 | 320.95 | 231.40 | 248.11 | 258.14 |
| | | N | 967 | 969 | 969 | 966 | 966 | 969 | 967 | 966 |
| | Non-Hispanic White | Mean | 284.48 | 204.71 | 225.05 | 315.90 | 300.41 | 215.14 | 228.63 | 240.51 |
| | | N | 20,220 | 20,220 | 20,219 | 20,220 | 20,220 | 20,218 | 20,220 | 20,218 |
| | Missing | Mean | 264.83 | 190.13 | 204.94 | 296.52 | 280.88 | 197.77 | 212.55 | 222.53 |
| | | N | 3,818 | 3,820 | 3,820 | 3,818 | 3,818 | 3,820 | 3,818 | 3,818 |
| 1 | Non-Hispanic Asian | Mean | 304.16 | 291.83 | 275.91 | 341.54 | 323.09 | 284.16 | 295.51 | 295.60 |
| | | N | 25,105 | 25,087 | 25,094 | 25,083 | 25,075 | 25,078 | 25,084 | 25,054 |
| | Non-Hispanic Pacific Islander | Mean | 292.97 | 279.06 | 268.01 | 334.01 | 313.73 | 273.81 | 283.25 | 285.52 |
| | | N | 1,821 | 1,817 | 1,818 | 1,816 | 1,816 | 1,815 | 1,817 | 1,811 |
| | Non-Hispanic Black | Mean | 300.03 | 285.52 | 268.60 | 344.74 | 322.63 | 277.35 | 289.87 | 290.69 |
| | | N | 8,790 | 8,780 | 8,784 | 8,788 | 8,781 | 8,773 | 8,777 | 8,765 |
| | Hispanic (Of Any Race) | Mean | 297.63 | 281.42 | 264.59 | 334.96 | 316.52 | 273.30 | 286.30 | 286.03 |
| | | N | 139,025 | 138,920 | 138,969 | 138,967 | 138,892 | 138,838 | 138,879 | 138,698 |
| | Non-Hispanic American Indian | Mean | 296.11 | 280.90 | 262.45 | 333.94 | 315.30 | 271.98 | 285.45 | 284.78 |
| | | N | 3,601 | 3,596 | 3,595 | 3,601 | 3,594 | 3,588 | 3,593 | 3,579 |
| | Non-Hispanic Multi-racial | Mean | 310.06 | 292.17 | 273.39 | 356.23 | 333.39 | 283.06 | 297.49 | 297.93 |
| | | N | 887 | 888 | 888 | 888 | 886 | 888 | 886 | 885 |
| | Non-Hispanic White | Mean | 304.26 | 287.56 | 270.87 | 346.83 | 325.75 | 279.51 | 292.56 | 293.16 |
| | | N | 18,721 | 18,706 | 18,721 | 18,721 | 18,711 | 18,702 | 18,701 | 18,690 |
| | Missing | Mean | 295.11 | 281.10 | 265.95 | 334.66 | 315.11 | 273.84 | 285.30 | 286.01 |
| | | N | 3,386 | 3,383 | 3,380 | 3,380 | 3,377 | 3,377 | 3,382 | 3,371 |

Return to Summary of Assessment Records Claims, Actions, and Evidence

| Grade | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Non-Hispanic Asian | Mean | 329.40 | 318.52 | 293.33 | 357.35 | 343.65 | 306.21 | 321.94 | 317.20 |
| | | N | 22,931 | 22,928 | 22,915 | 22,917 | 22,907 | 22,906 | 22,920 | 22,886 |
| | Non-Hispanic Pacific Islander | Mean | 318.84 | 307.12 | 286.94 | 350.55 | 335.07 | 297.31 | 310.88 | 308.47 |
| | | N | 1,646 | 1,643 | 1,646 | 1,643 | 1,642 | 1,642 | 1,641 | 1,638 |
| | Non-Hispanic Black | Mean | 324.59 | 310.72 | 286.09 | 358.46 | 341.81 | 298.71 | 314.99 | 311.40 |
| | | N | 8,572 | 8,569 | 8,570 | 8,570 | 8,566 | 8,563 | 8,566 | 8,556 |
| | Hispanic (Of Any Race) | Mean | 324.48 | 308.60 | 284.19 | 355.11 | 340.09 | 296.68 | 313.51 | 309.46 |
| | | N | 130,097 | 129,995 | 130,066 | 130,081 | 130,004 | 129,925 | 129,961 | 129,821 |
| | Non-Hispanic American Indian | Mean | 320.60 | 304.73 | 281.16 | 346.63 | 333.91 | 293.25 | 309.63 | 305.22 |
| | | N | 3,490 | 3,489 | 3,478 | 3,488 | 3,485 | 3,478 | 3,486 | 3,471 |
| | Non-Hispanic Multi-racial | Mean | 336.33 | 321.37 | 291.86 | 372.77 | 354.87 | 306.91 | 325.97 | 321.10 |
| | | N | 788 | 788 | 788 | 786 | 786 | 788 | 788 | 786 |
| | Non-Hispanic White | Mean | 329.45 | 315.48 | 289.84 | 362.68 | 346.36 | 302.95 | 319.81 | 315.73 |
| | | N | 17,352 | 17,349 | 17,344 | 17,339 | 17,336 | 17,339 | 17,345 | 17,323 |
| | Missing | Mean | 319.51 | 306.86 | 283.59 | 350.78 | 335.50 | 295.52 | 310.77 | 307.25 |
| | | N | 2,732 | 2,731 | 2,730 | 2,734 | 2,726 | 2,728 | 2,731 | 2,722 |
| 3 | Non-Hispanic Asian | Mean | 355.31 | 337.60 | 347.41 | 354.62 | 355.23 | 342.72 | 343.10 | 346.30 |
| | | N | 19,362 | 19,360 | 19,345 | 19,350 | 19,345 | 19,342 | 19,355 | 19,328 |
| | Non-Hispanic Pacific Islander | Mean | 344.70 | 329.59 | 342.20 | 349.43 | 347.28 | 336.09 | 334.25 | 339.28 |
| | | N | 1,635 | 1,634 | 1,630 | 1,635 | 1,633 | 1,630 | 1,633 | 1,629 |
| | Non-Hispanic Black | Mean | 350.25 | 330.62 | 340.02 | 357.48 | 354.10 | 335.53 | 336.66 | 340.94 |
| | | N | 8,072 | 8,056 | 8,058 | 8,063 | 8,061 | 8,048 | 8,054 | 8,042 |
| | Hispanic (Of Any Race) | Mean | 347.95 | 330.49 | 339.60 | 355.38 | 351.94 | 335.26 | 335.90 | 340.09 |
| | | N | 114154 | 114050 | 114077 | 114120 | 114060 | 113951 | 114021 | 113857 |
| | Non-Hispanic American Indian | Mean | 338.85 | 324.04 | 330.47 | 345.47 | 342.47 | 327.48 | 328.66 | 331.77 |
| | | N | 3,331 | 3,333 | 3,325 | 3,327 | 3,324 | 3,324 | 3,331 | 3,320 |
| | Non-Hispanic Multi-racial | Mean | 363.15 | 338.68 | 345.86 | 371.25 | 367.48 | 342.48 | 346.20 | 349.79 |
| | | N | 619 | 620 | 620 | 619 | 619 | 620 | 619 | 619 |
| | Non-Hispanic White | Mean | 355.19 | 335.64 | 344.63 | 361.95 | 358.84 | 340.36 | 341.67 | 345.71 |
| | | N | 14,065 | 14,049 | 14,048 | 14,055 | 14,053 | 14,035 | 14,048 | 14,030 |
| | Missing | Mean | 350.43 | 331.30 | 339.05 | 352.12 | 351.57 | 335.41 | 337.23 | 340.09 |
| | | N | 2,973 | 2,970 | 2,968 | 2,972 | 2,971 | 2,966 | 2,969 | 2,964 |
| 4 | Non-Hispanic Asian | Mean | 365.94 | 347.16 | 352.10 | 355.10 | 360.75 | 349.91 | 352.95 | 352.96 |
| | | N | 12,601 | 12,594 | 12,595 | 12,598 | 12,595 | 12,587 | 12,592 | 12,581 |
| | Non-Hispanic Pacific Islander | Mean | 354.77 | 337.62 | 347.13 | 353.26 | 354.26 | 342.63 | 342.87 | 345.89 |
| | | N | 1,013 | 1,012 | 1,010 | 1,012 | 1,012 | 1,010 | 1,012 | 1,009 |
| | Non-Hispanic Black | Mean | 359.69 | 338.86 | 344.64 | 359.23 | 359.70 | 342.05 | 345.27 | 347.16 |
| | | N | 5,812 | 5,794 | 5,805 | 5,811 | 5,807 | 5,787 | 5,792 | 5,782 |
| | Hispanic (Of Any Race) | Mean | 360.10 | 339.93 | 346.27 | 358.39 | 359.48 | 343.39 | 346.12 | 348.00 |
| | | N | 71,623 | 71,534 | 71,592 | 71,577 | 71,540 | 71,482 | 71,509 | 71,390 |
| | Non-Hispanic American Indian | Mean | 351.26 | 334.09 | 339.83 | 351.60 | 351.66 | 337.32 | 339.40 | 341.43 |
| | | N | 2,838 | 2,837 | 2,825 | 2,836 | 2,834 | 2,824 | 2,836 | 2,821 |
| | Non-Hispanic Multi-racial | Mean | 373.56 | 348.27 | 351.64 | 373.48 | 373.76 | 350.24 | 355.97 | 357.02 |
| | | N | 361 | 361 | 361 | 361 | 361 | 361 | 361 | 361 |
| | Non-Hispanic White | Mean | 364.25 | 343.93 | 349.08 | 362.64 | 363.67 | 346.80 | 350.18 | 351.64 |
| | | N | 8,775 | 8,769 | 8,773 | 8,772 | 8,769 | 8,766 | 8,766 | 8,759 |
| | Missing | Mean | 352.98 | 335.73 | 340.78 | 350.79 | 352.11 | 338.51 | 341.07 | 342.41 |
| | | N | 2,130 | 2,131 | 2,123 | 2,127 | 2,125 | 2,123 | 2,130 | 2,119 |

| Grade | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | Non-Hispanic Asian | Mean | 374.06 | 355.11 | 355.92 | 356.22 | 365.38 | 355.79 | 360.87 | 358.43 |
| | | N | 10,222 | 10,215 | 10,217 | 10,220 | 10,216 | 10,210 | 10,212 | 10,204 |
| | Non-Hispanic Pacific Islander | Mean | 365.66 | 345.25 | 352.88 | 355.98 | 361.06 | 349.43 | 351.69 | 352.84 |
| | | N | 805 | 809 | 808 | 805 | 805 | 808 | 805 | 805 |
| | Non-Hispanic Black | Mean | 369.85 | 348.41 | 350.62 | 362.68 | 366.59 | 349.80 | 354.92 | 354.61 |
| | | N | 4,893 | 4,885 | 4,890 | 4,894 | 4,889 | 4,880 | 4,882 | 4,876 |
| | Hispanic (Of Any Race) | Mean | 370.37 | 348.44 | 352.54 | 362.81 | 366.86 | 350.77 | 355.10 | 355.38 |
| | | N | 55,641 | 55,579 | 55,599 | 55,591 | 55,568 | 55,522 | 55,559 | 55,450 |
| | Non-Hispanic American Indian | Mean | 364.31 | 343.98 | 348.20 | 357.03 | 360.96 | 346.40 | 350.14 | 350.57 |
| | | N | 2,365 | 2,363 | 2,364 | 2,362 | 2,359 | 2,360 | 2,361 | 2,354 |
| | Non-Hispanic Multi-racial | Mean | 384.27 | 357.85 | 358.61 | 371.55 | 378.19 | 358.47 | 365.87 | 364.18 |
| | | N | 268 | 268 | 268 | 268 | 268 | 268 | 268 | 268 |
| | Non-Hispanic White | Mean | 374.91 | 353.48 | 354.82 | 366.77 | 371.10 | 354.44 | 359.99 | 359.20 |
| | | N | 6,887 | 6,876 | 6,883 | 6,886 | 6,884 | 6,872 | 6,875 | 6,870 |
| | Missing | Mean | 366.12 | 345.12 | 347.12 | 351.61 | 359.17 | 346.41 | 351.49 | 350.03 |
| | | N | 1,844 | 1,843 | 1,841 | 1,844 | 1,842 | 1,841 | 1,842 | 1,839 |
| 6 | Non-Hispanic Asian | Mean | 378.42 | 356.27 | 354.40 | 363.66 | 371.31 | 355.59 | 362.95 | 360.12 |
| | | N | 8,843 | 8,845 | 8,843 | 8,841 | 8,834 | 8,838 | 8,841 | 8,827 |
| | Non-Hispanic Pacific Islander | Mean | 369.48 | 347.86 | 348.18 | 364.48 | 367.26 | 348.26 | 354.38 | 353.78 |
| | | N | 826 | 826 | 822 | 825 | 825 | 822 | 826 | 821 |
| | Non-Hispanic Black | Mean | 375.66 | 350.69 | 347.20 | 369.41 | 372.85 | 349.20 | 358.22 | 356.09 |
| | | N | 4,671 | 4,660 | 4,667 | 4,667 | 4,661 | 4,656 | 4,658 | 4,648 |
| | Hispanic (Of Any Race) | Mean | 377.50 | 350.15 | 348.58 | 370.88 | 374.44 | 349.62 | 358.42 | 356.87 |
| | | N | 49,237 | 49,176 | 49,201 | 49,174 | 49,131 | 49,118 | 49,149 | 49,021 |
| | Non-Hispanic American Indian | Mean | 370.14 | 347.71 | 345.79 | 367.38 | 369.05 | 347.02 | 354.47 | 353.47 |
| | | N | 2,316 | 2,315 | 2,306 | 2,310 | 2,306 | 2,304 | 2,313 | 2,294 |
| | Non-Hispanic Multi-racial | Mean | 383.88 | 357.91 | 352.83 | 375.44 | 379.62 | 355.55 | 365.73 | 362.50 |
| | | N | 211 | 211 | 212 | 211 | 210 | 211 | 211 | 210 |
| | Non-Hispanic White | Mean | 378.96 | 354.61 | 352.00 | 372.58 | 376.04 | 353.58 | 361.96 | 360.12 |
| | | N | 6,060 | 6,055 | 6,057 | 6,059 | 6,056 | 6,050 | 6,053 | 6,045 |
| | Missing | Mean | 369.30 | 348.39 | 346.12 | 358.58 | 364.25 | 347.53 | 354.76 | 352.36 |
| | | N | 1,727 | 1,729 | 1,725 | 1,718 | 1,718 | 1,723 | 1,726 | 1,716 |
| 7 | Non-Hispanic Asian | Mean | 387.19 | 364.85 | 359.40 | 367.08 | 377.39 | 362.39 | 371.65 | 366.67 |
| | | N | 9,243 | 9,245 | 9,240 | 9,242 | 9,237 | 9,236 | 9,237 | 9,226 |
| | Non-Hispanic Pacific Islander | Mean | 376.54 | 352.63 | 352.22 | 367.59 | 372.43 | 352.74 | 360.08 | 358.55 |
| | | N | 816 | 819 | 818 | 820 | 815 | 817 | 816 | 814 |
| | Non-Hispanic Black | Mean | 384.55 | 359.18 | 354.03 | 372.93 | 379.04 | 356.87 | 366.86 | 363.31 |
| | | N | 4,796 | 4,783 | 4,798 | 4,797 | 4,789 | 4,781 | 4,778 | 4,770 |
| | Hispanic (Of Any Race) | Mean | 385.57 | 357.65 | 354.51 | 373.97 | 380.05 | 356.34 | 366.11 | 363.24 |
| | | N | 50,144 | 50,123 | 50,112 | 50,081 | 50,017 | 50,051 | 50,075 | 49,913 |
| | Non-Hispanic American Indian | Mean | 375.12 | 353.19 | 351.44 | 369.45 | 372.57 | 352.63 | 359.87 | 358.36 |
| | | N | 2,235 | 2,232 | 2,229 | 2,230 | 2,226 | 2,225 | 2,230 | 2,215 |
| | Non-Hispanic Multi-racial | Mean | 392.05 | 365.55 | 359.62 | 382.50 | 387.54 | 362.96 | 373.60 | 370.17 |
| | | N | 235 | 235 | 235 | 235 | 235 | 234 | 235 | 234 |
| | Non-Hispanic White | Mean | 387.15 | 363.28 | 358.08 | 375.50 | 381.60 | 360.93 | 370.57 | 366.93 |
| | | N | 5,884 | 5,882 | 5,879 | 5,882 | 5,877 | 5,877 | 5,878 | 5,869 |
| | Missing | Mean | 371.19 | 354.06 | 350.05 | 354.22 | 363.07 | 352.33 | 359.32 | 355.37 |
| | | N | 1,528 | 1,525 | 1,529 | 1,532 | 1,526 | 1,523 | 1,524 | 1,521 |

| Grade | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 8 | Non-Hispanic Asian | Mean | 394.04 | 372.38 | 364.04 | 370.81 | 382.71 | 368.49 | 378.94 | 372.54 |
| | | N | 9,287 | 9,285 | 9,277 | 9,279 | 9,277 | 9,273 | 9,282 | 9,262 |
| | Non-Hispanic Pacific Islander | Mean | 386.26 | 360.52 | 358.53 | 376.09 | 381.46 | 359.79 | 368.33 | 366.07 |
| | | N | 811 | 810 | 809 | 814 | 809 | 807 | 809 | 806 |
| | Non-Hispanic Black | Mean | 391.25 | 365.71 | 358.40 | 376.97 | 384.43 | 362.33 | 373.46 | 368.75 |
| | | N | 4,774 | 4,767 | 4,770 | 4,771 | 4,766 | 4,761 | 4,764 | 4,754 |
| | Hispanic (Of Any Race) | Mean | 391.40 | 364.15 | 359.10 | 375.22 | 383.62 | 361.89 | 372.40 | 368.20 |
| | | N | 46,617 | 46,575 | 46,580 | 46,523 | 46,475 | 46,513 | 46,549 | 46,385 |
| | Non-Hispanic American Indian | Mean | 385.67 | 361.02 | 357.51 | 374.08 | 380.15 | 359.53 | 368.48 | 365.51 |
| | | N | 2,068 | 2,068 | 2,065 | 2,060 | 2,051 | 2,062 | 2,065 | 2,043 |
| | Non-Hispanic Multi-racial | Mean | 405.59 | 375.56 | 364.69 | 384.11 | 395.25 | 370.40 | 384.58 | 377.68 |
| | | N | 201 | 201 | 200 | 199 | 199 | 200 | 201 | 198 |
| | Non-Hispanic White | Mean | 393.71 | 369.93 | 362.73 | 378.68 | 386.50 | 366.59 | 377.12 | 372.30 |
| | | N | 5,528 | 5,529 | 5,529 | 5,525 | 5,521 | 5,524 | 5,524 | 5,517 |
| | Missing | Mean | 375.27 | 357.83 | 351.52 | 355.42 | 365.72 | 354.94 | 363.19 | 358.02 |
| | | N | 1,420 | 1,424 | 1,424 | 1,424 | 1,418 | 1,422 | 1,419 | 1,416 |
| 9 | Non-Hispanic Asian | Mean | 387.92 | 380.49 | 401.04 | 378.51 | 383.47 | 391.03 | 382.84 | 388.56 |
| | | N | 10,632 | 10,642 | 10,640 | 10,629 | 10,605 | 10,630 | 10,624 | 10,589 |
| | Non-Hispanic Pacific Islander | Mean | 382.79 | 370.65 | 396.30 | 385.49 | 384.52 | 383.85 | 374.43 | 383.99 |
| | | N | 1,059 | 1,066 | 1,058 | 1,064 | 1,046 | 1,054 | 1,055 | 1,038 |
| | Non-Hispanic Black | Mean | 378.58 | 373.39 | 393.70 | 375.03 | 377.09 | 383.82 | 375.03 | 381.60 |
| | | N | 5,876 | 5,882 | 5,883 | 5,867 | 5,852 | 5,875 | 5,868 | 5,839 |
| | Hispanic (Of Any Race) | Mean | 380.35 | 370.97 | 393.64 | 375.40 | 378.17 | 382.61 | 373.88 | 381.10 |
| | | N | 51,569 | 51,575 | 51,488 | 51,476 | 51,244 | 51,371 | 51,478 | 51,066 |
| | Non-Hispanic American Indian | Mean | 388.33 | 368.12 | 395.70 | 382.78 | 385.78 | 382.24 | 374.26 | 383.10 |
| | | N | 1,953 | 1,952 | 1,949 | 1,939 | 1,931 | 1,946 | 1,951 | 1,927 |
| | Non-Hispanic Multi-racial | Mean | 393.65 | 381.73 | 400.95 | 391.09 | 392.56 | 392.03 | 385.17 | 391.94 |
| | | N | 240 | 241 | 241 | 241 | 240 | 240 | 240 | 239 |
| | Non-Hispanic White | Mean | 387.99 | 377.04 | 397.44 | 385.46 | 387.00 | 387.50 | 380.44 | 387.16 |
| | | N | 5893 | 5895 | 5896 | 5893 | 5881 | 5891 | 5888 | 5876 |
| | Missing | Mean | 358.81 | 363.00 | 379.66 | 347.80 | 353.58 | 371.64 | 361.82 | 366.03 |
| | | N | 2,346 | 2,349 | 2,344 | 2,343 | 2,329 | 2,342 | 2,345 | 2,326 |
| 10 | Non-Hispanic Asian | Mean | 386.84 | 379.63 | 400.32 | 373.75 | 380.55 | 390.25 | 381.88 | 387.12 |
| | | N | 7,855 | 7,853 | 7,855 | 7,848 | 7,834 | 7,842 | 7,842 | 7,817 |
| | Non-Hispanic Pacific Islander | Mean | 384.93 | 373.54 | 400.16 | 388.82 | 387.10 | 387.14 | 377.09 | 387.00 |
| | | N | 607 | 609 | 605 | 607 | 602 | 605 | 605 | 597 |
| | Non-Hispanic Black | Mean | 382.90 | 374.94 | 396.07 | 378.68 | 381.02 | 385.79 | 377.38 | 384.11 |
| | | N | 4,323 | 4,326 | 4,321 | 4,316 | 4,304 | 4,313 | 4,314 | 4,288 |
| | Hispanic (Of Any Race) | Mean | 383.52 | 373.04 | 395.67 | 379.34 | 381.71 | 384.64 | 376.27 | 383.58 |
| | | N | 30,017 | 30,025 | 29,989 | 29,954 | 29,829 | 29,921 | 29,966 | 29,729 |
| | Non-Hispanic American Indian | Mean | 392.25 | 373.42 | 398.55 | 384.53 | 388.73 | 386.31 | 379.16 | 386.82 |
| | | N | 1,298 | 1,296 | 1,293 | 1,284 | 1,280 | 1,289 | 1,296 | 1,275 |
| | Non-Hispanic Multi-racial | Mean | 398.30 | 385.66 | 404.32 | 394.08 | 396.47 | 395.26 | 389.52 | 395.40 |
| | | N | 178 | 178 | 178 | 178 | 178 | 178 | 178 | 178 |
| | Non-Hispanic White | Mean | 394.03 | 379.87 | 400.08 | 388.69 | 391.59 | 390.26 | 384.21 | 390.45 |
| | | N | 4,765 | 4,764 | 4,763 | 4,767 | 4,759 | 4,759 | 4,761 | 4,752 |
| | Missing | Mean | 374.53 | 371.29 | 389.94 | 365.05 | 370.01 | 380.97 | 372.31 | 377.40 |
| | | N | 1,097 | 1,097 | 1,092 | 1,097 | 1,089 | 1,091 | 1,095 | 1,083 |

| Grade | Ethnicity | | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
|---|---|---|---|---|---|---|---|---|---|---|
| 11 | Non-Hispanic Asian | Mean | 391.05 | 383.15 | 403.46 | 377.06 | 384.35 | 393.58 | 385.59 | 390.64 |
| | | N | 7,092 | 7,094 | 7,092 | 7,081 | 7,071 | 7,086 | 7,085 | 7,059 |
| | Non-Hispanic Pacific Islander | Mean | 384.40 | 374.38 | 399.05 | 383.95 | 384.66 | 387.09 | 377.65 | 386.20 |
| | | N | 419 | 422 | 419 | 417 | 414 | 419 | 419 | 413 |
| | Non-Hispanic Black | Mean | 385.81 | 378.35 | 399.98 | 380.91 | 383.61 | 389.47 | 380.65 | 387.52 |
| | | N | 3,709 | 3,714 | 3,707 | 3,711 | 3,696 | 3,700 | 3,702 | 3,683 |
| | Hispanic (Of Any Race) | Mean | 388.78 | 377.92 | 399.97 | 384.61 | 387.00 | 389.24 | 381.25 | 388.37 |
| | | N | 21,229 | 21,261 | 21,211 | 21,221 | 21,087 | 21,161 | 21,192 | 21,015 |
| | Non-Hispanic American Indian | Mean | 393.96 | 375.15 | 399.74 | 385.13 | 389.74 | 387.83 | 380.81 | 388.11 |
| | | N | 954 | 953 | 945 | 943 | 937 | 943 | 950 | 929 |
| | Non-Hispanic Multi-racial | Mean | 396.06 | 380.41 | 403.01 | 387.35 | 391.95 | 391.92 | 385.15 | 391.71 |
| | | N | 101 | 101 | 101 | 101 | 101 | 101 | 101 | 101 |
| | Non-Hispanic White | Mean | 399.28 | 384.31 | 403.85 | 394.71 | 397.23 | 394.33 | 388.87 | 395.01 |
| | | N | 3,800 | 3,798 | 3,802 | 3,795 | 3,783 | 3,790 | 3,791 | 3,773 |
| | Missing | Mean | 383.08 | 377.63 | 397.94 | 371.98 | 378.05 | 388.12 | 379.37 | 385.10 |
| | | N | 896 | 902 | 893 | 895 | 889 | 892 | 896 | 884 |
| 12 | Non-Hispanic Asian | Mean | 389.92 | 382.36 | 403.21 | 378.15 | 384.40 | 393.05 | 384.73 | 390.35 |
| | | N | 5,704 | 5,703 | 5,702 | 5,692 | 5,667 | 5,691 | 5,692 | 5,647 |
| | Non-Hispanic Pacific Islander | Mean | 384.99 | 375.37 | 402.31 | 390.08 | 388.31 | 389.26 | 378.60 | 389.10 |
| | | N | 299 | 302 | 299 | 297 | 292 | 299 | 298 | 290 |
| | Non-Hispanic Black | Mean | 386.14 | 378.75 | 400.72 | 384.77 | 385.85 | 390.01 | 381.06 | 388.63 |
| | | N | 3,154 | 3,160 | 3,144 | 3,153 | 3,133 | 3,138 | 3,143 | 3,112 |
| | Hispanic (Of Any Race) | Mean | 390.78 | 379.76 | 401.42 | 389.12 | 390.26 | 390.87 | 383.16 | 390.48 |
| | | N | 16,175 | 16,205 | 16,143 | 16,161 | 16,010 | 16,098 | 16,128 | 15,924 |
| | Non-Hispanic American Indian | Mean | 391.79 | 375.97 | 402.07 | 387.97 | 390.37 | 389.30 | 380.80 | 389.52 |
| | | N | 816 | 813 | 808 | 805 | 799 | 807 | 812 | 795 |
| | Non-Hispanic Multi-racial | Mean | 400.94 | 386.94 | 407.29 | 391.41 | 396.68 | 397.89 | 391.24 | 397.11 |
| | | N | 86 | 86 | 86 | 86 | 85 | 85 | 86 | 84 |
| | Non-Hispanic White | Mean | 401.44 | 385.57 | 404.92 | 399.68 | 400.90 | 395.51 | 390.45 | 396.93 |
| | | N | 3,038 | 3,042 | 3,040 | 3,033 | 3,023 | 3,035 | 3,036 | 3,020 |
| | Missing | Mean | 387.14 | 380.54 | 399.78 | 378.36 | 383.72 | 390.47 | 382.67 | 388.44 |
| | | N | 727 | 731 | 725 | 721 | 713 | 724 | 726 | 711 |

## 4.2.3 Correlations among Scale Scores by Grade-level Cluster

**Table 4.2.3A**
Correlations Among Scale Scores: K S302

|  |  | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **Listening** | Pearson Correlation | 1 | .537 | .555 | .783 |
|  | N | 203,841 | 203,837 | 203,824 | 203,823 |
| **Reading** | Pearson Correlation |  | 1 | .720 | .496 |
|  | N |  | 203,853 | 203,837 | 203,823 |
| **Writing** | Pearson Correlation |  |  | 1 | .553 |
|  | N |  |  | 203,840 | 203,813 |
| **Speaking** | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 203,827 |

**Table 4.2.3B**
Correlations Among Scale Scores: 1-2 S302

|  |  | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **Listening** | Pearson Correlation | 1 | .688 | .569 | .497 |
|  | N | 388,944 | 388,557 | 388,626 | 388,584 |
| **Reading** | Pearson Correlation |  | 1 | .671 | .447 |
|  | N |  | 388,669 | 388,428 | 388,302 |
| **Writing** | Pearson Correlation |  |  | 1 | .467 |
|  | N |  |  | 388,786 | 388,461 |
| **Speaking** | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 388,802 |

**Table 4.2.3C**
Correlations Among Scale Scores: 3-5 S302

|  |  | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **Listening** | Pearson Correlation | 1 | .727 | .611 | .484 |
|  | N | 352,289 | 351,832 | 351,878 | 351,940 |
| **Reading** | Pearson Correlation |  | 1 | .676 | .472 |
|  | N |  | 351,942 | 351,617 | 351,579 |
| **Writing** | Pearson Correlation |  |  | 1 | .509 |
|  | N |  |  | 352,025 | 351,709 |
| **Speaking** | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 352,105 |

**Table 4.2.3D**

Correlations Among Scale Scores: 6-8 S302

| | | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **Listening** | Pearson Correlation | 1 | .672 | .518 | .564 |
| | N | 219,478 | 219,163 | 219,123 | 218,979 |
| **Reading** | Pearson Correlation | | 1 | .593 | .479 |
| | N | | 219,320 | 219,028 | 218,791 |
| **Writing** | Pearson Correlation | | | 1 | .470 |
| | N | | | 219,327 | 218,869 |
| **Speaking** | Pearson Correlation | | | | 1 |
| | N | | | | 219,219 |

**Table 4.2.3E**

Correlations Among Scale Scores: 9-12 S302

| | | Listening | Reading | Writing | Speaking |
|---|---|---|---|---|---|
| **Listening** | Pearson Correlation | 1 | .700 | .648 | .600 |
| | N | 197,907 | 197,563 | 197,133 | 196,703 |
| **Reading** | Pearson Correlation | | 1 | .694 | .530 |
| | N | | 198,037 | 197,316 | 196,711 |
| **Writing** | Pearson Correlation | | | 1 | .597 |
| | N | | | 197,712 | 196,576 |
| **Speaking** | Pearson Correlation | | | | 1 |
| | N | | | | 197,615 |

## 4.3 Proficiency Level Results

### 4.3.1 Listening

#### 4.3.1.1 By Cluster by Tier

**Table 4.3.1.1A**

Proficiency Level by Cluster by Tier (Count): Listening S302

| Cluster | Tier | Listening Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 24,626 | 11,794 | 20,693 | 32,677 | 62,805 | 51,246 | 203,841 |
| K (accountability) | - | 51,112 | 20,746 | 17,932 | 11,923 | 32,179 | 69,949 | 203,841 |
| 1-2 | A | 6,345 | 11,393 | 21,368 | 46,277 | n/a | n/a | 85,383 |
| | B | 967 | 2,783 | 8,754 | 8,427 | 173,061 | n/a | 193,992 |
| | C | 456 | 2,920 | 17,081 | 10,457 | 29,631 | 49,024 | 109,569 |
| 3-5 | A | 2,267 | 8,556 | 8,057 | 10,364 | n/a | n/a | 29,244 |
| | B | 753 | 6,116 | 21,657 | 21,837 | 98,554 | n/a | 148,917 |
| | C | 259 | 1,627 | 13,043 | 16,456 | 46,912 | 95,831 | 174,128 |
| 6-8 | A | 5,666 | 9,476 | 4,678 | 3,931 | n/a | n/a | 23,751 |
| | B | 986 | 10,662 | 18,978 | 21,744 | 30,003 | n/a | 82,373 |
| | C | 251 | 695 | 8,043 | 18,368 | 40,420 | 45,577 | 113,354 |
| 9-12 | A | 13,996 | 10,413 | 2,561 | 2,107 | n/a | n/a | 29,077 |
| | B | 2,893 | 9,402 | 20,146 | 19,391 | 24,225 | n/a | 76,057 |
| | C | 1,287 | 3,986 | 14,401 | 29,040 | 24,490 | 19,569 | 92,773 |

**Table 4.3.1.1B**

Proficiency Level by Cluster by Tier (Percent): Listening S302

| Cluster | Tier | Listening Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 12.1% | 5.8% | 10.2% | 16.0% | 30.8% | 25.1% | 100.0% |
| K (accountability) | - | 25.1% | 10.2% | 8.8% | 5.8% | 15.8% | 34.3% | 100.0% |
| 1-2 | A | 7.4% | 13.3% | 25.0% | 54.2% | n/a | n/a | 100.0% |
| | B | 0.5% | 1.4% | 4.5% | 4.3% | 89.2% | n/a | 100.0% |
| | C | 0.4% | 2.7% | 15.6% | 9.5% | 27.0% | 44.7% | 100.0% |
| 3-5 | A | 7.8% | 29.3% | 27.6% | 35.4% | n/a | n/a | 100.0% |
| | B | 0.5% | 4.1% | 14.5% | 14.7% | 66.2% | n/a | 100.0% |
| | C | 0.1% | 0.9% | 7.5% | 9.5% | 26.9% | 55.0% | 100.0% |
| 6-8 | A | 23.9% | 39.9% | 19.7% | 16.6% | n/a | n/a | 100.0% |
| | B | 1.2% | 12.9% | 23.0% | 26.4% | 36.4% | n/a | 100.0% |
| | C | 0.2% | 0.6% | 7.1% | 16.2% | 35.7% | 40.2% | 100.0% |
| 9-12 | A | 48.1% | 35.8% | 8.8% | 7.2% | n/a | n/a | 100.0% |
| | B | 3.8% | 12.4% | 26.5% | 25.5% | 31.9% | n/a | 100.0% |
| | C | 1.4% | 4.3% | 15.5% | 31.3% | 26.4% | 21.1% | 100.0% |

Return to Summary of Assessment Records Claims, Actions, and Evidence

## 4.3.1.2    By Grade by Tier

**Table 4.3.1.2A**

Proficiency Level by Grade by Tier (Count): Listening S302

| Grade | Tier | Listening Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| K (instructional) | - | 24,626 | 11,794 | 20,693 | 32,677 | 62,805 | 51,246 | 203,841 |
| K (accountability) | - | 51,112 | 20,746 | 17,932 | 11,923 | 32,179 | 69,949 | 203,841 |
| 1 | A | 4,076 | 7,378 | 17,731 | 38,308 | n/a | n/a | 67,493 |
| | B | 569 | 1,874 | 3,972 | 4,022 | 80,709 | n/a | 91,146 |
| | C | 247 | 1,283 | 9,165 | 4,588 | 11,235 | 16,179 | 42,697 |
| 2 | A | 2,269 | 4,015 | 3,637 | 7,969 | n/a | n/a | 17,890 |
| | B | 398 | 909 | 4,782 | 4,405 | 92,352 | n/a | 102,846 |
| | C | 209 | 1,637 | 7,916 | 5,869 | 18,396 | 32,845 | 66,872 |
| 3 | A | 574 | 3,328 | 3,803 | 5,094 | n/a | n/a | 12,799 |
| | B | 260 | 2,851 | 11,149 | 6,756 | 53,928 | n/a | 74,944 |
| | C | 105 | 506 | 4,445 | 3,448 | 19,283 | 48,681 | 76,468 |
| 4 | A | 704 | 2,700 | 2,363 | 3,025 | n/a | n/a | 8,792 |
| | B | 244 | 1,798 | 5,931 | 8,507 | 26,038 | n/a | 42,518 |
| | C | 94 | 578 | 3,719 | 6,189 | 17,031 | 26,232 | 53,843 |
| 5 | A | 989 | 2,528 | 1,891 | 2,245 | n/a | n/a | 7,653 |
| | B | 249 | 1,467 | 4,577 | 6,574 | 18,588 | n/a | 31,455 |
| | C | 60 | 543 | 4,879 | 6,819 | 10,598 | 20,918 | 43,817 |
| 6 | A | 1,218 | 3,140 | 1,987 | 1,592 | n/a | n/a | 7,937 |
| | B | 202 | 2,876 | 6,589 | 7,043 | 11,305 | n/a | 28,015 |
| | C | 55 | 177 | 3,212 | 5,685 | 14,516 | 14,294 | 37,939 |
| 7 | A | 1,888 | 3,213 | 1,649 | 1,117 | n/a | n/a | 7,867 |
| | B | 322 | 3,564 | 7,565 | 7,664 | 9,042 | n/a | 28,157 |
| | C | 118 | 226 | 3,158 | 5,609 | 15,787 | 13,959 | 38,857 |
| 8 | A | 2,560 | 3,123 | 1,042 | 1,222 | n/a | n/a | 7,947 |
| | B | 462 | 4,222 | 4,824 | 7,037 | 9,656 | n/a | 26,201 |
| | C | 78 | 292 | 1,673 | 7,074 | 10,117 | 17,324 | 36,558 |
| 9 | A | 7,018 | 6,560 | 876 | 1,289 | n/a | n/a | 15,743 |
| | B | 540 | 3,178 | 7,540 | 6,755 | 10,514 | n/a | 28,527 |
| | C | 264 | 705 | 3,423 | 9,577 | 13,387 | 7,942 | 35,298 |
| 10 | A | 3,371 | 2,445 | 931 | 395 | n/a | n/a | 7,142 |
| | B | 666 | 2,299 | 5,573 | 4,669 | 6,688 | n/a | 19,895 |
| | C | 264 | 1,139 | 4,034 | 7,906 | 5,261 | 4,499 | 23,103 |
| 11 | A | 2,188 | 954 | 588 | 254 | n/a | n/a | 3,984 |
| | B | 721 | 2,482 | 3,093 | 5,263 | 3,984 | n/a | 15,543 |
| | C | 313 | 849 | 3,202 | 6,095 | 4,187 | 4,027 | 18,673 |
| 12 | A | 1,419 | 454 | 166 | 169 | n/a | n/a | 2,208 |
| | B | 966 | 1,443 | 3,940 | 2,704 | 3,039 | n/a | 12,092 |
| | C | 446 | 1,293 | 3,742 | 5,462 | 1,655 | 3,101 | 15,699 |

**Table 4.3.1.2B**
Proficiency Level by Grade by Tier (Percent): Listening S302

| Grade | Tier | Listening Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 12.1% | 5.8% | 10.2% | 16.0% | 30.8% | 25.1% | 100.0% |
| K (accountability) | - | 25.1% | 10.2% | 8.8% | 5.8% | 15.8% | 34.3% | 100.0% |
| 1 | A | 6.0% | 10.9% | 26.3% | 56.8% | n/a | n/a | 100.0% |
| | B | 0.6% | 2.1% | 4.4% | 4.4% | 88.5% | n/a | 100.0% |
| | C | 0.6% | 3.0% | 21.5% | 10.7% | 26.3% | 37.9% | 100.0% |
| 2 | A | 12.7% | 22.4% | 20.3% | 44.5% | n/a | n/a | 100.0% |
| | B | 0.4% | 0.9% | 4.6% | 4.3% | 89.8% | n/a | 100.0% |
| | C | 0.3% | 2.4% | 11.8% | 8.8% | 27.5% | 49.1% | 100.0% |
| 3 | A | 4.5% | 26.0% | 29.7% | 39.8% | n/a | n/a | 100.0% |
| | B | 0.3% | 3.8% | 14.9% | 9.0% | 72.0% | n/a | 100.0% |
| | C | 0.1% | 0.7% | 5.8% | 4.5% | 25.2% | 63.7% | 100.0% |
| 4 | A | 8.0% | 30.7% | 26.9% | 34.4% | n/a | n/a | 100.0% |
| | B | 0.6% | 4.2% | 13.9% | 20.0% | 61.2% | n/a | 100.0% |
| | C | 0.2% | 1.1% | 6.9% | 11.5% | 31.6% | 48.7% | 100.0% |
| 5 | A | 12.9% | 33.0% | 24.7% | 29.3% | n/a | n/a | 100.0% |
| | B | 0.8% | 4.7% | 14.6% | 20.9% | 59.1% | n/a | 100.0% |
| | C | 0.1% | 1.2% | 11.1% | 15.6% | 24.2% | 47.7% | 100.0% |
| 6 | A | 15.3% | 39.6% | 25.0% | 20.1% | n/a | n/a | 100.0% |
| | B | 0.7% | 10.3% | 23.5% | 25.1% | 40.4% | n/a | 100.0% |
| | C | 0.1% | 0.5% | 8.5% | 15.0% | 38.3% | 37.7% | 100.0% |
| 7 | A | 24.0% | 40.8% | 21.0% | 14.2% | n/a | n/a | 100.0% |
| | B | 1.1% | 12.7% | 26.9% | 27.2% | 32.1% | n/a | 100.0% |
| | C | 0.3% | 0.6% | 8.1% | 14.4% | 40.6% | 35.9% | 100.0% |
| 8 | A | 32.2% | 39.3% | 13.1% | 15.4% | n/a | n/a | 100.0% |
| | B | 1.8% | 16.1% | 18.4% | 26.9% | 36.9% | n/a | 100.0% |
| | C | 0.2% | 0.8% | 4.6% | 19.4% | 27.7% | 47.4% | 100.0% |
| 9 | A | 44.6% | 41.7% | 5.6% | 8.2% | n/a | n/a | 100.0% |
| | B | 1.9% | 11.1% | 26.4% | 23.7% | 36.9% | n/a | 100.0% |
| | C | 0.7% | 2.0% | 9.7% | 27.1% | 37.9% | 22.5% | 100.0% |
| 10 | A | 47.2% | 34.2% | 13.0% | 5.5% | n/a | n/a | 100.0% |
| | B | 3.3% | 11.6% | 28.0% | 23.5% | 33.6% | n/a | 100.0% |
| | C | 1.1% | 4.9% | 17.5% | 34.2% | 22.8% | 19.5% | 100.0% |
| 11 | A | 54.9% | 23.9% | 14.8% | 6.4% | n/a | n/a | 100.0% |
| | B | 4.6% | 16.0% | 19.9% | 33.9% | 25.6% | n/a | 100.0% |
| | C | 1.7% | 4.5% | 17.1% | 32.6% | 22.4% | 21.6% | 100.0% |
| 12 | A | 64.3% | 20.6% | 7.5% | 7.7% | n/a | n/a | 100.0% |
| | B | 8.0% | 11.9% | 32.6% | 22.4% | 25.1% | n/a | 100.0% |
| | C | 2.8% | 8.2% | 23.8% | 34.8% | 10.5% | 19.8% | 100.0% |

## 4.3.1.3　By Grade

**Table 4.3.1.3A**

Proficiency Level by Grade (Count): Listening S302

| | Listening Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 24,626 | 11,794 | 20,693 | 32,677 | 62,805 | 51,246 | 203,841 |
| K (accountability) | 51,112 | 20,746 | 17,932 | 11,923 | 32,179 | 69,949 | 203,841 |
| 1 | 4,892 | 10,535 | 30,868 | 46,918 | 91,944 | 16,179 | 201,336 |
| 2 | 2,876 | 6,561 | 16,335 | 18,243 | 110,748 | 32,845 | 187,608 |
| 3 | 939 | 6,685 | 19,397 | 15,298 | 73,211 | 48,681 | 164,211 |
| 4 | 1,042 | 5,076 | 12,013 | 17,721 | 43,069 | 26,232 | 105,153 |
| 5 | 1,298 | 4,538 | 11,347 | 15,638 | 29,186 | 20,918 | 82,925 |
| 6 | 1,475 | 6,193 | 11,788 | 14,320 | 25,821 | 14,294 | 73,891 |
| 7 | 2,328 | 7,003 | 12,372 | 14,390 | 24,829 | 13,959 | 74,881 |
| 8 | 3,100 | 7,637 | 7,539 | 15,333 | 19,773 | 17,324 | 70,706 |
| 9 | 7,822 | 10,443 | 11,839 | 17,621 | 23,901 | 7,942 | 79,568 |
| 10 | 4,301 | 5,883 | 10,538 | 12,970 | 11,949 | 4,499 | 50,140 |
| 11 | 3,222 | 4,285 | 6,883 | 11,612 | 8,171 | 4,027 | 38,200 |
| 12 | 2,831 | 3,190 | 7,848 | 8,335 | 4,694 | 3,101 | 29,999 |

**Table 4.3.1.3B**

Proficiency Level by Grade (Percent): Listening S302

| | Listening Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 12.1% | 5.8% | 10.2% | 16.0% | 30.8% | 25.1% | 100.0% |
| K (accountability) | 25.1% | 10.2% | 8.8% | 5.8% | 15.8% | 34.3% | 100.0% |
| 1 | 2.4% | 5.2% | 15.3% | 23.3% | 45.7% | 8.0% | 100.0% |
| 2 | 1.5% | 3.5% | 8.7% | 9.7% | 59.0% | 17.5% | 100.0% |
| 3 | 0.6% | 4.1% | 11.8% | 9.3% | 44.6% | 29.6% | 100.0% |
| 4 | 1.0% | 4.8% | 11.4% | 16.9% | 41.0% | 24.9% | 100.0% |
| 5 | 1.6% | 5.5% | 13.7% | 18.9% | 35.2% | 25.2% | 100.0% |
| 6 | 2.0% | 8.4% | 16.0% | 19.4% | 34.9% | 19.3% | 100.0% |
| 7 | 3.1% | 9.4% | 16.5% | 19.2% | 33.2% | 18.6% | 100.0% |
| 8 | 4.4% | 10.8% | 10.7% | 21.7% | 28.0% | 24.5% | 100.0% |
| 9 | 9.8% | 13.1% | 14.9% | 22.1% | 30.0% | 10.0% | 100.0% |
| 10 | 8.6% | 11.7% | 21.0% | 25.9% | 23.8% | 9.0% | 100.0% |
| 11 | 8.4% | 11.2% | 18.0% | 30.4% | 21.4% | 10.5% | 100.0% |
| 12 | 9.4% | 10.6% | 26.2% | 27.8% | 15.6% | 10.3% | 100.0% |

## 4.3.2    Reading

### 4.3.2.1    By Cluster by Tier

**Table 4.3.2.1A**

Proficiency Level by Cluster by Tier (Count): Reading S302

| Cluster | Tier | Reading Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** | **6** | |
| K (instructional) | - | 44,927 | 28,198 | 39,920 | 17,097 | 20,656 | 53,055 | 203,853 |
| K (accountability) | - | 136,612 | 14,186 | 9,412 | 10,849 | 32,794 | 0 | 203,853 |
| 1-2 | A | 20,639 | 20,757 | 16,980 | 26,948 | n/a | n/a | 85,324 |
| | B | 2,772 | 7,977 | 33,604 | 28,812 | 120,726 | n/a | 193,891 |
| | C | 1,444 | 4,440 | 16,953 | 14,111 | 29,797 | 42,709 | 109,454 |
| 3-5 | A | 9,265 | 9,527 | 4,175 | 6,202 | n/a | n/a | 29,169 |
| | B | 2,858 | 15,465 | 36,434 | 13,600 | 80,382 | n/a | 148,739 |
| | C | 661 | 4,241 | 22,564 | 17,162 | 56,469 | 72,937 | 174,034 |
| 6-8 | A | 9,353 | 9,416 | 2,900 | 2,060 | n/a | n/a | 23,729 |
| | B | 3,800 | 21,123 | 29,125 | 5,607 | 22,636 | n/a | 82,291 |
| | C | 1,488 | 20,293 | 36,822 | 15,335 | 21,116 | 18,246 | 113,300 |
| 9-12 | A | 11,232 | 10,882 | 3,455 | 3,537 | n/a | n/a | 29,106 |
| | B | 9,437 | 29,538 | 13,788 | 6,258 | 17,051 | n/a | 76,072 |
| | C | 1,598 | 13,021 | 14,112 | 12,626 | 20,505 | 30,997 | 92,859 |

**Table 4.3.2.1B**

Proficiency Level by Cluster by Tier (Percent): Reading S302

| Cluster | Tier | Reading Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** | **6** | |
| K (instructional) | - | 22.0% | 13.8% | 19.6% | 8.4% | 10.1% | 26.0% | 100.0% |
| K (accountability) | - | 67.0% | 7.0% | 4.6% | 5.3% | 16.1% | 0.0% | 100.0% |
| 1-2 | A | 24.2% | 24.3% | 19.9% | 31.6% | n/a | n/a | 100.0% |
| | B | 1.4% | 4.1% | 17.3% | 14.9% | 62.3% | n/a | 100.0% |
| | C | 1.3% | 4.1% | 15.5% | 12.9% | 27.2% | 39.0% | 100.0% |
| 3-5 | A | 31.8% | 32.7% | 14.3% | 21.3% | n/a | n/a | 100.0% |
| | B | 1.9% | 10.4% | 24.5% | 9.1% | 54.0% | n/a | 100.0% |
| | C | 0.4% | 2.4% | 13.0% | 9.9% | 32.4% | 41.9% | 100.0% |
| 6-8 | A | 39.4% | 39.7% | 12.2% | 8.7% | n/a | n/a | 100.0% |
| | B | 4.6% | 25.7% | 35.4% | 6.8% | 27.5% | n/a | 100.0% |
| | C | 1.3% | 17.9% | 32.5% | 13.5% | 18.6% | 16.1% | 100.0% |
| 9-12 | A | 38.6% | 37.4% | 11.9% | 12.2% | n/a | n/a | 100.0% |
| | B | 12.4% | 38.8% | 18.1% | 8.2% | 22.4% | n/a | 100.0% |
| | C | 1.7% | 14.0% | 15.2% | 13.6% | 22.1% | 33.4% | 100.0% |

#### 4.3.2.2 By Grade by Tier

**Table 4.3.2.2A**

Proficiency Level by Grade by Tier (Count): Reading S302

| Grade | Tier | Reading Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** | **6** | |
| K (instructional) | - | 44,927 | 28,198 | 39,920 | 17,097 | 20,656 | 53,055 | 203,853 |
| K (accountability) | - | 136,612 | 14,186 | 9,412 | 10,849 | 32,794 | 0 | 203,853 |
| 1 | A | 14,711 | 16,370 | 14,832 | 21,554 | n/a | n/a | 67,467 |
| | B | 1,642 | 2,562 | 15,264 | 17,879 | 53,726 | n/a | 91,073 |
| | C | 778 | 1,578 | 7,254 | 7,083 | 11,510 | 14,434 | 42,637 |
| 2 | A | 5,928 | 4,387 | 2,148 | 5,394 | n/a | n/a | 17,857 |
| | B | 1,130 | 5,415 | 18,340 | 10,933 | 67,000 | n/a | 102,818 |
| | C | 666 | 2,862 | 9,699 | 7,028 | 18,287 | 28,275 | 66,817 |
| 3 | A | 2,812 | 4,716 | 2,106 | 3,126 | n/a | n/a | 12,760 |
| | B | 664 | 4,488 | 17,361 | 5,752 | 46,621 | n/a | 74,886 |
| | C | 184 | 953 | 4,827 | 4,793 | 29,796 | 35,873 | 76,426 |
| 4 | A | 3,065 | 2,857 | 1,032 | 1,818 | n/a | n/a | 8,772 |
| | B | 1,166 | 5,812 | 9,982 | 5,648 | 19,836 | n/a | 42,444 |
| | C | 244 | 1,074 | 7,425 | 9,001 | 14,114 | 21,958 | 53,816 |
| 5 | A | 3,388 | 1,954 | 1,037 | 1,258 | n/a | n/a | 7,637 |
| | B | 1,028 | 5,165 | 9,091 | 2,200 | 13,925 | n/a | 31,409 |
| | C | 233 | 2,214 | 10,312 | 3,368 | 12,559 | 15,106 | 43,792 |
| 6 | A | 2,138 | 3,826 | 1,151 | 811 | n/a | n/a | 7,926 |
| | B | 1,001 | 5,776 | 11,188 | 1,962 | 8,048 | n/a | 27,975 |
| | C | 354 | 5,457 | 14,100 | 6,390 | 6,520 | 5,095 | 37,916 |
| 7 | A | 3,108 | 2,966 | 1,124 | 666 | n/a | n/a | 7,864 |
| | B | 1,291 | 7,271 | 10,507 | 1,850 | 7,218 | n/a | 28,137 |
| | C | 532 | 7,263 | 12,179 | 6,052 | 6,964 | 5,853 | 38,843 |
| 8 | A | 4,107 | 2,624 | 625 | 583 | n/a | n/a | 7,939 |
| | B | 1,508 | 8,076 | 7,430 | 1,795 | 7,370 | n/a | 26,179 |
| | C | 602 | 7,573 | 10,543 | 2,893 | 7,632 | 7,298 | 36,541 |
| 9 | A | 6,219 | 6,002 | 2,001 | 1,525 | n/a | n/a | 15,747 |
| | B | 2,382 | 9,946 | 7,692 | 1,717 | 6,802 | n/a | 28,539 |
| | C | 298 | 3,492 | 5,588 | 4,817 | 7,704 | 13,417 | 35,316 |
| 10 | A | 2,518 | 2,816 | 741 | 1,069 | n/a | n/a | 7,144 |
| | B | 2,416 | 9,077 | 2,560 | 2,052 | 3,787 | n/a | 19,892 |
| | C | 345 | 3,638 | 4,397 | 3,260 | 4,658 | 6,814 | 23,112 |
| 11 | A | 1,503 | 1,451 | 465 | 581 | n/a | n/a | 4,000 |
| | B | 2,299 | 5,816 | 1,984 | 1,785 | 3,657 | n/a | 15,541 |
| | C | 491 | 3,223 | 2,222 | 2,535 | 3,878 | 6,355 | 18,704 |
| 12 | A | 992 | 613 | 248 | 362 | n/a | n/a | 2,215 |
| | B | 2,340 | 4,699 | 1,552 | 704 | 2,805 | n/a | 12,100 |
| | C | 464 | 2,668 | 1,905 | 2,014 | 4,265 | 4,411 | 15,727 |

**Table 4.3.2.2B**

Proficiency Level by Grade by Tier (Percent): Reading S302

| Grade | Tier | Reading Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 22.0% | 13.8% | 19.6% | 8.4% | 10.1% | 26.0% | 100.0% |
| K (accountability) | - | 67.0% | 7.0% | 4.6% | 5.3% | 16.1% | 0.0% | 100.0% |
| 1 | A | 21.8% | 24.3% | 22.0% | 31.9% | n/a | n/a | 100.0% |
| | B | 1.8% | 2.8% | 16.8% | 19.6% | 59.0% | n/a | 100.0% |
| | C | 1.8% | 3.7% | 17.0% | 16.6% | 27.0% | 33.9% | 100.0% |
| 2 | A | 33.2% | 24.6% | 12.0% | 30.2% | n/a | n/a | 100.0% |
| | B | 1.1% | 5.3% | 17.8% | 10.6% | 65.2% | n/a | 100.0% |
| | C | 1.0% | 4.3% | 14.5% | 10.5% | 27.4% | 42.3% | 100.0% |
| 3 | A | 22.0% | 37.0% | 16.5% | 24.5% | n/a | n/a | 100.0% |
| | B | 0.9% | 6.0% | 23.2% | 7.7% | 62.3% | n/a | 100.0% |
| | C | 0.2% | 1.2% | 6.3% | 6.3% | 39.0% | 46.9% | 100.0% |
| 4 | A | 34.9% | 32.6% | 11.8% | 20.7% | n/a | n/a | 100.0% |
| | B | 2.7% | 13.7% | 23.5% | 13.3% | 46.7% | n/a | 100.0% |
| | C | 0.5% | 2.0% | 13.8% | 16.7% | 26.2% | 40.8% | 100.0% |
| 5 | A | 44.4% | 25.6% | 13.6% | 16.5% | n/a | n/a | 100.0% |
| | B | 3.3% | 16.4% | 28.9% | 7.0% | 44.3% | n/a | 100.0% |
| | C | 0.5% | 5.1% | 23.5% | 7.7% | 28.7% | 34.5% | 100.0% |
| 6 | A | 27.0% | 48.3% | 14.5% | 10.2% | n/a | n/a | 100.0% |
| | B | 3.6% | 20.6% | 40.0% | 7.0% | 28.8% | n/a | 100.0% |
| | C | 0.9% | 14.4% | 37.2% | 16.9% | 17.2% | 13.4% | 100.0% |
| 7 | A | 39.5% | 37.7% | 14.3% | 8.5% | n/a | n/a | 100.0% |
| | B | 4.6% | 25.8% | 37.3% | 6.6% | 25.7% | n/a | 100.0% |
| | C | 1.4% | 18.7% | 31.4% | 15.6% | 17.9% | 15.1% | 100.0% |
| 8 | A | 51.7% | 33.1% | 7.9% | 7.3% | n/a | n/a | 100.0% |
| | B | 5.8% | 30.8% | 28.4% | 6.9% | 28.2% | n/a | 100.0% |
| | C | 1.6% | 20.7% | 28.9% | 7.9% | 20.9% | 20.0% | 100.0% |
| 9 | A | 39.5% | 38.1% | 12.7% | 9.7% | n/a | n/a | 100.0% |
| | B | 8.3% | 34.9% | 27.0% | 6.0% | 23.8% | n/a | 100.0% |
| | C | 0.8% | 9.9% | 15.8% | 13.6% | 21.8% | 38.0% | 100.0% |
| 10 | A | 35.2% | 39.4% | 10.4% | 15.0% | n/a | n/a | 100.0% |
| | B | 12.1% | 45.6% | 12.9% | 10.3% | 19.0% | n/a | 100.0% |
| | C | 1.5% | 15.7% | 19.0% | 14.1% | 20.2% | 29.5% | 100.0% |
| 11 | A | 37.6% | 36.3% | 11.6% | 14.5% | n/a | n/a | 100.0% |
| | B | 14.8% | 37.4% | 12.8% | 11.5% | 23.5% | n/a | 100.0% |
| | C | 2.6% | 17.2% | 11.9% | 13.6% | 20.7% | 34.0% | 100.0% |
| 12 | A | 44.8% | 27.7% | 11.2% | 16.3% | n/a | n/a | 100.0% |
| | B | 19.3% | 38.8% | 12.8% | 5.8% | 23.2% | n/a | 100.0% |
| | C | 3.0% | 17.0% | 12.1% | 12.8% | 27.1% | 28.0% | 100.0% |

### 4.3.2.3 By Grade

**Table 4.3.2.3A**

Proficiency Level by Grade (Count): Reading S302

| | Reading Proficiency Range | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 44,927 | 28,198 | 39,920 | 17,097 | 20,656 | 53,055 | 203,853 |
| K (accountability) | 136,612 | 14,186 | 9,412 | 10,849 | 32,794 | 0 | 203,853 |
| 1 | 17,131 | 20,510 | 37,350 | 46,516 | 65,236 | 14,434 | 201,177 |
| 2 | 7,724 | 12,664 | 30,187 | 23,355 | 85,287 | 28,275 | 187,492 |
| 3 | 3,660 | 10,157 | 24,294 | 13,671 | 76,417 | 35,873 | 164,072 |
| 4 | 4,475 | 9,743 | 18,439 | 16,467 | 33,950 | 21,958 | 105,032 |
| 5 | 4,649 | 9,333 | 20,440 | 6,826 | 26,484 | 15,106 | 82,838 |
| 6 | 3,493 | 15,059 | 26,439 | 9,163 | 14,568 | 5,095 | 73,817 |
| 7 | 4,931 | 17,500 | 23,810 | 8,568 | 14,182 | 5,853 | 74,844 |
| 8 | 6,217 | 18,273 | 18,598 | 5,271 | 15,002 | 7,298 | 70,659 |
| 9 | 8,899 | 19,440 | 15,281 | 8,059 | 14,506 | 13,417 | 79,602 |
| 10 | 5,279 | 15,531 | 7,698 | 6,381 | 8,445 | 6,814 | 50,148 |
| 11 | 4,293 | 10,490 | 4,671 | 4,901 | 7,535 | 6,355 | 38,245 |
| 12 | 3,796 | 7,980 | 3,705 | 3,080 | 7,070 | 4,411 | 30,042 |

**Table 4.3.2.3B**

Proficiency Level by Grade (Percent): Reading S302

| | Reading Proficiency Range | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 22.0% | 13.8% | 19.6% | 8.4% | 10.1% | 26.0% | 100.0% |
| K (accountability) | 67.0% | 7.0% | 4.6% | 5.3% | 16.1% | 0.0% | 100.0% |
| 1 | 8.5% | 10.2% | 18.6% | 23.1% | 32.4% | 7.2% | 100.0% |
| 2 | 4.1% | 6.8% | 16.1% | 12.5% | 45.5% | 15.1% | 100.0% |
| 3 | 2.2% | 6.2% | 14.8% | 8.3% | 46.6% | 21.9% | 100.0% |
| 4 | 4.3% | 9.3% | 17.6% | 15.7% | 32.3% | 20.9% | 100.0% |
| 5 | 5.6% | 11.3% | 24.7% | 8.2% | 32.0% | 18.2% | 100.0% |
| 6 | 4.7% | 20.4% | 35.8% | 12.4% | 19.7% | 6.9% | 100.0% |
| 7 | 6.6% | 23.4% | 31.8% | 11.4% | 18.9% | 7.8% | 100.0% |
| 8 | 8.8% | 25.9% | 26.3% | 7.5% | 21.2% | 10.3% | 100.0% |
| 9 | 11.2% | 24.4% | 19.2% | 10.1% | 18.2% | 16.9% | 100.0% |
| 10 | 10.5% | 31.0% | 15.4% | 12.7% | 16.8% | 13.6% | 100.0% |
| 11 | 11.2% | 27.4% | 12.2% | 12.8% | 19.7% | 16.6% | 100.0% |
| 12 | 12.6% | 26.6% | 12.3% | 10.3% | 23.5% | 14.7% | 100.0% |

## 4.3.3 Writing

### 4.3.3.1 By Cluster by Tier

**Table 4.3.3.1A**

Proficiency Level by Cluster by Tier (Count): Writing S302

| Cluster | Tier | Writing Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 36,149 | 60,841 | 34,266 | 28,805 | 38,130 | 5,649 | 203,840 |
| K (accountability) | - | 119,172 | 40,889 | 25,262 | 12,868 | 5,649 | 0 | 203,840 |
| 1-2 | A | 11,849 | 54,991 | 18,534 | 0 | 1 | 0 | 85,375 |
| | B | 13,720 | 85,270 | 93,667 | 1,251 | 1 | 0 | 193,909 |
| | C | 2,667 | 21,936 | 73,696 | 11,159 | 43 | 1 | 109,502 |
| 3-5 | A | 3,727 | 7,522 | 11,201 | 6,506 | 218 | 1 | 29,175 |
| | B | 1,782 | 10,591 | 44,693 | 85,068 | 6,653 | 53 | 148,840 |
| | C | 361 | 2,034 | 13,244 | 110,455 | 46,309 | 1,607 | 174,010 |
| 6-8 | A | 4,493 | 8,027 | 9,442 | 1,735 | 30 | 2 | 23,729 |
| | B | 4,037 | 12,607 | 45,557 | 19,731 | 390 | 3 | 82,325 |
| | C | 1,955 | 9,078 | 68,680 | 33,087 | 468 | 5 | 113,273 |
| 9-12 | A | 3,991 | 8,656 | 14,473 | 1,902 | 45 | 1 | 29,068 |
| | B | 3,864 | 4,983 | 25,079 | 34,648 | 6,998 | 411 | 75,983 |
| | C | 1,812 | 1,864 | 13,041 | 41,755 | 30,897 | 3,292 | 92,661 |

**Table 4.3.3.1B**

Proficiency Level by Cluster by Tier (Percent): Writing S302

| Cluster | Tier | Writing Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 17.7% | 29.8% | 16.8% | 14.1% | 18.7% | 2.8% | 100.0% |
| K (accountability) | - | 58.5% | 20.1% | 12.4% | 6.3% | 2.8% | 0.0% | 100.0% |
| 1-2 | A | 13.9% | 64.4% | 21.7% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 7.1% | 44.0% | 48.3% | 0.6% | 0.0% | 0.0% | 100.0% |
| | C | 2.4% | 20.0% | 67.3% | 10.2% | 0.0% | 0.0% | 100.0% |
| 3-5 | A | 12.8% | 25.8% | 38.4% | 22.3% | 0.7% | 0.0% | 100.0% |
| | B | 1.2% | 7.1% | 30.0% | 57.2% | 4.5% | 0.0% | 100.0% |
| | C | 0.2% | 1.2% | 7.6% | 63.5% | 26.6% | 0.9% | 100.0% |
| 6-8 | A | 18.9% | 33.8% | 39.8% | 7.3% | 0.1% | 0.0% | 100.0% |
| | B | 4.9% | 15.3% | 55.3% | 24.0% | 0.5% | 0.0% | 100.0% |
| | C | 1.7% | 8.0% | 60.6% | 29.2% | 0.4% | 0.0% | 100.0% |
| 9-12 | A | 13.7% | 29.8% | 49.8% | 6.5% | 0.2% | 0.0% | 100.0% |
| | B | 5.1% | 6.6% | 33.0% | 45.6% | 9.2% | 0.5% | 100.0% |
| | C | 2.0% | 2.0% | 14.1% | 45.1% | 33.3% | 3.6% | 100.0% |

### 4.3.3.2    By Grade by Tier

**Table 4.3.3.2A**

Proficiency Level by Grade by Tier (Count): Writing S302

| Grade | Tier | Writing Proficiency Range 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 36,149 | 60,841 | 34,266 | 28,805 | 38,130 | 5,649 | 203,840 |
| K (accountability) | - | 119,172 | 40,889 | 25,262 | 12,868 | 5,649 | 0 | 203,840 |
| 1 | A | 8,045 | 41,655 | 17,787 | 0 | 1 | 0 | 67,488 |
| 1 | B | 8,283 | 44,277 | 37,824 | 707 | 1 | 0 | 91,092 |
| 1 | C | 1,656 | 11,473 | 25,425 | 4,101 | 13 | 1 | 42,669 |
| 2 | A | 3,804 | 13,336 | 747 | 0 | 0 | 0 | 17,887 |
| 2 | B | 5,437 | 40,993 | 55,843 | 544 | 0 | 0 | 102,817 |
| 2 | C | 1,011 | 10,463 | 48,271 | 7,058 | 30 | 0 | 66,833 |
| 3 | A | 1,169 | 2,883 | 4,554 | 3,982 | 168 | 1 | 12,757 |
| 3 | B | 672 | 4,888 | 17,131 | 46,798 | 5,376 | 43 | 74,908 |
| 3 | C | 141 | 657 | 4,768 | 43,355 | 26,437 | 1,048 | 76,406 |
| 4 | A | 1,224 | 2,468 | 3,440 | 1,610 | 36 | 0 | 8,778 |
| 4 | B | 522 | 3,329 | 13,836 | 23,815 | 983 | 6 | 42,491 |
| 4 | C | 111 | 796 | 3,481 | 35,493 | 13,534 | 400 | 53,815 |
| 5 | A | 1,334 | 2,171 | 3,207 | 914 | 14 | 0 | 7,640 |
| 5 | B | 588 | 2,374 | 13,726 | 14,455 | 294 | 4 | 31,441 |
| 5 | C | 109 | 581 | 4,995 | 31,607 | 6,338 | 159 | 43,789 |
| 6 | A | 708 | 2,904 | 3,246 | 1,040 | 20 | 1 | 7,919 |
| 6 | B | 1,170 | 3,743 | 13,623 | 9,226 | 239 | 1 | 28,002 |
| 6 | C | 648 | 2,574 | 18,512 | 15,870 | 307 | 1 | 37,912 |
| 7 | A | 1,551 | 2,578 | 3,204 | 523 | 7 | 1 | 7,864 |
| 7 | B | 1,387 | 3,920 | 15,450 | 7,262 | 119 | 2 | 28,140 |
| 7 | C | 722 | 2,755 | 24,577 | 10,670 | 111 | 1 | 38,836 |
| 8 | A | 2,234 | 2,545 | 2,992 | 172 | 3 | 0 | 7,946 |
| 8 | B | 1,480 | 4,944 | 16,484 | 3,243 | 32 | 0 | 26,183 |
| 8 | C | 585 | 3,749 | 25,591 | 6,547 | 50 | 3 | 36,525 |
| 9 | A | 2,232 | 4,803 | 7,318 | 1,340 | 33 | 1 | 15,727 |
| 9 | B | 940 | 1,611 | 6,324 | 14,667 | 4,687 | 283 | 28,512 |
| 9 | C | 409 | 500 | 2,492 | 9,710 | 19,693 | 2,456 | 35,260 |
| 10 | A | 865 | 2,019 | 3,854 | 387 | 10 | 0 | 7,135 |
| 10 | B | 1,148 | 1,354 | 6,418 | 9,476 | 1,404 | 80 | 19,880 |
| 10 | C | 425 | 489 | 3,077 | 11,725 | 6,847 | 518 | 23,081 |
| 11 | A | 501 | 1,022 | 2,336 | 138 | 2 | 0 | 3,999 |
| 11 | B | 919 | 973 | 6,159 | 6,724 | 712 | 35 | 15,522 |
| 11 | C | 398 | 398 | 3,402 | 11,114 | 3,109 | 228 | 18,649 |
| 12 | A | 393 | 812 | 965 | 37 | 0 | 0 | 2,207 |
| 12 | B | 857 | 1,045 | 6,178 | 3,781 | 195 | 13 | 12,069 |
| 12 | C | 580 | 477 | 4,070 | 9,206 | 1,248 | 90 | 15,671 |

**Table 4.3.3.2B**
Proficiency Level by Grade by Tier (Percent): Writing S302

| Grade | Tier | Writing Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 17.7% | 29.8% | 16.8% | 14.1% | 18.7% | 2.8% | 100.0% |
| K (accountability) | - | 58.5% | 20.1% | 12.4% | 6.3% | 2.8% | 0.0% | 100.0% |
| 1 | A | 11.9% | 61.7% | 26.4% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 9.1% | 48.6% | 41.5% | 0.8% | 0.0% | 0.0% | 100.0% |
| | C | 3.9% | 26.9% | 59.6% | 9.6% | 0.0% | 0.0% | 100.0% |
| 2 | A | 21.3% | 74.6% | 4.2% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 5.3% | 39.9% | 54.3% | 0.5% | 0.0% | 0.0% | 100.0% |
| | C | 1.5% | 15.7% | 72.2% | 10.6% | 0.0% | 0.0% | 100.0% |
| 3 | A | 9.2% | 22.6% | 35.7% | 31.2% | 1.3% | 0.0% | 100.0% |
| | B | 0.9% | 6.5% | 22.9% | 62.5% | 7.2% | 0.1% | 100.0% |
| | C | 0.2% | 0.9% | 6.2% | 56.7% | 34.6% | 1.4% | 100.0% |
| 4 | A | 13.9% | 28.1% | 39.2% | 18.3% | 0.4% | 0.0% | 100.0% |
| | B | 1.2% | 7.8% | 32.6% | 56.0% | 2.3% | 0.0% | 100.0% |
| | C | 0.2% | 1.5% | 6.5% | 66.0% | 25.1% | 0.7% | 100.0% |
| 5 | A | 17.5% | 28.4% | 42.0% | 12.0% | 0.2% | 0.0% | 100.0% |
| | B | 1.9% | 7.6% | 43.7% | 46.0% | 0.9% | 0.0% | 100.0% |
| | C | 0.2% | 1.3% | 11.4% | 72.2% | 14.5% | 0.4% | 100.0% |
| 6 | A | 8.9% | 36.7% | 41.0% | 13.1% | 0.3% | 0.0% | 100.0% |
| | B | 4.2% | 13.4% | 48.7% | 32.9% | 0.9% | 0.0% | 100.0% |
| | C | 1.7% | 6.8% | 48.8% | 41.9% | 0.8% | 0.0% | 100.0% |
| 7 | A | 19.7% | 32.8% | 40.7% | 6.7% | 0.1% | 0.0% | 100.0% |
| | B | 4.9% | 13.9% | 54.9% | 25.8% | 0.4% | 0.0% | 100.0% |
| | C | 1.9% | 7.1% | 63.3% | 27.5% | 0.3% | 0.0% | 100.0% |
| 8 | A | 28.1% | 32.0% | 37.7% | 2.2% | 0.0% | 0.0% | 100.0% |
| | B | 5.7% | 18.9% | 63.0% | 12.4% | 0.1% | 0.0% | 100.0% |
| | C | 1.6% | 10.3% | 70.1% | 17.9% | 0.1% | 0.0% | 100.0% |
| 9 | A | 14.2% | 30.5% | 46.5% | 8.5% | 0.2% | 0.0% | 100.0% |
| | B | 3.3% | 5.7% | 22.2% | 51.4% | 16.4% | 1.0% | 100.0% |
| | C | 1.2% | 1.4% | 7.1% | 27.5% | 55.9% | 7.0% | 100.0% |
| 10 | A | 12.1% | 28.3% | 54.0% | 5.4% | 0.1% | 0.0% | 100.0% |
| | B | 5.8% | 6.8% | 32.3% | 47.7% | 7.1% | 0.4% | 100.0% |
| | C | 1.8% | 2.1% | 13.3% | 50.8% | 29.7% | 2.2% | 100.0% |
| 11 | A | 12.5% | 25.6% | 58.4% | 3.5% | 0.1% | 0.0% | 100.0% |
| | B | 5.9% | 6.3% | 39.7% | 43.3% | 4.6% | 0.2% | 100.0% |
| | C | 2.1% | 2.1% | 18.2% | 59.6% | 16.7% | 1.2% | 100.0% |
| 12 | A | 17.8% | 36.8% | 43.7% | 1.7% | 0.0% | 0.0% | 100.0% |
| | B | 7.1% | 8.7% | 51.2% | 31.3% | 1.6% | 0.1% | 100.0% |
| | C | 3.7% | 3.0% | 26.0% | 58.7% | 8.0% | 0.6% | 100.0% |

### 4.3.3.3    By Grade

**Table 4.3.3.3A**
Proficiency Level by Grade (Count): Writing S302

| | Writing Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 36,149 | 60,841 | 34,266 | 28,805 | 38,130 | 5,649 | 203,840 |
| K (accountability) | 119,172 | 40,889 | 25,262 | 12,868 | 5,649 | 0 | 203,840 |
| 1 | 17,984 | 97,405 | 81,036 | 4,808 | 15 | 1 | 201,249 |
| 2 | 10,252 | 64,792 | 104,861 | 7,602 | 30 | 0 | 187,537 |
| 3 | 1,982 | 8,428 | 26,453 | 94,135 | 31,981 | 1,092 | 164,071 |
| 4 | 1,857 | 6,593 | 20,757 | 60,918 | 14,553 | 406 | 105,084 |
| 5 | 2,031 | 5,126 | 21,928 | 46,976 | 6,646 | 163 | 82,870 |
| 6 | 2,526 | 9,221 | 35,381 | 26,136 | 566 | 3 | 73,833 |
| 7 | 3,660 | 9,253 | 43,231 | 18,455 | 237 | 4 | 74,840 |
| 8 | 4,299 | 11,238 | 45,067 | 9,962 | 85 | 3 | 70,654 |
| 9 | 3,581 | 6,914 | 16,134 | 25,717 | 24,413 | 2,740 | 79,499 |
| 10 | 2,438 | 3,862 | 13,349 | 21,588 | 8,261 | 598 | 50,096 |
| 11 | 1,818 | 2,393 | 11,897 | 17,976 | 3,823 | 263 | 38,170 |
| 12 | 1,830 | 2,334 | 11,213 | 13,024 | 1,443 | 103 | 29,947 |

**Table 4.3.3.3B**
Proficiency Level by Grade (Percent): Writing S302

| | Writing Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 17.7% | 29.8% | 16.8% | 14.1% | 18.7% | 2.8% | 100.0% |
| K (accountability) | 58.5% | 20.1% | 12.4% | 6.3% | 2.8% | 0.0% | 100.0% |
| 1 | 8.9% | 48.4% | 40.3% | 2.4% | 0.0% | 0.0% | 100.0% |
| 2 | 5.5% | 34.5% | 55.9% | 4.1% | 0.0% | 0.0% | 100.0% |
| 3 | 1.2% | 5.1% | 16.1% | 57.4% | 19.5% | 0.7% | 100.0% |
| 4 | 1.8% | 6.3% | 19.8% | 58.0% | 13.8% | 0.4% | 100.0% |
| 5 | 2.5% | 6.2% | 26.5% | 56.7% | 8.0% | 0.2% | 100.0% |
| 6 | 3.4% | 12.5% | 47.9% | 35.4% | 0.8% | 0.0% | 100.0% |
| 7 | 4.9% | 12.4% | 57.8% | 24.7% | 0.3% | 0.0% | 100.0% |
| 8 | 6.1% | 15.9% | 63.8% | 14.1% | 0.1% | 0.0% | 100.0% |
| 9 | 4.5% | 8.7% | 20.3% | 32.3% | 30.7% | 3.4% | 100.0% |
| 10 | 4.9% | 7.7% | 26.6% | 43.1% | 16.5% | 1.2% | 100.0% |
| 11 | 4.8% | 6.3% | 31.2% | 47.1% | 10.0% | 0.7% | 100.0% |
| 12 | 6.1% | 7.8% | 37.4% | 43.5% | 4.8% | 0.3% | 100.0% |

## 4.3.4    Speaking

### 4.3.4.1    By Cluster by Tier

**Table 4.3.4.1A**

Proficiency Level by Cluster by Tier (Count): Speaking S302

| Cluster | Tier | Speaking Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 47,008 | 16,092 | 31,076 | 34,064 | 22,635 | 52,952 | 203,827 |
| K (accountability) | - | 47,008 | 47,168 | 34,064 | 22,635 | 52,952 | 0 | 203,827 |
| 1-2 | A | 18,932 | 30,606 | 13,802 | 5,141 | 3,658 | 13,208 | 85,347 |
| | B | 7,889 | 39,061 | 39,135 | 19,772 | 17,716 | 70,325 | 193,898 |
| | C | 1,499 | 8,893 | 13,748 | 9,641 | 10,576 | 65,200 | 109,557 |
| 3-5 | A | 14,926 | 7,166 | 2,867 | 1,113 | 811 | 2,347 | 29,230 |
| | B | 9,034 | 31,055 | 33,191 | 17,558 | 15,577 | 42,432 | 148,847 |
| | C | 3,618 | 17,324 | 27,975 | 21,930 | 24,741 | 78,440 | 174,028 |
| 6-8 | A | 12,653 | 4,911 | 2,789 | 1,545 | 523 | 1,325 | 23,746 |
| | B | 3,759 | 9,173 | 16,851 | 18,296 | 9,133 | 25,055 | 82,267 |
| | C | 1,466 | 4,543 | 13,536 | 23,871 | 16,236 | 53,554 | 113,206 |
| 9-12 | A | 18,440 | 4,779 | 2,503 | 1,285 | 532 | 1,513 | 29,052 |
| | B | 7,673 | 10,819 | 13,106 | 10,991 | 7,316 | 26,066 | 75,971 |
| | C | 2,245 | 4,409 | 10,136 | 13,026 | 11,697 | 51,079 | 92,592 |

**Table 4.3.4.1B**

Proficiency Level by Cluster by Tier (Percent): Speaking S302

| Cluster | Tier | Speaking Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 23.1% | 7.9% | 15.2% | 16.7% | 11.1% | 26.0% | 100.0% |
| K (accountability) | - | 23.1% | 23.1% | 16.7% | 11.1% | 26.0% | 0.0% | 100.0% |
| 1-2 | A | 22.2% | 35.9% | 16.2% | 6.0% | 4.3% | 15.5% | 100.0% |
| | B | 4.1% | 20.1% | 20.2% | 10.2% | 9.1% | 36.3% | 100.0% |
| | C | 1.4% | 8.1% | 12.5% | 8.8% | 9.7% | 59.5% | 100.0% |
| 3-5 | A | 51.1% | 24.5% | 9.8% | 3.8% | 2.8% | 8.0% | 100.0% |
| | B | 6.1% | 20.9% | 22.3% | 11.8% | 10.5% | 28.5% | 100.0% |
| | C | 2.1% | 10.0% | 16.1% | 12.6% | 14.2% | 45.1% | 100.0% |
| 6-8 | A | 53.3% | 20.7% | 11.7% | 6.5% | 2.2% | 5.6% | 100.0% |
| | B | 4.6% | 11.2% | 20.5% | 22.2% | 11.1% | 30.5% | 100.0% |
| | C | 1.3% | 4.0% | 12.0% | 21.1% | 14.3% | 47.3% | 100.0% |
| 9-12 | A | 63.5% | 16.4% | 8.6% | 4.4% | 1.8% | 5.2% | 100.0% |
| | B | 10.1% | 14.2% | 17.3% | 14.5% | 9.6% | 34.3% | 100.0% |
| | C | 2.4% | 4.8% | 10.9% | 14.1% | 12.6% | 55.2% | 100.0% |

### 4.3.4.2 By Grade by Tier

**Table 4.3.4.2A**

Proficiency Level by Grade by Tier (Count): Speaking S302

| Grade | Tier | Speaking Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 47,008 | 16,092 | 31,076 | 34,064 | 22,635 | 52,952 | 203,827 |
| K (accountability) | - | 47,008 | 47,168 | 34,064 | 22,635 | 52,952 | 0 | 203,827 |
| 1 | A | 12,401 | 26,010 | 11,295 | 4,205 | 2,968 | 10,577 | 67,456 |
| | B | 3,750 | 23,312 | 19,287 | 9,069 | 7,789 | 27,888 | 91,095 |
| | C | 643 | 5,161 | 6,585 | 4,156 | 4,077 | 22,071 | 42,693 |
| 2 | A | 6,531 | 4,596 | 2,507 | 936 | 690 | 2,631 | 17,891 |
| | B | 4,139 | 15,749 | 19,848 | 10,703 | 9,927 | 42,437 | 102,803 |
| | C | 856 | 3,732 | 7,163 | 5,485 | 6,499 | 43,129 | 66,864 |
| 3 | A | 5,858 | 3,787 | 1,298 | 478 | 339 | 1,025 | 12,785 |
| | B | 4,063 | 17,954 | 17,376 | 8,593 | 7,483 | 19,450 | 74,919 |
| | C | 1,374 | 8,945 | 13,159 | 9,631 | 10,663 | 32,665 | 76,437 |
| 4 | A | 4,830 | 1,840 | 821 | 332 | 258 | 702 | 8,783 |
| | B | 3,012 | 8,006 | 9,248 | 5,171 | 4,601 | 12,454 | 42,492 |
| | C | 1,375 | 5,224 | 8,759 | 7,095 | 7,795 | 23,571 | 53,819 |
| 5 | A | 4,238 | 1,539 | 748 | 303 | 214 | 620 | 7,662 |
| | B | 1,959 | 5,095 | 6,567 | 3,794 | 3,493 | 10,528 | 31,436 |
| | C | 869 | 3,155 | 6,057 | 5,204 | 6,283 | 22,204 | 43,772 |
| 6 | A | 3,657 | 1,923 | 1,043 | 654 | 181 | 468 | 7,926 |
| | B | 1,048 | 2,997 | 5,659 | 7,550 | 2,939 | 7,796 | 27,989 |
| | C | 501 | 1,650 | 4,729 | 9,990 | 5,385 | 15,635 | 37,890 |
| 7 | A | 4,452 | 1,218 | 1,001 | 628 | 170 | 407 | 7,876 |
| | B | 1,430 | 2,332 | 5,277 | 7,299 | 3,146 | 8,638 | 28,122 |
| | C | 524 | 1,065 | 3,962 | 9,291 | 5,614 | 18,365 | 38,821 |
| 8 | A | 4,544 | 1,770 | 745 | 263 | 172 | 450 | 7,944 |
| | B | 1,281 | 3,844 | 5,915 | 3,447 | 3,048 | 8,621 | 26,156 |
| | C | 441 | 1,828 | 4,845 | 4,590 | 5,237 | 19,554 | 36,495 |
| 9 | A | 10,938 | 1,836 | 1,250 | 809 | 239 | 653 | 15,725 |
| | B | 2,624 | 2,404 | 4,117 | 6,005 | 3,040 | 10,312 | 28,502 |
| | C | 672 | 783 | 2,616 | 6,381 | 4,520 | 20,253 | 35,225 |
| 10 | A | 4,304 | 1,477 | 616 | 214 | 148 | 386 | 7,145 |
| | B | 2,382 | 3,549 | 3,725 | 2,084 | 1,789 | 6,335 | 19,864 |
| | C | 577 | 1,341 | 3,010 | 2,666 | 3,021 | 12,427 | 23,042 |
| 11 | A | 2,145 | 930 | 402 | 159 | 80 | 272 | 3,988 |
| | B | 1,554 | 2,865 | 2,989 | 1,636 | 1,384 | 5,108 | 15,536 |
| | C | 512 | 1,179 | 2,484 | 2,185 | 2,266 | 10,014 | 18,640 |
| 12 | A | 1,053 | 536 | 235 | 103 | 65 | 202 | 2,194 |
| | B | 1,113 | 2,001 | 2,275 | 1,266 | 1,103 | 4,311 | 12,069 |
| | C | 484 | 1,106 | 2,026 | 1,794 | 1,890 | 8,385 | 15,685 |

**Table 4.3.4.2B**
Proficiency Level by Grade by Tier (Percent): Speaking S302

| Grade | Tier | Speaking Proficiency Range | | | | | | Total |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 23.1% | 7.9% | 15.2% | 16.7% | 11.1% | 26.0% | 100.0% |
| K (accountability) | - | 23.1% | 23.1% | 16.7% | 11.1% | 26.0% | 0.0% | 100.0% |
| 1 | A | 18.4% | 38.6% | 16.7% | 6.2% | 4.4% | 15.7% | 100.0% |
| | B | 4.1% | 25.6% | 21.2% | 10.0% | 8.6% | 30.6% | 100.0% |
| | C | 1.5% | 12.1% | 15.4% | 9.7% | 9.5% | 51.7% | 100.0% |
| 2 | A | 36.5% | 25.7% | 14.0% | 5.2% | 3.9% | 14.7% | 100.0% |
| | B | 4.0% | 15.3% | 19.3% | 10.4% | 9.7% | 41.3% | 100.0% |
| | C | 1.3% | 5.6% | 10.7% | 8.2% | 9.7% | 64.5% | 100.0% |
| 3 | A | 45.8% | 29.6% | 10.2% | 3.7% | 2.7% | 8.0% | 100.0% |
| | B | 5.4% | 24.0% | 23.2% | 11.5% | 10.0% | 26.0% | 100.0% |
| | C | 1.8% | 11.7% | 17.2% | 12.6% | 14.0% | 42.7% | 100.0% |
| 4 | A | 55.0% | 20.9% | 9.3% | 3.8% | 2.9% | 8.0% | 100.0% |
| | B | 7.1% | 18.8% | 21.8% | 12.2% | 10.8% | 29.3% | 100.0% |
| | C | 2.6% | 9.7% | 16.3% | 13.2% | 14.5% | 43.8% | 100.0% |
| 5 | A | 55.3% | 20.1% | 9.8% | 4.0% | 2.8% | 8.1% | 100.0% |
| | B | 6.2% | 16.2% | 20.9% | 12.1% | 11.1% | 33.5% | 100.0% |
| | C | 2.0% | 7.2% | 13.8% | 11.9% | 14.4% | 50.7% | 100.0% |
| 6 | A | 46.1% | 24.3% | 13.2% | 8.3% | 2.3% | 5.9% | 100.0% |
| | B | 3.7% | 10.7% | 20.2% | 27.0% | 10.5% | 27.9% | 100.0% |
| | C | 1.3% | 4.4% | 12.5% | 26.4% | 14.2% | 41.3% | 100.0% |
| 7 | A | 56.5% | 15.5% | 12.7% | 8.0% | 2.2% | 5.2% | 100.0% |
| | B | 5.1% | 8.3% | 18.8% | 26.0% | 11.2% | 30.7% | 100.0% |
| | C | 1.3% | 2.7% | 10.2% | 23.9% | 14.5% | 47.3% | 100.0% |
| 8 | A | 57.2% | 22.3% | 9.4% | 3.3% | 2.2% | 5.7% | 100.0% |
| | B | 4.9% | 14.7% | 22.6% | 13.2% | 11.7% | 33.0% | 100.0% |
| | C | 1.2% | 5.0% | 13.3% | 12.6% | 14.3% | 53.6% | 100.0% |
| 9 | A | 69.6% | 11.7% | 7.9% | 5.1% | 1.5% | 4.2% | 100.0% |
| | B | 9.2% | 8.4% | 14.4% | 21.1% | 10.7% | 36.2% | 100.0% |
| | C | 1.9% | 2.2% | 7.4% | 18.1% | 12.8% | 57.5% | 100.0% |
| 10 | A | 60.2% | 20.7% | 8.6% | 3.0% | 2.1% | 5.4% | 100.0% |
| | B | 12.0% | 17.9% | 18.8% | 10.5% | 9.0% | 31.9% | 100.0% |
| | C | 2.5% | 5.8% | 13.1% | 11.6% | 13.1% | 53.9% | 100.0% |
| 11 | A | 53.8% | 23.3% | 10.1% | 4.0% | 2.0% | 6.8% | 100.0% |
| | B | 10.0% | 18.4% | 19.2% | 10.5% | 8.9% | 32.9% | 100.0% |
| | C | 2.7% | 6.3% | 13.3% | 11.7% | 12.2% | 53.7% | 100.0% |
| 12 | A | 48.0% | 24.4% | 10.7% | 4.7% | 3.0% | 9.2% | 100.0% |
| | B | 9.2% | 16.6% | 18.8% | 10.5% | 9.1% | 35.7% | 100.0% |
| | C | 3.1% | 7.1% | 12.9% | 11.4% | 12.0% | 53.5% | 100.0% |

### 4.3.4.3 By Grade

**Table 4.3.4.3A**
Proficiency Level by Grade (Count): Speaking S302

| | Speaking Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 47,008 | 16,092 | 31,076 | 34,064 | 22,635 | 52,952 | 203,827 |
| K (accountability) | 47,008 | 47,168 | 34,064 | 22,635 | 52,952 | 0 | 203,827 |
| 1 | 16,794 | 54,483 | 37,167 | 17,430 | 14,834 | 60,536 | 201,244 |
| 2 | 11,526 | 24,077 | 29,518 | 17,124 | 17,116 | 88,197 | 187,558 |
| 3 | 11,295 | 30,686 | 31,833 | 18,702 | 18,485 | 53,140 | 164,141 |
| 4 | 9,217 | 15,070 | 18,828 | 12,598 | 12,654 | 36,727 | 105,094 |
| 5 | 7,066 | 9,789 | 13,372 | 9,301 | 9,990 | 33,352 | 82,870 |
| 6 | 5,206 | 6,570 | 11,431 | 18,194 | 8,505 | 23,899 | 73,805 |
| 7 | 6,406 | 4,615 | 10,240 | 17,218 | 8,930 | 27,410 | 74,819 |
| 8 | 6,266 | 7,442 | 11,505 | 8,300 | 8,457 | 28,625 | 70,595 |
| 9 | 14,234 | 5,023 | 7,983 | 13,195 | 7,799 | 31,218 | 79,452 |
| 10 | 7,263 | 6,367 | 7,351 | 4,964 | 4,958 | 19,148 | 50,051 |
| 11 | 4,211 | 4,974 | 5,875 | 3,980 | 3,730 | 15,394 | 38,164 |
| 12 | 2,650 | 3,643 | 4,536 | 3,163 | 3,058 | 12,898 | 29,948 |

**Table 4.3.4.3B**
Proficiency Level by Grade (Percent): Speaking S302

| | Speaking Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 23.1% | 7.9% | 15.2% | 16.7% | 11.1% | 26.0% | 100.0% |
| K (accountability) | 23.1% | 23.1% | 16.7% | 11.1% | 26.0% | 0.0% | 100.0% |
| 1 | 8.3% | 27.1% | 18.5% | 8.7% | 7.4% | 30.1% | 100.0% |
| 2 | 6.1% | 12.8% | 15.7% | 9.1% | 9.1% | 47.0% | 100.0% |
| 3 | 6.9% | 18.7% | 19.4% | 11.4% | 11.3% | 32.4% | 100.0% |
| 4 | 8.8% | 14.3% | 17.9% | 12.0% | 12.0% | 34.9% | 100.0% |
| 5 | 8.5% | 11.8% | 16.1% | 11.2% | 12.1% | 40.2% | 100.0% |
| 6 | 7.1% | 8.9% | 15.5% | 24.7% | 11.5% | 32.4% | 100.0% |
| 7 | 8.6% | 6.2% | 13.7% | 23.0% | 11.9% | 36.6% | 100.0% |
| 8 | 8.9% | 10.5% | 16.3% | 11.8% | 12.0% | 40.5% | 100.0% |
| 9 | 17.9% | 6.3% | 10.0% | 16.6% | 9.8% | 39.3% | 100.0% |
| 10 | 14.5% | 12.7% | 14.7% | 9.9% | 9.9% | 38.3% | 100.0% |
| 11 | 11.0% | 13.0% | 15.4% | 10.4% | 9.8% | 40.3% | 100.0% |
| 12 | 8.8% | 12.2% | 15.1% | 10.6% | 10.2% | 43.1% | 100.0% |

## 4.3.5 Oral Language Composite

### 4.3.5.1 By Cluster by Tier

**Table 4.3.5.1A**

Proficiency Level by Cluster by Tier (Count): Oral S302

| Cluster | Tier | Oral Language Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 29,747 | 18,588 | 24,444 | 41,756 | 45,977 | 43,311 | 203,823 |
| K (accountability) | - | 51,598 | 29,375 | 33,562 | 17,908 | 28,069 | 43,311 | 203,823 |
| 1-2 | A | 11,305 | 21,318 | 31,714 | 8,565 | 12,372 | 0 | 85,274 |
| | B | 1,821 | 9,002 | 59,496 | 37,141 | 86,351 | 0 | 193,811 |
| | C | 598 | 3,383 | 13,116 | 15,331 | 35,904 | 41,167 | 109,499 |
| 3-5 | A | 8,285 | 9,772 | 7,050 | 1,896 | 2,195 | 0 | 29,198 |
| | B | 1,761 | 10,930 | 37,649 | 43,908 | 54,537 | 0 | 148,785 |
| | C | 499 | 3,017 | 14,673 | 28,869 | 57,223 | 69,676 | 173,957 |
| 6-8 | A | 10,390 | 6,767 | 3,874 | 1,649 | 1,024 | 0 | 23,704 |
| | B | 1,670 | 8,283 | 17,127 | 28,130 | 26,969 | 0 | 82,179 |
| | C | 482 | 1,005 | 6,366 | 20,153 | 38,503 | 46,587 | 113,096 |
| 9-12 | A | 16,354 | 7,830 | 2,692 | 1,331 | 727 | 0 | 28,934 |
| | B | 3,167 | 11,033 | 15,884 | 19,721 | 25,809 | 0 | 75,614 |
| | C | 1,506 | 2,739 | 9,730 | 22,616 | 31,411 | 24,153 | 92,155 |

**Table 4.3.5.1B**

Proficiency Level by Cluster by Tier (Percent): Oral S302

| Cluster | Tier | Oral Language Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 14.6% | 9.1% | 12.0% | 20.5% | 22.6% | 21.2% | 100.0% |
| K (accountability) | - | 25.3% | 14.4% | 16.5% | 8.8% | 13.8% | 21.2% | 100.0% |
| 1-2 | A | 13.3% | 25.0% | 37.2% | 10.0% | 14.5% | 0.0% | 100.0% |
| | B | 0.9% | 4.6% | 30.7% | 19.2% | 44.6% | 0.0% | 100.0% |
| | C | 0.5% | 3.1% | 12.0% | 14.0% | 32.8% | 37.6% | 100.0% |
| 3-5 | A | 28.4% | 33.5% | 24.1% | 6.5% | 7.5% | 0.0% | 100.0% |
| | B | 1.2% | 7.3% | 25.3% | 29.5% | 36.7% | 0.0% | 100.0% |
| | C | 0.3% | 1.7% | 8.4% | 16.6% | 32.9% | 40.1% | 100.0% |
| 6-8 | A | 43.8% | 28.5% | 16.3% | 7.0% | 4.3% | 0.0% | 100.0% |
| | B | 2.0% | 10.1% | 20.8% | 34.2% | 32.8% | 0.0% | 100.0% |
| | C | 0.4% | 0.9% | 5.6% | 17.8% | 34.0% | 41.2% | 100.0% |
| 9-12 | A | 56.5% | 27.1% | 9.3% | 4.6% | 2.5% | 0.0% | 100.0% |
| | B | 4.2% | 14.6% | 21.0% | 26.1% | 34.1% | 0.0% | 100.0% |
| | C | 1.6% | 3.0% | 10.6% | 24.5% | 34.1% | 26.2% | 100.0% |

## 4.3.5.2 By Grade by Tier

**Table 4.3.5.2A**

Proficiency Level by Grade by Tier (Count): Oral S302

| Grade | Tier | Oral Language Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 29,747 | 18,588 | 24,444 | 41,756 | 45,977 | 43,311 | 203,823 |
| K (accountability) | - | 51,598 | 29,375 | 33,562 | 17,908 | 28,069 | 43,311 | 203,823 |
| 1 | A | 7,278 | 16,702 | 26,494 | 6,957 | 9,978 | 0 | 67,409 |
| | B | 1,009 | 5,353 | 32,356 | 17,304 | 35,034 | 0 | 91,056 |
| | C | 276 | 1,896 | 6,752 | 6,244 | 12,444 | 15,055 | 42,667 |
| 2 | A | 4,027 | 4,616 | 5,220 | 1,608 | 2,394 | 0 | 17,865 |
| | B | 812 | 3,649 | 27,140 | 19,837 | 51,317 | 0 | 102,755 |
| | C | 322 | 1,487 | 6,364 | 9,087 | 23,460 | 26,112 | 66,832 |
| 3 | A | 2,925 | 4,645 | 3,420 | 805 | 982 | 0 | 12,777 |
| | B | 694 | 5,310 | 20,769 | 22,342 | 25,774 | 0 | 74,889 |
| | C | 190 | 1,175 | 5,793 | 12,445 | 22,323 | 34,474 | 76,400 |
| 4 | A | 2,554 | 2,865 | 2,111 | 587 | 655 | 0 | 8,772 |
| | B | 565 | 2,955 | 10,322 | 12,684 | 15,948 | 0 | 42,474 |
| | C | 163 | 982 | 5,276 | 9,021 | 17,472 | 20,883 | 53,797 |
| 5 | A | 2,806 | 2,262 | 1,519 | 504 | 558 | 0 | 7,649 |
| | B | 502 | 2,665 | 6,558 | 8,882 | 12,815 | 0 | 31,422 |
| | C | 146 | 860 | 3,604 | 7,403 | 17,428 | 14,319 | 43,760 |
| 6 | A | 2,895 | 2,468 | 1,446 | 705 | 402 | 0 | 7,916 |
| | B | 430 | 2,410 | 5,532 | 10,408 | 9,188 | 0 | 27,968 |
| | C | 125 | 365 | 2,062 | 7,251 | 13,726 | 14,328 | 37,857 |
| 7 | A | 3,589 | 2,128 | 1,295 | 518 | 328 | 0 | 7,858 |
| | B | 565 | 2,886 | 6,161 | 8,886 | 9,590 | 0 | 28,088 |
| | C | 177 | 298 | 2,291 | 7,031 | 12,412 | 16,567 | 38,776 |
| 8 | A | 3,906 | 2,171 | 1,133 | 426 | 294 | 0 | 7,930 |
| | B | 675 | 2,987 | 5,434 | 8,836 | 8,191 | 0 | 26,123 |
| | C | 180 | 342 | 2,013 | 5,871 | 12,365 | 15,692 | 36,463 |
| 9 | A | 9,155 | 4,252 | 1,213 | 603 | 440 | 0 | 15,663 |
| | B | 899 | 3,216 | 4,682 | 7,992 | 11,579 | 0 | 28,368 |
| | C | 412 | 504 | 2,156 | 6,751 | 12,871 | 12,403 | 35,097 |
| 10 | A | 3,825 | 1,990 | 829 | 275 | 201 | 0 | 7,120 |
| | B | 860 | 3,457 | 4,099 | 4,935 | 6,453 | 0 | 19,804 |
| | C | 356 | 614 | 2,645 | 5,774 | 8,518 | 5,044 | 22,951 |
| 11 | A | 2,156 | 1,082 | 384 | 261 | 86 | 0 | 3,969 |
| | B | 710 | 2,419 | 3,993 | 3,699 | 4,642 | 0 | 15,463 |
| | C | 332 | 800 | 2,307 | 5,368 | 5,362 | 4,377 | 18,546 |
| 12 | A | 1,218 | 506 | 266 | 192 | 0 | 0 | 2,182 |
| | B | 698 | 1,941 | 3,110 | 3,095 | 3,135 | 0 | 11,979 |
| | C | 406 | 821 | 2,622 | 4,723 | 4,660 | 2,329 | 15,561 |

**Table 4.3.5.2B**

Proficiency Level by Grade by Tier (Percent): Oral S302

| Grade | Tier | Oral Language Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 14.6% | 9.1% | 12.0% | 20.5% | 22.6% | 21.2% | 100.0% |
| K (accountability) | - | 25.3% | 14.4% | 16.5% | 8.8% | 13.8% | 21.2% | 100.0% |
| 1 | A | 10.8% | 24.8% | 39.3% | 10.3% | 14.8% | 0.0% | 100.0% |
| | B | 1.1% | 5.9% | 35.5% | 19.0% | 38.5% | 0.0% | 100.0% |
| | C | 0.6% | 4.4% | 15.8% | 14.6% | 29.2% | 35.3% | 100.0% |
| 2 | A | 22.5% | 25.8% | 29.2% | 9.0% | 13.4% | 0.0% | 100.0% |
| | B | 0.8% | 3.6% | 26.4% | 19.3% | 49.9% | 0.0% | 100.0% |
| | C | 0.5% | 2.2% | 9.5% | 13.6% | 35.1% | 39.1% | 100.0% |
| 3 | A | 22.9% | 36.4% | 26.8% | 6.3% | 7.7% | 0.0% | 100.0% |
| | B | 0.9% | 7.1% | 27.7% | 29.8% | 34.4% | 0.0% | 100.0% |
| | C | 0.2% | 1.5% | 7.6% | 16.3% | 29.2% | 45.1% | 100.0% |
| 4 | A | 29.1% | 32.7% | 24.1% | 6.7% | 7.5% | 0.0% | 100.0% |
| | B | 1.3% | 7.0% | 24.3% | 29.9% | 37.5% | 0.0% | 100.0% |
| | C | 0.3% | 1.8% | 9.8% | 16.8% | 32.5% | 38.8% | 100.0% |
| 5 | A | 36.7% | 29.6% | 19.9% | 6.6% | 7.3% | 0.0% | 100.0% |
| | B | 1.6% | 8.5% | 20.9% | 28.3% | 40.8% | 0.0% | 100.0% |
| | C | 0.3% | 2.0% | 8.2% | 16.9% | 39.8% | 32.7% | 100.0% |
| 6 | A | 36.6% | 31.2% | 18.3% | 8.9% | 5.1% | 0.0% | 100.0% |
| | B | 1.5% | 8.6% | 19.8% | 37.2% | 32.9% | 0.0% | 100.0% |
| | C | 0.3% | 1.0% | 5.4% | 19.2% | 36.3% | 37.8% | 100.0% |
| 7 | A | 45.7% | 27.1% | 16.5% | 6.6% | 4.2% | 0.0% | 100.0% |
| | B | 2.0% | 10.3% | 21.9% | 31.6% | 34.1% | 0.0% | 100.0% |
| | C | 0.5% | 0.8% | 5.9% | 18.1% | 32.0% | 42.7% | 100.0% |
| 8 | A | 49.3% | 27.4% | 14.3% | 5.4% | 3.7% | 0.0% | 100.0% |
| | B | 2.6% | 11.4% | 20.8% | 33.8% | 31.4% | 0.0% | 100.0% |
| | C | 0.5% | 0.9% | 5.5% | 16.1% | 33.9% | 43.0% | 100.0% |
| 9 | A | 58.4% | 27.1% | 7.7% | 3.8% | 2.8% | 0.0% | 100.0% |
| | B | 3.2% | 11.3% | 16.5% | 28.2% | 40.8% | 0.0% | 100.0% |
| | C | 1.2% | 1.4% | 6.1% | 19.2% | 36.7% | 35.3% | 100.0% |
| 10 | A | 53.7% | 27.9% | 11.6% | 3.9% | 2.8% | 0.0% | 100.0% |
| | B | 4.3% | 17.5% | 20.7% | 24.9% | 32.6% | 0.0% | 100.0% |
| | C | 1.6% | 2.7% | 11.5% | 25.2% | 37.1% | 22.0% | 100.0% |
| 11 | A | 54.3% | 27.3% | 9.7% | 6.6% | 2.2% | 0.0% | 100.0% |
| | B | 4.6% | 15.6% | 25.8% | 23.9% | 30.0% | 0.0% | 100.0% |
| | C | 1.8% | 4.3% | 12.4% | 28.9% | 28.9% | 23.6% | 100.0% |
| 12 | A | 55.8% | 23.2% | 12.2% | 8.8% | 0.0% | 0.0% | 100.0% |
| | B | 5.8% | 16.2% | 26.0% | 25.8% | 26.2% | 0.0% | 100.0% |
| | C | 2.6% | 5.3% | 16.8% | 30.4% | 29.9% | 15.0% | 100.0% |

### 4.3.5.3 By Grade

**Table 4.3.5.3A**

Proficiency Level by Grade (Count): Oral S302

| | Oral Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 29,747 | 18,588 | 24,444 | 41,756 | 45,977 | 43,311 | 203,823 |
| K (accountability) | 51,598 | 29,375 | 33,562 | 17,908 | 28,069 | 43,311 | 203,823 |
| 1 | 8,563 | 23,951 | 65,602 | 30,505 | 57,456 | 15,055 | 201,132 |
| 2 | 5,161 | 9,752 | 38,724 | 30,532 | 77,171 | 26,112 | 187,452 |
| 3 | 3,809 | 11,130 | 29,982 | 35,592 | 49,079 | 34,474 | 164,066 |
| 4 | 3,282 | 6,802 | 17,709 | 22,292 | 34,075 | 20,883 | 105,043 |
| 5 | 3,454 | 5,787 | 11,681 | 16,789 | 30,801 | 14,319 | 82,831 |
| 6 | 3,450 | 5,243 | 9,040 | 18,364 | 23,316 | 14,328 | 73,741 |
| 7 | 4,331 | 5,312 | 9,747 | 16,435 | 22,330 | 16,567 | 74,722 |
| 8 | 4,761 | 5,500 | 8,580 | 15,133 | 20,850 | 15,692 | 70,516 |
| 9 | 10,466 | 7,972 | 8,051 | 15,346 | 24,890 | 12,403 | 79,128 |
| 10 | 5,041 | 6,061 | 7,573 | 10,984 | 15,172 | 5,044 | 49,875 |
| 11 | 3,198 | 4,301 | 6,684 | 9,328 | 10,090 | 4,377 | 37,978 |
| 12 | 2,322 | 3,268 | 5,998 | 8,010 | 7,795 | 2,329 | 29,722 |

**Table 4.3.5.3B**

Proficiency Level by Grade (Percent): Oral S302

| | Oral Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 14.6% | 9.1% | 12.0% | 20.5% | 22.6% | 21.2% | 100.0% |
| K (accountability) | 25.3% | 14.4% | 16.5% | 8.8% | 13.8% | 21.2% | 100.0% |
| 1 | 4.3% | 11.9% | 32.6% | 15.2% | 28.6% | 7.5% | 100.0% |
| 2 | 2.8% | 5.2% | 20.7% | 16.3% | 41.2% | 13.9% | 100.0% |
| 3 | 2.3% | 6.8% | 18.3% | 21.7% | 29.9% | 21.0% | 100.0% |
| 4 | 3.1% | 6.5% | 16.9% | 21.2% | 32.4% | 19.9% | 100.0% |
| 5 | 4.2% | 7.0% | 14.1% | 20.3% | 37.2% | 17.3% | 100.0% |
| 6 | 4.7% | 7.1% | 12.3% | 24.9% | 31.6% | 19.4% | 100.0% |
| 7 | 5.8% | 7.1% | 13.0% | 22.0% | 29.9% | 22.2% | 100.0% |
| 8 | 6.8% | 7.8% | 12.2% | 21.5% | 29.6% | 22.3% | 100.0% |
| 9 | 13.2% | 10.1% | 10.2% | 19.4% | 31.5% | 15.7% | 100.0% |
| 10 | 10.1% | 12.2% | 15.2% | 22.0% | 30.4% | 10.1% | 100.0% |
| 11 | 8.4% | 11.3% | 17.6% | 24.6% | 26.6% | 11.5% | 100.0% |
| 12 | 7.8% | 11.0% | 20.2% | 26.9% | 26.2% | 7.8% | 100.0% |

## 4.3.6 Literacy Composite

### 4.3.6.1 By Cluster by Tier

**Table 4.3.6.1A**

Proficiency Level by Cluster by Tier (Count): Literacy S302

| Cluster | Tier | Literacy Proficiency Range | | | | | | Total |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 35,756 | 50,231 | 37,594 | 25,398 | 41,480 | 13,378 | 203,837 |
| K (accountability) | - | 130,672 | 24,411 | 24,653 | 16,739 | 7,362 | 0 | 203,837 |
| 1-2 | A | 14,422 | 40,333 | 30,528 | 1 | 0 | 0 | 85,284 |
| | B | 3,770 | 41,938 | 141,538 | 6,522 | 1 | 0 | 193,769 |
| | C | 1,080 | 12,036 | 41,307 | 33,761 | 18,698 | 2,493 | 109,375 |
| 3-5 | A | 5,056 | 10,071 | 9,198 | 4,723 | 52 | 0 | 29,100 |
| | B | 1,460 | 10,991 | 46,374 | 82,937 | 6,863 | 1 | 148,626 |
| | C | 301 | 911 | 16,158 | 57,529 | 74,659 | 24,333 | 173,891 |
| 6-8 | A | 6,265 | 9,640 | 6,796 | 978 | 9 | 0 | 23,688 |
| | B | 2,795 | 17,396 | 42,224 | 19,417 | 352 | 1 | 82,185 |
| | C | 885 | 10,954 | 59,272 | 31,887 | 8,517 | 1,640 | 113,155 |
| 9-12 | A | 6,172 | 12,513 | 8,620 | 1,686 | 26 | 0 | 29,017 |
| | B | 4,007 | 13,933 | 27,815 | 23,782 | 6,253 | 27 | 75,817 |
| | C | 1,166 | 3,011 | 15,659 | 28,418 | 30,496 | 13,732 | 92,482 |

**Table 4.3.6.1B**

Proficiency Level by Cluster by Tier (Percent): Literacy S302

| Cluster | Tier | Literacy Proficiency Range | | | | | | Total |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 17.5% | 24.6% | 18.4% | 12.5% | 20.3% | 6.6% | 100.0% |
| K (accountability) | - | 64.1% | 12.0% | 12.1% | 8.2% | 3.6% | 0.0% | 100.0% |
| 1-2 | A | 16.9% | 47.3% | 35.8% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 1.9% | 21.6% | 73.0% | 3.4% | 0.0% | 0.0% | 100.0% |
| | C | 1.0% | 11.0% | 37.8% | 30.9% | 17.1% | 2.3% | 100.0% |
| 3-5 | A | 17.4% | 34.6% | 31.6% | 16.2% | 0.2% | 0.0% | 100.0% |
| | B | 1.0% | 7.4% | 31.2% | 55.8% | 4.6% | 0.0% | 100.0% |
| | C | 0.2% | 0.5% | 9.3% | 33.1% | 42.9% | 14.0% | 100.0% |
| 6-8 | A | 26.4% | 40.7% | 28.7% | 4.1% | 0.0% | 0.0% | 100.0% |
| | B | 3.4% | 21.2% | 51.4% | 23.6% | 0.4% | 0.0% | 100.0% |
| | C | 0.8% | 9.7% | 52.4% | 28.2% | 7.5% | 1.4% | 100.0% |
| 9-12 | A | 21.3% | 43.1% | 29.7% | 5.8% | 0.1% | 0.0% | 100.0% |
| | B | 5.3% | 18.4% | 36.7% | 31.4% | 8.2% | 0.0% | 100.0% |
| | C | 1.3% | 3.3% | 16.9% | 30.7% | 33.0% | 14.8% | 100.0% |

### 4.3.6.2 By Grade by Tier

**Table 4.3.6.2A**

Proficiency Level by Grade by Tier (Count): Literacy S302

| Grade | Tier | Literacy Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 35,756 | 50,231 | 37,594 | 25,398 | 41,480 | 13,378 | 203,837 |
| K (accountability) | - | 130,672 | 24,411 | 24,653 | 16,739 | 7,362 | 0 | 203,837 |
| 1 | A | 9,558 | 32,337 | 25,543 | 1 | 0 | 0 | 67,439 |
| | B | 1,764 | 21,038 | 64,581 | 3,631 | 1 | 0 | 91,015 |
| | C | 570 | 5,997 | 17,610 | 11,653 | 5,845 | 930 | 42,605 |
| 2 | A | 4,864 | 7,996 | 4,985 | 0 | 0 | 0 | 17,845 |
| | B | 2,006 | 20,900 | 76,957 | 2,891 | 0 | 0 | 102,754 |
| | C | 510 | 6,039 | 23,697 | 22,108 | 12,853 | 1,563 | 66,770 |
| 3 | A | 1,352 | 4,327 | 4,216 | 2,779 | 45 | 0 | 12,719 |
| | B | 445 | 3,928 | 19,574 | 45,675 | 5,208 | 0 | 74,830 |
| | C | 111 | 183 | 3,185 | 18,513 | 40,755 | 13,620 | 76,367 |
| 4 | A | 1,590 | 3,150 | 2,754 | 1,259 | 5 | 0 | 8,758 |
| | B | 429 | 3,547 | 13,706 | 23,439 | 1,287 | 1 | 42,409 |
| | C | 118 | 295 | 5,157 | 21,574 | 20,092 | 6,537 | 53,773 |
| 5 | A | 2,114 | 2,594 | 2,228 | 685 | 2 | 0 | 7,623 |
| | B | 586 | 3,516 | 13,094 | 13,823 | 368 | 0 | 31,387 |
| | C | 72 | 433 | 7,816 | 17,442 | 13,812 | 4,176 | 43,751 |
| 6 | A | 1,239 | 3,324 | 2,756 | 581 | 6 | 0 | 7,906 |
| | B | 599 | 5,066 | 13,762 | 8,308 | 211 | 0 | 27,946 |
| | C | 212 | 2,955 | 19,653 | 11,707 | 2,777 | 566 | 37,870 |
| 7 | A | 2,165 | 3,124 | 2,265 | 298 | 2 | 0 | 7,854 |
| | B | 955 | 5,761 | 14,383 | 6,888 | 108 | 1 | 28,096 |
| | C | 312 | 3,691 | 20,084 | 11,176 | 2,967 | 564 | 38,794 |
| 8 | A | 2,861 | 3,192 | 1,775 | 99 | 1 | 0 | 7,928 |
| | B | 1,241 | 6,569 | 14,079 | 4,221 | 33 | 0 | 26,143 |
| | C | 361 | 4,308 | 19,535 | 9,004 | 2,773 | 510 | 36,491 |
| 9 | A | 3,462 | 6,715 | 4,366 | 1,132 | 20 | 0 | 15,695 |
| | B | 991 | 3,594 | 10,104 | 9,685 | 4,060 | 13 | 28,447 |
| | C | 269 | 583 | 3,836 | 9,101 | 14,167 | 7,251 | 35,207 |
| 10 | A | 1,333 | 3,127 | 2,246 | 411 | 6 | 0 | 7,123 |
| | B | 1,102 | 3,980 | 7,428 | 6,017 | 1,307 | 9 | 19,843 |
| | C | 247 | 757 | 4,177 | 7,534 | 7,527 | 2,790 | 23,032 |
| 11 | A | 770 | 1,746 | 1,367 | 111 | 0 | 0 | 3,994 |
| | B | 949 | 3,371 | 5,639 | 4,816 | 705 | 4 | 15,484 |
| | C | 275 | 722 | 3,745 | 6,409 | 5,192 | 2,271 | 18,614 |
| 12 | A | 607 | 925 | 641 | 32 | 0 | 0 | 2,205 |
| | B | 965 | 2,988 | 4,644 | 3,264 | 181 | 1 | 12,043 |
| | C | 375 | 949 | 3,901 | 5,374 | 3,610 | 1,420 | 15,629 |

**Table 4.3.6.2B**
Proficiency Level by Grade by Tier (Percent): Literacy S302

| Grade | Tier | Literacy Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 17.5% | 24.6% | 18.4% | 12.5% | 20.3% | 6.6% | 100.0% |
| K (accountability) | - | 64.1% | 12.0% | 12.1% | 8.2% | 3.6% | 0.0% | 100.0% |
| 1 | A | 14.2% | 47.9% | 37.9% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 1.9% | 23.1% | 71.0% | 4.0% | 0.0% | 0.0% | 100.0% |
| | C | 1.3% | 14.1% | 41.3% | 27.4% | 13.7% | 2.2% | 100.0% |
| 2 | A | 27.3% | 44.8% | 27.9% | 0.0% | 0.0% | 0.0% | 100.0% |
| | B | 2.0% | 20.3% | 74.9% | 2.8% | 0.0% | 0.0% | 100.0% |
| | C | 0.8% | 9.0% | 35.5% | 33.1% | 19.2% | 2.3% | 100.0% |
| 3 | A | 10.6% | 34.0% | 33.1% | 21.8% | 0.4% | 0.0% | 100.0% |
| | B | 0.6% | 5.2% | 26.2% | 61.0% | 7.0% | 0.0% | 100.0% |
| | C | 0.1% | 0.2% | 4.2% | 24.2% | 53.4% | 17.8% | 100.0% |
| 4 | A | 18.2% | 36.0% | 31.4% | 14.4% | 0.1% | 0.0% | 100.0% |
| | B | 1.0% | 8.4% | 32.3% | 55.3% | 3.0% | 0.0% | 100.0% |
| | C | 0.2% | 0.5% | 9.6% | 40.1% | 37.4% | 12.2% | 100.0% |
| 5 | A | 27.7% | 34.0% | 29.2% | 9.0% | 0.0% | 0.0% | 100.0% |
| | B | 1.9% | 11.2% | 41.7% | 44.0% | 1.2% | 0.0% | 100.0% |
| | C | 0.2% | 1.0% | 17.9% | 39.9% | 31.6% | 9.5% | 100.0% |
| 6 | A | 15.7% | 42.0% | 34.9% | 7.3% | 0.1% | 0.0% | 100.0% |
| | B | 2.1% | 18.1% | 49.2% | 29.7% | 0.8% | 0.0% | 100.0% |
| | C | 0.6% | 7.8% | 51.9% | 30.9% | 7.3% | 1.5% | 100.0% |
| 7 | A | 27.6% | 39.8% | 28.8% | 3.8% | 0.0% | 0.0% | 100.0% |
| | B | 3.4% | 20.5% | 51.2% | 24.5% | 0.4% | 0.0% | 100.0% |
| | C | 0.8% | 9.5% | 51.8% | 28.8% | 7.6% | 1.5% | 100.0% |
| 8 | A | 36.1% | 40.3% | 22.4% | 1.2% | 0.0% | 0.0% | 100.0% |
| | B | 4.7% | 25.1% | 53.9% | 16.1% | 0.1% | 0.0% | 100.0% |
| | C | 1.0% | 11.8% | 53.5% | 24.7% | 7.6% | 1.4% | 100.0% |
| 9 | A | 22.1% | 42.8% | 27.8% | 7.2% | 0.1% | 0.0% | 100.0% |
| | B | 3.5% | 12.6% | 35.5% | 34.0% | 14.3% | 0.0% | 100.0% |
| | C | 0.8% | 1.7% | 10.9% | 25.8% | 40.2% | 20.6% | 100.0% |
| 10 | A | 18.7% | 43.9% | 31.5% | 5.8% | 0.1% | 0.0% | 100.0% |
| | B | 5.6% | 20.1% | 37.4% | 30.3% | 6.6% | 0.0% | 100.0% |
| | C | 1.1% | 3.3% | 18.1% | 32.7% | 32.7% | 12.1% | 100.0% |
| 11 | A | 19.3% | 43.7% | 34.2% | 2.8% | 0.0% | 0.0% | 100.0% |
| | B | 6.1% | 21.8% | 36.4% | 31.1% | 4.6% | 0.0% | 100.0% |
| | C | 1.5% | 3.9% | 20.1% | 34.4% | 27.9% | 12.2% | 100.0% |
| 12 | A | 27.5% | 42.0% | 29.1% | 1.5% | 0.0% | 0.0% | 100.0% |
| | B | 8.0% | 24.8% | 38.6% | 27.1% | 1.5% | 0.0% | 100.0% |
| | C | 2.4% | 6.1% | 25.0% | 34.4% | 23.1% | 9.1% | 100.0% |

### 4.3.6.3    By Grade

**Table 4.3.6.3A**

Proficiency Level by Grade (Count): Literacy S302

| | Literacy Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 35,756 | 50,231 | 37,594 | 25,398 | 41,480 | 13,378 | 203,837 |
| K (accountability) | 130,672 | 24,411 | 24,653 | 16,739 | 7,362 | 0 | 203,837 |
| 1 | 11,892 | 59,372 | 107,734 | 15,285 | 5,846 | 930 | 201,059 |
| 2 | 7,380 | 34,935 | 105,639 | 24,999 | 12,853 | 1,563 | 187,369 |
| 3 | 1,908 | 8,438 | 26,975 | 66,967 | 46,008 | 13,620 | 163,916 |
| 4 | 2,137 | 6,992 | 21,617 | 46,272 | 21,384 | 6,538 | 104,940 |
| 5 | 2,772 | 6,543 | 23,138 | 31,950 | 14,182 | 4,176 | 82,761 |
| 6 | 2,050 | 11,345 | 36,171 | 20,596 | 2,994 | 566 | 73,722 |
| 7 | 3,432 | 12,576 | 36,732 | 18,362 | 3,077 | 565 | 74,744 |
| 8 | 4,463 | 14,069 | 35,389 | 13,324 | 2,807 | 510 | 70,562 |
| 9 | 4,722 | 10,892 | 18,306 | 19,918 | 18,247 | 7,264 | 79,349 |
| 10 | 2,682 | 7,864 | 13,851 | 13,962 | 8,840 | 2,799 | 49,998 |
| 11 | 1,994 | 5,839 | 10,751 | 11,336 | 5,897 | 2,275 | 38,092 |
| 12 | 1,947 | 4,862 | 9,186 | 8,670 | 3,791 | 1,421 | 29,877 |

**Table 4.3.6.3B**

Proficiency Level by Grade (Percent): Literacy S302

| | Literacy Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 17.5% | 24.6% | 18.4% | 12.5% | 20.3% | 6.6% | 100.0% |
| K (accountability) | 64.1% | 12.0% | 12.1% | 8.2% | 3.6% | 0.0% | 100.0% |
| 1 | 5.9% | 29.5% | 53.6% | 7.6% | 2.9% | 0.5% | 100.0% |
| 2 | 3.9% | 18.6% | 56.4% | 13.3% | 6.9% | 0.8% | 100.0% |
| 3 | 1.2% | 5.1% | 16.5% | 40.9% | 28.1% | 8.3% | 100.0% |
| 4 | 2.0% | 6.7% | 20.6% | 44.1% | 20.4% | 6.2% | 100.0% |
| 5 | 3.3% | 7.9% | 28.0% | 38.6% | 17.1% | 5.0% | 100.0% |
| 6 | 2.8% | 15.4% | 49.1% | 27.9% | 4.1% | 0.8% | 100.0% |
| 7 | 4.6% | 16.8% | 49.1% | 24.6% | 4.1% | 0.8% | 100.0% |
| 8 | 6.3% | 19.9% | 50.2% | 18.9% | 4.0% | 0.7% | 100.0% |
| 9 | 6.0% | 13.7% | 23.1% | 25.1% | 23.0% | 9.2% | 100.0% |
| 10 | 5.4% | 15.7% | 27.7% | 27.9% | 17.7% | 5.6% | 100.0% |
| 11 | 5.2% | 15.3% | 28.2% | 29.8% | 15.5% | 6.0% | 100.0% |
| 12 | 6.5% | 16.3% | 30.7% | 29.0% | 12.7% | 4.8% | 100.0% |

## 4.3.7 Comprehension Composite

### 4.3.7.1 By Cluster by Tier

**Table 4.3.7.1A**

Proficiency Level by Cluster by Tier (Count): Comprehension S302

| Cluster | Tier | Comprehension Proficiency Range | | | | | | Total |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 25,081 | 30,698 | 43,899 | 29,078 | 30,330 | 44,751 | 203,837 |
| K (accountability) | - | 120,340 | 14,043 | 14,430 | 13,518 | 25,262 | 16,244 | 203,837 |
| 1-2 | A | 10,391 | 23,289 | 30,315 | 21,285 | n/a | n/a | 85,280 |
| | B | 1,393 | 3,083 | 32,275 | 43,076 | 114,006 | n/a | 193,833 |
| | C | 499 | 2,560 | 14,327 | 15,359 | 35,404 | 41,295 | 109,444 |
| 3-5 | A | 4,744 | 11,712 | 7,478 | 5,207 | n/a | n/a | 29,141 |
| | B | 890 | 9,118 | 37,231 | 33,343 | 68,114 | n/a | 148,696 |
| | C | 257 | 1,113 | 15,253 | 18,246 | 62,063 | 77,063 | 173,995 |
| 6-8 | A | 7,642 | 10,240 | 4,431 | 1,391 | n/a | n/a | 23,704 |
| | B | 1,217 | 16,036 | 31,877 | 18,351 | 14,748 | n/a | 82,229 |
| | C | 364 | 4,247 | 27,849 | 20,784 | 36,308 | 23,678 | 113,230 |
| 9-12 | A | 12,065 | 11,721 | 4,111 | 1,136 | n/a | n/a | 29,033 |
| | B | 4,461 | 21,858 | 22,397 | 15,256 | 11,925 | n/a | 75,897 |
| | C | 1,102 | 6,373 | 16,936 | 19,683 | 25,277 | 23,262 | 92,633 |

**Table 4.3.7.1B**

Proficiency Level by Cluster by Tier (Percent): Comprehension S302

| Cluster | Tier | Comprehension Proficiency Range | | | | | | Total |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| K (instructional) | - | 12.3% | 15.1% | 21.5% | 14.3% | 14.9% | 22.0% | 100.0% |
| K (accountability) | - | 59.0% | 6.9% | 7.1% | 6.6% | 12.4% | 8.0% | 100.0% |
| 1-2 | A | 12.2% | 27.3% | 35.5% | 25.0% | n/a | n/a | 100.0% |
| | B | 0.7% | 1.6% | 16.7% | 22.2% | 58.8% | n/a | 100.0% |
| | C | 0.5% | 2.3% | 13.1% | 14.0% | 32.3% | 37.7% | 100.0% |
| 3-5 | A | 16.3% | 40.2% | 25.7% | 17.9% | n/a | n/a | 100.0% |
| | B | 0.6% | 6.1% | 25.0% | 22.4% | 45.8% | n/a | 100.0% |
| | C | 0.1% | 0.6% | 8.8% | 10.5% | 35.7% | 44.3% | 100.0% |
| 6-8 | A | 32.2% | 43.2% | 18.7% | 5.9% | n/a | n/a | 100.0% |
| | B | 1.5% | 19.5% | 38.8% | 22.3% | 17.9% | n/a | 100.0% |
| | C | 0.3% | 3.8% | 24.6% | 18.4% | 32.1% | 20.9% | 100.0% |
| 9-12 | A | 41.6% | 40.4% | 14.2% | 3.9% | n/a | n/a | 100.0% |
| | B | 5.9% | 28.8% | 29.5% | 20.1% | 15.7% | n/a | 100.0% |
| | C | 1.2% | 6.9% | 18.3% | 21.2% | 27.3% | 25.1% | 100.0% |

### 4.3.7.2    By Grade by Tier

**Table 4.3.7.2A**

Proficiency Level by Grade by Tier (Count): Comprehension S302

| Grade | Tier | Comprehension Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** | **6** | |
| K (instructional) | - | 25,081 | 30,698 | 43,899 | 29,078 | 30,330 | 44,751 | 203,837 |
| K (accountability) | - | 120,340 | 14,043 | 14,430 | 13,518 | 25,262 | 16,244 | 203,837 |
| 1 | A | 6,765 | 18,250 | 25,773 | 16648 | n/a | n/a | 67,436 |
| | B | 1,002 | 1,274 | 13,373 | 25,386 | 50,017 | n/a | 91,052 |
| | C | 310 | 880 | 6,125 | 6,973 | 14,848 | 13,495 | 42,631 |
| 2 | A | 3,626 | 5,039 | 4,542 | 4637 | n/a | n/a | 17,844 |
| | B | 391 | 1,809 | 18,902 | 17,690 | 63,989 | n/a | 102,781 |
| | C | 189 | 1,680 | 8,202 | 8,386 | 20,556 | 27,800 | 66,813 |
| 3 | A | 1,125 | 5,270 | 3,768 | 2,587 | n/a | n/a | 12,750 |
| | B | 273 | 2,679 | 15,218 | 16,825 | 39,874 | n/a | 74,869 |
| | C | 109 | 165 | 2,277 | 5,312 | 28,119 | 40,429 | 76,411 |
| 4 | A | 1,491 | 3,687 | 2,036 | 1,547 | n/a | n/a | 8,761 |
| | B | 283 | 3,034 | 12,542 | 9,945 | 16,628 | n/a | 42,432 |
| | C | 94 | 314 | 5,814 | 7,155 | 19,057 | 21,371 | 53,805 |
| 5 | A | 2,128 | 2,755 | 1,674 | 1,073 | n/a | n/a | 7,630 |
| | B | 334 | 3,405 | 9,471 | 6,573 | 11,612 | n/a | 31,395 |
| | C | 54 | 634 | 7,162 | 5,779 | 14,887 | 15,263 | 43,779 |
| 6 | A | 1,656 | 3,733 | 1,953 | 579 | n/a | n/a | 7,921 |
| | B | 202 | 4,111 | 11,540 | 6,576 | 5,533 | n/a | 27,962 |
| | C | 99 | 910 | 9,878 | 7,608 | 12,464 | 6,935 | 37,894 |
| 7 | A | 2,705 | 3,309 | 1,412 | 426 | n/a | n/a | 7,852 |
| | B | 395 | 5,590 | 11,337 | 6,364 | 4,421 | n/a | 28,107 |
| | C | 144 | 1,509 | 10,016 | 6,900 | 12,015 | 8,230 | 38,814 |
| 8 | A | 3,281 | 3,198 | 1,066 | 386 | n/a | n/a | 7,931 |
| | B | 620 | 6,335 | 9,000 | 5,411 | 4,794 | n/a | 26,160 |
| | C | 121 | 1,828 | 7,955 | 6,276 | 11,829 | 8,513 | 36,522 |
| 9 | A | 6,697 | 6,294 | 2,089 | 639 | n/a | n/a | 15,719 |
| | B | 880 | 6,471 | 10,543 | 5,506 | 5,069 | n/a | 28,469 |
| | C | 273 | 1,151 | 5,834 | 6,876 | 10,990 | 10,137 | 35,261 |
| 10 | A | 2,650 | 3,095 | 1,151 | 236 | n/a | n/a | 7,132 |
| | B | 987 | 6,285 | 6,000 | 3,727 | 2,864 | n/a | 19,863 |
| | C | 233 | 1,662 | 4,964 | 4,976 | 6,228 | 4,999 | 23,062 |
| 11 | A | 1,609 | 1,647 | 572 | 151 | n/a | n/a | 3,979 |
| | B | 1,193 | 5,298 | 3,187 | 3,552 | 2,278 | n/a | 15,508 |
| | C | 257 | 1,818 | 3,188 | 4,421 | 4,363 | 4,602 | 18,649 |
| 12 | A | 1,109 | 685 | 299 | 110 | n/a | n/a | 2,203 |
| | B | 1,401 | 3,804 | 2,667 | 2,471 | 1,714 | n/a | 12,057 |
| | C | 339 | 1,742 | 2,950 | 3,410 | 3,696 | 3,524 | 15,661 |

Return to Summary of Assessment Records Claims, Actions, and Evidence

**Table 4.3.7.2B**
Proficiency Level by Grade by Tier (Percent): Comprehension S302

| Grade | Tier | Comprehension Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 12.3% | 15.1% | 21.5% | 14.3% | 14.9% | 22.0% | 100.0% |
| K (accountability) | - | 59.0% | 6.9% | 7.1% | 6.6% | 12.4% | 8.0% | 100.0% |
| 1 | A | 10.0% | 27.1% | 38.2% | 24.7% | n/a | n/a | 100.0% |
| | B | 1.1% | 1.4% | 14.7% | 27.9% | 54.9% | n/a | 100.0% |
| | C | 0.7% | 2.1% | 14.4% | 16.4% | 34.8% | 31.7% | 100.0% |
| 2 | A | 20.3% | 28.2% | 25.5% | 26.0% | n/a | n/a | 100.0% |
| | B | 0.4% | 1.8% | 18.4% | 17.2% | 62.3% | n/a | 100.0% |
| | C | 0.3% | 2.5% | 12.3% | 12.6% | 30.8% | 41.6% | 100.0% |
| 3 | A | 8.8% | 41.3% | 29.6% | 20.3% | n/a | n/a | 100.0% |
| | B | 0.4% | 3.6% | 20.3% | 22.5% | 53.3% | n/a | 100.0% |
| | C | 0.1% | 0.2% | 3.0% | 7.0% | 36.8% | 52.9% | 100.0% |
| 4 | A | 17.0% | 42.1% | 23.2% | 17.7% | n/a | n/a | 100.0% |
| | B | 0.7% | 7.2% | 29.6% | 23.4% | 39.2% | n/a | 100.0% |
| | C | 0.2% | 0.6% | 10.8% | 13.3% | 35.4% | 39.7% | 100.0% |
| 5 | A | 27.9% | 36.1% | 21.9% | 14.1% | n/a | n/a | 100.0% |
| | B | 1.1% | 10.8% | 30.2% | 20.9% | 37.0% | n/a | 100.0% |
| | C | 0.1% | 1.4% | 16.4% | 13.2% | 34.0% | 34.9% | 100.0% |
| 6 | A | 20.9% | 47.1% | 24.7% | 7.3% | n/a | n/a | 100.0% |
| | B | 0.7% | 14.7% | 41.3% | 23.5% | 19.8% | n/a | 100.0% |
| | C | 0.3% | 2.4% | 26.1% | 20.1% | 32.9% | 18.3% | 100.0% |
| 7 | A | 34.4% | 42.1% | 18.0% | 5.4% | n/a | n/a | 100.0% |
| | B | 1.4% | 19.9% | 40.3% | 22.6% | 15.7% | n/a | 100.0% |
| | C | 0.4% | 3.9% | 25.8% | 17.8% | 31.0% | 21.2% | 100.0% |
| 8 | A | 41.4% | 40.3% | 13.4% | 4.9% | n/a | n/a | 100.0% |
| | B | 2.4% | 24.2% | 34.4% | 20.7% | 18.3% | n/a | 100.0% |
| | C | 0.3% | 5.0% | 21.8% | 17.2% | 32.4% | 23.3% | 100.0% |
| 9 | A | 42.6% | 40.0% | 13.3% | 4.1% | n/a | n/a | 100.0% |
| | B | 3.1% | 22.7% | 37.0% | 19.3% | 17.8% | n/a | 100.0% |
| | C | 0.8% | 3.3% | 16.5% | 19.5% | 31.2% | 28.7% | 100.0% |
| 10 | A | 37.2% | 43.4% | 16.1% | 3.3% | n/a | n/a | 100.0% |
| | B | 5.0% | 31.6% | 30.2% | 18.8% | 14.4% | n/a | 100.0% |
| | C | 1.0% | 7.2% | 21.5% | 21.6% | 27.0% | 21.7% | 100.0% |
| 11 | A | 40.4% | 41.4% | 14.4% | 3.8% | n/a | n/a | 100.0% |
| | B | 7.7% | 34.2% | 20.6% | 22.9% | 14.7% | n/a | 100.0% |
| | C | 1.4% | 9.7% | 17.1% | 23.7% | 23.4% | 24.7% | 100.0% |
| 12 | A | 50.3% | 31.1% | 13.6% | 5.0% | n/a | n/a | 100.0% |
| | B | 11.6% | 31.6% | 22.1% | 20.5% | 14.2% | n/a | 100.0% |
| | C | 2.2% | 11.1% | 18.8% | 21.8% | 23.6% | 22.5% | 100.0% |

### 4.3.7.3  By Grade

**Table 4.3.7.3A**

Proficiency Level by Grade (Count): Comprehension S302

| | Comprehension Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 25,081 | 30,698 | 43,899 | 29,078 | 30,330 | 44,751 | 203,837 |
| K (accountability) | 120,340 | 14,043 | 14,430 | 13,518 | 25,262 | 16,244 | 203,837 |
| 1 | 8,077 | 20,404 | 45,271 | 49,007 | 64,865 | 13,495 | 201,119 |
| 2 | 4,206 | 8,528 | 31,646 | 30,713 | 84,545 | 27,800 | 187,438 |
| 3 | 1,507 | 8,114 | 21,263 | 24,724 | 67,993 | 40,429 | 164,030 |
| 4 | 1,868 | 7,035 | 20,392 | 18,647 | 35,685 | 21,371 | 104,998 |
| 5 | 2,516 | 6,794 | 18,307 | 13,425 | 26,499 | 15,263 | 82,804 |
| 6 | 1,957 | 8,754 | 23,371 | 14,763 | 17,997 | 6,935 | 73,777 |
| 7 | 3,244 | 10,408 | 22,765 | 13,690 | 16,436 | 8,230 | 74,773 |
| 8 | 4,022 | 11,361 | 18,021 | 12,073 | 16,623 | 8,513 | 70,613 |
| 9 | 7,850 | 13,916 | 18,466 | 13,021 | 16,059 | 10,137 | 79,449 |
| 10 | 3,870 | 11,042 | 12,115 | 8,939 | 9,092 | 4,999 | 50,057 |
| 11 | 3,059 | 8,763 | 6,947 | 8,124 | 6,641 | 4,602 | 38,136 |
| 12 | 2,849 | 6,231 | 5,916 | 5,991 | 5,410 | 3,524 | 29,921 |

**Table 4.3.7.3B**

Proficiency Level by Grade (Percent): Comprehension S302

| | Comprehension Proficiency Range | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| K (instructional) | 12.3% | 15.1% | 21.5% | 14.3% | 14.9% | 22.0% | 100.0% |
| K (accountability) | 59.0% | 6.9% | 7.1% | 6.6% | 12.4% | 8.0% | 100.0% |
| 1 | 4.0% | 10.1% | 22.5% | 24.4% | 32.3% | 6.7% | 100.0% |
| 2 | 2.2% | 4.5% | 16.9% | 16.4% | 45.1% | 14.8% | 100.0% |
| 3 | 0.9% | 4.9% | 13.0% | 15.1% | 41.5% | 24.6% | 100.0% |
| 4 | 1.8% | 6.7% | 19.4% | 17.8% | 34.0% | 20.4% | 100.0% |
| 5 | 3.0% | 8.2% | 22.1% | 16.2% | 32.0% | 18.4% | 100.0% |
| 6 | 2.7% | 11.9% | 31.7% | 20.0% | 24.4% | 9.4% | 100.0% |
| 7 | 4.3% | 13.9% | 30.4% | 18.3% | 22.0% | 11.0% | 100.0% |
| 8 | 5.7% | 16.1% | 25.5% | 17.1% | 23.5% | 12.1% | 100.0% |
| 9 | 9.9% | 17.5% | 23.2% | 16.4% | 20.2% | 12.8% | 100.0% |
| 10 | 7.7% | 22.1% | 24.2% | 17.9% | 18.2% | 10.0% | 100.0% |
| 11 | 8.0% | 23.0% | 18.2% | 21.3% | 17.4% | 12.1% | 100.0% |
| 12 | 9.5% | 20.8% | 19.8% | 20.0% | 18.1% | 11.8% | 100.0% |

## 4.3.8 Overall Composite

### 4.3.8.1 By Cluster by Tier

**Table 4.3.8.1A**

Proficiency Level by Cluster by Tier (Count): Overall S302

| Cluster | Tier | Overall Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 28,669 | 41,422 | 40,481 | 36,103 | 43,506 | 13,628 | 203,809 |
| K (accountability) | - | 108,070 | 32,479 | 29,796 | 19,836 | 11,866 | 1,762 | 203,809 |
| 1-2 | A | 10,171 | 35,943 | 38,921 | 122 | 0 | 0 | 85,157 |
| | B | 1,603 | 20,388 | 122,352 | 49,220 | 17 | 0 | 193,580 |
| | C | 389 | 4,881 | 30,572 | 39,444 | 29,432 | 4,601 | 109,319 |
| 3-5 | A | 6,102 | 10,070 | 8,917 | 3,822 | 135 | 0 | 29,046 |
| | B | 859 | 8,816 | 44,265 | 79,344 | 15,204 | 0 | 148,488 |
| | C | 192 | 664 | 11,764 | 48,616 | 79,469 | 33,038 | 173,743 |
| 6-8 | A | 7,956 | 8,780 | 5,607 | 1,279 | 11 | 0 | 23,633 |
| | B | 1,313 | 12,243 | 35,405 | 30,846 | 2,170 | 0 | 81,977 |
| | C | 318 | 2,905 | 30,883 | 50,791 | 24,572 | 3,446 | 112,915 |
| 9-12 | A | 9,643 | 12,176 | 5,767 | 1,222 | 42 | 0 | 28,850 |
| | B | 2,628 | 12,216 | 25,408 | 24,990 | 10,081 | 14 | 75,337 |
| | C | 835 | 2,136 | 12,591 | 29,239 | 32,837 | 14,234 | 91,872 |

**Table 4.3.8.1B**

Proficiency Level by Cluster by Tier (Percent): Overall S302

| Cluster | Tier | Overall Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 14.1% | 20.3% | 19.9% | 17.7% | 21.3% | 6.7% | 100.0% |
| K (accountability) | - | 53.0% | 15.9% | 14.6% | 9.7% | 5.8% | 0.9% | 100.0% |
| 1-2 | A | 11.9% | 42.2% | 45.7% | 0.1% | 0.0% | 0.0% | 100.0% |
| | B | 0.8% | 10.5% | 63.2% | 25.4% | 0.0% | 0.0% | 100.0% |
| | C | 0.4% | 4.5% | 28.0% | 36.1% | 26.9% | 4.2% | 100.0% |
| 3-5 | A | 21.0% | 34.7% | 30.7% | 13.2% | 0.5% | 0.0% | 100.0% |
| | B | 0.6% | 5.9% | 29.8% | 53.4% | 10.2% | 0.0% | 100.0% |
| | C | 0.1% | 0.4% | 6.8% | 28.0% | 45.7% | 19.0% | 100.0% |
| 6-8 | A | 33.7% | 37.2% | 23.7% | 5.4% | 0.0% | 0.0% | 100.0% |
| | B | 1.6% | 14.9% | 43.2% | 37.6% | 2.6% | 0.0% | 100.0% |
| | C | 0.3% | 2.6% | 27.4% | 45.0% | 21.8% | 3.1% | 100.0% |
| 9-12 | A | 33.4% | 42.2% | 20.0% | 4.2% | 0.1% | 0.0% | 100.0% |
| | B | 3.5% | 16.2% | 33.7% | 33.2% | 13.4% | 0.0% | 100.0% |
| | C | 0.9% | 2.3% | 13.7% | 31.8% | 35.7% | 15.5% | 100.0% |

### 4.3.8.2 By Grade by Tier

**Table 4.3.8.2A**

Proficiency Level by Grade by Tier (Count): Overall S302

| Grade | Tier | Overall Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 28,669 | 41,422 | 40,481 | 36,103 | 43,506 | 13,628 | 203,809 |
| K (accountability) | - | 108,070 | 32,479 | 29,796 | 19,836 | 11,866 | 1,762 | 203,809 |
| 1 | A | 6,224 | 28,518 | 32,494 | 106 | 0 | 0 | 67,342 |
| | B | 957 | 10,892 | 61,866 | 17,210 | 6 | 0 | 90,931 |
| | C | 187 | 2,615 | 14,904 | 13,940 | 9,225 | 1,709 | 42,580 |
| 2 | A | 3,947 | 7,425 | 6,427 | 16 | 0 | 0 | 17,815 |
| | B | 646 | 9,496 | 60,486 | 32,010 | 11 | 0 | 102,649 |
| | C | 202 | 2,266 | 15,668 | 25,504 | 20,207 | 2,892 | 66,739 |
| 3 | A | 1,814 | 4,455 | 4,299 | 2,021 | 110 | 0 | 12,699 |
| | B | 259 | 3,161 | 20,013 | 40,919 | 10,425 | 0 | 74,777 |
| | C | 79 | 167 | 2,644 | 15,709 | 38,945 | 18,769 | 76,313 |
| 4 | A | 1,999 | 3,059 | 2,581 | 1,079 | 15 | 0 | 8,733 |
| | B | 292 | 2,904 | 13,315 | 22,528 | 3,321 | 0 | 42,360 |
| | C | 67 | 238 | 4,291 | 17,451 | 23,098 | 8,584 | 53,729 |
| 5 | A | 2,289 | 2,556 | 2,037 | 722 | 10 | 0 | 7,614 |
| | B | 308 | 2,751 | 10,937 | 15,897 | 1,458 | 0 | 31,351 |
| | C | 46 | 259 | 4,829 | 15,456 | 17,426 | 5,685 | 43,701 |
| 6 | A | 1,984 | 2,984 | 2,255 | 656 | 8 | 0 | 7,887 |
| | B | 255 | 3,484 | 11,343 | 11,733 | 1,085 | 0 | 27,900 |
| | C | 75 | 785 | 9,955 | 17,919 | 7,823 | 1,238 | 37,795 |
| 7 | A | 2,715 | 2,898 | 1,847 | 376 | 2 | 0 | 7,838 |
| | B | 439 | 4,091 | 12,467 | 10,185 | 828 | 0 | 28,010 |
| | C | 112 | 979 | 10,971 | 16,486 | 8,998 | 1,168 | 38,714 |
| 8 | A | 3,257 | 2,898 | 1,505 | 247 | 1 | 0 | 7,908 |
| | B | 619 | 4,668 | 11,595 | 8,928 | 257 | 0 | 26,067 |
| | C | 131 | 1,141 | 9,957 | 16,386 | 7,751 | 1,040 | 36,406 |
| 9 | A | 5,577 | 6,337 | 2,951 | 711 | 37 | 0 | 15,613 |
| | B | 646 | 3,120 | 8,603 | 10,139 | 5,752 | 8 | 28,268 |
| | C | 209 | 416 | 2,729 | 8,741 | 15,112 | 7,812 | 35,019 |
| 10 | A | 2,053 | 3,240 | 1,486 | 313 | 5 | 0 | 7,097 |
| | B | 691 | 3,564 | 6,811 | 6,243 | 2,429 | 4 | 19,742 |
| | C | 176 | 492 | 3,289 | 7,835 | 8,203 | 2,885 | 22,880 |
| 11 | A | 1,205 | 1,749 | 874 | 137 | 0 | 0 | 3,965 |
| | B | 621 | 3,052 | 5,358 | 4,942 | 1,428 | 2 | 15,403 |
| | C | 200 | 528 | 3,156 | 6,684 | 5,697 | 2,224 | 18,489 |
| 12 | A | 808 | 850 | 456 | 61 | 0 | 0 | 2,175 |
| | B | 670 | 2,480 | 4,636 | 3,666 | 472 | 0 | 11,924 |
| | C | 250 | 700 | 3,417 | 5,979 | 3,825 | 1,313 | 15,484 |

Return to Summary of Assessment Records Claims, Actions, and Evidence

**Table 4.3.8.2B**
Proficiency Level by Grade by Tier (Percent): Overall S302

| Grade | Tier | Overall Proficiency Range | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| K (instructional) | - | 14.1% | 20.3% | 19.9% | 17.7% | 21.3% | 6.7% | 100.0% |
| K (accountability) | - | 53.0% | 15.9% | 14.6% | 9.7% | 5.8% | 0.9% | 100.0% |
| 1 | A | 9.2% | 42.3% | 48.3% | 0.2% | 0.0% | 0.0% | 100.0% |
| | B | 1.1% | 12.0% | 68.0% | 18.9% | 0.0% | 0.0% | 100.0% |
| | C | 0.4% | 6.1% | 35.0% | 32.7% | 21.7% | 4.0% | 100.0% |
| 2 | A | 22.2% | 41.7% | 36.1% | 0.1% | 0.0% | 0.0% | 100.0% |
| | B | 0.6% | 9.3% | 58.9% | 31.2% | 0.0% | 0.0% | 100.0% |
| | C | 0.3% | 3.4% | 23.5% | 38.2% | 30.3% | 4.3% | 100.0% |
| 3 | A | 14.3% | 35.1% | 33.9% | 15.9% | 0.9% | 0.0% | 100.0% |
| | B | 0.3% | 4.2% | 26.8% | 54.7% | 13.9% | 0.0% | 100.0% |
| | C | 0.1% | 0.2% | 3.5% | 20.6% | 51.0% | 24.6% | 100.0% |
| 4 | A | 22.9% | 35.0% | 29.6% | 12.4% | 0.2% | 0.0% | 100.0% |
| | B | 0.7% | 6.9% | 31.4% | 53.2% | 7.8% | 0.0% | 100.0% |
| | C | 0.1% | 0.4% | 8.0% | 32.5% | 43.0% | 16.0% | 100.0% |
| 5 | A | 30.1% | 33.6% | 26.8% | 9.5% | 0.1% | 0.0% | 100.0% |
| | B | 1.0% | 8.8% | 34.9% | 50.7% | 4.7% | 0.0% | 100.0% |
| | C | 0.1% | 0.6% | 11.1% | 35.4% | 39.9% | 13.0% | 100.0% |
| 6 | A | 25.2% | 37.8% | 28.6% | 8.3% | 0.1% | 0.0% | 100.0% |
| | B | 0.9% | 12.5% | 40.7% | 42.1% | 3.9% | 0.0% | 100.0% |
| | C | 0.2% | 2.1% | 26.3% | 47.4% | 20.7% | 3.3% | 100.0% |
| 7 | A | 34.6% | 37.0% | 23.6% | 4.8% | 0.0% | 0.0% | 100.0% |
| | B | 1.6% | 14.6% | 44.5% | 36.4% | 3.0% | 0.0% | 100.0% |
| | C | 0.3% | 2.5% | 28.3% | 42.6% | 23.2% | 3.0% | 100.0% |
| 8 | A | 41.2% | 36.6% | 19.0% | 3.1% | 0.0% | 0.0% | 100.0% |
| | B | 2.4% | 17.9% | 44.5% | 34.3% | 1.0% | 0.0% | 100.0% |
| | C | 0.4% | 3.1% | 27.3% | 45.0% | 21.3% | 2.9% | 100.0% |
| 9 | A | 35.7% | 40.6% | 18.9% | 4.6% | 0.2% | 0.0% | 100.0% |
| | B | 2.3% | 11.0% | 30.4% | 35.9% | 20.3% | 0.0% | 100.0% |
| | C | 0.6% | 1.2% | 7.8% | 25.0% | 43.2% | 22.3% | 100.0% |
| 10 | A | 28.9% | 45.7% | 20.9% | 4.4% | 0.1% | 0.0% | 100.0% |
| | B | 3.5% | 18.1% | 34.5% | 31.6% | 12.3% | 0.0% | 100.0% |
| | C | 0.8% | 2.2% | 14.4% | 34.2% | 35.9% | 12.6% | 100.0% |
| 11 | A | 30.4% | 44.1% | 22.0% | 3.5% | 0.0% | 0.0% | 100.0% |
| | B | 4.0% | 19.8% | 34.8% | 32.1% | 9.3% | 0.0% | 100.0% |
| | C | 1.1% | 2.9% | 17.1% | 36.2% | 30.8% | 12.0% | 100.0% |
| 12 | A | 37.1% | 39.1% | 21.0% | 2.8% | 0.0% | 0.0% | 100.0% |
| | B | 5.6% | 20.8% | 38.9% | 30.7% | 4.0% | 0.0% | 100.0% |
| | C | 1.6% | 4.5% | 22.1% | 38.6% | 24.7% | 8.5% | 100.0% |

### 4.3.8.3    By Grade

**Table 4.3.8.3A**
Proficiency Level by Grade (Count): Overall S302

| | Overall Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 28,669 | 41,422 | 40,481 | 36,103 | 43,506 | 13,628 | 203,809 |
| K (accountability) | 108,070 | 32,479 | 29,796 | 19,836 | 11,866 | 1,762 | 203,809 |
| 1 | 7,368 | 42,025 | 109,264 | 31,256 | 9,231 | 1,709 | 200,853 |
| 2 | 4,795 | 19,187 | 82,581 | 57,530 | 20,218 | 2,892 | 187,203 |
| 3 | 2,152 | 7,783 | 26,956 | 58,649 | 49,480 | 18,769 | 163,789 |
| 4 | 2,358 | 6,201 | 20,187 | 41,058 | 26,434 | 8,584 | 104,822 |
| 5 | 2,643 | 5,566 | 17,803 | 32,075 | 18,894 | 5,685 | 82,666 |
| 6 | 2,314 | 7,253 | 23,553 | 30,308 | 8,916 | 1,238 | 73,582 |
| 7 | 3,266 | 7,968 | 25,285 | 27,047 | 9,828 | 1,168 | 74,562 |
| 8 | 4,007 | 8,707 | 23,057 | 25,561 | 8,009 | 1,040 | 70,381 |
| 9 | 6,432 | 9,873 | 14,283 | 19,591 | 20,901 | 7,820 | 78,900 |
| 10 | 2,920 | 7,296 | 11,586 | 14,391 | 10,637 | 2,889 | 49,719 |
| 11 | 2,026 | 5,329 | 9,388 | 11,763 | 7,125 | 2,226 | 37,857 |
| 12 | 1,728 | 4,030 | 8,509 | 9,706 | 4,297 | 1,313 | 29,583 |

**Table 4.3.8.3B**
Proficiency Level by Grade (Percent): Overall S302

| | Overall Proficiency Range | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Total** |
| K (instructional) | 14.1% | 20.3% | 19.9% | 17.7% | 21.3% | 6.7% | 100.0% |
| K (accountability) | 53.0% | 15.9% | 14.6% | 9.7% | 5.8% | 0.9% | 100.0% |
| 1 | 3.7% | 20.9% | 54.4% | 15.6% | 4.6% | 0.9% | 100.0% |
| 2 | 2.6% | 10.2% | 44.1% | 30.7% | 10.8% | 1.5% | 100.0% |
| 3 | 1.3% | 4.8% | 16.5% | 35.8% | 30.2% | 11.5% | 100.0% |
| 4 | 2.2% | 5.9% | 19.3% | 39.2% | 25.2% | 8.2% | 100.0% |
| 5 | 3.2% | 6.7% | 21.5% | 38.8% | 22.9% | 6.9% | 100.0% |
| 6 | 3.1% | 9.9% | 32.0% | 41.2% | 12.1% | 1.7% | 100.0% |
| 7 | 4.4% | 10.7% | 33.9% | 36.3% | 13.2% | 1.6% | 100.0% |
| 8 | 5.7% | 12.4% | 32.8% | 36.3% | 11.4% | 1.5% | 100.0% |
| 9 | 8.2% | 12.5% | 18.1% | 24.8% | 26.5% | 9.9% | 100.0% |
| 10 | 5.9% | 14.7% | 23.3% | 28.9% | 21.4% | 5.8% | 100.0% |
| 11 | 5.4% | 14.1% | 24.8% | 31.1% | 18.8% | 5.9% | 100.0% |
| 12 | 5.8% | 13.6% | 28.8% | 32.8% | 14.5% | 4.4% | 100.0% |

## 4.4 Mean Raw Score Results by Standards

### 4.4.1 Comprehension Composite

#### 4.4.1.1 By Cluster

**Table 4.4.1.1**

Mean Raw Score by Cluster by Tier by Standard: Comprehension S302

| Cluster | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---------|------|----------|---------------|------------|--------------------|
| 1-2 | A | Social Instructional Language | 12 | 7.77 | 64.74% |
| | | Language of Language Arts | 9 | 5.97 | 66.34% |
| | | Language of Math | 9 | 5.27 | 58.54% |
| | | Language of Science | 6 | 4.16 | 69.33% |
| | | Language of Social Studies | 6 | 3.66 | 60.97% |
| | B | Social Instructional Language | 6 | 4.51 | 75.15% |
| | | Language of Language Arts | 12 | 7.36 | 61.33% |
| | | Language of Math | 12 | 7.49 | 62.43% |
| | | Language of Science | 9 | 6.65 | 73.89% |
| | | Language of Social Studies | 9 | 6.25 | 69.47% |
| | C | Social Instructional Language | 6 | 4.89 | 81.53% |
| | | Language of Language Arts | 12 | 8.86 | 73.83% |
| | | Language of Math | 12 | 6.95 | 57.94% |
| | | Language of Science | 9 | 5.71 | 63.46% |
| | | Language of Social Studies | 9 | 6.24 | 69.35% |
| 3-5 | A | Social Instructional Language | 12 | 7.10 | 59.15% |
| | | Language of Language Arts | 9 | 4.03 | 44.76% |
| | | Language of Math | 9 | 4.79 | 53.18% |
| | | Language of Science | 6 | 3.22 | 53.60% |
| | | Language of Social Studies | 6 | 3.26 | 54.33% |
| | B | Social Instructional Language | 6 | 4.81 | 80.09% |
| | | Language of Language Arts | 12 | 7.73 | 64.38% |
| | | Language of Math | 12 | 6.89 | 57.42% |
| | | Language of Science | 9 | 5.25 | 58.32% |
| | | Language of Social Studies | 9 | 5.42 | 60.22% |
| | C | Social Instructional Language | 6 | 3.73 | 62.17% |
| | | Language of Language Arts | 12 | 7.71 | 64.27% |
| | | Language of Math | 12 | 5.10 | 42.49% |
| | | Language of Science | 9 | 5.05 | 56.06% |
| | | Language of Social Studies | 9 | 4.33 | 48.12% |

| Cluster | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 6-8 | A | Social Instructional Language | 12 | 5.97 | 49.77% |
| | | Language of Language Arts | 9 | 4.59 | 50.95% |
| | | Language of Math | 9 | 4.67 | 51.89% |
| | | Language of Science | 6 | 3.56 | 59.32% |
| | | Language of Social Studies | 6 | 3.08 | 51.26% |
| | B | Social Instructional Language | 6 | 3.93 | 65.43% |
| | | Language of Language Arts | 12 | 8.50 | 70.81% |
| | | Language of Math | 12 | 7.15 | 59.62% |
| | | Language of Science | 9 | 4.78 | 53.10% |
| | | Language of Social Studies | 9 | 5.50 | 61.06% |
| | C | Social Instructional Language | 6 | 4.22 | 70.35% |
| | | Language of Language Arts | 12 | 8.01 | 66.75% |
| | | Language of Math | 12 | 7.34 | 61.14% |
| | | Language of Science | 9 | 5.55 | 61.65% |
| | | Language of Social Studies | 9 | 4.54 | 50.41% |
| 9-12 | A | Social Instructional Language | 12 | 5.84 | 48.67% |
| | | Language of Language Arts | 9 | 5.17 | 57.45% |
| | | Language of Math | 9 | 4.87 | 54.07% |
| | | Language of Science | 6 | 3.32 | 55.41% |
| | | Language of Social Studies | 6 | 3.67 | 61.21% |
| | B | Social Instructional Language | 6 | 4.33 | 72.21% |
| | | Language of Language Arts | 12 | 8.08 | 67.35% |
| | | Language of Math | 12 | 7.22 | 60.14% |
| | | Language of Science | 9 | 4.91 | 54.57% |
| | | Language of Social Studies | 9 | 5.18 | 57.61% |
| | C | Social Instructional Language | 6 | 3.79 | 63.13% |
| | | Language of Language Arts | 12 | 7.44 | 61.98% |
| | | Language of Math | 12 | 7.40 | 61.68% |
| | | Language of Science | 9 | 5.18 | 57.55% |
| | | Language of Social Studies | 9 | 4.75 | 52.81% |

## 4.4.1.2    By Grade

**Table 4.4.1.2**

Mean Raw Score by Grade by Tier by Standard: Comprehension S302

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 1 | A | Social Instructional Language | 12 | 7.72 | 64.30% |
| | | Language of Language Arts | 9 | 5.88 | 65.34% |
| | | Language of Math | 9 | 5.17 | 57.49% |
| | | Language of Science | 6 | 4.14 | 69.01% |
| | | Language of Social Studies | 6 | 3.61 | 60.23% |
| | B | Social Instructional Language | 6 | 4.19 | 69.84% |
| | | Language of Language Arts | 12 | 6.57 | 54.75% |
| | | Language of Math | 12 | 6.64 | 55.32% |
| | | Language of Science | 9 | 6.03 | 66.98% |
| | | Language of Social Studies | 9 | 5.58 | 62.02% |
| | C | Social Instructional Language | 6 | 4.48 | 74.75% |
| | | Language of Language Arts | 12 | 7.85 | 65.44% |
| | | Language of Math | 12 | 6.07 | 50.58% |
| | | Language of Science | 9 | 5.11 | 56.80% |
| | | Language of Social Studies | 9 | 5.41 | 60.13% |
| 2 | A | Social Instructional Language | 12 | 7.96 | 66.35% |
| | | Language of Language Arts | 9 | 6.31 | 70.09% |
| | | Language of Math | 9 | 5.62 | 62.47% |
| | | Language of Science | 6 | 4.23 | 70.54% |
| | | Language of Social Studies | 6 | 3.82 | 63.72% |
| | B | Social Instructional Language | 6 | 4.79 | 79.86% |
| | | Language of Language Arts | 12 | 8.06 | 67.17% |
| | | Language of Math | 12 | 8.25 | 68.74% |
| | | Language of Science | 9 | 7.20 | 80.02% |
| | | Language of Social Studies | 9 | 6.85 | 76.08% |
| | C | Social Instructional Language | 6 | 5.15 | 85.86% |
| | | Language of Language Arts | 12 | 9.50 | 79.19% |
| | | Language of Math | 12 | 7.52 | 62.63% |
| | | Language of Science | 9 | 6.09 | 67.72% |
| | | Language of Social Studies | 9 | 6.77 | 75.25% |

Return to Summary of Assessment Records Claims, Actions, and Evidence

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 3 | A | Social Instructional Language | 12 | 6.92 | 57.63% |
| | | Language of Language Arts | 9 | 3.81 | 42.30% |
| | | Language of Math | 9 | 4.58 | 50.87% |
| | | Language of Science | 6 | 3.06 | 51.00% |
| | | Language of Social Studies | 6 | 3.03 | 50.50% |
| | B | Social Instructional Language | 6 | 4.64 | 77.37% |
| | | Language of Language Arts | 12 | 7.26 | 60.51% |
| | | Language of Math | 12 | 6.52 | 54.30% |
| | | Language of Science | 9 | 4.95 | 55.02% |
| | | Language of Social Studies | 9 | 5.11 | 56.76% |
| | C | Social Instructional Language | 6 | 3.53 | 58.77% |
| | | Language of Language Arts | 12 | 7.31 | 60.88% |
| | | Language of Math | 12 | 4.66 | 38.80% |
| | | Language of Science | 9 | 4.73 | 52.59% |
| | | Language of Social Studies | 9 | 3.96 | 44.03% |
| 4 | A | Social Instructional Language | 12 | 7.13 | 59.39% |
| | | Language of Language Arts | 9 | 4.10 | 45.50% |
| | | Language of Math | 9 | 4.87 | 54.11% |
| | | Language of Science | 6 | 3.25 | 54.21% |
| | | Language of Social Studies | 6 | 3.33 | 55.57% |
| | B | Social Instructional Language | 6 | 4.89 | 81.49% |
| | | Language of Language Arts | 12 | 7.95 | 66.25% |
| | | Language of Math | 12 | 7.06 | 58.85% |
| | | Language of Science | 9 | 5.39 | 59.85% |
| | | Language of Social Studies | 9 | 5.54 | 61.57% |
| | C | Social Instructional Language | 6 | 3.76 | 62.66% |
| | | Language of Language Arts | 12 | 7.75 | 64.59% |
| | | Language of Math | 12 | 5.18 | 43.13% |
| | | Language of Science | 9 | 5.07 | 56.39% |
| | | Language of Social Studies | 9 | 4.35 | 48.37% |

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 5 | A | Social Instructional Language | 12 | 7.37 | 61.40% |
| | | Language of Language Arts | 9 | 4.32 | 48.01% |
| | | Language of Math | 9 | 5.04 | 55.94% |
| | | Language of Science | 6 | 3.43 | 57.23% |
| | | Language of Social Studies | 6 | 3.56 | 59.31% |
| | B | Social Instructional Language | 6 | 5.08 | 84.66% |
| | | Language of Language Arts | 12 | 8.53 | 71.06% |
| | | Language of Math | 12 | 7.55 | 62.90% |
| | | Language of Science | 9 | 5.77 | 64.13% |
| | | Language of Social Studies | 9 | 6.00 | 66.63% |
| | C | Social Instructional Language | 6 | 4.05 | 67.49% |
| | | Language of Language Arts | 12 | 8.37 | 69.79% |
| | | Language of Math | 12 | 5.78 | 48.13% |
| | | Language of Science | 9 | 5.55 | 61.70% |
| | | Language of Social Studies | 9 | 4.94 | 54.93% |
| 6 | A | Social Instructional Language | 12 | 5.84 | 48.69% |
| | | Language of Language Arts | 9 | 4.47 | 49.64% |
| | | Language of Math | 9 | 4.51 | 50.07% |
| | | Language of Science | 6 | 3.47 | 57.84% |
| | | Language of Social Studies | 6 | 2.97 | 49.43% |
| | B | Social Instructional Language | 6 | 3.76 | 62.62% |
| | | Language of Language Arts | 12 | 8.16 | 68.04% |
| | | Language of Math | 12 | 6.79 | 56.56% |
| | | Language of Science | 9 | 4.53 | 50.29% |
| | | Language of Social Studies | 9 | 5.09 | 56.52% |
| | C | Social Instructional Language | 6 | 3.93 | 65.51% |
| | | Language of Language Arts | 12 | 7.23 | 60.29% |
| | | Language of Math | 12 | 6.81 | 56.77% |
| | | Language of Science | 9 | 5.14 | 57.10% |
| | | Language of Social Studies | 9 | 3.94 | 43.78% |

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 7 | A | Social Instructional Language | 12 | 5.98 | 49.83% |
| | | Language of Language Arts | 9 | 4.58 | 50.86% |
| | | Language of Math | 9 | 4.67 | 51.91% |
| | | Language of Science | 6 | 3.57 | 59.52% |
| | | Language of Social Studies | 6 | 3.05 | 50.90% |
| | B | Social Instructional Language | 6 | 3.93 | 65.52% |
| | | Language of Language Arts | 12 | 8.52 | 71.04% |
| | | Language of Math | 12 | 7.17 | 59.72% |
| | | Language of Science | 9 | 4.81 | 53.44% |
| | | Language of Social Studies | 9 | 5.51 | 61.25% |
| | C | Social Instructional Language | 6 | 4.25 | 70.86% |
| | | Language of Language Arts | 12 | 8.05 | 67.11% |
| | | Language of Math | 12 | 7.35 | 61.26% |
| | | Language of Science | 9 | 5.56 | 61.77% |
| | | Language of Social Studies | 9 | 4.56 | 50.62% |
| 8 | A | Social Instructional Language | 12 | 6.10 | 50.80% |
| | | Language of Language Arts | 9 | 4.71 | 52.37% |
| | | Language of Math | 9 | 4.83 | 53.70% |
| | | Language of Science | 6 | 3.64 | 60.60% |
| | | Language of Social Studies | 6 | 3.21 | 53.45% |
| | B | Social Instructional Language | 6 | 4.10 | 68.35% |
| | | Language of Language Arts | 12 | 8.82 | 73.53% |
| | | Language of Math | 12 | 7.53 | 62.78% |
| | | Language of Science | 9 | 5.02 | 55.76% |
| | | Language of Social Studies | 9 | 5.91 | 65.70% |
| | C | Social Instructional Language | 6 | 4.49 | 74.82% |
| | | Language of Language Arts | 12 | 8.77 | 73.09% |
| | | Language of Math | 12 | 7.87 | 65.55% |
| | | Language of Science | 9 | 5.96 | 66.24% |
| | | Language of Social Studies | 9 | 5.14 | 57.06% |

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 9 | A | Social Instructional Language | 12 | 5.52 | 46.01% |
| | | Language of Language Arts | 9 | 4.94 | 54.86% |
| | | Language of Math | 9 | 4.70 | 52.22% |
| | | Language of Science | 6 | 3.19 | 53.21% |
| | | Language of Social Studies | 6 | 3.56 | 59.32% |
| | B | Social Instructional Language | 6 | 4.38 | 73.04% |
| | | Language of Language Arts | 12 | 8.10 | 67.51% |
| | | Language of Math | 12 | 7.12 | 59.32% |
| | | Language of Science | 9 | 4.86 | 54.04% |
| | | Language of Social Studies | 9 | 5.16 | 57.37% |
| | C | Social Instructional Language | 6 | 3.89 | 64.85% |
| | | Language of Language Arts | 12 | 7.57 | 63.11% |
| | | Language of Math | 12 | 7.43 | 61.90% |
| | | Language of Science | 9 | 5.20 | 57.82% |
| | | Language of Social Studies | 9 | 4.74 | 52.64% |
| 10 | A | Social Instructional Language | 12 | 6.05 | 50.44% |
| | | Language of Language Arts | 9 | 5.36 | 59.57% |
| | | Language of Math | 9 | 5.00 | 55.58% |
| | | Language of Science | 6 | 3.45 | 57.55% |
| | | Language of Social Studies | 6 | 3.77 | 62.85% |
| | B | Social Instructional Language | 6 | 4.28 | 71.30% |
| | | Language of Language Arts | 12 | 7.96 | 66.35% |
| | | Language of Math | 12 | 7.15 | 59.60% |
| | | Language of Science | 9 | 4.85 | 53.90% |
| | | Language of Social Studies | 9 | 5.12 | 56.91% |
| | C | Social Instructional Language | 6 | 3.75 | 62.43% |
| | | Language of Language Arts | 12 | 7.33 | 61.10% |
| | | Language of Math | 12 | 7.27 | 60.60% |
| | | Language of Science | 9 | 5.08 | 56.44% |
| | | Language of Social Studies | 9 | 4.66 | 51.80% |

| Grade | Tier | Standard | Maximum Score | Mean Score | Percent of Maximum |
|---|---|---|---|---|---|
| 11 | A | Social Instructional Language | 12 | 6.44 | 53.63% |
| | | Language of Language Arts | 9 | 5.55 | 61.71% |
| | | Language of Math | 9 | 5.18 | 57.59% |
| | | Language of Science | 6 | 3.56 | 59.34% |
| | | Language of Social Studies | 6 | 3.90 | 65.00% |
| | B | Social Instructional Language | 6 | 4.33 | 72.23% |
| | | Language of Language Arts | 12 | 8.17 | 68.10% |
| | | Language of Math | 12 | 7.40 | 61.66% |
| | | Language of Science | 9 | 5.02 | 55.79% |
| | | Language of Social Studies | 9 | 5.26 | 58.50% |
| | C | Social Instructional Language | 6 | 3.73 | 62.21% |
| | | Language of Language Arts | 12 | 7.43 | 61.92% |
| | | Language of Math | 12 | 7.49 | 62.46% |
| | | Language of Science | 9 | 5.26 | 58.43% |
| | | Language of Social Studies | 9 | 4.85 | 53.85% |
| 12 | A | Social Instructional Language | 12 | 6.30 | 52.52% |
| | | Language of Language Arts | 9 | 5.49 | 61.03% |
| | | Language of Math | 9 | 5.02 | 55.74% |
| | | Language of Science | 6 | 3.40 | 56.74% |
| | | Language of Social Studies | 6 | 3.74 | 62.26% |
| | B | Social Instructional Language | 6 | 4.30 | 71.74% |
| | | Language of Language Arts | 12 | 8.12 | 67.67% |
| | | Language of Math | 12 | 7.32 | 60.98% |
| | | Language of Science | 9 | 4.98 | 55.32% |
| | | Language of Social Studies | 9 | 5.23 | 58.15% |
| | C | Social Instructional Language | 6 | 3.68 | 61.39% |
| | | Language of Language Arts | 12 | 7.30 | 60.82% |
| | | Language of Math | 12 | 7.42 | 61.83% |
| | | Language of Science | 9 | 5.18 | 57.55% |
| | | Language of Social Studies | 9 | 4.81 | 53.42% |

## 4.4.2 Writing

### 4.4.2.1 By Cluster

**Table 4.4.2.1**

Mean Raw Score by Cluster by Tier by Standard: Writing S302

| Cluster | Tier | Standard | Mean Raw Score | | | Total | Percent of Maximum |
|---|---|---|---|---|---|---|---|
| | | | Linguistic Complexity | Vocabulary Usage | Language Control | | |
| 1-2 | A | Social Instructional Language | 5.18 | 4.81 | 4.24 | 14.23 | 19.76% |
| | B | Social Instructional Language | 1.18 | 1.92 | 1.20 | 4.30 | 23.88% |
| | | Language of Math / Science | 2.37 | 2.23 | 1.92 | 6.52 | 36.24% |
| | | Language of Language Arts / Social Studies | 2.48 | 2.31 | 1.74 | 6.54 | 36.32% |
| | C | Social Instructional Language | 2.72 | 2.45 | 2.08 | 7.25 | 40.26% |
| | | Language of Math / Science | 2.85 | 2.73 | 2.19 | 7.76 | 43.13% |
| | | Language of Language Arts / Social Studies | 2.94 | 2.63 | 2.19 | 7.75 | 43.07% |
| 3-5 | A | Social Instructional Language | 2.05 | 1.88 | 1.54 | 5.47 | 30.38% |
| | | Language of Math / Science | 2.12 | 2.22 | 1.73 | 6.06 | 33.69% |
| | | Language of Language Arts | 2.14 | 1.91 | 1.64 | 5.70 | 31.68% |
| | B | Social Instructional Language | 2.96 | 3.12 | 2.46 | 8.53 | 47.41% |
| | | Language of Math / Science | 2.94 | 3.06 | 2.56 | 8.55 | 47.50% |
| | | Language of Language Arts / Social Studies | 2.84 | 2.30 | 2.29 | 7.43 | 41.25% |
| | C | Social Instructional Language | 3.23 | 3.45 | 2.78 | 9.46 | 52.58% |
| | | Language of Math / Science | 3.12 | 2.97 | 2.67 | 8.76 | 48.68% |
| | | Language of Language Arts / Social Studies | 3.13 | 2.58 | 2.61 | 8.32 | 46.22% |
| 6-8 | A | Social Instructional Language | 2.23 | 1.94 | 1.81 | 5.98 | 33.21% |
| | | Language of Math / Science | 2.12 | 1.67 | 1.74 | 5.53 | 30.70% |
| | | Language of Language Arts | 2.29 | 2.07 | 1.73 | 6.10 | 33.88% |
| | B | Social Instructional Language | 3.30 | 2.92 | 2.68 | 8.90 | 49.43% |
| | | Language of Math / Science | 3.18 | 3.32 | 2.65 | 9.15 | 50.85% |
| | | Language of Language Arts / Social Studies | 3.18 | 2.65 | 2.56 | 8.38 | 46.58% |
| | C | Social Instructional Language | 3.63 | 3.12 | 3.01 | 9.76 | 54.25% |
| | | Language of Math / Science | 3.66 | 3.73 | 3.06 | 10.45 | 58.05% |
| | | Language of Language Arts / Social Studies | 3.55 | 2.98 | 2.91 | 9.45 | 52.48% |
| 9-12 | A | Social Instructional Language | 2.17 | 2.03 | 1.85 | 6.05 | 33.61% |
| | | Language of Math / Science | 2.23 | 2.05 | 1.69 | 5.98 | 33.20% |
| | | Language of Language Arts | 2.42 | 2.16 | 1.76 | 6.34 | 35.22% |
| | B | Social Instructional Language | 3.52 | 2.93 | 3.01 | 9.45 | 52.50% |
| | | Language of Math / Science | 3.35 | 2.96 | 2.84 | 9.15 | 50.86% |
| | | Language of Language Arts / Social Studies | 3.29 | 3.17 | 2.72 | 9.18 | 50.99% |
| | C | Social Instructional Language | 3.82 | 3.24 | 3.39 | 10.45 | 58.07% |
| | | Language of Math / Science | 3.36 | 3.64 | 3.04 | 10.04 | 55.78% |
| | | Language of Language Arts / Social Studies | 3.73 | 3.61 | 3.17 | 10.52 | 58.45% |

## 4.4.2.2 By Grade

**Table 4.4.2.2**
Mean Raw Score by Grade by Tier by Standard: Writing S302

| Grade | Tier | Standard | Linguistic Complexity | Vocabulary Usage | Language Control | Total | Percent of Maximum |
|---|---|---|---|---|---|---|---|
| 1 | A | Social Instructional Language | 5.13 | 4.76 | 4.15 | 14.04 | 19.50% |
| | B | Social Instructional Language | 1.21 | 1.88 | 1.11 | 4.19 | 23.27% |
| | | Language of Math / Science | 2.12 | 2.04 | 1.69 | 5.86 | 32.53% |
| | | Language of Language Arts / Social Studies | 2.15 | 2.06 | 1.44 | 5.65 | 31.38% |
| | C | Social Instructional Language | 2.51 | 2.22 | 1.82 | 6.55 | 36.38% |
| | | Language of Math / Science | 2.51 | 2.53 | 1.87 | 6.91 | 38.39% |
| | | Language of Language Arts / Social Studies | 2.60 | 2.38 | 1.85 | 6.82 | 37.91% |
| 2 | A | Social Instructional Language | 5.39 | 4.99 | 4.55 | 14.93 | 20.74% |
| | B | Social Instructional Language | 1.16 | 1.96 | 1.28 | 4.40 | 24.42% |
| | | Language of Math / Science | 2.60 | 2.40 | 2.12 | 7.11 | 39.52% |
| | | Language of Language Arts / Social Studies | 2.78 | 2.54 | 2.01 | 7.33 | 40.71% |
| | C | Social Instructional Language | 2.85 | 2.61 | 2.24 | 7.69 | 42.74% |
| | | Language of Math / Science | 3.07 | 2.85 | 2.39 | 8.31 | 46.15% |
| | | Language of Language Arts / Social Studies | 3.15 | 2.78 | 2.41 | 8.35 | 46.37% |
| 3 | A | Social Instructional Language | 1.98 | 1.80 | 1.45 | 5.23 | 29.05% |
| | | Language of Math / Science | 2.06 | 2.15 | 1.65 | 5.85 | 32.53% |
| | | Language of Language Arts | 2.07 | 1.86 | 1.56 | 5.49 | 30.49% |
| | B | Social Instructional Language | 2.86 | 2.99 | 2.34 | 8.19 | 45.52% |
| | | Language of Math / Science | 2.87 | 2.93 | 2.46 | 8.26 | 45.86% |
| | | Language of Language Arts / Social Studies | 2.71 | 2.16 | 2.16 | 7.03 | 39.08% |
| | C | Social Instructional Language | 3.13 | 3.33 | 2.67 | 9.12 | 50.66% |
| | | Language of Math / Science | 3.02 | 2.91 | 2.57 | 8.50 | 47.20% |
| | | Language of Language Arts / Social Studies | 2.99 | 2.41 | 2.48 | 7.89 | 43.85% |
| 4 | A | Social Instructional Language | 2.05 | 1.89 | 1.55 | 5.49 | 30.51% |
| | | Language of Math / Science | 2.13 | 2.22 | 1.74 | 6.09 | 33.83% |
| | | Language of Language Arts | 2.16 | 1.92 | 1.67 | 5.74 | 31.88% |
| | B | Social Instructional Language | 3.00 | 3.18 | 2.51 | 8.68 | 48.23% |
| | | Language of Math / Science | 2.97 | 3.12 | 2.60 | 8.68 | 48.24% |
| | | Language of Language Arts / Social Studies | 2.89 | 2.35 | 2.33 | 7.57 | 42.08% |
| | C | Social Instructional Language | 3.25 | 3.47 | 2.79 | 9.50 | 52.80% |
| | | Language of Math / Science | 3.14 | 2.97 | 2.68 | 8.80 | 48.88% |
| | | Language of Language Arts / Social Studies | 3.15 | 2.60 | 2.62 | 8.37 | 46.50% |

| Grade | Tier | Standard | Mean Raw Score | | | | Percent of Maximum |
| | | | Linguistic Complexity | Vocabulary Usage | Language Control | Total | |
|---|---|---|---|---|---|---|---|
| 5 | A | Social Instructional Language | 2.17 | 2.00 | 1.67 | 5.84 | 32.44% |
| | | Language of Math / Science | 2.23 | 2.31 | 1.84 | 6.39 | 35.48% |
| | | Language of Language Arts | 2.26 | 2.01 | 1.75 | 6.01 | 33.41% |
| | B | Social Instructional Language | 3.13 | 3.34 | 2.67 | 9.15 | 50.81% |
| | | Language of Math / Science | 3.06 | 3.28 | 2.73 | 9.07 | 50.40% |
| | | Language of Language Arts / Social Studies | 3.07 | 2.57 | 2.52 | 8.16 | 45.31% |
| | C | Social Instructional Language | 3.40 | 3.63 | 2.98 | 10.01 | 55.64% |
| | | Language of Math / Science | 3.29 | 3.06 | 2.84 | 9.19 | 51.04% |
| | | Language of Language Arts / Social Studies | 3.33 | 2.85 | 2.82 | 9.00 | 50.00% |
| 6 | A | Social Instructional Language | 2.17 | 1.88 | 1.75 | 5.80 | 32.23% |
| | | Language of Math / Science | 2.05 | 1.62 | 1.67 | 5.34 | 29.65% |
| | | Language of Language Arts | 2.25 | 2.05 | 1.70 | 6.00 | 33.32% |
| | B | Social Instructional Language | 3.18 | 2.83 | 2.54 | 8.55 | 47.49% |
| | | Language of Math / Science | 3.03 | 3.17 | 2.52 | 8.71 | 48.40% |
| | | Language of Language Arts / Social Studies | 3.01 | 2.49 | 2.41 | 7.91 | 43.94% |
| | C | Social Instructional Language | 3.46 | 3.01 | 2.83 | 9.30 | 51.67% |
| | | Language of Math / Science | 3.48 | 3.59 | 2.88 | 9.94 | 55.23% |
| | | Language of Language Arts / Social Studies | 3.34 | 2.77 | 2.70 | 8.81 | 48.97% |
| 7 | A | Social Instructional Language | 2.23 | 1.93 | 1.80 | 5.95 | 33.07% |
| | | Language of Math / Science | 2.12 | 1.67 | 1.74 | 5.54 | 30.75% |
| | | Language of Language Arts | 2.29 | 2.07 | 1.72 | 6.08 | 33.75% |
| | B | Social Instructional Language | 3.31 | 2.92 | 2.68 | 8.92 | 49.55% |
| | | Language of Math / Science | 3.19 | 3.35 | 2.67 | 9.21 | 51.18% |
| | | Language of Language Arts / Social Studies | 3.19 | 2.65 | 2.57 | 8.41 | 46.72% |
| | C | Social Instructional Language | 3.64 | 3.12 | 3.02 | 9.79 | 54.40% |
| | | Language of Math / Science | 3.67 | 3.75 | 3.07 | 10.49 | 58.29% |
| | | Language of Language Arts / Social Studies | 3.56 | 2.98 | 2.92 | 9.47 | 52.62% |
| 8 | A | Social Instructional Language | 2.30 | 2.01 | 1.87 | 6.18 | 34.34% |
| | | Language of Math / Science | 2.19 | 1.71 | 1.80 | 5.71 | 31.70% |
| | | Language of Language Arts | 2.35 | 2.10 | 1.77 | 6.22 | 34.56% |
| | B | Social Instructional Language | 3.43 | 3.01 | 2.81 | 9.25 | 51.37% |
| | | Language of Math / Science | 3.31 | 3.46 | 2.78 | 9.56 | 53.11% |
| | | Language of Language Arts / Social Studies | 3.34 | 2.81 | 2.71 | 8.87 | 49.26% |
| | C | Social Instructional Language | 3.79 | 3.24 | 3.18 | 10.22 | 56.77% |
| | | Language of Math / Science | 3.84 | 3.85 | 3.23 | 10.93 | 60.70% |
| | | Language of Language Arts / Social Studies | 3.76 | 3.20 | 3.11 | 10.07 | 55.97% |

| Grade | Tier | Standard | Mean Raw Score | | | | Percent of Maximum |
| | | | Linguistic Complexity | Vocabulary Usage | Language Control | Total | |
|---|---|---|---|---|---|---|---|
| 9 | A | Social Instructional Language | 2.06 | 1.95 | 1.75 | 5.77 | 32.04% |
| | | Language of Math / Science | 2.08 | 1.93 | 1.57 | 5.58 | 31.01% |
| | | Language of Language Arts | 2.26 | 2.05 | 1.62 | 5.94 | 33.00% |
| | B | Social Instructional Language | 3.50 | 2.89 | 3.02 | 9.41 | 52.30% |
| | | Language of Math / Science | 3.35 | 2.93 | 2.87 | 9.15 | 50.86% |
| | | Language of Language Arts / Social Studies | 3.29 | 3.15 | 2.73 | 9.16 | 50.89% |
| | C | Social Instructional Language | 3.84 | 3.22 | 3.44 | 10.49 | 58.30% |
| | | Language of Math / Science | 3.40 | 3.68 | 3.09 | 10.17 | 56.53% |
| | | Language of Language Arts / Social Studies | 3.76 | 3.63 | 3.22 | 10.60 | 58.91% |
| 10 | A | Social Instructional Language | 2.25 | 2.09 | 1.93 | 6.28 | 34.87% |
| | | Language of Math / Science | 2.36 | 2.14 | 1.79 | 6.28 | 34.90% |
| | | Language of Language Arts | 2.55 | 2.25 | 1.86 | 6.66 | 37.00% |
| | B | Social Instructional Language | 3.49 | 2.90 | 2.97 | 9.35 | 51.93% |
| | | Language of Math / Science | 3.32 | 2.94 | 2.81 | 9.07 | 50.40% |
| | | Language of Language Arts / Social Studies | 3.23 | 3.12 | 2.66 | 9.01 | 50.03% |
| | C | Social Instructional Language | 3.80 | 3.21 | 3.36 | 10.38 | 57.65% |
| | | Language of Math / Science | 3.33 | 3.61 | 3.00 | 9.93 | 55.17% |
| | | Language of Language Arts / Social Studies | 3.70 | 3.58 | 3.14 | 10.42 | 57.91% |
| 11 | A | Social Instructional Language | 2.37 | 2.17 | 2.02 | 6.56 | 36.44% |
| | | Language of Math / Science | 2.50 | 2.26 | 1.90 | 6.66 | 37.01% |
| | | Language of Language Arts | 2.68 | 2.35 | 1.98 | 7.01 | 38.94% |
| | B | Social Instructional Language | 3.55 | 2.97 | 3.03 | 9.55 | 53.08% |
| | | Language of Math / Science | 3.39 | 3.00 | 2.86 | 9.25 | 51.36% |
| | | Language of Language Arts / Social Studies | 3.34 | 3.23 | 2.75 | 9.32 | 51.78% |
| | C | Social Instructional Language | 3.84 | 3.28 | 3.37 | 10.49 | 58.26% |
| | | Language of Math / Science | 3.38 | 3.65 | 3.03 | 10.06 | 55.89% |
| | | Language of Language Arts / Social Studies | 3.75 | 3.64 | 3.17 | 10.56 | 58.68% |
| 12 | A | Social Instructional Language | 2.31 | 2.09 | 1.98 | 6.38 | 35.43% |
| | | Language of Math / Science | 2.45 | 2.18 | 1.87 | 6.49 | 36.08% |
| | | Language of Language Arts | 2.64 | 2.26 | 1.97 | 6.88 | 38.24% |
| | B | Social Instructional Language | 3.55 | 2.99 | 3.02 | 9.57 | 53.15% |
| | | Language of Math / Science | 3.37 | 2.97 | 2.83 | 9.17 | 50.96% |
| | | Language of Language Arts / Social Studies | 3.34 | 3.24 | 2.75 | 9.32 | 51.78% |
| | C | Social Instructional Language | 3.81 | 3.28 | 3.34 | 10.43 | 57.96% |
| | | Language of Math / Science | 3.32 | 3.58 | 2.97 | 9.88 | 54.87% |
| | | Language of Language Arts / Social Studies | 3.71 | 3.60 | 3.12 | 10.43 | 57.95% |

## 4.4.3 Speaking

### 4.4.3.1 By Cluster

**Table 4.4.3.2**
Mean Raw Score by Cluster by Tier by Standard: Speaking S302

| Cluster | Tier | Standard | Maximum Score | Mean Raw Score | Percentage of Maximum |
|---------|------|----------|---------------|----------------|------------------------|
| 1-2 | A | Social and Instructional Language | 3 | 2.35 | 78.46% |
| | | Language of Language Arts/Social Studies | 5 | 2.87 | 57.47% |
| | | Language of Mathematics/Science | 5 | 2.41 | 48.27% |
| | B | Social and Instructional Language | 3 | 2.84 | 94.72% |
| | | Language of Language Arts/Social Studies | 5 | 3.98 | 79.53% |
| | | Language of Mathematics/Science | 5 | 3.57 | 71.44% |
| | C | Social and Instructional Language | 3 | 2.93 | 97.68% |
| | | Language of Language Arts/Social Studies | 5 | 4.50 | 90.01% |
| | | Language of Mathematics/Science | 5 | 4.24 | 84.82% |
| 3-5 | A | Social and Instructional Language | 3 | 1.72 | 57.27% |
| | | Language of Language Arts/Social Studies | 5 | 2.01 | 40.21% |
| | | Language of Mathematics/Science | 5 | 1.62 | 32.39% |
| | B | Social and Instructional Language | 3 | 2.82 | 93.96% |
| | | Language of Language Arts/Social Studies | 5 | 3.73 | 74.52% |
| | | Language of Mathematics/Science | 5 | 3.49 | 69.81% |
| | C | Social and Instructional Language | 3 | 2.92 | 97.36% |
| | | Language of Language Arts/Social Studies | 5 | 4.24 | 84.82% |
| | | Language of Mathematics/Science | 5 | 4.10 | 81.95% |
| 6-8 | A | Social and Instructional Language | 3 | 1.69 | 56.24% |
| | | Language of Language Arts/Social Studies | 5 | 1.80 | 35.93% |
| | | Language of Mathematics/Science | 5 | 1.46 | 29.30% |
| | B | Social and Instructional Language | 3 | 2.83 | 94.24% |
| | | Language of Language Arts/Social Studies | 5 | 3.97 | 79.48% |
| | | Language of Mathematics/Science | 5 | 3.55 | 70.91% |
| | C | Social and Instructional Language | 3 | 2.93 | 97.55% |
| | | Language of Language Arts/Social Studies | 5 | 4.44 | 88.88% |
| | | Language of Mathematics/Science | 5 | 4.11 | 82.21% |
| 9-12 | A | Social and Instructional Language | 3 | 1.65 | 55.00% |
| | | Language of Language Arts/Social Studies | 5 | 1.63 | 32.61% |
| | | Language of Mathematics/Science | 5 | 1.47 | 29.33% |
| | B | Social and Instructional Language | 3 | 2.77 | 92.17% |
| | | Language of Language Arts/Social Studies | 5 | 3.83 | 76.56% |
| | | Language of Mathematics/Science | 5 | 3.44 | 68.77% |
| | C | Social and Instructional Language | 3 | 2.90 | 96.73% |
| | | Language of Language Arts/Social Studies | 5 | 4.50 | 89.99% |
| | | Language of Mathematics/Science | 5 | 4.16 | 83.23% |

## 4.4.3.2    By Grade

**Table 4.4.3.1**

Mean Raw Score by Grade by Tier by Standard: Speaking S302

| Grade | Tier | Standard | Maximum Score | Mean Raw Score | Percent of Maximum |
|-------|------|----------|---------------|----------------|--------------------|
| 1 | A | Social and Instructional Language | 3 | 2.40 | 80.05% |
| | | Language of Language Arts/Social Studies | 5 | 2.93 | 58.66% |
| | | Language of Mathematics/Science | 5 | 2.47 | 49.42% |
| | B | Social and Instructional Language | 3 | 2.80 | 93.47% |
| | | Language of Language Arts/Social Studies | 5 | 3.79 | 75.90% |
| | | Language of Mathematics/Science | 5 | 3.39 | 67.75% |
| | C | Social and Instructional Language | 3 | 2.90 | 96.79% |
| | | Language of Language Arts/Social Studies | 5 | 4.33 | 86.66% |
| | | Language of Mathematics/Science | 5 | 4.05 | 80.90% |
| 2 | A | Social and Instructional Language | 3 | 2.18 | 72.52% |
| | | Language of Language Arts/Social Studies | 5 | 2.65 | 53.03% |
| | | Language of Mathematics/Science | 5 | 2.20 | 43.96% |
| | B | Social and Instructional Language | 3 | 2.87 | 95.82% |
| | | Language of Language Arts/Social Studies | 5 | 4.14 | 82.74% |
| | | Language of Mathematics/Science | 5 | 3.74 | 74.72% |
| | C | Social and Instructional Language | 3 | 2.95 | 98.25% |
| | | Language of Language Arts/Social Studies | 5 | 4.61 | 92.14% |
| | | Language of Mathematics/Science | 5 | 4.37 | 87.33% |
| 3 | A | Social and Instructional Language | 3 | 1.77 | 58.98% |
| | | Language of Language Arts/Social Studies | 5 | 2.06 | 41.23% |
| | | Language of Mathematics/Science | 5 | 1.62 | 32.46% |
| | B | Social and Instructional Language | 3 | 2.81 | 93.75% |
| | | Language of Language Arts/Social Studies | 5 | 3.65 | 73.04% |
| | | Language of Mathematics/Science | 5 | 3.39 | 67.87% |
| | C | Social and Instructional Language | 3 | 2.92 | 97.31% |
| | | Language of Language Arts/Social Studies | 5 | 4.18 | 83.62% |
| | | Language of Mathematics/Science | 5 | 4.03 | 80.60% |
| 4 | A | Social and Instructional Language | 3 | 1.68 | 56.05% |
| | | Language of Language Arts/Social Studies | 5 | 1.97 | 39.47% |
| | | Language of Mathematics/Science | 5 | 1.60 | 32.03% |
| | B | Social and Instructional Language | 3 | 2.82 | 93.94% |
| | | Language of Language Arts/Social Studies | 5 | 3.76 | 75.14% |
| | | Language of Mathematics/Science | 5 | 3.53 | 70.61% |
| | C | Social and Instructional Language | 3 | 2.92 | 97.21% |
| | | Language of Language Arts/Social Studies | 5 | 4.23 | 84.59% |
| | | Language of Mathematics/Science | 5 | 4.08 | 81.53% |

| Grade | Tier | Standard | Maximum Score | Mean Raw Score | Percent of Maximum |
|---|---|---|---|---|---|
| 5 | A | Social and Instructional Language | 3 | 1.67 | 55.81% |
| | | Language of Language Arts/Social Studies | 5 | 1.97 | 39.38% |
| | | Language of Mathematics/Science | 5 | 1.63 | 32.69% |
| | B | Social and Instructional Language | 3 | 2.84 | 94.51% |
| | | Language of Language Arts/Social Studies | 5 | 3.86 | 77.20% |
| | | Language of Mathematics/Science | 5 | 3.67 | 73.36% |
| | C | Social and Instructional Language | 3 | 2.93 | 97.62% |
| | | Language of Language Arts/Social Studies | 5 | 4.36 | 87.19% |
| | | Language of Mathematics/Science | 5 | 4.24 | 84.81% |
| 6 | A | Social and Instructional Language | 3 | 1.73 | 57.81% |
| | | Language of Language Arts/Social Studies | 5 | 1.86 | 37.24% |
| | | Language of Mathematics/Science | 5 | 1.52 | 30.42% |
| | B | Social and Instructional Language | 3 | 2.82 | 93.91% |
| | | Language of Language Arts/Social Studies | 5 | 3.91 | 78.29% |
| | | Language of Mathematics/Science | 5 | 3.47 | 69.46% |
| | C | Social and Instructional Language | 3 | 2.91 | 97.09% |
| | | Language of Language Arts/Social Studies | 5 | 4.34 | 86.80% |
| | | Language of Mathematics/Science | 5 | 3.97 | 79.45% |
| 7 | A | Social and Instructional Language | 3 | 1.67 | 55.64% |
| | | Language of Language Arts/Social Studies | 5 | 1.76 | 35.22% |
| | | Language of Mathematics/Science | 5 | 1.43 | 28.67% |
| | B | Social and Instructional Language | 3 | 2.83 | 94.31% |
| | | Language of Language Arts/Social Studies | 5 | 3.98 | 79.60% |
| | | Language of Mathematics/Science | 5 | 3.55 | 71.07% |
| | C | Social and Instructional Language | 3 | 2.93 | 97.66% |
| | | Language of Language Arts/Social Studies | 5 | 4.45 | 89.09% |
| | | Language of Mathematics/Science | 5 | 4.12 | 82.34% |
| 8 | A | Social and Instructional Language | 3 | 1.66 | 55.25% |
| | | Language of Language Arts/Social Studies | 5 | 1.77 | 35.33% |
| | | Language of Mathematics/Science | 5 | 1.44 | 28.79% |
| | B | Social and Instructional Language | 3 | 2.84 | 94.51% |
| | | Language of Language Arts/Social Studies | 5 | 4.03 | 80.63% |
| | | Language of Mathematics/Science | 5 | 3.61 | 72.29% |
| | C | Social and Instructional Language | 3 | 2.94 | 97.93% |
| | | Language of Language Arts/Social Studies | 5 | 4.54 | 90.83% |
| | | Language of Mathematics/Science | 5 | 4.25 | 84.94% |
| 9 | A | Social and Instructional Language | 3 | 1.50 | 50.07% |
| | | Language of Language Arts/Social Studies | 5 | 1.46 | 29.12% |
| | | Language of Mathematics/Science | 5 | 1.32 | 26.32% |
| | B | Social and Instructional Language | 3 | 2.78 | 92.76% |
| | | Language of Language Arts/Social Studies | 5 | 3.94 | 78.86% |
| | | Language of Mathematics/Science | 5 | 3.53 | 70.68% |
| | C | Social and Instructional Language | 3 | 2.92 | 97.31% |
| | | Language of Language Arts/Social Studies | 5 | 4.58 | 91.65% |
| | | Language of Mathematics/Science | 5 | 4.24 | 84.73% |

| Grade | Tier | Standard | Maximum Score | Mean Raw Score | Percent of Maximum |
|---|---|---|---|---|---|
| 10 | A | Social and Instructional Language | 3 | 1.74 | 58.00% |
| | | Language of Language Arts/Social Studies | 5 | 1.72 | 34.30% |
| | | Language of Mathematics/Science | 5 | 1.55 | 30.97% |
| | B | Social and Instructional Language | 3 | 2.74 | 91.38% |
| | | Language of Language Arts/Social Studies | 5 | 3.72 | 74.33% |
| | | Language of Mathematics/Science | 5 | 3.34 | 66.73% |
| | C | Social and Instructional Language | 3 | 2.90 | 96.65% |
| | | Language of Language Arts/Social Studies | 5 | 4.49 | 89.74% |
| | | Language of Mathematics/Science | 5 | 4.14 | 82.86% |
| 11 | A | Social and Instructional Language | 3 | 1.88 | 62.77% |
| | | Language of Language Arts/Social Studies | 5 | 1.91 | 38.12% |
| | | Language of Mathematics/Science | 5 | 1.69 | 33.83% |
| | B | Social and Instructional Language | 3 | 2.77 | 92.20% |
| | | Language of Language Arts/Social Studies | 5 | 3.77 | 75.34% |
| | | Language of Mathematics/Science | 5 | 3.38 | 67.68% |
| | C | Social and Instructional Language | 3 | 2.90 | 96.58% |
| | | Language of Language Arts/Social Studies | 5 | 4.45 | 88.90% |
| | | Language of Mathematics/Science | 5 | 4.12 | 82.43% |
| 12 | A | Social and Instructional Language | 3 | 1.97 | 65.55% |
| | | Language of Language Arts/Social Studies | 5 | 2.07 | 41.41% |
| | | Language of Mathematics/Science | 5 | 1.84 | 36.83% |
| | B | Social and Instructional Language | 3 | 2.76 | 92.04% |
| | | Language of Language Arts/Social Studies | 5 | 3.82 | 76.39% |
| | | Language of Mathematics/Science | 5 | 3.45 | 69.02% |
| | C | Social and Instructional Language | 3 | 2.87 | 95.76% |
| | | Language of Language Arts/Social Studies | 5 | 4.40 | 87.95% |
| | | Language of Mathematics/Science | 5 | 4.07 | 81.36% |

World-Class Instructional Design and Assessment

# Annual Technical Report for
# ACCESS for ELLs English Language Proficiency Test,
# Series 302, 2013-2014 Administration

**Annual Technical Report No. 10**
**Volume 2 of 3: Analyses of Test Forms**

Prepared by:

Center for Applied Linguistics

CAL/WIDA Partnership Activities
Psychometrics/Research Team

May 30, 2015

## Table of Contents

Volume 2

# 5. Analyses of Test Forms: Overview

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

## 5.1 Background

### 5.1.1 Measurement Models Used

The measurement model that forms the basis of the analysis for the development of ACCESS for ELLs is the Rasch measurement model (Wright & Stone, 1979). Additional information on its use in the development of the test is available in WIDA Technical Report 1, *Development and Field Test of ACCESS for ELLs* (Kenyon, 2006). The test was developed using Rasch measurement principles, and in that sense the Rasch model guided all decisions throughout the development of the assessment and was not just a tool for the statistical analysis of the data. Thus, for example, data based on Rasch fit statistics guided the inclusion, revision, or deletion of items during the development and field testing of the test forms, and will continue to guide the refinement and further development of the test.

For Listening, Reading, and Speaking, 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 a correct response "1" by person "n" on item "i"

$P_{ni0}$ = probability of an incorrect response "0" by person "n" on item "i"

$B_n$ = ability of person "n"

$D_i$ = difficulty of item "i"

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

For the Writing tasks, a Rasch Rating Scale model was used. Mathematically, this can be represented as

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

where

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

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

$B_n$ = ability of person "n"

$D_i$ = difficulty of task "i"

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

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

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

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

Linacre (2002, Autumn), the author of the Winsteps program, provides more guidance on how to interpret these statistics for test items. He writes:

- values greater than 2.0 "distort or degrade the measurement system"
- values between 1.5 and 2.0 are "unproductive for construction of measurement, but not degrading"

- values between 0.5 and 1.5 should be considered "productive for measurement"
- values below 0.5 Linacre calls "less productive for measurement, but not degrading"

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

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

## 5.1.2 Sampling

The results presented in most of the tables in Chapter 6 are based on the full data set of all students who were administered operational Series 302 of ACCESS for ELLs in the academic year 2013–2014. Exceptions are Tables E, G, H, and I. The equating summary tables (Table E) use data from a sample of about 1,000 students rather than the entire population of students, because the equating was done in the midst of the operational scoring. The item or task analysis summary tables (Table G), the complete item analysis tables (Table H), and the raw score to scale score conversion tables (Table I) use item and task difficulties from this equating.

## 5.1.3 Equating and Scaling

Complete information on the horizontal and vertical scaling of ACCESS for ELLs scores is provided in Technical Report 1, *Development and Field Test of ACCESS for ELLs®*. In brief, this scaling was accomplished during the field test based on an elaborate common item design, both across tiers and across grade-level clusters, which spanned two series of complete test forms. Concurrent calibration was used to determine item difficulty measures. These item difficulty measures were used to create the ACCESS for ELLs scale scores used for reporting results on the test. Table D in Section 6 for each form provides the equation for converting Rasch ability measures in logits to ACCESS for ELLs scale scores.

The operational test forms in Series 302 represent a partial refreshment of Series 301. That is, while many items were common on both forms, certain folders on Series 301 were replaced with new items (see Chapter 1.4). Thus, to place results on Series 302 onto the ACCESS for ELLs scale score, items that were not revised or otherwise changed were anchored to the difficulty values from Series 301, which itself had been anchored to Series 203. Table E in Section 6 for each test form provides explicit information on the anchor items used for equating Series 302 results to those of Series 301.

## 5.1.4 DIF Analyses

Differential item analyses (DIF) attempt to investigate whether performances on items were influenced by factors extraneous to English language proficiency (i.e., the construct being measured on the test). In other words, it attempts to find items that may be functioning differently for different groups based on criteria irrelevant to what is being tested. The performance of students on the ACCESS for ELLs items was compared by dividing students into two different groupings: first, males versus females; second, students of Hispanic ethnic background versus students of all other backgrounds. (For both analyses, students for whom gender or ethnicity was missing were excluded.) Two commonly used procedures for detecting

DIF were used: one for dichotomously scored items (Listening, Reading, and Speaking) and one for polytomously scored items (Writing).

### 5.1.4.1 Dichotomous Items

Following procedures originally proposed by the Educational Testing Service (ETS), the Mantel-Haenszel Chi-square statistic was used for dichotomous items. This procedure compares item-level performances of students in the two groups (e.g., males versus females) who are divided into subgroups based on their performance on the total test. It is assumed that, if there is no DIF, at any ability level (based on performance on the total test), a similar percentage of students in each group should get the item correct. The Mantel-Haenszel Chi-square statistic is used to check the probability that the two groups performed the same on each item across the ability groupings. The statistic is transformed into a scale called 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 is worrying. We follow the guidance provided by ETS to classify items into DIF levels as follows:

- A (no DIF), when the absolute value of delta was less than 1.0

- B (weak DIF), when the absolute value of delta was between 1.0 and 1.5

- C (strong DIF), when the absolute value of the delta was greater than 1.5

The software program *EZDIF* (Waller, n.d.) was used 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 statistics to 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.)

(For information on procedures for dealing with items with C-level DIF, see Section 1.4.5.)

### 5.1.4.2 Polytomous Items

For polytomous items (i.e., the Writing tasks), a similar approach is used. It is based on the Mantel Chi-square statistic and the standardized mean difference following procedures again developed by ETS. As with dichotomous items, the underlying assumption is that students who performed similarly overall on the test should perform similarly on the individual tasks. To test this assumption, students are placed into 6 groups based on their total raw score on the Writing test. We determined these categories by calculating what the total raw score of a student scoring WIDA Proficiency Levels 1, 2, 3, 4, 5, or 6 in each category would be. For example, a student consistently scoring 1 would have a total score of 18 on a Tier B or Tier C form. A student consistently scoring 2 would score a 36.

To divide the students into performance groups in this way, cut points were made halfway between the above totals, such that students in Group 1 would have a total score of 0 to 27;

Group 2 totaled 28 to 45; Group 3 totaled 46 to 63; Group 4 totaled 64 to 81; and Group 5 totaled 82 to 108. (Note that Group 5 contained students scoring in the 6 range. These two groups were combined because there are so few students in that category.)

For each Writing task, performance was similarly categorized according to the scoring rubric. Thus, raw scores of 0 to 4 were category 1 (i.e., up to a score totaling 4, such as 2-1-1, which is a high 1 but not yet a 2); the raw scores of 5 to 7 were category 2; the raw scores of 8 to 10 were category 3; the raw scores of 11 to 13 were category 4; the raw scores of 14 to 16 were category 5; and the raw scores of 17 to 18 were category 6. (The only exception to this was Kindergarten Writing tasks, where there was much smaller spread of scores on the Writing tasks. In such cases, total raw scores were used to determine categories.)

Following formulae provided by Zwick, Donoghue, and Grima (1993), an Excel spreadsheet was programmed to take cross-tabulated data output by SPSS and calculate the Mantel statistic and determine its probability of significance. This statistic gives an indication of the probability that observed differences are the result of chance but does not indicate how significant that difference is. To indicate how significant the difference is, the standardized mean difference (SMD) between the performances of the two groups being compared is calculated. The standardized mean difference compares the means of the two groups, adjusting for differences in the distribution of the two groups being compared across the values of the matching variable. To standardize the outcome, this difference is divided by the standard deviation (SD) of the item for the total group. This calculation is also programmed into the Excel spreadsheet.

Following guidance proposed by ETS, polytomously scaled items are classified into DIF levels as follows:

- AA (no DIF), when the Mantel Chi-square statistic is not significant; or, when it is significant and the absolute value of (SMD/SD) is less than or equal to .17

- BB (weak DIF), when the Mantel Chi-square statistic is significant and the absolute value of (SMD/SD) is greater than .17 but less than or equal to .25

- CC (strong DIF), when the Mantel Chi-square statistic is significant and the absolute value of (SMD/SD) is greater than .25

## *5.2      Descriptions*

The following paragraphs describe the tables that follow and are repeated for each test form in each domain.

### 5.2.1      Raw Score Information (Figure A and Table A)

Figure A and Table A relate to the *raw scores* on each test form. Listening, Reading, and Speaking were scored dichotomously (i.e., right or wrong). Thus, the highest possible score was the number of items on the test form. Each Writing task, however, could be awarded up to 18 points. Additionally, certain Writing tasks are weighted because of their potential to elicit higher levels of writing ability. For cluster 1–2, Tier A has a weight of 3 for the fourth task. For clusters 1–2, 3–5, 6–8, and 9–12, Tiers B and C have a weight of 2 for the second task and a weight of 3 for the third task. Thus, the maximum number of points on each Writing test form varies from 54 for the Tier A forms for clusters 3–5, 6–8, and 9–12 to 108 for the Tier B and C forms and cluster 1–2 Tier A (see Chapter 1.7.2).

For each test form, Figure A shows the distribution of the raw scores. The horizontal axis shows the raw scores. The vertical axis shows the number of students (count). Each bar shows how many students were awarded each raw score.

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

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

- The minimum observed raw score

- The maximum observed raw score

- The mean (average) raw score

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

### 5.2.2      Scale Score Information (Figure B and Table B)

Figure B and Table B relate to the ACCESS for ELLs *scale scores* on each test form. For each test form, raw scores were converted to vertically-equated scale scores. (The raw score to scale score conversion table for each test form is given as the last table―Table I―in each section.)

Thus, for each test form, Figure B shows the distribution of the scale scores. The horizontal axis shows the scale scores based on performances on the test form. To provide full perspective, it extends somewhat below and above the range of possible or observed scale scores. The vertical axis shows the number of students (count). Each bar shows how many students were awarded each scale score.

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

- Number of students in the analyses

- The minimum observed scale score

- The maximum observed scale score

- The mean (average) scale score

- The standard deviation (std. dev.) of the scale scores

Note that scale scores for Tier A and Tier B in Listening and Reading are capped. Within each grade, the highest possible scale score for Tier A is the scale score corresponding to the cut score for Proficiency Level 4 (i.e., proficiency level score of 4.0). For Tier B, the highest possible scale score within each grade is the score corresponding to the cut score for Proficiency Level 5 (i.e., proficiency level score of 5.0). Because of these grade-level cut scores, the scale score associated with a given proficiency level score increases by grade level within a cluster, and so the cap also increases by grade level. For example, for Listening 3–5A, the scale score is capped at 325 for Grade 3, 338 for Grade 4, and 350 for Grade 5 (see Table 6.3.1.1B). Thus, a third grade student with a raw score of 18 (out of 18) on that test will have a scale score of 325, a fourth grader with the same raw score will have a scale score of 338, and a fifth grader with the same raw score will have a scale score of 350. However, all three students would have a proficiency level score of 4.0. For more information, see WIDA Technical Report 1, *Development and Field Test of ACCESS for ELLs* (Kenyon, 2006).

Also note that, because the scale is vertically equated, the range of scale scores moves up the scale from one cluster to the next. Thus, a second grade student with a raw score of 0 on the Listening Tier A test would have a scale score of 108, while a fifth grade student with a raw score of 0 on the Listening Tier A test would have a scale score of 120.

Similarly, scale scores at the lower end may be truncated so that the lowest achievable proficiency level score is 1.0. Again, this results in a lower minimum scale score for students in lower grade levels within a grade-level cluster.

The influence of these cuts will also be noticed in Figure B, as well as in many other tables throughout the report.

## 5.2.3    Proficiency Level Information (Figure C and Table C)

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

Each row of Table C shows, by each grade in the cluster and by total for the cluster:

- The WIDA Proficiency Level designation (1 to 6)

- The number of students (count) whose performance on the test form placed them into that Proficiency Level in the domain being tested

- The percent of students, out of the total number of students taking the form (by grade or by total for the cluster), who were placed into that Proficiency Level in the domain being tested

(Note that for some domains for Kindergarten and Tier A tests, it was not possible to place into all proficiency levels. Figure C and Table C also clearly show the effect of the scoring cap on Tiers A and B.)

For Kindergarten this information is provided for scores based on both the Accountability cut scores and the Instructional cut scores.

## 5.2.4    Scaling Equation Table (Table D)

For each test form, Table D provides the scaling equation for that domain. This equation is used to convert an examinee's ability measure into the scale score. Because ACCESS for ELLs is vertically equated (see 5.1.3 above), though each domain has its own equation, the same equation is used across all tiers and grade-level clusters within each domain.

## 5.2.5    Equating Summary (Table E)

Each year a certain percentage of items on each ACCESS for ELLs test form are refreshed. A post-equating procedure known as common item equating is used to equate results on new forms to the older forms. This means that the difficulty measure of items appearing on the new form that are the same as those on the older form are kept constant across both forms. Thus, performances on the newer form may be interpreted in the same frame of reference.

Many items appearing on ACCESS for ELLs Series 302 also appeared on Series 301. All items common to both forms were anchored in the first equating run. After the first equating run, some items that were originally 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 its difficulty value would have been had it not been anchored. For Listening and Reading items, and for Writing and Speaking tasks, if this value was large (i.e., usually above .30 or below -.30), that item was unanchored in the final equating run (i.e., it was treated as if it were a new item).

Table E presents a summary of the common item equating procedures. The first section of the table compares the current test (i.e., the Series 302 version of that test form) to the previous year's test (i.e., the Series 301 version of that test form). The number of items, the average item difficulty, the standard deviation of the item difficulty values, as well as the difficulty value of the easiest and hardest item on each test form is presented. These values are in terms of logits used in the Rasch measurement model.

The second section of the table presents information on the anchoring items. The total number of possible anchors (i.e., all common items) is shown, as well as the standard deviation of those items. Next, the number of items that were actually anchored (i.e., in general, those items whose displacement values were below .30 or above -.30) in the final equating run is shown, again with the average item difficulty and standard deviation. Finally, the percentage of items that served as anchors and the average displacement value is given. Generally speaking, the greater the number of tasks anchored and the closer the average displacement is to 0.00, the more trustworthy the equating results will be.

The final section of Table E shows the location of the anchor items or tasks, both by order on the test form and by order of difficulty. It is desirable that the anchored items appear throughout the test form in order to ensure that no systematic bias affects performance on them (e.g., if they all appear at the end of a test form, there may be a fatigue effect). It is also desirable that the anchor items represent a wide range of difficulties across the entire spectrum of the item difficulty values on a test form. The greater the representation across the difficulty range, the more trustworthy the equating results will be. This section also provides information on displacement;

that is, the difference between the difficulty value of the anchored item and what that difficulty value would have been had the item not been anchored. Smaller displacement statistics indicate more consistency between the item's difficulty value on the Series 302 test form and on the Series 301 test form. Typically, random displacements of less than 0.5 logits are unlikely to have much impact on measurement in a test instrument (Linacre, n.d.).

Note that for the Writing tasks, this table also provides the anchored step measures for the total score on each task. For the ACCESS Writing tasks, a rating scale model is used (see 5.1.1 above). Because a single generic rubric based on the generic WIDA Performance Level definitions is used to score all of the Writing tasks across all of the grade-level clusters, we modeled a rating scale that has the same step difficulty values across all Writing tasks across all grade-level clusters. Thus, these values are the same for every Writing task on ACCESS. These constant step difficulty values help to provide anchors in the calibration of new Writing tasks onto the common WIDA score scale each year.

Note that because the Kindergarten test form was newly created for Series 200, it was not equated to the Series 103 test. Therefore, Table E is not included for Kindergarten. For technical details on the Kindergarten test, see MacGregor, Kenyon, Gibson, and Evans, (2009). In addition, in the other grade-level clusters, scores for the Speaking test are based on a content analysis rather than on equating to previous forms; therefore, Table E is included only to verify that the raw score to scale score conversion remains within reasonable parameters.

Note that for Series 302, no equating was performed for Writing Tiers A for all grade clusters. Also, no equating was performed for all Reading grade clusters. The results of the unequated tests needed for certain tables were taken from the results of Series 301.

## 5.2.6    Test Characteristic Curve (Figure D)

For each test form, Figure D graphically shows the relationship between the ability measure (in logits) on the horizontal axis and the expected raw score on the vertical axis. Five vertical lines indicate the five cut scores for the highest grade in the cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (Levels 1–6) for the domain being tested. (Note that for some domains for Kindergarten and Tier A tests, it was not possible to place into all six language proficiency levels. As would be expected, higher raw scores are required to be placed into higher language proficiency levels. The relative width of each section between the cut score lines, however, gives an indication of how many items on that form must be answered correctly (or points on the Writing section must be earned) to be placed into a WIDA language Proficiency Level.

## 5.2.7    Test Information Function (Figure E)

With the Rasch measurement model, as with any measurement model following Item Response Theory (IRT), the relationship between the ability measure (in logits) and the accuracy of test scores can be modeled. It is recognized that tests measure most accurately when the abilities of the examinees and the difficulty of the items are most appropriate for each other. If a test is too difficult for an examinee (i.e., the examinee scores close to zero), or if the test is too easy for an examinee (i.e., the examinee "tops out"), accurate measurement of the examinee's ability cannot be made. The test information function shows graphically how well the test is measuring across

the ability measure spectrum. High values indicate more accuracy in measurement. Thus, for each test form, Figure E shows the relationship between the ability measure (in logits) on the horizontal axis and measurement accuracy, represented as the Fisher information value (which is the inverse squared of the standard error), on the vertical axis. The test information function, then, reflects the conditional standard error of measurement.

Again, as in Figure D, five vertical lines in Figure E indicate the five cut scores for the highest grade in the cluster for the test form, dividing the figure into six sections for each of the WIDA language proficiency levels (1–6) for the domain being tested. (Note that for some domains for Kindergarten and Tier A tests, it was not possible to place into all six language proficiency levels. Note also that, although Listening and Reading scores on Tiers A and B were capped, all 5 horizontal lines indicating the cut points remain in this figure.) It is important that each test form measure most accurately in the areas for which it is primarily used to make classification decisions. In other words, optimally the test information function should be high for the cuts between 1/2 and 2/3 for Tier A test forms; between 2/3, 3/4, and 4/5 for Tier B test forms; and between 3/4, 4/5, and 5/6 for Tier C test forms.

## 5.2.8    Reliability (Table F)

In contrast to Figure E, which is based on the Rasch measurement model, Table F presents reliability and accuracy information based on Classical Test Theory. It shows:

- The number of students

- The number of items

- Cronbach's coefficient alpha (as a measure of internal consistency)

- The classical standard error of measurement (SEM) in terms of ***raw scores***

Cronbach's coefficient alpha is widely used as an estimate of reliability, particularly of the internal consistency of test items. It expresses how well the items on a test appear to measure the same construct. Conceptually, it may be thought of as the correlation obtained between performances on two halves of the test, if every possibility of dividing the test items in two were attempted. Thus, Cronbach's alpha may be low if some items are measuring something other than what the majority of the items are measuring. As with any reliability index, it is affected by the number of test items (or test score points that may be awarded). That is, all things being equal, the greater the number of items, the higher the reliability.

Cronbach's alpha is also affected by the distribution of ability within the group of students tested. All things being equal, the greater the heterogeneity of abilities within the group of students tested (i.e., the more widely the scores are distributed), the higher the reliability. In this sense, Cronbach's alpha is *sample dependent.* It is widely recognized that reliability can be as much a function of the test as of the sample of students tested. That is, the exact same test can produce widely disparate reliability indices based on ability distribution of the group of students tested. Because ACCESS for ELLs is a tiered test (that is, because each form in Tier A, B, or C targets only a certain range of the entire ability distribution), results for reliability on any one form, particularly for the shorter Listening test, may at times be lower than typically expected.

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 $i$

$\sigma_i^2$ = variance of score on item $i$

$\sigma_t^2$ = variance of total score

Table F also presents the *standard error of measurement* (SEM) based on classical test theory. Unlike IRT, in this approach, SEM is seen as a constant across the spread of test scores (ability continuum). Thus, it is **not** conditional on ability being measured. It is, however, a function of two statistics: the reliability of the test and the (observed) standard deviation of the test scores. It is calculated as

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

Traditionally, SEM has been used to create a band around an examinee's observed score, with the assertion in the view of classical test theory, that the examinee's true score (i.e., what the examinee's score would be if it could be measured without error) would lie with a certain degree of probability within this band. Statistically speaking, then, there is an expectation that an examinee's true score has a 68% probability of lying within the band extending from the observed score minus 1 SEM to the observed score plus 1 SEM.

For the Writing tests (except Kindergarten, which is scored by the test administrator), information on inter-rater reliability is also provided in Table F. This portion of the table shows, for each of the three or four Writing tasks, the percent of agreement between two raters in terms of the three features being rated: Linguistic Complexity (LX), Vocabulary Usage (VU), and Language Control (LC). In this part of the table, the first column shows the Writing task (i.e., the first, second, third, or fourth, if applicable). The second column shows the number of Writing papers that were double scored. This number is generally 25% of all papers scored, chosen at random during the operational scoring process. The next column shows the feature, while the following columns show the rates of agreement: exact, adj (adjacent), and total sum of exact and adjacent. When the two raters agreed on the score, an exact agreement was counted. If the two raters were different in that feature by one point, an adjacent agreement was counted.

## 5.2.9    Item/Task Analysis Summary (Table G)

Table G provides a summary of the analyses of the items (for Listening and Reading) or the tasks (for Writing and Speaking). The top part of the table gives an item or task summary. The first column in this part states the type of item (MC for multiple choice or ECR for extended constructed response). The next column shows the number of items or tasks on the test form. The next column gives the average item or task difficulty value in logits. For the multiple-choice items, the next column shows the average p-value. This is the average percent of correct items. The last two columns give information on the Rasch model fit statistics (see 5.1.1). The first is

the average infit mean square statistic; the second is the average outfit mean square statistic. Optimally, these values should be close to 1.00.

The next section of Table G provides a summary of the findings of the DIF analyses (see 5.1.4). 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 tasks). The next major columns show the contrasting groups in the DIF analyses: either male versus female (Male/Female) or Hispanic versus other ethnicities (Hispanic/Other). Even though DIF may be negligible (category A or AA), this table shows the number of items that were favoring one group or the other at all levels of DIF. Optimally, even when items are all in category A or AA, there should be roughly an even number of items favoring each of the two groups to ensure that there is no systematic biasing test effect across items.

For the Writing tasks, the last part of this table shows the distribution of the raw scores on each task by total score category. (Recall that the total score for a task equals the sum of three feature scores, which are scored from 1 to 6, for a maximum total of 18; however, papers that are written in languages other than English or are totally incomprehensible may receive a score of 0, while papers that demonstrate the ability to copy or write a few words in English may be awarded a score of 1. The total score of 2 is impossible to achieve.)

## 5.2.10   Complete Item Analysis Table (Table H)

Table H presents results of the analyses of all of the items or tasks on the test form. The first column provides a descriptive name of the item or task. The item or task names vary slightly across domains and grade-level clusters, but they usually consist of characters that represent the domain (e.g., "R" for Reading), the grade-level cluster (e.g., "g91" for Grades 9–12), the tier (e.g., C, if applicable), the unique number in the item database (e.g., 3820), the WIDA Standard (e.g., "MA" for the Language of Mathematics), the language proficiency level targeted (e.g., "p3"), the thematic folder name (e.g., "Cafeteria"), and the test series (e.g., 302). Note that for Writing, "IT" stands for the "integrated" task, which requires more extensive writing and that integrates Model Performance Indicators for WIDA ELD Standards SI, LA, and SS. Also, note that for some Speaking and Kindergarten tasks, the naming system is a bit simpler, e.g., "1.S_A1_K_302", which contains the item order, domain, the folder, the proficiency level, the grade-level cluster, and the test series.

The second column in Table H presents the item difficulty in logits, while the third column indicates whether that item served as a common item (except for Kindergarten), anchoring the measurement scale to the results of the field test. For dichotomously scored items (Listening, Reading, and Speaking), the fourth column shows the p-value (percent of correct answers on that item or, in the case of Speaking, percent of students meeting the expectations of that task). The next two columns show the Rasch fit statistics for the item or task, while the following columns show the results of the two DIF analyses for that item or task. These last columns are interpreted just as in Table G.

Note that in previous years, many of the Speaking tasks had high outfit values. This was especially true for the easier tasks that appeared early in a folder. An investigation into the response patterns to the Speaking test revealed a number of cases where either the test was administered incorrectly or one or more ratings were wrongly recorded. As explained in 1.2.5, if a student cannot meet the expectations of a task in a folder, the remaining tasks in that folder are not administered, and are assigned a score of 0. However, we found many cases in which

Return to Summary of Assessment Records Claims, Actions, and Evidence

students received a score of 0 for one task in a folder, and a score of 1 for a later task in that same folder. As a result, it appears that some students who would be expected to meet the expectations of certain low-level tasks did not meet those expectations; the existence of these outliers would increase the outfit value. Because these patterns indicate that either the test administrator did not follow the administration procedures, or that one or more responses were incorrectly recorded, we removed these responses from the data set when analyzing fit for the Speaking test. Table 5.2.10 shows how many such cases were removed from the analysis for each cluster.

**Table 5.2.10**

Rate of Speaking responses removed from fit analysis S302

| Cluster | No. of responses | No. of responses removed | Percent of responses removed |
|---|---|---|---|
| 1-2 | 388,802 | 15,829 | 4.1% |
| 3-5 | 352,105 | 17,787 | 5.1% |
| 6-8 | 219,219 | 5,977 | 2.7% |
| 9-12 | 197,615 | 7,169 | 3.6% |

Removing these items from the analysis helped to lower the outfit value for many of the Speaking items. However, there are still some items with high outfit values. We continue to investigate potential sources for these high outfit values.

Note also that the Kindergarten test used a new format starting with Series 200 (2008-2009). It was equated to Series 103 through a separate study, reported on in MacGregor, Kenyon, Gibson, and Evans (2009). Thus, the column labeled "Anchored?" is not included in Table H for the Kindergarten test.

## 5.2.11 Complete Raw Score to Scale Score Conversion Chart (Table I)

The next table in this section, Table I, presents the raw score to scale score conversion table for the test form. The first column shows all possible raw scores. The next one to four columns show the corresponding scale score for each grade level in the grade-level cluster. Note that for Listening and Reading items on Tier A, these have been capped to the scale score that represents the Proficiency Level score of 4.0. On Tier B, these have been capped to the scale score representing the Proficiency Level score of 5.0.

The next column shows the *conditional* standard error (i.e., from the Rasch analysis) in the metric of the scale score. The last two columns show a lower bound (i.e., the scale score minus one standard error) and an upper bound (i.e., the scale score plus one standard error) around the scale score. In some cases the resulting lower bound is below 100, which has been set as the lowest score on the scale. In those cases, the lower bound has been set at 100.

As can be clearly seen from the table, on any dichotomously scored test form, standard errors are very large at the lowest and highest ends of the raw score scale. Because of this phenomenon and because the scale scores are combined to form composite scores, the top scale scores for the Listening and Reading forms were often adjusted for an end-of-scale effect on Tier C by allowing the top scale scores to increase only at the same rate as the preceding scale scores. If they were not adjusted, their effect in the composite scores might be excessive.

Thus, if the scale scores towards the high end of the raw score scale were increasing with each raw score by 9 scale points before the group of adjusted scores, then each of the adjusted scores would increase by only 9 scale points each. Because the lower and upper bounds were calculated

based on the original logit scores, these adjusted scores do not fall in the middle of the range; they fall toward the lower end of the range, but they always fall *within* the range. In other words, the adjusted scale score is a very possible observed score for that number of raw score points obtained.

Because on Tiers A and B the highest possible scores have been capped before the escalation of scale scores due to large standard errors at the highest end of the raw score scale inflates them, there has been no need to make any other adjustment to the scale scores for these tiers at the extreme high end of the raw score range. Since the point at which scale scores are capped depends on the proficiency level associated with the score, the caps take effect at lower scores for lower grades within a cluster. In this case the scores have been marked in Table I as capped, and the standard error, and low and high bound for the capped scale score, has been repeated in the final rows of the table. In addition, at the lower end of the raw score scale, scale scores are truncated when necessary so that the lowest scale score given is the scale score corresponding to a proficiency level score of 1.0. As with the adjusted scores, the standard error and the lower and upper bounds reported in Table I reflect the true scale score, not the truncated score.

## 5.2.12    Raw Score to Proficiency Level Score Conversion Table (Table J)

The final table, Table J, shows the interpretive Proficiency Level score associated with each raw score. (Note that in previous annual technical reports, some of this information was included in Table I; however, with the grade-level cut scores in effect, we have put this information in a separate table for ease of reading.) The first column in Table J shows the raw score. The remaining columns show the Proficiency Level score associated with each raw score/scale score for each grade in the grade-level cluster, along with the percentage of students in that grade who scored at that raw score/scale score/proficiency level score.

There are two things to note about this table. First, unlike scale scores, which are determined psychometrically and have a one-to-one correspondence to raw scores regardless of the grade level of the student, Proficiency Level scores are interpretations of the scale score. In Series 100 and 101, cut scores between proficiency levels were determined at the grade-level cluster level; thus, for example, in the 3–5 grade-level cluster, a given scale score was associated with the same Proficiency Level score for students in Grades 3, 4, and 5. Such a system, however, fails to take into account that older children can be expected to perform better on the test due to general cognitive growth over and above growth in English language proficiency. This effect can clearly be seen in Tables A and B, where average scores on any test form tend to rise, albeit slightly, by grade level. In order words, we would expect a fifth grader to perform better on the 3–5 grade-level cluster test form than a third grader at the same underlying level of English proficiency. To account for this effect, the WIDA Consortium adopted grade-level cut scores beginning with Series 102 so that, for any given raw score/scale score, the Proficiency Level score now associated with it differs according to the grade level of the student. (For details on how grade-level cut scores were determined, see Kenyon et al., 2013.) The effect of this for Table J is to require a separate column for each grade.

Second, because scale scores are capped on Listening and Reading for Tiers A and B at the scale score corresponding to the proficiency level score of 4.0 (for Tier A) and 5.0 (for Tier B), beginning with Series 102, this capped score is now dependent on the grade level (rather than

dependent on the grade-level cluster level). These differences in the cap are also shown in Table J on Tiers A and B for Listening and Reading.

For Kindergarten the Proficiency Level scores are provided based on both the Accountability cut scores and the Instructional cut scores.

# 6. Analyses of Test Forms: Results

Chapter 6 contains proprietary test information and is not publicly available. State educational agencies (SEAs) may request this information; please contact us at help@wida.us.

World-Class Instructional Design and Assessment

**Annual Technical Report for
ACCESS for ELLs English Language Proficiency Test,
Series 302, 2013-2014 Administration**

**Annual Technical Report No. 10
Volume 3 of 3: Analyses Across Tiers**

Prepared by:

Center for Applied Linguistics

CAL/WIDA Partnership Activities
Psychometrics/Research Team

May 30, 2015

*Table of Contents*

Volume 3

# 7. Analysis Across Tiers: Overview

## *7.1 Background*

### 7.1.1 Reliability of Composites

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

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

where

$k$ = number of components $j$

$w_j$ = weight of component $j$

$\sigma_j^2$ = variance of component $j$

$\sigma_c^2$ = variance of composite

$\rho_j$ = reliability coefficient of component $j$.

The data to compute the stratified Cronbach's alpha is provided in the appropriate tables in Chapter 8.

### 7.1.2 Accuracy and Consistency of Classification

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

In the approach of Livingston and Lewis (1995), the accuracy of a decision is the extent to which decisions made on the basis of the administered test (i.e., the observed scores) would agree with the decisions that would be made if each student could somehow be tested with all possible parallel forms of the assessments; that is, decisions based on the examinees' "true score." On the

Return to Summary of Assessment Records Claims, Actions, and Evidence

other hand, the consistency of a decision is the extent to which decisions made on the basis of the administered test would agree with the decisions that would be made if the students had taken a different but parallel form of the test. Thus, in every analysis of classification, two parallel analyses are made: accuracy (that is, vis-à-vis "true scores") and consistency (that is, vis-à-vis a second form).

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

True scores are, of course, unknown. 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, we modeled the distribution of the true scores and of scores on a parallel form. Overall accuracy and consistency indices are produced by comparing the percentage of students classified across all categories the same way by both the observed distribution and modeled distribution. These indices indicate the percent of all students who would be classified into the same language Proficiency Level by both the administered test and either the true score distribution (accuracy) or a parallel test (consistency). (Our tables also provide an estimate of Cohen's kappa statistic, which is a very conservative estimate of the overall classification since it corrects for chance.)

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

Finally, we look at what may be the most important set of indices, which are the indices at the cut points. That is, at every cut point, using the true score distribution (e.g., accuracy), we provide the percent of students who are consistently placed above and below the cut score, as well as those who are false positives and false negatives. For consistency, only the percent of students classified consistently above and below the cut score is calculated. Thus, for example, to evaluate the degree of confidence that one can have in a decision made based on the Overall Composite score as to whether students are being accurately classified into WIDA language proficiency level 5 ("bridging") or not, one can look at the accuracy index provided in the table for the cut score 4/5.

## *7.2    Descriptions*

### 7.2.1    Scale Score Information (Figure A and Table A)

Figure A and Table A relate to the ACCESS for ELLs *scale scores* that were achieved by students in the grade-level cluster. Figure A shows the distribution of the scale scores. The horizontal axis shows the full range of all scale scores observed for the grade-level cluster. To provide a full perspective, it extends somewhat below and above the range of observed scale scores. The vertical axis shows the number of students (count). Each bar shows how many

students were awarded each scale score. Note that for Listening and Reading, the effects of capping the scores for Tier A and Tier B can often be clearly detected in this figure.

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

- Number of students in the analyses (the number students who were not absent, invalid, refused, exempt, or in the wrong cluster)

- Minimum observed scale score

- Maximum observed scale score

- The mean (average) scale score

- The standard deviation (std. dev.) of the scale scores

## 7.2.2    Proficiency Level Information (Figure B and Table B)

Figure B and Table B provide information on the proficiency level distribution of the students in the grade-level cluster. Figure B shows the distribution of the proficiency levels. The horizontal axis shows the six WIDA proficiency levels. The vertical axis shows the percent of students. Each bar shows the percent of students who were placed into each language proficiency level.

Each row of Table B shows, by each grade in the cluster and by total for the cluster:

- The WIDA Proficiency Level designation (1 to 6)

- The number of students (count) whose performance on the test form placed them into that proficiency level in the domain being tested (the number students who were not absent, invalid, refused, exempt, or in the wrong cluster)

- The percent of students, out of the total number of students taking the form within a grade or within the total of students in the grade-level cluster, who were placed into that Proficiency Level in the domain being tested

For Kindergarten, this information is provided for scores based on both the Accountability cut scores and the Instructional cut scores.

## 7.2.3    Conditional Standard Error of Measurement (Table C and Figures C and D)

Table C and Figures C and D provide information across the three overlapping tier forms within a grade-level cluster and on the comparative conditional standard error of measurement. (Note that this information applies only to the domain scores; this information is not applicable to the composite scores.)

Table C presents information on the conditional standard error of measurement at the most important points at which decisions are made about students on the basis of performances on ACCESS for ELLs, the cut points between Language Proficiency levels. Because the cut points depend on the grade level, information is provided for each grade level within the cluster. The leftmost column shows the cut (e.g., 1/2, which is the cut score between level 1 and level 2). The next column shows the grade level. The next column shows the cut score in the scale score metric (e.g., 305). In the last column(s), the corresponding conditional standard error of

measurement is given for each cut score in the scale score metric. For Kindergarten, the SEMs are provided in separate tables for the accountability and instructional cut scores. For each of the other grade-level clusters, the SEMs for the cut scores are provided in one table for the Tiers (A, B, and C).

From this table it is possible to examine how well the different tiers are targeted for making decisions about students at the various cut scores. 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 conditional SEM of any Tier at the 1/2 cut point, and a relatively low one at the 2/3 cut point. At the other end, Tier C forms should optimally have the lowest conditional SEM at the 5/6 cut point, and also a relatively low one at the 4/5 cut point. Tier B should have low SEM in the mid range. Information from this table provides easily comparable information on how well the three Tier forms are targeted to provide the most accurate measure to place their intended examinees into the language proficiency levels that they target. (Note that because of the capping of scores on Tiers A and B, there is no information given for some of the cuts.)

Figure C shows the test characteristic curve across the entire test for Kindergarten and across the three tiers for the other grade-level clusters. It shows graphically how the tiers differ in difficulty. Tier A is represented by a dotted curve, Tier B by a light solid curve, and Tier C by a dark solid curve. Note that not all tiers have the same number of items. Thus, some curves for Listening and Reading in this figure may not end at the top horizontal line. Five vertical lines in the graphic indicate the cut scores at the highest grade in each cluster only.

Figure D compares the test information function across the entire test for Kindergarten and across the three tiers for the other grade-level clusters. This figure reflects the "SEM" columns in Table C. Again, Tier A is represented by a dotted curve, Tier B by a light solid curve, and Tier C by a dark solid curve. As in Figure C, the cut scores at the highest grade in each cluster are indicated by vertical lines. These lines make it easy to see which form measures most accurately at which cut score.

## 7.2.4    Reliability Information (Table D)

In order to produce accuracy and consistency of classification tables, it was necessary to produce a single reliability estimate across the three tiers. For the domains, this was a weighted reliability estimate (Cronbach's alpha). In other words, it is the average reliability weighted by the number of students who were administered that tier form. Thus, Table D, based on the information from Table F in Chapter 6, provides the number of students and the reliability estimate for each tier. The final column presents the weighted reliability, an estimate of the reliability of the scale scores across the tiers.

For the composite scores, Table D presents the data used to calculate an estimate of the reliability of the composite using stratified Cronbach's alpha (see Chapter 7.1.1). The first column shows the components forming the composite, the second column the weight of the composite in the total score, the third the variance of the scale scores, and the fourth the reliability of the composite. (Note that these are the weighted reliabilities across the tiers.) Unlike the weighted composite, which is an average, the stratified alpha reflects the fact that there are two or four measures being combined into one single measure. Thus, the reliability of the composite score will be higher than the reliability of any single subscore within the composite.

## 7.2.5    Accuracy and Consistency of Classification Tables (Table E)

Table E presents three rows of information related to the accuracy and consistency of placement into Proficiency Level categories based on WIDA ACCESS (see Chapter 7.1.2). With the adoption of grade-level cut scores with Series 102, placement within a Proficiency Level now depends on the grade level of the student. Therefore, we provide a separate table for each grade level in a grade-level cluster. The first row provides overall indices related to the accuracy and consistency of classification, as well as Cohen's kappa. The second row of information shows accuracy and consistency information conditional on level. The third provides indices of classification accuracy and consistency at the cut points. These indices are perhaps the most important of all when using any of these as an absolute cut-point (i.e., asking the question which students have reached level six and which have not). Note that the consistency is generally higher at the cut points than over the levels. For practical purposes, the primary score used for such decisions are the Overall Composite scores.

Note that because of the scoring caps now imposed on Tier A and Tier B in Listening and Reading, in several cases only a very small percentage of test takers get placed into Proficiency Level 6. This outcome, combined with the range of observed scale scores, (which may be very close to the 5/6 cut), and the reliability of the test, means that the accuracy conditional on level for level 6 cannot be estimated. In such cases a hyphen (-) has been placed in the table. For Writing, this result can also occur for both levels 5 and 6.

For Kindergarten, these tables are provided for both the Accountability cut scores and the Instructional cut scores.

# 8. Analysis Across Tiers: Results

## 8.1　Grade: K

### 8.1.1　Listening K

**Figure 8.1.1A**
**Scale Scores: List K S302**
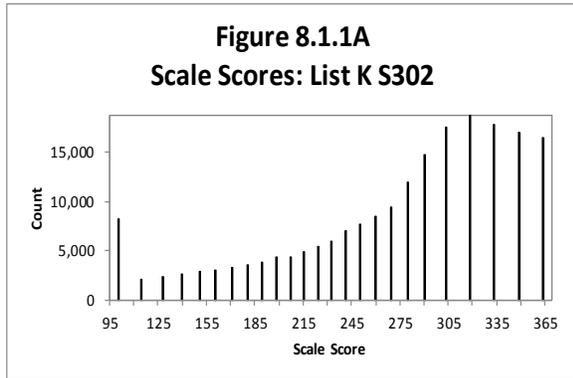


**Table 8.1.1A**

Scale Score Descriptive Statistics: List K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,841 | 100 | 363 | 269.99 | 70.70 |

**Figure 8.1.1Bi**
**Proficiency Level: List K S302**
**(Accountability)**
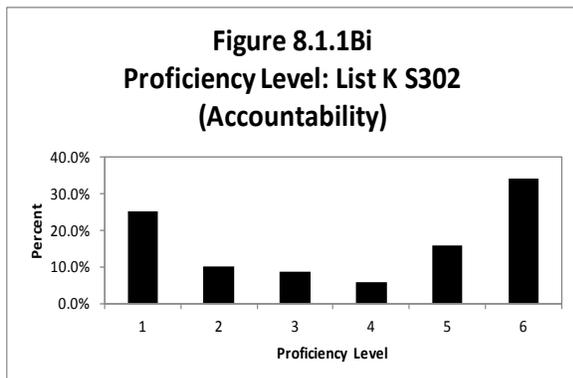


**Table 8.1.1Bi**

Proficiency Level Distribution: List K S302
(Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 51,112 | 25.1% |
| 2 | 20,746 | 10.2% |
| 3 | 17,932 | 8.8% |
| 4 | 11,923 | 5.8% |
| 5 | 32,179 | 15.8% |
| 6 | 69,949 | 34.3% |
| Total | 203,841 | 100.0% |

**Figure 8.1.1Bii**
**Proficiency Level: List K S302**
**(Instructional)**
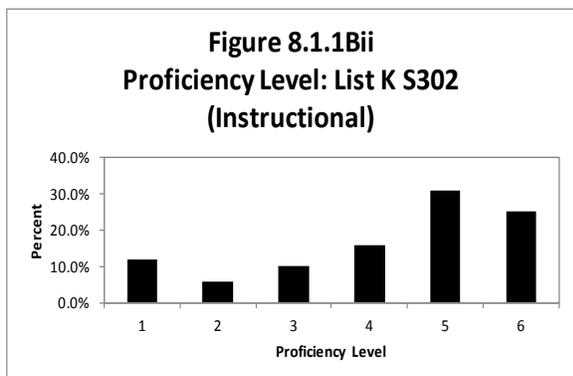


**Table 8.1.1Bii**

Proficiency Level Distribution: List K S302
(Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 24,626 | 12.1% |
| K2 | 11,794 | 5.8% |
| K3 | 20,693 | 10.2% |
| K4 | 32,677 | 16.0% |
| K5 | 62,805 | 30.8% |
| K6 | 51,246 | 25.1% |
| Total | 203,841 | 100.0% |

**Table 8.1.1Ci**
Conditional Standard Error of
Measurement at Cut Scores: List K
S302 (Accountability)

| Proficiency Level | 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 |

**Table 8.1.1Cii**
Conditional Standard Error of
Measurement at Cut Scores: List K
S302 (Instructional)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 175 | 17.28 |
| 2/3 | 204 | 16.91 |
| 3/4 | 240 | 17.66 |
| 4/5 | 279 | 20.66 |
| 5/6 | 322 | 27.43 |

# Figure 8.1.1D
## Test Characteristic Curve: List K S302
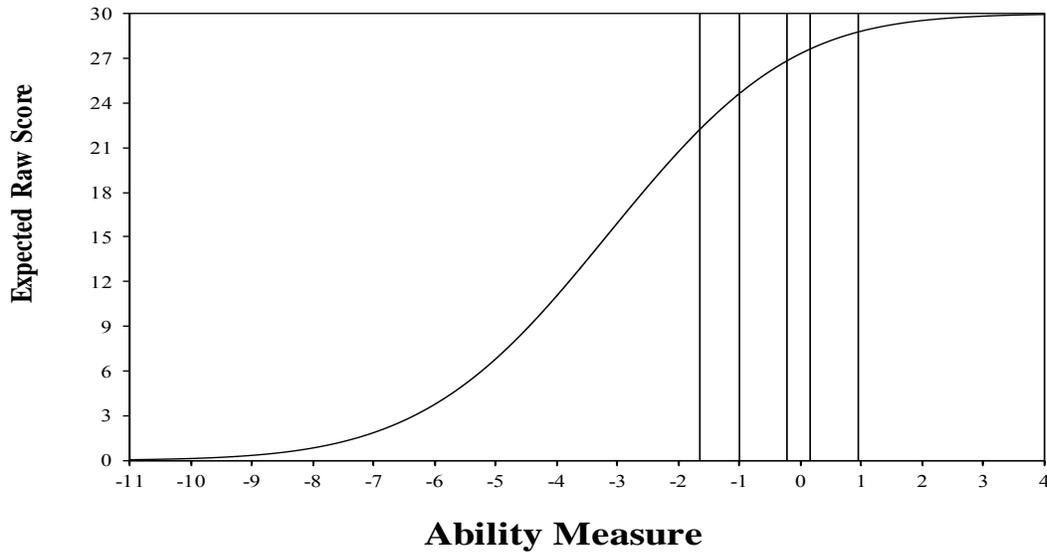


# Figure 8.1.1D
## Test Information Function: List K S302
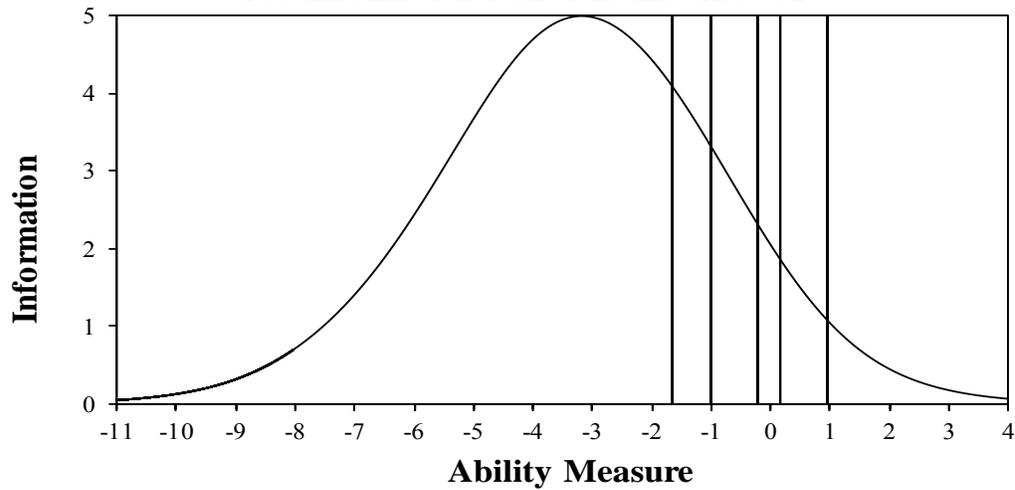
**Table 8.1.1D**

Reliability: List K S302

| Tiers | No. of Students | Reliability |
|-------|-----------------|-------------|
| - | 203,841 | 0.934 |

**Table 8.1.1E**

Accuracy and Consistency of Classification Indices: List  (Grade K) S302 (Accountability)

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| Indices | 0.679 | 0.614 | | 0.495 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.872 | | 0.817 | |
| | 2 | 0.469 | | 0.351 | |
| | 3 | 0.327 | | 0.247 | |
| | 4 | 0.211 | | 0.155 | |
| | 5 | 0.474 | | 0.363 | |
| | 6 | 0.824 | | 0.770 | |

| Indices at Cut Points | | Accuracy | | | |
|------------------------|-----------|----------|-----------------|-----------------|-------------|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.941 | 0.033 | 0.026 | 0.918 |
| | 2/3 | 0.930 | 0.027 | 0.043 | 0.902 |
| | 3/4 | 0.916 | 0.049 | 0.035 | 0.883 |
| | 4/5 | 0.907 | 0.045 | 0.048 | 0.874 |
| | 5/6 | 0.899 | 0.035 | 0.066 | 0.860 |

**Table 8.1.1E**

Accuracy and Consistency of Classification Indices: List  (Grade K) S302 (Instructional)

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| Indices | 0.677 | 0.574 | | 0.462 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.886 | | 0.811 | |
| | 2 | 0.442 | | 0.325 | |
| | 3 | 0.526 | | 0.401 | |
| | 4 | 0.566 | | 0.446 | |
| | 5 | 0.702 | | 0.580 | |
| | 6 | 0.734 | | 0.656 | |

| Indices at Cut Points | | Accuracy | | | |
|------------------------|-----------|----------|-----------------|-----------------|-------------|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.969 | 0.013 | 0.018 | 0.955 |
| | 2/3 | 0.959 | 0.020 | 0.021 | 0.940 |
| | 3/4 | 0.940 | 0.030 | 0.030 | 0.914 |
| | 4/5 | 0.916 | 0.042 | 0.042 | 0.882 |
| | 5/6 | 0.884 | 0.040 | 0.077 | 0.838 |

## 8.1.2    Reading K

**Figure 8.1.2A**
**Scale Scores: Read K S302**
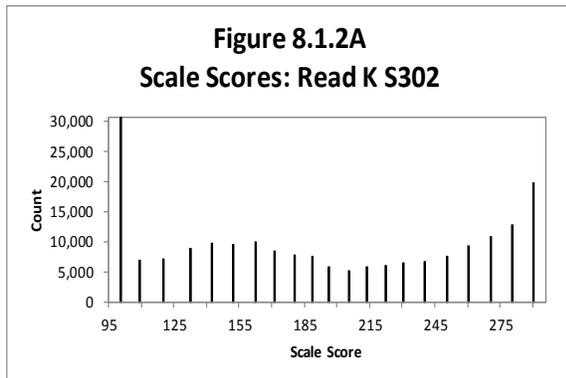


**Table 8.1.2A**
Scale Score Descriptive Statistics: Read K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,853 | 100 | 290 | 192.18 | 66.09 |

**Figure 8.1.2Bi**
**Proficiency Level: Read K S302**
**(Accountability)**
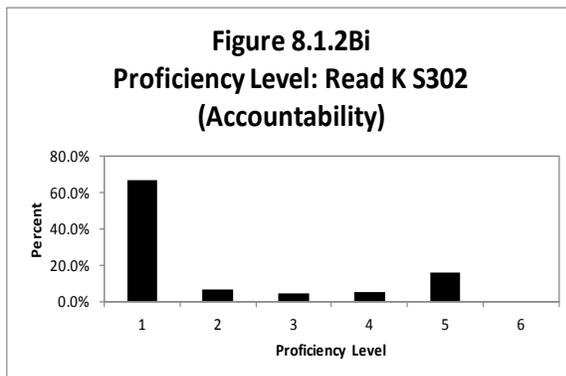


**Table 8.1.2Bi**
Proficiency Level Distribution: Read K S302
(Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 136,612 | 67.0% |
| 2 | 14,186 | 7.0% |
| 3 | 9,412 | 4.6% |
| 4 | 10,849 | 5.3% |
| 5 | 32,794 | 16.1% |
| 6 | 0 | 0.0% |
| Total | 203,853 | 100.0% |

**Figure 8.1.2Bii**
**Proficiency Level: Read K S302**
**(Instructional)**
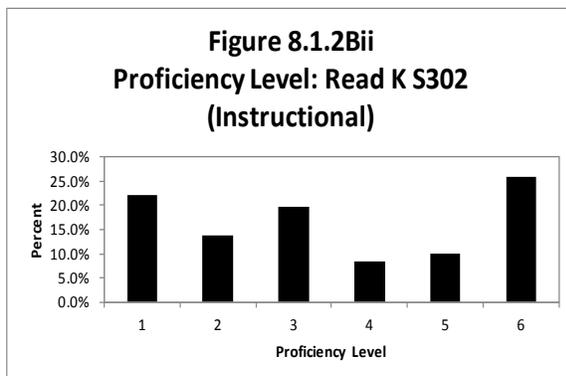


**Table 8.1.2Bii**
Proficiency Level Distribution: Read K S302
(Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 44,927 | 22.0% |
| K2 | 28,198 | 13.8% |
| K3 | 39,920 | 19.6% |
| K4 | 17,097 | 8.4% |
| K5 | 20,656 | 10.1% |
| K6 | 53,055 | 26.0% |
| Total | 203,853 | 100.0% |

**Table 8.1.2Ci**
Conditional Standard Error of Measurement
at Cut Scores: Read K S302 (Accountability)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 238 | 15.08 |
| 2/3 | 251 | 16.90 |
| 3/4 | 261 | 18.98 |
| 4/5 | 274 | 22.10 |
| 5/6 | 295 | 30.68 |

**Table 8.1.2Cii**
Conditional Standard Error of Measurement
at Cut Scores: Read K S302 (Instructional)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 121 | 14.04 |
| 2/3 | 159 | 13.52 |
| 3/4 | 204 | 13.00 |
| 4/5 | 228 | 14.04 |
| 5/6 | 255 | 17.68 |

**Figure 8.1.2C**
Test Characteristic Curve: Read K S302
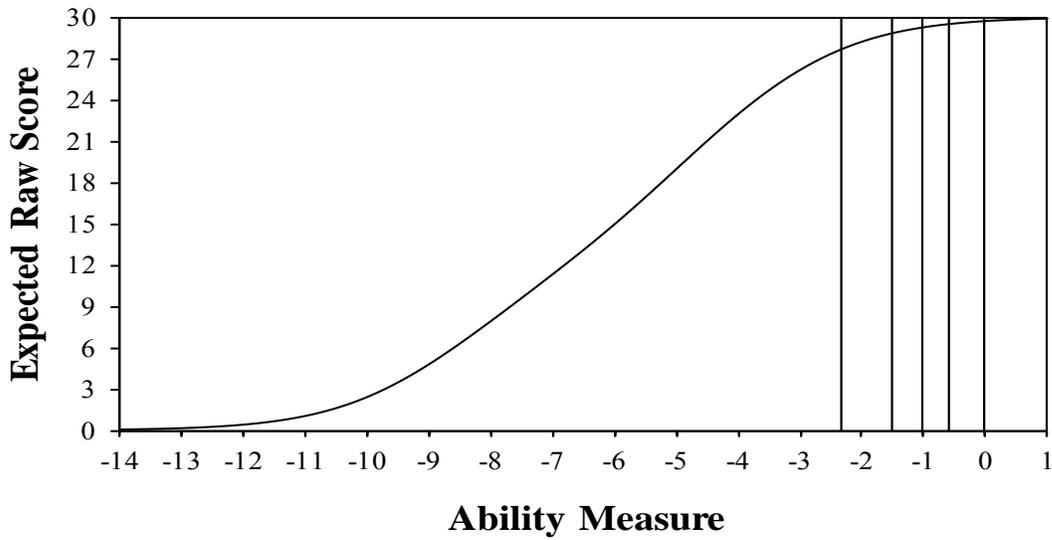


**Figure 8.1.2D**
Test Information Function: Read K S302
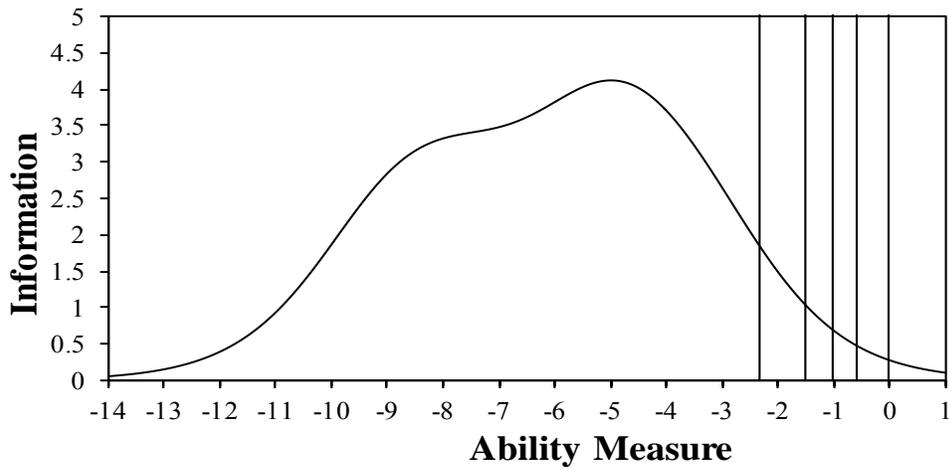
**Table 8.1.2D**

Reliability: Read K S302

| Tiers | No. of Students | Reliability |
|-------|-----------------|-------------|
| - | 203,853 | 0.947 |

**Table 8.1.2E**

Accuracy and Consistency of Classification Indices: Read (Grade K) S302 (Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|---|---|---|---|---|---|---|
| | 0.821 | 0.787 | | 0.583 | | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | | |
| | 1 | 0.943 | | 0.929 | | |
| | 2 | 0.337 | | 0.252 | | |
| | 3 | 0.240 | | 0.176 | | |
| | 4 | 0.293 | | 0.213 | | |
| | 5 | 0.872 | | 0.772 | | |
| **Indices at Cut Points** | | **Accuracy** | | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** | |
| | 1/2 | 0.934 | 0.039 | 0.027 | 0.910 | |
| | 2/3 | 0.937 | 0.032 | 0.031 | 0.913 | |
| | 3/4 | 0.943 | 0.029 | 0.028 | 0.919 | |
| | 4/5 | 0.948 | 0.033 | 0.019 | 0.926 | |

**Table 8.1.2E**

Accuracy and Consistency of Classification Indices: Read (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|---|---|---|---|---|---|---|
| | 0.771 | 0.699 | | 0.603 | | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | | |
| | 1 | 0.902 | | 0.835 | | |
| | 2 | 0.574 | | 0.459 | | |
| | 3 | 0.700 | | 0.589 | | |
| | 4 | 0.388 | | 0.289 | | |
| | 5 | 0.922 | | 0.881 | | |
| **Indices at Cut Points** | | **Accuracy** | | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** | |
| | 1/2 | 0.944 | 0.020 | 0.036 | 0.922 | |
| | 2/3 | 0.940 | 0.033 | 0.026 | 0.915 | |
| | 3/4 | 0.936 | 0.029 | 0.034 | 0.910 | |
| | 4/5 | 0.940 | 0.032 | 0.028 | 0.914 | |

## 8.1.3 Writing K

**Figure 8.1.3A**
**Scale Scores: Writ K S302**
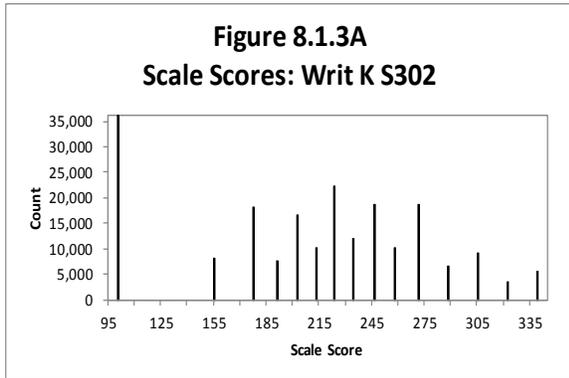


**Table 8.1.3A**
Scale Score Descriptive Statistics: Writ K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,840 | 100 | 339 | 210.63 | 65.58 |

**Figure 8.1.3Bi**
**Proficiency Level: Writ K S302**
**(Accountability)**
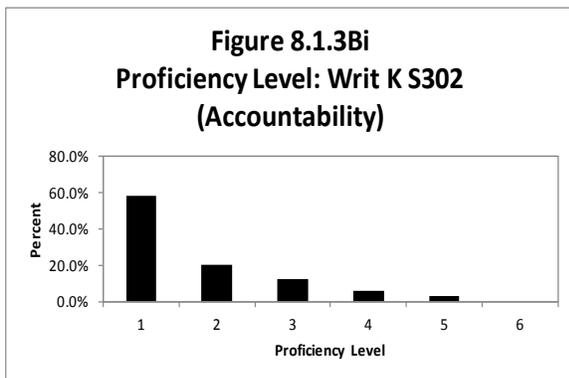


**Table 8.1.3Bi**
Proficiency Level Distribution: Writ K S302 (Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 119,172 | 58.5% |
| 2 | 40,889 | 20.1% |
| 3 | 25,262 | 12.4% |
| 4 | 12,868 | 6.3% |
| 5 | 5,649 | 2.8% |
| 6 | 0 | 0.0% |
| Total | 203,840 | 100.0% |

**Figure 8.1.3Bii**
**Proficiency Level: Writ K S302**
**(Instructional)**
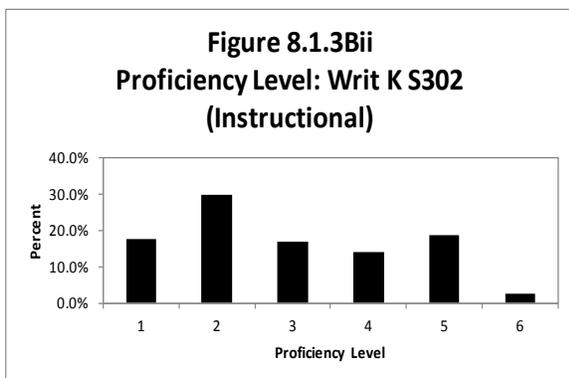


**Table 8.1.3Bii**
Proficiency Level Distribution: Writ K S302 (Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 36,149 | 17.7% |
| K2 | 60,841 | 29.8% |
| K3 | 34,266 | 16.8% |
| K4 | 28,805 | 14.1% |
| K5 | 38,130 | 18.7% |
| K6 | 5,649 | 2.8% |
| Total | 203,840 | 100.0% |

**Table 8.1.3Ci**
Conditional Standard Error of
Measurement at Cut Scores: Writ K
S302 (Accountability)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 225 | 18.35 |
| 2/3 | 259 | 19.90 |
| 3/4 | 295 | 26.43 |
| 4/5 | 323 | 33.90 |
| 5/6 | 350 | 38.87 |

**Table 8.1.3Cii**
Conditional Standard Error of
Measurement at Cut Scores: Writ K
S302 (Instructional)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 145 | 31.10 |
| 2/3 | 218 | 18.04 |
| 3/4 | 244 | 19.28 |
| 4/5 | 269 | 20.83 |
| 5/6 | 326 | 34.52 |

**Figure 8.1.3C**
Test Characteristic Curve: Writ K S302
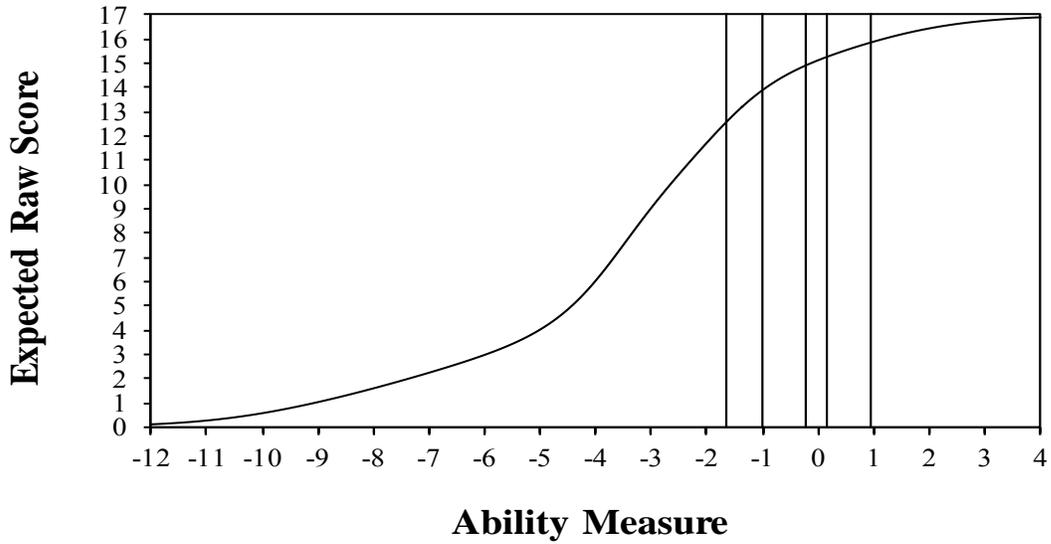


**Figure 8.1.3D**
Test Information Function: Writ K S302
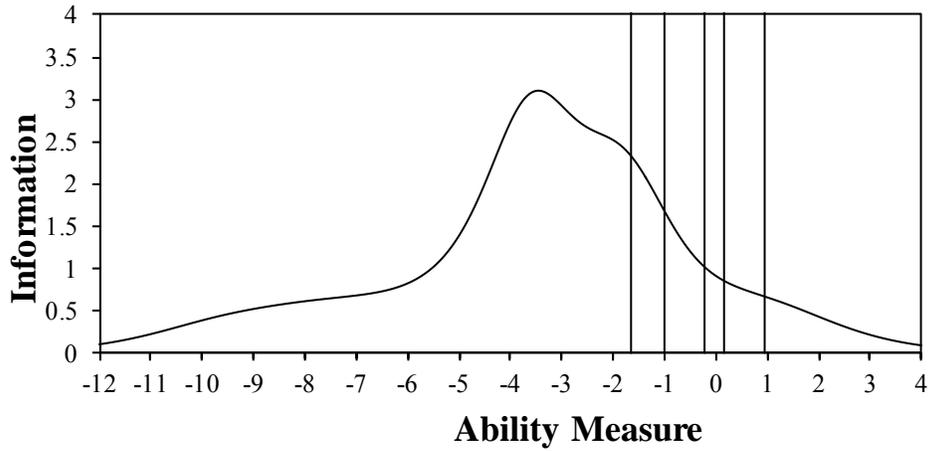
**Table 8.1.3D**

Reliability: Writ K S302

| Tiers | No. of Students | Reliability |
|---|---|---|
| - | 203,840 | 0.922 |

**Table 8.1.3E**

Accuracy and Consistency of Classification Indices: Writ (Grade K) S302 (Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.740 | 0.689 | | 0.482 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.941 | | 0.914 | |
| | 2 | 0.610 | | 0.465 | |
| | 3 | 0.386 | | 0.344 | |
| | 4 | - | | 0.259 | |
| | 5 | - | | 0.139 | |
| **Indices at Cut Points** | | | **Accuracy** | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.922 | 0.034 | 0.044 | 0.893 |
| | 2/3 | 0.902 | 0.021 | 0.077 | 0.863 |
| | 3/4 | 0.909 | 0.091 | 0.000 | 0.898 |
| | 4/5 | 0.972 | 0.028 | 0.000 | 0.971 |

**Table 8.1.3E**

Accuracy and Consistency of Classification Indices: Writ (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.680 | 0.588 | | 0.476 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.865 | | 0.792 | |
| | 2 | 0.791 | | 0.707 | |
| | 3 | 0.515 | | 0.388 | |
| | 4 | 0.365 | | 0.281 | |
| | 5 | 0.245 | | 0.603 | |
| **Indices at Cut Points** | | | **Accuracy** | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.955 | 0.025 | 0.020 | 0.935 |
| | 2/3 | 0.916 | 0.041 | 0.043 | 0.885 |
| | 3/4 | 0.898 | 0.031 | 0.071 | 0.860 |
| | 4/5 | 0.879 | 0.045 | 0.077 | 0.832 |

## 8.1.4   Speaking K



**Figure 8.1.4A**
**Scale Scores: Spek K S302**
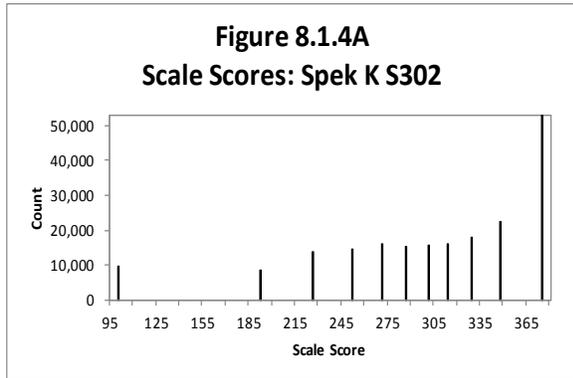
**Table 8.1.4A**
Scale Score Descriptive Statistics: Spek K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,827 | 100 | 375 | 302.51 | 69.67 |



**Figure 8.1.4Bi**
**Proficiency Level: Spek K S302**
**(Accountability)**
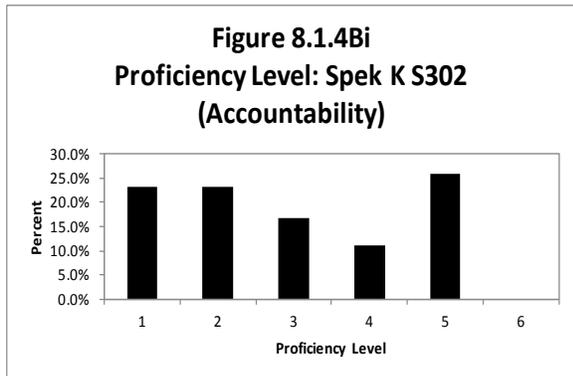
**Table 8.1.4Bi**
Proficiency Level Distribution: Spek K S302
(Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 47,008 | 23.1% |
| 2 | 47,168 | 23.1% |
| 3 | 34,064 | 16.7% |
| 4 | 22,635 | 11.1% |
| 5 | 52,952 | 26.0% |
| 6 | 0 | 0.0% |
| Total | 203,827 | 100.0% |



**Figure 8.1.4Bii**
**Proficiency Level: Spek K S302**
**(Instructional)**
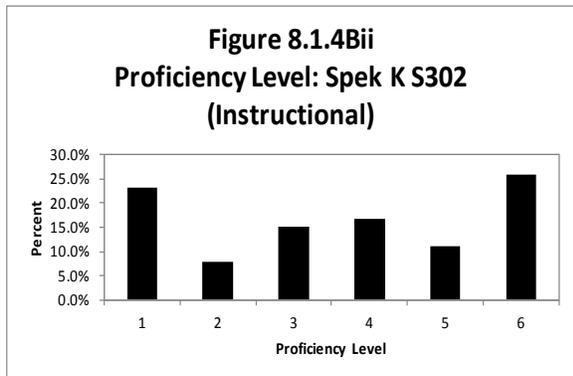
**Table 8.1.4Bii**
Proficiency Level Distribution: Spek K S302
(Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 47,008 | 23.1% |
| K2 | 16,092 | 7.9% |
| K3 | 31,076 | 15.2% |
| K4 | 34,064 | 16.7% |
| K5 | 22,635 | 11.1% |
| K6 | 52,952 | 26.0% |
| Total | 203,827 | 100.0% |

**Table 8.1.4Ci**
Conditional Standard Error of
Measurement at Cut Scores: Spek K
S302 (Accountability)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 269 | 18.68 |
| 2/3 | 314 | 16.27 |
| 3/4 | 343 | 20.89 |
| 4/5 | 366 | 31.33 |
| 5/6 | 383 | 44.99 |

**Table 8.1.4Cii**
Conditional Standard Error of
Measurement at Cut Scores: Spek K
S302 (Instructional)

| Proficiency Level | Cut Score | SEM |
|---|---|---|
| 1/2 | 256 | 20.89 |
| 2/3 | 285 | 17.07 |
| 3/4 | 308 | 16.27 |
| 4/5 | 342 | 20.49 |
| 5/6 | 365 | 30.53 |

**Figure 8.1.4C**
Test Characteristic Curve: Spek K S302



**Figure 8.1.4D**
Test Information Function: Spek K S302
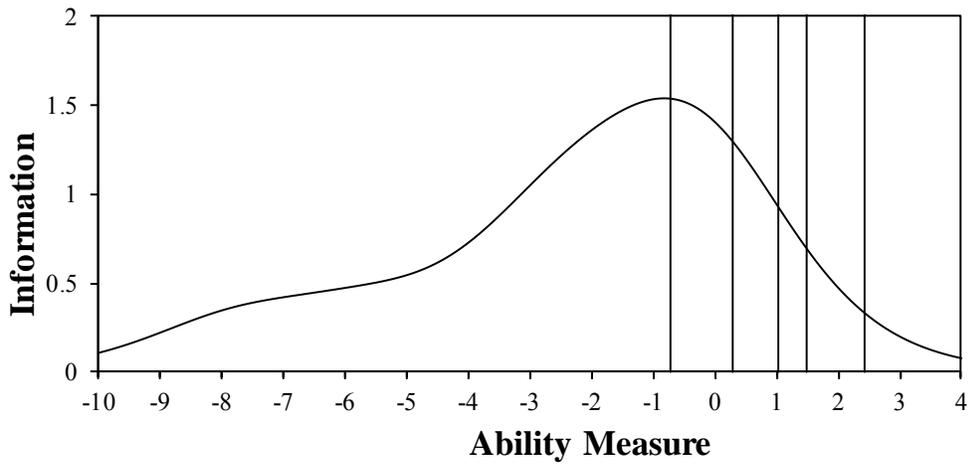
**Table 8.1.4D**

Reliability: Spek K S302

| Tiers | No. of Students | Reliability |
|-------|-----------------|-------------|
| - | 203,827 | 0.894 |

**Table 8.1.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade K) S302 (Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|-----------------|----------|-------------|--|-----------|--|--|
| | 0.468 | 0.451 | | 0.321 | | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | | |
| | 1 | 0.830 | | 0.760 | | |
| | 2 | 0.662 | | 0.533 | | |
| | 3 | 0.377 | | 0.260 | | |
| | 4 | 0.212 | | 0.194 | | |
| | 5 | - | | 0.563 | | |
| **Indices at Cut Points** | | **Accuracy** | | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** | |
| | 1/2 | 0.928 | 0.041 | 0.031 | 0.900 | |
| | 2/3 | 0.891 | 0.032 | 0.077 | 0.850 | |
| | 3/4 | 0.865 | 0.046 | 0.089 | 0.786 | |
| | 4/5 | 0.740 | 0.260 | 0.000 | 0.755 | |

**Table 8.1.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|-----------------|----------|-------------|--|-----------|--|--|
| | 0.652 | 0.563 | | 0.419 | | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | | |
| | 1 | 0.871 | | 0.797 | | |
| | 2 | 0.312 | | 0.234 | | |
| | 3 | 0.474 | | 0.357 | | |
| | 4 | 0.360 | | 0.264 | | |
| | 5 | 0.794 | | 0.721 | | |
| **Indices at Cut Points** | | **Accuracy** | | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** | |
| | 1/2 | 0.938 | 0.029 | 0.033 | 0.910 | |
| | 2/3 | 0.914 | 0.047 | 0.039 | 0.883 | |
| | 3/4 | 0.888 | 0.031 | 0.081 | 0.850 | |
| | 4/5 | 0.859 | 0.061 | 0.080 | 0.790 | |

## 8.1.5    Oral Language Composite K

**Figure 8.1.5A**
**Scale Scores: Oral K S302**
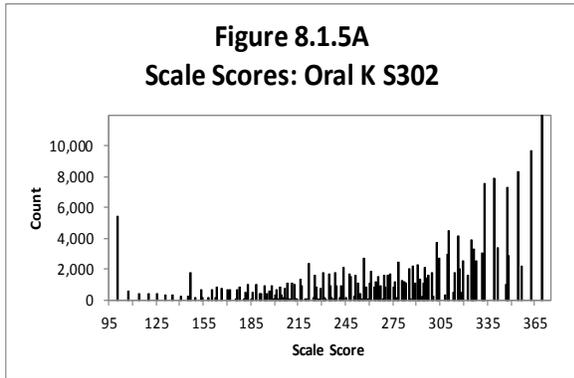


**Table 8.1.5A**
Scale Score Descriptive Statistics: Oral K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,823 | 100 | 369 | 286.48 | 66.28 |

**Figure 8.1.5Bi**
**Proficiency Level: Oral K S302**
**(Accountability)**
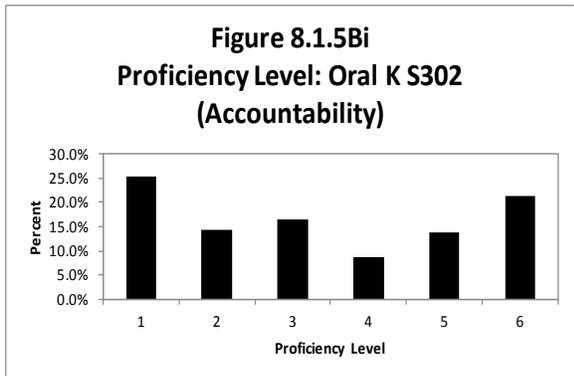


**Table 8.1.5Bi**
Proficiency Level Distribution: Oral K S302
(Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 51,598 | 25.3% |
| 2 | 29,375 | 14.4% |
| 3 | 33,562 | 16.5% |
| 4 | 17,908 | 8.8% |
| 5 | 28,069 | 13.8% |
| 6 | 43,311 | 21.2% |
| Total | 203,823 | 100.0% |

**Figure 8.1.5Bii**
**Proficiency Level: Oral K S302**
**(Instructional)**
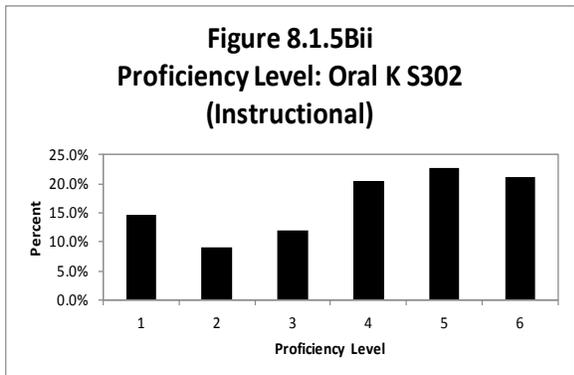


**Table 8.1.5Bii**
Proficiency Level Distribution: Oral K S302
(Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 29,747 | 14.6% |
| K2 | 18,588 | 9.1% |
| K3 | 24,444 | 12.0% |
| K4 | 41,756 | 20.5% |
| K5 | 45,977 | 22.6% |
| K6 | 43,311 | 21.2% |
| Total | 203,823 | 100.0% |

**Table 8.1.5C**

n/a

**Figure 8.1.5C**

n/a

**Figure 8.1.5D**

n/a

**Table 8.1.5D**

Oral Composite Reliability: Oral K S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.50 | 4997.123 | 0.934 |
| Speaking | 0.50 | 4852.854 | 0.894 |
| Oral | | 4393.165 | 0.952 |

**Table 8.1.5E**

Accuracy and Consistency of Classification Indices: Oral (Grade K) S302 (Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.628 | 0.546 | | 0.447 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.908 | | 0.862 | |
| | 2 | 0.633 | | 0.514 | |
| | 3 | 0.609 | | 0.488 | |
| | 4 | 0.338 | | 0.226 | |
| | 5 | 0.359 | | 0.297 | |
| | 6 | 0.728 | | 0.625 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.955 | 0.023 | 0.021 | 0.936 |
| | 2/3 | 0.935 | 0.030 | 0.035 | 0.910 |
| | 3/4 | 0.927 | 0.025 | 0.047 | 0.899 |
| | 4/5 | 0.924 | 0.027 | 0.049 | 0.881 |
| | 5/6 | 0.857 | 0.101 | 0.042 | 0.827 |

**Table 8.1.5E**

Accuracy and Consistency of Classification Indices: Oral (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.660 | 0.559 | | 0.459 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.883 | | 0.821 | |
| | 2 | 0.578 | | 0.451 | |
| | 3 | 0.560 | | 0.443 | |
| | 4 | 0.695 | | 0.575 | |
| | 5 | 0.546 | | 0.447 | |
| | 6 | 0.708 | | 0.615 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.968 | 0.017 | 0.015 | 0.954 |
| | 2/3 | 0.951 | 0.023 | 0.026 | 0.931 |
| | 3/4 | 0.938 | 0.028 | 0.034 | 0.914 |
| | 4/5 | 0.929 | 0.024 | 0.047 | 0.899 |
| | 5/6 | 0.868 | 0.075 | 0.057 | 0.829 |

## 8.1.6    Literacy Composite K

**Figure 8.1.6A**
**Scale Scores: Litr K S302**
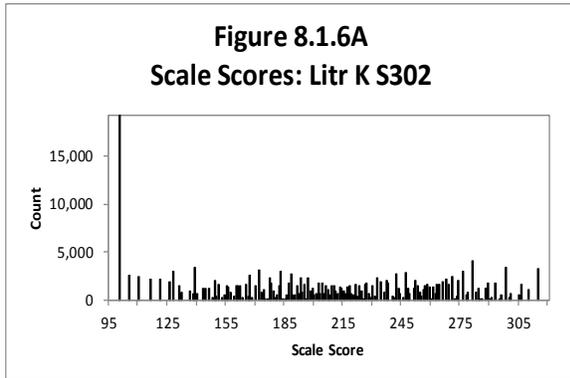


**Table 8.1.6A**
Scale Score Descriptive Statistics: Litr K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,837 | 100 | 315 | 201.65 | 61.11 |

**Figure 8.1.6Bi**
**Proficiency Level: Litr K S302**
**(Accountability)**
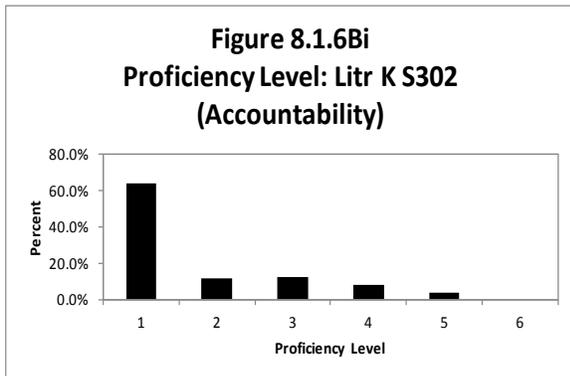


**Table 8.1.6Bi**
Proficiency Level Distribution: Litr K S302
(Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 130,672 | 64.1% |
| 2 | 24,411 | 12.0% |
| 3 | 24,653 | 12.1% |
| 4 | 16,739 | 8.2% |
| 5 | 7,362 | 3.6% |
| 6 | 0 | 0.0% |
| Total | 203,837 | 100.0% |

**Figure 8.1.6Bii**
**Proficiency Level: Litr K S302**
**(Instructional)**
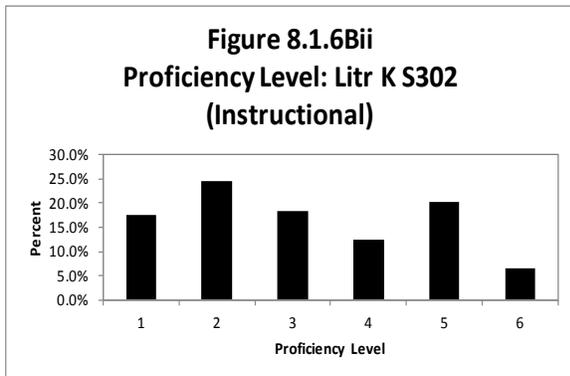


**Table 8.1.6Bii**
Proficiency Level Distribution: Litr K S302
(Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 35,756 | 17.5% |
| K2 | 50,231 | 24.6% |
| K3 | 37,594 | 18.4% |
| K4 | 25,398 | 12.5% |
| K5 | 41,480 | 20.3% |
| K6 | 13,378 | 6.6% |
| Total | 203,837 | 100.0% |

**Table 8.1.6C**
n/a

**Figure 8.1.6C**
n/a

**Figure 8.1.6D**
n/a

**Table 8.1.6D**
Literacy Composite Reliability: Litr K S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Reading | 0.50 | 4368.024 | 0.947 |
| Writing | 0.50 | 4299.672 | 0.922 |
| Literacy | | 3733.607 | 0.962 |

**Table 8.1.6E**
Accuracy and Consistency of Classification Indices: Litr (Grade K) S302
(Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.797 | 0.747 | | 0.542 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.961 | | 0.943 | |
| | 2 | 0.562 | | 0.434 | |
| | 3 | 0.513 | | 0.399 | |
| | 4 | 0.452 | | 0.390 | |
| | 5 | - | | 0.264 | |
| | 6 | - | | 0.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.947 | 0.025 | 0.028 | 0.925 |
| | 2/3 | 0.942 | 0.024 | 0.034 | 0.918 |
| | 3/4 | 0.933 | 0.034 | 0.033 | 0.906 |
| | 4/5 | 0.964 | 0.036 | 0.000 | 0.955 |
| | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

**Table 8.1.6E**

Accuracy and Consistency of Classification Indices: Litr (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.747 | 0.671 | | 0.594 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.917 | | 0.870 | |
| | 2 | 0.835 | | 0.769 | |
| | 3 | 0.730 | | 0.624 | |
| | 4 | 0.575 | | 0.450 | |
| | 5 | 0.650 | | 0.619 | |
| | 6 | - | | 0.410 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.969 | 0.014 | 0.017 | 0.956 |
| | 2/3 | 0.952 | 0.025 | 0.024 | 0.932 |
| | 3/4 | 0.944 | 0.024 | 0.032 | 0.922 |
| | 4/5 | 0.946 | 0.023 | 0.032 | 0.922 |
| | 5/6 | 0.934 | 0.066 | 0.000 | 0.926 |

## 8.1.7    Comprehension Composite K

### Figure 8.1.7A
### Scale Scores: Cphn K S302
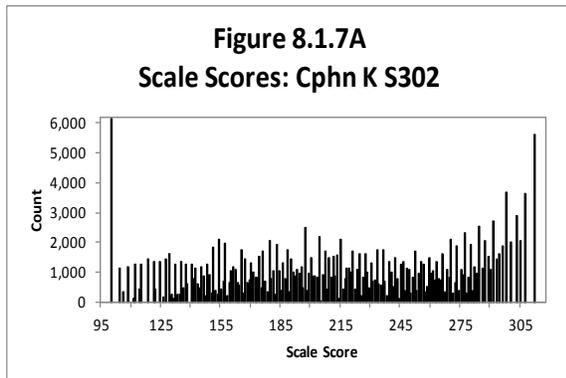


**Table 8.1.7A**

Scale Score Descriptive Statistics: Cphn K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,837 | 100 | 312 | 215.52 | 60.34 |

### Figure 8.1.7Bi
### Proficiency Level: Cphn K S302
### (Accountability)
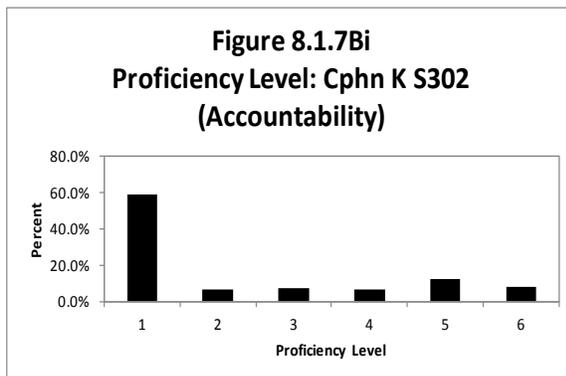


**Table 8.1.7Bi**

Proficiency Level Distribution: Cphn K S302 (Accountability)

| Level | Count | Percent |
|---|---|---|
| 1 | 120,340 | 59.0% |
| 2 | 14,043 | 6.9% |
| 3 | 14,430 | 7.1% |
| 4 | 13,518 | 6.6% |
| 5 | 25,262 | 12.4% |
| 6 | 16,244 | 8.0% |
| Total | 203,837 | 100.0% |

### Figure 8.1.7Bii
### Proficiency Level: Cphn K S302
### (Instructional)
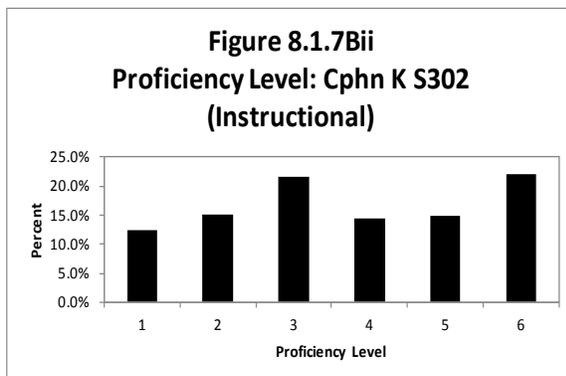


**Table 8.1.7Bii**

Proficiency Level Distribution: Cphn K S302 (Instructional)

| Level | Count | Percent |
|---|---|---|
| K1 | 25,081 | 12.3% |
| K2 | 30,698 | 15.1% |
| K3 | 43,899 | 21.5% |
| K4 | 29,078 | 14.3% |
| K5 | 30,330 | 14.9% |
| K6 | 44,751 | 22.0% |
| Total | 203,837 | 100.0% |

**Table 8.1.7C**
n/a

**Figure 8.1.7C**
n/a

**Figure 8.1.7D**
n/a

**Table 8.1.7D**
Comprehension Composite Reliability: Cphn K S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.30 | 4997.123 | 0.934 |
| Reading | 0.70 | 4368.024 | 0.947 |
| Comprehension | | 3641.051 | 0.961 |

**Table 8.1.7E-1**
Accuracy and Consistency of Classification Indices: Cphn (Grade K) S302
(Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.763 | 0.708 | | 0.529 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.961 | | 0.943 | |
| | 2 | 0.391 | | 0.285 | |
| | 3 | 0.392 | | 0.285 | |
| | 4 | 0.350 | | 0.257 | |
| | 5 | 0.560 | | 0.462 | |
| | 6 | 0.671 | | 0.531 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.949 | 0.023 | 0.028 | 0.928 |
| | 2/3 | 0.950 | 0.025 | 0.024 | 0.928 |
| | 3/4 | 0.946 | 0.029 | 0.025 | 0.923 |
| | 4/5 | 0.941 | 0.031 | 0.028 | 0.918 |
| | 5/6 | 0.945 | 0.032 | 0.023 | 0.925 |

**Table 8.1.7E-1**

Accuracy and Consistency of Classification Indices: Cphn (Grade K) S302 (Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.767 | 0.678 | | 0.610 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.894 | | 0.830 | |
| | 2 | 0.757 | | 0.659 | |
| | 3 | 0.775 | | 0.686 | |
| | 4 | 0.635 | | 0.514 | |
| | 5 | 0.631 | | 0.511 | |
| | 6 | 0.884 | | 0.832 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.972 | 0.013 | 0.016 | 0.960 |
| | 2/3 | 0.955 | 0.021 | 0.023 | 0.937 |
| | 3/4 | 0.946 | 0.024 | 0.030 | 0.924 |
| | 4/5 | 0.949 | 0.024 | 0.028 | 0.927 |
| | 5/6 | 0.944 | 0.032 | 0.025 | 0.920 |

## 8.1.8    Overall Composite K

**Figure 8.1.8A**
**Scale Scores: Over K S302**
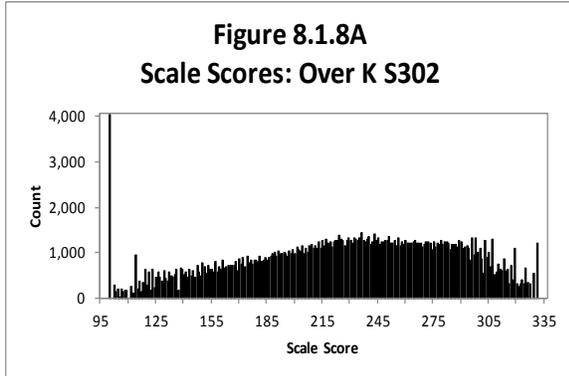


**Table 8.1.8A**
Scale Score Descriptive Statistics: Over K S302

| No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|
| 203,809 | 100 | 331 | 226.90 | 57.10 |
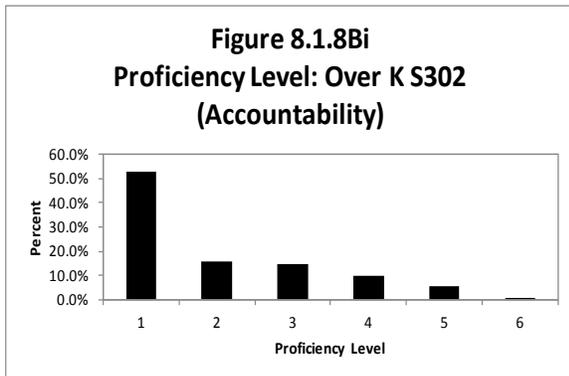
**Table 8.1.8Bi**
Proficiency Level Distribution: Over K S302
(Accountability)

**Figure 8.1.8Bi**
**Proficiency Level: Over K S302**
**(Accountability)**



| Level | Count | Percent |
|---|---|---|
| 1 | 108,070 | 53.0% |
| 2 | 32,479 | 15.9% |
| 3 | 29,796 | 14.6% |
| 4 | 19,836 | 9.7% |
| 5 | 11,866 | 5.8% |
| 6 | 1,762 | 0.9% |
| Total | 203,809 | 100.0% |
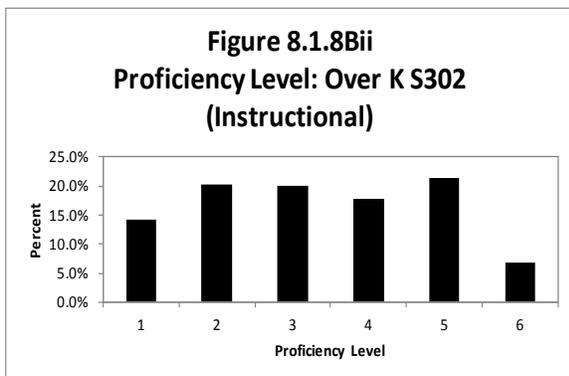
**Table 8.1.8Bii**
Proficiency Level Distribution: Over K S302
(Instructional)

**Figure 8.1.8Bii**
**Proficiency Level: Over K S302**
**(Instructional)**



| Level | Count | Percent |
|---|---|---|
| K1 | 28,669 | 14.1% |
| K2 | 41,422 | 20.3% |
| K3 | 40,481 | 19.9% |
| K4 | 36,103 | 17.7% |
| K5 | 43,506 | 21.3% |
| K6 | 13,628 | 6.7% |
| Total | 203,809 | 100.0% |

**Table 8.1.8C**
n/a

**Figure 8.1.8C**
n/a

**Figure 8.1.8D**
n/a

**Table 8.1.8D**
Overall Composite Reliability: Over K S302

| Component | Weight | Variance | Reliability |
|-----------|--------|----------|-------------|
| Listening | 0.15 | 4997.123 | 0.934 |
| Reading | 0.35 | 4368.024 | 0.947 |
| Speaking | 0.15 | 4852.854 | 0.894 |
| Writing | 0.35 | 4299.672 | 0.922 |
| Overall Composite | | 3260.417 | 0.973 |

**Table 8.1.8E**
Accuracy and Consistency of Classification Indices: Over (Grade K) S302 (Accountability)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.807 | 0.747 | | 0.616 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.956 | | 0.936 | |
| | 2 | 0.704 | | 0.593 | |
| | 3 | 0.680 | | 0.561 | |
| | 4 | 0.533 | | 0.444 | |
| | 5 | 0.253 | | 0.477 | |
| | 6 | - | | 0.149 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.953 | 0.023 | 0.024 | 0.933 |
| | 2/3 | 0.951 | 0.023 | 0.026 | 0.931 |
| | 3/4 | 0.955 | 0.021 | 0.024 | 0.935 |
| | 4/5 | 0.953 | 0.035 | 0.011 | 0.941 |
| | 5/6 | 0.991 | 0.009 | 0.000 | 0.991 |

**Table 8.1.8E**

Accuracy and Consistency of Classification Indices: Over (Grade K) S302
(Instructional)

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.787 | 0.710 | | 0.645 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.921 | | 0.878 | |
| | 2 | 0.835 | | 0.768 | |
| | 3 | 0.776 | | 0.684 | |
| | 4 | 0.726 | | 0.621 | |
| | 5 | 0.738 | | 0.681 | |
| | 6 | 0.710 | | 0.551 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.976 | 0.011 | 0.013 | 0.966 |
| | 2/3 | 0.959 | 0.021 | 0.020 | 0.941 |
| | 3/4 | 0.950 | 0.023 | 0.027 | 0.929 |
| | 4/5 | 0.953 | 0.021 | 0.026 | 0.933 |
| | 5/6 | 0.950 | 0.038 | 0.012 | 0.938 |

## 8.2    Grades: 1–2

### 8.2.1    Listening 1-2



**Figure 8.2.1A**
**Scale Scores: List 1-2 S302**



**Figure 8.2.1B**
**Proficiency Level: List 1-2 S302**
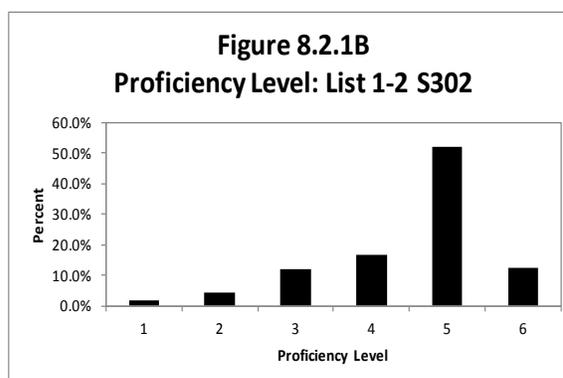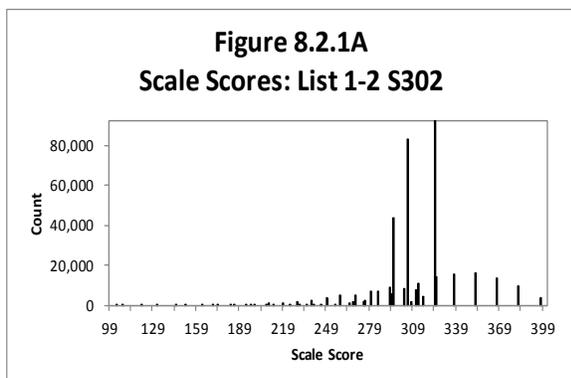
**Table 8.2.1A**

Scale Score Descriptive Statistics: List 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 1 | 201,336 | 104 | 397 | 299.11 | 27.35 |
| 2 | 187,608 | 108 | 397 | 325.40 | 28.51 |
| Total | 388,944 | 104 | 397 | 311.79 | 30.85 |

**Table 8.2.1B**

Proficiency Level Distribution: List 1-2 S302

| Level | Grade 1 | | Grade 2 | | Total | |
|-------|---------|---------|---------|---------|---------|---------|
| | Count | Percent | Count | Percent | Count | Percent |
| 1 | 4,892 | 2.4% | 2,876 | 1.5% | 7,768 | 2.0% |
| 2 | 10,535 | 5.2% | 6,561 | 3.5% | 17,096 | 4.4% |
| 3 | 30,868 | 15.3% | 16,335 | 8.7% | 47,203 | 12.1% |
| 4 | 46,918 | 23.3% | 18,243 | 9.7% | 65,161 | 16.8% |
| 5 | 91,944 | 45.7% | 110,748 | 59.0% | 202,692 | 52.1% |
| 6 | 16,179 | 8.0% | 32,845 | 17.5% | 49,024 | 12.6% |
| Total | 201,336 | 100.0% | 187,608 | 100.0% | 388,944 | 100.0% |

**Table 8.2.1C**

Conditional Standard Error of Measurement at Cut Scores: List 1-2 S302

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 1 | 238 | 19.16 | 19.91 | 19.16 |
| | 2 | 247 | 19.16 | 19.54 | 18.79 |
| 2/3 | 1 | 267 | 19.91 | 19.16 | 18.03 |
| | 2 | 281 | 21.04 | 19.16 | 18.41 |
| 3/4 | 1 | 295 | 22.92 | 19.54 | 18.79 |
| | 2 | 311 | 25.55 | 20.29 | 19.91 |
| 4/5 | 1 | 305 | n/a | 19.91 | 19.16 |
| | 2 | 324 | n/a | 21.42 | 21.04 |
| 5/6 | 1 | 330 | n/a | n/a | 21.79 |
| | 2 | 350 | n/a | n/a | 24.80 |

**Figure 8.2.1C**

Test Characteristic Curve: List 1-2ABC S302
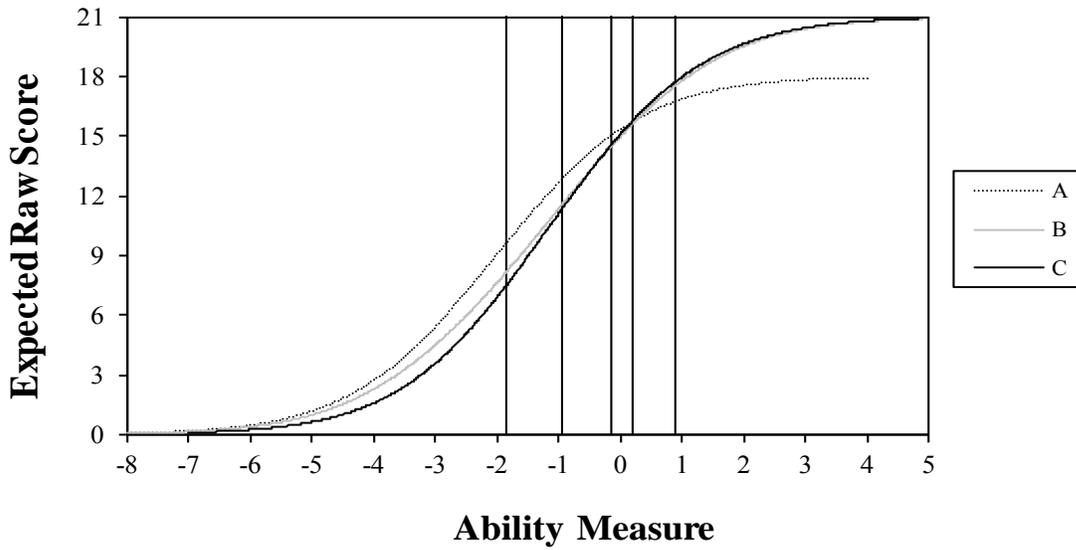


**Figure 8.2.1D**

Test Information Function: List 1-2ABC S302
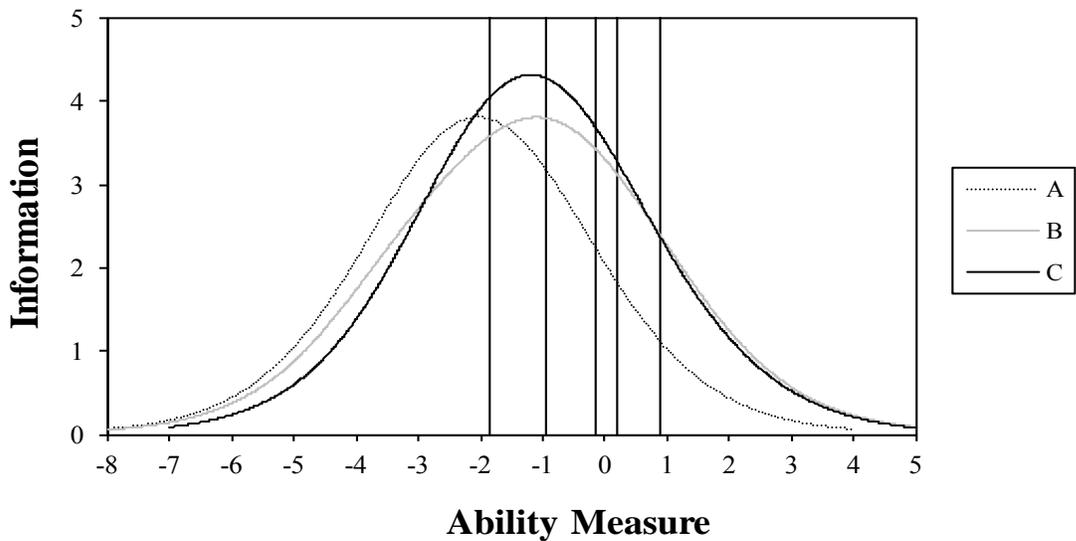
**Table 8.2.1D**
Weighted Reliability: List 1-2 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|-----------------|-------------|-------------|
| A | 85,383 | 0.781 | |
| B | 193,992 | 0.672 | 0.688 |
| C | 109,569 | 0.645 | |

**Table 8.2.1E-1**
Accuracy and Consistency of Classification Indices: List (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.534 | 0.395 | | 0.160 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.808 | | 0.615 | |
| | 2 | 0.534 | | 0.342 | |
| | 3 | 0.429 | | 0.274 | |
| | 4 | 0.354 | | 0.271 | |
| | 5 | 0.604 | | 0.561 | |
| | 6 | - | | 0.141 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.987 | 0.003 | 0.010 | 0.981 |
| | 2/3 | 0.962 | 0.013 | 0.025 | 0.939 |
| | 3/4 | 0.853 | 0.077 | 0.070 | 0.770 |
| | 4/5 | 0.740 | 0.097 | 0.163 | 0.658 |
| | 5/6 | 0.920 | 0.080 | 0.000 | 0.847 |

**Table 8.2.1E-2**
Accuracy and Consistency of Classification Indices: List (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.598 | 0.439 | | 0.149 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.816 | | 0.618 | |
| | 2 | 0.548 | | 0.371 | |
| | 3 | 0.444 | | 0.257 | |
| | 4 | 0.217 | | 0.135 | |
| | 5 | 0.675 | | 0.659 | |
| | 6 | - | | 0.265 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.991 | 0.002 | 0.007 | 0.988 |
| | 2/3 | 0.975 | 0.007 | 0.017 | 0.962 |
| | 3/4 | 0.923 | 0.036 | 0.042 | 0.866 |
| | 4/5 | 0.840 | 0.086 | 0.074 | 0.748 |
| | 5/6 | 0.825 | 0.175 | 0.000 | 0.731 |

## 8.2.2    Reading 1-2

**Figure 8.2.2A**
**Scale Scores: Read 1-2 S302**



**Figure 8.2.2B**
**Proficiency Level: Read 1-2 S302**
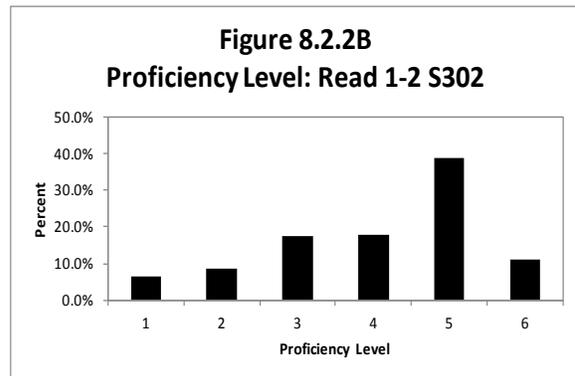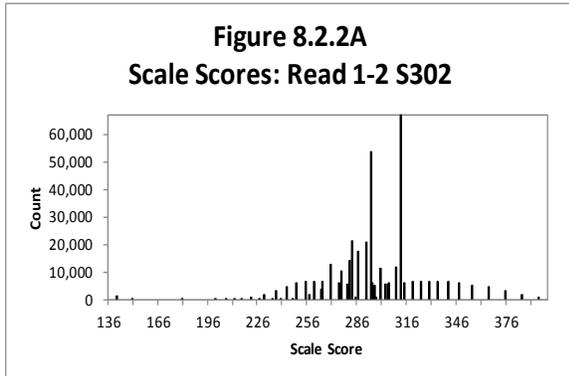


**Table 8.2.2A**

Scale Score Descriptive Statistics: Read 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 1 | 201,177 | 141 | 395 | 283.48 | 24.82 |
| 2 | 187,492 | 150 | 395 | 310.49 | 25.84 |
| Total | 388,669 | 141 | 395 | 296.51 | 28.69 |

**Table 8.2.2B**

Proficiency Level Distribution: Read 1-2 S302

| Level | Grade 1 Count | Grade 1 Percent | Grade 2 Count | Grade 2 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|
| 1 | 17,131 | 8.5% | 7,724 | 4.1% | 24,855 | 6.4% |
| 2 | 20,510 | 10.2% | 12,664 | 6.8% | 33,174 | 8.5% |
| 3 | 37,350 | 18.6% | 30,187 | 16.1% | 67,537 | 17.4% |
| 4 | 46,516 | 23.1% | 23,355 | 12.5% | 69,871 | 18.0% |
| 5 | 65,236 | 32.4% | 85,287 | 45.5% | 150,523 | 38.7% |
| 6 | 14,434 | 7.2% | 28,275 | 15.1% | 42,709 | 11.0% |
| Total | 201,177 | 100.0% | 187,492 | 100.0% | 388,669 | 100.0% |

**Table 8.2.2C**
Conditional Standard Error of Measurement at Cut Scores: Read 1-2 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 1 | 253 | 11.96 | 15.34 | 14.30 |
| | 2 | 267 | 11.44 | 13.00 | 12.74 |
| 2/3 | 1 | 269 | 11.44 | 12.74 | 12.48 |
| | 2 | 286 | 11.70 | 11.18 | 11.18 |
| 3/4 | 1 | 283 | 11.70 | 11.44 | 11.44 |
| | 2 | 303 | 13.00 | 10.66 | 10.92 |
| 4/5 | 1 | 294 | n/a | 10.92 | 10.92 |
| | 2 | 312 | n/a | 10.66 | 10.92 |
| 5/6 | 1 | 314 | n/a | n/a | 11.18 |
| | 2 | 331 | n/a | n/a | 11.96 |

* No equating was performed for S302

# Figure 8.2.2C
## Test Characteristic Curve: Read 1-2ABC S302
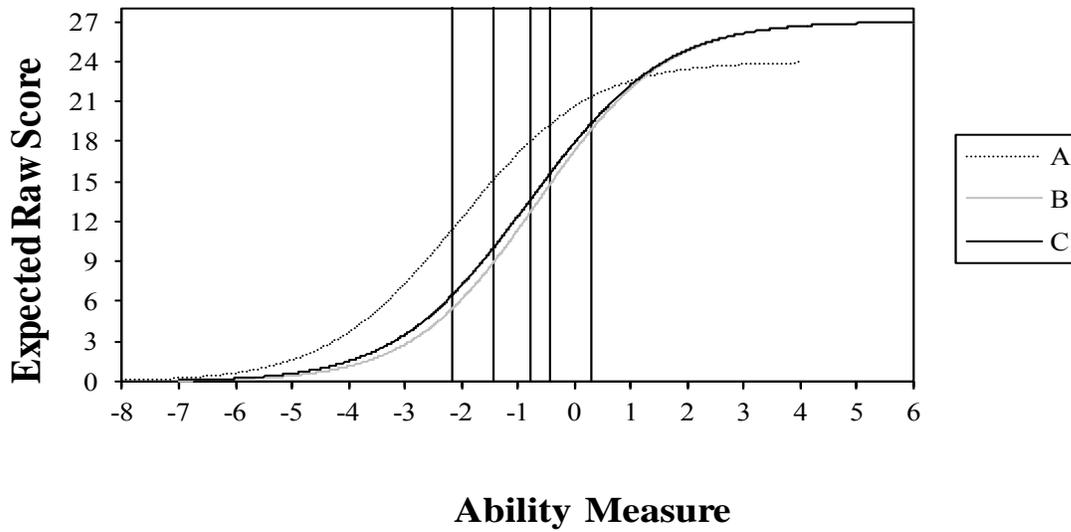


# Figure 8.2.2D
## Test Information Function: Read 1-2ABC S302
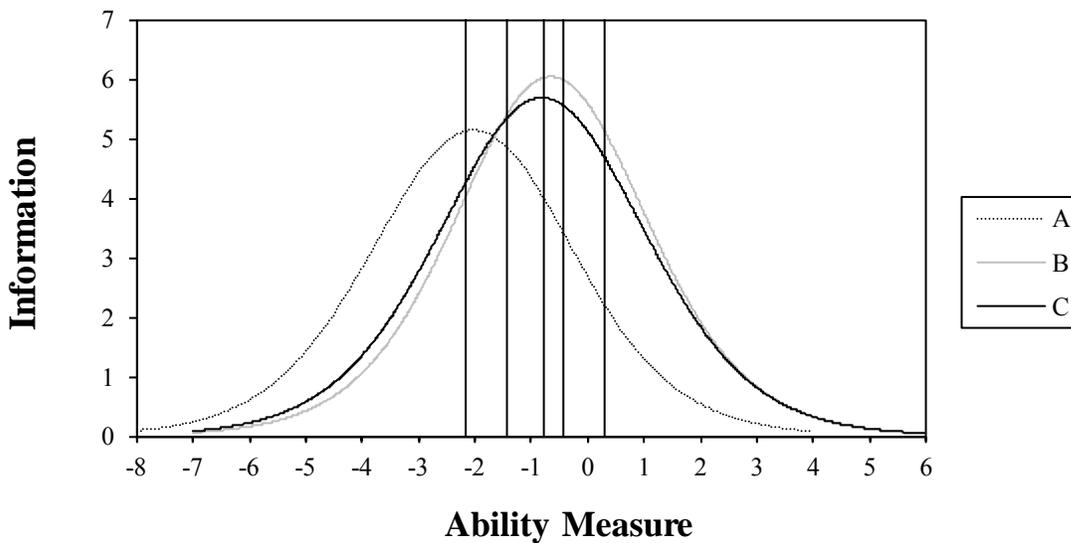
**Table 8.2.2D**
Weighted Reliability: Read 1-2 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|-----------------|-------------|-------------|
| A | 85,324 | 0.801 | |
| B | 193,891 | 0.838 | 0.828 |
| C | 109,454 | 0.832 | |

**Table 8.2.2E-1**
Accuracy and Consistency of Classification Indices: Read (Grade 1) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| **Indices** | 0.544 | 0.428 | | 0.272 | |
| **Conditional** | **Level** | **Accuracy** | | **Consistency** | |
| **on Level** | 1 | 0.814 | | 0.671 | |
| | 2 | 0.482 | | 0.340 | |
| | 3 | 0.473 | | 0.354 | |
| | 4 | 0.433 | | 0.339 | |
| | 5 | 0.596 | | 0.527 | |
| | 6 | - | | 0.239 | |
| **Indices at** | | | **Accuracy** | | |
| **Cut Points** | | | **False** | **False** | |
| | **Cut Point** | **Accuracy** | **Positives** | **Negatives** | **Consistency** |
| | 1/2 | 0.964 | 0.015 | 0.021 | 0.946 |
| | 2/3 | 0.923 | 0.036 | 0.041 | 0.886 |
| | 3/4 | 0.860 | 0.064 | 0.076 | 0.806 |
| | 4/5 | 0.825 | 0.073 | 0.102 | 0.761 |
| | 5/6 | 0.928 | 0.072 | 0.000 | 0.889 |

**Table 8.2.2E-2**
Accuracy and Consistency of Classification Indices: Read (Grade 2) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| **Indices** | 0.576 | 0.459 | | 0.291 | |
| **Conditional** | **Level** | **Accuracy** | | **Consistency** | |
| **on Level** | 1 | 0.802 | | 0.592 | |
| | 2 | 0.453 | | 0.307 | |
| | 3 | 0.484 | | 0.349 | |
| | 4 | 0.259 | | 0.199 | |
| | 5 | 0.737 | | 0.651 | |
| | 6 | 0.624 | | 0.474 | |
| **Indices at** | | | **Accuracy** | | |
| **Cut Points** | | | **False** | **False** | |
| | **Cut Point** | **Accuracy** | **Positives** | **Negatives** | **Consistency** |
| | 1/2 | 0.976 | 0.006 | 0.018 | 0.966 |
| | 2/3 | 0.944 | 0.026 | 0.030 | 0.913 |
| | 3/4 | 0.878 | 0.067 | 0.055 | 0.822 |
| | 4/5 | 0.828 | 0.110 | 0.061 | 0.771 |
| | 5/6 | 0.892 | 0.043 | 0.065 | 0.838 |

## 8.2.3    Writing 1-2



**Figure 8.2.3A**
**Scale Scores: Writ 1-2 S302**



**Figure 8.2.3B**
**Proficiency Level: Writ 1-2 S302**
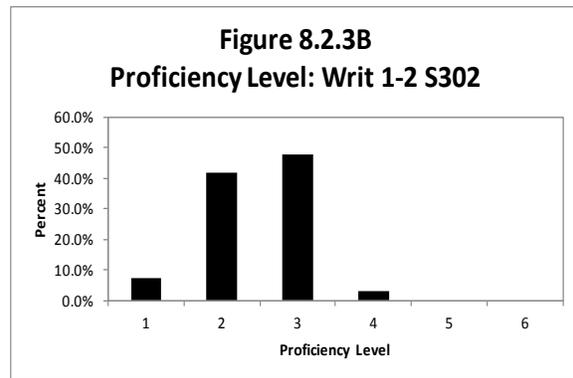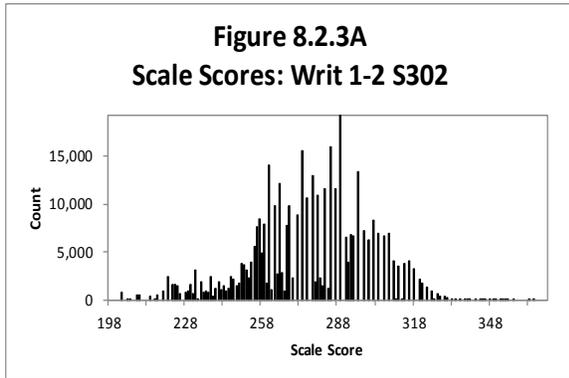
**Table 8.2.3A**

Scale Score Descriptive Statistics: Writ 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 1 | 201,249 | 203 | 365 | 266.82 | 20.66 |
| 2 | 187,537 | 209 | 363 | 285.90 | 21.44 |
| Total | 388,786 | 203 | 365 | 276.02 | 23.10 |

**Table 8.2.3B**

Proficiency Level Distribution: Writ 1-2 S302

| Level | Grade 1 Count | Grade 1 Percent | Grade 2 Count | Grade 2 Percent | Total Count | Total Percent |
|-------|---------------|-----------------|---------------|-----------------|-------------|---------------|
| 1 | 17,984 | 8.9% | 10,252 | 5.5% | 28,236 | 7.3% |
| 2 | 97,405 | 48.4% | 64,792 | 34.5% | 162,197 | 41.7% |
| 3 | 81,036 | 40.3% | 104,861 | 55.9% | 185,897 | 47.8% |
| 4 | 4,808 | 2.4% | 7,602 | 4.1% | 12,410 | 3.2% |
| 5 | 15 | 0.0% | 30 | 0.0% | 45 | 0.0% |
| 6 | 1 | 0.0% | 0 | 0.0% | 1 | 0.0% |
| Total | 201,249 | 100.0% | 187,537 | 100.0% | 388,786 | 100.0% |

**Table 8.2.3C**

Conditional Standard Error of Measurement at Cut Scores: Writ 1-2 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 1 | 238 | 7.15 | 6.84 | 6.53 |
| | 2 | 251 | 6.84 | 7.46 | 7.46 |
| 2/3 | 1 | 272 | 7.77 | 8.09 | 8.40 |
| | 2 | 285 | 8.40 | 8.09 | 8.40 |
| 3/4 | 1 | 308 | 8.09 | 7.77 | 7.77 |
| | 2 | 320 | 7.77 | 7.46 | 7.46 |
| 4/5 | 1 | 336 | n/a | 6.84 | 6.84 |
| | 2 | 348 | n/a | 6.53 | 6.53 |
| 5/6 | 1 | 362 | n/a | n/a | 6.53 |
| | 2 | 373 | n/a | n/a | 7.46 |

* No equating was performed for Writing Tier A S302

## Figure 8.2.3C
### Test Characteristic Curve: Writ 1-2ABC S302
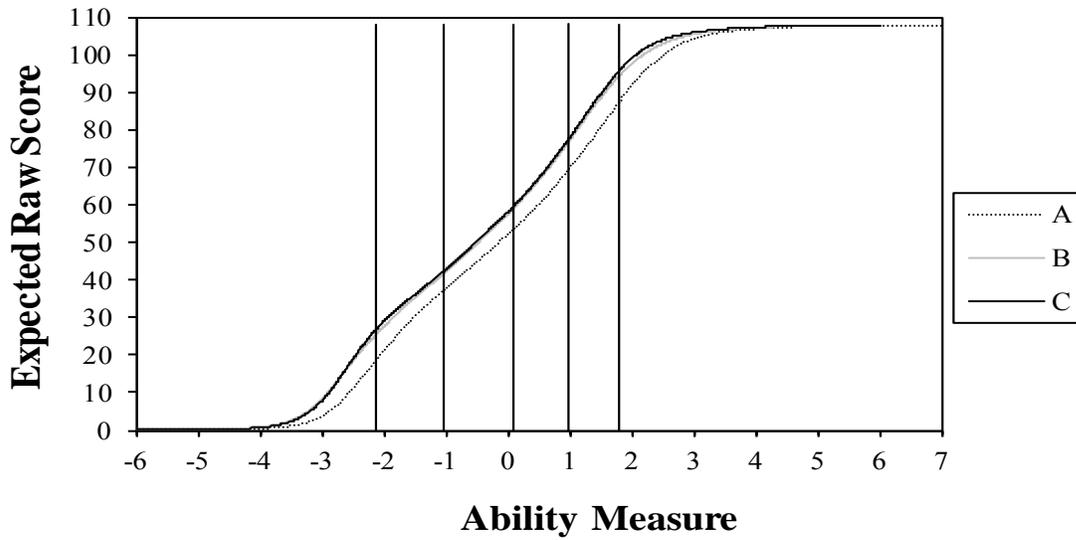


## Figure 8.2.3D
### Test Information Function: Writ 1-2ABC S302
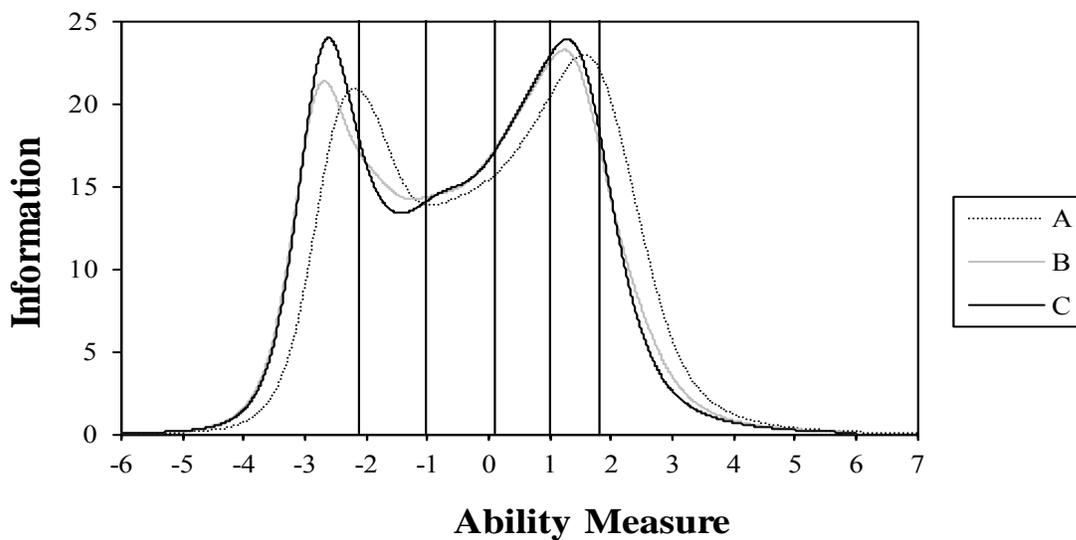
**Table 8.2.3D**
Weighted Reliability: Writ 1-2 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|-----------------|-------------|-------------|
| A | 85,375 | 0.896 | |
| B | 193,909 | 0.926 | 0.925 |
| C | 109,502 | 0.945 | |

**Table 8.2.3E-1**
Accuracy and Consistency of Classification Indices: Writ (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.863 | 0.815 | | 0.686 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.851 | | 0.750 | |
| | 2 | 0.881 | | 0.838 | |
| | 3 | 0.846 | | 0.812 | |
| | 4 | 0.791 | | 0.458 | |
| | 5 | - | | 1.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.970 | 0.013 | 0.018 | 0.956 |
| | 2/3 | 0.917 | 0.040 | 0.043 | 0.883 |
| | 3/4 | 0.977 | 0.023 | 0.000 | 0.975 |
| | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |

**Table 8.2.3E-2**
Accuracy and Consistency of Classification Indices: Writ (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.864 | 0.816 | | 0.668 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.831 | | 0.727 | |
| | 2 | 0.865 | | 0.805 | |
| | 3 | 0.866 | | 0.847 | |
| | 4 | - | | 0.300 | |
| | 5 | - | | 1.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.981 | 0.009 | 0.010 | 0.972 |
| | 2/3 | 0.923 | 0.036 | 0.041 | 0.892 |
| | 3/4 | 0.959 | 0.041 | 0.000 | 0.953 |
| | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |

## 8.2.4　Speaking 1-2



**Figure 8.2.4A**
**Scale Scores: Spek 1-2 S302**



**Figure 8.2.4B**
**Proficiency Level: Spek 1-2 S302**
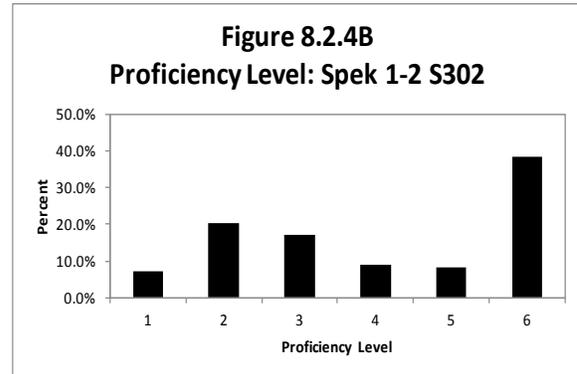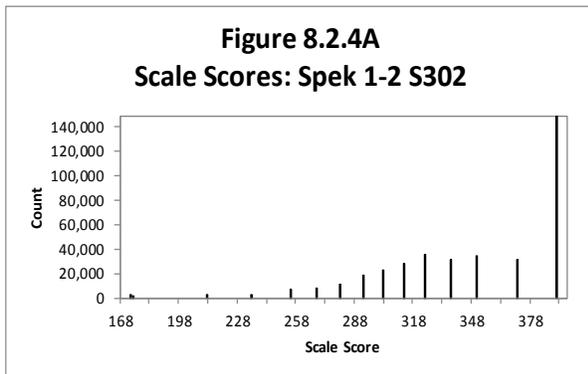
**Table 8.2.4A**

Scale Score Descriptive Statistics: Spek 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 1 | 201,244 | 173 | 391 | 337.37 | 47.84 |
| 2 | 187,558 | 174 | 391 | 356.05 | 42.90 |
| Total | 388,802 | 173 | 391 | 346.38 | 46.47 |

**Table 8.2.4B**

Proficiency Level Distribution: Spek 1-2 S302

| Level | Grade 1 Count | Grade 1 Percent | Grade 2 Count | Grade 2 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|
| 1 | 16,794 | 8.3% | 11,526 | 6.1% | 28,320 | 7.3% |
| 2 | 54,483 | 27.1% | 24,077 | 12.8% | 78,560 | 20.2% |
| 3 | 37,167 | 18.5% | 29,518 | 15.7% | 66,685 | 17.2% |
| 4 | 17,430 | 8.7% | 17,124 | 9.1% | 34,554 | 8.9% |
| 5 | 14,834 | 7.4% | 17,116 | 9.1% | 31,950 | 8.2% |
| 6 | 60,536 | 30.1% | 88,197 | 47.0% | 148,733 | 38.3% |
| Total | 201,244 | 100.0% | 187,558 | 100.0% | 388,802 | 100.0% |

**Table 8.2.4C**
Conditional Standard Error of Measurement at
Cut Scores: Spek 1-2 S302*

| Proficiency Level | Grade | Cut Score | SEM |
|---|---|---|---|
| 1/2 | 1 | 278 | 20.89 |
| | 2 | 286 | 19.88 |
| 2/3 | 1 | 318 | 18.28 |
| | 2 | 322 | 18.28 |
| 3/4 | 1 | 344 | 19.08 |
| | 2 | 345 | 19.08 |
| 4/5 | 1 | 367 | 20.08 |
| | 2 | 368 | 20.08 |
| 5/6 | 1 | 385 | 20.69 |
| | 2 | 386 | 20.69 |

*No equating was performed for S302

**Figure 8.2.4C**
Test Characteristic Curve: Spek 1-2 S302
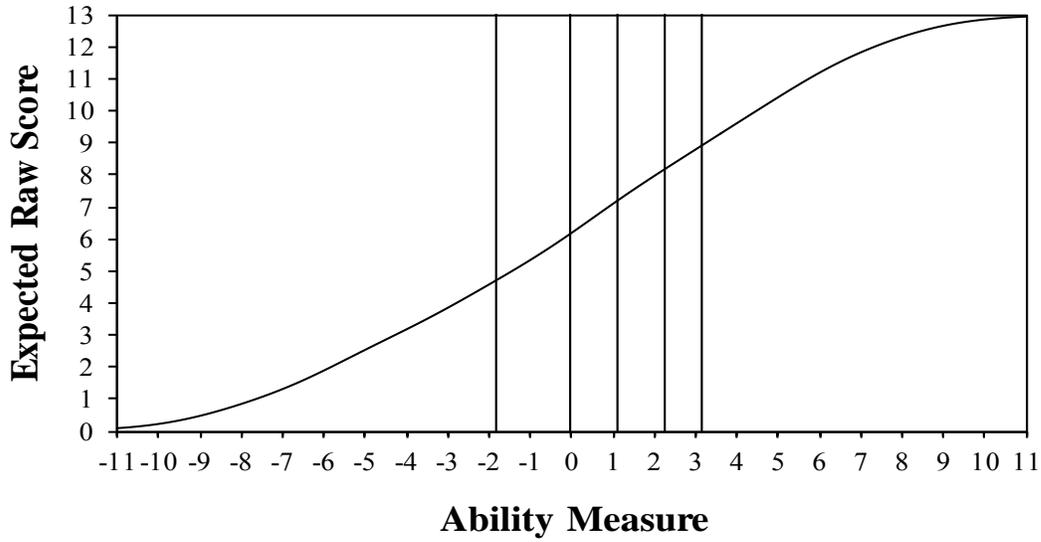


**Figure 8.2.4D**
Test Information Function: Spek 1-2 S302
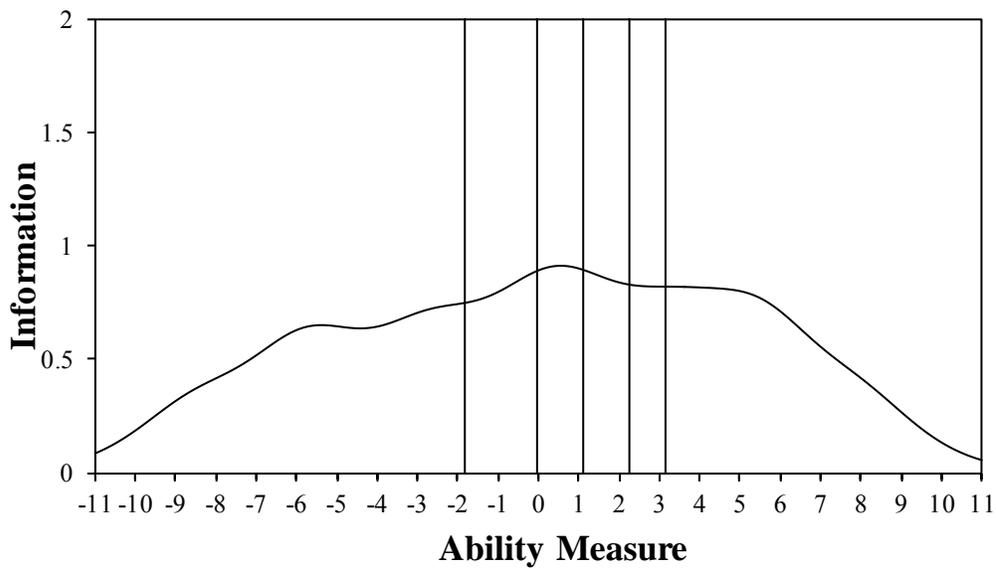
**Table 8.2.4D**

Reliability: Spek 1-2 S302

| Tiers | No. of Students | Reliability |
|---|---|---|
| -- | 388,802 | 0.891 |

**Table 8.2.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.610 | 0.519 | | 0.409 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.599 | | 0.467 | |
| | 2 | 0.717 | | 0.623 | |
| | 3 | 0.530 | | 0.427 | |
| | 4 | 0.366 | | 0.249 | |
| | 5 | 0.277 | | 0.207 | |
| | 6 | 0.947 | | 0.883 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.938 | 0.043 | 0.020 | 0.908 |
| | 2/3 | 0.888 | 0.047 | 0.064 | 0.853 |
| | 3/4 | 0.921 | 0.024 | 0.055 | 0.884 |
| | 4/5 | 0.949 | 0.028 | 0.023 | 0.916 |
| | 5/6 | 0.890 | 0.098 | 0.011 | 0.873 |

**Table 8.2.4E-2**

Accuracy and Consistency of Classification Indices: Spek (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.653 | 0.582 | | 0.441 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.687 | | 0.542 | |
| | 2 | 0.576 | | 0.468 | |
| | 3 | 0.561 | | 0.465 | |
| | 4 | 0.370 | | 0.269 | |
| | 5 | 0.305 | | 0.223 | |
| | 6 | 0.955 | | 0.906 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.964 | 0.021 | 0.015 | 0.945 |
| | 2/3 | 0.922 | 0.039 | 0.039 | 0.897 |
| | 3/4 | 0.918 | 0.023 | 0.060 | 0.888 |
| | 4/5 | 0.945 | 0.026 | 0.029 | 0.910 |
| | 5/6 | 0.880 | 0.103 | 0.017 | 0.856 |

## 8.2.5    Oral Language Composite 1-2



**Figure 8.2.5A**
**Scale Scores: Oral 1-2 S302**



**Figure 8.2.5B**
**Proficiency Level: Oral 1-2 S302**
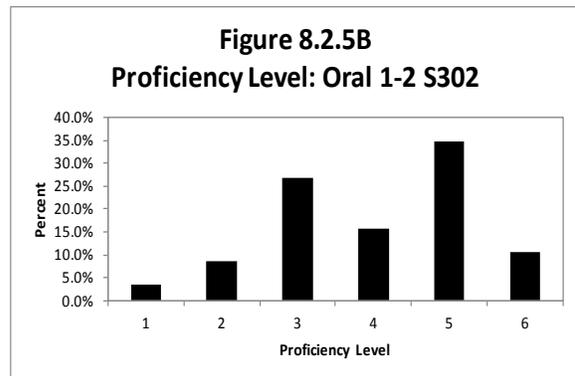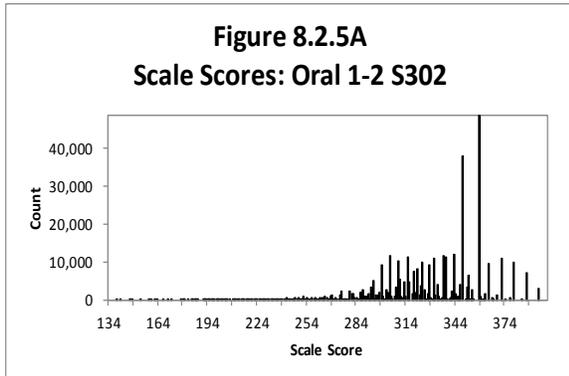
**Table 8.2.5A**

Scale Score Descriptive Statistics: Oral 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 1 | 201,132 | 139 | 394 | 318.47 | 32.54 |
| 2 | 187,452 | 141 | 394 | 341.02 | 30.79 |
| Total | 388,584 | 139 | 394 | 329.35 | 33.65 |

**Table 8.2.5B**

Proficiency Level Distribution: Oral 1-2 S302

| Level | Grade 1 | | Grade 2 | | Total | |
|-------|---------|---------|---------|---------|---------|---------|
| | Count | Percent | Count | Percent | Count | Percent |
| 1 | 8,563 | 4.3% | 5,161 | 2.8% | 13,724 | 3.5% |
| 2 | 23,951 | 11.9% | 9,752 | 5.2% | 33,703 | 8.7% |
| 3 | 65,602 | 32.6% | 38,724 | 20.7% | 104,326 | 26.8% |
| 4 | 30,505 | 15.2% | 30,532 | 16.3% | 61,037 | 15.7% |
| 5 | 57,456 | 28.6% | 77,171 | 41.2% | 134,627 | 34.6% |
| 6 | 15,055 | 7.5% | 26,112 | 13.9% | 41,167 | 10.6% |
| Total | 201,132 | 100.0% | 187,452 | 100.0% | 388,584 | 100.0% |

**Table 8.2.5C**
n/a

**Figure 8.2.5C**
n/a

**Figure 8.2.5D**
n/a

**Table 8.2.5D**
Oral Composite Reliability: Oral 1-2 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.50 | 950.266 | 0.688 |
| Speaking | 0.50 | 2155.715 | 0.891 |
| Oral | | 1131.001 | 0.882 |

*Variances from students who had results in all four domains

**Table 8.2.5E-1**
Accuracy and Consistency of Classification Indices: Oral (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.630 | 0.516 | | 0.379 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.838 | | 0.717 | |
| | 2 | 0.665 | | 0.528 | |
| | 3 | 0.773 | | 0.665 | |
| | 4 | 0.366 | | 0.272 | |
| | 5 | 0.637 | | 0.574 | |
| | 6 | - | | 0.300 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.984 | 0.006 | 0.010 | 0.976 |
| | 2/3 | 0.939 | 0.033 | 0.029 | 0.911 |
| | 3/4 | 0.885 | 0.035 | 0.080 | 0.841 |
| | 4/5 | 0.882 | 0.063 | 0.054 | 0.828 |
| | 5/6 | 0.925 | 0.075 | 0.000 | 0.898 |

**Table 8.2.5E-2**

Accuracy and Consistency of Classification Indices: Oral (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.631 | 0.514 | | 0.347 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.877 | | 0.777 | |
| | 2 | 0.578 | | 0.431 | |
| | 3 | 0.775 | | 0.654 | |
| | 4 | 0.448 | | 0.326 | |
| | 5 | 0.643 | | 0.611 | |
| | 6 | - | | 0.328 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.991 | 0.003 | 0.006 | 0.987 |
| | 2/3 | 0.967 | 0.020 | 0.013 | 0.950 |
| | 3/4 | 0.916 | 0.024 | 0.059 | 0.887 |
| | 4/5 | 0.889 | 0.047 | 0.064 | 0.834 |
| | 5/6 | 0.861 | 0.139 | 0.000 | 0.813 |

## 8.2.6    Literacy Composite 1-2



**Figure 8.2.6A**
**Scale Scores: Litr 1-2 S302**



**Figure 8.2.6B**
**Proficiency Level: Litr 1-2 S302**

**Table 8.2.6A**
Scale Score Descriptive Statistics: Litr 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 1 | 201,059 | 172 | 369 | 275.44 | 19.98 |
| 2 | 187,369 | 180 | 376 | 298.49 | 21.47 |
| Total | 388,428 | 172 | 376 | 286.56 | 23.70 |

**Table 8.2.6B**
Proficiency Level Distribution: Litr 1-2 S302

| Level | Grade 1 Count | Grade 1 Percent | Grade 2 Count | Grade 2 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|---------|---------|
| 1 | 11,892 | 5.9% | 7,380 | 3.9% | 19,272 | 5.0% |
| 2 | 59,372 | 29.5% | 34,935 | 18.6% | 94,307 | 24.3% |
| 3 | 107,734 | 53.6% | 105,639 | 56.4% | 213,373 | 54.9% |
| 4 | 15,285 | 7.6% | 24,999 | 13.3% | 40,284 | 10.4% |
| 5 | 5,846 | 2.9% | 12,853 | 6.9% | 18,699 | 4.8% |
| 6 | 930 | 0.5% | 1,563 | 0.8% | 2,493 | 0.6% |
| Total | 201,059 | 100.0% | 187,369 | 100.0% | 388,428 | 100.0% |

**Table 8.2.6C**
n/a

**Figure 8.2.6C**
n/a

**Figure 8.2.6D**
n/a

**Table 8.2.6D**
Literacy Composite Reliability: Litr 1-2 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Reading | 0.50 | 822.466 | 0.828 |
| Writing | 0.50 | 532.856 | 0.925 |
| Literacy | | 561.317 | 0.919 |

**Table 8.2.6E-1**
Accuracy and Consistency of Classification Indices: Litr (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.813 | 0.740 | | 0.590 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.795 | | 0.699 | |
| | 2 | 0.802 | | 0.721 | |
| | 3 | 0.880 | | 0.840 | |
| | 4 | 0.527 | | 0.407 | |
| | 5 | 0.763 | | 0.673 | |
| | 6 | - | | 0.997 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.975 | 0.012 | 0.013 | 0.966 |
| | 2/3 | 0.910 | 0.049 | 0.042 | 0.875 |
| | 3/4 | 0.945 | 0.023 | 0.032 | 0.922 |
| | 4/5 | 0.978 | 0.021 | 0.001 | 0.978 |
| | 5/6 | 0.995 | 0.005 | 0.000 | 0.999 |

**Table 8.2.6E-2**

Accuracy and Consistency of Classification Indices: Litr (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.794 | 0.715 | | 0.560 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.827 | | 0.724 | |
| | 2 | 0.752 | | 0.648 | |
| | 3 | 0.891 | | 0.851 | |
| | 4 | 0.562 | | 0.451 | |
| | 5 | 0.745 | | 0.642 | |
| | 6 | - | | 0.872 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.984 | 0.006 | 0.010 | 0.978 |
| | 2/3 | 0.928 | 0.040 | 0.032 | 0.898 |
| | 3/4 | 0.923 | 0.027 | 0.050 | 0.890 |
| | 4/5 | 0.964 | 0.030 | 0.006 | 0.955 |
| | 5/6 | 0.992 | 0.008 | 0.000 | 0.993 |

## 8.2.7　Comprehension Composite 1-2



**Figure 8.2.7A**
**Scale Scores: Cphn 1-2 S302**



**Figure 8.2.7B**
**Proficiency Level: Cphn 1-2 S302**
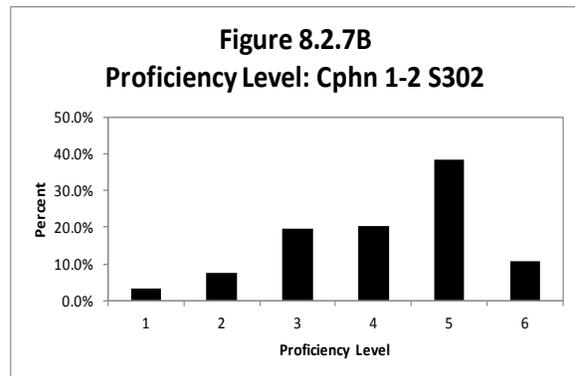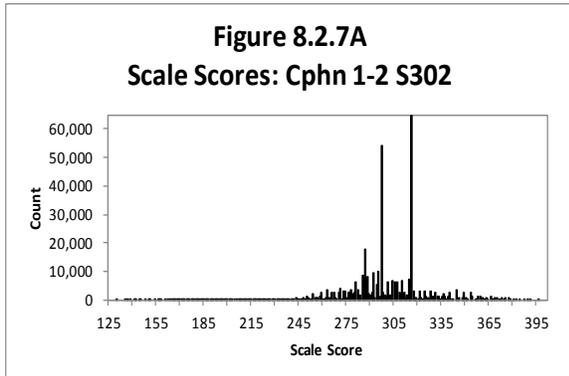
**Table 8.2.7A**

Scale Score Descriptive Statistics: Cphn 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 1 | 201,119 | 130 | 396 | 288.17 | 23.06 |
| 2 | 187,438 | 137 | 396 | 315.11 | 24.50 |
| Total | 388,557 | 130 | 396 | 301.17 | 27.31 |

**Table 8.2.7B**

Proficiency Level Distribution: Cphn 1-2 S302

| Level | Grade 1 Count | Grade 1 Percent | Grade 2 Count | Grade 2 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|---------|---------|
| 1 | 8,077 | 4.0% | 4,206 | 2.2% | 12,283 | 3.2% |
| 2 | 20,404 | 10.1% | 8,528 | 4.5% | 28,932 | 7.4% |
| 3 | 45,271 | 22.5% | 31,646 | 16.9% | 76,917 | 19.8% |
| 4 | 49,007 | 24.4% | 30,713 | 16.4% | 79,720 | 20.5% |
| 5 | 64,865 | 32.3% | 84,545 | 45.1% | 149,410 | 38.5% |
| 6 | 13,495 | 6.7% | 27,800 | 14.8% | 41,295 | 10.6% |
| Total | 201,119 | 100.0% | 187,438 | 100.0% | 388,557 | 100.0% |

**Table 8.2.7C**

n/a

**Figure 8.2.7C**

n/a

**Figure 8.2.7D**

n/a

**Table 8.2.7D**

Comprehension Composite Reliability: Cphn 1-2 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.30 | 950.266 | 0.688 |
| Reading | 0.70 | 822.466 | 0.828 |
| Comprehension | | 745.518 | 0.871 |

*Variances from students who had results in all four domains

**Table 8.2.7E-1**

Accuracy and Consistency of Classification Indices: Cphn (Grade 1) S302
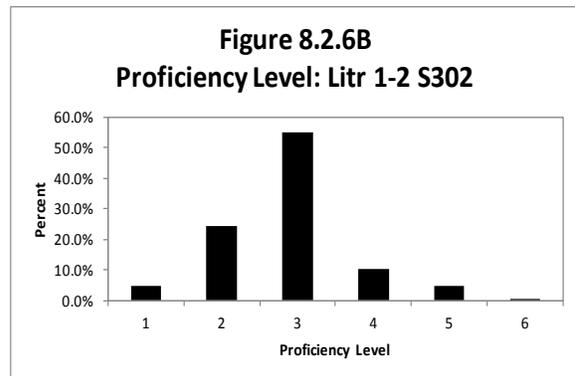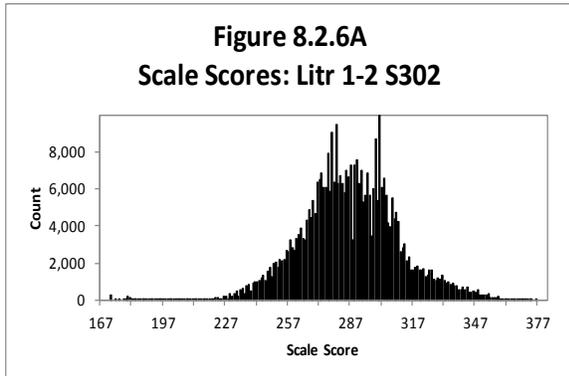
| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.623 | 0.510 | | 0.363 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.798 | | 0.651 | |
| | 2 | 0.655 | | 0.505 | |
| | 3 | 0.605 | | 0.486 | |
| | 4 | 0.526 | | 0.420 | |
| | 5 | 0.662 | | 0.584 | |
| | 6 | 0.632 | | 0.388 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.982 | 0.007 | 0.011 | 0.973 |
| | 2/3 | 0.946 | 0.023 | 0.031 | 0.922 |
| | 3/4 | 0.879 | 0.066 | 0.055 | 0.831 |
| | 4/5 | 0.858 | 0.059 | 0.083 | 0.807 |
| | 5/6 | 0.942 | 0.045 | 0.013 | 0.918 |

**Table 8.2.7E-2**

Accuracy and Consistency of Classification Indices: Cphn (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.656 | 0.540 | | 0.377 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.875 | | 0.733 | |
| | 2 | 0.539 | | 0.387 | |
| | 3 | 0.641 | | 0.495 | |
| | 4 | 0.411 | | 0.313 | |
| | 5 | 0.763 | | 0.683 | |
| | 6 | 0.680 | | 0.532 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.990 | 0.002 | 0.009 | 0.986 |
| | 2/3 | 0.969 | 0.016 | 0.015 | 0.951 |
| | 3/4 | 0.908 | 0.046 | 0.046 | 0.867 |
| | 4/5 | 0.865 | 0.076 | 0.059 | 0.814 |
| | 5/6 | 0.908 | 0.043 | 0.050 | 0.864 |

## 8.2.8    Overall Composite 1-2



**Figure 8.2.8A**
**Scale Scores: Over 1-2 S302**



**Figure 8.2.8B**
**Proficiency Level: Over 1-2 S302**
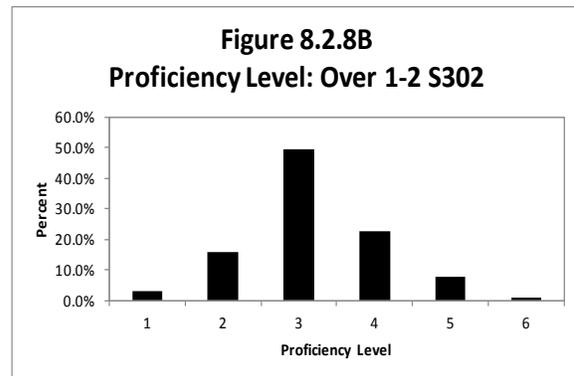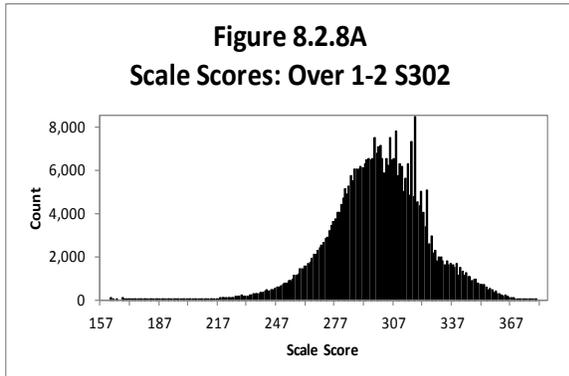
**Table 8.2.8A**

Scale Score Descriptive Statistics: Over 1-2 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 1 | 200,853 | 162 | 376 | 288.11 | 21.19 |
| 2 | 187,203 | 168 | 380 | 311.00 | 22.06 |
| Total | 388,056 | 162 | 380 | 299.16 | 24.46 |

**Table 8.2.8B**

Proficiency Level Distribution: Over 1-2 S302

| Level | Grade 1 | | Grade 2 | | Total | |
|-------|---------|---------|---------|---------|---------|---------|
| | Count | Percent | Count | Percent | Count | Percent |
| 1 | 7,368 | 3.7% | 4,795 | 2.6% | 12,163 | 3.1% |
| 2 | 42,025 | 20.9% | 19,187 | 10.2% | 61,212 | 15.8% |
| 3 | 109,264 | 54.4% | 82,581 | 44.1% | 191,845 | 49.4% |
| 4 | 31,256 | 15.6% | 57,530 | 30.7% | 88,786 | 22.9% |
| 5 | 9,231 | 4.6% | 20,218 | 10.8% | 29,449 | 7.6% |
| 6 | 1,709 | 0.9% | 2,892 | 1.5% | 4,601 | 1.2% |
| Total | 200,853 | 100.0% | 187,203 | 100.0% | 388,056 | 100.0% |

**Table 8.2.8C**
n/a


**Figure 8.2.8C**
n/a

**Figure 8.2.8D**
n/a


**Table 8.2.8D**
Overall Composite Reliability: Over 1-2 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.15 | 950.266 | 0.688 |
| Reading | 0.35 | 822.466 | 0.828 |
| Speaking | 0.15 | 2155.715 | 0.891 |
| Writing | 0.35 | 532.856 | 0.925 |
| Overall Composite | | 598.100 | 0.943 |

\*Variances from students who had results in all four domains


**Table 8.2.8E-1**
Accuracy and Consistency of Classification Indices: Over (Grade 1) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.829 | 0.765 | | 0.636 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.787 | | 0.763 | |
| | 2 | 0.820 | | 0.741 | |
| | 3 | 0.896 | | 0.860 | |
| | 4 | 0.674 | | 0.573 | |
| | 5 | 0.723 | | 0.591 | |
| | 6 | - | | 0.995 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.985 | 0.008 | 0.007 | 0.983 |
| | 2/3 | 0.938 | 0.035 | 0.028 | 0.915 |
| | 3/4 | 0.930 | 0.031 | 0.040 | 0.905 |
| | 4/5 | 0.970 | 0.024 | 0.006 | 0.966 |
| | 5/6 | 0.991 | 0.009 | 0.000 | 0.995 |

**Table 8.2.8E-2**

Accuracy and Consistency of Classification Indices: Over (Grade 2) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.802 | 0.734 | | 0.617 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.679 | | 0.809 | |
| | 2 | 0.767 | | 0.658 | |
| | 3 | 0.883 | | 0.834 | |
| | 4 | 0.748 | | 0.674 | |
| | 5 | 0.718 | | 0.599 | |
| | 6 | - | | 0.997 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.985 | 0.010 | 0.005 | 0.989 |
| | 2/3 | 0.956 | 0.028 | 0.016 | 0.946 |
| | 3/4 | 0.911 | 0.039 | 0.050 | 0.884 |
| | 4/5 | 0.936 | 0.042 | 0.022 | 0.922 |
| | 5/6 | 0.985 | 0.015 | 0.000 | 0.992 |

## 8.3  Grades: 3–5

### 8.3.1  Listening 3-5



**Figure 8.3.1A**
**Scale Scores: List 3-5 S302**



**Figure 8.3.1B**
**Proficiency Level: List 3-5 S302**
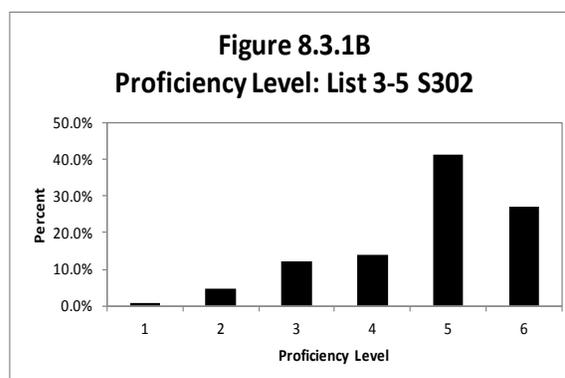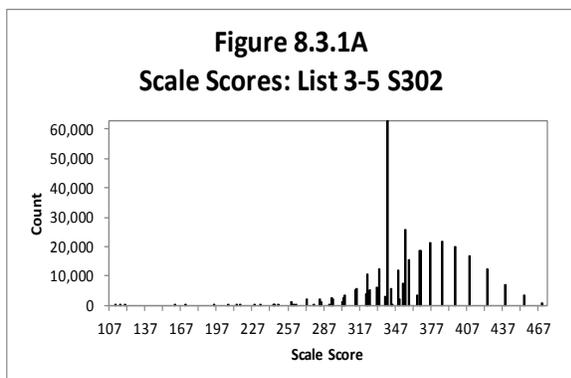
**Table 8.3.1A**
Scale Score Descriptive Statistics: List 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 3 | 164,211 | 112 | 469 | 349.43 | 34.80 |
| 4 | 105,153 | 116 | 469 | 360.74 | 36.54 |
| 5 | 82,925 | 120 | 469 | 370.90 | 38.07 |
| Total | 352,289 | 112 | 469 | 357.86 | 37.15 |

**Table 8.3.1B**
Proficiency Level Distribution: List 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|---|---|
| 1 | 939 | 0.6% | 1,042 | 1.0% | 1,298 | 1.6% | 3,279 | 0.9% |
| 2 | 6,685 | 4.1% | 5,076 | 4.8% | 4,538 | 5.5% | 16,299 | 4.6% |
| 3 | 19,397 | 11.8% | 12,013 | 11.4% | 11,347 | 13.7% | 42,757 | 12.1% |
| 4 | 15,298 | 9.3% | 17,721 | 16.9% | 15,638 | 18.9% | 48,657 | 13.8% |
| 5 | 73,211 | 44.6% | 43,069 | 41.0% | 29,186 | 35.2% | 145,466 | 41.3% |
| 6 | 48,681 | 29.6% | 26,232 | 24.9% | 20,918 | 25.2% | 95,831 | 27.2% |
| Total | 164,211 | 100.0% | 105,153 | 100.0% | 82,925 | 100.0% | 352,289 | 100.0% |

**Table 8.3.1C**
Conditional Standard Error of Measurement at Cut Scores: List 3-5 S302

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 3 | 255 | 22.54 | 22.17 | 27.05 |
| | 4 | 264 | 21.79 | 21.42 | 25.17 |
| | 5 | 274 | 20.66 | 20.66 | 23.29 |
| 2/3 | 3 | 295 | 19.54 | 19.54 | 20.66 |
| | 4 | 307 | 19.16 | 19.16 | 19.54 |
| | 5 | 318 | 19.54 | 19.16 | 18.79 |
| 3/4 | 3 | 325 | 19.54 | 19.16 | 18.41 |
| | 4 | 338 | 20.29 | 19.16 | 18.03 |
| | 5 | 350 | 21.42 | 19.91 | 18.03 |
| 4/5 | 3 | 340 | n/a | 19.54 | 18.03 |
| | 4 | 355 | n/a | 20.29 | 18.41 |
| | 5 | 368 | n/a | 21.42 | 18.79 |
| 5/6 | 3 | 367 | n/a | n/a | 18.41 |
| | 4 | 383 | n/a | n/a | 19.54 |
| | 5 | 397 | n/a | n/a | 20.66 |

# Figure 8.3.1C
## Test Characteristic Curve: List 3-5ABC S302
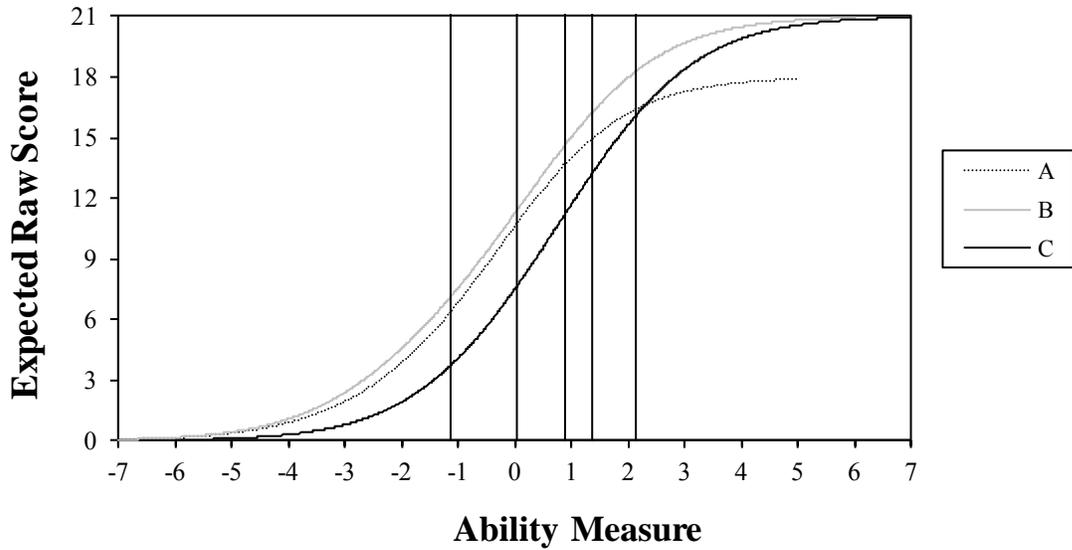


# Figure 8.3.1D
## Test Information Function: List 3-5ABC S302
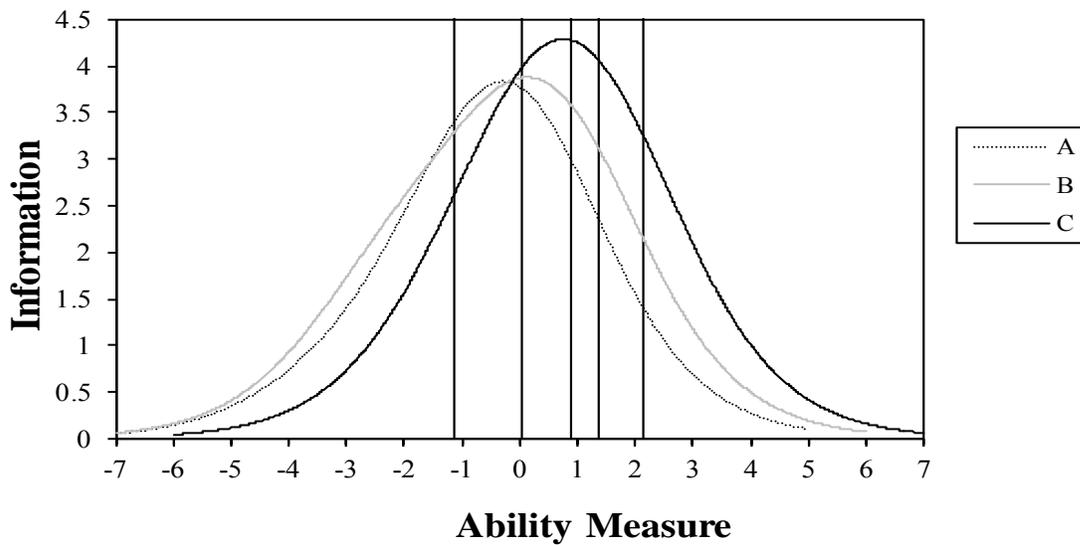
**Table 8.3.1D**
Weighted Reliability: List 3-5 S302

| Tiers | No. of Students | Reliability | Weighted Reliability |
|---|---|---|---|
| A | 29,244 | 0.761 | |
| B | 148,917 | 0.663 | 0.657 |
| C | 174,128 | 0.634 | |

**Table 8.3.1E-1**
Accuracy and Consistency of Classification Indices: List (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.537 | 0.428 | | 0.213 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.711 | | 0.258 | |
| | 2 | 0.509 | | 0.270 | |
| | 3 | 0.414 | | 0.267 | |
| | 4 | 0.183 | | 0.137 | |
| | 5 | 0.616 | | 0.540 | |
| | 6 | 0.677 | | 0.544 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.994 | 0.000 | 0.005 | 0.993 |
| | 2/3 | 0.962 | 0.007 | 0.031 | 0.939 |
| | 3/4 | 0.883 | 0.051 | 0.066 | 0.818 |
| | 4/5 | 0.811 | 0.120 | 0.069 | 0.739 |
| | 5/6 | 0.806 | 0.101 | 0.093 | 0.734 |

**Table 8.3.1E-2**
Accuracy and Consistency of Classification Indices: List (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.505 | 0.393 | | 0.186 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.744 | | 0.396 | |
| | 2 | 0.524 | | 0.303 | |
| | 3 | 0.381 | | 0.249 | |
| | 4 | 0.313 | | 0.237 | |
| | 5 | 0.555 | | 0.485 | |
| | 6 | 0.607 | | 0.464 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.992 | 0.001 | 0.007 | 0.988 |
| | 2/3 | 0.959 | 0.011 | 0.029 | 0.933 |
| | 3/4 | 0.880 | 0.058 | 0.062 | 0.817 |
| | 4/5 | 0.795 | 0.100 | 0.105 | 0.726 |
| | 5/6 | 0.804 | 0.099 | 0.097 | 0.729 |

**Table 8.3.1E-3**

Accuracy and Consistency of Classification Indices: List (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.483 | 0.378 | | 0.182 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.764 | | 0.447 | |
| | 2 | 0.486 | | 0.295 | |
| | 3 | 0.415 | | 0.284 | |
| | 4 | 0.341 | | 0.260 | |
| | 5 | 0.486 | | 0.421 | |
| | 6 | 0.625 | | 0.468 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.988 | 0.002 | 0.010 | 0.983 |
| | 2/3 | 0.953 | 0.015 | 0.032 | 0.924 |
| | 3/4 | 0.868 | 0.059 | 0.073 | 0.808 |
| | 4/5 | 0.794 | 0.088 | 0.119 | 0.723 |
| | 5/6 | 0.802 | 0.117 | 0.081 | 0.730 |

## 8.3.2    Reading 3-5



**Figure 8.3.2A**
**Scale Scores: Read 3-5 S302**



**Figure 8.3.2B**
**Proficiency Level: Read 3-5 S302**
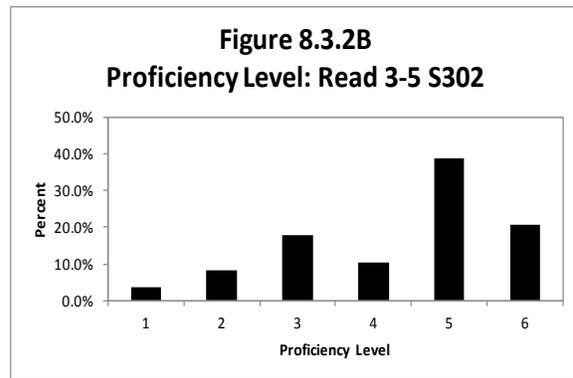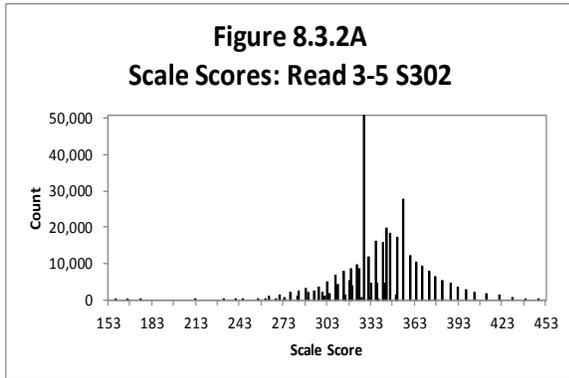
**Table 8.3.2A**

Scale Score Descriptive Statistics: Read 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 3 | 164,072 | 158 | 448 | 331.68 | 25.63 |
| 4 | 105,032 | 166 | 448 | 340.84 | 27.35 |
| 5 | 82,838 | 175 | 448 | 349.48 | 29.27 |
| Total | 351,942 | 158 | 448 | 338.60 | 27.98 |

**Table 8.3.2B**

Proficiency Level Distribution: Read 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|-------|---------|
| 1 | 3,660 | 2.2% | 4,475 | 4.3% | 4,649 | 5.6% | 12,784 | 3.6% |
| 2 | 10,157 | 6.2% | 9,743 | 9.3% | 9,333 | 11.3% | 29,233 | 8.3% |
| 3 | 24,294 | 14.8% | 18,439 | 17.6% | 20,440 | 24.7% | 63,173 | 17.9% |
| 4 | 13,671 | 8.3% | 16,467 | 15.7% | 6,826 | 8.2% | 36,964 | 10.5% |
| 5 | 76,417 | 46.6% | 33,950 | 32.3% | 26,484 | 32.0% | 136,851 | 38.9% |
| 6 | 35,873 | 21.9% | 21,958 | 20.9% | 15,106 | 18.2% | 72,937 | 20.7% |
| Total | 164,072 | 100.0% | 105,032 | 100.0% | 82,838 | 100.0% | 351,942 | 100.0% |

**Table 8.3.2C**
Conditional Standard Error of Measurement at Cut Scores: Read 3-5 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 3 | 279 | 12.48 | 15.34 | 22.88 |
| | 4 | 291 | 11.96 | 13.52 | 19.24 |
| | 5 | 302 | 11.70 | 12.48 | 16.38 |
| 2/3 | 3 | 302 | 11.70 | 12.48 | 16.38 |
| | 4 | 316 | 11.96 | 11.44 | 14.04 |
| | 5 | 328 | 12.48 | 10.92 | 12.48 |
| 3/4 | 3 | 320 | 11.96 | 11.18 | 13.52 |
| | 4 | 336 | 13.00 | 10.92 | 11.70 |
| | 5 | 350 | 14.56 | 11.18 | 10.92 |
| 4/5 | 3 | 328 | n/a | 10.92 | 12.48 |
| | 4 | 343 | n/a | 10.92 | 11.18 |
| | 5 | 355 | n/a | 11.44 | 10.66 |
| 5/6 | 3 | 347 | n/a | n/a | 10.92 |
| | 4 | 360 | n/a | n/a | 10.66 |
| | 5 | 372 | n/a | n/a | 10.66 |

* No equating was performed for S302

**Figure 8.3.2C**
Test Characteristic Curve: Read 3-5ABC S302
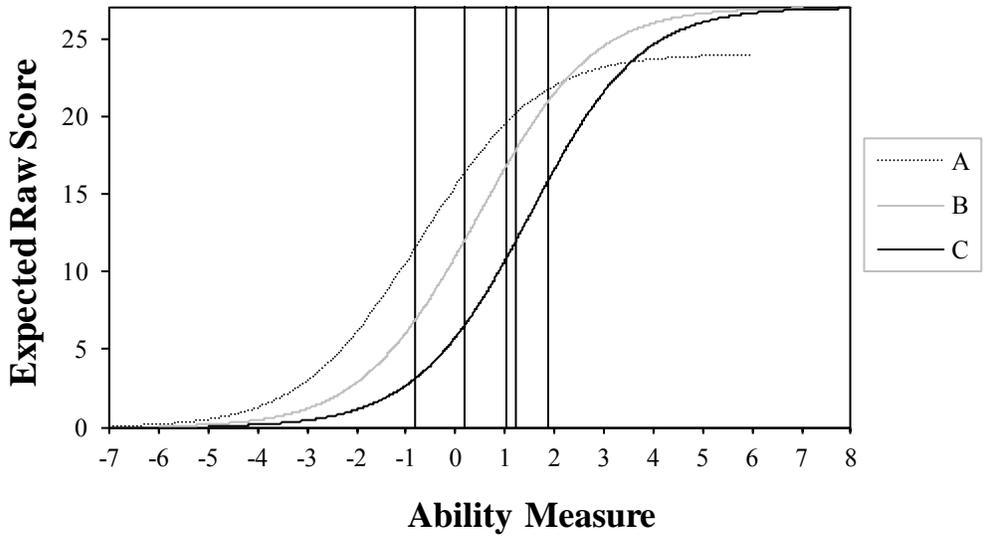


**Figure 8.3.2D**
Test Information Function: Read 3-5ABC S302
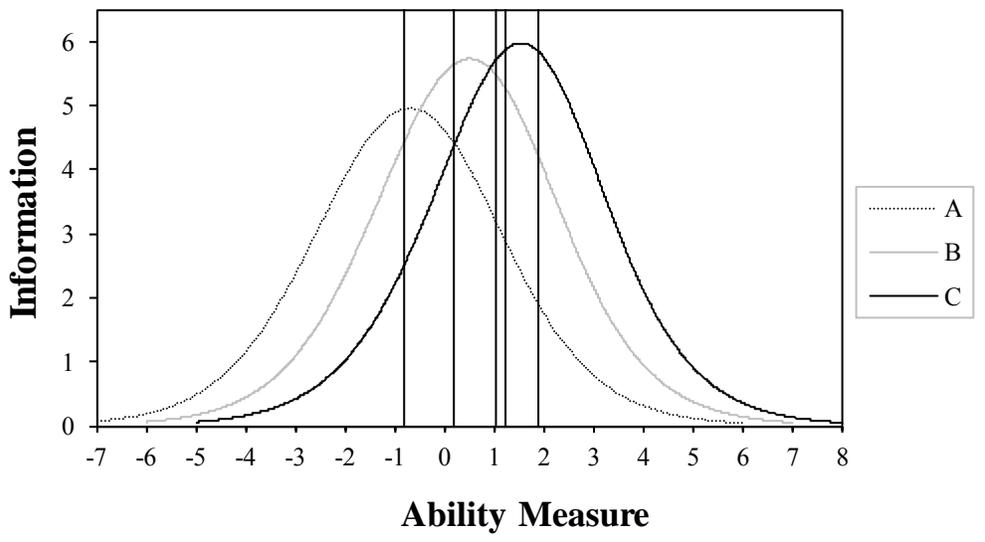
**Table 8.3.2D**

Weighted Reliability: Read 3-5 S302

| Tiers | No. of Students | Reliability | Weighted Reliability |
|---|---|---|---|
| A | 29,169 | 0.838 | |
| B | 148,739 | 0.805 | 0.779 |
| C | 174,034 | 0.748 | |

**Table 8.3.2E-1**

Accuracy and Consistency of Classification Indices: Read (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.567 | 0.450 | | 0.260 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.770 | | 0.540 | |
| | 2 | 0.526 | | 0.348 | |
| | 3 | 0.490 | | 0.341 | |
| | 4 | 0.182 | | 0.134 | |
| | 5 | 0.691 | | 0.605 | |
| | 6 | 0.597 | | 0.478 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.986 | 0.004 | 0.010 | 0.980 |
| | 2/3 | 0.953 | 0.020 | 0.028 | 0.926 |
| | 3/4 | 0.885 | 0.053 | 0.062 | 0.829 |
| | 4/5 | 0.840 | 0.093 | 0.067 | 0.777 |
| | 5/6 | 0.832 | 0.063 | 0.105 | 0.768 |

**Table 8.3.2E-2**

Accuracy and Consistency of Classification Indices: Read (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.505 | 0.400 | | 0.240 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.795 | | 0.608 | |
| | 2 | 0.549 | | 0.380 | |
| | 3 | 0.446 | | 0.329 | |
| | 4 | 0.297 | | 0.229 | |
| | 5 | 0.507 | | 0.432 | |
| | 6 | 0.599 | | 0.463 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.977 | 0.007 | 0.016 | 0.967 |
| | 2/3 | 0.936 | 0.026 | 0.038 | 0.902 |
| | 3/4 | 0.851 | 0.079 | 0.069 | 0.792 |
| | 4/5 | 0.809 | 0.080 | 0.112 | 0.749 |
| | 5/6 | 0.836 | 0.075 | 0.090 | 0.772 |

Return to Visual Table of Tables and Figures

**Table 8.3.2E-3**

Accuracy and Consistency of Classification Indices: Read (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.516 | 0.408 | | 0.251 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.789 | | 0.592 | |
| | 2 | 0.509 | | 0.358 | |
| | 3 | 0.515 | | 0.401 | |
| | 4 | 0.151 | | 0.118 | |
| | 5 | 0.534 | | 0.449 | |
| | 6 | 0.593 | | 0.447 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.969 | 0.009 | 0.022 | 0.954 |
| | 2/3 | 0.916 | 0.038 | 0.046 | 0.873 |
| | 3/4 | 0.829 | 0.086 | 0.085 | 0.767 |
| | 4/5 | 0.813 | 0.095 | 0.092 | 0.751 |
| | 5/6 | 0.853 | 0.069 | 0.078 | 0.793 |

## 8.3.3    Writing 3-5



**Figure 8.3.3A**
**Scale Scores: Writ 3-5 S302**



**Figure 8.3.3B**
**Proficiency Level: Writ 3-5 S302**
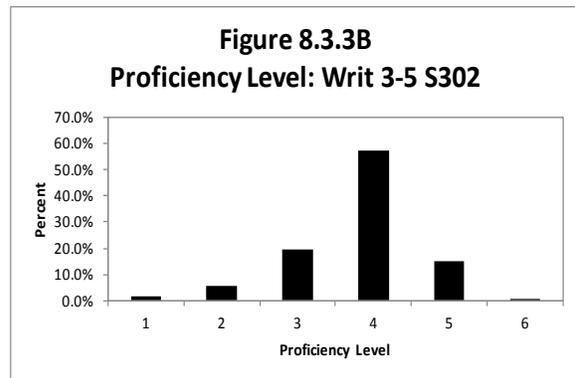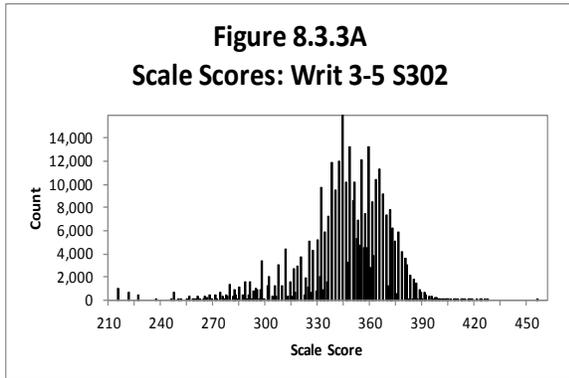
**Table 8.3.3A**

Scale Score Descriptive Statistics: Writ 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 3 | 164,071 | 215 | 409 | 340.82 | 25.58 |
| 4 | 105,084 | 221 | 427 | 346.85 | 25.47 |
| 5 | 82,870 | 227 | 456 | 352.81 | 25.34 |
| Total | 352,025 | 215 | 456 | 345.45 | 25.95 |

**Table 8.3.3B**

Proficiency Level Distribution: Writ 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|-------|---------|
| 1 | 1,982 | 1.2% | 1,857 | 1.8% | 2,031 | 2.5% | 5,870 | 1.7% |
| 2 | 8,428 | 5.1% | 6,593 | 6.3% | 5,126 | 6.2% | 20,147 | 5.7% |
| 3 | 26,453 | 16.1% | 20,757 | 19.8% | 21,928 | 26.5% | 69,138 | 19.6% |
| 4 | 94,135 | 57.4% | 60,918 | 58.0% | 46,976 | 56.7% | 202,029 | 57.4% |
| 5 | 31,981 | 19.5% | 14,553 | 13.8% | 6,646 | 8.0% | 53,180 | 15.1% |
| 6 | 1,092 | 0.7% | 406 | 0.4% | 163 | 0.2% | 1,661 | 0.5% |
| Total | 164,071 | 100.0% | 105,084 | 100.0% | 82,870 | 100.0% | 352,025 | 100.0% |

**Table 8.3.3C**
Conditional Standard Error of Measurement at Cut Scores: Writ 3-5 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 3 | 264 | 10.88 | 7.77 | 11.82 |
| | 4 | 275 | 9.02 | 6.53 | 8.09 |
| | 5 | 287 | 9.64 | 6.84 | 6.53 |
| 2/3 | 3 | 297 | 11.19 | 7.77 | 6.53 |
| | 4 | 308 | 11.82 | 8.40 | 7.46 |
| | 5 | 319 | 11.82 | 8.40 | 8.40 |
| 3/4 | 3 | 330 | 11.51 | 8.09 | 8.40 |
| | 4 | 340 | 11.51 | 8.09 | 8.40 |
| | 5 | 350 | 11.19 | 7.77 | 8.09 |
| 4/5 | 3 | 360 | n/a | 7.46 | 7.77 |
| | 4 | 371 | n/a | 7.15 | 7.77 |
| | 5 | 381 | n/a | 6.84 | 7.15 |
| 5/6 | 3 | 384 | n/a | n/a | 7.15 |
| | 4 | 394 | n/a | n/a | 6.84 |
| | 5 | 403 | n/a | n/a | 6.53 |

* No equating was performed for Writing Tier A S302

**Figure 8.3.3C**
Test Characteristic Curve: Writ 3-5ABC S302
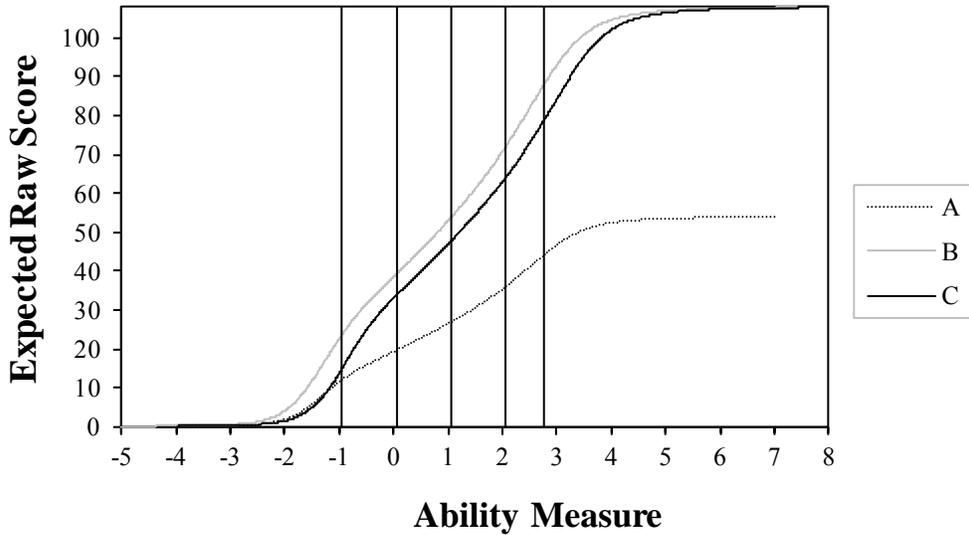


**Figure 8.3.3D**
Test Information Function: Writ 3-5ABC S302
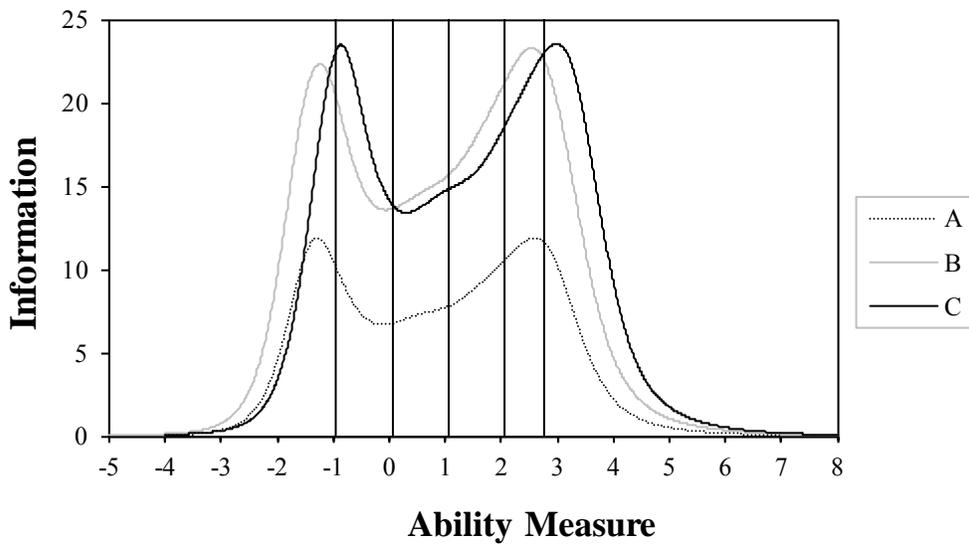
**Table 8.3.3D**
Weighted Reliability: Writ 3-5 S302

| Tiers | No. of Students | Reliability | Weighted Reliability |
|---|---|---|---|
| A | 29,175 | 0.918 | |
| B | 148,840 | 0.935 | 0.924 |
| C | 174,010 | 0.915 | |

**Table 8.3.3E-3**
Accuracy and Consistency of Classification Indices: Writ (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.742 | 0.670 | | 0.461 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.785 | | 0.756 | |
| | 2 | 0.811 | | 0.716 | |
| | 3 | 0.799 | | 0.702 | |
| | 4 | 0.826 | | 0.750 | |
| | 5 | 0.481 | | 0.426 | |
| | 6 | - | | 0.973 | |
| **Indices at** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.995 | 0.003 | 0.002 | 0.994 |
| | 2/3 | 0.983 | 0.008 | 0.009 | 0.977 |
| | 3/4 | 0.949 | 0.024 | 0.027 | 0.928 |
| | 4/5 | 0.834 | 0.070 | 0.097 | 0.774 |
| | 5/6 | 0.993 | 0.007 | 0.000 | 0.994 |

**Table 8.3.3E-2**
Accuracy and Consistency of Classification Indices: Writ (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.779 | 0.717 | | 0.507 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.747 | | 0.784 | |
| | 2 | 0.812 | | 0.715 | |
| | 3 | 0.839 | | 0.749 | |
| | 4 | 0.762 | | 0.758 | |
| | 5 | - | | 0.304 | |
| | 6 | - | | 1.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.992 | 0.005 | 0.003 | 0.993 |
| | 2/3 | 0.978 | 0.012 | 0.010 | 0.973 |
| | 3/4 | 0.944 | 0.022 | 0.034 | 0.924 |
| | 4/5 | 0.858 | 0.142 | 0.000 | 0.828 |
| | 5/6 | 0.996 | 0.004 | 0.000 | 0.999 |

**Table 8.3.3E-3**

Accuracy and Consistency of Classification Indices: Writ (Grade 5) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| Indices | 0.820 | 0.766 | | 0.588 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.837 | | 0.835 | |
| | 2 | 0.768 | | 0.658 | |
| | 3 | 0.868 | | 0.785 | |
| | 4 | 0.808 | | 0.792 | |
| | 5 | - | | 0.198 | |
| | 6 | - | | 1.000 | |
| **Indices at** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.992 | 0.004 | 0.003 | 0.992 |
| | 2/3 | 0.977 | 0.013 | 0.010 | 0.968 |
| | 3/4 | 0.929 | 0.023 | 0.048 | 0.903 |
| | 4/5 | 0.918 | 0.082 | 0.000 | 0.903 |
| | 5/6 | 0.998 | 0.002 | 0.000 | 1.000 |

## 8.3.4    Speaking 3-5



**Figure 8.3.4A
Scale Scores: Spek 3-5 S302**



**Figure 8.3.4B
Proficiency Level: Spek 3-5 S302**
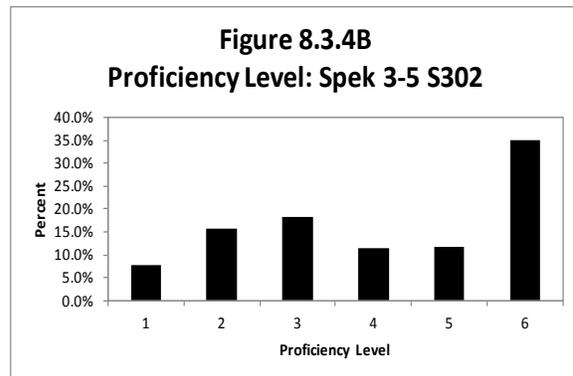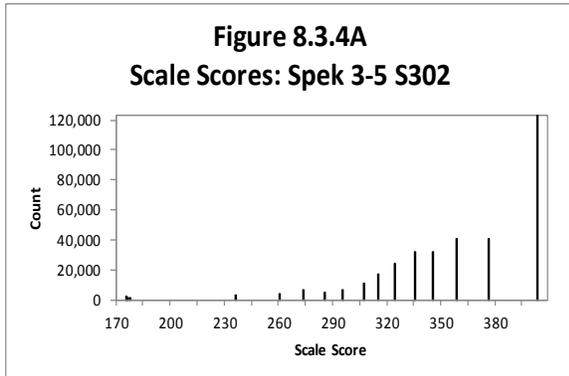
**Table 8.3.4A**

Scale Score Descriptive Statistics: Spek 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 3     | 164,141         | 175  | 403  | 355.70 | 44.80     |
| 4     | 105,094         | 176  | 403  | 358.06 | 45.60     |
| 5     | 82,870          | 177  | 403  | 361.87 | 46.17     |
| Total | 352,105         | 175  | 403  | 357.86 | 45.43     |

**Table 8.3.4B**

Proficiency Level Distribution: Spek 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1     | 11,295  | 6.9%    | 9,217   | 8.8%    | 7,066   | 8.5%    | 27,578  | 7.8%    |
| 2     | 30,686  | 18.7%   | 15,070  | 14.3%   | 9,789   | 11.8%   | 55,545  | 15.8%   |
| 3     | 31,833  | 19.4%   | 18,828  | 17.9%   | 13,372  | 16.1%   | 64,033  | 18.2%   |
| 4     | 18,702  | 11.4%   | 12,598  | 12.0%   | 9,301   | 11.2%   | 40,601  | 11.5%   |
| 5     | 18,485  | 11.3%   | 12,654  | 12.0%   | 9,990   | 12.1%   | 41,129  | 11.7%   |
| 6     | 53,140  | 32.4%   | 36,727  | 34.9%   | 33,352  | 40.2%   | 123,219 | 35.0%   |
| Total | 164,141 | 100.0%  | 105,094 | 100.0%  | 82,870  | 100.0%  | 352,105 | 100.0%  |

**Table 8.3.4C**
Conditional Standard Error of Measurement at
Cut Scores: Spek 3-5 S302

| Proficiency Level | Grade | Cut Score | SEM |
|---|---|---|---|
| 1/2 | 3 | 293 | 19.08 |
| | 4 | 299 | 19.48 |
| | 5 | 305 | 19.68 |
| 2/3 | 3 | 326 | 20.89 |
| | 4 | 329 | 21.09 |
| | 5 | 333 | 21.49 |
| 3/4 | 3 | 346 | 22.29 |
| | 4 | 348 | 22.49 |
| | 5 | 350 | 22.69 |
| 4/5 | 3 | 369 | 24.90 |
| | 4 | 371 | 25.31 |
| | 5 | 374 | 25.71 |
| 5/6 | 3 | 389 | 27.31 |
| | 4 | 391 | 27.52 |
| | 5 | 394 | 27.52 |

**Figure 8.3.4C**

Test Characteristic Curve: Spek 3-5 S302



Expected Raw Score vs. Ability Measure
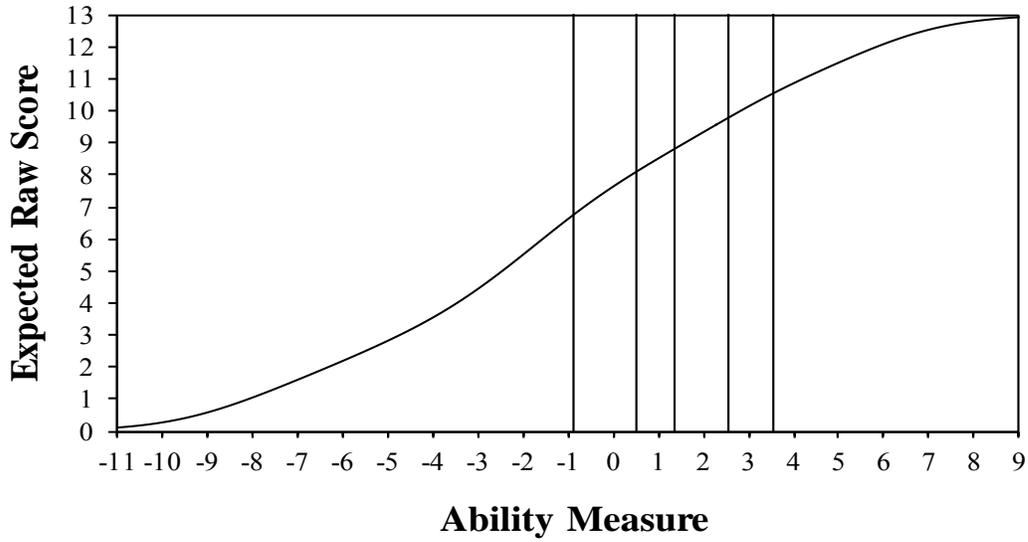
**Figure 8.3.4D**

Test Information Function: Spek 3-5 S302



Information vs. Ability Measure
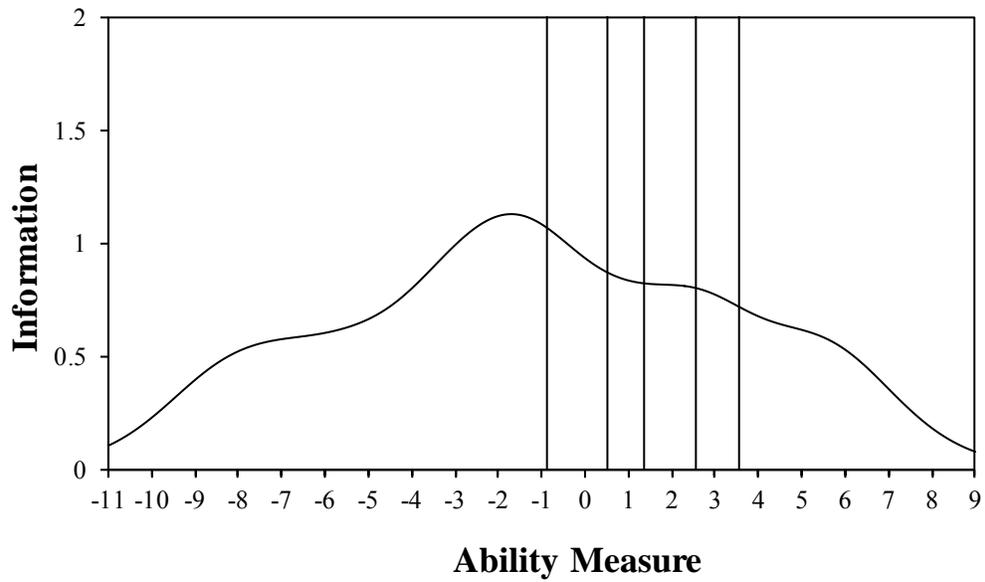
**Table 8.3.4D**

Reliability: Spek 3-5 S302

| Tiers | No. of Students | Reliability |
|---|---|---|
| -- | 352,105 | 0.891 |

**Table 8.3.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade 3) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| Indices | 0.603 | 0.512 | | 0.397 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.650 | | 0.493 | |
| | 2 | 0.614 | | 0.506 | |
| | 3 | 0.519 | | 0.431 | |
| | 4 | 0.353 | | 0.261 | |
| | 5 | 0.379 | | 0.274 | |
| | 6 | 0.915 | | 0.844 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.955 | 0.028 | 0.017 | 0.930 |
| | 2/3 | 0.890 | 0.050 | 0.060 | 0.857 |
| | 3/4 | 0.890 | 0.029 | 0.082 | 0.851 |
| | 4/5 | 0.930 | 0.033 | 0.037 | 0.887 |
| | 5/6 | 0.903 | 0.073 | 0.023 | 0.872 |

**Table 8.3.4E-2**

Accuracy and Consistency of Classification Indices: Spek (Grade 4) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| Indices | 0.591 | 0.501 | | 0.380 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.718 | | 0.576 | |
| | 2 | 0.524 | | 0.420 | |
| | 3 | 0.513 | | 0.426 | |
| | 4 | 0.360 | | 0.269 | |
| | 5 | 0.350 | | 0.247 | |
| | 6 | 0.901 | | 0.825 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.954 | 0.027 | 0.019 | 0.929 |
| | 2/3 | 0.898 | 0.050 | 0.052 | 0.868 |
| | 3/4 | 0.890 | 0.029 | 0.082 | 0.854 |
| | 4/5 | 0.923 | 0.033 | 0.044 | 0.877 |
| | 5/6 | 0.884 | 0.087 | 0.029 | 0.844 |

**Table 8.3.4E-3**

Accuracy and Consistency of Classification Indices: Spek (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.576 | 0.480 | | 0.345 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.716 | | 0.578 | |
| | 2 | 0.475 | | 0.378 | |
| | 3 | 0.505 | | 0.420 | |
| | 4 | 0.347 | | 0.259 | |
| | 5 | 0.297 | | 0.206 | |
| | 6 | 0.886 | | 0.812 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.956 | 0.027 | 0.018 | 0.932 |
| | 2/3 | 0.905 | 0.049 | 0.046 | 0.879 |
| | 3/4 | 0.893 | 0.026 | 0.082 | 0.862 |
| | 4/5 | 0.926 | 0.031 | 0.043 | 0.879 |
| | 5/6 | 0.848 | 0.115 | 0.037 | 0.792 |

## 8.3.5    Oral 3-5



Figure 8.3.5A
Scale Scores: Oral 3-5 S302



Figure 8.3.5B
Proficiency Level: Oral 3-5 S302
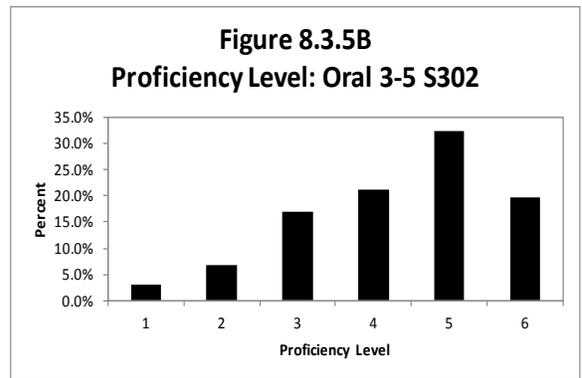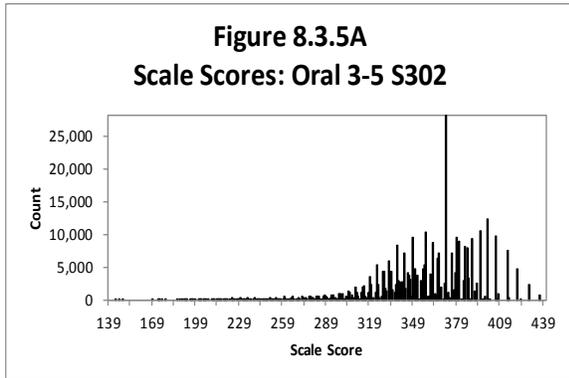
**Table 8.3.5A**

Scale Score Descriptive Statistics: Oral 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 3 | 164,066 | 144 | 436 | 352.84 | 34.11 |
| 4 | 105,043 | 146 | 436 | 359.63 | 35.50 |
| 5 | 82,831 | 149 | 436 | 366.65 | 36.84 |
| Total | 351,940 | 144 | 436 | 358.12 | 35.62 |

**Table 8.3.5B**

Proficiency Level Distribution: Oral 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|---------|---------|
| 1 | 3,809 | 2.3% | 3,282 | 3.1% | 3,454 | 4.2% | 10,545 | 3.0% |
| 2 | 11,130 | 6.8% | 6,802 | 6.5% | 5,787 | 7.0% | 23,719 | 6.7% |
| 3 | 29,982 | 18.3% | 17,709 | 16.9% | 11,681 | 14.1% | 59,372 | 16.9% |
| 4 | 35,592 | 21.7% | 22,292 | 21.2% | 16,789 | 20.3% | 74,673 | 21.2% |
| 5 | 49,079 | 29.9% | 34,075 | 32.4% | 30,801 | 37.2% | 113,955 | 32.4% |
| 6 | 34,474 | 21.0% | 20,883 | 19.9% | 14,319 | 17.3% | 69,676 | 19.8% |
| Total | 164,066 | 100.0% | 105,043 | 100.0% | 82,831 | 100.0% | 351,940 | 100.0% |

**Table 8.3.5C**

n/a

**Figure 8.3.5C**

n/a

**Figure 8.3.5D**

n/a

**Table 8.3.5D**

Oral Composite Reliability: Oral 3-5 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.50 | 1377.820 | 0.657 |
| Speaking | 0.50 | 2058.602 | 0.891 |
| Oral | | 1266.567 | 0.863 |

**Table 8.3.5E-1**

Accuracy and Consistency of Classification Indices: Oral (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.612 | 0.497 | | 0.358 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.824 | | 0.657 | |
| | 2 | 0.586 | | 0.429 | |
| | 3 | 0.619 | | 0.487 | |
| | 4 | 0.506 | | 0.402 | |
| | 5 | 0.596 | | 0.492 | |
| | 6 | 0.727 | | 0.604 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.989 | 0.003 | 0.008 | 0.984 |
| | 2/3 | 0.956 | 0.023 | 0.022 | 0.933 |
| | 3/4 | 0.896 | 0.044 | 0.060 | 0.857 |
| | 4/5 | 0.870 | 0.056 | 0.074 | 0.819 |
| | 5/6 | 0.885 | 0.057 | 0.058 | 0.838 |

**Table 8.3.5E-2**
Accuracy and Consistency of Classification Indices: Oral (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.591 | 0.475 | | 0.326 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.836 | | 0.690 | |
| | 2 | 0.557 | | 0.405 | |
| | 3 | 0.609 | | 0.475 | |
| | 4 | 0.507 | | 0.397 | |
| | 5 | 0.588 | | 0.492 | |
| | 6 | 0.649 | | 0.517 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.986 | 0.004 | 0.009 | 0.980 |
| | 2/3 | 0.957 | 0.023 | 0.020 | 0.935 |
| | 3/4 | 0.902 | 0.040 | 0.058 | 0.865 |
| | 4/5 | 0.869 | 0.052 | 0.078 | 0.818 |
| | 5/6 | 0.860 | 0.071 | 0.069 | 0.809 |

**Table 8.3.5E-3**
Accuracy and Consistency of Classification Indices: Oral (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.581 | 0.468 | | 0.308 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.844 | | 0.714 | |
| | 2 | 0.561 | | 0.415 | |
| | 3 | 0.558 | | 0.425 | |
| | 4 | 0.500 | | 0.384 | |
| | 5 | 0.616 | | 0.538 | |
| | 6 | 0.560 | | 0.424 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.983 | 0.006 | 0.011 | 0.976 |
| | 2/3 | 0.955 | 0.022 | 0.022 | 0.934 |
| | 3/4 | 0.910 | 0.037 | 0.052 | 0.876 |
| | 4/5 | 0.869 | 0.054 | 0.077 | 0.816 |
| | 5/6 | 0.847 | 0.082 | 0.071 | 0.797 |

## 8.3.6    Literacy Composite 3-5



**Figure 8.3.6A**
**Scale Scores: Litr 3-5 S302**



**Figure 8.3.6B**
**Proficiency Level: Litr 3-5 S302**
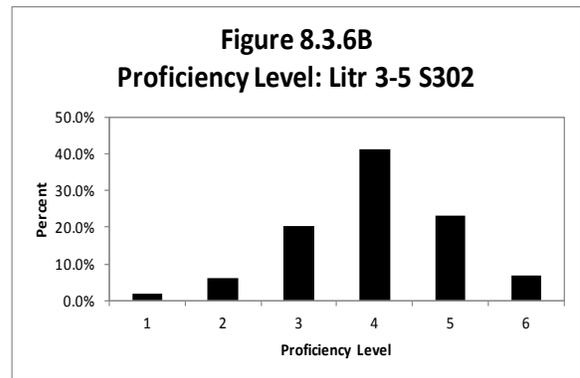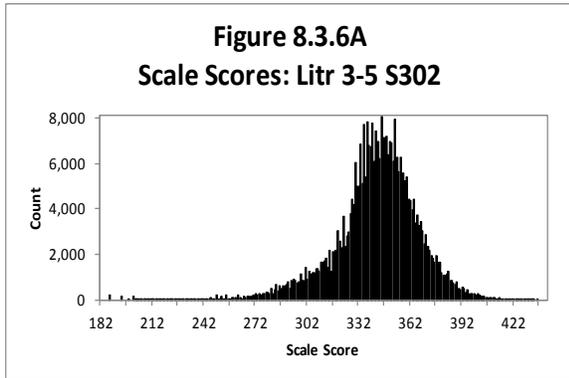
**Table 8.3.6A**
Scale Score Descriptive Statistics: Litr 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 3 | 163,916 | 187 | 422 | 336.47 | 23.14 |
| 4 | 104,940 | 194 | 433 | 344.14 | 24.14 |
| 5 | 82,761 | 201 | 436 | 351.43 | 25.19 |
| **Total** | 351,617 | 187 | 436 | 342.28 | 24.68 |

**Table 8.3.6B**
Proficiency Level Distribution: Litr 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|--------|---------|
| 1 | 1,908 | 1.2% | 2,137 | 2.0% | 2,772 | 3.3% | 6,817 | 1.9% |
| 2 | 8,438 | 5.1% | 6,992 | 6.7% | 6,543 | 7.9% | 21,973 | 6.2% |
| 3 | 26,975 | 16.5% | 21,617 | 20.6% | 23,138 | 28.0% | 71,730 | 20.4% |
| 4 | 66,967 | 40.9% | 46,272 | 44.1% | 31,950 | 38.6% | 145,189 | 41.3% |
| 5 | 46,008 | 28.1% | 21,384 | 20.4% | 14,182 | 17.1% | 81,574 | 23.2% |
| 6 | 13,620 | 8.3% | 6,538 | 6.2% | 4,176 | 5.0% | 24,334 | 6.9% |
| Total | 163,916 | 100.0% | 104,940 | 100.0% | 82,761 | 100.0% | 351,617 | 100.0% |

**Table 8.3.6C**
n/a


**Figure 8.3.6C**
n/a

**Figure 8.3.6D**
n/a


**Table 8.3.6D**
Literacy Composite Reliability: Litr 3-5 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Reading | 0.50 | 781.583 | 0.779 |
| Writing | 0.50 | 672.252 | 0.924 |
| Literacy | | 608.817 | 0.908 |

*Variances from students who had results in all four domains


**Table 8.3.6E-1**
Accuracy and Consistency of Classification Indices: Litr (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.705 | 0.617 | | 0.471 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.846 | | 0.725 | |
| | 2 | 0.759 | | 0.636 | |
| | 3 | 0.698 | | 0.575 | |
| | 4 | 0.808 | | 0.723 | |
| | 5 | 0.606 | | 0.556 | |
| | 6 | 0.714 | | 0.472 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.995 | 0.002 | 0.003 | 0.993 |
| | 2/3 | 0.979 | 0.009 | 0.012 | 0.970 |
| | 3/4 | 0.926 | 0.041 | 0.033 | 0.893 |
| | 4/5 | 0.885 | 0.036 | 0.079 | 0.842 |
| | 5/6 | 0.919 | 0.079 | 0.002 | 0.914 |

**Table 8.3.6E-2**
Accuracy and Consistency of Classification Indices: Litr (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.691 | 0.596 | | 0.444 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.863 | | 0.763 | |
| | 2 | 0.752 | | 0.632 | |
| | 3 | 0.724 | | 0.607 | |
| | 4 | 0.802 | | 0.708 | |
| | 5 | 0.520 | | 0.455 | |
| | 6 | - | | 0.336 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.993 | 0.003 | 0.005 | 0.990 |
| | 2/3 | 0.974 | 0.012 | 0.014 | 0.962 |
| | 3/4 | 0.915 | 0.044 | 0.041 | 0.878 |
| | 4/5 | 0.870 | 0.034 | 0.095 | 0.821 |
| | 5/6 | 0.938 | 0.062 | 0.000 | 0.932 |

**Table 8.3.6E-3**
Accuracy and Consistency of Classification Indices: Litr (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.688 | 0.588 | | 0.445 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.881 | | 0.794 | |
| | 2 | 0.695 | | 0.566 | |
| | 3 | 0.763 | | 0.659 | |
| | 4 | 0.743 | | 0.641 | |
| | 5 | 0.513 | | 0.439 | |
| | 6 | - | | 0.334 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.989 | 0.004 | 0.007 | 0.985 |
| | 2/3 | 0.964 | 0.018 | 0.017 | 0.948 |
| | 3/4 | 0.899 | 0.047 | 0.054 | 0.858 |
| | 4/5 | 0.885 | 0.037 | 0.078 | 0.837 |
| | 5/6 | 0.950 | 0.050 | 0.000 | 0.946 |

## 8.3.7 Comprehension Composite 3-5



**Figure 8.3.7A**
**Scale Scores: Cphn 3-5 S302**



**Figure 8.3.7B**
**Proficiency Level: Cphn 3-5 S302**
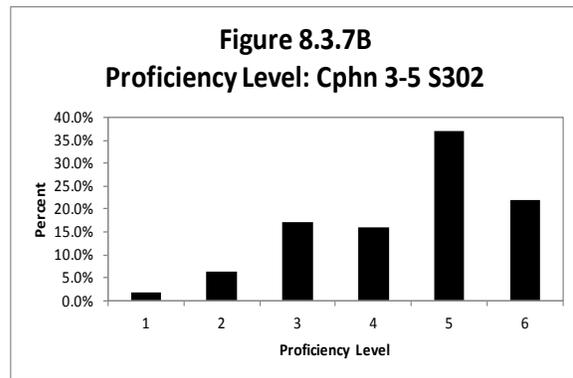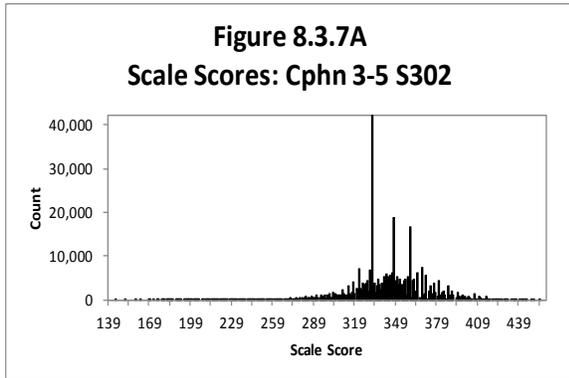
**Table 8.3.7A**

Scale Score Descriptive Statistics: Cphn 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 3 | 164,030 | 144 | 454 | 337.18 | 26.39 |
| 4 | 104,998 | 151 | 454 | 346.95 | 27.94 |
| 5 | 82,804 | 159 | 454 | 355.99 | 29.73 |
| Total | 351,832 | 144 | 454 | 344.52 | 28.70 |

**Table 8.3.7B**

Proficiency Level Distribution: Cphn 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1 | 1,507 | 0.9% | 1,868 | 1.8% | 2,516 | 3.0% | 5,891 | 1.7% |
| 2 | 8,114 | 4.9% | 7,035 | 6.7% | 6,794 | 8.2% | 21,943 | 6.2% |
| 3 | 21,263 | 13.0% | 20,392 | 19.4% | 18,307 | 22.1% | 59,962 | 17.0% |
| 4 | 24,724 | 15.1% | 18,647 | 17.8% | 13,425 | 16.2% | 56,796 | 16.1% |
| 5 | 67,993 | 41.5% | 35,685 | 34.0% | 26,499 | 32.0% | 130,177 | 37.0% |
| 6 | 40,429 | 24.6% | 21,371 | 20.4% | 15,263 | 18.4% | 77,063 | 21.9% |
| Total | 164,030 | 100.0% | 104,998 | 100.0% | 82,804 | 100.0% | 351,832 | 100.0% |

**Table 8.3.7C**

n/a

**Figure 8.3.7C**

n/a

**Figure 8.3.7D**

n/a

**Table 8.3.7D**

Comprehension Composite Reliability: Cphn 3-5 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.30 | 1377.820 | 0.657 |
| Reading | 0.70 | 781.583 | 0.779 |
| Comprehension | | 822.711 | 0.845 |

*Variances from students who had results in all four domains

**Table 8.3.7E-1**

Accuracy and Consistency of Classification Indices: Cphn (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.634 | 0.524 | | 0.356 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.781 | | 0.548 | |
| | 2 | 0.653 | | 0.471 | |
| | 3 | 0.533 | | 0.388 | |
| | 4 | 0.379 | | 0.286 | |
| | 5 | 0.691 | | 0.606 | |
| | 6 | 0.739 | | 0.624 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.994 | 0.001 | 0.005 | 0.992 |
| | 2/3 | 0.970 | 0.011 | 0.019 | 0.954 |
| | 3/4 | 0.908 | 0.051 | 0.041 | 0.863 |
| | 4/5 | 0.856 | 0.074 | 0.070 | 0.805 |
| | 5/6 | 0.877 | 0.056 | 0.067 | 0.824 |

**Table 8.3.7E-2**

Accuracy and Consistency of Classification Indices: Cphn (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.597 | 0.485 | | 0.334 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.801 | | 0.613 | |
| | 2 | 0.624 | | 0.454 | |
| | 3 | 0.594 | | 0.462 | |
| | 4 | 0.398 | | 0.309 | |
| | 5 | 0.609 | | 0.516 | |
| | 6 | 0.701 | | 0.570 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.990 | 0.003 | 0.007 | 0.986 |
| | 2/3 | 0.959 | 0.018 | 0.023 | 0.937 |
| | 3/4 | 0.882 | 0.061 | 0.057 | 0.835 |
| | 4/5 | 0.847 | 0.066 | 0.086 | 0.795 |
| | 5/6 | 0.881 | 0.055 | 0.063 | 0.830 |


**Table 8.3.7E-3**

Accuracy and Consistency of Classification Indices: Cphn (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.587 | 0.477 | | 0.332 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.826 | | 0.656 | |
| | 2 | 0.592 | | 0.435 | |
| | 3 | 0.593 | | 0.471 | |
| | 4 | 0.356 | | 0.276 | |
| | 5 | 0.595 | | 0.502 | |
| | 6 | 0.708 | | 0.564 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.985 | 0.004 | 0.012 | 0.978 |
| | 2/3 | 0.949 | 0.024 | 0.027 | 0.922 |
| | 3/4 | 0.870 | 0.068 | 0.062 | 0.820 |
| | 4/5 | 0.846 | 0.068 | 0.086 | 0.793 |
| | 5/6 | 0.891 | 0.056 | 0.053 | 0.842 |

## 8.3.8    Overall Composite 3-5



**Figure 8.3.8A**
**Scale Scores: Over 3-5 S302**



**Figure 8.3.8B**
**Proficiency Level: Over 3-5 S302**
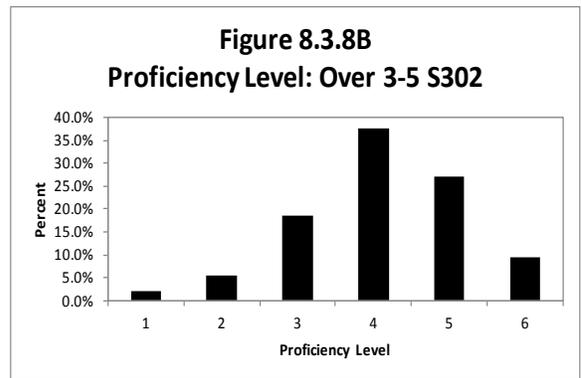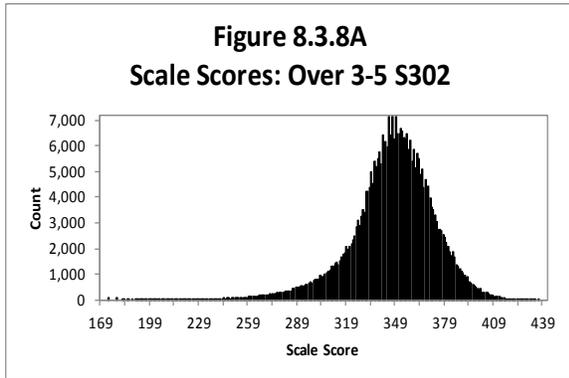
**Table 8.3.8A**

Scale Score Descriptive Statistics: Over 3-5 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 3 | 163,789 | 174 | 426 | 341.21 | 24.49 |
| 4 | 104,822 | 179 | 433 | 348.57 | 25.59 |
| 5 | 82,666 | 185 | 436 | 355.77 | 26.81 |
| Total | 351,277 | 174 | 436 | 346.83 | 26.05 |

**Table 8.3.8B**

Proficiency Level Distribution: Over 3-5 S302

| Level | Grade 3 Count | Grade 3 Percent | Grade 4 Count | Grade 4 Percent | Grade 5 Count | Grade 5 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|---------|---------|
| 1 | 2,152 | 1.3% | 2,358 | 2.2% | 2,643 | 3.2% | 7,153 | 2.0% |
| 2 | 7,783 | 4.8% | 6,201 | 5.9% | 5,566 | 6.7% | 19,550 | 5.6% |
| 3 | 26,956 | 16.5% | 20,187 | 19.3% | 17,803 | 21.5% | 64,946 | 18.5% |
| 4 | 58,649 | 35.8% | 41,058 | 39.2% | 32,075 | 38.8% | 131,782 | 37.5% |
| 5 | 49,480 | 30.2% | 26,434 | 25.2% | 18,894 | 22.9% | 94,808 | 27.0% |
| 6 | 18,769 | 11.5% | 8,584 | 8.2% | 5,685 | 6.9% | 33,038 | 9.4% |
| Total | 163,789 | 100.0% | 104,822 | 100.0% | 82,666 | 100.0% | 351,277 | 100.0% |

**Table 8.3.8C**

n/a

**Figure 8.3.8C**

n/a

**Figure 8.3.8D**

n/a

**Table 8.3.8D**

Overall Composite Reliability: Over 3-5 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.15 | 1377.820 | 0.657 |
| Reading | 0.35 | 781.583 | 0.779 |
| Speaking | 0.15 | 2058.602 | 0.891 |
| Writing | 0.35 | 672.252 | 0.924 |
| Overall Composite | | 678.626 | 0.937 |

**Table 8.3.8E-1**

Accuracy and Consistency of Classification Indices: Over (Grade 3) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.756 | 0.665 | | 0.548 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.897 | | 0.816 | |
| | 2 | 0.764 | | 0.655 | |
| | 3 | 0.751 | | 0.643 | |
| | 4 | 0.818 | | 0.739 | |
| | 5 | 0.695 | | 0.621 | |
| | 6 | 0.774 | | 0.626 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.996 | 0.001 | 0.003 | 0.994 |
| | 2/3 | 0.982 | 0.009 | 0.009 | 0.974 |
| | 3/4 | 0.938 | 0.033 | 0.028 | 0.912 |
| | 4/5 | 0.910 | 0.031 | 0.059 | 0.874 |
| | 5/6 | 0.930 | 0.052 | 0.018 | 0.909 |

**Table 8.3.8E-2**

Accuracy and Consistency of Classification Indices: Over (Grade 4) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.722 | 0.642 | | 0.515 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.900 | | 0.832 | |
| | 2 | 0.759 | | 0.651 | |
| | 3 | 0.761 | | 0.656 | |
| | 4 | 0.825 | | 0.744 | |
| | 5 | 0.592 | | 0.546 | |
| | 6 | - | | 0.451 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.994 | 0.002 | 0.004 | 0.992 |
| | 2/3 | 0.979 | 0.011 | 0.010 | 0.969 |
| | 3/4 | 0.931 | 0.037 | 0.032 | 0.901 |
| | 4/5 | 0.900 | 0.028 | 0.073 | 0.862 |
| | 5/6 | 0.918 | 0.082 | 0.000 | 0.915 |

**Table 8.3.8E-3**

Accuracy and Consistency of Classification Indices: Over (Grade 5) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.720 | 0.635 | | 0.509 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.904 | | 0.841 | |
| | 2 | 0.735 | | 0.623 | |
| | 3 | 0.766 | | 0.662 | |
| | 4 | 0.808 | | 0.720 | |
| | 5 | 0.584 | | 0.529 | |
| | 6 | - | | 0.409 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.992 | 0.003 | 0.006 | 0.989 |
| | 2/3 | 0.974 | 0.014 | 0.012 | 0.963 |
| | 3/4 | 0.925 | 0.038 | 0.037 | 0.893 |
| | 4/5 | 0.897 | 0.029 | 0.074 | 0.858 |
| | 5/6 | 0.931 | 0.069 | 0.000 | 0.928 |

## 8.4　Grades: 6–8

### 8.4.1　Listening 6-8



**Figure 8.4.1A**
**Scale Scores: List 6-8 S302**



**Figure 8.4.1B**
**Proficiency Level: List 6-8 S302**

**Table 8.4.1A**

Scale Score Descriptive Statistics: List 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 6 | 73,891 | 124 | 473 | 377.12 | 41.48 |
| 7 | 74,881 | 128 | 473 | 385.15 | 44.36 |
| 8 | 70,706 | 132 | 473 | 391.41 | 46.32 |
| Total | 219,478 | 124 | 473 | 384.46 | 44.45 |

**Table 8.4.1B**

Proficiency Level Distribution: List 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|-------|------|---------|------|---------|------|---------|------|---------|
| 1 | 1,475 | 2.0% | 2,328 | 3.1% | 3,100 | 4.4% | 6,903 | 3.1% |
| 2 | 6,193 | 8.4% | 7,003 | 9.4% | 7,637 | 10.8% | 20,833 | 9.5% |
| 3 | 11,788 | 16.0% | 12,372 | 16.5% | 7,539 | 10.7% | 31,699 | 14.4% |
| 4 | 14,320 | 19.4% | 14,390 | 19.2% | 15,333 | 21.7% | 44,043 | 20.1% |
| 5 | 25,821 | 34.9% | 24,829 | 33.2% | 19,773 | 28.0% | 70,423 | 32.1% |
| 6 | 14,294 | 19.3% | 13,959 | 18.6% | 17,324 | 24.5% | 45,577 | 20.8% |
| Total | 73,891 | 100.0% | 74,881 | 100.0% | 70,706 | 100.0% | 219,478 | 100.0% |

**Table 8.4.1C**

Conditional Standard Error of Measurement at Cut Scores: List 6-8 S302

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 6 | 283 | 21.04 | 19.54 | 23.67 |
| | 7 | 293 | 20.66 | 18.79 | 21.79 |
| | 8 | 302 | 20.29 | 18.41 | 20.66 |
| 2/3 | 6 | 328 | 20.29 | 18.03 | 18.41 |
| | 7 | 337 | 20.66 | 18.41 | 18.03 |
| | 8 | 345 | 21.04 | 18.79 | 17.66 |
| 3/4 | 6 | 359 | 22.17 | 19.54 | 17.66 |
| | 7 | 368 | 23.29 | 20.29 | 18.03 |
| | 8 | 375 | 24.05 | 21.04 | 18.03 |
| 4/5 | 6 | 380 | n/a | 21.79 | 18.41 |
| | 7 | 390 | n/a | 23.29 | 19.16 |
| | 8 | 399 | n/a | 24.42 | 19.91 |
| 5/6 | 6 | 409 | n/a | n/a | 21.04 |
| | 7 | 418 | n/a | n/a | 22.54 |
| | 8 | 426 | n/a | n/a | 23.67 |

**Figure 8.4.1C**
Test Characteristic Curve: List 6-8ABC S302
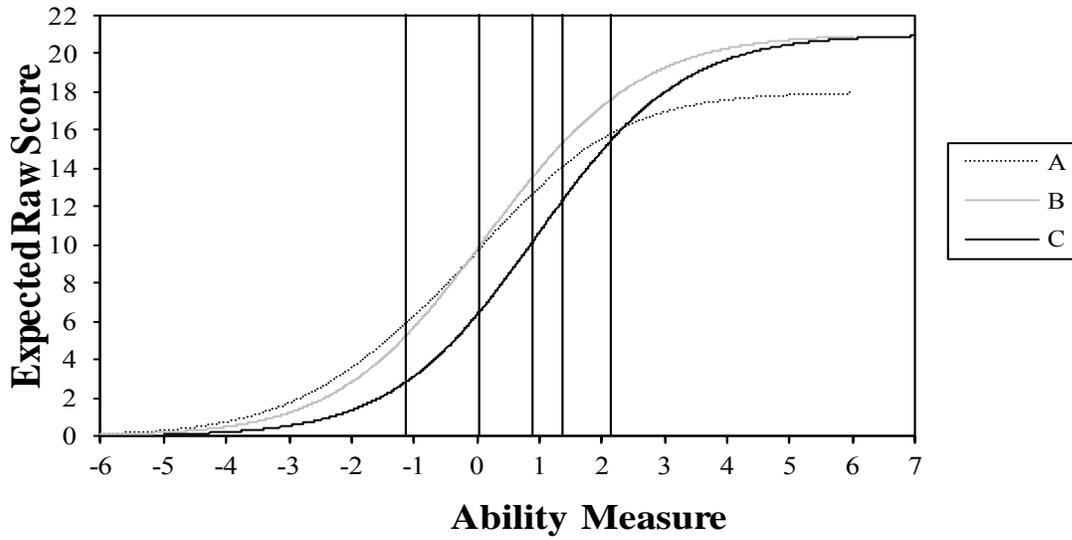


**Figure 8.4.1D**
Test Information Function: List 6-8ABC S302
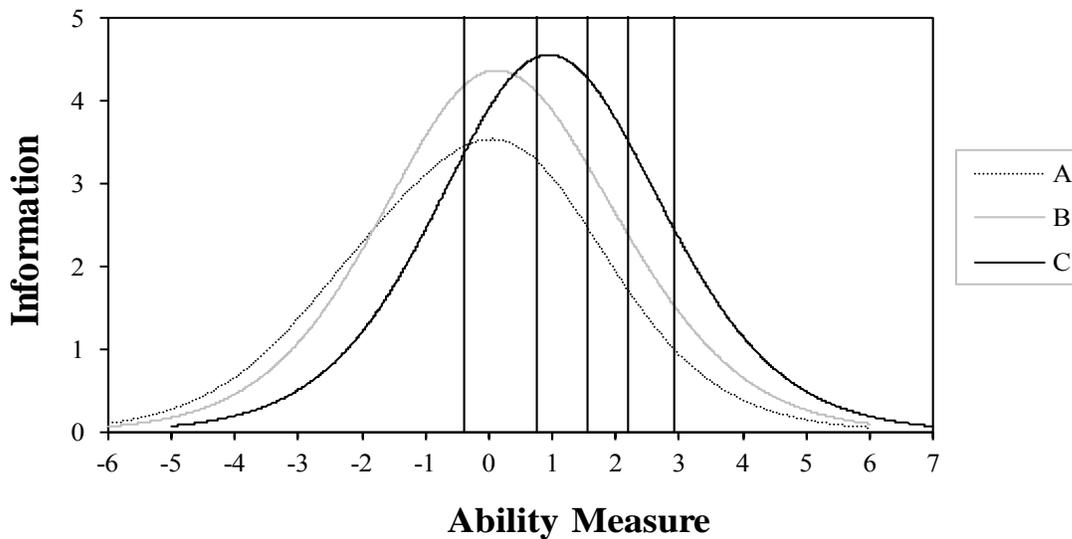
**Table 8.4.1D**

Weighted Reliability: List 6-8 S302

| Tiers | No. of Students | Reliability | Reliability |
|---|---|---|---|
| A | 23,751 | 0.750 | |
| B | 82,373 | 0.663 | 0.645 |
| C | 113,354 | 0.611 | |

**Table 8.4.1E-1**

Accuracy and Consistency of Classification Indices: List (Grade 6) S302

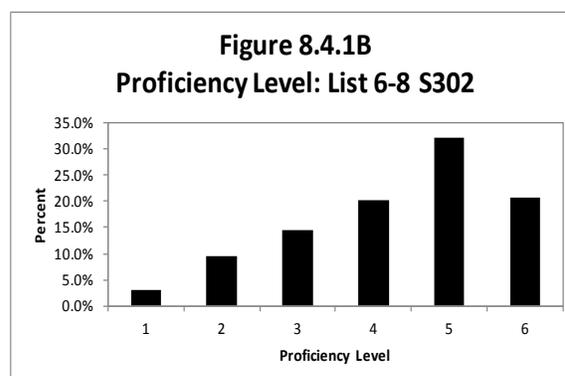| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.446 | 0.343 | | 0.155 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.671 | | 0.376 | |
| | 2 | 0.519 | | 0.331 | |
| | 3 | 0.406 | | 0.286 | |
| | 4 | 0.314 | | 0.242 | |
| | 5 | 0.480 | | 0.425 | |
| | 6 | 0.512 | | 0.353 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.984 | 0.004 | 0.012 | 0.974 |
| | 2/3 | 0.930 | 0.023 | 0.047 | 0.895 |
| | 3/4 | 0.848 | 0.058 | 0.093 | 0.784 |
| | 4/5 | 0.777 | 0.096 | 0.128 | 0.698 |
| | 5/6 | 0.810 | 0.124 | 0.067 | 0.739 |

**Table 8.4.1E-2**
Accuracy and Consistency of Classification Indices: List (Grade 7) S302

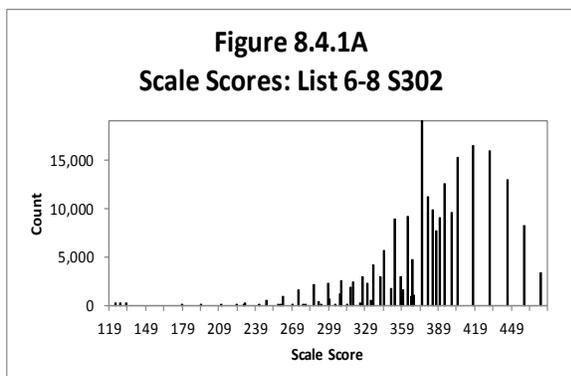| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.427 | 0.328 | | 0.148 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.683 | | 0.418 | |
| | 2 | 0.489 | | 0.321 | |
| | 3 | 0.396 | | 0.280 | |
| | 4 | 0.303 | | 0.234 | |
| | 5 | 0.458 | | 0.409 | |
| | 6 | 0.462 | | 0.322 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.977 | 0.007 | 0.016 | 0.963 |
| | 2/3 | 0.920 | 0.027 | 0.053 | 0.882 |
| | 3/4 | 0.841 | 0.056 | 0.102 | 0.777 |
| | 4/5 | 0.773 | 0.090 | 0.137 | 0.692 |
| | 5/6 | 0.805 | 0.134 | 0.061 | 0.730 |

**Table 8.4.1E-3**
Accuracy and Consistency of Classification Indices: List (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.400 | 0.326 | | 0.150 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.699 | | 0.456 | |
| | 2 | 0.508 | | 0.344 | |
| | 3 | 0.268 | | 0.184 | |
| | 4 | 0.353 | | 0.265 | |
| | 5 | 0.378 | | 0.341 | |
| | 6 | 0.556 | | 0.406 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.968 | 0.009 | 0.023 | 0.952 |
| | 2/3 | 0.912 | 0.023 | 0.064 | 0.875 |
| | 3/4 | 0.861 | 0.060 | 0.078 | 0.793 |
| | 4/5 | 0.772 | 0.094 | 0.133 | 0.690 |
| | 5/6 | 0.768 | 0.180 | 0.052 | 0.709 |

## 8.4.2    Reading 6-8



**Figure 8.4.2A**
**Scale Scores: Read 6-8 S302**



**Figure 8.4.2B**
**Proficiency Level: Read 6-8 S302**
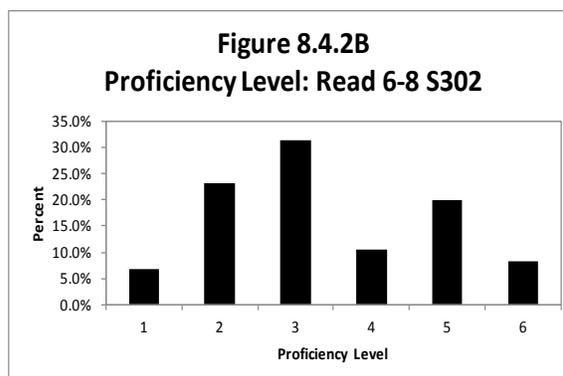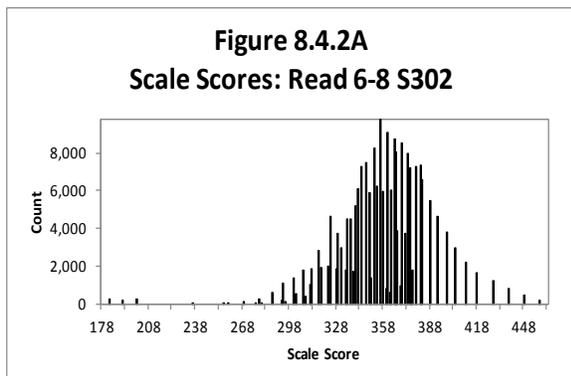
**Table 8.4.2A**

Scale Score Descriptive Statistics: Read 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 6 | 73,817 | 183 | 458 | 351.16 | 24.27 |
| 7 | 74,844 | 191 | 458 | 358.84 | 26.35 |
| 8 | 70,659 | 200 | 458 | 365.56 | 28.63 |
| **Total** | 219,320 | 183 | 458 | 358.42 | 27.09 |

**Table 8.4.2B**

Proficiency Level Distribution: Read 6-8 S302

| Level | Grade 6 | | Grade 7 | | Grade 8 | | Total | |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| 1 | 3,493 | 4.7% | 4,931 | 6.6% | 6,217 | 8.8% | 14,641 | 6.7% |
| 2 | 15,059 | 20.4% | 17,500 | 23.4% | 18,273 | 25.9% | 50,832 | 23.2% |
| 3 | 26,439 | 35.8% | 23,810 | 31.8% | 18,598 | 26.3% | 68,847 | 31.4% |
| 4 | 9,163 | 12.4% | 8,568 | 11.4% | 5,271 | 7.5% | 23,002 | 10.5% |
| 5 | 14,568 | 19.7% | 14,182 | 18.9% | 15,002 | 21.2% | 43,752 | 19.9% |
| 6 | 5,095 | 6.9% | 5,853 | 7.8% | 7,298 | 10.3% | 18,246 | 8.3% |
| Total | 73,817 | 100.0% | 74,844 | 100.0% | 70,659 | 100.0% | 219,320 | 100.0% |

**Table 8.4.2C**
Conditional Standard Error of Measurement at Cut Scores: Read 6-8 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 6 | 312 | 11.96 | 13.78 | 15.60 |
| | 7 | 321 | 11.70 | 12.74 | 14.04 |
| | 8 | 329 | 11.70 | 11.96 | 12.74 |
| 2/3 | 6 | 340 | 11.70 | 11.18 | 11.70 |
| | 7 | 349 | 12.22 | 10.92 | 11.18 |
| | 8 | 358 | 12.74 | 10.92 | 10.92 |
| 3/4 | 6 | 360 | 13.00 | 10.92 | 10.92 |
| | 7 | 369 | 13.78 | 10.92 | 10.66 |
| | 8 | 376 | 14.82 | 11.44 | 10.92 |
| 4/5 | 6 | 366 | n/a | 10.92 | 10.66 |
| | 7 | 375 | n/a | 11.18 | 10.92 |
| | 8 | 382 | n/a | 11.70 | 10.92 |
| 5/6 | 6 | 382 | n/a | n/a | 10.92 |
| | 7 | 391 | n/a | n/a | 11.44 |
| | 8 | 398 | n/a | n/a | 11.96 |

* No equating was performed for S302

**Figure 8.4.2C**
Test Characteristic Curve: Read 6-8ABC S302
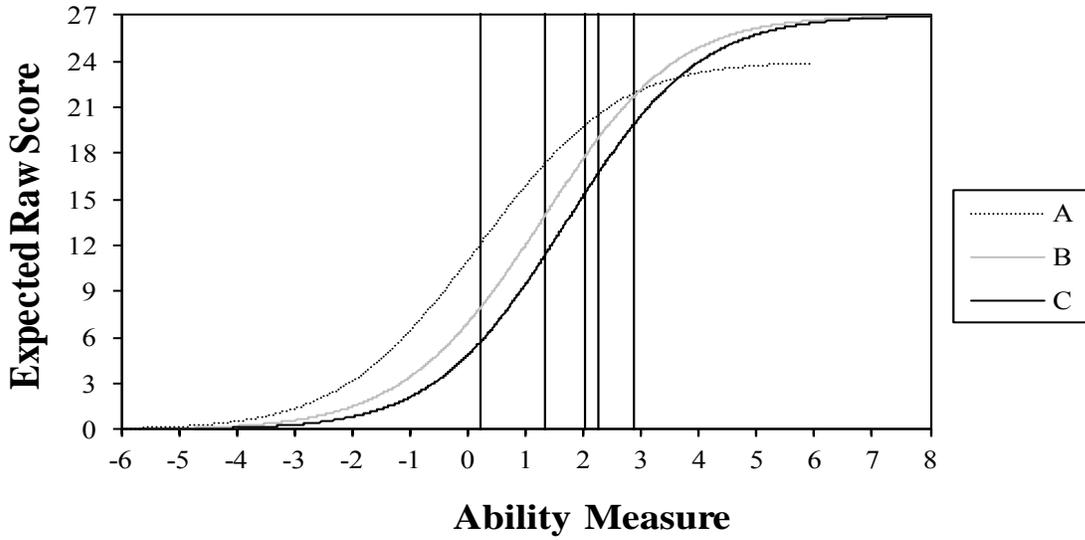


**Figure 8.4.2D**
Test Information Function: Read 6-8ABC S302
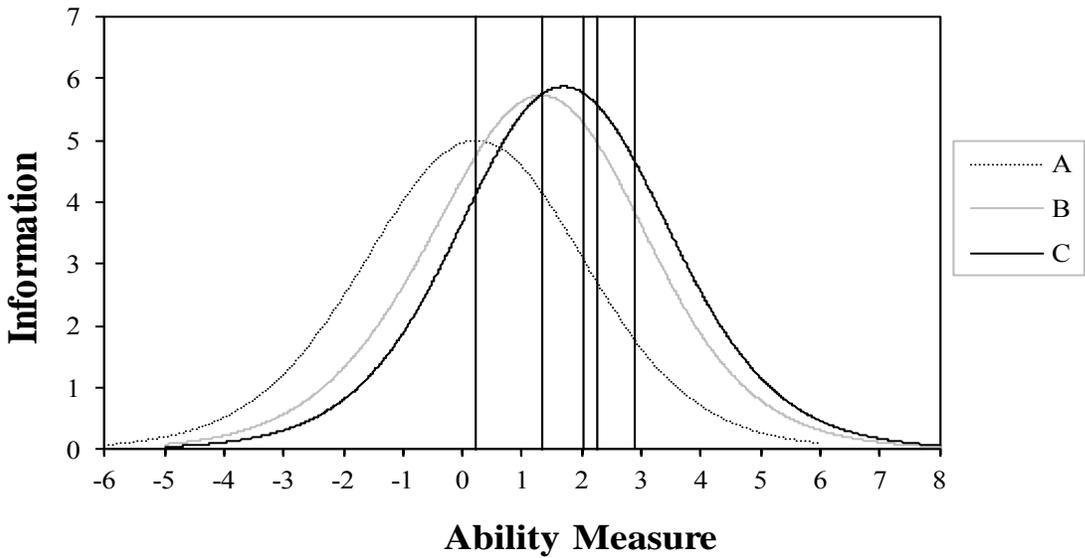
**Table 8.4.2D**

Weighted Reliability: Read 6-8 S302

| Tiers | No. of Students | Reliability | Reliability |
|---|---|---|---|
| A | 23,729 | 0.775 | |
| B | 82,291 | 0.780 | 0.770 |
| C | 113,300 | 0.761 | |

**Table 8.4.2E-1**

Accuracy and Consistency of Classification Indices: Read (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | | Kappa (k) | |
|---|---|---|---|---|---|---|
| | 0.497 | 0.400 | | | 0.223 | |
| Conditional on Level | Level | Accuracy | | | Consistency | |
| | 1 | 0.730 | | | 0.549 | |
| | 2 | 0.686 | | | 0.527 | |
| | 3 | 0.578 | | | 0.465 | |
| | 4 | 0.205 | | | 0.170 | |
| | 5 | 0.412 | | | 0.338 | |
| | 6 | - | | | 0.177 | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.974 | 0.012 | 0.014 | 0.958 | |
| | 2/3 | 0.887 | 0.042 | 0.071 | 0.840 | |
| | 3/4 | 0.790 | 0.077 | 0.133 | 0.717 | |
| | 4/5 | 0.794 | 0.092 | 0.114 | 0.727 | |
| | 5/6 | 0.931 | 0.069 | 0.000 | 0.897 | |

**Table 8.4.2E-2**
Accuracy and Consistency of Classification Indices: Read (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.500 | 0.399 | | 0.236 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.727 | | 0.540 | |
| | 2 | 0.657 | | 0.514 | |
| | 3 | 0.533 | | 0.434 | |
| | 4 | 0.210 | | 0.166 | |
| | 5 | 0.422 | | 0.343 | |
| | 6 | - | | 0.252 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.961 | 0.016 | 0.022 | 0.940 |
| | 2/3 | 0.869 | 0.051 | 0.080 | 0.817 |
| | 3/4 | 0.814 | 0.084 | 0.102 | 0.747 |
| | 4/5 | 0.824 | 0.083 | 0.093 | 0.759 |
| | 5/6 | 0.922 | 0.078 | 0.000 | 0.884 |

**Table 8.4.2E-3**
Accuracy and Consistency of Classification Indices: Read (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.490 | 0.392 | | 0.241 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.736 | | 0.559 | |
| | 2 | 0.648 | | 0.514 | |
| | 3 | 0.467 | | 0.369 | |
| | 4 | 0.140 | | 0.108 | |
| | 5 | 0.443 | | 0.368 | |
| | 6 | 0.536 | | 0.310 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.950 | 0.021 | 0.029 | 0.924 |
| | 2/3 | 0.857 | 0.056 | 0.087 | 0.803 |
| | 3/4 | 0.825 | 0.076 | 0.099 | 0.757 |
| | 4/5 | 0.827 | 0.092 | 0.081 | 0.760 |
| | 5/6 | 0.898 | 0.095 | 0.007 | 0.858 |

## 8.4.3 Writing 6-8



Figure 8.4.3A
Scale Scores: Writ 6-8 S302



Figure 8.4.3B
Proficiency Level: Writ 6-8 S302
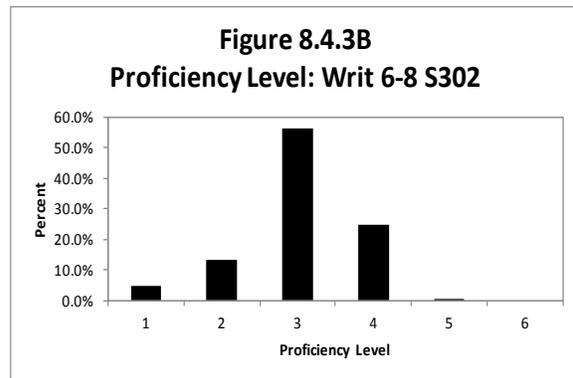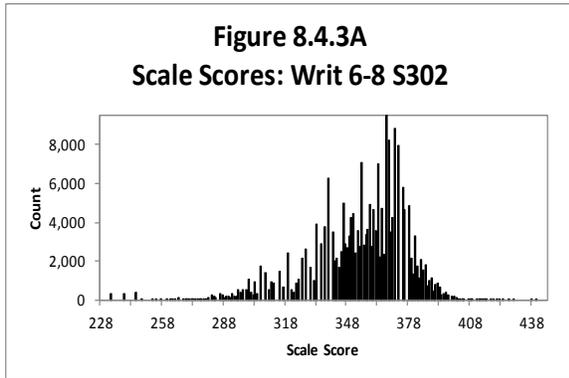
**Table 8.4.3A**

Scale Score Descriptive Statistics: Writ 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 6 | 73,833 | 233 | 416 | 349.34 | 23.45 |
| 7 | 74,840 | 239 | 440 | 355.17 | 23.13 |
| 8 | 70,654 | 245 | 438 | 359.80 | 23.10 |
| Total | 219,327 | 233 | 440 | 354.70 | 23.62 |

**Table 8.4.3B**

Proficiency Level Distribution: Writ 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|---------|---------|
| 1 | 2,526 | 3.4% | 3,660 | 4.9% | 4,299 | 6.1% | 10,485 | 4.8% |
| 2 | 9,221 | 12.5% | 9,253 | 12.4% | 11,238 | 15.9% | 29,712 | 13.5% |
| 3 | 35,381 | 47.9% | 43,231 | 57.8% | 45,067 | 63.8% | 123,679 | 56.4% |
| 4 | 26,136 | 35.4% | 18,455 | 24.7% | 9,962 | 14.1% | 54,553 | 24.9% |
| 5 | 566 | 0.8% | 237 | 0.3% | 85 | 0.1% | 888 | 0.4% |
| 6 | 3 | 0.0% | 4 | 0.0% | 3 | 0.0% | 10 | 0.0% |
| Total | 73,833 | 100.0% | 74,840 | 100.0% | 70,654 | 100.0% | 219,327 | 100.0% |

**Table 8.4.3C**
Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 6 | 298 | 9.02 | 7.15 | 7.77 |
| | 7 | 308 | 9.64 | 8.09 | 8.40 |
| | 8 | 318 | 11.19 | 8.40 | 8.40 |
| 2/3 | 6 | 329 | 11.82 | 8.40 | 8.40 |
| | 7 | 339 | 12.13 | 8.09 | 8.09 |
| | 8 | 348 | 11.82 | 8.09 | 8.09 |
| 3/4 | 6 | 361 | 11.51 | 7.77 | 7.46 |
| | 7 | 371 | 11.19 | 7.46 | 7.15 |
| | 8 | 381 | 10.57 | 7.15 | 6.84 |
| 4/5 | 6 | 391 | n/a | 6.84 | 6.53 |
| | 7 | 399 | n/a | 6.53 | 6.53 |
| | 8 | 408 | n/a | 6.53 | 6.53 |
| 5/6 | 6 | 412 | n/a | n/a | 6.84 |
| | 7 | 420 | n/a | n/a | 8.09 |
| | 8 | 428 | n/a | n/a | 9.95 |

* No equating was performed for Writing Tier A S302

**Figure 8.4.3C**

Test Characteristic Curve: Writ 6-8ABC S302
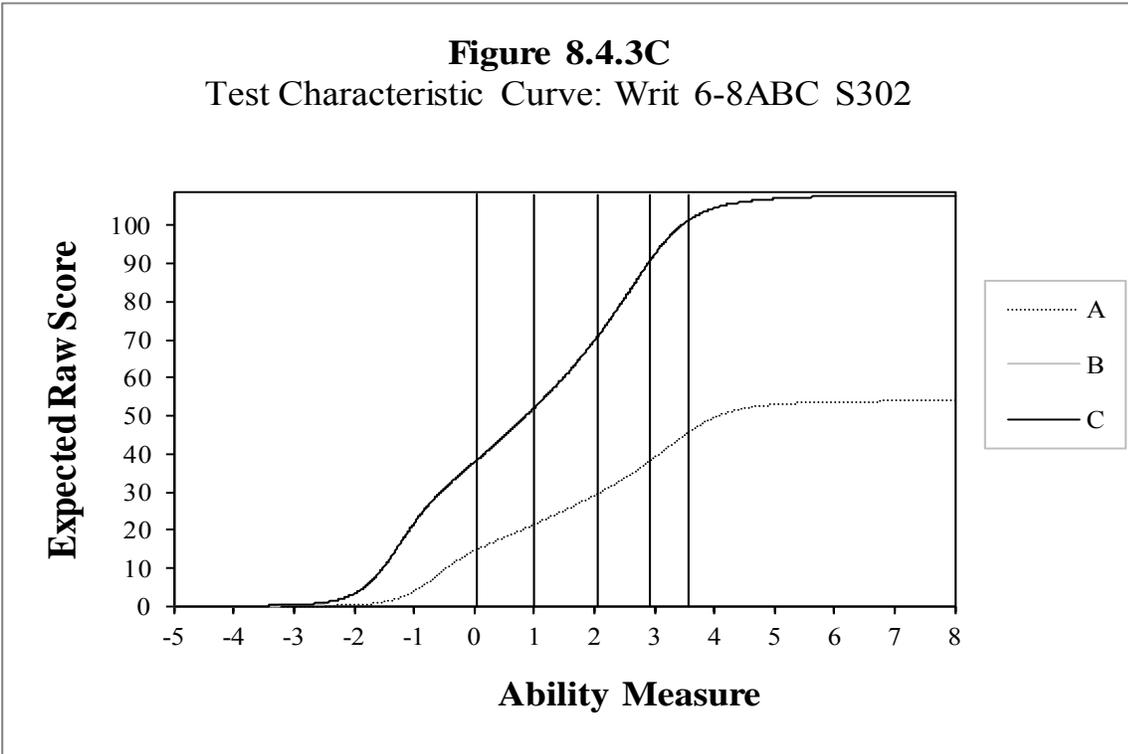


**Figure 8.4.3D**

Test Information Function: Writ 6-8ABC S302
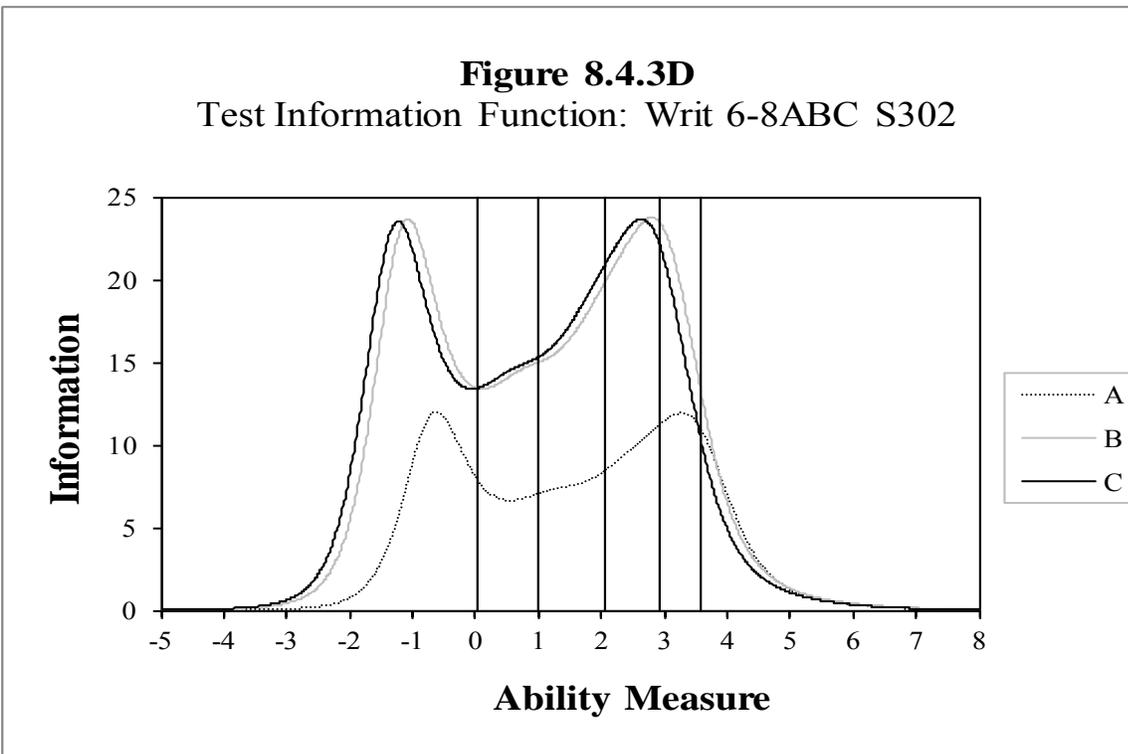
**Table 8.4.3D**

Weighted Reliability: Writ 6-8 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|-----------------|-------------|-------------|
| A | 23,729 | 0.889 | |
| B | 82,325 | 0.936 | 0.920 |
| C | 113,273 | 0.915 | |

**Table 8.4.3E-1**

Accuracy and Consistency of Classification Indices: Writ (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|--|-----------|--|
| | 0.812 | 0.739 | | 0.585 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.859 | | 0.767 | |
| | 2 | 0.796 | | 0.698 | |
| | 3 | 0.848 | | 0.764 | |
| | 4 | 0.774 | | 0.719 | |
| | 5 | - | | 0.000 | |
| | 6 | - | | 1.000 | |
| Indices at Cut Points | | | Accuracy | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.989 | 0.005 | 0.006 | 0.985 |
| | 2/3 | 0.959 | 0.019 | 0.021 | 0.942 |
| | 3/4 | 0.871 | 0.047 | 0.082 | 0.819 |
| | 4/5 | 0.992 | 0.008 | 0.000 | 0.992 |
| | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

**Table 8.4.3E-2**
Accuracy and Consistency of Classification Indices: Writ (Grade 7) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| Indices | 0.766 | 0.692 | | 0.482 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.899 | | 0.834 | |
| | 2 | 0.791 | | 0.693 | |
| | 3 | 0.835 | | 0.760 | |
| | 4 | 0.608 | | 0.523 | |
| | 5 | - | | - | |
| | 6 | - | | 1.000 | |
| Indices at | | Accuracy | | | |
| Cut Points | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.988 | 0.005 | 0.007 | 0.983 |
| | 2/3 | 0.961 | 0.019 | 0.020 | 0.944 |
| | 3/4 | 0.820 | 0.069 | 0.111 | 0.767 |
| | 4/5 | 0.997 | 0.003 | 0.000 | 0.997 |
| | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

**Table 8.3.3E-3**
Accuracy and Consistency of Classification Indices: Writ (Grade 8) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| Indices | 0.798 | 0.732 | | 0.491 | |
| Conditional | Level | Accuracy | | Consistency | |
| on Level | 1 | 0.885 | | 0.815 | |
| | 2 | 0.826 | | 0.738 | |
| | 3 | 0.786 | | 0.791 | |
| | 4 | - | | 0.311 | |
| | 5 | - | | - | |
| | 6 | - | | 1.000 | |
| Indices at | | Accuracy | | | |
| Cut Points | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.985 | 0.007 | 0.008 | 0.979 |
| | 2/3 | 0.955 | 0.019 | 0.026 | 0.936 |
| | 3/4 | 0.858 | 0.142 | 0.000 | 0.817 |
| | 4/5 | 0.999 | 0.001 | 0.000 | 0.999 |
| | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

## 8.4.4    Speaking 6-8



Figure 8.4.4A
Scale Scores: Spek 6-8 S302



Figure 8.4.4B
Proficiency Level: Spek 6-8 S302
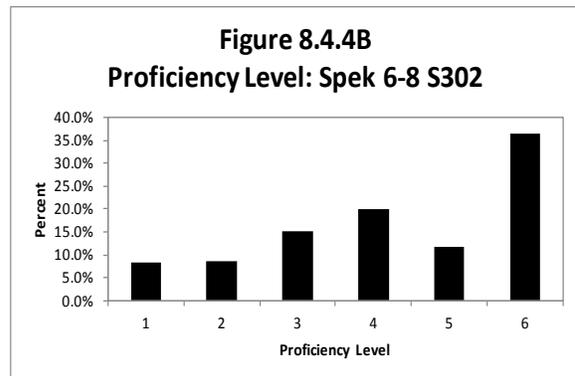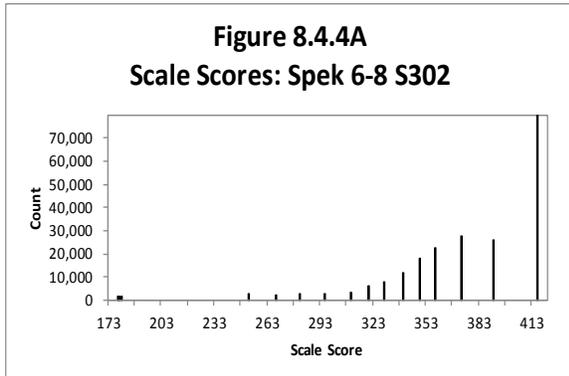
**Table 8.4.4A**

Scale Score Descriptive Statistics: Spek 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 6 | 73,805 | 178 | 416 | 369.61 | 47.71 |
| 7 | 74,819 | 179 | 416 | 372.59 | 48.65 |
| 8 | 70,595 | 180 | 416 | 374.63 | 49.97 |
| Total | 219,219 | 178 | 416 | 372.24 | 48.81 |

**Table 8.4.4B**

Proficiency Level Distribution: Spek 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|--------|---------|
| 1 | 5,206 | 7.1% | 6,406 | 8.6% | 6,266 | 8.9% | 17,878 | 8.2% |
| 2 | 6,570 | 8.9% | 4,615 | 6.2% | 7,442 | 10.5% | 18,627 | 8.5% |
| 3 | 11,431 | 15.5% | 10,240 | 13.7% | 11,505 | 16.3% | 33,176 | 15.1% |
| 4 | 18,194 | 24.7% | 17,218 | 23.0% | 8,300 | 11.8% | 43,712 | 19.9% |
| 5 | 8,505 | 11.5% | 8,930 | 11.9% | 8,457 | 12.0% | 25,892 | 11.8% |
| 6 | 23,899 | 32.4% | 27,410 | 36.6% | 28,625 | 40.5% | 79,934 | 36.5% |
| Total | 73,805 | 100.0% | 74,819 | 100.0% | 70,595 | 100.0% | 219,219 | 100.0% |

**Table 8.4.4C**
Conditional Standard Error of Measurement at
Cut Scores: Spek 6-8 S302

| Proficiency Level | Grade | Cut Score | SEM |
|---|---|---|---|
| 1/2 | 6 | 310 | 22.09 |
| | 7 | 314 | 22.29 |
| | 8 | 317 | 22.69 |
| 2/3 | 6 | 337 | 23.50 |
| | 7 | 340 | 23.50 |
| | 8 | 344 | 23.70 |
| 3/4 | 6 | 353 | 23.50 |
| | 7 | 358 | 23.30 |
| | 8 | 361 | 23.30 |
| 4/5 | 6 | 377 | 22.69 |
| | 7 | 380 | 22.29 |
| | 8 | 384 | 22.09 |
| 5/6 | 6 | 397 | 21.49 |
| | 7 | 400 | 21.49 |
| | 8 | 404 | 21.49 |

**Figure 8.4.4C**
Test Characteristic Curve: Spek 6-8 S302
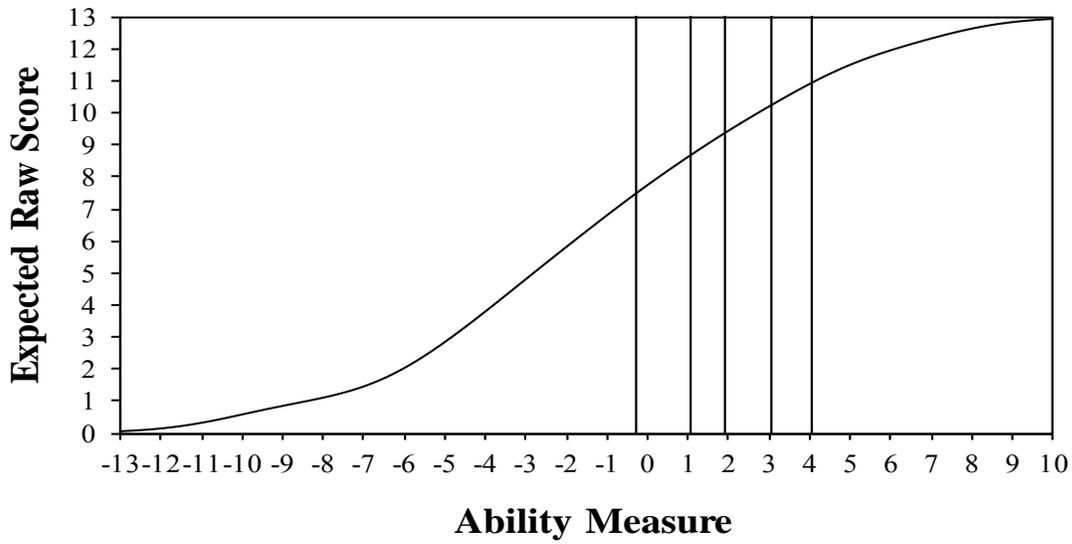


**Figure 8.4.4D**
Test Information Function: Spek 6-8 S302
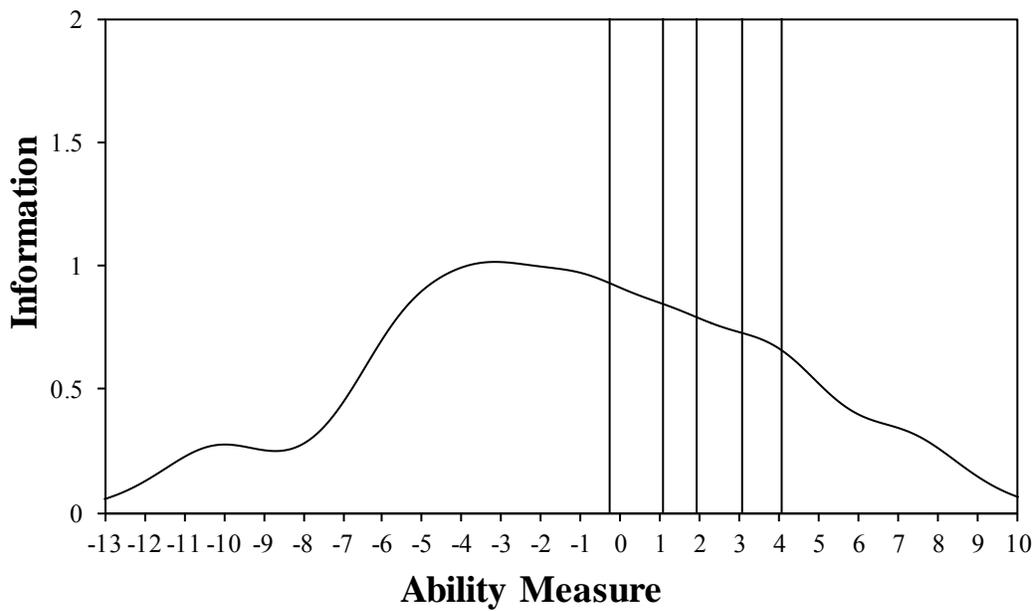
**Table 8.4.4D**

Reliability: Spek 6-8 S302

| Tiers | No. of Students | Reliability |
|-------|-----------------|-------------|
| -- | 219,219 | 0.904 |

**Table 8.4.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|---|-----------|---|
| | 0.563 | 0.456 | | 0.322 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.725 | | 0.576 | |
| | 2 | 0.405 | | 0.299 | |
| | 3 | 0.453 | | 0.370 | |
| | 4 | 0.552 | | 0.455 | |
| | 5 | 0.250 | | 0.179 | |
| | 6 | 0.804 | | 0.710 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.966 | 0.022 | 0.012 | 0.944 |
| | 2/3 | 0.913 | 0.054 | 0.034 | 0.882 |
| | 3/4 | 0.874 | 0.047 | 0.079 | 0.842 |
| | 4/5 | 0.876 | 0.022 | 0.102 | 0.827 |
| | 5/6 | 0.872 | 0.064 | 0.063 | 0.802 |

**Table 8.4.4E-2**

Accuracy and Consistency of Classification Indices: Spek (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.566 | 0.454 | | 0.310 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.798 | | 0.667 | |
| | 2 | 0.334 | | 0.239 | |
| | 3 | 0.457 | | 0.368 | |
| | 4 | 0.553 | | 0.442 | |
| | 5 | 0.226 | | 0.171 | |
| | 6 | 0.791 | | 0.700 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.968 | 0.018 | 0.014 | 0.949 |
| | 2/3 | 0.929 | 0.046 | 0.025 | 0.902 |
| | 3/4 | 0.888 | 0.046 | 0.066 | 0.859 |
| | 4/5 | 0.872 | 0.021 | 0.108 | 0.824 |
| | 5/6 | 0.844 | 0.081 | 0.076 | 0.760 |

**Table 8.4.4E-3**

Accuracy and Consistency of Classification Indices: Spek (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.533 | 0.434 | | 0.291 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.718 | | 0.595 | |
| | 2 | 0.445 | | 0.356 | |
| | 3 | 0.526 | | 0.443 | |
| | 4 | 0.320 | | 0.228 | |
| | 5 | 0.225 | | 0.175 | |
| | 6 | 0.805 | | 0.732 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.957 | 0.029 | 0.014 | 0.934 |
| | 2/3 | 0.908 | 0.050 | 0.042 | 0.884 |
| | 3/4 | 0.891 | 0.022 | 0.087 | 0.864 |
| | 4/5 | 0.910 | 0.021 | 0.069 | 0.852 |
| | 5/6 | 0.800 | 0.135 | 0.066 | 0.731 |

## 8.4.5    Oral Language Composite 6-8



**Figure 8.4.5A**
**Scale Scores: Oral 6-8 S302**



**Figure 8.4.5B**
**Proficiency Level: Oral 6-8 S302**
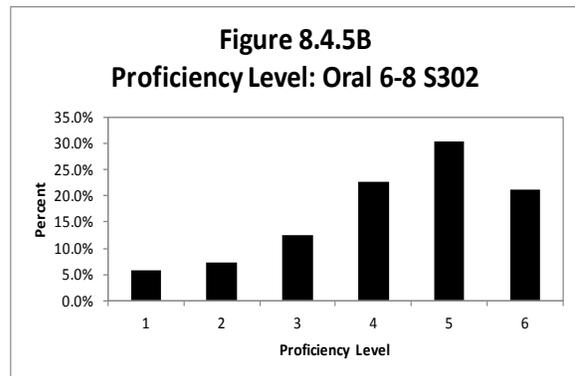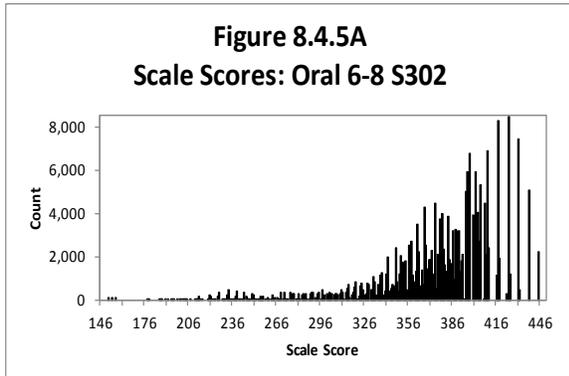
**Table 8.4.5A**

Scale Score Descriptive Statistics: Oral 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 6 | 73,741 | 151 | 445 | 373.63 | 39.11 |
| 7 | 74,722 | 154 | 445 | 379.15 | 41.12 |
| 8 | 70,516 | 156 | 445 | 383.33 | 42.98 |
| **Total** | 218,979 | 151 | 445 | 378.64 | 41.26 |

**Table 8.4.5B**

Proficiency Level Distribution: Oral 6-8 S302

| | Grade 6 | | Grade 7 | | Grade 8 | | Total | |
|---|---|---|---|---|---|---|---|---|
| **Level** | **Count** | **Percent** | **Count** | **Percent** | **Count** | **Percent** | **Count** | **Percent** |
| 1 | 3,450 | 4.7% | 4,331 | 5.8% | 4,761 | 6.8% | 12,542 | 5.7% |
| 2 | 5,243 | 7.1% | 5,312 | 7.1% | 5,500 | 7.8% | 16,055 | 7.3% |
| 3 | 9,040 | 12.3% | 9,747 | 13.0% | 8,580 | 12.2% | 27,367 | 12.5% |
| 4 | 18,364 | 24.9% | 16,435 | 22.0% | 15,133 | 21.5% | 49,932 | 22.8% |
| 5 | 23,316 | 31.6% | 22,330 | 29.9% | 20,850 | 29.6% | 66,496 | 30.4% |
| 6 | 14,328 | 19.4% | 16,567 | 22.2% | 15,692 | 22.3% | 46,587 | 21.3% |
| Total | 73,741 | 100.0% | 74,722 | 100.0% | 70,516 | 100.0% | 218,979 | 100.0% |

**Table 8.4.5C**
n/a

**Figure 8.4.5C**
n/a

**Figure 8.4.5D**
n/a

**Table 8.4.5D**
Oral Composite Reliability: Oral 6-8 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.50 | 1972.790 | 0.645 |
| Speaking | 0.50 | 2377.192 | 0.904 |
| Oral | | 1700.473 | 0.863 |

*Variances from students who had results in all four domains

**Table 8.4.5E-1**
Accuracy and Consistency of Classification Indices: Oral (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.554 | 0.447 | | 0.290 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.840 | | 0.713 | |
| | 2 | 0.555 | | 0.405 | |
| | 3 | 0.489 | | 0.364 | |
| | 4 | 0.571 | | 0.447 | |
| | 5 | 0.524 | | 0.453 | |
| | 6 | 0.571 | | 0.441 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.982 | 0.007 | 0.012 | 0.973 |
| | 2/3 | 0.954 | 0.023 | 0.023 | 0.932 |
| | 3/4 | 0.909 | 0.042 | 0.050 | 0.874 |
| | 4/5 | 0.856 | 0.047 | 0.097 | 0.803 |
| | 5/6 | 0.831 | 0.093 | 0.076 | 0.781 |

**Table 8.4.5E-2**

Accuracy and Consistency of Classification Indices: Oral (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.536 | 0.436 | | 0.283 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.845 | | 0.722 | |
| | 2 | 0.514 | | 0.373 | |
| | 3 | 0.503 | | 0.380 | |
| | 4 | 0.523 | | 0.398 | |
| | 5 | 0.488 | | 0.423 | |
| | 6 | 0.590 | | 0.468 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.978 | 0.008 | 0.014 | 0.968 |
| | 2/3 | 0.951 | 0.025 | 0.024 | 0.927 |
| | 3/4 | 0.907 | 0.040 | 0.053 | 0.872 |
| | 4/5 | 0.859 | 0.045 | 0.096 | 0.805 |
| | 5/6 | 0.813 | 0.109 | 0.078 | 0.764 |

**Table 8.4.5E-3**

Accuracy and Consistency of Classification Indices: Oral (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.525 | 0.429 | | 0.277 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.838 | | 0.716 | |
| | 2 | 0.507 | | 0.372 | |
| | 3 | 0.468 | | 0.349 | |
| | 4 | 0.506 | | 0.378 | |
| | 5 | 0.479 | | 0.420 | |
| | 6 | 0.581 | | 0.458 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.974 | 0.010 | 0.016 | 0.962 |
| | 2/3 | 0.945 | 0.027 | 0.028 | 0.921 |
| | 3/4 | 0.908 | 0.038 | 0.054 | 0.873 |
| | 4/5 | 0.859 | 0.050 | 0.092 | 0.800 |
| | 5/6 | 0.806 | 0.121 | 0.073 | 0.759 |

## 8.4.6    Literacy Composite 6-8



**Figure 8.4.6A**
**Scale Scores: Litr 6-8 S302**



**Figure 8.4.6B**
**Proficiency Level: Litr 6-8 S302**
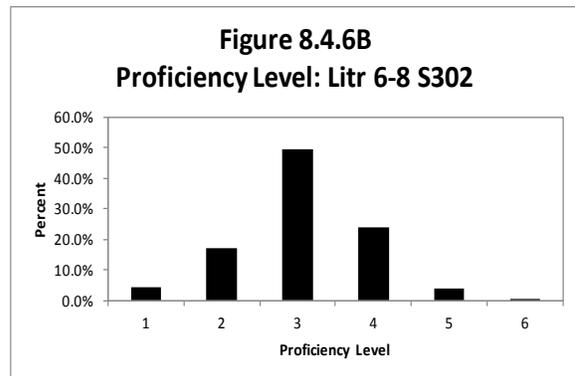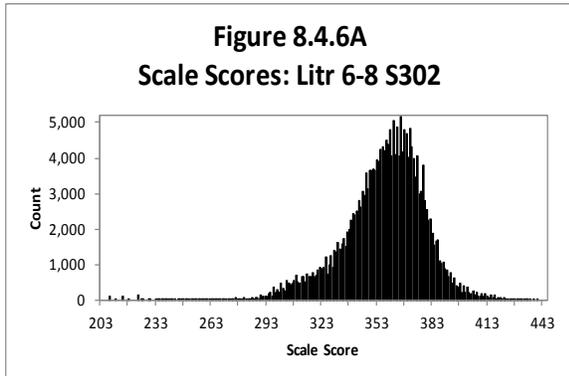
**Table 8.4.6A**
Scale Score Descriptive Statistics: Litr 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 6 | 73,722 | 208 | 436 | 350.51 | 20.98 |
| 7 | 74,744 | 215 | 435 | 357.27 | 21.99 |
| 8 | 70,562 | 223 | 440 | 362.95 | 23.21 |
| **Total** | 219,028 | 208 | 440 | 356.82 | 22.63 |

**Table 8.4.6B**
Proficiency Level Distribution: Litr 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|--------|---------|
| 1 | 2,050 | 2.8% | 3,432 | 4.6% | 4,463 | 6.3% | 9,945 | 4.5% |
| 2 | 11,345 | 15.4% | 12,576 | 16.8% | 14,069 | 19.9% | 37,990 | 17.3% |
| 3 | 36,171 | 49.1% | 36,732 | 49.1% | 35,389 | 50.2% | 108,292 | 49.4% |
| 4 | 20,596 | 27.9% | 18,362 | 24.6% | 13,324 | 18.9% | 52,282 | 23.9% |
| 5 | 2,994 | 4.1% | 3,077 | 4.1% | 2,807 | 4.0% | 8,878 | 4.1% |
| 6 | 566 | 0.8% | 565 | 0.8% | 510 | 0.7% | 1,641 | 0.7% |
| Total | 73,722 | 100.0% | 74,744 | 100.0% | 70,562 | 100.0% | 219,028 | 100.0% |

**Table 8.4.6C**
n/a

**Figure 8.4.6C**
n/a

**Figure 8.4.6D**
n/a

**Table 8.4.6D**
Literacy Composite Reliability: Litr 6-8 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Reading | 0.50 | 733.056 | 0.770 |
| Writing | 0.50 | 556.239 | 0.920 |
| Literacy | | 511.916 | 0.896 |

*Variances from students who had results in all four domains

**Table 8.4.6E-1**
Accuracy and Consistency of Classification Indices: Litr (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.754 | 0.673 | | 0.504 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.660 | | 0.698 | |
| | 2 | 0.787 | | 0.679 | |
| | 3 | 0.831 | | 0.752 | |
| | 4 | 0.650 | | 0.583 | |
| | 5 | - | | 0.201 | |
| | 6 | - | | 0.999 | |

| Indices at Cut Points | | Accuracy | | | |
|---|---|---|---|---|---|
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.983 | 0.011 | 0.006 | 0.984 |
| | 2/3 | 0.937 | 0.033 | 0.030 | 0.919 |
| | 3/4 | 0.863 | 0.054 | 0.083 | 0.818 |
| | 4/5 | 0.952 | 0.048 | 0.000 | 0.946 |
| | 5/6 | 0.992 | 0.008 | 0.000 | 0.999 |

**Table 8.4.6E-2**
Accuracy and Consistency of Classification Indices: Litr (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.744 | 0.659 | | 0.492 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.734 | | 0.739 | |
| | 2 | 0.766 | | 0.653 | |
| | 3 | 0.822 | | 0.744 | |
| | 4 | 0.619 | | 0.545 | |
| | 5 | - | | 0.214 | |
| | 6 | - | | 1.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.977 | 0.013 | 0.010 | 0.976 |
| | 2/3 | 0.931 | 0.036 | 0.034 | 0.909 |
| | 3/4 | 0.866 | 0.055 | 0.078 | 0.821 |
| | 4/5 | 0.951 | 0.049 | 0.000 | 0.944 |
| | 5/6 | 0.992 | 0.008 | 0.000 | 0.999 |

**Table 8.4.6E-3**
Accuracy and Consistency of Classification Indices: Litr (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.728 | 0.643 | | 0.472 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.770 | | 0.745 | |
| | 2 | 0.768 | | 0.660 | |
| | 3 | 0.811 | | 0.731 | |
| | 4 | 0.536 | | 0.456 | |
| | 5 | - | | 0.196 | |
| | 6 | - | | 1.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.972 | 0.015 | 0.012 | 0.969 |
| | 2/3 | 0.921 | 0.039 | 0.040 | 0.896 |
| | 3/4 | 0.862 | 0.057 | 0.081 | 0.815 |
| | 4/5 | 0.953 | 0.047 | 0.000 | 0.949 |
| | 5/6 | 0.993 | 0.007 | 0.000 | 0.999 |

## 8.4.7    Comprehension Composite 6-8



**Figure 8.4.7A**
**Scale Scores: Cphn 6-8 S302**



**Figure 8.4.7B**
**Proficiency Level: Cphn 6-8 S302**
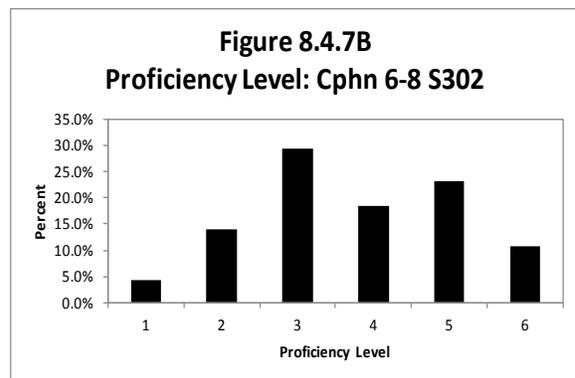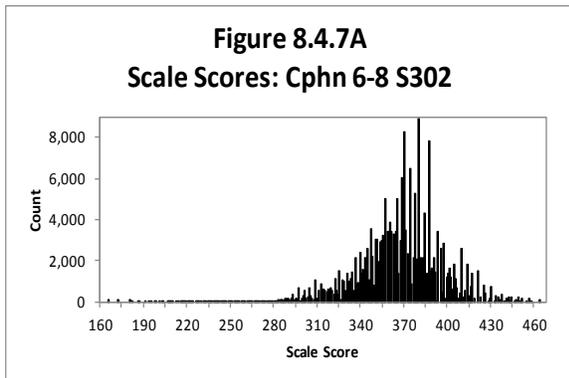
**Table 8.4.7A**

Scale Score Descriptive Statistics: Cphn 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 6 | 73,777 | 165 | 463 | 359.00 | 26.64 |
| 7 | 74,773 | 172 | 463 | 366.83 | 29.04 |
| 8 | 70,613 | 180 | 463 | 373.39 | 31.31 |
| Total | 219,163 | 165 | 463 | 366.31 | 29.61 |

**Table 8.4.7B**

Proficiency Level Distribution: Cphn 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|---|---|
| 1 | 1,957 | 2.7% | 3,244 | 4.3% | 4,022 | 5.7% | 9,223 | 4.2% |
| 2 | 8,754 | 11.9% | 10,408 | 13.9% | 11,361 | 16.1% | 30,523 | 13.9% |
| 3 | 23,371 | 31.7% | 22,765 | 30.4% | 18,021 | 25.5% | 64,157 | 29.3% |
| 4 | 14,763 | 20.0% | 13,690 | 18.3% | 12,073 | 17.1% | 40,526 | 18.5% |
| 5 | 17,997 | 24.4% | 16,436 | 22.0% | 16,623 | 23.5% | 51,056 | 23.3% |
| 6 | 6,935 | 9.4% | 8,230 | 11.0% | 8,513 | 12.1% | 23,678 | 10.8% |
| Total | 73,777 | 100.0% | 74,773 | 100.0% | 70,613 | 100.0% | 219,163 | 100.0% |

**Table 8.4.7C**
n/a

**Figure 8.4.7C**
n/a

**Figure 8.4.7D**
n/a

**Table 8.4.7D**
Comprehension Composite Reliability: Cphn 6-8 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.30 | 1972.790 | 0.645 |
| Reading | 0.70 | 733.056 | 0.770 |
| Comprehension | | 876.066 | 0.834 |

*Variances from students who had results in all four domains

**Table 8.4.7E-1**
Accuracy and Consistency of Classification Indices: Cphn (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.557 | 0.456 | | 0.300 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.781 | | 0.611 | |
| | 2 | 0.668 | | 0.513 | |
| | 3 | 0.673 | | 0.558 | |
| | 4 | 0.393 | | 0.305 | |
| | 5 | 0.498 | | 0.431 | |
| | 6 | 0.597 | | 0.347 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.987 | 0.005 | 0.008 | 0.980 |
| | 2/3 | 0.934 | 0.031 | 0.034 | 0.903 |
| | 3/4 | 0.847 | 0.063 | 0.090 | 0.794 |
| | 4/5 | 0.843 | 0.057 | 0.101 | 0.782 |
| | 5/6 | 0.907 | 0.091 | 0.002 | 0.882 |

**Table 8.4.7E-2**
Accuracy and Consistency of Classification Indices: Cphn (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.545 | 0.442 | | 0.297 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.792 | | 0.630 | |
| | 2 | 0.648 | | 0.501 | |
| | 3 | 0.646 | | 0.533 | |
| | 4 | 0.373 | | 0.287 | |
| | 5 | 0.471 | | 0.397 | |
| | 6 | 0.625 | | 0.395 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.979 | 0.008 | 0.013 | 0.969 |
| | 2/3 | 0.922 | 0.036 | 0.042 | 0.888 |
| | 3/4 | 0.851 | 0.061 | 0.088 | 0.798 |
| | 4/5 | 0.852 | 0.059 | 0.089 | 0.791 |
| | 5/6 | 0.899 | 0.088 | 0.013 | 0.870 |

**Table 8.4.7E-3**
Accuracy and Consistency of Classification Indices: Cphn (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.532 | 0.426 | | 0.288 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.789 | | 0.630 | |
| | 2 | 0.646 | | 0.505 | |
| | 3 | 0.566 | | 0.452 | |
| | 4 | 0.350 | | 0.269 | |
| | 5 | 0.487 | | 0.407 | |
| | 6 | 0.608 | | 0.402 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.972 | 0.011 | 0.017 | 0.959 |
| | 2/3 | 0.912 | 0.039 | 0.049 | 0.874 |
| | 3/4 | 0.854 | 0.064 | 0.083 | 0.800 |
| | 4/5 | 0.847 | 0.065 | 0.088 | 0.788 |
| | 5/6 | 0.894 | 0.079 | 0.027 | 0.856 |

## 8.4.8　Overall Composite 6-8



**Figure 8.4.8A**
**Scale Scores: Over 6-8 S302**



**Figure 8.4.8B**
**Proficiency Level: Over 6-8 S302**
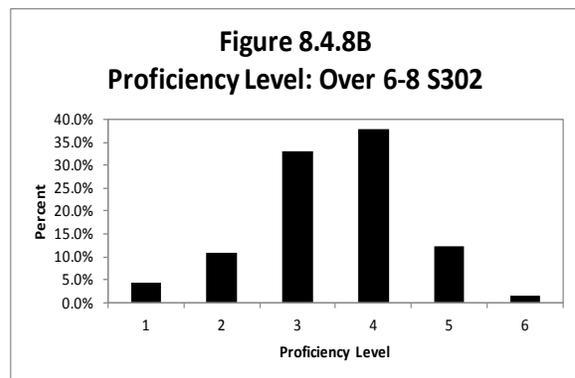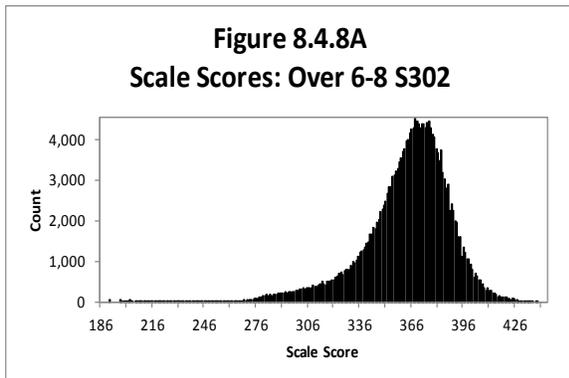
**Table 8.4.8A**

Scale Score Descriptive Statistics: Over 6-8 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 6 | 73,582 | 191 | 439 | 357.25 | 23.91 |
| 7 | 74,562 | 197 | 438 | 363.63 | 25.38 |
| 8 | 70,381 | 203 | 439 | 368.85 | 26.97 |
| Total | 218,525 | 191 | 439 | 363.16 | 25.86 |

**Table 8.4.8B**

Proficiency Level Distribution: Over 6-8 S302

| Level | Grade 6 Count | Grade 6 Percent | Grade 7 Count | Grade 7 Percent | Grade 8 Count | Grade 8 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|-------|---------|
| 1 | 2,314 | 3.1% | 3,266 | 4.4% | 4,007 | 5.7% | 9,587 | 4.4% |
| 2 | 7,253 | 9.9% | 7,968 | 10.7% | 8,707 | 12.4% | 23,928 | 10.9% |
| 3 | 23,553 | 32.0% | 25,285 | 33.9% | 23,057 | 32.8% | 71,895 | 32.9% |
| 4 | 30,308 | 41.2% | 27,047 | 36.3% | 25,561 | 36.3% | 82,916 | 37.9% |
| 5 | 8,916 | 12.1% | 9,828 | 13.2% | 8,009 | 11.4% | 26,753 | 12.2% |
| 6 | 1,238 | 1.7% | 1,168 | 1.6% | 1,040 | 1.5% | 3,446 | 1.6% |
| Total | 73,582 | 100.0% | 74,562 | 100.0% | 70,381 | 100.0% | 218,525 | 100.0% |

**Table 8.4.8C**
n/a

**Figure 8.4.8C**
n/a

**Figure 8.4.8D**
n/a

**Table 8.4.8D**
Overall Composite Reliability: Over 6-8 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.15 | 1972.790 | 0.645 |
| Reading | 0.35 | 733.056 | 0.770 |
| Speaking | 0.15 | 2377.192 | 0.904 |
| Writing | 0.35 | 556.239 | 0.920 |
| Overall Composite | | 668.610 | 0.930 |

*Variances from students who had results in all four domains

**Table 8.4.8E-1**
Accuracy and Consistency of Classification Indices: Over (Grade 6) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.745 | 0.662 | | 0.519 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.788 | | 0.814 | |
| | 2 | 0.771 | | 0.664 | |
| | 3 | 0.832 | | 0.748 | |
| | 4 | 0.739 | | 0.672 | |
| | 5 | 0.512 | | 0.402 | |
| | 6 | - | | 0.982 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.988 | 0.007 | 0.005 | 0.988 |
| | 2/3 | 0.962 | 0.022 | 0.016 | 0.949 |
| | 3/4 | 0.908 | 0.037 | 0.055 | 0.876 |
| | 4/5 | 0.889 | 0.066 | 0.045 | 0.854 |
| | 5/6 | 0.983 | 0.017 | 0.000 | 0.987 |

**Table 8.4.8E-2**
Accuracy and Consistency of Classification Indices: Over (Grade 7) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.728 | 0.640 | | 0.503 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.818 | | 0.819 | |
| | 2 | 0.753 | | 0.643 | |
| | 3 | 0.828 | | 0.743 | |
| | 4 | 0.700 | | 0.616 | |
| | 5 | 0.524 | | 0.424 | |
| | 6 | - | | 0.984 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.985 | 0.008 | 0.007 | 0.984 |
| | 2/3 | 0.958 | 0.024 | 0.018 | 0.943 |
| | 3/4 | 0.903 | 0.039 | 0.058 | 0.868 |
| | 4/5 | 0.886 | 0.064 | 0.050 | 0.848 |
| | 5/6 | 0.984 | 0.016 | 0.000 | 0.988 |

**Table 8.4.8E-3**
Accuracy and Consistency of Classification Indices: Over (Grade 8) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.719 | 0.633 | | 0.495 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.832 | | 0.815 | |
| | 2 | 0.749 | | 0.640 | |
| | 3 | 0.810 | | 0.716 | |
| | 4 | 0.708 | | 0.617 | |
| | 5 | 0.335 | | 0.369 | |
| | 6 | - | | 0.987 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.981 | 0.010 | 0.009 | 0.979 |
| | 2/3 | 0.951 | 0.027 | 0.022 | 0.935 |
| | 3/4 | 0.900 | 0.039 | 0.061 | 0.865 |
| | 4/5 | 0.908 | 0.067 | 0.024 | 0.855 |
| | 5/6 | 0.985 | 0.015 | 0.000 | 0.989 |

## 8.5 Grades: 9–12

### 8.5.1 Listening 9-12



Figure 8.5.1A
Scale Scores: List 9-12 S302



Figure 8.5.1B
Proficiency Level: List 9-12 S302
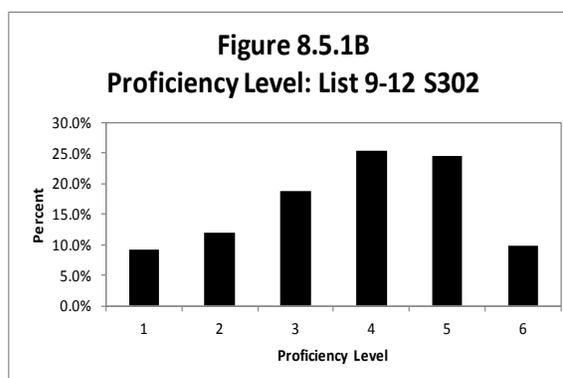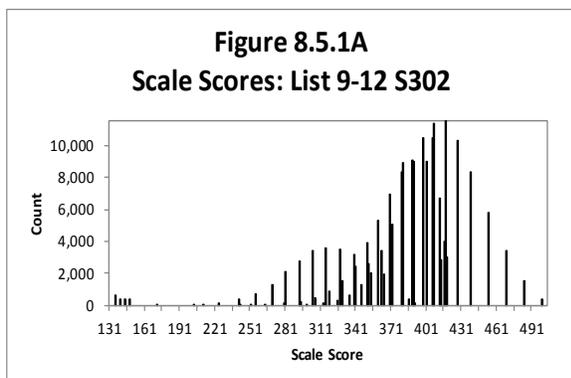
**Table 8.5.1A**
Scale Score Descriptive Statistics: List 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 9 | 79,568 | 136 | 499 | 381.43 | 50.87 |
| 10 | 50,140 | 140 | 499 | 385.08 | 46.50 |
| 11 | 38,200 | 144 | 499 | 389.92 | 45.23 |
| 12 | 29,999 | 148 | 499 | 391.12 | 45.82 |
| Total | 197,907 | 136 | 499 | 385.46 | 48.13 |

**Table 8.5.1B**
Proficiency Level Distribution: List 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1 | 7,822 | 9.8% | 4,301 | 8.6% | 3,222 | 8.4% | 2,831 | 9.4% | 18,176 | 9.2% |
| 2 | 10,443 | 13.1% | 5,883 | 11.7% | 4,285 | 11.2% | 3,190 | 10.6% | 23,801 | 12.0% |
| 3 | 11,839 | 14.9% | 10,538 | 21.0% | 6,883 | 18.0% | 7,848 | 26.2% | 37,108 | 18.8% |
| 4 | 17,621 | 22.1% | 12,970 | 25.9% | 11,612 | 30.4% | 8,335 | 27.8% | 50,538 | 25.5% |
| 5 | 23,901 | 30.0% | 11,949 | 23.8% | 8,171 | 21.4% | 4,694 | 15.6% | 48,715 | 24.6% |
| 6 | 7,942 | 10.0% | 4,499 | 9.0% | 4,027 | 10.5% | 3,101 | 10.3% | 19,569 | 9.9% |
| Total | 79,568 | 100.0% | 50,140 | 100.0% | 38,200 | 100.0% | 29,999 | 100.0% | 197,907 | 100.0% |

**Table 8.5.1C**
Conditional Standard Error of Measurement at Cut Scores: List 9-12 S302

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 9 | 312 | 20.66 | 21.04 | 22.17 |
| | 10 | 322 | 20.66 | 20.29 | 21.04 |
| | 11 | 332 | 21.04 | 19.91 | 19.91 |
| | 12 | 343 | 21.42 | 19.54 | 19.16 |
| 2/3 | 9 | 352 | 22.17 | 19.54 | 18.41 |
| | 10 | 358 | 22.92 | 19.54 | 18.41 |
| | 11 | 363 | 23.29 | 19.54 | 18.03 |
| | 12 | 366 | 23.67 | 19.54 | 18.03 |
| 3/4 | 9 | 381 | 25.92 | 19.54 | 18.03 |
| | 10 | 386 | 26.68 | 19.91 | 18.03 |
| | 11 | 389 | 27.05 | 19.91 | 18.03 |
| | 12 | 391 | 27.80 | 19.91 | 18.03 |
| 4/5 | 9 | 406 | n/a | 20.66 | 18.79 |
| | 10 | 412 | n/a | 21.42 | 19.16 |
| | 11 | 416 | n/a | 21.79 | 19.54 |
| | 12 | 418 | n/a | 21.79 | 19.54 |
| 5/6 | 9 | 432 | n/a | n/a | 21.04 |
| | 10 | 436 | n/a | n/a | 21.42 |
| | 11 | 438 | n/a | n/a | 21.79 |
| | 12 | 439 | n/a | n/a | 22.17 |

# Figure 8.5.1C
## Test Characteristic Curve: List 9-12ABC S302
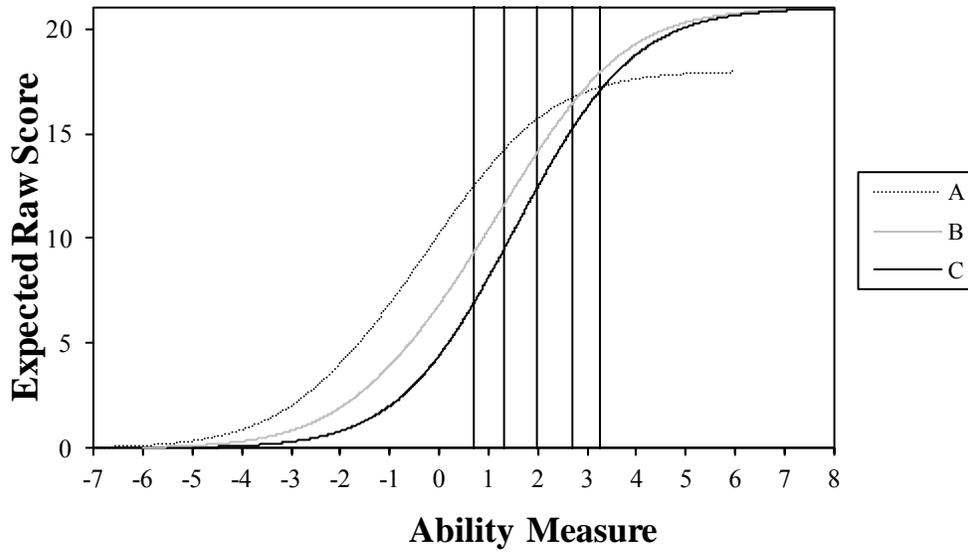


# Figure 8.5.1D
## Test Information Function: List 9-12ABC S302
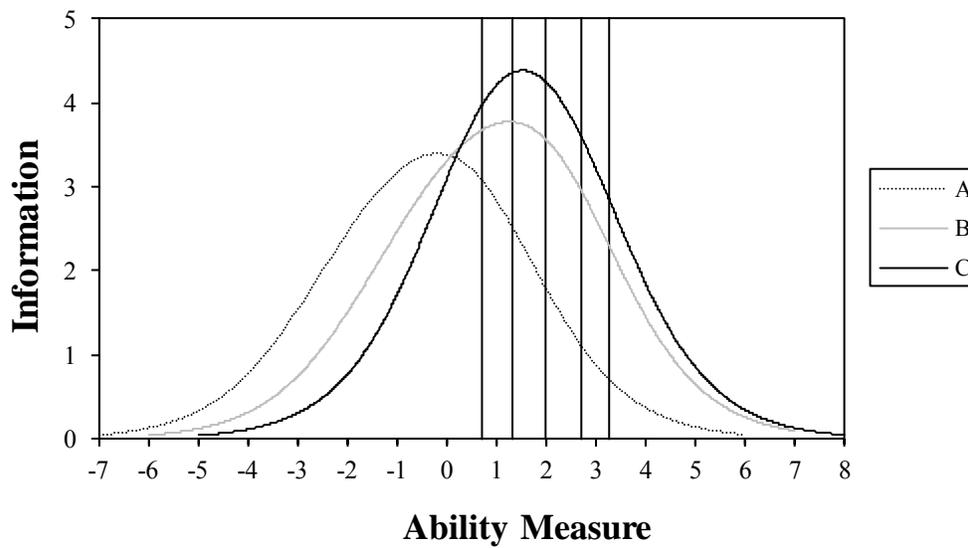
**Table 8.5.1D**

Weighted Reliability: List 9-12 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|-----------------|-------------|-------------|
| A | 29,077 | 0.682 | |
| B | 76,057 | 0.720 | 0.688 |
| C | 92,773 | 0.665 | |

**Table 8.5.1E-1**

Accuracy and Consistency of Classification Indices: List  (Grade 9) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|--|-----------|--|
| | 0.410 | 0.316 | | 0.155 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.819 | | 0.684 | |
| | 2 | 0.511 | | 0.327 | |
| | 3 | 0.289 | | 0.186 | |
| | 4 | 0.287 | | 0.261 | |
| | 5 | 0.437 | | 0.396 | |
| | 6 | - | | 0.149 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.957 | 0.016 | 0.028 | 0.938 |
| | 2/3 | 0.899 | 0.023 | 0.078 | 0.853 |
| | 3/4 | 0.817 | 0.048 | 0.134 | 0.716 |
| | 4/5 | 0.681 | 0.169 | 0.150 | 0.630 |
| | 5/6 | 0.900 | 0.100 | 0.000 | 0.822 |

**Table 8.5.1E-2**

Accuracy and Consistency of Classification Indices: List (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|-----------------|----------|-------------|--|-----------|--|
| | 0.370 | 0.309 | | 0.134 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.825 | | 0.687 | |
| | 2 | 0.484 | | 0.281 | |
| | 3 | 0.366 | | 0.244 | |
| | 4 | 0.312 | | 0.301 | |
| | 5 | - | | 0.305 | |
| | 6 | - | | 0.124 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False | False | |
| | Cut Point | Accuracy | Positives | Negatives | Consistency |
| | 1/2 | 0.964 | 0.013 | 0.022 | 0.947 |
| | 2/3 | 0.902 | 0.022 | 0.076 | 0.853 |
| | 3/4 | 0.768 | 0.041 | 0.192 | 0.667 |
| | 4/5 | 0.672 | 0.328 | 0.000 | 0.631 |
| | 5/6 | 0.910 | 0.090 | 0.000 | 0.855 |

**Table 8.5.1E-3**

Accuracy and Consistency of Classification Indices: List (Grade 11) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|---|---|---|---|---|---|---|
| | 0.398 | 0.308 | | 0.126 | | |
| Conditional on Level | Level | Accuracy | | Consistency | | |
| | 1 | 0.826 | | 0.692 | | |
| | 2 | 0.477 | | 0.256 | | |
| | 3 | 0.318 | | 0.206 | | |
| | 4 | 0.359 | | 0.348 | | |
| | 5 | - | | 0.265 | | |
| | 6 | - | | 0.138 | | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.966 | 0.013 | 0.021 | 0.950 | |
| | 2/3 | 0.902 | 0.017 | 0.081 | 0.854 | |
| | 3/4 | 0.778 | 0.035 | 0.187 | 0.670 | |
| | 4/5 | 0.681 | 0.319 | 0.000 | 0.626 | |
| | 5/6 | 0.895 | 0.105 | 0.000 | 0.840 | |

**Table 8.5.1E-4**

Accuracy and Consistency of Classification Indices: List (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|---|---|---|---|---|---|---|
| | 0.368 | 0.295 | | 0.109 | | |
| Conditional on Level | Level | Accuracy | | Consistency | | |
| | 1 | 0.860 | | 0.694 | | |
| | 2 | 0.367 | | 0.165 | | |
| | 3 | 0.377 | | 0.284 | | |
| | 4 | 0.319 | | 0.311 | | |
| | 5 | - | | 0.183 | | |
| | 6 | - | | 0.125 | | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.964 | 0.011 | 0.025 | 0.943 | |
| | 2/3 | 0.884 | 0.014 | 0.102 | 0.826 | |
| | 3/4 | 0.673 | 0.021 | 0.306 | 0.599 | |
| | 4/5 | 0.740 | 0.260 | 0.000 | 0.643 | |
| | 5/6 | 0.897 | 0.103 | 0.000 | 0.848 | |

## 8.5.2    Reading 9-12



**Figure 8.5.2A**
**Scale Scores: Read 9-12 S302**



**Figure 8.5.2B**
**Proficiency Level: Read 9-12 S302**
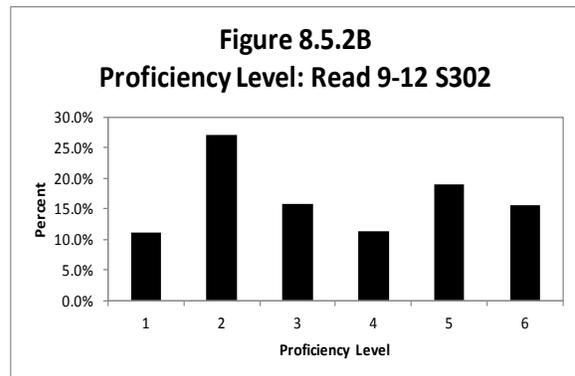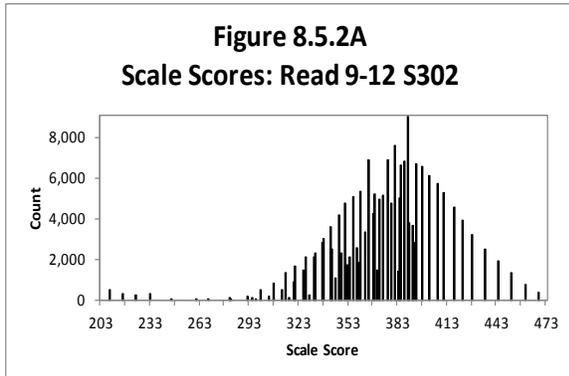
**Table 8.5.2A**

Scale Score Descriptive Statistics: Read 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 9 | 79,602 | 208 | 468 | 372.59 | 32.55 |
| 10 | 50,148 | 216 | 468 | 374.91 | 30.92 |
| 11 | 38,245 | 224 | 468 | 379.46 | 30.85 |
| 12 | 30,042 | 233 | 468 | 380.63 | 31.45 |
| Total | 198,037 | 208 | 468 | 375.73 | 31.82 |

**Table 8.5.2B**

Proficiency Level Distribution: Read 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 8,899 | 11.2% | 5,279 | 10.5% | 4,293 | 11.2% | 3,796 | 12.6% | 22,267 | 11.2% |
| 2 | 19,440 | 24.4% | 15,531 | 31.0% | 10,490 | 27.4% | 7,980 | 26.6% | 53,441 | 27.0% |
| 3 | 15,281 | 19.2% | 7,698 | 15.4% | 4,671 | 12.2% | 3,705 | 12.3% | 31,355 | 15.8% |
| 4 | 8,059 | 10.1% | 6,381 | 12.7% | 4,901 | 12.8% | 3,080 | 10.3% | 22,421 | 11.3% |
| 5 | 14,506 | 18.2% | 8,445 | 16.8% | 7,535 | 19.7% | 7,070 | 23.5% | 37,556 | 19.0% |
| 6 | 13,417 | 16.9% | 6,814 | 13.6% | 6,355 | 16.6% | 4,411 | 14.7% | 30,997 | 15.7% |
| Total | 79,602 | 100.0% | 50,148 | 100.0% | 38,245 | 100.0% | 30,042 | 100.0% | 198,037 | 100.0% |

**Table 8.5.2C**
Conditional Standard Error of Measurement at Cut Scores: Read 9-12 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 9 | 336 | 11.96 | 11.70 | 14.82 |
| | 10 | 341 | 11.70 | 11.44 | 14.04 |
| | 11 | 346 | 11.70 | 11.18 | 13.26 |
| | 12 | 350 | 11.70 | 10.92 | 12.74 |
| 2/3 | 9 | 364 | 11.96 | 10.92 | 11.44 |
| | 10 | 370 | 12.48 | 10.92 | 10.92 |
| | 11 | 374 | 12.74 | 11.18 | 10.66 |
| | 12 | 376 | 13.00 | 11.18 | 10.66 |
| 3/4 | 9 | 381 | 13.52 | 11.44 | 10.40 |
| | 10 | 383 | 13.78 | 11.70 | 10.40 |
| | 11 | 384 | 13.78 | 11.70 | 10.40 |
| | 12 | 385 | 14.04 | 11.70 | 10.40 |
| 4/5 | 9 | 387 | n/a | 11.96 | 10.40 |
| | 10 | 390 | n/a | 12.22 | 10.40 |
| | 11 | 392 | n/a | 12.48 | 10.40 |
| | 12 | 393 | n/a | 12.48 | 10.40 |
| 5/6 | 9 | 402 | n/a | n/a | 10.66 |
| | 10 | 406 | n/a | n/a | 10.92 |
| | 11 | 407 | n/a | n/a | 10.92 |
| | 12 | 408 | n/a | n/a | 11.18 |

* No equating was performed for S302

# Figure 8.5.2C
## Test Characteristic Curve: Read 9-12ABC S302
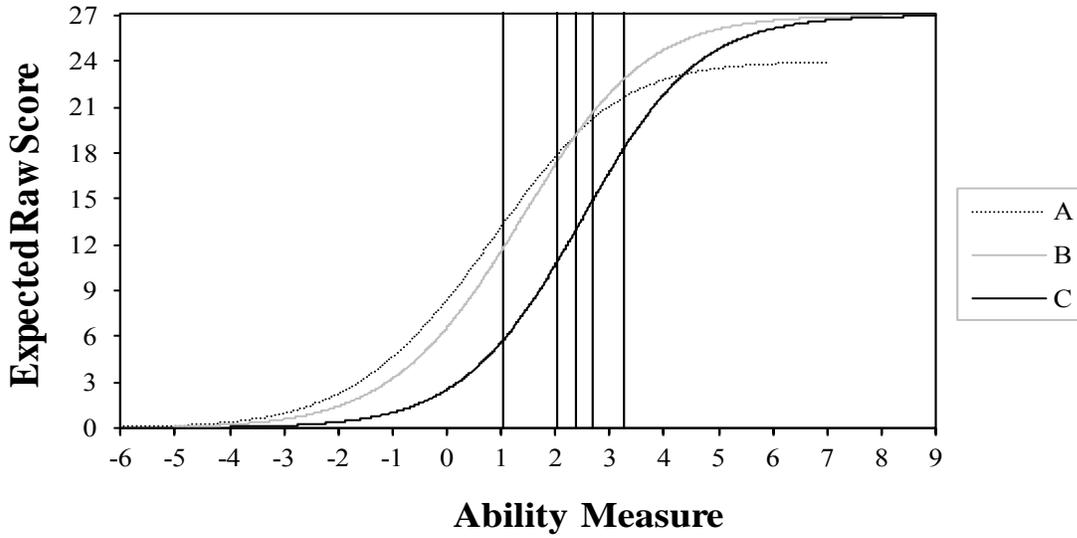


# Figure 8.5.2D
## Test Information Function: Read 9-12ABC S302
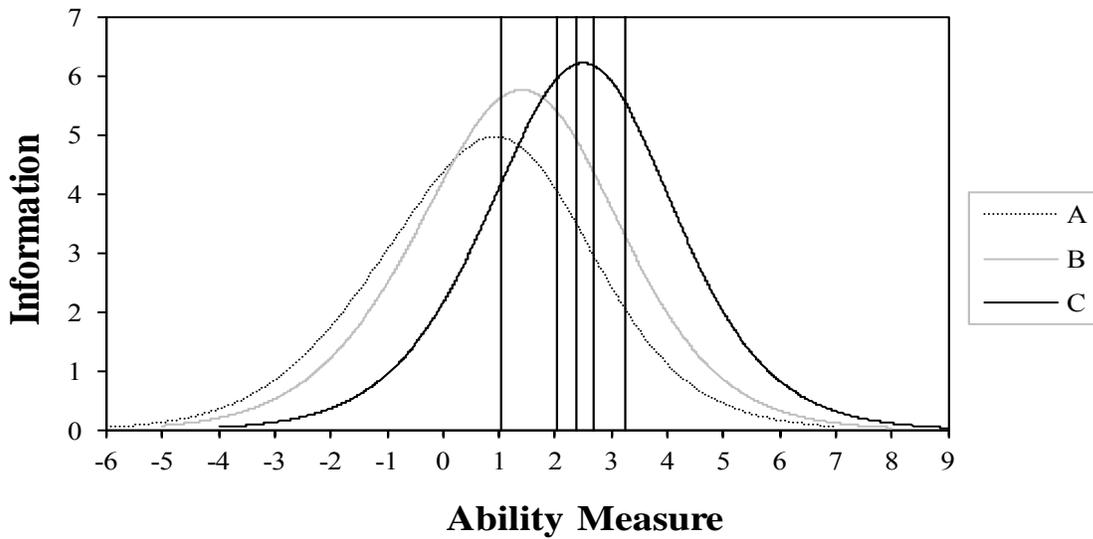
**Table 8.5.2D**
Weighted Reliability: Read 9-12 S302

| Tiers | No. of Students | Reliability | Reliability |
|---|---|---|---|
| A | 29,106 | 0.794 | |
| B | 76,072 | 0.816 | 0.800 |
| C | 92,859 | 0.789 | |

**Table 8.5.2E-1**
Accuracy and Consistency of Classification Indices: Read (Grade 9) S302

| Overall | Accuracy | Consistency | | | Kappa (k) | |
|---|---|---|---|---|---|---|
| Indices | 0.491 | 0.397 | | | 0.265 | |
| Conditional | Level | Accuracy | | | Consistency | |
| on Level | 1 | 0.752 | | | 0.597 | |
| | 2 | 0.630 | | | 0.499 | |
| | 3 | 0.379 | | | 0.294 | |
| | 4 | 0.203 | | | 0.154 | |
| | 5 | 0.369 | | | 0.298 | |
| | 6 | 0.635 | | | 0.463 | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.942 | 0.026 | 0.032 | 0.913 | |
| | 2/3 | 0.866 | 0.052 | 0.083 | 0.814 | |
| | 3/4 | 0.839 | 0.078 | 0.083 | 0.777 | |
| | 4/5 | 0.834 | 0.074 | 0.092 | 0.773 | |
| | 5/6 | 0.868 | 0.082 | 0.050 | 0.818 | |

**Table 8.5.2E-2**
Accuracy and Consistency of Classification Indices: Read (Grade 10) 302

| Overall | Accuracy | Consistency | | | Kappa (k) | |
|---|---|---|---|---|---|---|
| Indices | 0.488 | 0.396 | | | 0.258 | |
| Conditional | Level | Accuracy | | | Consistency | |
| on Level | 1 | 0.715 | | | 0.554 | |
| | 2 | 0.691 | | | 0.568 | |
| | 3 | 0.299 | | | 0.231 | |
| | 4 | 0.256 | | | 0.193 | |
| | 5 | 0.365 | | | 0.294 | |
| | 6 | 0.627 | | | 0.422 | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.940 | 0.030 | 0.030 | 0.909 | |
| | 2/3 | 0.850 | 0.055 | 0.096 | 0.795 | |
| | 3/4 | 0.838 | 0.076 | 0.086 | 0.776 | |
| | 4/5 | 0.841 | 0.066 | 0.094 | 0.779 | |
| | 5/6 | 0.886 | 0.083 | 0.032 | 0.843 | |

**Table 8.5.2E-3**

Accuracy and Consistency of Classification Indices: Read (Grade 11) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.493 | 0.399 | | 0.266 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.737 | | 0.577 | |
| | 2 | 0.644 | | 0.519 | |
| | 3 | 0.250 | | 0.190 | |
| | 4 | 0.252 | | 0.193 | |
| | 5 | 0.400 | | 0.320 | |
| | 6 | 0.642 | | 0.471 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.939 | 0.028 | 0.033 | 0.908 |
| | 2/3 | 0.855 | 0.059 | 0.086 | 0.801 |
| | 3/4 | 0.840 | 0.075 | 0.085 | 0.779 |
| | 4/5 | 0.836 | 0.075 | 0.089 | 0.774 |
| | 5/6 | 0.874 | 0.075 | 0.051 | 0.823 |

**Table 8.5.2E-4**

Accuracy and Consistency of Classification Indices: Read (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.475 | 0.384 | | 0.247 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.751 | | 0.605 | |
| | 2 | 0.633 | | 0.501 | |
| | 3 | 0.245 | | 0.184 | |
| | 4 | 0.189 | | 0.144 | |
| | 5 | 0.434 | | 0.365 | |
| | 6 | 0.510 | | 0.358 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.937 | 0.031 | 0.032 | 0.905 |
| | 2/3 | 0.851 | 0.054 | 0.095 | 0.798 |
| | 3/4 | 0.834 | 0.061 | 0.106 | 0.770 |
| | 4/5 | 0.823 | 0.076 | 0.101 | 0.755 |
| | 5/6 | 0.856 | 0.087 | 0.058 | 0.803 |

## 8.5.3    Writing 9-12

Figure 8.5.3A
Scale Scores: Writ 9-12 S302

Figure 8.5.3B
Proficiency Level: Writ 9-12 S302
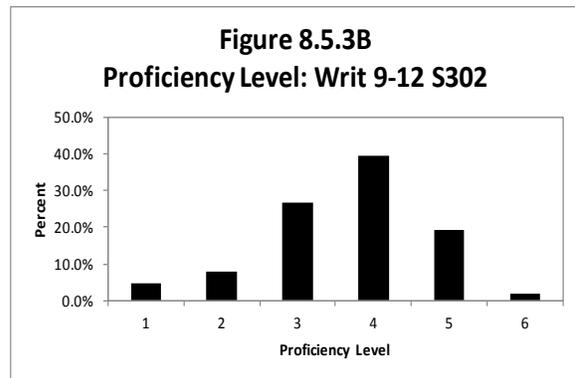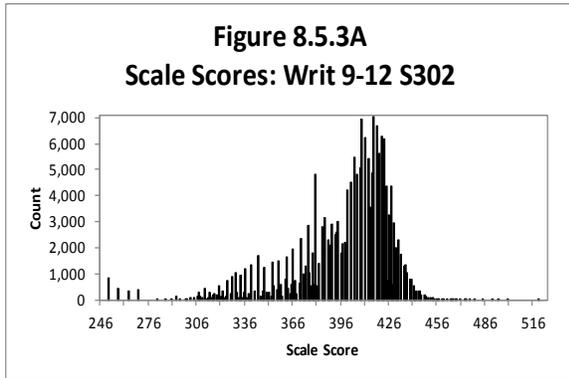
**Table 8.5.3A**

Scale Score Descriptive Statistics: Writ 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 9 | 79,499 | 251 | 473 | 394.61 | 33.76 |
| 10 | 50,096 | 257 | 494 | 396.88 | 30.82 |
| 11 | 38,170 | 263 | 500 | 400.95 | 29.00 |
| 12 | 29,947 | 269 | 519 | 402.05 | 29.45 |
| Total | 197,712 | 251 | 519 | 397.54 | 31.64 |

**Table 8.5.3B**

Proficiency Level Distribution: Writ 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3,581 | 4.5% | 2,438 | 4.9% | 1,818 | 4.8% | 1,830 | 6.1% | 9,667 | 4.9% |
| 2 | 6,914 | 8.7% | 3,862 | 7.7% | 2,393 | 6.3% | 2,334 | 7.8% | 15,503 | 7.8% |
| 3 | 16,134 | 20.3% | 13,349 | 26.6% | 11,897 | 31.2% | 11,213 | 37.4% | 52,593 | 26.6% |
| 4 | 25,717 | 32.3% | 21,588 | 43.1% | 17,976 | 47.1% | 13,024 | 43.5% | 78,305 | 39.6% |
| 5 | 24,413 | 30.7% | 8,261 | 16.5% | 3,823 | 10.0% | 1,443 | 4.8% | 37,940 | 19.2% |
| 6 | 2,740 | 3.4% | 598 | 1.2% | 263 | 0.7% | 103 | 0.3% | 3,704 | 1.9% |
| Total | 79,499 | 100.0% | 50,096 | 100.0% | 38,170 | 100.0% | 29,947 | 100.0% | 197,712 | 100.0% |

**Table 8.5.3C**
Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 S302*

| Proficiency Level | Grade | Cut Score | SEM | | |
|---|---|---|---|---|---|
| | | | Tier A | Tier B | Tier C |
| 1/2 | 9 | 327 | 9.02 | 6.53 | 6.53 |
| | 10 | 336 | 10.26 | 7.46 | 6.84 |
| | 11 | 344 | 11.19 | 8.09 | 7.77 |
| | 12 | 352 | 11.82 | 8.40 | 8.40 |
| 2/3 | 9 | 356 | 12.13 | 8.40 | 8.40 |
| | 10 | 363 | 12.13 | 8.40 | 8.40 |
| | 11 | 370 | 11.82 | 8.40 | 8.40 |
| | 12 | 377 | 11.51 | 8.09 | 8.09 |
| 3/4 | 9 | 389 | 11.19 | 8.09 | 8.09 |
| | 10 | 397 | 11.19 | 7.77 | 7.77 |
| | 11 | 404 | 10.88 | 7.46 | 7.77 |
| | 12 | 410 | 10.57 | 7.15 | 7.46 |
| 4/5 | 9 | 415 | n/a | 7.15 | 7.15 |
| | 10 | 422 | n/a | 6.84 | 7.15 |
| | 11 | 428 | n/a | 6.53 | 6.84 |
| | 12 | 434 | n/a | 6.53 | 6.53 |
| 5/6 | 9 | 435 | n/a | n/a | 6.53 |
| | 10 | 441 | n/a | n/a | 6.53 |
| | 11 | 447 | n/a | n/a | 6.22 |
| | 12 | 452 | n/a | n/a | 6.53 |

* No equating was performed for Writing Tier A S302

**Figure 8.5.3C**
Test Characteristic Curve: Writ 9-12ABC S302
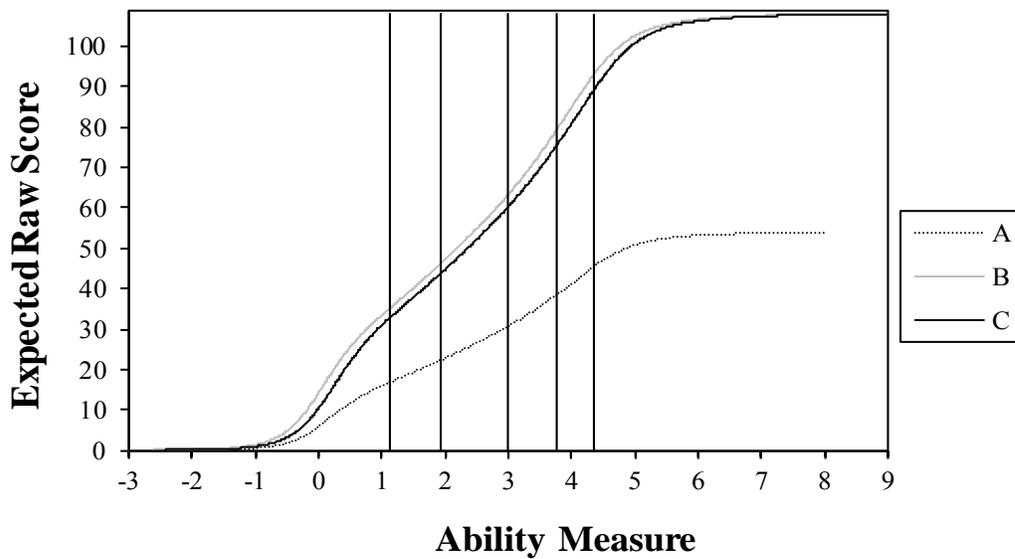


**Figure 8.5.3D**
Test Information Function: Writ 9-12ABC S302
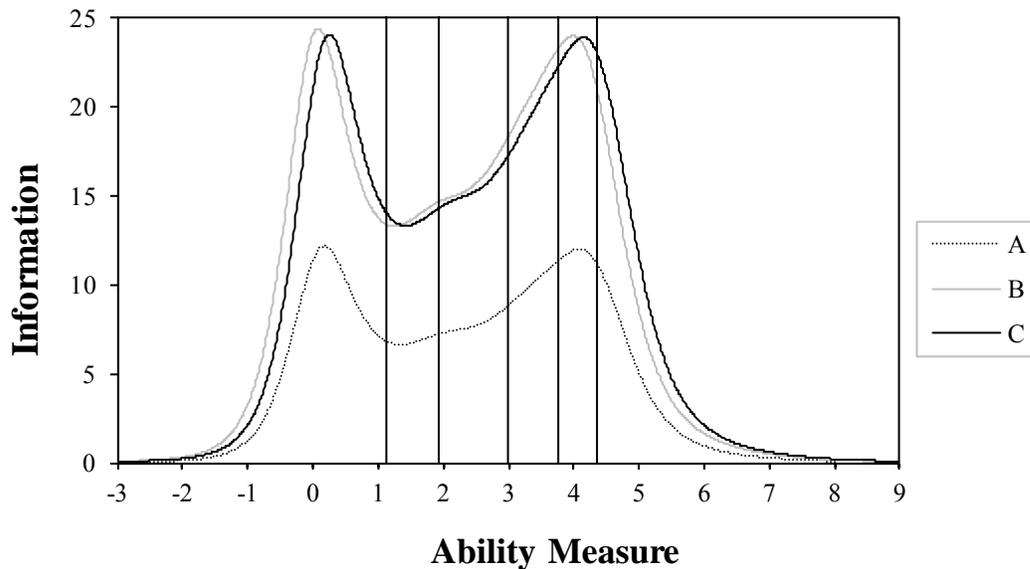
**Table 8.5.3D**

Weighted Reliability: Writ 9-12 S302

| Tiers | No. of Students | Reliability | Reliability |
|-------|----------------|-------------|-------------|
| A | 29,068 | 0.867 | |
| B | 75,983 | 0.936 | 0.916 |
| C | 92,661 | 0.915 | |

**Table 8.5.3E-1**

Accuracy and Consistency of Classification Indices: Writ (Grade 9) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| **Indices** | 0.676 | 0.567 | | 0.417 | |
| **Conditional** | **Level** | **Accuracy** | | **Consistency** | |
| **on Level** | 1 | 0.840 | | 0.746 | |
| | 2 | 0.735 | | 0.614 | |
| | 3 | 0.774 | | 0.668 | |
| | 4 | 0.632 | | 0.494 | |
| | 5 | 0.631 | | 0.576 | |
| | 6 | - | | 0.089 | |
| **Indices at** | | | **Accuracy** | | |
| **Cut Points** | | | **False** | **False** | |
| | **Cut Point** | **Accuracy** | **Positives** | **Negatives** | **Consistency** |
| | 1/2 | 0.986 | 0.007 | 0.007 | 0.979 |
| | 2/3 | 0.966 | 0.015 | 0.019 | 0.951 |
| | 3/4 | 0.934 | 0.023 | 0.043 | 0.907 |
| | 4/5 | 0.822 | 0.071 | 0.107 | 0.755 |
| | 5/6 | 0.966 | 0.034 | 0.000 | 0.952 |

**Table 8.5.3E-2**

Accuracy and Consistency of Classification Indices: Writ (Grade 10) S302

| Overall | Accuracy | Consistency | | Kappa (k) | |
|---------|----------|-------------|--|-----------|--|
| **Indices** | 0.695 | 0.607 | | 0.437 | |
| **Conditional** | **Level** | **Accuracy** | | **Consistency** | |
| **on Level** | 1 | 0.875 | | 0.797 | |
| | 2 | 0.717 | | 0.591 | |
| | 3 | 0.819 | | 0.713 | |
| | 4 | 0.633 | | 0.618 | |
| | 5 | - | | 0.354 | |
| | 6 | - | | 0.038 | |
| **Indices at** | | | **Accuracy** | | |
| **Cut Points** | | | **False** | **False** | |
| | **Cut Point** | **Accuracy** | **Positives** | **Negatives** | **Consistency** |
| | 1/2 | 0.987 | 0.006 | 0.007 | 0.982 |
| | 2/3 | 0.969 | 0.015 | 0.016 | 0.955 |
| | 3/4 | 0.916 | 0.026 | 0.058 | 0.882 |
| | 4/5 | 0.823 | 0.177 | 0.000 | 0.787 |
| | 5/6 | 0.988 | 0.012 | 0.000 | 0.986 |

Return to Visual Table of Tables and Figures

**Table 8.5.3E-3**
Accuracy and Consistency of Classification Indices: Writ (Grade 11) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.749 | 0.655 | | 0.470 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.895 | | 0.831 | |
| | 2 | 0.688 | | 0.557 | |
| | 3 | 0.845 | | 0.722 | |
| | 4 | 0.705 | | 0.680 | |
| | 5 | - | | 0.212 | |
| | 6 | - | | 0.227 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.989 | 0.005 | 0.006 | 0.985 |
| | 2/3 | 0.973 | 0.015 | 0.013 | 0.960 |
| | 3/4 | 0.894 | 0.027 | 0.078 | 0.851 |
| | 4/5 | 0.893 | 0.107 | 0.000 | 0.853 |
| | 5/6 | 0.993 | 0.007 | 0.000 | 0.993 |

**Table 8.5.3E-4**
Accuracy and Consistency of Classification Indices: Writ (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.723 | 0.633 | | 0.430 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.892 | | 0.830 | |
| | 2 | 0.742 | | 0.619 | |
| | 3 | 0.819 | | 0.621 | |
| | 4 | 0.664 | | 0.638 | |
| | 5 | - | | 0.091 | |
| | 6 | - | | 0.000 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.987 | 0.007 | 0.006 | 0.982 |
| | 2/3 | 0.970 | 0.013 | 0.016 | 0.958 |
| | 3/4 | 0.816 | 0.030 | 0.154 | 0.747 |
| | 4/5 | 0.948 | 0.052 | 0.000 | 0.933 |
| | 5/6 | 0.997 | 0.003 | 0.000 | 0.997 |

## 8.5.4    Speaking 9-12



**Figure 8.5.4A**
**Scale Scores: Spek 9-12 S302**



**Figure 8.5.4B**
**Proficiency Level: Spek 9-12 S302**
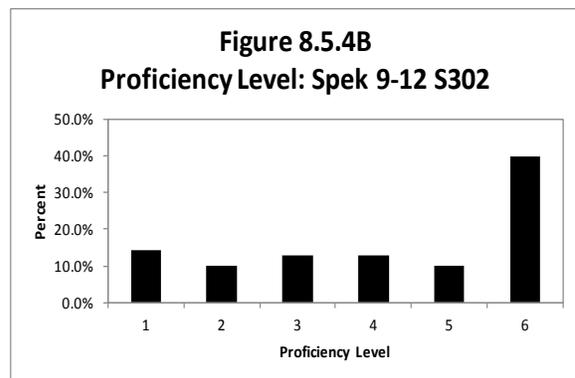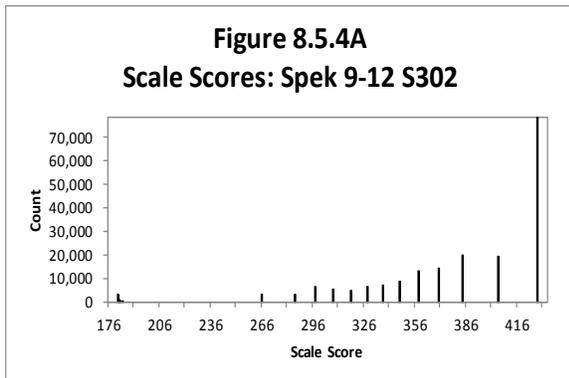
**Table 8.5.4A**

Scale Score Descriptive Statistics: Spek 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 9 | 79,452 | 181 | 428 | 376.09 | 60.72 |
| 10 | 50,051 | 182 | 428 | 379.28 | 53.97 |
| 11 | 38,164 | 183 | 428 | 383.57 | 50.35 |
| 12 | 29,948 | 184 | 428 | 387.37 | 48.45 |
| **Total** | 197,615 | 181 | 428 | 380.05 | 55.52 |

**Table 8.5.4B**

Proficiency Level Distribution: Spek 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|-------|------|---------|------|---------|------|---------|------|---------|------|---------|
| 1 | 14,234 | 17.9% | 7,263 | 14.5% | 4,211 | 11.0% | 2,650 | 8.8% | 28,358 | 14.4% |
| 2 | 5,023 | 6.3% | 6,367 | 12.7% | 4,974 | 13.0% | 3,643 | 12.2% | 20,007 | 10.1% |
| 3 | 7,983 | 10.0% | 7,351 | 14.7% | 5,875 | 15.4% | 4,536 | 15.1% | 25,745 | 13.0% |
| 4 | 13,195 | 16.6% | 4,964 | 9.9% | 3,980 | 10.4% | 3,163 | 10.6% | 25,302 | 12.8% |
| 5 | 7,799 | 9.8% | 4,958 | 9.9% | 3,730 | 9.8% | 3,058 | 10.2% | 19,545 | 9.9% |
| 6 | 31,218 | 39.3% | 19,148 | 38.3% | 15,394 | 40.3% | 12,898 | 43.1% | 78,658 | 39.8% |
| Total | 79,452 | 100.0% | 50,051 | 100.0% | 38,164 | 100.0% | 29,948 | 100.0% | 197,615 | 100.0% |

**Table 8.5.4C**
Conditional Standard Error of Measurement at
Cut Scores: Spek 9-12 S302

| Proficiency Level | Grade | Cut Score | SEM |
|---|---|---|---|
| 1/2 | 9 | 319 | 20.49 |
| | 10 | 321 | 20.69 |
| | 11 | 322 | 20.69 |
| | 12 | 323 | 20.89 |
| 2/3 | 9 | 347 | 22.49 |
| | 10 | 351 | 22.69 |
| | 11 | 354 | 22.90 |
| | 12 | 357 | 23.10 |
| 3/4 | 9 | 366 | 23.90 |
| | 10 | 371 | 24.10 |
| | 11 | 377 | 24.90 |
| | 12 | 384 | 25.91 |
| 4/5 | 9 | 388 | 26.51 |
| | 10 | 393 | 27.92 |
| | 11 | 399 | 29.32 |
| | 12 | 405 | 30.53 |
| 5/6 | 9 | 407 | 30.93 |
| | 10 | 412 | 32.13 |
| | 11 | 416 | 32.94 |
| | 12 | 421 | 33.54 |

**Figure 8.5.4C**

Test Characteristic Curve: Spek 9-12 S302
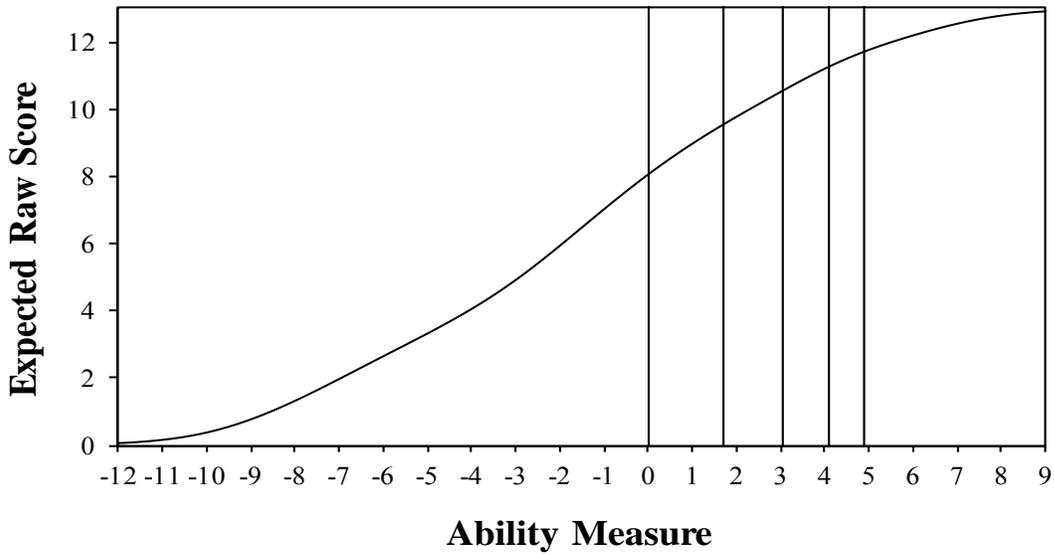


**Figure 8.5.4D**

Test Information Function: Spek 9-12 S302
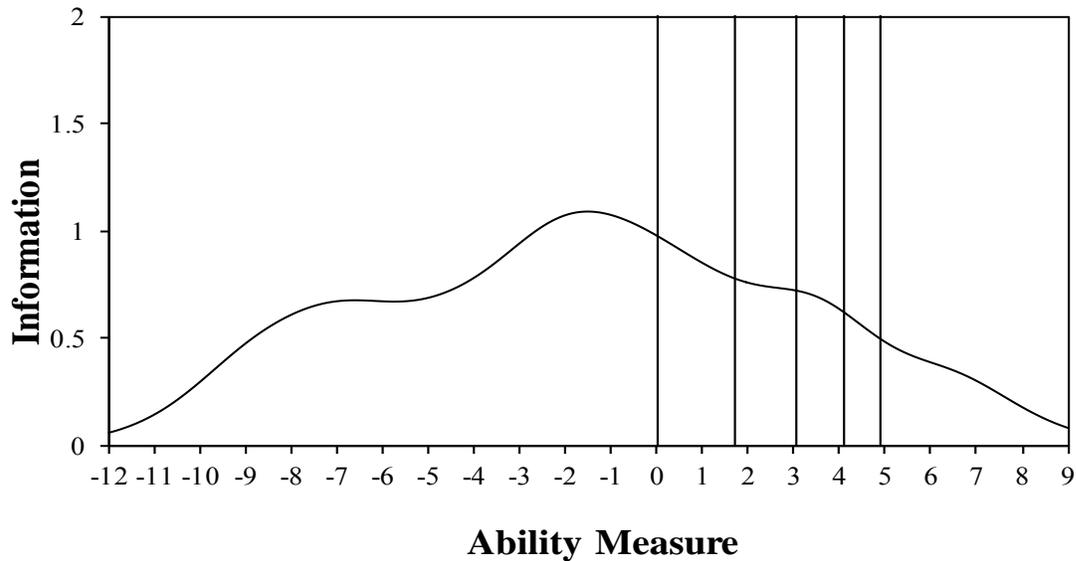
**Table 8.5.4D**

Reliability: Spek 9-12 S302

| Tiers | No. of Students | Reliability |
|---|---|---|
| -- | 197,615 | 0.924 |

**Table 8.5.4E-1**

Accuracy and Consistency of Classification Indices: Spek (Grade 9) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.656 | 0.574 | | 0.444 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.885 | | 0.809 | |
| | 2 | 0.319 | | 0.231 | |
| | 3 | 0.377 | | 0.293 | |
| | 4 | 0.505 | | 0.402 | |
| | 5 | 0.259 | | 0.179 | |
| | 6 | 0.858 | | 0.790 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | | | Consistency |
| | 1/2 | 0.956 | 0.020 | 0.024 | 0.934 |
| | 2/3 | 0.932 | 0.041 | 0.027 | 0.904 |
| | 3/4 | 0.906 | 0.045 | 0.049 | 0.877 |
| | 4/5 | 0.900 | 0.027 | 0.073 | 0.862 |
| | 5/6 | 0.896 | 0.047 | 0.057 | 0.836 |

**Table 8.5.4E-2**

Accuracy and Consistency of Classification Indices: Spek (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.659 | 0.578 | | 0.463 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.817 | | 0.722 | |
| | 2 | 0.508 | | 0.405 | |
| | 3 | 0.497 | | 0.403 | |
| | 4 | 0.361 | | 0.268 | |
| | 5 | 0.359 | | 0.248 | |
| | 6 | 0.914 | | 0.859 | |
| Indices at Cut Points | | Accuracy | | | |
| | | | False Positives | False Negatives | |
| | Cut Point | Accuracy | | | Consistency |
| | 1/2 | 0.949 | 0.027 | 0.023 | 0.926 |
| | 2/3 | 0.914 | 0.043 | 0.044 | 0.886 |
| | 3/4 | 0.910 | 0.027 | 0.063 | 0.879 |
| | 4/5 | 0.934 | 0.028 | 0.037 | 0.898 |
| | 5/6 | 0.912 | 0.057 | 0.031 | 0.874 |

**Table 8.5.4E-3**

Accuracy and Consistency of Classification Indices: Spek (Grade 11) S302

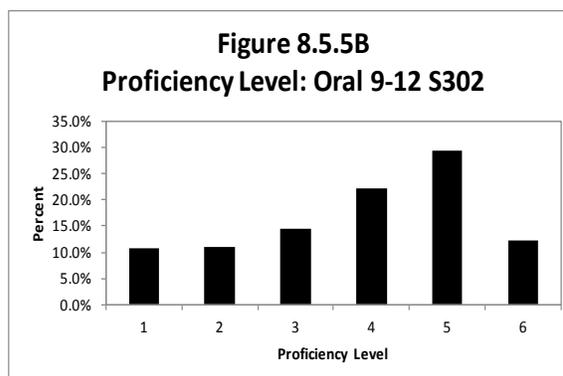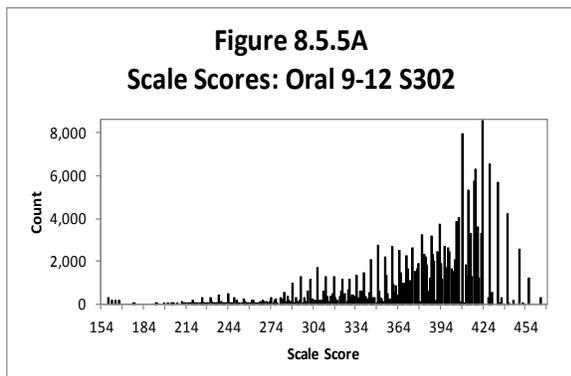| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.666 | 0.582 | | 0.463 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.796 | | 0.691 | |
| | 2 | 0.566 | | 0.455 | |
| | 3 | 0.543 | | 0.443 | |
| | 4 | 0.397 | | 0.292 | |
| | 5 | 0.346 | | 0.242 | |
| | 6 | 0.925 | | 0.873 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.958 | 0.024 | 0.019 | 0.938 |
| | 2/3 | 0.922 | 0.037 | 0.041 | 0.896 |
| | 3/4 | 0.918 | 0.025 | 0.056 | 0.888 |
| | 4/5 | 0.936 | 0.029 | 0.035 | 0.901 |
| | 5/6 | 0.905 | 0.068 | 0.027 | 0.868 |


**Table 8.5.4E-4**

Accuracy and Consistency of Classification Indices: Spek (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.640 | 0.535 | | 0.407 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.789 | | 0.683 | |
| | 2 | 0.607 | | 0.490 | |
| | 3 | 0.579 | | 0.475 | |
| | 4 | 0.432 | | 0.307 | |
| | 5 | 0.272 | | 0.186 | |
| | 6 | 0.905 | | 0.847 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.965 | 0.020 | 0.015 | 0.949 |
| | 2/3 | 0.933 | 0.030 | 0.037 | 0.910 |
| | 3/4 | 0.928 | 0.024 | 0.048 | 0.901 |
| | 4/5 | 0.938 | 0.027 | 0.035 | 0.901 |
| | 5/6 | 0.853 | 0.113 | 0.033 | 0.796 |

## 8.5.5    Oral Language Composite 9-12

**Figure 8.5.5A**
**Scale Scores: Oral 9-12 S302**



**Figure 8.5.5B**
**Proficiency Level: Oral 9-12 S302**



**Table 8.5.5A**

Scale Score Descriptive Statistics: Oral 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 9 | 79,128 | 159 | 464 | 379.05 | 50.90 |
| 10 | 49,875 | 161 | 464 | 382.45 | 44.66 |
| 11 | 37,978 | 164 | 464 | 387.04 | 41.83 |
| 12 | 29,722 | 166 | 464 | 389.60 | 40.57 |
| Total | 196,703 | 159 | 464 | 383.05 | 46.37 |

**Table 8.5.5B**

Proficiency Level Distribution: Oral 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|-------|------|---------|------|---------|------|---------|------|---------|------|---------|
| 1 | 10,466 | 13.2% | 5,041 | 10.1% | 3,198 | 8.4% | 2,322 | 7.8% | 21,027 | 10.7% |
| 2 | 7,972 | 10.1% | 6,061 | 12.2% | 4,301 | 11.3% | 3,268 | 11.0% | 21,602 | 11.0% |
| 3 | 8,051 | 10.2% | 7,573 | 15.2% | 6,684 | 17.6% | 5,998 | 20.2% | 28,306 | 14.4% |
| 4 | 15,346 | 19.4% | 10,984 | 22.0% | 9,328 | 24.6% | 8,010 | 26.9% | 43,668 | 22.2% |
| 5 | 24,890 | 31.5% | 15,172 | 30.4% | 10,090 | 26.6% | 7,795 | 26.2% | 57,947 | 29.5% |
| 6 | 12,403 | 15.7% | 5,044 | 10.1% | 4,377 | 11.5% | 2,329 | 7.8% | 24,153 | 12.3% |
| Total | 79,128 | 100.0% | 49,875 | 100.0% | 37,978 | 100.0% | 29,722 | 100.0% | 196,703 | 100.0% |

**Table 8.5.5C**
n/a

**Figure 8.5.5C**
n/a

**Figure 8.5.5D**
n/a

**Table 8.5.5D**
Oral Composite Reliability: Oral 9-12 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.50 | 2311.716 | 0.688 |
| Speaking | 0.50 | 3077.424 | 0.924 |
| Oral | | 2149.264 | 0.889 |

*Variances from students who had results in all four domains

**Table 8.5.5E-1**
Accuracy and Consistency of Classification Indices: Oral (Grade 9) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.550 | 0.449 | | 0.318 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.879 | | 0.801 | |
| | 2 | 0.548 | | 0.416 | |
| | 3 | 0.403 | | 0.288 | |
| | 4 | 0.478 | | 0.350 | |
| | 5 | 0.547 | | 0.489 | |
| | 6 | 0.474 | | 0.359 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.962 | 0.015 | 0.023 | 0.946 |
| | 2/3 | 0.943 | 0.024 | 0.034 | 0.917 |
| | 3/4 | 0.916 | 0.040 | 0.043 | 0.879 |
| | 4/5 | 0.859 | 0.058 | 0.083 | 0.803 |
| | 5/6 | 0.836 | 0.090 | 0.074 | 0.789 |

**Table 8.5.5E-2**
Accuracy and Consistency of Classification Indices: Oral (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.580 | 0.456 | | 0.324 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.844 | | 0.746 | |
| | 2 | 0.608 | | 0.478 | |
| | 3 | 0.521 | | 0.396 | |
| | 4 | 0.493 | | 0.368 | |
| | 5 | 0.576 | | 0.523 | |
| | 6 | - | | 0.263 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.966 | 0.015 | 0.019 | 0.951 |
| | 2/3 | 0.935 | 0.028 | 0.037 | 0.909 |
| | 3/4 | 0.907 | 0.036 | 0.057 | 0.869 |
| | 4/5 | 0.854 | 0.062 | 0.083 | 0.793 |
| | 5/6 | 0.899 | 0.101 | 0.000 | 0.843 |

**Table 8.5.5E-3**
Accuracy and Consistency of Classification Indices: Oral (Grade 11) S302
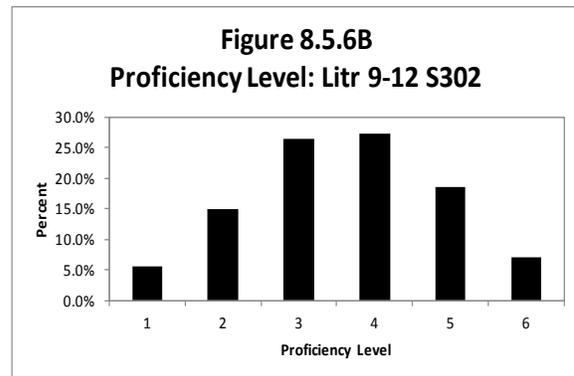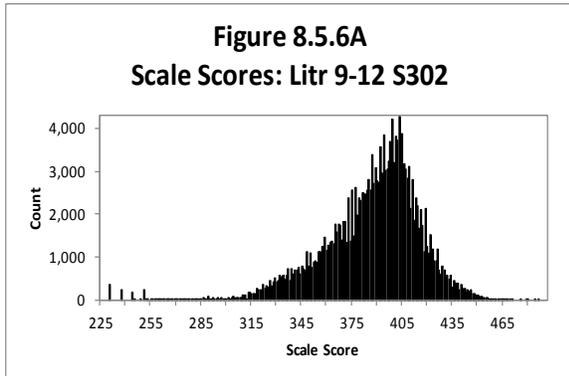
| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.558 | 0.448 | | 0.310 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.834 | | 0.730 | |
| | 2 | 0.603 | | 0.471 | |
| | 3 | 0.584 | | 0.457 | |
| | 4 | 0.511 | | 0.382 | |
| | 5 | 0.510 | | 0.463 | |
| | 6 | - | | 0.282 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.971 | 0.014 | 0.015 | 0.957 |
| | 2/3 | 0.936 | 0.029 | 0.035 | 0.911 |
| | 3/4 | 0.904 | 0.032 | 0.064 | 0.867 |
| | 4/5 | 0.845 | 0.062 | 0.093 | 0.775 |
| | 5/6 | 0.885 | 0.115 | 0.000 | 0.844 |

**Table 8.5.5E-4**

Accuracy and Consistency of Classification Indices: Oral (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.547 | 0.442 | | 0.290 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.814 | | 0.710 | |
| | 2 | 0.620 | | 0.487 | |
| | 3 | 0.632 | | 0.489 | |
| | 4 | 0.462 | | 0.369 | |
| | 5 | 0.500 | | 0.449 | |
| | 6 | - | | 0.165 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.973 | 0.015 | 0.012 | 0.959 |
| | 2/3 | 0.938 | 0.026 | 0.036 | 0.916 |
| | 3/4 | 0.896 | 0.026 | 0.079 | 0.855 |
| | 4/5 | 0.791 | 0.089 | 0.120 | 0.722 |
| | 5/6 | 0.922 | 0.078 | 0.000 | 0.890 |

## 8.5.6    Literacy Composite 9-12



**Figure 8.5.6A**
**Scale Scores: Litr 9-12 S302**



**Figure 8.5.6B**
**Proficiency Level: Litr 9-12 S302**

**Table 8.5.6A**

Scale Score Descriptive Statistics: Litr 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|------|-----------|
| 9 | 79,349 | 230 | 471 | 383.90 | 30.76 |
| 10 | 49,998 | 237 | 479 | 386.19 | 28.25 |
| 11 | 38,092 | 244 | 484 | 390.50 | 27.26 |
| 12 | 29,877 | 251 | 486 | 391.62 | 27.62 |
| Total | 197,316 | 230 | 486 | 386.92 | 29.18 |

**Table 8.5.6B**

Proficiency Level Distribution: Litr 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|-------|-------|---------|-------|---------|-------|---------|-------|---------|-------|---------|
| 1 | 4,722 | 6.0% | 2,682 | 5.4% | 1,994 | 5.2% | 1,947 | 6.5% | 11,345 | 5.7% |
| 2 | 10,892 | 13.7% | 7,864 | 15.7% | 5,839 | 15.3% | 4,862 | 16.3% | 29,457 | 14.9% |
| 3 | 18,306 | 23.1% | 13,851 | 27.7% | 10,751 | 28.2% | 9,186 | 30.7% | 52,094 | 26.4% |
| 4 | 19,918 | 25.1% | 13,962 | 27.9% | 11,336 | 29.8% | 8,670 | 29.0% | 53,886 | 27.3% |
| 5 | 18,247 | 23.0% | 8,840 | 17.7% | 5,897 | 15.5% | 3,791 | 12.7% | 36,775 | 18.6% |
| 6 | 7,264 | 9.2% | 2,799 | 5.6% | 2,275 | 6.0% | 1,421 | 4.8% | 13,759 | 7.0% |
| Total | 79,349 | 100.0% | 49,998 | 100.0% | 38,092 | 100.0% | 29,877 | 100.0% | 197,316 | 100.0% |

**Table 8.5.6C**
n/a

**Figure 8.5.6C**
n/a

**Figure 8.5.6D**
n/a

**Table 8.5.6D**
Literacy Composite Reliability: Litr 9-12 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Reading | 0.50 | 1010.989 | 0.800 |
| Writing | 0.50 | 996.222 | 0.916 |
| Literacy | | 850.309 | 0.916 |

*Variances from students who had results in all four domains

**Table 8.5.6E-1**
Accuracy and Consistency of Classification Indices: Litr (Grade 9) S302

| Overall Indices | Accuracy | | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|---|
| | 0.636 | | 0.535 | | 0.416 | |
| Conditional on Level | Level | | Accuracy | | Consistency | |
| | 1 | | 0.834 | | 0.733 | |
| | 2 | | 0.751 | | 0.636 | |
| | 3 | | 0.691 | | 0.577 | |
| | 4 | | 0.613 | | 0.487 | |
| | 5 | | 0.537 | | 0.484 | |
| | 6 | | - | | 0.366 | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.980 | 0.010 | 0.010 | 0.970 | |
| | 2/3 | 0.948 | 0.022 | 0.030 | 0.927 | |
| | 3/4 | 0.909 | 0.041 | 0.050 | 0.873 | |
| | 4/5 | 0.885 | 0.043 | 0.072 | 0.837 | |
| | 5/6 | 0.908 | 0.092 | 0.000 | 0.892 | |

**Table 8.5.6E-2**

Accuracy and Consistency of Classification Indices: Litr (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.656 | 0.553 | | 0.429 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.822 | | 0.716 | |
| | 2 | 0.770 | | 0.658 | |
| | 3 | 0.724 | | 0.616 | |
| | 4 | 0.628 | | 0.506 | |
| | 5 | 0.518 | | 0.450 | |
| | 6 | - | | 0.280 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.981 | 0.010 | 0.009 | 0.972 |
| | 2/3 | 0.942 | 0.026 | 0.032 | 0.918 |
| | 3/4 | 0.899 | 0.041 | 0.060 | 0.860 |
| | 4/5 | 0.885 | 0.043 | 0.072 | 0.836 |
| | 5/6 | 0.944 | 0.056 | 0.000 | 0.936 |

**Table 8.5.6E-3**

Accuracy and Consistency of Classification Indices: Litr (Grade 11) S302
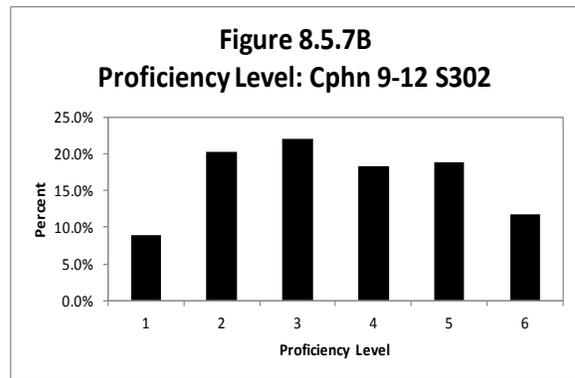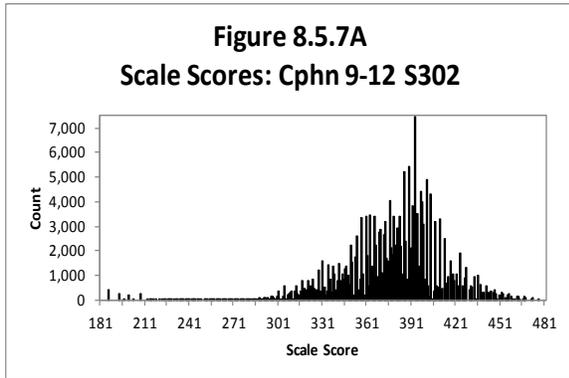
| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.650 | 0.548 | | 0.420 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.822 | | 0.717 | |
| | 2 | 0.766 | | 0.653 | |
| | 3 | 0.731 | | 0.622 | |
| | 4 | 0.643 | | 0.523 | |
| | 5 | 0.465 | | 0.400 | |
| | 6 | - | | 0.289 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.982 | 0.010 | 0.009 | 0.973 |
| | 2/3 | 0.943 | 0.026 | 0.031 | 0.919 |
| | 3/4 | 0.897 | 0.041 | 0.062 | 0.858 |
| | 4/5 | 0.882 | 0.042 | 0.076 | 0.832 |
| | 5/6 | 0.940 | 0.060 | 0.000 | 0.933 |

**Table 8.5.6E-4**

Accuracy and Consistency of Classification Indices: Litr (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | | |
|---|---|---|---|---|---|---|
| | 0.632 | 0.536 | | 0.397 | | |
| Conditional on Level | Level | Accuracy | | Consistency | | |
| | 1 | 0.827 | | 0.732 | | |
| | 2 | 0.757 | | 0.640 | | |
| | 3 | 0.742 | | 0.623 | | |
| | 4 | 0.558 | | 0.478 | | |
| | 5 | 0.396 | | 0.328 | | |
| | 6 | - | | 0.208 | | |
| Indices at Cut Points | | Accuracy | | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency | |
| | 1/2 | 0.979 | 0.012 | 0.009 | 0.969 | |
| | 2/3 | 0.937 | 0.028 | 0.035 | 0.911 | |
| | 3/4 | 0.884 | 0.035 | 0.080 | 0.841 | |
| | 4/5 | 0.861 | 0.083 | 0.056 | 0.817 | |
| | 5/6 | 0.952 | 0.048 | 0.000 | 0.950 | |

## 8.5.7    Comprehension Composite 9-12



**Figure 8.5.7A**
**Scale Scores: Cphn 9-12 S302**



**Figure 8.5.7B**
**Proficiency Level: Cphn 9-12 S302**

**Table 8.5.7A**

Scale Score Descriptive Statistics: Cphn 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|
| 9 | 79,449 | 186 | 477 | 375.34 | 35.56 |
| 10 | 50,057 | 193 | 477 | 378.04 | 32.81 |
| 11 | 38,136 | 200 | 477 | 382.67 | 32.26 |
| 12 | 29,921 | 208 | 477 | 383.88 | 32.75 |
| Total | 197,563 | 186 | 477 | 378.73 | 34.01 |

**Table 8.5.7B**

Proficiency Level Distribution: Cphn 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7,850 | 9.9% | 3,870 | 7.7% | 3,059 | 8.0% | 2,849 | 9.5% | 17,628 | 8.9% |
| 2 | 13,916 | 17.5% | 11,042 | 22.1% | 8,763 | 23.0% | 6,231 | 20.8% | 39,952 | 20.2% |
| 3 | 18,466 | 23.2% | 12,115 | 24.2% | 6,947 | 18.2% | 5,916 | 19.8% | 43,444 | 22.0% |
| 4 | 13,021 | 16.4% | 8,939 | 17.9% | 8,124 | 21.3% | 5,991 | 20.0% | 36,075 | 18.3% |
| 5 | 16,059 | 20.2% | 9,092 | 18.2% | 6,641 | 17.4% | 5,410 | 18.1% | 37,202 | 18.8% |
| 6 | 10,137 | 12.8% | 4,999 | 10.0% | 4,602 | 12.1% | 3,524 | 11.8% | 23,262 | 11.8% |
| Total | 79,449 | 100.0% | 50,057 | 100.0% | 38,136 | 100.0% | 29,921 | 100.0% | 197,563 | 100.0% |

**Table 8.5.7C**
n/a

**Figure 8.5.7C**
n/a

**Figure 8.5.7D**
n/a

**Table 8.5.7D**
Comprehension Composite Reliability: Cphn 9-12 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.30 | 2311.716 | 0.688 |
| Reading | 0.70 | 1010.989 | 0.800 |
| Comprehension | | 1155.625 | 0.858 |

*Variances from students who had results in all four domains

**Table 8.5.7E-1**
Accuracy and Consistency of Classification Indices: Cphn (Grade 9) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.525 | 0.426 | | 0.302 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.821 | | 0.699 | |
| | 2 | 0.631 | | 0.498 | |
| | 3 | 0.541 | | 0.425 | |
| | 4 | 0.351 | | 0.266 | |
| | 5 | 0.437 | | 0.367 | |
| | 6 | 0.591 | | 0.403 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.961 | 0.017 | 0.023 | 0.943 |
| | 2/3 | 0.907 | 0.041 | 0.051 | 0.869 |
| | 3/4 | 0.862 | 0.056 | 0.081 | 0.811 |
| | 4/5 | 0.854 | 0.059 | 0.087 | 0.796 |
| | 5/6 | 0.885 | 0.087 | 0.028 | 0.848 |

**Table 8.5.7E-2**
Accuracy and Consistency of Classification Indices: Cphn (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.520 | 0.424 | | 0.293 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.765 | | 0.625 | |
| | 2 | 0.698 | | 0.571 | |
| | 3 | 0.534 | | 0.421 | |
| | 4 | 0.371 | | 0.282 | |
| | 5 | 0.424 | | 0.362 | |
| | 6 | - | | 0.337 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.964 | 0.019 | 0.017 | 0.946 |
| | 2/3 | 0.894 | 0.044 | 0.061 | 0.854 |
| | 3/4 | 0.857 | 0.050 | 0.093 | 0.805 |
| | 4/5 | 0.857 | 0.057 | 0.086 | 0.795 |
| | 5/6 | 0.900 | 0.100 | 0.000 | 0.876 |

**Table 8.5.7E-3**
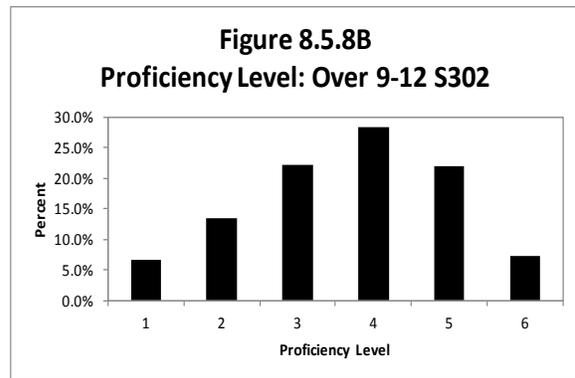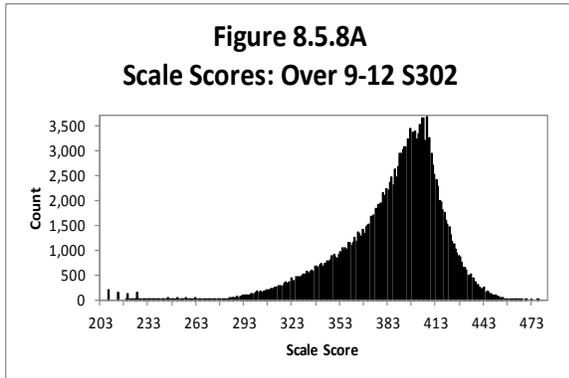Accuracy and Consistency of Classification Indices: Cphn (Grade 11) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.497 | 0.414 | | 0.286 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.749 | | 0.609 | |
| | 2 | 0.710 | | 0.581 | |
| | 3 | 0.425 | | 0.328 | |
| | 4 | 0.439 | | 0.336 | |
| | 5 | 0.380 | | 0.327 | |
| | 6 | 3.214 | | 0.388 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.962 | 0.021 | 0.017 | 0.941 |
| | 2/3 | 0.892 | 0.042 | 0.066 | 0.852 |
| | 3/4 | 0.863 | 0.055 | 0.082 | 0.811 |
| | 4/5 | 0.852 | 0.050 | 0.098 | 0.792 |
| | 5/6 | 0.879 | 0.121 | 0.000 | 0.857 |

**Table 8.5.7E-4**
Accuracy and Consistency of Classification Indices: Cphn (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.475 | 0.388 | | 0.256 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.761 | | 0.632 | |
| | 2 | 0.666 | | 0.533 | |
| | 3 | 0.443 | | 0.336 | |
| | 4 | 0.369 | | 0.285 | |
| | 5 | 0.378 | | 0.326 | |
| | 6 | - | | 0.302 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.958 | 0.024 | 0.018 | 0.936 |
| | 2/3 | 0.888 | 0.041 | 0.071 | 0.850 |
| | 3/4 | 0.850 | 0.041 | 0.109 | 0.796 |
| | 4/5 | 0.828 | 0.059 | 0.113 | 0.758 |
| | 5/6 | 0.882 | 0.118 | 0.000 | 0.850 |

## 8.5.8    Overall Composite 9-12



**Figure 8.5.8A**
**Scale Scores: Over 9-12 S302**



**Figure 8.5.8B**
**Proficiency Level: Over 9-12 S302**

**Table 8.5.8A**

Scale Score Descriptive Statistics: Over 9-12 S302

| Grade | No. of Students | Min. | Max. | Mean | Std. Dev. |
|-------|-----------------|------|------|--------|-----------|
| 9 | 78,900 | 208 | 465 | 382.26 | 34.84 |
| 10 | 49,719 | 214 | 469 | 384.87 | 30.90 |
| 11 | 37,857 | 220 | 476 | 389.27 | 29.24 |
| 12 | 29,583 | 226 | 477 | 390.85 | 28.98 |
| Total | 196,059 | 208 | 477 | 385.57 | 32.15 |

**Table 8.5.8B**

Proficiency Level Distribution: Over 9-12 S302

| Level | Grade 9 Count | Grade 9 Percent | Grade 10 Count | Grade 10 Percent | Grade 11 Count | Grade 11 Percent | Grade 12 Count | Grade 12 Percent | Total Count | Total Percent |
|-------|---------|---------|----------|----------|----------|----------|----------|----------|---------|---------|
| 1 | 6,432 | 8.2% | 2,920 | 5.9% | 2,026 | 5.4% | 1,728 | 5.8% | 13,106 | 6.7% |
| 2 | 9,873 | 12.5% | 7,296 | 14.7% | 5,329 | 14.1% | 4,030 | 13.6% | 26,528 | 13.5% |
| 3 | 14,283 | 18.1% | 11,586 | 23.3% | 9,388 | 24.8% | 8,509 | 28.8% | 43,766 | 22.3% |
| 4 | 19,591 | 24.8% | 14,391 | 28.9% | 11,763 | 31.1% | 9,706 | 32.8% | 55,451 | 28.3% |
| 5 | 20,901 | 26.5% | 10,637 | 21.4% | 7,125 | 18.8% | 4,297 | 14.5% | 42,960 | 21.9% |
| 6 | 7,820 | 9.9% | 2,889 | 5.8% | 2,226 | 5.9% | 1,313 | 4.4% | 14,248 | 7.3% |
| Total | 78,900 | 100.0% | 49,719 | 100.0% | 37,857 | 100.0% | 29,583 | 100.0% | 196,059 | 100.0% |

**Table 8.5.8C**
n/a

**Figure 8.5.8C**
n/a

**Figure 8.5.8D**
n/a

**Table 8.5.8D**
Overall Composite Reliability: Over 9-12 S302

| Component | Weight | Variance | Reliability |
|---|---|---|---|
| Listening | 0.15 | 2311.716 | 0.688 |
| Reading | 0.35 | 1010.989 | 0.800 |
| Speaking | 0.15 | 3077.424 | 0.924 |
| Writing | 0.35 | 996.222 | 0.916 |
| Overall Composite | | 1033.737 | 0.945 |

**Table 8.5.8E-1**
Accuracy and Consistency of Classification Indices: Over (Grade 9) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.679 | 0.588 | | 0.485 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.889 | | 0.822 | |
| | 2 | 0.767 | | 0.665 | |
| | 3 | 0.693 | | 0.579 | |
| | 4 | 0.693 | | 0.572 | |
| | 5 | 0.594 | | 0.556 | |
| | 6 | - | | 0.411 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.980 | 0.009 | 0.011 | 0.972 |
| | 2/3 | 0.960 | 0.017 | 0.023 | 0.943 |
| | 3/4 | 0.931 | 0.035 | 0.034 | 0.903 |
| | 4/5 | 0.906 | 0.036 | 0.058 | 0.868 |
| | 5/6 | 0.901 | 0.099 | 0.000 | 0.887 |

**Table 8.5.8E-2**

Accuracy and Consistency of Classification Indices: Over (Grade 10) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.711 | 0.618 | | 0.512 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.858 | | 0.777 | |
| | 2 | 0.805 | | 0.713 | |
| | 3 | 0.747 | | 0.645 | |
| | 4 | 0.719 | | 0.601 | |
| | 5 | 0.601 | | 0.544 | |
| | 6 | - | | 0.316 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.984 | 0.008 | 0.008 | 0.976 |
| | 2/3 | 0.956 | 0.020 | 0.025 | 0.937 |
| | 3/4 | 0.925 | 0.034 | 0.042 | 0.894 |
| | 4/5 | 0.904 | 0.035 | 0.061 | 0.862 |
| | 5/6 | 0.942 | 0.058 | 0.000 | 0.936 |

**Table 8.5.8E-3**

Accuracy and Consistency of Classification Indices: Over (Grade 11) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.703 | 0.611 | | 0.500 | |
| Conditional on Level | Level | Accuracy | | Consistency | |
| | 1 | 0.849 | | 0.768 | |
| | 2 | 0.803 | | 0.708 | |
| | 3 | 0.768 | | 0.669 | |
| | 4 | 0.725 | | 0.603 | |
| | 5 | 0.548 | | 0.492 | |
| | 6 | - | | 0.307 | |
| Indices at Cut Points | | Accuracy | | | |
| | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
| | 1/2 | 0.985 | 0.008 | 0.007 | 0.978 |
| | 2/3 | 0.956 | 0.020 | 0.024 | 0.938 |
| | 3/4 | 0.923 | 0.032 | 0.045 | 0.893 |
| | 4/5 | 0.896 | 0.035 | 0.069 | 0.851 |
| | 5/6 | 0.941 | 0.059 | 0.000 | 0.938 |

**Table 8.5.8E-4**
Accuracy and Consistency of Classification Indices: Over (Grade 12) S302

| Overall Indices | Accuracy | Consistency | | Kappa (k) | |
|---|---|---|---|---|---|
| | 0.677 | 0.594 | | 0.463 | |
| **Conditional on Level** | **Level** | **Accuracy** | | **Consistency** | |
| | 1 | 0.840 | | 0.771 | |
| | 2 | 0.790 | | 0.686 | |
| | 3 | 0.809 | | 0.714 | |
| | 4 | 0.649 | | 0.548 | |
| | 5 | 0.395 | | 0.376 | |
| | 6 | - | | 0.548 | |
| **Indices at Cut Points** | | **Accuracy** | | | |
| | **Cut Point** | **Accuracy** | **False Positives** | **False Negatives** | **Consistency** |
| | 1/2 | 0.984 | 0.010 | 0.006 | 0.977 |
| | 2/3 | 0.954 | 0.022 | 0.024 | 0.936 |
| | 3/4 | 0.915 | 0.025 | 0.060 | 0.884 |
| | 4/5 | 0.868 | 0.075 | 0.057 | 0.813 |
| | 5/6 | 0.956 | 0.044 | 0.000 | 0.956 |

# References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.

Bauman, J., Boals, T., Cranley, E., Gottlieb, M., and Kenyon, D.M. (2007). The Newly Developed English Language Tests (World-Class Instructional Design and Assessment – WIDA). In Abedi, Jamal (Ed.), *English Language Proficiency Assessment in the Nation: Current Status and Future Practice.* Davis: University of California.

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

Cook, H. G. (2007). Alignment Study Report: The WIDA Consortium's English Language Proficiency Standards for English Language Learners in Kindergarten through Grade 12 to ACCESS for ELLs® Assessment. Madison, WI: WIDA Consortium.

Cook, H. G., Boals, T., Wilmes, C., and Santos, M. (2007). Issues in the Development of Annual Measurable Achievement Objectives (AMAOs) for WIDA Consortium States. Madison, WI: WIDA Consortium.

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

Fairbairn, S., & Fox, J. (2009). Inclusive achievement testing for linguistically and culturally diverse test takers: Essential considerations for test developers and decision makers**.** *Educational Measurement: Issues and Practice*, *28*(1), 10–24.

Fox, J. (2011). Test review: ACCESS for ELLs®. *Language Testing 28* (3): 425-431.

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.

Gottlieb, M. (2012). *ACCESS for ELLs® Interpretive Guide for Score Reports* (WIDA Consortium)*.* Madison, WI: The Board of Regents of the University of Wisconsin System.

Gottlieb, M., and Boals, T. (2005). Considerations in Reconfiguring Cohorts and Resetting Annual Measurable Achievement Objectives (AMAOs) based on ACCESS for ELLs® Data (WIDA Consortium Technical Report No. 3).

Gottlieb, M. and Kenyon, D.M. (2006). The Bridge Study between Tests of English Language Proficiency and ACCESS for ELLs® (WIDA Consortium Technical Report No. 2).

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

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.

Keng, L., Miller, G.E., O'Malley, K. & Turhan, A. (2008). A Generalization of Stratified Alpha that Allows for Correlated Measurement Errors between Subtests. Paper presented at the annual meeting of the American Educational Research Association, New York City, New York.

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

Kenyon, D.M., MacGregor, D., Li, D., and Cook, H. G. (2011). Issues in vertical scaling of a K-12 English language proficiency test. *Language Testing 28* (3): 383-400.

Kenyon, D.M., MacGregor, D., Louguit, M. Cho, B., and Ryu, J.R. (Willow), (2007). *Annual Technical Report for ACCESS for ELLs*® *English Language Proficiency Test, Series 101, 2005-2006 Administration* (WIDA Consortium Annual Technical Report No. 2).

Kenyon, D.M., MacGregor, D., Ryu, J.R. (Willow), Cho, B., and Louguit, M. (2006). *Annual Technical Report for ACCESS for ELLs*® *English Language Proficiency Test, Series 100, 2004-2005 Administration* (WIDA Consortium Annual Technical Report No. 1)

Kenyon, D.M., Ryu, J.R. (Willow), and MacGregor, D. (2013). *Setting Grade Level Cut Scores for ACCESS for ELLs*®. (WIDA Consortium Technical Report No. 4).

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. (n.d). Displacement measures. Retrieved from http://www.winsteps.com/winman/displacement.htm.

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

Linacre, J.M. (2006). *Winsteps Rasch analysis* ver. 3.60.1, date 02-02-2006 [Computer Program]. 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.*

MacGregor, D., Kenyon, D.M., Gibson, S., and Evans, E. (2009). *Development and Field Test of Kindergarten ACCESS for ELLs*®. (WIDA Consortium).

MacGregor, D., Louguit, M., Yanosky, T., Fidelman, C. G., Pan, M., Huang, X., and Kenyon, D.M. (2010). *Annual Technical Report for ACCESS for ELLs*® *English Language Proficiency Test, Series 200, 2008-2009 Administration* (WIDA Consortium Annual Technical Report No. 5).

MacGregor, D., Louguit, M., Huang, X., and Kenyon, D.M. (2009). Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 103, 2007-2008 Administration (WIDA Consortium Annual Technical Report No. 4).

MacGregor, D., Louguit, M., Ryu, J.R. (Willow), Kenyon, D.M., and Li, D. (2008). *Annual Technical Report for ACCESS for ELLs*® *English Language Proficiency Test, Series 102, 2006-2007 Administration* (WIDA Consortium Annual Technical Report No. 3).

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

National Research Council. (2011). Allocating federal funds for state programs for English language learners. Panel to Review Alternative Data Sources for the Limited-English Proficiency Allocation Formula under Title III, Part A, Elementary and Secondary Education Act, Committee on National Statistics and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Parker, C. E., Louie, J., and O'Dwyer, L. (2009). New measures of English language proficiency and their relationship to performance on large-scale content assessments (Issues & Answers Report, REL 2009–No. 066). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs, January 29, 2009.

Römhild, A., Kenyon, D. M., and MacGregor, D. (2011). Exploring domain-general and domain-specific linguistic knowledge in the assessment of academic English language proficiency. *Language Assessment Quarterly*, 8:213-228.

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.

Wang, H., Choi, I., Schmidgall, J., & Bachman, L. F. (2012). Review of Pearson Test of English Academic: Building an assessment use argument. Language Testing, 29(4), 603-619.

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

WIDA Consortium. (2012a). 2012 Amplification of the English Language Development Standards Kindergarten–Grade 12. Madison, Wisconsin: Board of Regents of the University of Wisconsin System.

WIDA Consortium (2012b). *WIDA ACCESS for ELLs™ Test Administration Manual.* Retrieved from http://www.wida.us/assessment/ACCESS/#about.

WIDA Consortium. (2013). *Interpretive Guide for Score Reports Spring 2013* (WIDA Consortium)*.* Madison, WI: The 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.

Wolf, M., Kao, J., Griffin, N., Herman, J., Bachman, P., Chang, S., and Farnsworth, T. (2008). *Issues In Assessing English Language Learners: English Language Proficiency Measures And Accommodation Uses. Practice Review* (Part 2 Of 3). Los Angeles, CA: CRESST Report 732, National Center for Research on Evaluation, Standards, and Student Testing.

Yanosky, T., Amos, M., Louguit, M., Olson, Cameron, C., Louguit, M., MacGregor, D., Yen, S., and Kenyon, D.M. (2013). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 203, 2011-2012 Administration* (WIDA Consortium Annual Technical Report No. 8).

Yanosky, T., Chong, A., Louguit, M., Olson, E., Choi, Y., MacGregor, D., Yen, S., Cameron, C., and Kenyon, D.M. (2012). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 202, 2010-2011 Administration* (WIDA Consortium Annual Technical Report No. 7).

Yanosky, T., Yen, S., Louguit, M., MacGregor, D., Zhang, Y., and Kenyon, D.M. (2011). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 201, 2009-2010 Administration* (WIDA Consortium Annual Technical Report No. 6).

Young, M.J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment* (CSE Tech. Rep. 475). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education and Information Studies.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. Holland & H. Wainer (Eds.), Differential Item Functioning. Hillsdale, NJ: Lawrence Erlbaum.

Zwick, R., Donoghue, J.R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30, 233-251.*

# Acknowledgements

We would like to extend our appreciation to the many CAL staff members who have supported this work, including:

Melissa Amos, M.S.
Catherine Cameron, M.A.
Deepak Ebenezer, M.S.
Simon Fong, B.A.
Dale Ingram, B.A.
Mohammed Louguit, Ph.D.
James Marcus, Ph.D.
Dorry M. Kenyon, Ph.D.
David MacGregor, Ph.D.
April Maddy, Ed.M.
Jennifer Renn, Ph.D.
Shauna Sweet, M.S.
Shu Jing Yen, Ph.D.
Xin Yu, M.A.