World-Class Instructional Design and Assessment

# WIDA <br> CONSORTIUM 

# Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ 2.0 Online English Language Proficiency Test, Series 401, 2016-2017 Administration 

Annual Technical Report No. 13A

Prepared by:
Center for Applied Linguistics
Language Assessment Division
Psychometrics and Quantitative Research Team

April 20, 2018
© 2018 Board of Regents of the University of Wisconsin System on behalf of the WIDA Consortium.

## The WIDA ACCESS for ELLs Technical Advisory Committee

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

- Jamal Abedi, Ph.D., Professor, Graduate School of Education, 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., Professor Emerita, Educational Psychology, University of Illinois at Chicago


## Executive Summary

This is the 13th annual technical report on the ACCESS for ELLs® English Language Proficiency Test, and the second report on the ACCESS for ELLs 2.0 assessment. ACCESS for ELLs 2.0 measures the same constructs as ACCESS for ELLs, but the assessment is now offered in an online, multi-stage adaptive format.

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

ACCESS for ELLs is intended to assess reliably and validly the English language development (ELD) of English language learners (ELLs) in Grades K-12 according to the WIDA 2012 Amplification of the English Language Development Standards Kindergarten-Grade 12 (WIDA Consortium, 2012). Results on ACCESS for ELLs are used by WIDA Consortium states for monitoring the progress of students, for making decisions about exiting students from language support services, and for accountability.

ACCESS for ELLs 2.0 Series 401 was administered in school year 2016-17 in 35 states, the District of Columbia, the Commonwealth of the Northern Marianas and the Virgin Islands of the United States, for a total of 38 state entities (henceforth "states"). ACCESS for ELLs 2.0 Series 401 was offered in two administrative formats, an online format (grades 1-12) and a paper format (kindergarten-grade 12). Table 0.1 summarizes the numbers of students, by state, who participated in the grades $1-12$ assessment online, in the grades $1-12$ assessment on paper, the total number of students who participated in the grades $1-12$ assessment, the total number who participated in the Kindergarten assessment (only offered in the paper format), and the total participants in ACCESS K-12. The current report (WIDA ACCESS Technical Report 13A) provides technical information pertaining to ACCESS for ELLs 2.0 Series 401 Online. A second report (WIDA ACCESS Technical Report 13B) provides technical information for the ACCESS for ELLs 2.0 Series 401 Paper assessment, including the Kindergarten assessment.

Table 0.1
Participation in ACCESS for ELLs Online and Paper, Series 401

| State | Participants in ACCESS for ELLs Grades 1-12 |  |  | Participants in Kindergarten | Total Participants in ACCESS for ELLs Grades K-12 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Participants in ACCESS for ELLs Online | Participants in ACCESS for ELLs Paper | Total Participants in ACCESS for ELLs |  |  |
| AK | 7,737 | 4,795 | 12,532 | 1,386 | 13,918 |
| AL | 11,649 | 5,741 | 17,390 | 3,487 | 20,877 |
| CO | 61,768 | 29,167 | 90,935 | 10,836 | 101,771 |
| DC | 5,610 | 75 | 5,685 | 1,089 | 6,774 |
| DE | 9,495 | 13 | 9,508 | 1,637 | 11,145 |
| FL | 0 | 243,736 | 243,736 | 35,774 | 279,510 |
| GA | 73,992 | 12,164 | 86,156 | 17,196 | 103,352 |
| HI | 7,807 | 3,014 | 10,821 | 1,876 | 12,697 |
| ID | 13,367 | 39 | 13,406 | 2,230 | 15,636 |
| IL | 130,134 | 31,024 | 161,158 | 26,670 | 187,828 |
| IN | 41,970 | 523 | 42,493 | 7,405 | 49,898 |
| KY | 19,390 | 347 | 19,737 | 3,377 | 23,114 |
| MA | 46,274 | 28,823 | 75,097 | 10,330 | 85,427 |
| MD | 60,015 | 122 | 60,137 | 10,675 | 70,812 |
| ME | 4,711 | 247 | 4,958 | 485 | 5,443 |
| MI | 84,524 | 3,774 | 88,298 | 10,371 | 98,669 |
| MN | 59,906 | 597 | 60,503 | 8,316 | 68,819 |
| MO | 26,122 | 69 | 26,191 | 4,899 | 31,090 |
| MP | 1,302 | 0 | 1,302 | 78 | 1,380 |
| MT | 2,581 | 0 | 2,581 | 137 | 2,718 |
| NC | 79,468 | 1,543 | 81,011 | 11,957 | 92,968 |
| ND | 2,725 | 41 | 2,766 | 384 | 3,150 |
| NH | 3,594 | 294 | 3,888 | 441 | 4,329 |
| NJ | 60,066 | 955 | 61,021 | 12,035 | 73,056 |
| NM | 38,249 | 2,796 | 41,045 | 4,717 | 45,762 |
| NV | 64,380 | 39 | 64,419 | 7,956 | 72,375 |
| OK | 24,430 | 13,433 | 37,863 | 6,902 | 44,765 |
| PA | 41,074 | 10,708 | 51,782 | 5,017 | 56,799 |
| RI | 8,744 | 1,117 | 9,861 | 1,092 | 10,953 |
| SC | 39,374 | 1,295 | 40,669 | 3,478 | 44,147 |
| SD | 3,653 | 192 | 3,845 | 742 | 4,587 |
| TN | 38,872 | 15 | 38,887 | 5,711 | 44,598 |
| UT | 34,945 | 6 | 34,951 | 4,975 | 39,926 |
| VA | 76,847 | 11,395 | 88,242 | 14,215 | 102,457 |
| VI | 1,023 | 0 | 1,023 | 96 | 1,119 |
| VT | 1,295 | 13 | 1,308 | 178 | 1,486 |
| WI | 42,080 | 180 | 42,260 | 5,531 | 47,791 |
| WY | 2,186 | 60 | 2,246 | 386 | 2,632 |
| Total | 1,231,359 | 408,352 | 1,639,711 | 244,067 | 1,883,778 |

## Summary Highlights

This report presents a wealth of data documenting the technical properties of ACCESS for ELLs 2.0 Series 401 Online, which cannot be fully summarized 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 and a separate table highlighting conditional standard errors around the cut scores. 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.

## Argument-based validation framework for ACCESS for ELLs

The purpose of this report is to provide evidence for the validity of the online version of ACCESS for ELLs 2.0 (hereafter ACCESS 2.0 Online), when used for its intended purposes. This report is structured using a validation approach developed at the Center for Applied Linguistics (CAL), and based on Bachman and Palmer's (2010) Assessment Use Argument, integrated with the Evidence Centered Design principles outlined in Mislevy, Almond, and Lukas (2004). CAL's validation framework, shown in Figure 2 of Part I of this report, consists of 7 steps, leading the line of argumentation from Plan (Step 7) through Consequences (Step 1).

Part I of this report, Foundations, covers Steps 7-5 of CAL's Validation Framework (Plan, Design, and Assessment Performance).

Part II of the report, Assessment Records covers Step 4 in the Validation Framework. Part II has three subsections:

II:1 Assessment Records for ACCESS 2.0 Online
II:2 Background and Descriptions for the Presentation of Results
II:3 Results by Grade Cluster

The first subsection provides the Assessment Use Argument (AUA), a set of claims which allow stakeholders to better interpret and use ACCESS for ELLs. These claims are each supported by evidence, much of which is drawn from the tables and figures presented in this report. The second subsection provides technical detail on the analyses conducted to provide evidence for the AUA claims, and the third subsection contains the tables and figures with the results of that analysis.

## Demographic data

The Series 401 Online data set included the results of $1,231,359$ students. The largest grade was Grade 3 with 182,698 students, while the smallest was Grade 12 with 33,370 students. Of the participating WIDA states, the largest was Illinois with 130,134 students, while the smallest was the United States Virgin Islands with 1,023 students.

## Reliability and accuracy data

For most test users, the Overall Composite proficiency score, based on performances in Listening, Reading, Writing and Speaking, is the major score used for making decisions about gains in student proficiency and exiting from language support services.

Results indicate that the reliability (stratified Crohnbach's alpha) of the Overall Composite score for Series 401 Online, presented in Table C of Section 3.4 (see also section 2.3.3. of Part II) is very high across all grade-level clusters. For Grade 1 it was .937 ; for Grades 2-3, .947; for Grades 4-5, .950; for Grades 6-8, .959; and for Grades 9-12, . 950 .

Likewise, as Table 0.2 shows, the accuracy of classification for student placement using the Overall Composite score around the proficiency level cut scores is very high across 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 0.2
Accuracy of Classification of Overall Score at Cut Points (Proficiency Level Score)

| Grade | $\mathbf{1 / 2}$ Cut <br> $\mathbf{( 2 . 0})$ | $\mathbf{2 / 3}$ Cut <br> $\mathbf{( 3 . 0 )}$ | $\mathbf{3 / 4} \mathbf{C u t}$ <br> $\mathbf{( 4 . 0 )}$ | $\mathbf{4 / 5}$ Cut <br> $\mathbf{( 5 . 0 )}$ | $\mathbf{5 / 6}$ Cut <br> $(\mathbf{6 . 0})$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.970 | 0.923 | 0.952 | 0.988 | 0.998 |
| 2 | 0.977 | 0.936 | 0.939 | 0.979 | 0.999 |
| 3 | 0.980 | 0.949 | 0.920 | 0.959 | 0.999 |
| 4 | 0.982 | 0.961 | 0.917 | 0.947 | 0.994 |
| 5 | 0.978 | 0.956 | 0.921 | 0.930 | 0.994 |
| 6 | 0.970 | 0.942 | 0.944 | 0.982 | 0.999 |
| 7 | 0.966 | 0.941 | 0.941 | 0.974 | 0.998 |
| 8 | 0.965 | 0.944 | 0.938 | 0.968 | 0.998 |
| 9 | 0.958 | 0.938 | 0.942 | 0.981 | 0.998 |
| 10 | 0.959 | 0.935 | 0.949 | 0.987 | 0.999 |
| 11 | 0.956 | 0.932 | 0.955 | 0.988 | 0.999 |
| 12 | 0.956 | 0.929 | 0.961 | 0.991 | 1.000 |

## Series 401 Online: Special Considerations

## Data Exclusion: State of Michigan

Data for the production of the Annual Technical Report were received by CAL in late September of 2017. Data for the state of Michigan were not available in the system at the time of the initial data pull. Data from the state of Michigan were received in mid-November. In order to allow for the timely production of the report, analyses which pertain to the technical properties of test forms (the analyses included in section 3.3 and 3.4) were conducted using the original September data. These analyses do not include data from the state of Michigan. Students from the state of Michigan constitute 84,524 of 1,231,539 total students in the ACCESS Online population, a proportion of $6.86 \%$. Students from the state of Michigan are included in summary tables which pertain to the counts of students participating in the assessment (the tables in section 3.2).

## Annotated Bibliography

## Technical Reports

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

Center for Applied Linguistics (2015). ACCESS for ELLs Series 302 Media-Based Listening Field Test Technical Brief. (WIDA Consortium).

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

Center for Applied Linguistics (2016). ACCESS for ELLs ${ }^{\circledR}$ Series 400 Listening and Reading Scale Maintenance: Technical Brief. (WIDA Consortium).

This brief summarizes the results of two scaling studies which were conducted to ensure that scores on ACCESS 2.0 in the domains of Listening and Reading remained on the original ACCESS score scale as the test transitioned from paper-based format to online delivery.

Center for Applied Linguistics (2017). ACCESS for ELLs® 2.0 Speaking and Writing Score Scale Reconstruction: Technical Brief. (WIDA Consortium).

The purpose of this report is to summarize the results of two scaling studies that were conducted to the reconstruct ACCESS Speaking and Writing score scales as the transition was made to ACCESS 2.0 in the domains of Speaking and Writing.

Cook, H. G. and MacGregor, D. (2017). The ACCESS for ELLs 2.0 2016 Standard Setting Study (Technical Report). Madison, WI: Board of Regents of the University of Wisconsin System.

This report summarizes the processes and procedures employed to set the new ACCESS 2.0 cut scores in the summer of 2016.

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., \& 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 ELLs, listed by year of publication. These reports provide extensive analysis of the results from the operational administrations of ACCESS for ELLs. They provide detailed information on student results broken down by gradelevel cluster, grade, and tier. They also provide detailed information on test and item characteristics.

Kenyon, D. M., MacGregor, D., Ryu, J. R., Cho, B., \& Louguit, M. (2006). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 100, 20042005 Administration (WIDA Consortium Annual Technical Report No. 1).

Kenyon, D. M., MacGregor, D., Louguit, M., Cho, B., \& Ryu, J. R. (2007). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 101, 20052006 Administration (WIDA Consortium Annual Technical Report No. 2).

MacGregor, D., Louguit, M., Ryu, J. R., Kenyon, D. M., \& Li, D. (2008). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 102, 20062007 Administration (WIDA Consortium Annual Technical Report No. 3).

MacGregor, D., Louguit, M., Huang, X., \& Kenyon, D. M. (2009). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ 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., \& Kenyon, D. M. (2010). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ 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., \& Kenyon, D. M. (2011). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ 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., . . .Kenyon, D. M. (2012). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 202, 2010-2011 Administration (WIDA Consortium Annual Technical Report No. 7).

Yanosky, T., Amos, M., Cameron, C., Louguit, M., MacGregor, D., Yen, S., \& Kenyon, D. M. (2013). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ 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 ${ }^{\circledR}$ English Language Proficiency Test, Series 301, 2012-2013 Administration (WIDA Consortium Annual Technical Report No. 9).

Center for Applied Linguistics (2015). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 302, 2013-2014 Administration (WIDA Consortium Annual Technical Report No. 10).

Center for Applied Linguistics (2016). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR}$ English Language Proficiency Test, Series 303, 2014-2015 Administration (WIDA Consortium Annual Technical Report No. 11).

Center for Applied Linguistics (2017a). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR} 2.0$ Online English Language Proficiency Test, Series 400, 2015-2016 Administration (WIDA Consortium Annual Technical Report No. 12A).

Center for Applied Linguistics (2017b). Annual Technical Report for ACCESS for ELLs ${ }^{\circledR} 2.0$ Paper English Language Proficiency Test, Series 400, 2015-2016 Administration (WIDA Consortium Annual Technical Report No. 12B).

## 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., \& Kenyon, D. M. (2007). The Newly Developed English Language Tests (World-Class Instructional Design and Assessment - WIDA). In J. Abedi (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.K. \& 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 ${ }^{\mathrm{TM}}$ 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.

Chapelle, C. A., Enright, M. K., \& Jamieson, J. (2010). Does an argument-based approach to validity make a difference? Educational Measurement: Issues and Practice, 29(1), 313.

Drawing on experience between 2000 and 2007 in developing a validity argument for the high-stakes Test of English as a Foreign Language ${ }^{\mathrm{TM}}$, 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.

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 ${ }^{\circledR}$ 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., \& 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. \& Fairbairn, S. (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 (4 ${ }^{\text {th }}$ 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., \& Cook, H. G. (2011). Issues in vertical scaling of a K12 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. 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., \& 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., \& MacGregor, D. (2011). Exploring domain-general and domainspecific 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. (2012). 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. (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., \& Farnsworth, T. (2008). Issues in assessing English language learners: English language proficiency measures and accommodation uses-Practice review (Part 2 of 3) (CRESST Report 732). Los Angeles, CA: 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 20052006, 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 (Eds.), Differential item functioning (pp. 337-347). Hillsdale, NJ: Lawrence Erlbaum Associates.

This book chapter describes procedures for conducting DIF analysis.

## Table of Contents

## Part I

1 The Validation Framework for ACCESS 2.0 Online ..... 1
1.1 Development of the Validation Framework for ACCESS 2.0 Online ..... 1
1.2 Description of the Framework ..... 2
2 The Plan for ACCESS 2.0 Online ..... 5
2.1 Purpose Statement: What are the intended decisions and consequences of using ACCESS? ..... 5
2.2 Domain Analysis: What is WIDA's conceptualization of the development of academic English language proficiency? ..... 6
2.2.1 The WIDA Standards ..... 6
2.2.2 Grade-Level Clusters ..... 7
2.2.3 Language Domains ..... 7
2.2.4 Language Proficiency Levels ..... 7
2.3 Domain Modeling: How are the components of the ACCESS assessment program interrelated? ..... 8
3 The Design of ACCESS 2.0 Online ..... 9
3.1 The Student Model: What knowledge, skills, and abilities does a student possess? ..... 9
3.2 The Task Model: What do assessment tasks for ACCESS 2.0 Online look like? ..... 9
3.2.1 Listening Items ..... 9
3.2.2 Reading Items ..... 10
3.2.3 Writing Tasks ..... 10
3.2.4 Speaking Tasks ..... 10
3.3 The Evidence Model ..... 11
3.3.1 How are student performances on ACCESS 2.0 Online scored? ..... 12
3.3.2 How are measures of student performances on ACCESS 2.0 Online calculated? ..... 16
3.4 The Assembly Model: How are the assessment components for ACCESS put together?19
3.4.1 Listening. ..... 20
3.4.2 Reading ..... 21
3.4.3 Writing ..... 21
3.4.4 Speaking ..... 22
4 Assessment Performance: The Implementation of ACCESS 2.0 Online ..... 24
4.1 How is ACCESS implemented? ..... 24
4.1.1 Listening and Reading ..... 24
4.1.2 Writing ..... 25
4.1.3 Speaking ..... 26
4.2 What is the assessment delivery experience for students taking ACCESS 2.0 Online? 27 ..... 27
4.2.1 Listening and Reading ..... 27
4.2.2 Writing ..... 27
4.2.3 Speaking ..... 27
4.3 Assessment performance-interaction between test and student ..... 28

## Part II

1 Assessment Records for ACCESS 2.0 Online ..... 29
1.1 Claims for the Assessment Records for ACCESS ..... 31
1.2 Evidence for Assessment Records Claims of ACCESS 2.0 Online ..... 32
1.3 Summary of Assessment Records Claims, Actions, and Evidence. ..... 42
2 Background and Descriptions for the Presentation of Results ..... 45
2.1 Student Participation and Performance ..... 45
2.1.1 Participation ..... 46
2.1.2 Scale Score Results ..... 47
2.1.3 Proficiency Level Results ..... 48
2.2 Analyses of Domain Scores ..... 48
2.2.1 Complete Item or Task Analysis and Summary ..... 48
2.2.2 DIF Analysis and Summary ..... 50
2.2.3 Raw score distribution for Speaking and Writing. ..... 53
2.2.4 Scale Score Distribution. ..... 53
2.2.5 Proficiency Level distribution ..... 54
2.2.6 Raw to Scale Score Conversion for Speaking and Writing ..... 54
2.2.7 Equating Summary ..... 54
2.2.8 Test Characteristic Curve ..... 57
2.2.9 Test Information Function ..... 57
2.2.10 Reliability of Domain Scores ..... 58
2.2.11 Conditional Standard Errors of Measurement at Cut Score ..... 61
2.2.12 Accuracy and consistency ..... 62
2.3 Analyses of Composite Scores ..... 64
2.3.1 Scale Score distribution for Composites ..... 64
2.3.2 Proficiency Level distribution for Composites ..... 64
2.3.3 Reliability of Composites ..... 65
2.3.4 Accuracy and Consistency of Composites ..... 65
3 Results By Grade Cluster ..... 67
3.1 Guide to Tables and Figures ..... 67
3.1.1 Guide to 3.2, Student Participation and Performance ..... 67
3.1.2 Guide to 3.3, Analysis of Domain Scores. ..... 68
3.1.3 Guide to 3.4, Analysis of Composite Scores ..... 70
3.2 Student Participation and Performance ..... 70
3.2.1 Participation ..... 70
3.2.2 Scale Score Results ..... 78
3.2.3 Proficiency Level Results ..... 94
3.3. Analyses of Domain Scores: Results ..... 111
3.3.1 Grade: 1 ..... 111
3.3.1.1 Listening 1 ..... 111
3.3.1.2 Reading 1 ..... 119
3.3.1.3 Writing 1 ..... 129
3.3.1.4 Speaking 1 ..... 144
3.3.2 Grades: 2-3 ..... 160
3.3.2.1 Listening 2-3 ..... 160
3.3.2.2 Reading 2-3 ..... 169
3.3.2.3 Writing 2-3 ..... 180
3.3.2.4 Speaking 2-3 ..... 196
3.3.3 Grades: 4-5 ..... 213
3.3.3.1 Listening 4-5 ..... 213
3.3.3.2 Reading 4-5 ..... 222
3.3.3.3 Writing 4-5 ..... 233
3.3.3.4 Speaking 4-5 ..... 249
3.3.4 Grades: 6-8 ..... 266
3.3.4.1 Listening 6-8 ..... 266
3.3.4.2 Reading 6-8 ..... 276
3.3.4.3 Writing 6-8 ..... 288
3.3.4.4 Speaking 6-8 ..... 305
3.3.5 Grades: 9-12 ..... 323
3.3.5.1 Listening 9-12 ..... 323
3.3.5.2 Reading 9-12 ..... 333
3.3.5.3 Writing 9-12 ..... 345
3.3.5.4 Speaking 9-12 ..... 362
3.4.Analyses of Composite Scores: Results ..... 380
3.4.1 Grade: 1 ..... 380
3.4.1.1 Oral Language Composite 1 ..... 380
3.4.1.2 Literacy Language Composite 1 ..... 382
3.4.1.3 Comprehension Language Composite 1 ..... 384
3.4.1.4 Overall Language Composite 1 ..... 386
3.4.2 Grades: 2-3 ..... 388
3.4.2.1 Oral Language Composite 2-3 ..... 388
3.4.2.2 Literacy Language Composite 2-3 ..... 391
3.4.2.3 Comprehension Language Composite 2-3 ..... 393
3.4.2.4 Overall Language Composite 2-3 ..... 395
3.4.3 Grades: 4-5 ..... 398
3.4.3.1 Oral Language Composite 4-5 ..... 398
3.4.3.2 Literacy Language Composite 4-5 ..... 401
3.4.3.3 Comprehension Language Composite 4-5 ..... 403
3.4.3.4 Overall Language Composite 4-5 ..... 405
3.4.4 Grades: 6-8 ..... 408
3.4.4.1 Oral Language Composite 6-8 ..... 408
3.4.4.2 Literacy Language Composite 6-8 ..... 411
3.4.4.3 Comprehension Language Composite 6-8 ..... 414
3.4.4.4 Overall Language Composite 6-8 ..... 417
3.4.5 Grades: 9-12 ..... 420
3.4.5.1 Oral Language Composite 9-12 ..... 420
3.4.5.2 Literacy Language Composite 9-12 ..... 424
3.4.5.3 Comprehension Language Composite 9-12 ..... 428
3.4.5.4 Overall Language Composite 9-12 ..... 432
References ..... 436
Acknowledgements ..... 439

## Part I: Foundations

ACCESS for ELLs 2.0 Online is a secure, large-scale English language proficiency assessment administered to students in grades $1-12$ who have been identified as English language learners (ELLs). It is administered annually in WIDA Consortium member states to monitor students' progress in acquiring academic English. ACCESS 2.0 Online is a standards-based English language proficiency test designed to measure both social and academic language proficiency of ELLs in English in a school context. It assesses social and instructional language, as well as the academic language associated with language arts, mathematics, science, and social studies, across the four language domains (Listening, Reading, Writing, and Speaking).

## 1 The Validation Framework for ACCESS 2.0 Online

### 1.1 Development of the Validation Framework for ACCESS 2.0 Online

As with any assessment, an important consideration during the development of ACCESS 2.0 was determining how to establish its validity. 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 consider the evidence that supports the interpretations and decisions made about students on the basis of their performance on a test, and the appropriateness and adequacy of such interpretations. A fully developed validation framework, including an Assessment Use Argument (AUA; Bachman \& Palmer, 2010), consists of several steps that connect test design and administration to intended and actual score interpretation and consequences. The present section contextualizes the information presented in this Annual Technical Report within an argumentbased approach to addressing validity (Bachman \& Palmer, 2010; Chapelle, Enright, \& Jamieson, 2008; Kane, 2002, 2013; Mislevy, Almond, \& Lukas, 2004) for ACCESS 2.0 Online.

An argument-based approach to the ACCESS 2.0 Online validation framework organizes the information in the present report to support claims about Assessment Records (i.e., test scores and proficiency level descriptions collected via ACCESS 2.0 Online). Specifically, tables and figures from this report explicitly address questions related to assessment data. Chapelle, Enright, and Jamieson (2010) support using such a structure for presenting 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).

The validity argument for ACCESS 2.0 Online shows the path from test design to student 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. Evidence includes results from analyses of test data, outside documentation, and other resources. In the validation argument, this process of identifying evidence to support claims encompasses the entire testing process, from the commencement of test design to the consequences of test use (Bachman \& Palmer, 2010; Llosa, 2008); Figure 1 shows the process by which evidence supports validation actions, which are used to establish larger claims about ACCESS 2.0 Online. The figure shows the generic structure of the line of argumentation for validity.


Figure 1. General Argument Structure for Assessment Validation (simplified from Toulmin, 2003).

### 1.2 Description of the Framework

The generic validation framework applied to the entire ACCESS 2.0 Online 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, combines models for both test development (i.e., Evidence-Centered Design (ECD) [Mislevy, Almond, \& Lukas, 2004]) and assessment validation (i.e., the AUA from Bachman and Palmer [2010]) to cover the
assessment development and implementation process from initial conceptualization to score interpretations and consequences of using the assessment. This framework constantly looks both forward and backward, and each subsequent step depends upon the strength of the step below it; for this reason, the steps going from bottom to top are numbered from seven to one. For example, during the initial Plan step (Step 7), test developers state the anticipated decisions and consequences of implementing the assessment program, which then drive the development and implementation of the assessment (Steps 6 through 4). Assessment results are then used to formulate Interpretations (Step 3) and to make Decisions (Step 2). Consequences (Step 1) represents the culmination of all previous steps. This structure highlights the fact that any weakness in a lower step affects the steps above it.

In CAL's Validation Framework, Plan involves an examination of possible decisions that state educational agencies might make and consequences that might result from use of the assessment. This leads to the consideration of several models during Design, 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 students' 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).
Part I of this report presents evidence regarding the Planning, Designing, and Operationalization of the test, while information related to Assessment Records is found in Part II.


Figure 2. CAL’s Validation Framework (based on Bachman \& Palmer, 2010; Mislevy, Almond, \& Lukas, 2004).

## 2 The Plan for ACCESS 2.0 Online

This section focuses on Plan (Step 7) of CAL's Validation Framework. This section details the decisions that the test is intended to inform, along with the consequences of those decisions. It then describes the domain analysis and modeling that undergirds WIDA's conceptualization of academic English language proficiency.

### 2.1 Purpose Statement: What are the intended decisions and consequences of using ACCESS?

The overarching purpose of ACCESS for ELLs 2.0 is to assess the developing English language proficiency of English language learners (ELLs) in Grades K-12 in the United States as defined by the multi-state WIDA Consortium, first in the English Language Proficiency Standards (Gottlieb, 2004; WIDA Consortium, 2007), then in the amplified 2012 English Language Development (ELD) Standards (WIDA Consortium, 2012). The WIDA ELD Standards, which correspond to the academic language identified in state academic content standards, describe six levels of developing English language proficiency and form the core of the WIDA Consortium's approach to instructing and testing ELLs. ACCESS 2.0 may thus be described as a standardsbased English language proficiency test designed to measure the social and academic language proficiency of ELLs in English. It assesses social and instructional English as well as the academic language associated with language arts, mathematics, science, and social studies within the school context across the four language domains (Listening, Reading, Writing, and Speaking).

Other major purposes of ACCESS 2.0 include:

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

ACCESS 2.0 is offered in two formats: ACCESS 2.0 Online, described in this report, and ACCESS 2.0 Paper, described in a companion report.

### 2.2 Domain Analysis: What is WIDA's conceptualization of the development of academic English language proficiency?

The Domain Analysis aspect of the Plan step in CAL's Validation Framework defines what ACCESS 2.0 is assessing as a measure of English language proficiency. In ECD (Mislevy Almond, \& Lukas, 2004), Domain Analysis involves compiling and synthesizing all of the relevant information about what will be assessed, namely, academic language proficiency. WIDA's conceptualization of academic English language proficiency is encapsulated in the 2012 Amplification of the ELD Standards (WIDA, 2012), which built upon previous editions of the WIDA ELD Standards (Gottlieb, 2004; WIDA, 2007). The five WIDA ELD Standards form the basis of this conceptualization. In order to capture the language development of ELLs, the Standards include the following layers of organization: Grade-level clusters, Language Domains, and Language Proficiency Levels. Domain Analysis therefore also incorporates more granular information about the characteristics of a task and/or response for these various organizational levels.

### 2.2.1 The WIDA Standards

The five foundational WIDA ELD Standards, which inform the design, structure, and content of ACCESS 2.0 Online, are:

- Standard 1: ELLs communicate in English for Social and Instructional purposes within the school setting.
- Standard 2: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of Language Arts.
- Standard 3: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of Mathematics.
- Standard 4: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of Science.
- Standard 5: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of Social Studies.

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

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

Every selected response item and every performance-based task on ACCESS 2.0 Online targets at least one of these five Standards. In the cases of some test items and tasks, the Standards are combined as follows:

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


### 2.2.2 Grade-Level Clusters

The WIDA ELD Standards describe developing English language proficiency within six gradelevel clusters. These are $\mathrm{K}, 1,2-3,4-5,6-8$, and $9-12$. Test forms follow this grade-level clustering.

### 2.2.3 Language Domains

The WIDA ELD Standards describe developing English language proficiency for each of the four language domains: Listening, Speaking, Reading, and Writing. Thus, ACCESS 2.0 Online contains four sections, each assessing an individual language domain.

### 2.2.4 Language Proficiency Levels

The WIDA ELD Standards describe the continuum of language development via five language proficiency levels (PLs) that are fully delineated in the WIDA ELD Standards document (WIDA 2012), with scores indicating progression through each level. These levels are Entering, Emerging, Developing, Expanding, and Bridging. There is also a final stage known as Reaching which is used to describe students who have progressed across the entire WIDA English language proficiency continuum; as such, scores do not indicate progression through this level. The proficiency levels are shown graphically in Figure 3.


Figure 3. The Language Proficiency Levels of the WIDA ELD Standards.

### 2.3 Domain Modeling: How are the components of the ACCESS assessment program interrelated?

The Domain Model aspect of the Plan step in CAL's Validation Framework formulates the argument between the evidence collected about the test taker and the intended inferences to be made about them. In other words, in the Domain Model, we ask what evidence is necessary and sufficient to make the target inferences. In the case of ACCESS 2.0 Online, evidence is collected in order to make inferences about the test takers' language proficiency. This argument has been operationalized within ACCESS 2.0 Online in terms of the Model Performance Indicator (MPI). The MPIs convey what the test taker should be able to do with language. Each MPI is mapped to a PL, providing examples of how a test taker at each level would accomplish the task. This Domain Model serves as the basis for the Task Model in the Design step (Step 6) of CAL's Validation Framework, where the characteristics of individual items or tasks are defined. In ACCESS 2.0 Online, therefore, the overall enterprise of mapping evidence to inferences is mediated through the theoretical notion of the MPI in the Domain Model, while specific MPIs for actual test items are developed at a later stage.

## 3 The Design of ACCESS 2.0 Online

Step 6 in CAL's Validation Framework is the Design step, which has four components, derived from ECD (Mislevy, Almond, \& Lukas, 2004): The Student Model; the Evidence Model; the Task Model; and the Assembly Model. For the benefit of the reader, the Task Model is presented prior to the Evidence Model in this section, as our description of the evidence derived from scoring is dependent upon a clear understanding of the nature of the tasks.

### 3.1 The Student Model: What knowledge, skills, and abilities does a student possess?

The Student Model defines the knowledge, skills, and abilities that a student possesses and that are going to be assessed. The Student Model for ACCESS 2.0 is operationalized in terms of the WIDA ELD Standards; the Standards define what ELLs process (in the Reading and Listening domains) or produce (in the Writing and Speaking domains) for a given grade-level cluster and proficiency level.

### 3.2 The Task Model: What do assessment tasks for ACCESS 2.0 Online look like?

This section describes how items and tasks are designed to reflect the elements of the domain analysis described in Section 2.2 in order to collect the necessary evidence required for later decision-making. Data Recognition Corporation (DRC), the vendor responsible for the online implementation of the assessment, administers the tasks in the online environment and carries out the automated scoring of the Listening and Reading tasks and the hand scoring of the Writing and Speaking tasks. Items and tasks are discussed by language domain.

### 3.2.1 Listening Items

All Listening items include a pre-recorded stimulus passage and question stem. Listening items are selected-response items, with one key and two distractors as answer choices. Answer choices are primarily illustrations; for Grades $2-12$, items that test listening proficiency at PLs $3-5$ may consist of short written text response options that are written to be about two PLs lower than the targeted PL of the Listening item. All operational Listening items are traditional multiple-choice items, though some items embedded for field-testing purposes involved enhanced itempresentations (see Section 4.1.1.).

Each item on the Listening test is written to reflect the language of one of the five WIDA ELD Standards and to test a student's ability to process language at one of the five fully delineated proficiency levels. Folders group together three test items that are written around a common theme, with each item targeting a progressively higher proficiency level. Thematic folders are targeted as A, B, or C, with A folders encompassing PLs 1-3, B folders encompassing PLs 2-4, and C folders encompassing PLs 3-5.

### 3.2.2 Reading Items

Reading items are similar in format to Listening items. The stimulus for Reading items is written text and answer choices primarily are also written text, though for Grades 1-12 graphic response options may be used for items targeting PLs 1 and 2 . As with Listening items, Reading items are grouped into thematic folders of three test items each. Most items on the operational Reading assessment are traditional multiple choice, though some operational items and some items embedded for field-testing purposes involved enhanced item-presentations (see Section 4.1.1.).

### 3.2.3 Writing Tasks

Writing tasks are designed to elicit language from one or more of the WIDA ELD Standards. Tasks appearing on the Tier A test form (see Section 3.4.3) are designed to give students the opportunity to produce writing samples that fulfill linguistic expectations up to PL 3, while those appearing on the Tier $\mathrm{B} / \mathrm{C}$ form are designed to give students the opportunity to produce writing samples that fulfill linguistic expectations up to PL 6.

With the exception of students in Grades 1-3 and those taking the paper-based accommodation, writing prompts appear on the computer screen. In the spirit of providing maximal support and making every provision to ensure that students are given the opportunity to demonstrate the full extent of their English language proficiency, modeling is sometimes used to make task expectations as clear as possible to students. For example, the first of a series of questions may already be partially completed, or a sentence starter may be provided.

Students in Grades 4-5 provide either handwritten or keyboarded responses, with the default response mode determined in advance at the state or district level. For students in Grades 6-12, keyboarding is the default response mode, with a handwriting option offered as an accommodation. For students in Grades $1-3$, the test is not administered via computer. Rather, the familiar format from ACCESS 1.0 is utilized, where the test administrator reads from a script and the students respond in a printed test booklet.

### 3.2.4 Speaking Tasks

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

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

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

Figure 4 shows the generic screen layout of the Speaking test.

## Speaking



Figure 4. Visualization of the Speaking test screen layout.

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

### 3.3 The Evidence Model

In determining what evidence should be sought at the Design phase of ACCESS 2.0 Online, two questions were articulated: (a) How are student performances on ACCESS 2.0 Online scored? and (b) How are measures of student performance on ACCESS 2.0 Online calculated? This section describes the scoring procedures and the methodologies used to score student performances in each domain.

### 3.3.1 How are student performances on ACCESS 2.0 Online scored?

### 3.3.1.1 Multiple Choice Scoring: Listening and Reading

Listening and Reading items are scored dichotomously, as correct or incorrect. Scale scores for each domain are calculated based on the items that are administered to the test taker and the number of those items that the student answers correctly. For details on how scale scores for Listening and Reading are calculated, see Section 3.3.2.1 below.

### 3.3.1.2 Performance-based Tasks: Writing and Speaking

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

## Rater Training and Qualifying

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


## Calculating Score Agreement for Scoring Monitoring

- For writing, agreement is defined as two adjacent scores. (See section 3.3.1.3 for a description of the writing scoring scale.) For example, using the writing scoring scale,
scores of 2 and $2+$ would be considered agreement as would scores of 2 and 2 or scores of $2+$ and 3 . Scores of 2 and 3 on the writing scoring scale would be considered adjacent and scores of 2 and $3+$ would be considered non-adjacent.
- For speaking, agreement is defined as two scores being exactly the same. (See section 3.3.1.4 for a description of the speaking scoring scale.)


## Routing Responses to Ensure "Blind" Second Ratings

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


## Monitoring Scoring (Quality Control)

- Ongoing quality control checks and procedures help monitor and maintain the quality of the scoring sessions. DRC monitors rater reliability with a $20 \%$ read-behind protocol. Read-behind data are monitored daily.
- Responses can be retrieved on-demand (e.g., specific grade-level clusters, specific students) should the need arise during or subsequent to the scoring process.
- If needed, responses can be rescored based on task- or response-level information, such as task number, date, score value assigned, or rater ID.
- For Writing, DRC uses both recalibration and validation sets. For each of the first five days that raters score a task, they take one recalibration set of five responses per task. After the raters take the recalibration sets, the Scoring Director or Team Leader reviews them using descriptors from the Writing Scoring Scale and the anchor responses to confirm the rationale behind each response's score. Starting on the $6^{\text {th }}$ day of scoring, DRC uses validity sets to monitor rater performance. These are sets of items seeded into the operational sets that, on a daily basis, monitor how raters are doing when compared to the known ratings of the validity sets. The raters do not know which items are operational and which are from a validation set.
- For Speaking, DRC uses recalibration sets. At the start of the scoring window, raters take these sets every day to ensure that they are calibrated, and raters' performances on recalibration sets are used for monitoring purposes. Later in the scoring window, recalibration sets are used on a weekly basis to monitor scoring.


## Handling Unusual Responses

- Raters can forward responses to Team Leaders for assistance.
- Responses requiring special attention, including nonscorable responses, are routed to Scoring Directors for review and resolution.


### 3.3.1.3 The ACCESS 2.0 Writing Scoring Scale

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

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

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

To calculate a raw score for the Writing test, raters' scores for each Writing task are converted to whole numbers ranging from $0-9$, as shown in Table 1. On Tier A tests, for all grade-level clusters except for Grade 1, the scores from the three tasks are added to calculate a total raw score, which can range from 0-27. The exception to this rule is the Grade 1 Tier A test. On this form, there are four Writing tasks. The first two of these tasks use a modified version of the scoring scale and have score ranges of $0-1$ and $0-3$ respectively. The third and fourth task use the full scoring scale from $0-9$; additionally, the last task is weighted as 3 . Therefore, the possible final raw scores for Grade 1 Tier A range from 0-40.

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

Table 1.
Rating to raw score conversion (Writing).

| Rating | Raw Score |
| :---: | :---: |
| nonscorable | 0 |
| 1 | 1 |
| $1+$ | 2 |
| 2 | 3 |
| $2+$ | 4 |
| 3 | 5 |
| $3+$ | 6 |
| 4 | 7 |
| $4+$ | 8 |
| 5 | 9 |
| $5+$ | 9 |
| 6 | 9 |

The ACCESS 2.0 Writing Scoring Scale is distinct from the WIDA Writing Rubric, which is a tool for evaluating student writing in classrooms and for interpreting student scores from ACCESS 2.0 Online. The Writing Scoring Scale was designed specifically as a scoring tool and is not appropriate for any other purposes.

### 3.3.1.4 The ACCESS 2.0 Speaking Scoring Scale

The Speaking Scoring Scale defines five score points: Exemplary, Strong, Adequate, Attempted, and No Response (the final score point only applies if the rater uses one of three non-scorable codes: $\mathrm{B}=$ Blank response; $\mathrm{F}=$ Foreign language response; $\mathrm{I}=$ Indecipherable response). These score points are applied based on the proficiency level expectations of each task; that is, the level of language proficiency that each task is designed to elicit. These expectations are exemplified by the model student response (See Section 3.2.4). In this way, the model response serves as a scoring benchmark. Raters listen to the model response and score test taker responses relative to the model. A score of Exemplary means that the student response demonstrates English language use that is equal to or beyond the English language use illustrated by the model student's response.

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

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

To calculate a raw score for the Speaking test, the five score points are converted to whole numbers, as shown in Table 2. To calculate a total raw score, the raw scores for each task are added together; additionally, in Tier B/C, six points are added to the total raw score, representing a score of Adequate and Above for three tasks targeting language at PL 1. Though a Tier B/C student would not be administered any tasks targeting the PL 1 level, it is assumed that a score of Adequate and Above would be applicable to such tasks. Thus, on the pre-A test, scores can range from 0-6; on the A test, from $0-18$; and on the $B / C$ test, from 6-30.

Table 2.

| Rating to raw score conversion (Speaking). |  |
| :--- | :--- |
| Rating | Raw Score |
| No Response (B, F, or I)* | 0 |
| Attempted | 1 |
| Adequate/Adequate and Above | 2 |
| Strong | 3 |
| Exemplary | 4 |

* $\mathrm{B}=\mathrm{Blank}$ response; $\mathrm{F}=$ Foreign language response; $\mathrm{I}=$ Indecipherable response

Speaking tasks are scored using the ACCESS 2.0 Speaking Scoring Scale. The Speaking Scoring Scale is distinct from the WIDA Speaking Rubric, which is a tool for classroom use and score interpretation. The Speaking Scoring Scale was designed specifically for test scoring use and is not intended for classroom purposes.

### 3.3.2 How are measures of student performances on ACCESS 2.0 Online calculated?

The measurement model that forms the basis of the analysis for the development of ACCESS 2.0 Online is the Rasch measurement model (Wright \& Stone, 1979). Additional information on its use in the development of the ACCESS assessment program is available in the WIDA Consortium Technical Report No. 1, Development and Field Test of ACCESS for ELLs (Kenyon, 2006). The original ACCESS test was developed using Rasch measurement principles, and in that sense, the Rasch model guided all decisions throughout the development of the assessment and was not just a tool for the statistical analysis of the data. Thus, for example, data based on Rasch fit statistics guides the inclusion, revision, or deletion of items during the development and field testing of the test forms, and will continue to guide the refinement and further development
of the test. All Rasch analyses are conducted using the Rasch measurement software program Winsteps (Linacre, 2006).

### 3.3.2.1 Rasch Model for Dichotomous Scoring

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

$$
\log \left(\frac{P_{n i 1}}{P_{n i 0}}\right)=B_{n}-D_{i}
$$

where
$P_{n i 1}=$ probability of a correct response " 1 " by person " n " on item " i "
$P_{n i 0}=$ 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), $P_{n i 1} / P_{n i 0}$ 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.

### 3.3.2.2. Rasch Model for Polytomous Scoring

For the Writing and Speaking tasks, a Rasch-grouped rating scale model is used. Mathematically, this can be represented as
$\log \left(\frac{P_{n g i k}}{1-P_{n g i(k-1)}}\right)=\beta_{n}-D_{g i}-F_{g k}$
where
$P_{n g i k}=$ probability of person " n " on task " i " receiving a rating at level " k " on rating scale " g "
$P_{n g i(k-1)}=$ probability of person " n " on task " i " receiving a rating at level " $\mathrm{k}-1$ " on rating scale "g" (i.e., the next lowest rating)
$\beta_{n}=$ ability of person " n "
$D_{g i}=$ difficulty of task " i " specific to rating scale " g "
$F_{g k}=$ calibration of step "k" on rating scale " g "
The subscript " g " is a group index specifying the group of tasks to which task " i " belongs. It also identifies the scoring scale that was used for the group of tasks.

As described in section 3.3.1.3, ratings on the ACCESS 2.0 Writing Scoring Scale range from 0, $1,1+, \ldots, 6$ and the possible raw scores range from $0-9$. All Writing tasks are scored using this
scoring scale except for Grade 1 Tier A Task 1 and 2. The profiles of the responses to these two tasks do not fit the generic scoring scale well, so additional task-specific instructions are provided to raters. These instructions guide raters in applying a limited number of score points on the scoring scale to responses elicited by these two tasks. The possible ratings for Grade 1 Tier A Task 1 are 0 or 1 and the possible ratings for Grade 1 Tier A Task 2 are $0,1,1+$, or 2 . To simplify the year-to-year linking process, the Grade 1 Writing Tier A Task 1 is treated as a dichotomously-scored task. The Grade 1 Writing Tier A Task 2 is modeled using a rating scale with possible raw score of 0 to 3 . All other Writing tasks are modeled using a rating scale with possible raw scores of 0 to 9 . Thus there are total of two rating scales being modeled for ACCESS Writing. One rating scale is associated with the Grade 1 Writing Tier A Task 2, and the other rating scale is associated with all Writing tasks that are scored using the rating scale with raw score values 0 to 9 .

For Speaking, PL 1 tasks are modeled as a group on a $0-2$ scale and PL 3 and PL 5 tasks are modeled as a group on a $0-4$ scale (see section 3.3.1.4).

### 3.3.2.2 Scale Scores and Proficiency Level Scores

Scale scores are calculated by transforming the person ability estimate via a scaling equation. The scaling equations for each domain are provided in II.1.2, under Claim 4.3 in the CAL Validation Framework. In the domains of Listening and Reading, the ACCESS scale was maintained through the transition from ACCESS 1.0 to ACCESS 2.0 in Series 400, and is continued to Series 401 (evidence for scale maintenance from ACCES 1.0 to ACCESS 2.0 can be found in Center for Applied Linguistics [2016]). In the domains of Writing and Speaking, a study was conducted in the summer of 2016 to reconstruct the logit scale (see Center for Applied Linguistics [2017]).

Proficiency Level (PL) scores are interpretations of these scale scores in terms of the PLs described in the WIDA ELD Standards. These interpretations derive from a series of standard setting studies, in which educators reviewed evidence from the test, either in the form of items for the selected response sections (Listening and Reading) or student portfolios for the constructed response sections (Writing and Speaking), to establish cut scores between the PLs. The first standard setting study for ACCESS took place in 2005; it established cut scores for all four domains by grade-level cluster (Kenyon, 2006). The second cut score study took place in 2007; it established cut scores for all four domains by grade level (Kenyon, Ryu, \& MacGregor, 2013). These cut scores were used to derive PL scores through Series 400 of ACCESS 2.0 Online. A third cut score study was conducted in summer 2016 (Cook and MacGregor, 2017). The purpose of this study was to re-examine cut scores for each of the PLs on the new ACCESS 2.0 assessment in light of the migration from the paper-and-pencil only assessment, the revision of the Speaking test, and the influence of college- and career-ready standards.

Test Series 401 is the first series which employed these newly revised proficiency level cut scores.

A PL score consists of a two-digit decimal number (e.g., 4.5). The first digit represents the student's overall PL range based on the student's scale score. The number to the right of the decimal is an indication of the proportion of the range between cut scores that the student's scale score represents. A score of 4.5, for example, tells us that the student is in PL4 and that his/her scale score is halfway between the cut scores for Levels 4 and 5.

Unlike the scale scores, which form an interval scale and are continuous across grades from Kindergarten to Grade 12, PL scores are dependent upon which grade a student was in when ACCESS 2.0 Online was administered. Using the cut scores newly in effect for Series 401, if a Grade 2 student receives a 350 in Listening, it would be interpreted as a PL score of 5.8; if a Grade 5 student receives a 350 in Listening, it would be a 3.8 ; if a Grade 8 student receives a 350 in Listening, it would be a 3.1; and if a Grade 12 student receives a 350 in Listening, it would be a 2.3.

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

### 3.3.2.3 Composite Scores

Four composite scores are calculated for ACCESS 2.0 Online: Oral language, Literacy, Comprehension, and Overall. Composite scores are calculated as weighted averages of domain scale scores, as follows:

- Oral Language: 50\% Listening + 50\% Speaking
- Literacy: 50\% Reading + 50\% Writing
- Comprehension: 30\% Listening + 70\% Reading
- Overall Composite: $15 \%$ Listening $+15 \%$ Speaking $+35 \%$ Reading $+35 \%$ Writing


### 3.4 The Assembly Model: How are the assessment components for ACCESS put together?

This section describes how ACCESS 2.0 Online is assembled to ensure that the evidence collected is (a) sufficient to make the intended decisions, and (b) appropriate for the student's level of proficiency. In order to tailor the test closely to student ability levels while still including items and tasks that assess all of the Standards, adaptivity has been built into the test. The Listening and Reading tests both use a multistage adaptive test design. The Writing and Speaking tests are tiered, and placement into the tiers depends on performance on the Listening and Reading tests. Details are presented below.

### 3.4.1 Listening

The Listening test uses a multistage adaptive design, as illustrated in Figure 5. All students begin the Listening test with two entry folders (with three items each) at Stage 1 and Stage 2, both targeting SIL (See Section 2.2.1 for the WIDA ELD Standards and their abbreviations). At that point, the student's ability is estimated based on performance on those six items, and that ability estimate is used to determine which of the three leveled LoLA folders in Stage 3 is administered next. Students whose ability estimate predicts a PL score of 5.0 or higher are routed into the folder at the highest level (C in Figure 5); students whose ability estimate predicts a PL score of 2.5 or lower are routed into the folder at the lowest level (A in Figure 5); all others are routed into the B folder. ${ }^{1}$ Throughout the test, a student's underlying measure of ability is re-estimated with the completion of each folder, and the level of the next folder to be administered is chosen accordingly, following the decision rules above. Thus, each student will trace a tailor-made path through the test according to ability level, but the order of the stages is invariant across students. In total, there are eight possible stages, but students whose ability estimate falls below PL 2.5 after the sixth stage end the test at this point. The intent of this design is to ensure coverage of the Standards while delivering a test that closely matches the student's PL, thus minimizing measurement error.


Figure 5. Format of the Listening test.

[^0]
### 3.4.2 Reading

Figure 6 shows the format of the Reading test. The format and adaptivity are similar to the Listening test, but the Reading test consists of ten stages rather than eight. This reflects the greater weight given to Reading in calculating the composite scores, as well as the view that literacy skills are paramount in developing academic language proficiency. The greater weight afforded to Reading and Writing resulted from a policy decision by the WIDA Board before the first operational administration of ACCESS.


Figure 6. Format of the Reading test.

### 3.4.3 Writing

Figure 7 shows the format of the Writing test. As can be seen from the figure, Writing is tiered. Tier A consists of tasks written to elicit language at PLs $1-3$, while Tier B/C is designed to elicit language at PLs 4-6. With the exception of Grade 1 Tier A, both tiers consist of three tasks. Both tiers include tasks that target a single standard and tasks that integrate more than one WIDA Standard. For example, in the Tier A forms (except for Grade 1), one task integrates the Language of Math and the Language of Science. On the Tier B/C forms, one task integrates the Language of Math and the Language of Science, while another extended task integrates Social Instructional Language, the Language of Language Arts, and the Language of Social Studies. The ways in which the Standards are targeted by these tasks vary across grade levels and are spelled out in the generative item specifications.

| TaskI | Task II | Task III |
| :---: | :---: | :---: | :---: | :---: |

Figure 7. Format of the Writing test.
Note: Grade 1 Tier A follows a different model, and has four tasks targeting PLs 1, 1, 2, and 3. Numbers inside the boxes represent the targeted proficiency level of the task; the smaller numbers on the right edge of each box represent the range of proficiency levels that a task may elicit.

Placement into tiers on the Writing test depends on how students perform on the Listening and Reading tests, which receive computerized scores. To determine how to best place students into a tier, the previous year's test data for all students who were administered the assessment are analyzed to examine the relationship between how students perform on Listening and Reading and how they perform on Writing using logistic regression analyses. This information is used to program an algorithm into the ACCESS 2.0 Online test that will be used by the computer to determine which tier of the Writing test will be administered to each student. The purpose of the algorithm is to place students who are predicted to score above PL 3.0, based on their performances in Listening and Reading, into Tier B/C for Writing and Speaking, and all other students into Tier A.

### 3.4.4 Speaking

Figure 8 shows the format of the Speaking test. The Speaking test includes tasks that target language elicitation at three PLs: 1,3 , or 5 . The tasks are grouped into thematic folders, which are aligned to one or two of the WIDA Standards.

As shown in Figure 8, the Speaking test includes three tiers: Tier Pre-A, Tier A, and Tier B/C. Tier Pre-A includes tasks that target language elicitation at PL 1. Tier A includes tasks that target language elicitation at PLs 1 and 3. Tier B/C includes tasks that target language elicitation at PLs 3 and 5.

A thematic panel refers to the folders across all tiers within a grade-level cluster that relate to a particular WIDA ELD Standard. For example, the Tier B/C, Tier A, and Tier Pre-A folders that address SIL make up a single thematic panel. Ideally, within a thematic panel, tasks at PL 1 and PL 3 are the same across tiers. ${ }^{2}$ For example, within a SIL panel, the same PL 3 task appears on both the Tier A and the Tier B/C forms of the test.


Figure 8. Format of the Speaking test.

As with Writing, placement into the three tiers on the Speaking test shown in Figure 8 depends on performance on the Listening and Reading tests. An algorithm is applied to the results of the Listening and Reading test to determine which tier is optimal for the student. Unlike Writing, the Speaking test has one additional tier, Tier pre-A. Students are placed into Tier pre-A when their scores on Listening and Reading fall below chance performance. The Speaking pre-A tier is designed to meet the needs of students in the very early stages of English language development. As noted above, tasks are targeted to the P1 level and are scored using a modified version of the full Speaking rating scale.

[^1]
## 4 Assessment Performance: The Implementation of ACCESS 2.0 Online

This section focuses on Assessment Performance (Step 5) in CAL's Validation Framework. This section reviews how items and tasks for ACCESS 2.0 Online are developed, reviewed, revised, and chosen for inclusion in the operational test. It also describes the interaction between students and the test.

The development process for the first implementation of the ACCESS 2.0 Online represents a break from the regular cycle of item development and refreshment, as all of the items and tasks were developed specifically for the ACCESS 2.0 online environment.

### 4.1 How is ACCESS implemented?

### 4.1.1 Listening and Reading

Series 401 represents the second implementation of ACCESS 2.0. Prior to the transition to ACCESS 2.0, a refreshment cycle was in place in which roughly one-third of all Listening and one-third of all Reading items were targeted for refreshment annually. During the transition to ACCESS 2.0, item development in Listening and Reading followed a different trajectory.

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

The Listening items on Series 400 were developed entirely for ACCESS 2.0 Online. No refreshment took place between Series 400 and Series 401 in the domain of Listening. The Listening items used on Series 401 were field tested prior to the operational launch of Series 400. For detail on the field testing of these items, see Center for Applied Linguistics (2016).

During the operational administration of Series 401, students experienced operational Listening items plus embedded Series 402 field test items. The embedded field test items included innovative item formats, including hot spot and drag-and-drop items, where the student either clicked on an area of the screen or dragged an image/text to a specified screen area to respond.

In the domain of Reading, items on Series 400 were developed based on operational test items from previous test series, adapted for implementation in the online environment. Item format was adjusted to optimize the items for presentation on computer screens. For example, on the paper test, students might have read a single "theme passage" with multiple items related to the single passage. The test booklet was laid out so that the student could see the passage and all three
items simultaneously. In the online format, the student sees only one item per screen. Therefore, the format was adjusted so that each item has its own passage.
For Series 401, approximately one-sixth of Reading items were targeted for refreshment. An enhancement in development for Series 401 was the inclusion of innovative item formats, including hot spot and drag-and-drop items, where the student either clicked on an area of the screen or dragged an image/text to a specified screen area to respond.

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

At this point, items undergo external bias, sensitivity, and content reviews, after which they undergo further refinement. Items that reach this point are then administered as embedded field test items on students' operational assessments.

For Series 401, a total of 126 Reading items ( 42 folders) were field tested, across all five gradelevel clusters, embedded into the Series 400 operational assessment. Each student receives one Listening and one Reading field test folder embedded into their operational test. Field test folders are targeted to refresh a specific operational folder on the test, and field test folder specifications include the stage, standard, and tier pool target (A, B, or C) of the folder. Students are administered the embedded field test folder at the stage targeted for refreshment, with administration randomized so that half of the students see the field test folder before the corresponding operational folder, and half see the operational folder before the field test folder. Field test folders are administered to those students who are routed to take the Operational folder that is either at the same tier or adjacent to the tier that the field test folder targets. When field test samples are drawn, the sample includes $50 \%$ of students at the tier targeted by the field test folder, and $50 \%$ at adjacent tiers. Field test sample targets in Listening and Reading are set at a minimum of 3,000 responses per folder.

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

As is the case for Listening, students taking the operational Series 401 Reading test receive one embedded Series 402 field test folder during their operational assessment.

### 4.1.2 Writing

The development of Writing tasks is similar to that of Listening and Reading items. Writing tasks, however, do not currently undergo large-scale field testing. Instead, after external bias,
sensitivity and content reviews, they are subject to small-scale tryouts, led by CAL staff. In these tryouts, candidate folders for Grades 4-12 are administered to students using a mocked-up computer delivery; as noted above, students in Grades $1-3$ complete the Writing assessment with a traditional paper-and-pencil administration. Student responses, as well as observations and interviews, inform further revisions to the folders.

Then, a small-scale stand-alone field test of Writing folders is conducted. For Series 401, a total of 5 Writing tasks were field tested. A sample target of 75 students per task was established. The field test uses the online interface where applicable, and the field test is administered under standard testing conditions, with responses captured online where applicable. For the writing field test, responses are double-scored and adjudicated by CAL experts, and qualitative analysis of the collected responses is conducted. The main purposes of this small-scale field testing are (a) to confirm that the tasks are working as intended, (b) to identify anchor samples for rater training, and (c) to inform the rating of the tasks when they become operational.

The Writing items on Series 400 were primarily adapted to the computer from operational items from previous test series prior to the launch of Online ACCESS (Series 203, 301, and 302). Series 401 incorporates continuing items from Series 400 as well as items newly developed and field tested for Series 401. Major differences between Series 400 and Series 401 are that the extended task was refreshed on the Tier B/C form, with the introduction of tabbed browsing and click-to-enlarge functionalities in the input for Grades 4-12 tasks.

### 4.1.3 Speaking

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

Many of the folders used in Series 401 were previously field tested as part of Series 400 and then further revised and field tested again for Series 401. Much of the content of the Speaking items on Series 401 was adapted to the computer from both operational items from previous paperbased test series and from materials that were not developed to finality for previous test series. Some folder content was created specifically for ACCESS 2.0 Online.

All students are administered a Speaking field test folder appended to their operational Speaking assessment. A total of 84 tasks ( 42 folders) were field tested for Series 401, with a target sample size of 300 students per folder. Responses were double-scored by DRC trained raters, and adjudicated by CAL raters.

The Speaking test underwent a major overhaul between ACCESS 1.0 and ACCESS 2.0. The Speaking test was previously administered one-on-one with a live test administrator, who scored the test as it was administered. Each folder had tasks at each proficiency level (1-5). The test administrator used "stopping rules" to determine when the test taker could no longer answer
appropriately and when to move on to the next folder. The test administrator also was permitted to ask follow-up questions to elicit additional responses from the student.

For ACCESS 2.0 Online, folders were designed to target one or two proficiency levels: Tier PreA folders include one task which targets only PL 1. Tier A folders include tasks that target PLs 1 and 3. Tier B/C folders include tasks that target PLs 3 and 5. Students are routed into a tier based on their performance in the Reading and Listening sections of the test. The content is presented entirely on the computer, and the responses are recorded by the test engine and are transmitted to DRC for scoring.

### 4.2 What is the assessment delivery experience for students taking ACCESS 2.0 Online?

### 4.2.1 Listening and Reading

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

### 4.2.2 Writing

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

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

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


### 4.2.3 Speaking

Speaking tasks are delivered online. Students listen to prompts via headsets that are equipped with microphones to capture their responses. Extensive support is provided to the student through illustrations and written input designed to provide sufficient content for the response, as well as a model student response that is intended to provide guidance regarding the level of linguistic complexity required to respond adequately (see Section 3.2.4).

### 4.3 Assessment performance-interaction between test and student

Administration of ACCESS 2.0 Online takes place between December and April of the academic year, with testing windows determined at the state level. The Reading and Listening tests are administered first (in either order), followed by Writing and Speaking (in either order). The test may be administered in several sessions within one day or over a series of days. Student performance on the test forms the basis for developing Assessment Records, which are addressed in detail in Part II of this report.

## Part II: Assessment Records

In Part II of the Annual Technical Report, the focus is on the Assessment Records step in the CAL Validation Framework (see Part I.1.2, for a full description of the framework). Section 1 details the claims made regarding assessment records and provide references to evidence that supports those claims. Section 2 provides descriptions of the data and analyses presented in Section 3. In Section 3, detailed data and analyses are presented regarding the most recent operational administration of ACCESS 2.0 Online.

## 1 Assessment Records for ACCESS 2.0 Online

The complete validation framework for ACCESS for ELLs assessment program, as described in Part I of this report contains seven steps. Part I of this report focuses on the initial three steps (Plan, Design, and Assessment Performance). The argumentation and the data presented in this part (Part II) address Assessment Records, and present evidence specific to ACCESS 2.0 Online. By focusing on Assessment Records (i.e., test scores and proficiency level descriptions), the information here will be used to support claims related to the quality and consistency of the assessment data gathered and analyzed using ACCESS 2.0 Online. The claims in this step of the Assessment Use Argument (AUA) all pertain to the general question: How do we know that the reported language domain scores and composite scores on ACCESS 2.0 Online are consistent and dependable?

The diagram in Figure 1 shows a visual representation of an argument-based approach for supporting claims related to Assessment Records (Step 4). The figure shows how claims related to Assessment Records fit into the complete validation framework. Evidence in the form of data from this report or other sources will be presented to support these claims as they relate to ACCESS 2.0 Online. Section 1.2 provides an overview of the sources of evidence which support the argument.


Figure 1. Structure of the argument-based approach supporting Assessment Records (Step 4).

### 1.1 Claims for the Assessment Records for ACCESS

Assessment Records (Step 4) of the CAL 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 (ELD) Standards.

As shown in Figure 2, these claims depend upon each other, moving from (C4.6) down to (C4.1). Within this organizational structure, each successive claim requires that the previous claim be met in order for it to support the validation argument.

The claim that tasks and items are scored consistently (C4.5) does not support the overall validity argument unless the claim that all test takers are provided with comparable opportunities (C4.6) is also met. In other words, tasks and items may be scored consistently for all test takers, but if all test takers are not provided with comparable opportunities, then consistent scoring in and of itself does not support the validity argument. Likewise, support for the claim that test items or tasks work appropriately together to measure English language proficiency (C4.4) requires that those items or tasks be consistently scored (C4.5), otherwise C4.4 cannot support the validity of the assessment. C4.3 asserts that scale score interpretation remains consistent over time-one requirement for this to be true is that the assessment must be able to measure students across a broad range of English language proficiency abilities (as claimed at C 4.4 ). While comparability of opportunity is evinced by the steps taken to ensure that the implementation of the ACCESS test is equitable, C 4.2 looks at measurement, or how student performance is translated into a quantifiable outcome. In order for this to be done in a fair and unbiased manner across time, C4.3 must be met. Finally, the appropriate classification of test takers ( C 4.1 ) cannot be accomplished unless the performance of all test takers is measured in a fair and unbiased manner (C4.2).

Each prior claim alone does not constitute the entirety of the evidence for the successive claims, however; while each claim requires the evidence from its predecessor, it also requires additional evidence to be supported fully. Section 1.2 below provides a fully fleshed out structure of the line of argumentation for Assessment Records, including actions that are taken to ensure the
consistency and reliability of the assessment records, and evidence to demonstrate that those actions are taken.


Figure 2. Progression of claims for Assessment Records (Step 4).

### 1.2 Evidence for Assessment Records Claims of ACCESS 2.0 Online

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. In what follows, we outline the location within this Annual Technical Report or the external sources that provide evidence related to each action. A summary table of this information is presented in Section 1.3, below.

Because these claims relate to Assessment Records, which is 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. Individual actions to ensure each claim are denoted by the corresponding letter ( $\mathrm{a}, \mathrm{b}, \mathrm{c}$, and so on).

Note that the Assessment Records claims are claims for the ACCESS assessment program. The evidence provided for these claims in this report is evidence specific to ACCESS 2.0 Online Series 401.

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

Action 4.6a: Test design and student training procedures ensure that all students are able to interact with the technology of the test.

Evidence: CAL conducted extensive cognitive laboratories to ensure that students at all grade levels and at the lowest proficiency levels could successfully manipulate the student test interface.

A Test Demo video is available for all students to view prior to testing. This video walks the students through all aspects of testing.

The Test Practice items, which appear in the operational test prior to the operational items, are also available as stand-alone packages for students to familiarize themselves with the computer interface prior to testing.

Procedures for administering the test are documented in the Test Administration Manual.
All test domains contain an audio check prior to administration to ensure that students can hear the audio stimulus. In addition, the Speaking test, which requires that students speak into a microphone to capture their oral responses, contains a check to ensure that the students are speaking loudly enough for the interface to successfully record the response. This check occurs at the beginning of the Speaking section of the test. In addition, as the students record their responses, the interface detects the volume level as students respond, and prompts them to try again if they speak too softly. A further measure ensures that if the student does not speak loudly enough a second time, the test pauses and prompts the students to raise their hands for assistance.

Action 4.6b: Procedures are in place to address technical issues and interruptions.
Evidence: Procedures on handling technical issues and interruptions are detailed in the DRC INSIGHT Technology User Guide, Volume V (Data Recognition Corporation, 2016a), as well as in the WIDA AMS User Guide (Data Recognition Corporation, 2016b). The ACCESS for ELLs 2.0 Test Administrator Manual (WIDA Consortium, 2016) details the steps that test administrators can take during testing, and includes a troubleshooting chart, as well as information on how to contact DRC Customer Support. WIDA also offers a series of webinars
which focus on issue resolution during testing, including some with a specific focus on technological issues. The WIDA website also has a compilation of technology FAQs available. ${ }^{1}$

WIDA and DRC also collaborate to create documents and memos to address issues in the field. For example, on the ACCESS 2.0 Technology webpage, an iPad Troubleshooting Guide and a Whitelisting Memo have been added in response to common questions and concerns.

WIDA maintains a number of tools on their website in case of technical issues or interruptions.
WIDA has a system status page on their website to monitor and track system outages. A troubleshooting chart accompanies this page.

DRC maintains a customer service email account and phone number for technical issues. In the event of a systemic issue or outage, educators have access to the WIDA System Status Dashboard. ${ }^{2}$

Should an outage or technical issue occur, DRC notifies State Education Agencies (SEAs) via email as to when the systemic issue occurs as well as when the issue is resolved, noting which aspects of testing or testing devices were impacted. Additionally, for extended technical issues, WIDA posts general information pertaining to the outage on the main page of the WIDA website. In the event of extended technical issues, WIDA and DRC provide updates to SEAs via email as follows: (1) broadcast message/announcement of incident; (2) update(s) on the incident (if not resolved after two hours); (3) restoration of service message; (4) root cause analysis message; (5) solution confirmation message. In the event that DRC needs to schedule maintenance to fix the underlying issue, a final message is sent out once this maintenance occurs and a solution is implemented.

Action 4.6c: Administration procedures are in place to ensure consistency in test administration.
Evidence: Procedures for administering the test are documented in the Test Administrator Manual.

The Test Demo and Test Practice items (see Action 4.6a) are also available for teachers to familiarize themselves with the test prior to administration.

WIDA provides webinars and other training courses on their website to orient new test administrators to test administration procedures. The training courses include certification quizzes to ensure that test administrators properly understand the processes prior to administration.

[^2]Action 4.6d: Procedures are in place to ensure that items and tasks do not have issues with bias or sensitivity.
Evidence: As detailed in Part I.4.1.1-I.4.1.3 of this report, all test items and tasks are subject to bias and sensitivity reviews. These reviews examine items to ensure that they do not favor students from a particular socioeconomic status, geographic area, or educational background, or introduce other systematic biases.

Action 4.6e: Test administrators document and report any irregularities that may occur so that appropriate action may be taken.
Evidence: General processes and procedures for test irregularities due to student conditions, testing environment, or other unusual occurrences can be found in the District and School Test Coordinator Manual. Specific testing situations, including where to start and stop the test, when breaks can be taken, material management protocols in the case of damaged testing materials, and other detailed guidance can be found in the Test Administrator Manual. Both the District and School Test Coordinator Manual and the Test Administrator Manual can be found on WIDA's website. States each have a specific policy for Test Administrators to follow in the case of a testing irregularity, which can include steps such as documentation to use or notification procedures to follow. These state-specific steps can be found on the ACCESS for ELLs 2.0 State Checklists, found on the state pages ${ }^{3}$ of the WIDA website and within the training course. Frequently asked questions regarding interruptions can be found in the ACCESS for ELLs 2.0 FAQ section of the WIDA website. ${ }^{4}$ Additionally, the ACCESS for ELLs 2.0 Training Course highlights common testing irregularities and the resources to use in such circumstances.

Should the Test Administrator have additional questions about how to proceed in the event of a testing interruption or irregularity, the WIDA Client Services Center can be contacted via email or phone at help@wida.us or toll free at 1-866-276-7735.

## C4.5. All tasks and items 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: Part I.3.3.1.2 specifies the scoring procedure for performance-based tasks in ACCESS 2.0 Online. Raters of performance-based tasks are trained by DRC to appropriately use the Writing and Speaking scoring scales (detailed in Sections 3.3.1.3 and 3.3.1.4, respectively) to score performance-based tasks.

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

[^3]Evidence: Part I.3.3.1 specifies the scoring procedure for ACCESS 2.0 Online. Listening and Reading items are dichotomous and are scored electronically by DRC (see Part I.3.3.1.1).

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

Evidence: Part I.3.3.1.2 specifies the scoring procedure for ACCESS 2.0 Online. Writing and Speaking tasks are centrally scored at DRC, and all raters are pre-screened, trained, and subject to qualifying scoring tests before becoming operational raters. Once raters are qualified, they then undergo additional training on the grade-level cluster and specific tasks they will be scoring. Following this more intense training, they rate calibration sets to ensure that they are properly calibrated to the grade-level cluster and task(s).

Action 4.5d: Raters of performance-based tasks are monitored daily to ensure that they are scoring appropriately.

Evidence: DRC provides raters of performance-based tasks with specially prepared calibration sets each day to ensure that the scoring rubric is being applied consistently across scoring sessions (see Part I.3.3.1.2). For the Writing test, pre-rated and vetted validation sets are seeded into the operational items for scoring. The validation sets are utilized to ensure that raters are scoring accurately and consistently and that any drift is identified and promptly corrected. For the Speaking test, pre-rated and vetted recalibration sets are administered to raters. Raters take these sets every day to ensure that they are calibrated. Due to the nature of the Speaking test structure, validation sets cannot be seeded into the Speaking scoring queues, so the recalibration sets are needed.

Action 4.5e: Scoring data for performance-based tasks are analyzed for rater agreement to understand how closely raters agree.

Evidence: For a sample of $20 \%$ of responses to each task, interrater reliability is calculated for each of the Writing and Speaking tasks (see Section 2.2.10). During operational scoring, these data are monitored daily for quality control purposes.

Action 4.5f: Raters of performance-based tasks are monitored over time to ensure that they apply the scales in a consistent way (internal consistency).

Evidence: Part I.3.3.1.2 details the procedures used by DRC to monitor raters. This includes ongoing quality control checks and procedures, and investigation of any irregularities.

For the Writing test, pre-rated and vetted validation sets are seeded into the operational items for scoring. The validation sets are utilized to ensure that raters are scoring accurately and consistently and that any drift is identified and promptly corrected.

For the Speaking test, pre-rated and vetted recalibration sets are administered to raters. Raters take these sets every day to ensure that they are calibrated. Due to the nature of the Speaking test structure, validation sets cannot be seeded into the Speaking scoring queues, so the recalibration sets are needed.

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

Action 4.4a: For each domain and grade-level cluster (e.g., Reading 6-8), item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.

Evidence: Listening and Reading reliability are computed using the reliability coefficient described in Thissen (2000). For the Writing and Speaking domains, Cronbach's alpha is computed for each tier and also for each grade-level cluster, across tiers. Section 2.2.10 describes the ways in which test reliability is computed for the domains.

Action 4.4b: For each composite score, psychometric properties are evaluated to confirm that scores are internally consistent.

Evidence: To compute reliability for the composites, a stratified Cronbach's alpha is used. Section 2.3.3 describes the ways in which test reliability is computed for the composites.

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

Evidence: Section 2.2.1.1 describes the Rasch fit statistics that are computed for each item; the statistics are detailed in Table A, Complete Item/Task Analysis and Summary, in Section 3.3.

Action 4.4d: Items and tasks of appropriate difficulty are chosen for each domain.
Evidence: The Complete Item or Task Analysis and Summary tables provide information on the difficulty of each item or task. Section 2.2.1 describes the construction of these tables. When the test is assembled, task difficulty is one of several criteria used to select appropriate items for operational assessment from the pool of field tested items.

Action 4.4e: Items in folders aimed at higher proficiency levels within a stage of the multistage adaptive tests (Listening and Reading) are more difficult than items in folders aimed at lower proficiency levels within the same stage.
Evidence: The Complete Item or Task Analysis and Summary tables include information on item difficulty (see Section 2.2.1.2).

Action 4.4f: Routing and placement procedures are in place to ensure that students are administered a test appropriate to their proficiency level.

Evidence: Part I of this report describes routing rules for Listening (I.3.4.1) and Reading (I.3.4.2), and placement rules for Writing (I.3.4.3) and Speaking (I.3.4.4).

Quality control procedures are in place to ensure that routing rules are implemented with fidelity in the computerized assessment.

Placement rules place students into tiers for Writing (A or B/C) and Speaking (pre-A, A, or B/C) tests. Evidence of the effects of these rules can be found in figures and tables which present raw
score and scale score distributions by tier and across tiers. Descriptions of the raw score distribution and scale score distribution tables can be found in Section 2.2.3 and Section 2.2.4, respectively.

## C4.3. 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 2.0 Online test form is refreshed, a number of items and tasks are retained from the previous year's assessment for the purpose of scale maintenance. Section 2.2.7 of this report describes the equating procedures used, and Table G presents item-by-item information, including information on which items or tasks were used as anchor items or tasks.

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:

i. Section 2.2.7 describes the equating summary included in this report.
ii. Previously used items and tasks (i.e., anchor items) are included on each test form along with new items and 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 following scaling equations are used to convert ability measures in logits to scale scores:

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

For Listening and Reading, these equations have been in use from the first operational administration of ACCESS (Series 100). Evidence for scale maintenance in Listening and Reading is detailed in the ACCESS for ELLs Series 400 Listening and Reading Scale Maintenance: Technical Brief (Center for Applied Linguistics, 2016).

For Writing and Speaking, scaling equations are new for Series 401. A scaling study was conducted in summer 2016 (see Center for Applied Linguistics [2017]). The equations derived from this scaling study were used for the first time in Series 401 (2016-17 operational year).

## C4.2. 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 DIF analysis and summary table provides a summary of the findings of the DIF analyses, which look for measurement bias in test items (see Section 2.2.2). Ethnicity (Hispanic vs. non-Hispanic) and gender DIF analyses are conducted. In the domains of Listening and Reading, DIF analyses are conducted prior to operational testing, using data from the previous year's operational and embedded field test items. In the domains of Writing and Speaking, DIF analyses are conducted using population data, after the conclusion of operational testing.

Action 4.2b: Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not included in the pool of items selected for Listening and Reading, or are removed from future test forms for Speaking and Writing.

Evidence: If an item shows C-level DIF, a content review panel is convened to examine the content of the item. The panel is composed of diverse members and is chosen carefully so that panelists include male and female members as well as bilingual individuals who speak either English and Spanish or English and another language. The panel then comes to a consensus decision on whether or not the item content is likely to favor or disfavor specific subgroups of students.

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

Action 4.1a: Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that ACCESS 2.0 Online measures the performance of test takers across the range of English Language Proficiency levels defined by the WIDA ELD Standards. Distributions of raw scores are analyzed where appropriate.

## Evidence:

i. The distribution of test takers' raw scores on ACCESS 2.0 Online for the Writing and Speaking tests, organized by individual test form (e.g., Writing 45B/C), shows the extent to which ACCESS 2.0 Online measures the performance of test takers across the range of ELD abilities that each form was designed to assess (see Section 2.2.3).
ii. The distribution of test takers' scale scores on ACCESS 2.0 Online for each domain, organized by test form, shows that ACCESS 2.0 Online measures the
performance of test takers across the range of ELD abilities that each form was designed to assess (see Section 2.2.4).
iii. The proficiency level distribution of test takers' scores on ACCESS 2.0 Online, for each domain, organized by individual test form, shows that ACCESS 2.0 Online measures the performance of test takers across the range of proficiency levels that each form was designed to assess (see Section 2.2.5).
iv. The Test Characteristic Curve graphically shows the relationship between test takers' ability measures (calculated based on test performance using Rasch modeling) on the horizontal axis and expected raw scores on the vertical axis. Test Characteristic Curves are provided for each tier for Writing and Speaking (see Section 2.2.8). (Note that there is no Test Characteristic Curve for Listening and Reading, as the notion of "expected raw score" is meaningless on the adaptive assessment.)

Action 4.1b: Distributions of scale scores and proficiency levels, for each domain and each composite, organized by grade-level cluster, are analyzed to confirm that ACCESS 2.0 Online measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA ELD Standards.

## Evidence:

i. The distribution of test takers' scale scores on ACCESS 2.0 Online, for each domain and each composite, organized by grade-level cluster, shows that ACCESS 2.0 Online measures the performance of test takers across the range of ELD abilities as described by the WIDA ELD Standards (see Section 2.2.4 and Section 2.3.1).
ii. The proficiency level distribution of test takers' scores on ACCESS 2.0 Online, for each domain and each composite, organized by grade-level cluster, shows that ACCESS 2.0 Online measures the performance of test takers across the range of proficiency levels as defined by the WIDA ELD Standards (see Section 2.2.5 and Section 2.3.2).
iii. The Test Characteristic Curve graphically shows the relationship between test takers' ability measures (calculated based on test performance using Rasch modeling) on the horizontal axis and expected raw scores on the vertical axis. Test Characteristic Curves are provided across each grade-level cluster for Writing and Speaking (see Section 2.2.8). (Note that there is no Test Characteristic Curve for Listening and Reading, as the notion of "expected raw score" is meaningless on the adaptive assessment.)

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 grade.

## Evidence:

i. The Test Information Function graphically shows the relationship between ability measure and the accuracy of test scores (see Section 2.2.9). Cut points are marked on the Test Information Function figures.
ii. Tables provide information on the conditional standard error of measurement (CSEM) at the cut scores for Writing and Speaking (Section 2.2.11).

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

Evidence: Accuracy and consistency statistics are calculated for each domain for the grade-level cluster (see Section 2.2.12).

Action 4.1e: Students are placed into the appropriate proficiency level based on their test scores. Evidence: A 2016 standard setting study established new cut scores for ACCESS 2.0 (Cook \& MacGregor, 2017). A short history of ACCESS cut score setting can be found in Part I.3.3.2.3 of this report.

Action 4.1f: Items and tasks are aligned to the WIDA Standards.
Evidence: See Cook (2007) for evidence of alignment between the WIDA Standards and the ACCESS assessment program. Part I.3.2 details the continuing development of items and tasks for ACCESS 2.0 Online to maintain alignment.

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

Table 1
Summary of Assessment Records Claims, Actions, and Evidence.

| Claim | Actions | Evidence |
| :--- | :--- | :--- |
| 6. All test takers are <br> provided comparable <br> opportunities to <br> demonstrate their <br> English Language <br> Proficiency. | a. Test design and student training procedures <br> ensure that all students are able to interact with <br> the technology of the test. | a. Evidence summarized with <br> claim at 4.6a. |
|  | b. Procedures are in place to address technical <br> issues and interruptions. | b. Administration procedures are in place to <br> ensure consistency in test administration. <br> claim at 4.6b. |
|  | d. Procedures are in place to ensure that items and <br> tasks do not have issues with bias or sensitivity. | c. Test Administration <br> Manual, plus additional <br> evidence summarized with <br> claim at 4.6c. |
|  | e. Test administrators document and report any <br> irregularities that may occur so that appropriate <br> action may be taken. | e. Evidence summarized with |
|  |  |  |$|$


| 4. Test items/tasks work appropriately together to measure each test taker's English Language Proficiency. | a. For each domain and grade-level cluster (e.g., Reading 6-8), item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent. | a. Section 2.2.10 |
| :---: | :---: | :---: |
|  | b. For each composite score, psychometric properties are evaluated to confirm that scores are internally consistent. | b. Section 2.3.3 |
|  | c. Analyses of Rasch model fit statistics are conducted to show that individual tasks perform appropriately. | c. Section 2.2.1.1 |
|  | d. Items and task of appropriate difficulty are chosen for each domain. | d. Section 2.2.1 |
|  | e. Items in folders aimed at higher proficiency level within a stage of the multistage adaptive tests (Listening and Reading) are more difficult than items in folders aimed at lower proficiency levels within the same stage. | e. Section 2.2.1.2 |
|  | f. Routing and placement procedures are in place to ensure that students are administered a test appropriate to their proficiency level. | $\begin{aligned} & \text { f. Sections I.3.4.1, I.3.4.2., } \\ & \text { I.3.4.3, I.3.4.4. } \end{aligned}$ |
| 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. Section 2.2.7 |
|  | 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. Section 2.2.7 |
|  | c. The same scaling equation is applied from year to year to ensure that scale scores are obtained consistently over time | c. Evidence summarized with claim at 4.3c. |
| 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. <br> b. Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not included in the pool of items selected for Listening and Reading, or are removed from future test forms for Speaking and Writing. | a. Section 2.2.2 <br> b. Evidence summarized with claim at 4.2b. |


| 1. Test takers are classified appropriately according to the proficiency levels defined in the WIDA English Language Development (ELD) Standards. | a. Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that ACCESS 2.0 Online measures the performance of test takers across the range of English Language Proficiency levels defined by the WIDA ELD Standards. Distributions of raw scores are analyzed where appropriate. | a. Sections 2.2.3, 2.2.4, 2.2.5, 2.2.8 |
| :---: | :---: | :---: |
|  | b. Distributions of scale scores and proficiency levels, for each domain and each composite, organized by grade-level cluster, are analyzed to confirm that ACCESS 2.0 Online measures the performance of test takers across the range of English Language Proficiency levels as defined by the WIDA ELD Standards. | b. Sections 2.2.4, 2.2.5, 2.3.1, 2.3.2, 2.2.9 |
|  | 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 grade. | c. Sections 2.2.9, 2.2.11 |
|  | d. Classification and accuracy analyses are conducted by grade level to confirm that proficiency level classifications are reliable for all domain and composite scores. | d. Section 2.2.12 |
|  | e. Students are placed into the appropriate proficiency level based on their test scores | e. Cook \& MacGregor (2017) and Part I.3.3.2.3 |
|  | f. Items and tasks are aligned to the WIDA Standards. | f. Cook (2007) and Part I.3.2 |

## 2 Background and Descriptions for the Presentation of Results

This section describes the tables and figures included in Section 3 of this report.

### 2.1 Student Participation and Performance

Student participation and performance is detailed in Section 3.2, which has three subsections: Participation (3.2.1); Scale score results (3.2.2); and Proficiency level results (3.2.3).

There are two subsets of students who are included in the descriptions of student participation and performance but are excluded from subsequent analyses. There are a subset of students who were flagged as potentially having experienced test interruptions. Using telemetry data, three variables were selected which might potentially indicate interruption (that is, testing experiences that are outside of regular testing experiences). The interruption indicators WIDA used are 1) longer than expected testing time, 2) number of appearances (i.e. more than 1) of test items, and 3) number of log-ins. Records are flagged if they fall outside of established criteria for any of these three indicators. Students whose records are flagged as interrupted are included in the tables which describe participation in the assessment (these tables are described in Section 2.1.1) but are excluded from all subsequent analyses. Table 2.2 summarizes the numbers of students who are excluded from these analyses.

Table 2.2
Students Excluded from Analysis due to Test Interruptions by Domain and Cluster

| Domain | Cluster | No. of Students | Percent |
| :---: | :---: | :---: | :---: |
| List | 1 | 12,544 | 9.12\% |
|  | 2-3 | 32,446 | 23.60\% |
|  | 4-5 | 18,583 | 13.52\% |
|  | 6-8 | 35,991 | 26.18\% |
|  | 9-12 | 37,916 | 27.58\% |
|  | Total | 137,480 | 100.00\% |
| Read | 1 | 12,298 | 5.89\% |
|  | 2-3 | 39,689 | 19.01\% |
|  | 4-5 | 39,707 | 19.02\% |
|  | 6-8 | 44,875 | 21.50\% |
|  | 9-12 | 72,170 | 34.57\% |
|  | Total | 208,739 | 100.00\% |
| Writ | 1 | 0 | 0.00\% |
|  | 2-3 | 0 | 0.00\% |
|  | 4-5 | 15,581 | 25.21\% |
|  | 6-8 | 19,846 | 32.11\% |
|  | 9-12 | 26,379 | 42.68\% |
|  | Total | 61,806 | 100.00\% |
| Spek | 1 | 16,908 | 13.86\% |
|  | 2-3 | 37,582 | 30.80\% |
|  | 4-5 | 18,276 | 14.98\% |
|  | 6-8 | 25,507 | 20.91\% |
|  | 9-12 | 23,737 | 19.45\% |
|  | Total | 122,010 | 100.00\% |

Students from the state of Michigan are included in the analyses of participation and performance, but these students are excluded from subsequent analyses, as data from the state of Michigan arrived later than expected, and it was not possible to include these data in all analyses for this report.

### 2.1.1 Participation

Participation in ACCESS 2.0 Online is shown in three ways: by grade-level cluster, by grade, and, for Writing and Speaking only, by tier. This is the first subsection of Student Participation and Performance.

### 2.1.1.1 Grade-Level Cluster

Section 3.2.1.1 gives information on participation by grade-level cluster.
Table 3.2.1.1.1 shows participation across the 37 WIDA states that participated in the ACCESS 2.0 Online operational testing program in 2016-2017. The first row shows the grade-level
cluster, the next 37 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 37states.

Table 3.2.1.1.2 shows participation by grade-level cluster by gender across all 37 states combined, while Table 3.2.1.1.3 shows participation by grade-level cluster by ethnicity across all 36 states.

Table 3.2.1.1.4 shows participation by grade-level cluster and tier for all Writing and Speaking forms.

### 2.1.1.2 Grade

Section 3.2.1.2 gives similar data as in the previous section, but broken out by grade rather than by grade-level cluster.

### 2.1.2 Scale Score Results

The second subsection of Student Participation and Performance provides information on students' scale score results.

### 2.1.2.1 Mean Scale Scores Across Domain and Composite Scores Section

Section 3.2.2.1 shows mean (average) scale scores by grade-level cluster across the eight scores awarded, first for the four domains (Listening, Reading, Writing, and Speaking) and then for the four composites (Oral Language, Literacy, Comprehension, and Overall Composite). In this section, under each average, the number of students in each group is also given.

Table 3.2.2.1.1 shows mean scale scores by grade-level cluster, while Table 3.2.2.1.2 shows the same information broken down by gender, and Table 3.2.2.1.3 shows the same information broken down by ethnicity and race. Following the approach of the Census Bureau, ethnicity is a binary category (Hispanic or Non-Hispanic), with five categories for race (American Indian/Alaskan Native, Asian, Black/African American, Pacific Islander/Hawaiian, and White) that are not mutually exclusive. Thus, for example, Student A may be labeled as Hispanic for ethnicity and Asian for race, while Student B may be labeled as Non-Hispanic for ethnicity and both American Indian/Alaskan Native and Black/African American for race. Students who are labeled as Hispanic are included in the Hispanic (Of Any Race) category, regardless of how many racial categories they are included in. Students who are identified in one racial category (e.g., Asian) who have not been identified as Hispanic are identified in only one racial category; if they are identified in more than one racial category and have not been identified as Hispanic, they are labeled Non-Hispanic Multi-racial.

Section 3.2.2.2 shows the mean scale scores broken down by grade rather than by grade-level cluster. Table 3.2.2.2.1 shows mean scale scores by grade, while Table 3.2.2.2.2 shows the same information broken down by gender, and Table 3.2.2.2.3 shows the same information broken down by ethnicity and race.

### 2.1.2.2 Correlations

Tables 3.2.2.3A through 3.2.2.3E show correlations among the four domain scale scores by grade-level cluster across all tiers, as well as the number of students included in each correlation. Table 3.2.2.3A shows the results for Grade 1, Table 3.2.2.3B shows the results for Cluster 2-3, Table 3.2.2.3C shows the results for Cluster $4-5$, Table 3.2.2.3D shows the results for Cluster 6 8, and Table 3.2.2.3E shows the results for Cluster 9-12. Note that all correlations in Tables 3.2.2.3A through 3.2.2.3E are significant at the 0.01 level (2-tailed).

### 2.1.3 Proficiency Level Results

The third subsection of Student Participation and Performance covers results by proficiency level, and shows the distribution of students falling into the six language proficiency levels defined in the WIDA ELD Standards. Section 3.2.3.1 provides the results for Domains, while Section 3.2.3.2 provides the results for Composite scores.

Within each section, results are first presented by grade-level cluster, then by grade. For both, the first table shows the number of students classified into each language proficiency level (count), while the second table shows the results in terms of percentages within each row.

### 2.2 Analyses of Domain Scores

Section 3.3 presents a series of tables and figures pertaining to scores in the four domains. The tables and figures are organized by grade-level cluster, then by domain, then, where relevant, by tier. Tables and figures are numbered through the text according to their grade-level cluster and domain (and tier, where relevant); each table or figure is then labeled by a letter designation which indicates the table or figure type. Thus in Section 3.3, Table 3.3.1.1.A indicates that the table refers to the first grade-level cluster covered in the section (Grade 1) and the first domain covered (Listening). The letter designation, in this case, A, indicates that the table is a Complete Item Analysis and Summary table-so Table A appears for each relevant grade-level cluster, domain, and tier.

### 2.2.1 Complete Item or Task Analysis and Summary

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

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

For Listening and Reading, Table A provides information on every item in the grade-level cluster. For Writing, Table Ai provides information on Tier A for the grade-level cluster, and Table Aii provides information on Tier B/C for the grade-level cluster. For Speaking, Table A provides information on every task in the grade-level cluster.

### 2.2.1.1 Fit Statistics

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

Fit statistics for the Rasch model are calculated by comparing the observed empirical data with the data that would be expected to be produced by the Rasch model. Outfit mean square statistics are influenced by outliers. For example, a difficult item that some low-ability examinees get correct-for reasons unknown-will have a high outfit mean square statistic. Infit mean square statistics are influenced by unexpected patterns of observations by persons on items that are roughly targeted on them and generally indicate a more serious measurement problem. The expectation for both of these statistics is 1.00 and values near 1.00 are not of great concern. Values less than 1.00 indicate that the observations are too predictable and thus redundant, but are not of great concern. High values are of greater concern.

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

- values greater than 2.0 "distort or degrade the measurement system;"
- values between 1.5 and 2.0 are "unproductive for construction of measurement, but not degrading;"
- values between 0.5 and 1.5 should be considered "productive for measurement;" and
- values below 0.5 are "less productive for measurement, but not degrading."

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

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

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

### 2.2.1.2 Structure of Complete Item Analysis and Summary Table

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

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

The final section of Table A applies to Writing and Speaking only. This portion of the table provides raw score distributions by task.

### 2.2.2 DIF Analysis and Summary

Differential item analysis (DIF) attempts to investigate whether performances on items were influenced by factors extraneous to English language proficiency (i.e., the construct being measured on the test). In other words, DIF attempts to find items that may be functioning differently for different groups based on criteria irrelevant to what is being tested. The performance of students on ACCESS for ELLs items and tasks is compared by dividing students into two different groupings: first, males versus females; second, students of Hispanic ethnic background versus students of all other backgrounds. Students for whom gender or ethnicity ${ }^{5}$ was unknown were excluded from both analyses. Two commonly used procedures for detecting

[^4]DIF were used: one for dichotomously scored items (Listening and Reading), conducted prior to operational testing, and one for polytomously scored items (Writing and Speaking), conducted on population data subsequent to the close of operational testing.

### 2.2.2.1 Dichotomous Items

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

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

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

- A (no DIF), when the absolute value of delta is less than 1.0
- B (weak DIF), when the absolute value of delta is between 1.0 and 1.5
- $\quad$ ( strong DIF), when the absolute value of the delta is greater than 1.5


### 2.2.2.2 Polytomous Items

For polytomous items (i.e., Writing and Speaking tasks), a similar approach is used. It is based on the Mantel-Haenszel Chi-square statistic and the standardized mean difference following procedures again developed by ETS (Zwick, Donoghue, \& Grima, 1993; Allen, Carlson, \& Zalanak, 1999). The DIF procedures developed by ETS for polytomous items were used to identify tasks that exhibit DIF. JMetrik (Meyer, 2014), an open source computer program for psychometric analysis, was used in conducting the analyses. The procedures implemented in JMetrik first calculate the Cochran-Mantel-Haenszel Chi-square statistic 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 SMD compares the means of the two groups, adjusting for differences in the distribution of the groups across the values of the total raw scores. To standardize the outcome, this difference is divided by the item score range and serves as an effect size measure for the Cochran-Mantel-Haenszel Chi-square statistic. This effect size measure (reported as standardized P-DIF in JMetrik) ranges from -1 to 1 , which may present some interpretation challenges. To mitigate this, the absolute value is taken in JMetrik (Meyer, 2014), thereby restricting the range of the rescaled effect size (standardized P-DIF*) to fall between 0 and 1. The effect size flagging criterion for polytomous items, proposed by ETS (Allen, Carlson, \& Zalanak, 1999), is also rescaled to the standardized P-DIF* metric (Meyer, 2014).

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

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

Table B, DIF Analysis and Summary, provides a summary of the findings of the DIF analyses at the top, followed by detailed information for each item or task. The first column gives the DIF level: A, B, or C for dichotomous items or AA, BB, or CC for polytomous tasks (i.e., Writing and Speaking tasks). The next columns show the contrasting groups in the DIF analyses: either male versus female or Hispanic versus non-Hispanic other ethnicities. Even though DIF may be negligible (category A or AA), this table shows the number of items that favored 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.

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

### 2.2.3 Raw score distribution for Speaking and Writing

Figure C and Table C provide raw score information for Speaking and Writing only. Raw score distribution is presented by grade-level cluster and also by grade-level cluster and tier. For each test form, Figure C shows the distribution of the raw scores. The horizontal axis shows the raw scores. The vertical axis shows the number of students (count). Each bar shows how many students received each raw score.

Table C shows, by grade and by total for the grade-level cluster:

- the number of students in the analyses (the number of students who were not absent, invalid, refused, exempt, or in the wrong grade-level cluster),
- the minimum observed raw score,
- the maximum observed raw score,
- the mean (average) raw score, and
- the standard deviation (std. dev.) of the raw scores.


### 2.2.4 Scale Score Distribution

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

Thus, for each test form, Figure D shows the distribution of the scale scores. The horizontal axis shows the scale scores based on performances on the test form. To provide a 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 received each scale score.

Table D shows, by grade and by total for the grade-level cluster:

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


### 2.2.5 Proficiency Level distribution

Figure E and Table E provide information on the proficiency level distribution of the students who took the test form based on their performance. Proficiency level distribution is presented by grade-level cluster. For Writing and Speaking, it is also presented by grade-level cluster and tier. In Figure E, the horizontal axis shows the six WIDA proficiency levels. The vertical axis shows the percentage of students. Each bar shows the percentage of students who were placed into each proficiency level in the domain being tested on this test form.

Each row of Table E shows, by grade and by total for the grade-level 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, and
- the percentage of students, out of the total number of students taking the form who were placed into that proficiency level in the domain being tested.


### 2.2.6 Raw to Scale Score Conversion for Speaking and Writing

The next table in this section, Table F, presents the raw score to scale score conversion table for the test form for Speaking and Writing only.

The first column shows all possible raw scores. The following column(s) show the corresponding scale score for the highest grade in the grade-level cluster. The next column shows the conditional standard error of measurement (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 fell below 100. In such cases, the lower bound has been set at 100 , which has been determined to be the lowest score possible on the scale.

At the lower end of the raw score scale, scale scores are truncated where necessary so that the lowest scale score given is the scale score corresponding to a proficiency level score of 1.0. The standard error and the lower and upper bounds reported in Table F reflect the truncated score.

### 2.2.7 Equating Summary

Each year a certain percentage of items on each ACCESS for ELLs test form is refreshed, as determined by the refreshment plan for that Series. For Series 401, Reading, Writing, and Speaking were refreshed while Listening domain was not refreshed. An equating procedure known as common-item equating is used to equate the results on new forms to the older forms.

In this procedure, the difficulty measures for items that appear on both the new and the old forms are kept constant across both forms. Thus, performances on the newer form may be interpreted with the same frame of reference. Many items appearing on ACCESS 2.0 Online Series 401 also appeared on Series 400 . All items common to both forms were anchored to their 400 values in the first equating run. In addition, for the Speaking domain, difficulty measures for the new tasks were anchored to their initial calibrated values from the Speaking field test analysis. After the first equating run, some items that were originally anchored proved to have changed in their difficulty measure. This change is measured by the "Displacement" statistic. This statistic shows the difference between the difficulty value of the anchored item and what its difficulty value would have been had it not been anchored. Typically, random displacements of less than 0.5 logits are unlikely to have much impact on measurement in a test instrument (Linacre, n.d.). For Listening and Reading items, and for Writing tasks, if this value was large (i.e., usually above .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). For Speaking tasks, a slightly different displacement criterion (above .50 or below -.50) was used since anchored tasks from the Speaking domain have been shown to be less stable than items and tasks from other domains.

A pre equating design was used to conduct the annual equating for Listening and Reading domains. This design allows for Listening and Reading item parameters to be available for setting up the computer adaptive engine prior to operational administration. For the Listening domain, although the test was not refreshed in Series 401, student data from the Series 400 operational administration were used to refine the parameters. For the Reading domain, student data collected from the Series 401 Reading embedded field test were used to conduct the equating analyses. All available student data at the time the equating analyses were conducted were included in the analyses.
For the Writing domain, the equating analysis was conducted using Series 401 operational data collected during the early testing window. The Writing equating study was conducting a random sample drawn with a target sample size of 3,000 students per grade-level cluster. The Writing equating sample was drawn so that it was proportional to the Series 400 population for the Writing domain, by grade and tiered form.

For the Speaking domain, student data from the Series 401 appended Speaking field test administration were used to conduct the initial common-item equating. These initial item parameters were then verified using Series 401 operational data collected during the early testing window. The Speaking verification study was conducted using a random sample drawn with a target sample size of 3,000 students per grade-level cluster. The Speaking verification sample was drawn such that it was proportional to the Series 400 population for the Speaking domain, by grade and tier.

Table G presents a summary of the equating and verification procedures. The first section of the table compares the current test (i.e., the Series 401 version of that test form) to the previous year's test (i.e., the Series 400 version of that test form). The number of items, the average item
difficulty, the standard deviation of the item difficulty values, and the difficulty value of the easiest and hardest item on each test form are shown. These values are in terms of logits used in the Rasch measurement model.

The second section of the table presents information on the anchoring items. The total number of possible anchors (i.e., all common items) is shown, as well as the standard deviation of those items. For the Speaking domain, all item parameters were initially anchored to the field test measures. Next, the number of items that were actually anchored in the final equating run is shown, again with the average item difficulty and standard deviation. Finally, the percentage of items that served as anchors and 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 third section of Table G gives information about the anchor items or tasks, both by order of displacement statistics and by order of item difficulty. The displacement statistics provide information on the difference between the difficulty value of the anchored item and what that difficulty value would have been had the item not been anchored. Smaller displacement statistics indicate more consistency between the item's difficulty value on the Series 401 test form and on the Series 400 test form. It is desirable that the anchor items represent a wide range of difficulties across the entire spectrum of the item difficulty values on a test form.

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

For the Listening and Reading domains, the Series 400 parameters were derived from the scale maintenance studies in preparation for ACCESS transitioning from ACCESS 1.0 to 2.0 (see Center for Applied Linguistics, 2016). The scale maintenance studies used data from stand-alone field tests with small samples and the field tests were administered under a slightly different condition; therefore, there was a plan to refine these parameters using Series 400 operational data. However, due to technical issues related to interrupted test sessions which occurred during the Series 400 operational administration year, WIDA ultimately made the decision to retain the field test parameters in order to meet the reporting timeline (see the Annual Technical Report for Series 400 Online ACCESS for ELLs for further detail). Thus, it is not surprising that many of the repeated Listening and Reading items that were initially anchored needed to be re-estimated using the Series 401 Reading embedded field test data. In other words, had the refined parameters been used in the Series 400 Listening and Reading reporting, the final number of items being anchored in the Series 401 equating analyses would have been greater.

Note that, for the Writing and Speaking tasks, this table has a fourth section, which provides the anchored step measures for the score on each task. For the ACCESS Writing and Speaking tasks, a Rasch-grouped rating scale model is used (see Part I Section 3.3.2.2.). For Writing, the step difficulties values are the same for all the tasks that are scored on the $0-9$ raw score scale. These constant step difficulty values help to provide anchors in the calibration of new Writing tasks onto the common WIDA score scale each year. For Speaking, the step difficulty values for all P1 tasks are the same and the step difficulty values for all P3 and P5 tasks are the same. As with Writing, these constant step difficulty values help to provide anchors in the calibration of new Speaking tasks onto the common WIDA score scale each year.

### 2.2.8 Test Characteristic Curve

Test characteristic curves graphically show the relationship between the ability measure (in logits) on the horizontal axis and the expected raw score on the vertical axis. Five vertical lines indicate the five cut scores for the highest grade in the cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (Levels 1-6) for the domain being tested. (Note that for some domains for 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.

As the Listening and Reading assessments are multistage adaptive tests, raw scores are not a meaningful aspect of these tests, so no test characteristic curve is presented for these domains.

### 2.2.9 Test Information Function

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

Five vertical lines in Figure I indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1-6) for the domain being tested. The ACCESS cut scores lines are presented
along with the test information function to facilitate the interpretation of the test information curves. The test information curve and the corresponding ACCESS cut score lines are both expressed on the ACCESS logit scale. Note that for Speaking, in Tier pre-A, all scores fall in the PL 1.0 range, so there are no vertical lines expressing the cuts between proficiency levels.

### 2.2.10 Reliability of Domain Scores

### 2.2.10.1 Listening and Reading Domains

In the Listening and Reading domains, Table J presents reliability information based on Item Response Theory. The table shows:

- the number of students (count),
- the number of items,
- Rasch Reliability (as a measure of internal consistency)

For tests administered using a multistage adaptive method, a reliability coefficient based on classical test theory such as Cronbach's coefficient alpha cannot be applied because not all students take the same set of items. Reliability for Listening and Reading was estimated using a method by Thissen (2000) by grade-level cluster:
$\left.\bar{\rho}=\frac{\sigma_{\theta-\text { average }(\text { CSEM }}^{\text {observed }}}{2}\right)$
where
$\bar{\rho}$ is the average reliability,
$\sigma_{\theta}^{2}$ is the variance of the distribution of student measure,
CSEM ${ }_{\text {observed }}^{2}$ is the squared observed conditional standard errors of measurement for each student

This estimate is equivalent to the Rasch separation reliability coefficient (Linacre, 1999). Like Cronbach's alpha, the Rasch reliability coefficient is an estimate of the ratio of "true measure variance" to "observed measure variance." To obtain these values, item parameters and population student data were used as inputs in the Winsteps program. The Rasch separation reliability coefficient can be interpreted like Cronbach's coefficient alpha. It expresses how well the items on a test appear to measure the same construct.

### 2.2.10.2 Speaking and Writing Domains

In the Speaking and Writing domains, Table J presents reliability and accuracy information based on classical test theory. Table $\mathbf{J}$ is provided for each tier, and it is also provided, in a different format, to express weighted reliability for each grade-level cluster.

For each tier, the table shows:

- the number of students (count),
- the number of tasks,
- for Writing, the response mode (keyboarded or handwritten)
- Cronbach's coefficient alpha (as a measure of internal consistency), and
- 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
examinees (i.e., the more widely the scores are distributed), the higher the reliability. In this sense, Cronbach's alpha is sample dependent. It is widely recognized that reliability can be as much a function of the test as of the sample of students tested. That is, the exact same test can produce widely disparate reliability indices based on the ability distribution of the group of examinees.

The formula for Cronbach's alpha is
$\alpha=\frac{n}{n-1}\left[1-\frac{\sum_{i=1}^{n} \sigma_{i}^{2}}{\sigma_{t}^{2}}\right]$
where
$n=$ number of items $i$
$\sigma_{i}{ }^{2}=$ variance of score on item $i$
$\sigma_{t}^{2}=$ variance of total score

For the Writing test, a slight modification was made in the estimation of the Cronbach's alpha for tiered forms that have differential weighting across tasks. This modification is an attempt to take into account that some tasks are weighted more than others when deriving student's ability measure for these tiered forms. For writing tasks with weight greater than one, student's response to the tasks are replicated as a function of their weights. For example, the fourth task is weighted three in Writing G1A, therefore, student's response to this task was repeated three times when computing the Cronbach's alpha. This modification means that the number of pieces of information or Writing tasks that contribute to the estimation of the Cronbach's alpha for G1A is actually six, not four.

Table J also presents the SEM based on classical test theory for Speaking and Writing. Unlike Item Response Theory, in this approach, SEM is seen as a constant across the spread of test scores (ability continuum). Thus, it is not conditional on ability being measured. It is, however, a function of two statistics: the reliability of the test and the (observed) standard deviation (SD) of the test scores. It is calculated as
$\mathrm{SEM}=S D \sqrt{1-\text { reliability }}$
Traditionally, SEM has been used to create a band around an examinee's observed score, with the assertion in the view of classical test theory, that the examinee's true score (i.e., what the examinee's score would be if it could be measured without error) would lie with a certain degree of probability within this band. Statistically speaking then, there is an expectation that an examinee's true score has a $68 \%$ probability of falling within the band extending from the observed score minus 1 SEM to the observed score plus 1 SEM.

For the Writing and Speaking tests, information on interrater reliability for a sample of $20 \%$ of task raters is also provided in Table J. This portion of the table shows, for each of the tasks, the percent of agreement between two raters. In this part of the table, the first column shows the task and the second column shows the number of responses that were double scored. DRC selects a sample of $20 \%$ of all responses scored, chosen at random during the operational scoring process. The next column shows the rates of agreement: exact, adjacent, and non-adjacent. For Speaking, when the two raters agreed on the rating, an exact agreement was counted. If the two raters were different by one point, an adjacent agreement was counted. Otherwise, the raters are nonadjacent. For Writing, with 0-6 as defined levels and the possibility of awarding a "plus" score between levels (e.g., 3, 3+, or 4 are all valid scores), scores that match or are contiguous are categorized as agreement (for example, if Rater 1 assigns a $3+$, then a score of $3,3+$ or 4 from Rater 2 is categorized as agreement). Scores that are one whole score point apart are categorized as adjacent (for example, if Rater 1 assigns a $3+$, then a score of $2+$ or $4+$ from Rater 2 is categorized as adjacent). Note that for Writing, interrater reliability is computed independently between ratings of keyboarded and handwritten responses.

For each grade-level cluster in Writing and Speaking, Table J is presents a single reliability value for the grade-level cluster. To produce this single value, values for Cronbach's alpha for each of the tiers in the grade-level cluster are weighted by the number of students who were administered the tier form, and a weighted average is expressed in Table J.

### 2.2.11 Conditional Standard Errors of Measurement at Cut Score

Table $K$ presents information on the conditional standard errors of measurement (CSEM) at the most important points at which decisions are made about students based on performance on ACCESS-the cut points between language proficiency levels. Because the cut points depend on the grade level, information is provided for each grade level within a grade-level cluster. The leftmost column shows the cut (e.g., $1 / 2$, which is the cut score between Proficiency Level 1 and Proficiency Level 2).

The second column shows the grade level. The third column shows the cut score in the scale score metric (e.g., 305). In the last column(s), the corresponding CSEM is given for each cut score in the scale score metric for Writing and Speaking.

For Writing and Speaking, the values are presented by tier. From Table K, it is possible to identify how well the different Writing and Speaking 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 CSEM of any tier at the $1 / 2$ cut point, and a relatively low CSEM at the $2 / 3$ cut point. At the other end of the continuum, Tier B/C forms should optimally have the lowest CSEM at the $5 / 6$ cut point, and a relatively low CSEM at the $4 / 5$ cut point. Information from Table K provides comparable information on how well the two tier forms are targeted to provide the most accurate
measure in order to place their intended examinees into the language proficiency levels that they target.

Since the Listening and Reading tests are multistage adaptive tests, the CSEM will vary for the same scale score since students were routed to take different items; the mean, standard deviation, minimum, and maximum of the CSEM of all students at the cut scores are presented instead. Note that there are some rare cases where there are no observed scale scores corresponding to the cut score values, therefore these descriptive statistics cannot be provided.

### 2.2.12 Accuracy and consistency

Table L presents three sections of information related to the accuracy and consistency of placement into the WIDA language proficiency levels for each domain. A separate table is provided for each grade in a grade-level cluster. The first section provides overall indices related to the accuracy and consistency of classification, as well as Cohen's kappa. The second section of information shows accuracy and consistency information conditional on proficiency level. The third provides indices of classification accuracy, including the false positives and the false negatives, and consistency at the cut points. These indices are perhaps the most important of all when using any of these as an absolute cut-point (e.g., determining which students have reached Proficiency Level 5). Note that the consistency is generally higher at the cut points than over the proficiency levels.

There are several cases where there were no test takers who were placed into the proficiency level and accuracy of classification conditional on that level cannot be computed. In these cases, 'N/A' has been placed in the table. In addition, there are a few cases where due to the small percentage of test takers placed into the proficiency level and the range of observed scale scores, accuracy of classification conditional on that level cannot be estimated by the software program that is used (BB-CLASS, see below). In such cases, a hyphen (-) has been placed in the table.

For each domain, tables are provided that indicate estimates of the accuracy and consistency of classification of examinees into the WIDA language proficiency levels based on their performance 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 errors of the Rasch measurement model). However, because decisions about students are ultimately made on the basis of their classification into language proficiency levels according to their performance on ACCESS, it is important to know how well these classifications are made. The analyses that were used utilize the methods outlined and implemented in Livingston and Lewis (1995) and Young and Yoon (1998) as implemented in the software program BBCLASS (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 those made if each student could somehow be tested with all possible parallel forms of the
assessments; that is, the examinee's "true score." Meanwhile, the consistency of a decision is the extent to which decisions made on the basis of the administered test would agree with those made if each student were to take a different but parallel form of the test. Thus, in every analysis of classification, two parallel analyses are made: accuracy (vis-à-vis "true scores") and consistency (vis-à-vis a parallel test).

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 based on the observed score as being above the cut score, are considered to be false positives. Students who were above the proficiency level cut score (based on their "true score"), but were classified as being below a cut score based on the observed score, are considered to be false negatives. All other students are considered to be accurately placed either above or below the cut score.

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, the distribution of the true scores and of scores on a parallel form were modeled. 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 percentage 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). (These 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).

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

Finally, the most important set of indices may be the indices at the cut points. At every cut point, using the true score distribution (i.e., accuracy), the percentage of students who are consistently placed above and below the cut score is provided, as well as those who are false positives and false negatives. For consistency, only the percentage 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 or not students are being accurately classified into Proficiency Level 5 ("Bridging"), one can look at the accuracy index provided in Table $L$ for the cut score 4/5.

The Livingston and Lewis procedure requires that the reliability estimate of the test form be provided in estimating the classification consistency and accuracy statistics. For Listening and Reading, the Rasch reliability estimates by grade-level clusters were used in the procedure. Since the Writing and Speaking tests were tiered, it was necessary to produce a single reliability estimate across tiers for the Livingston and Lewis procedure. This is a weighted reliability
estimate across tiers. In other words, it is the average reliability weighted by the number of students who were administered that tier form. Thus, Table L, based on the information from Table J, 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.

### 2.3 Analyses of Composite Scores

In Section 3.4, analyses of the four composites-Oral Language, Literacy, Comprehension, and Overall Composite-are presented. Tables and figures pertaining to the composite scores are presented by grade-level cluster.

### 2.3.1 Scale Score distribution for Composites

Figure A and Table A provide scale score distributions for each of the composites, for each grade-level cluster.

Figure A 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 received each scale score.

Table A shows, by grade and by total for the grade-level cluster:

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


### 2.3.2 Proficiency Level distribution for Composites

Figure B and Table B provide information on the proficiency level distribution for each of the composites for each grade-level cluster.

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

Each row of Table B shows, by grade and by total for the grade-level 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, and
- the percentage of students, out of the total number of students taking the form who were placed into that proficiency level in the domain being tested.


### 2.3.3 Reliability of Composites

To estimate the reliability of the composite scores, a stratified Cronbach's alpha coefficient (e.g., Rudner, 2001; Kamata, Turhan, \& Darandari, 2003; Kane \& Case, 2004;) is computed, weighted by the contribution of each domain score into the composite. Specifically, the formula is
$\alpha_{c}=1-\frac{\sum_{j=1}^{k} w_{j}^{2} \sigma_{\sigma}^{2}\left(1-\rho_{j}\right)}{\sigma_{c}^{2}}$
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 used to compute the stratified Cronbach's alpha is provided in Table C. The first column shows the components forming the composite, the second column shows the weight of the composite in the total score, the third shows the variance of the scale scores, and the fourth shows the reliability of the domains forming the composite (note that these are the weighted reliabilities across the tiers for Speaking and Writing) and the reliability of the composite. Unlike the weighted composite, which is an average, the stratified alpha reflects the fact that there are two to four measures being combined into one single measure. Thus, the reliability of the composite score will be higher than the reliability of any single sub-score within the composite.

The stratified Cronbach's alpha, presented in Table C, was also used to produce the Accuracy and Consistency classification tables of the composites (Table D).

### 2.3.4 Accuracy and Consistency of Composites

Table D presents three sections of information related to the accuracy and consistency of placement into the WIDA language proficiency levels for each composite score. The first section provides overall indices related to the accuracy and consistency of classification, as well as Cohen's kappa. The second section shows accuracy and consistency information conditional per proficiency level. The third section provides indices of classification accuracy, including the false positives and false negatives, and consistency at the cut points. These indices are perhaps the most important of all when using any of these as an absolute cut-point (e.g., determining which students have reached Proficiency Level 5). Note that the consistency is generally higher
at the cut points than over the proficiency levels. For practical purposes, the primary score used for such decisions is the Overall Composite score.

As noted above in 2.2.12, there may be cases where there are no test takers placed into the proficiency level and accuracy of classification conditional on that level cannot be computed. In this case 'N/A' has been placed in the table. In addition, there may be cases where due to the small percentage of test takers placed into the proficiency level and the range of observed scale scores, accuracy of classification conditional on that level cannot be estimated by the software program that is used. In such cases, a hyphen (-) has been placed in the table.

## 3 Results By Grade Cluster

### 3.1 Guide to Tables and Figures

The remainder of the subsections of this report (3.2, 3.3, and 3.4) present tables and figures describing, respectively, students' participation and performance, analyses of the scores in the four language domains (Listening, Reading, Writing, and Speaking), and analyses of the scores in the four composites (Oral Language, Literacy, Comprehension, and Overall).

For ease of navigation through these subsequent sections, this section provides a visual overview of the numbered tables and figures. For readers who are reviewing this report in an electronic format, section headers are built into the document structure to assist the reader to navigate through the document.

### 3.1.1 Guide to 3.2, Student Participation and Performance

Tables 3.1A-C provide a visual overview of the tables included in Section 3.2. There are three subsections:
3.2.1 Participation presents distributions of students' participation by grade and by grade-level cluster. Student participation by grade and grade-level cluster is further broken down by state, by gender, by ethnicity, and finally by tier and domain combined. Table 3.1A presents the tables included in this subsection.

Table 3.1A
Table Numbering System for Section 3.2.1, Participation

|  | 3.2.1.1. By Grade-level Cluster | 3.2.1.2. By Grade |
| :--- | :--- | :--- |
| By State | Table 3.2.1.1.1 | Table 3.2.1.2.1 |
| By Gender | Table 3.2.1.1.2 | Table 3.2.1.2.2 |
| By Ethnicity | Table 3.2.1.1.3 | Table 3.2.1.2.3 |
| By Tier by Domain | Table 3.2.1.1.4 | Table 3.2.1.2.4 |

3.2.2 Scale Score Results presents distributions of students' scale score results. These are again presented by grade and grade-level cluster. Student scale score results by grade and grade-level cluster are further broken down by gender and by ethnicity, and correlations among scale score results are presented. Table 3.1B presents the section and table numbering system for this section.

Table 3.1B
Section and Table Numbering System for Section 3.2.2, Scale Score Results

| Mean Scale Scores Across Domain and Composite |  |  |
| :--- | :--- | :--- |
|  | Section 3.2.2.1. <br> By Grade-level Cluster | Section 3.2.2.2. <br> By Grade |
| Alone | Table 3.2.2.1.1 | Table 3.2.2.2.1 |
| And by Gender | Table 3.2.2.1.2 | Table 3.2.2.2.2 |
| And by Ethnicity | Table 3.2.2.1.3 | Table 3.2.2.2.3 |

Section 3.2.2.3 Correlations Among Scale Scores by Grade-level Cluster
3.2.3 Proficiency Level Results presents distributions of students' proficiency level results for the four domains and four composites, by grade and by grade-level cluster. Table 3.1C lists the numbers of subsections. Each subsection contains a table expressing descriptive statistics as counts (Table A) and as percentages (Table B).

Table 3.1C
Section Numbering System for Section 3.2.3, Proficiency Level Results

|  |  | By Grade-Level Cluster | By Grade |
| :--- | :--- | :--- | :--- |
|  | For each, distributions by count and by percent |  |  |
| 3.2.3.1 Domains |  |  |  |
| 3.2 .3 .1 .1 | Listening | 3.2 .3 .1 .1 .1 | 3.2 .3 .1 .1 .2 |
| 3.2 .3 .1 .2 | Reading | 3.2 .3 .1 .2 .1 | 3.2 .3 .1 .2 .2 |
| 3.2 .3 .1 .3 | Writing | 3.2 .3 .1 .3 .1 | 3.2 .3 .1 .3 .2 |
| 3.2 .3 .1 .4 | Speaking | 3.2 .3 .1 .4 .1 | 3.2 .3 .1 .4 .2 |
| 3.2.3.2 Composites |  | 3.2 .3 .2 .1 .2 |  |
| 3.2 .3 .2 .1 | Oral Composite | 3.2 .3 .2 .2 .2 .1 | 3.2 .3 .2 .3 .2 |
| 3.2 .3 .2 .2 | Literacy Composite | 3.2 .3 .2 .3 .1 | 3.2 .3 .2 .4 .2 |
| 3.2 .3 .2 .3 | Comprehension Composite | 3.2 .3 .2 .4 .1 |  |
| 3.2 .3 .2 .4 | Overall Composite |  |  |

### 3.1.2 Guide to 3.3, Analysis of Domain Scores

An overview of the tables and figures in Section 3.3 Analysis of Domain Scores is provided in Table 3.1D. This section is organized by grade-level cluster, and the figure provides an overview of the detail in any given grade-level cluster. Note that the headers within the figure include an "X" to denote the grade-level cluster-for example, "Reading 3.3.X.2" would read "Reading 3.3.1.2" for grade 1 , "Reading 3.3.2.2." for grades $2-3$, and so on.

Table 3.1D
Naming conventions for tables and figures in Section 3.3. Analysis of Domain Scores

|  | Table | Figure | 3.3.X.1. <br> Listening | 3.3.X.2. <br> Reading | 3.3.X.3. <br> Writing |  | 3.3.X.4. <br> Speaking |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | By tier | Across tiers | By tier | Across tiers |
| Complete Item/Task Analys is and Summary | Table A |  | V | V | V | n/a | n/a | V |
| DIF Analys is and Summary | Table B |  | V | V | V | n/a | V | n/a |
| Raw Score Distribution (Spek and Writ only) | Table C | Figure C | n/a | n/a | V | V | $\checkmark$ | $\checkmark$ |
| Scale Score Distribution | Table D | Figure D | V | $\checkmark$ | V | V | V | $\checkmark$ |
| Proficiency Level Distribution | Table E | Figure E | V | V | V | V | $\checkmark$ | $\checkmark$ |
| Raw Score to Scale Score Conversion with CSEM (Spek and Writ only) | Table F |  | n/a | n/a | V | n/a | V | n/a |
| Equating Summary | Table G |  | $\checkmark$ | V | $\checkmark$ | n/a | n/a | $\checkmark$ |
| Test Characteristic Curve |  | Figure H | n/a | n/a | V | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Test Information Function |  | Figure I | V | V | V | V | V | V |
| Reliability | Table J |  | V | V | V | V | V | V |
| CSEM at Cut Score Points | Table K |  | V | $\checkmark$ | n/a | V | n/a | $\checkmark$ |
| Accuracy and Consistency (by grade) | Table L |  | V | V | n/a | V | n/a | V |

Note: By tier means that a table is presented for each tier of the grade-level cluster. Across tiers means that one table is presented for the grade-level cluster, including information from all tiers.

The left column of Table 3.1D lists the content of the table of figure. The next two columns list the letter designations for the tables or figures as applicable. Check marks under a domain subheading indicate that the table and/or figure is included in the report or not applicable ( $\mathrm{n} / \mathrm{a}$ ).

For the adaptive domains (Listening and Reading), tables or figures related to raw scores are not provided.

For the tiered domains (Writing and Speaking), differing subsets of tables are provided either by tier or across tiers in a grade-level cluster.

If a table or figure is provided multiple times with the same grade-level cluster and domain, it is denoted with a roman numeral-e.g. Table 3.3.1.4.Di provides scale score distribution information for Speaking Grade 1 pre-A; Table 3.3.1.4.Dii provides the same information for Speaking Grade 1 Tier A, and so on. For Writing, the two tables describing the individual Writing tasks (Table A, Complete Task Analysis and Summary, and Table B, DIF Analysis and Summary) are provided once for Tier A and once for Tier B/C. For Speaking, due to the design of the tiered assessment, Table A, Complete Task Analysis and Summary, is provided once across the three tiers. Table B, DIF Analysis and Summary, is provided separately for each tier. For both Writing and Speaking,. Figures and Tables C-E (Raw Score Distribution, Scale Score Distribution, and Proficiency Level Distribution) are provided first for each tier and then for the entire grade-level cluster. Table F, Raw Score to Scale Score Conversion with SEM is provided by Tier only. Table G, Equating Summary, is provided by tier for Writing and across tiers for Speaking. Figures H and I (Test Characteristic Curve; Test Information Function), and Table J (Reliability) are provided by each tier and also for the entire grade-level cluster. Finally, Tables K and L (CSEM at Cut Score Points, Accuracy and Consistency) are provided for the grade-level cluster.

### 3.1.3 Guide to 3.4, Analysis of Composite Scores

As with Section 3.3, Section 3.4. is first organized by grade-level cluster, and then by each of the four composites (Oral Language, Literacy, Comprehension, Overall). For each grade-level cluster/composite combination (e.g. Grade 4-5, Comprehension), the figures and tables presented in Table 3.1E below are provided.

Table 3.1E
Naming conventions for tables and figures in Section 3.4. Analysis of Composite Scores

| Scale Score Distribution | Figure A | Table A |
| :--- | :--- | :---: |
| Proficiency Level Distribution | Figure B | Table B |
| Reliability |  | Table C |
| Accuracy and Consistency |  | Table D |

### 3.2 Student Participation and Performance

### 3.2.1 Participation

### 3.2.1.1 Participation by Grade-level Cluster

### 3.2.1.1.1 By State

Table 3.2.1.1.1
Participation by Cluster by State S 401 Online

| State | Cluster |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2-3 | 4-5 | 6-8 | 9-12 |  |
| AK | 1,015 | 2,118 | 1,401 | 1,604 | 1,599 | 7,737 |
| AL | 2,138 | 3,747 | 1,799 | 1,847 | 2,118 | 11,649 |
| CO | 7,819 | 16,357 | 11,831 | 13,238 | 12,523 | 61,768 |
| DC | 897 | 1,495 | 796 | 1,003 | 1,419 | 5,610 |
| DE | 1,920 | 3,517 | 1,439 | 1,167 | 1,452 | 9,495 |
| GA | 12,795 | 23,783 | 12,432 | 12,924 | 12,058 | 73,992 |
| HI | 1,245 | 2,134 | 1,070 | 1,814 | 1,544 | 7,807 |
| ID | 2,290 | 3,428 | 2,596 | 2,661 | 2,392 | 13,367 |
| IL | 15,958 | 45,157 | 25,092 | 23,025 | 20,902 | 130,134 |
| IN | 7,016 | 12,765 | 5,748 | 7,620 | 8,821 | 41,970 |
| KY | 3,321 | 5,680 | 3,035 | 3,207 | 4,147 | 19,390 |
| MA | 5,734 | 11,696 | 8,582 | 9,493 | 10,769 | 46,274 |
| MD | 9,779 | 17,899 | 9,007 | 9,510 | 13,820 | 60,015 |
| ME | 458 | 1,022 | 892 | 1,066 | 1,273 | 4,711 |
| MI | 9,948 | 20,976 | 15,429 | 18,828 | 19,343 | 84,524 |
| MN | 7,948 | 16,601 | 11,001 | 12,108 | 12,248 | 59,906 |
| MO | 4,075 | 7,841 | 4,905 | 4,885 | 4,416 | 26,122 |
| MP | 123 | 240 | 272 | 443 | 224 | 1,302 |
| MT | 242 | 781 | 570 | 648 | 340 | 2,581 |
| NC | 12,082 | 26,978 | 11,127 | 13,644 | 15,637 | 79,468 |
| ND | 370 | 651 | 408 | 581 | 715 | 2,725 |
| NH | 477 | 978 | 615 | 659 | 865 | 3,594 |
| NJ | 10,563 | 16,209 | 8,538 | 10,053 | 14,703 | 60,066 |
| NM | 4,751 | 10,358 | 7,270 | 8,287 | 7,583 | 38,249 |
| NV | 7,834 | 18,000 | 12,208 | 14,595 | 11,743 | 64,380 |
| OK | 3,840 | 7,368 | 3,780 | 4,586 | 4,856 | 24,430 |
| PA | 4,176 | 8,565 | 6,839 | 9,208 | 12,286 | 41,074 |
| RI | 1,051 | 2,223 | 1,354 | 1,685 | 2,431 | 8,744 |
| S C | 3,825 | 8,671 | 7,823 | 10,425 | 8,630 | 39,374 |
| SD | 561 | 1,120 | 491 | 628 | 853 | 3,653 |
| TN | 5,448 | 12,164 | 6,843 | 7,468 | 6,949 | 38,872 |
| UT | 5,309 | 10,905 | 6,583 | 6,443 | 5,705 | 34,945 |
| VA | 9,626 | 22,681 | 12,510 | 13,565 | 18,465 | 76,847 |
| VI | 132 | 231 | 192 | 245 | 223 | 1,023 |
| VT | 197 | 390 | 199 | 208 | 301 | 1,295 |
| WI | 5,668 | 11,841 | 8,750 | 8,410 | 7,411 | 42,080 |
| WY | 350 | 712 | 314 | 377 | 433 | 2,186 |
| Total | 170,981 | 357,282 | 213,741 | 238,158 | 251,197 | 1,231,359 |

### 3.2.1.1.2 By Gender

Table 3.2.1.1.2
Participation by Cluster by Gender S401 Online

| Cluster |  | Gender |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | F | M | Missing |  |
| 1 | Count | 79,220 | 88,591 | 3,170 | 170,981 |
|  | \% within Cluster | 46.3\% | 51.8\% | 1.9\% | 100.0\% |
| 2-3 | Count | 165,166 | 186,343 | 5,773 | 357,282 |
|  | \% within Cluster | 46.2\% | 52.2\% | 1.6\% | 100.0\% |
| 4-5 | Count | 92,320 | 118,209 | 3,212 | 213,741 |
|  | \% within Cluster | 43.2\% | 55.3\% | 1.5\% | 100.0\% |
| 6-8 | Count | 102,951 | 131,834 | 3,373 | 238,158 |
|  | \% within Cluster | 43.2\% | 55.4\% | 1.4\% | 100.0\% |
| 9-12 | Count | 107,652 | 138,931 | 4,614 | 251,197 |
|  | \% within Cluster | 42.9\% | 55.3\% | 1.8\% | 100.0\% |
| Total | Count | 547,309 | 663,908 | 20,142 | 1,231,359 |
|  | \% within Cluster | 44.4\% | 53.9\% | 1.6\% | 100.0\% |

### 3.2.1.1.3 By Ethnicity

Table 3.2.1.1.3
Participation by Cluster by Ethnicity S 401 Online

| Cluster |  | Hispanic/Non-Hispanic |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Hispanic | Other | Unknown |  |
| 1 | Count | 110,766 | 52,683 | 7,532 | 170,981 |
|  | \% within Cluster | $64.8 \%$ | $30.8 \%$ | $4.4 \%$ | $100.0 \%$ |
| $2 \mathbf{2}-3$ | Count | 240,919 | 102,640 | 13,723 | 357,282 |
|  | $\%$ within Cluster | $67.4 \%$ | $28.7 \%$ | $3.8 \%$ | $100.0 \%$ |
| $4-5$ | Count | 145,104 | 59,671 | 8,966 | 213,741 |
|  | \% within Cluster | $67.9 \%$ | $27.9 \%$ | $4.2 \%$ | $100.0 \%$ |
| $6-8$ | Count | 158,434 | 69,241 | 10,483 | 238,158 |
|  | \% within Cluster | $66.5 \%$ | $29.1 \%$ | $4.4 \%$ | $100.0 \%$ |
| $9-12$ | Count | 161,945 | 76,987 | 12,265 | 251,197 |
|  | \% within Cluster | $64.5 \%$ | $30.6 \%$ | $4.9 \%$ | $100.0 \%$ |
| Total | Count | 817,168 | 361,222 | 52,969 | $1,231,359$ |
|  | \% within Cluster | $66.4 \%$ | $29.3 \%$ | $4.3 \%$ | $100.0 \%$ |

### 3.2.1.1.4 By Tier by Domain

Table 3.2.1.1.4
Participation by Cluster by Tier by Domain S401 Online

| Cluster |  |  | Domain |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Writing | Speaking |
| 1 | Tier | Pre-A | - | 5,726 |
|  |  | A | 133,261 | 57,484 |
|  |  | BC | 37,698 | 107,764 |
|  | Total |  | 170,959 | 170,974 |
| 2-3 | Tier | Pre-A | - | 17,191 |
|  |  | A | 108,599 | 89,259 |
|  |  | BC | 248,613 | 250,817 |
|  | Total |  | 357,212 | 357,267 |
| 4-5 | Tier | Pre-A | - | 6,315 |
|  |  | A | 45,598 | 29,848 |
|  |  | BC | 168,132 | 177,570 |
|  | Total |  | 213,730 | 213,733 |
| 6-8 | Tier | Pre-A | - | 13,101 |
|  |  | A | 96,875 | 55,371 |
|  |  | BC | 141,270 | 169,669 |
|  | Total |  | 238,145 | 238,141 |
| 9-12 | Tier | Pre-A | - | 22,978 |
|  |  | A | 105,343 | 112,487 |
|  |  | BC | 145,823 | 115,703 |
|  | Total |  | 251,166 | 251,168 |

### 3.2.1.2 Participation by Grade

### 3.2.1.2.1 By State

Table 3.2.1.2.1
Participation by Grade by State S401 Online

| State | Grade |  |  |  |  |  |  |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |  |
| AK | 1,015 | 1,034 | 1,084 | 791 | 610 | 536 | 536 | 532 | 544 | 433 | 330 | 292 | 7,737 |
| AL | 2,138 | 1,986 | 1,761 | 1,143 | 656 | 586 | 661 | 600 | 817 | 585 | 449 | 267 | 11,649 |
| CO | 7,819 | 8,276 | 8,081 | 6,431 | 5,400 | 4,432 | 4,459 | 4,347 | 4,632 | 3,289 | 2,585 | 2,017 | 61,768 |
| DC | 897 | 764 | 731 | 488 | 308 | 318 | 324 | 361 | 679 | 346 | 240 | 154 | 5,610 |
| DE | 1,920 | 1,824 | 1,693 | 943 | 496 | 424 | 378 | 365 | 661 | 400 | 229 | 162 | 9,495 |
| GA | 12,795 | 12,205 | 11,578 | 7,506 | 4,926 | 4,345 | 4,451 | 4,128 | 6,036 | 3,253 | 1,675 | 1,094 | 73,992 |
| HI | 1,245 | 985 | 1,149 | 599 | 471 | 600 | 619 | 595 | 652 | 376 | 300 | 216 | 7,807 |
| ID | 2,290 | 1,723 | 1,705 | 1,428 | 1,168 | 1,126 | 810 | 725 | 781 | 664 | 523 | 424 | 13,367 |
| IL | 15,958 | 18,111 | 27,046 | 15,470 | 9,622 | 8,332 | 7,572 | 7,121 | 8,144 | 5,958 | 3,951 | 2,849 | 130,134 |
| IN | 7,016 | 6,751 | 6,014 | 3,438 | 2,310 | 2,182 | 2,703 | 2,735 | 3,374 | 2,257 | 1,888 | 1,302 | 41,970 |
| KY | 3,321 | 3,009 | 2,671 | 1,752 | 1,283 | 1,072 | 1,073 | 1,062 | 1,829 | 1,119 | 714 | 485 | 19,390 |
| MA | 5,734 | 5,941 | 5,755 | 4,686 | 3,896 | 3,100 | 3,196 | 3,197 | 4,118 | 2,771 | 2,367 | 1,513 | 46,274 |
| MD | 9,779 | 9,596 | 8,303 | 5,347 | 3,660 | 3,154 | 3,092 | 3,264 | 6,988 | 3,655 | 1,992 | 1,185 | 60,015 |
| ME | 458 | 513 | 509 | 474 | 418 | 391 | 347 | 328 | 380 | 341 | 282 | 270 | 4,711 |
| MI | 9,948 | 10,252 | 10,724 | 8,126 | 7,303 | 6,280 | 6,303 | 6,245 | 6,852 | 5,259 | 3,782 | 3,450 | 84,524 |
| MN | 7,948 | 8,162 | 8,439 | 6,203 | 4,798 | 4,001 | 4,130 | 3,977 | 4,652 | 3,233 | 2,425 | 1,938 | 59,906 |
| MO | 4,075 | 4,035 | 3,806 | 2,705 | 2,200 | 1,752 | 1,630 | 1,503 | 1,741 | 1,216 | 875 | 584 | 26,122 |
| MP | 123 | 104 | 136 | 134 | 138 | 136 | 139 | 168 | 98 | 51 | 38 | 37 | 1,302 |
| MT | 242 | 379 | 402 | 328 | 242 | 212 | 232 | 204 | 140 | 106 | 55 | 39 | 2,581 |
| NC | 12,082 | 12,969 | 14,009 | 6,630 | 4,497 | 4,018 | 4,610 | 5,016 | 7,108 | 4,270 | 2,596 | 1,663 | 79,468 |
| ND | 370 | 337 | 314 | 202 | 206 | 168 | 191 | 222 | 272 | 187 | 131 | 125 | 2,725 |
| NH | 477 | 492 | 486 | 348 | 267 | 218 | 214 | 227 | 355 | 223 | 166 | 121 | 3,594 |
| NJ | 10,563 | 9,018 | 7,191 | 4,951 | 3,587 | 3,161 | 3,352 | 3,540 | 5,109 | 4,177 | 3,297 | 2,120 | 60,066 |
| NM | 4,751 | 5,227 | 5,131 | 4,080 | 3,190 | 2,626 | 2,890 | 2,771 | 3,087 | 2,144 | 1,401 | 951 | 38,249 |
| NV | 7,834 | 8,767 | 9,233 | 6,611 | 5,597 | 4,844 | 4,821 | 4,930 | 4,539 | 3,228 | 2,425 | 1,551 | 64,380 |
| OK | 3,840 | 3,810 | 3,558 | 2,314 | 1,466 | 1,183 | 1,706 | 1,697 | 2,150 | 1,316 | 874 | 516 | 24,430 |
| PA | 4,176 | 4,372 | 4,193 | 3,500 | 3,339 | 3,055 | 3,051 | 3,102 | 3,871 | 3,365 | 2,729 | 2,321 | 41,074 |
| RI | 1,051 | 1,055 | 1,168 | 766 | 588 | 541 | 549 | 595 | 896 | 671 | 511 | 353 | 8,744 |
| SC | 3,825 | 3,789 | 4,882 | 4,332 | 3,491 | 3,581 | 3,416 | 3,428 | 3,918 | 2,284 | 1,378 | 1,050 | 39,374 |
| SD | 561 | 547 | 573 | 299 | 192 | 197 | 187 | 244 | 385 | 201 | 153 | 114 | 3,653 |
| TN | 5,448 | 5,832 | 6,332 | 4,154 | 2,689 | 2,597 | 2,631 | 2,240 | 3,061 | 1,938 | 1,218 | 732 | 38,872 |
| UT | 5,309 | 5,687 | 5,218 | 3,847 | 2,736 | 1,999 | 2,229 | 2,215 | 1,985 | 1,595 | 1,247 | 878 | 34,945 |
| VA | 9,626 | 10,369 | 12,312 | 7,531 | 4,979 | 4,230 | 4,597 | 4,738 | 7,933 | 5,080 | 4,087 | 1,365 | 76,847 |
| VI | 132 | 113 | 118 | 102 | 90 | 100 | 88 | 57 | 84 | 54 | 43 | 42 | 1,023 |
| VT | 197 | 193 | 197 | 116 | 83 | 73 | 66 | 69 | 116 | 60 | 77 | 48 | 1,295 |
| WI | 5,668 | 5,989 | 5,852 | 5,012 | 3,738 | 2,950 | 2,753 | 2,707 | 2,846 | 2,019 | 1,471 | 1,075 | 42,080 |
| WY | 350 | 368 | 344 | 185 | 129 | 132 | 122 | 123 | 175 | 118 | 73 | 67 | 2,186 |
| Total | 170,981 | 174,584 | 182,698 | 122,972 | 90,769 | 78,652 | 80,128 | 79,378 | 101,008 | 68,242 | 48,577 | 33,370 | 1,231,359 |

### 3.2.1.2.2 By Gender

Table 3.2.1.2.2
Participation by Grade by Gender S401 Online

| Grade |  | Gender |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | F | M | Missing |  |
| 1 | Count | 79,220 | 88,591 | 3,170 | 170,981 |
|  | \% within Grade | 46.3\% | 51.8\% | 1.9\% | 100.0\% |
| 2 | Count | 81,059 | 90,528 | 2,997 | 174,584 |
|  | \% within Grade | 46.4\% | 51.9\% | 1.7\% | 100.0\% |
| 3 | Count | 84,107 | 95,815 | 2,776 | 182,698 |
|  | \% within Grade | 46.0\% | 52.4\% | 1.5\% | 100.0\% |
| 4 | Count | 53,106 | 67,887 | 1,979 | 122,972 |
|  | \% within Grade | 43.2\% | 55.2\% | 1.6\% | 100.0\% |
| 5 | Count | 39,214 | 50,322 | 1,233 | 90,769 |
|  | \% within Grade | 43.2\% | 55.4\% | 1.4\% | 100.0\% |
| 6 | Count | 33,860 | 43,678 | 1,114 | 78,652 |
|  | \% within Grade | 43.1\% | 55.5\% | 1.4\% | 100.0\% |
| 7 | Count | 34,707 | 44,313 | 1,108 | 80,128 |
|  | \% within Grade | 43.3\% | 55.3\% | 1.4\% | 100.0\% |
| 8 | Count | 34,384 | 43,843 | 1,151 | 79,378 |
|  | \% within Grade | 43.3\% | 55.2\% | 1.5\% | 100.0\% |
| 9 | Count | 41,969 | 56,955 | 2,084 | 101,008 |
|  | \% within Grade | 41.6\% | 56.4\% | 2.1\% | 100.0\% |
| 10 | Count | 29,130 | 37,920 | 1,192 | 68,242 |
|  | \% within Grade | 42.7\% | 55.6\% | 1.7\% | 100.0\% |
| 11 | Count | 21,399 | 26,362 | 816 | 48,577 |
|  | \% within Grade | 44.1\% | 54.3\% | 1.7\% | 100.0\% |
| 12 | Count | 15,154 | 17,694 | 522 | 33,370 |
|  | \% within Grade | 45.4\% | 53.0\% | 1.6\% | 100.0\% |
| Total | Count | 547,309 | 663,908 | 20,142 | 1,231,359 |
|  | \% within Grade | 44.4\% | 53.9\% | 1.6\% | 100.0\% |

### 3.2.1.2.3 By Ethnicity

Table 3.2.1.2.3
Participation by Grade by Ethnicity S 401 Online

| Grade |  | Hispanic/Non-Hispanic |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Hispanic | Other | Unknown |  |
| 1 | Count | 110,766 | 52,683 | 7,532 | 170,981 |
|  | \% within Grade | 64.8\% | 30.8\% | 4.4\% | 100.0\% |
| 2 | Count | 116,261 | 51,200 | 7,123 | 174,584 |
|  | \% within Grade | 66.6\% | 29.3\% | 4.1\% | 100.0\% |
| 3 | Count | 124,658 | 51,440 | 6,600 | 182,698 |
|  | \% within Grade | 68.2\% | 28.2\% | 3.6\% | 100.0\% |
| 4 | Count | 84,293 | 33,526 | 5,153 | 122,972 |
|  | \% within Grade | 68.5\% | 27.3\% | 4.2\% | 100.0\% |
| 5 | Count | 60,811 | 26,145 | 3,813 | 90,769 |
|  | \% within Grade | 67.0\% | 28.8\% | 4.2\% | 100.0\% |
| 6 | Count | 52,484 | 22,680 | 3,488 | 78,652 |
|  | \% within Grade | 66.7\% | 28.8\% | 4.4\% | 100.0\% |
| 7 | Count | 53,426 | 23,239 | 3,463 | 80,128 |
|  | \% within Grade | 66.7\% | 29.0\% | 4.3\% | 100.0\% |
| 8 | Count | 52,524 | 23,322 | 3,532 | 79,378 |
|  | \% within Grade | 66.2\% | 29.4\% | 4.4\% | 100.0\% |
| 9 | Count | 66,886 | 28,181 | 5,941 | 101,008 |
|  | \% within Grade | 66.2\% | 27.9\% | 5.9\% | 100.0\% |
| 10 | Count | 45,317 | 19,877 | 3,048 | 68,242 |
|  | \% within Grade | 66.4\% | 29.1\% | 4.5\% | 100.0\% |
| 11 | Count | 30,369 | 16,192 | 2,016 | 48,577 |
|  | \% within Grade | 62.5\% | 33.3\% | 4.2\% | 100.0\% |
| 12 | Count | 19,373 | 12,737 | 1,260 | 33,370 |
|  | \% within Grade | 58.1\% | 38.2\% | 3.8\% | 100.0\% |
| Total | Count | 817,168 | 361,222 | 52,969 | 1,231,359 |
|  | \% within Grade | 66.4\% | 29.3\% | 4.3\% | 100.0\% |

### 3.2.1.2.4 By Tier by Domain

Table 3.2.1.2.4
Participation by Grade by Tier by Domain S401 Online

| Grade |  |  | Domain |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Writing | Speaking |
| 1 | Tier | Pre-A | - | 5,726 |
|  |  | A | 133,261 | 57,484 |
|  |  | BC | 37,698 | 107,764 |
|  | Total |  | 170,959 | 170,974 |
| 2 | Tier | Pre-A | - | 6,671 |
|  |  | A | 64,582 | 46,859 |
|  |  | BC | 109,956 | 121,049 |
|  | Total |  | 174,538 | 174,579 |
| 3 | Tier | Pre-A | - | 10,520 |
|  |  | A | 44,017 | 42,400 |
|  |  | BC | 138,657 | 129,768 |
|  | Total |  | 182,674 | 182,688 |
| 4 | Tier | Pre-A | - | 2,407 |
|  |  | A | 22,948 | 16,803 |
|  |  | BC | 100,019 | 103,760 |
|  | Total |  | 122,967 | 122,970 |
| 5 | Tier | Pre-A | - | 3,908 |
|  |  | A | 22,650 | 13,045 |
|  |  | BC | 68,113 | 73,810 |
|  | Total |  | 90,763 | 90,763 |
| 6 | Tier | Pre-A | - | 3,069 |
|  |  | A | 29,332 | 18,465 |
|  |  | BC | 49,315 | 57,114 |
|  | Total |  | 78,647 | 78,648 |
| 7 | Tier | Pre-A | - | 4,342 |
|  |  | A | 33,601 | 13,664 |
|  |  | BC | 46,522 | 62,115 |
|  | Total |  | 80,123 | 80,121 |
| 8 | Tier | Pre-A | - | 5,690 |
|  |  | A | 33,942 | 23,242 |
|  |  | BC | 45,433 | 50,440 |
|  | Total |  | 79,375 | 79,372 |


| Grade |  |  | Domain |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Writing | Speaking |
| 9 | Tier | Pre-A | - | 8,458 |
|  |  | A | 46,095 | 54,086 |
|  |  | BC | 54,900 | 38,452 |
|  | Total |  | 100,995 | 100,996 |
| 10 | Tier | Pre-A | - | 6,694 |
|  |  | A | 28,532 | 29,789 |
|  |  | BC | 39,701 | 31,750 |
|  | Total |  | 68,233 | 68,233 |
| 11 | Tier | Pre-A | - | 4,750 |
|  |  | A | 19,397 | 11,620 |
|  |  | BC | 29,176 | 32,202 |
|  | Total |  | 48,573 | 48,572 |
| 12 | Tier | Pre-A | - | 3,076 |
|  |  | A | 11,319 | 16,992 |
|  |  | BC | 22,046 | 13,299 |
|  | Total |  | 33,365 | 33,367 |

### 3.2.2 Scale Score Results

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

### 3.2.2.1.1 By Cluster

Table 3.2.2.1.1
Mean Scale Scores by Cluster S401 Online

| Cluster |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Mean | 328.49 | 290.64 | 260.96 | 271.80 | 300.96 | 275.63 | 301.76 | 283.21 |
|  | N | 155,365 | 155,144 | 170,761 | 135,531 | 125,541 | 155,093 | 144,040 | 116,652 |
| $2-3$ | Mean | 340.95 | 320.65 | 307.18 | 269.94 | 306.49 | 313.83 | 326.70 | 312.04 |
|  | N | 318,299 | 309,909 | 356,941 | 293,645 | 266,501 | 309,803 | 282,857 | 237,158 |
| $4-5$ | Mean | 403.89 | 345.40 | 328.57 | 311.83 | 358.84 | 335.79 | 362.28 | 342.49 |
|  | N | 191,571 | 168,465 | 140,492 | 175,273 | 159,860 | 113,502 | 154,642 | 87,235 |
| $6-8$ | Mean | 393.84 | 344.61 | 327.01 | 321.92 | 358.48 | 334.55 | 358.99 | 342.05 |
|  | N | 195,700 | 186,282 | 210,933 | 180,383 | 153,304 | 171,107 | 160,625 | 118,380 |
| $9-12$ | Mean | 386.17 | 368.94 | 351.35 | 314.68 | 351.01 | 358.44 | 373.54 | 356.52 |
|  | N | 208,374 | 170,634 | 217,998 | 193,155 | 165,040 | 155,139 | 147,903 | 108,773 |

### 3.2.2.1.2 By Cluster by Gender

Table 3.2.2.1.2
Mean Scale Scores by Cluster by Gender S401 Online

| Cluster | Gender |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | F | Mean | 331.85 | 292.39 | 264.82 | 276.16 | 304.78 | 278.35 | 303.91 | 286.14 |
|  |  | N | 72,356 | 71,329 | 79,119 | 63,893 | 59,372 | 71,307 | 66,458 | 54,696 |
|  | M | Mean | 325.15 | 288.81 | 257.21 | 267.47 | 297.12 | 272.94 | 299.55 | 280.25 |
|  |  | N | 80,376 | 81,212 | 88,474 | 69,470 | 64,179 | 81,185 | 75,183 | 60,123 |
|  | Missing | Mean | 338.48 | 299.97 | 269.43 | 281.81 | 311.11 | 284.96 | 311.50 | 293.02 |
|  |  | N | 2,633 | 2,603 | 3,168 | 2,168 | 1,990 | 2,601 | 2,399 | 1,833 |
| 2-3 | F | Mean | 342.68 | 322.84 | 313.05 | 272.74 | 308.65 | 317.87 | 328.72 | 315.34 |
|  |  | N | 147,678 | 141,278 | 165,018 | 137,238 | 124,828 | 141,243 | 129,344 | 109,532 |
|  | M | Mean | 339.28 | 318.61 | 301.76 | 267.23 | 304.35 | 310.18 | 324.81 | 308.97 |
|  |  | N | 165,809 | 164,093 | 186,152 | 152,119 | 137,809 | 164,023 | 149,386 | 124,288 |
|  | Missing | Mean | 345.52 | 326.45 | 313.80 | 276.58 | 312.70 | 320.28 | 332.15 | 318.53 |
|  |  | N | 4,812 | 4,538 | 5,771 | 4,288 | 3,864 | 4,537 | 4,127 | 3,338 |
| 4-5 | F | Mean | 403.86 | 346.25 | 333.30 | 315.17 | 360.49 | 338.49 | 362.71 | 344.65 |
|  |  | N | 82,963 | 70,943 | 61,000 | 76,024 | 69,520 | 48,060 | 65,374 | 37,236 |
|  | M | Mean | 404.09 | 344.82 | 325.08 | 309.28 | 357.65 | 333.94 | 362.05 | 341.03 |
|  |  | N | 105,916 | 95,146 | 78,128 | 96,855 | 88,193 | 64,324 | 87,110 | 49,202 |
|  | Missing | Mean | 397.34 | 343.39 | 316.50 | 308.78 | 354.16 | 326.31 | 358.76 | 331.58 |
|  |  | N | 2,692 | 2,376 | 1,364 | 2,394 | 2,147 | 1,118 | 2,158 | 797 |
| 6-8 | F | Mean | 394.03 | 347.71 | 332.67 | 324.27 | 359.68 | 338.90 | 361.11 | 345.24 |
|  |  | N | 85,939 | 79,055 | 91,255 | 78,072 | 67,257 | 72,711 | 69,242 | 51,099 |
|  | M | Mean | 394.09 | 342.53 | 322.91 | 320.38 | 357.87 | 331.55 | 357.65 | 339.86 |
|  |  | N | 107,421 | 104,941 | 117,223 | 100,189 | 84,329 | 96,476 | 89,547 | 66,041 |
|  | Missing | Mean | 375.05 | 332.71 | 312.31 | 307.86 | 341.56 | 320.47 | 344.64 | 326.89 |
|  |  | N | 2,340 | 2,286 | 2,455 | 2,122 | 1,718 | 1,920 | 1,836 | 1,240 |
| 9-12 | F | Mean | 388.18 | 372.04 | 356.34 | 317.01 | 353.23 | 362.55 | 376.36 | 360.00 |
|  |  | N | 90,670 | 71,047 | 92,873 | 82,653 | 71,612 | 64,445 | 62,413 | 45,722 |
|  | M | Mean | 384.93 | 366.93 | 347.85 | 313.25 | 349.60 | 355.74 | 371.72 | 354.24 |
|  |  | N | 114,293 | 96,921 | 121,598 | 107,479 | 90,989 | 88,420 | 83,355 | 61,676 |
|  | Missing | Mean | 374.00 | 359.41 | 340.60 | 301.90 | 338.42 | 346.48 | 362.56 | 343.01 |
|  |  | N | 3,411 | 2,666 | 3,527 | 3,023 | 2,439 | 2,274 | 2,135 | 1,375 |

### 3.2.2.1.3 By Cluster by Ethnicity

Table 3.2.2.1.3
Mean Scale Scores by Cluster by Ethnicity S401 Online

| Cluster | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Non-Hispanic Asian | Mean | 338.38 | 305.96 | 273.71 | 278.38 | 309.17 | 289.68 | 315.43 | 295.57 |
|  |  | N | 21,490 | 21,252 | 23,367 | 19,050 | 17,745 | 21,249 | 19,872 | 16,454 |
|  | Non-Hispanic Pacific Islander | Mean | 308.48 | 284.09 | 259.62 | 263.94 | 286.42 | 271.98 | 291.41 | 276.15 |
|  |  | N | 1,206 | 1,228 | 1,326 | 1,034 | 963 | 1,228 | 1,141 | 914 |
|  | Non-Hispanic Black | Mean | 318.47 | 291.23 | 259.36 | 277.18 | 299.20 | 275.20 | 299.40 | 282.73 |
|  |  | N | 7,395 | 7,524 | 8,353 | 6,657 | 6,021 | 7,522 | 6,807 | 5,571 |
|  | Hispanic (Of Any Race) | Mean | 326.66 | 286.42 | 257.86 | 269.36 | 298.70 | 271.99 | 298.30 | 279.88 |
|  |  | N | 101,226 | 101,400 | 110,678 | 87,788 | 81,434 | 101,363 | 94,323 | 76,004 |
|  | Non-Hispanic American Indian | Mean | 319.82 | 282.46 | 254.27 | 262.72 | 292.81 | 268.42 | 293.70 | 276.16 |
|  |  | N | 1,532 | 1,602 | 1,710 | 1,336 | 1,210 | 1,600 | 1,452 | 1,151 |
|  | Non-Hispanic <br> Multi-racial | Mean | 345.33 | 299.38 | 267.79 | 283.29 | 314.98 | 283.15 | 312.59 | 292.58 |
|  |  | N | 786 | 778 | 887 | 711 | 645 | 778 | 709 | 580 |
|  | Non-Hispanic White | Mean | 336.66 | 297.18 | 266.73 | 278.75 | 308.55 | 281.64 | 308.49 | 289.59 |
|  |  | N | 15,375 | 15,050 | 16,951 | 13,645 | 12,596 | 15,045 | 13,879 | 11,412 |
|  | Unknown | Mean | 320.01 | 292.80 | 256.77 | 266.05 | 294.66 | 274.68 | 300.59 | 280.86 |
|  |  | N | 6,355 | 6,310 | 7,489 | 5,310 | 4,927 | 6,308 | 5,857 | 4,566 |
| 2-3 | Non-Hispanic Asian | Mean | 352.26 | 334.43 | 316.83 | 275.23 | 314.71 | 325.84 | 339.78 | 322.91 |
|  |  | N | 38,534 | 38,276 | 42,587 | 35,911 | 32,929 | 38,273 | 35,268 | 30,252 |
|  | Non-Hispanic Pacific Islander | Mean | 324.90 | 312.42 | 305.72 | 260.04 | 292.47 | 309.12 | 316.00 | 303.99 |
|  |  | N | 2,262 | 2,219 | 2,524 | 2,024 | 1,861 | 2,219 | 2,032 | 1,671 |
|  | Non-Hispanic Black | Mean | 332.85 | 316.32 | 299.73 | 269.91 | 302.64 | 307.64 | 321.17 | 306.51 |
|  |  | N | 14,996 | 14,962 | 17,477 | 14,299 | 12,551 | 14,954 | 13,178 | 11,063 |
|  | Hispanic (Of Any Race) | Mean | 338.81 | 317.66 | 305.92 | 268.72 | 304.74 | 311.64 | 323.97 | 309.89 |
|  |  | N | 215,630 | 209,113 | 240,755 | 198,058 | 180,021 | 209,038 | 191,199 | 159,730 |
|  | Non-Hispanic American Indian | Mean | 335.84 | 311.20 | 301.75 | 266.47 | 302.04 | 306.31 | 318.61 | 305.72 |
|  |  | N | 3,768 | 3,754 | 4,332 | 3,474 | 3,056 | 3,750 | 3,338 | 2,717 |
|  | Non-Hispanic <br> Multi-racial | Mean | 356.28 | 329.19 | 310.87 | 278.01 | 317.93 | 320.26 | 337.12 | 319.86 |
|  |  | N | 1,506 | 1,445 | 1,670 | 1,401 | 1,286 | 1,444 | 1,335 | 1,151 |
|  | Non-Hispanic White | Mean | 351.74 | 328.78 | 311.88 | 276.20 | 314.98 | 320.21 | 335.57 | 319.12 |
|  |  | N | 30,110 | 29,021 | 33,919 | 28,203 | 25,427 | 29,011 | 26,316 | 22,238 |
|  | Unknown | Mean | 328.24 | 317.82 | 298.69 | 259.95 | 295.92 | 308.05 | 320.47 | 305.03 |
|  |  | N | 11,493 | 11,119 | 13,677 | 10,275 | 9,370 | 11,114 | 10,191 | 8,336 |


| Cluster | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 4-5 | Non-Hispanic Asian | Mean | 410.38 | 355.66 | 337.03 | 315.38 | 363.87 | 346.34 | 371.61 | 352.00 |
|  |  | N | 20,661 | 18,642 | 13,558 | 18,828 | 17,368 | 11,403 | 17,341 | 8,920 |
|  | Non-Hispanic Pacific Islander | Mean | 390.23 | 338.51 | 323.47 | 304.44 | 347.59 | 330.06 | 353.79 | 334.83 |
|  |  | N | 1,436 | 1,339 | 1,209 | 1,289 | 1,176 | 1,052 | 1,219 | 814 |
|  | Non-Hispanic Black | Mean | 392.78 | 338.70 | 317.68 | 311.41 | 353.21 | 325.91 | 353.86 | 333.82 |
|  |  | N | 10,781 | 9,671 | 6,394 | 10,246 | 9,113 | 5,026 | 8,636 | 3,798 |
|  | Hispanic (Of Any Race) | Mean | 404.65 | 344.18 | 329.05 | 311.81 | 359.11 | 335.39 | 361.70 | 342.20 |
|  |  | N | 130,639 | 114,056 | 97,228 | 119,270 | 108,993 | 78,192 | 104,852 | 60,189 |
|  | Non-Hispanic American Indian | Mean | 403.20 | 341.66 | 324.95 | 309.64 | 357.68 | 332.89 | 359.90 | 340.67 |
|  |  | N | 2,794 | 2,524 | 2,697 | 2,508 | 2,246 | 2,192 | 2,286 | 1,592 |
|  | Non-Hispanic Multi-racial | Mean | 418.91 | 353.43 | 335.93 | 320.53 | 370.96 | 343.30 | 372.37 | 351.53 |
|  |  | N | 729 | 648 | 456 | 663 | 619 | 381 | 603 | 299 |
|  | Non-Hispanic White | Mean | 409.15 | 350.77 | 331.66 | 317.13 | 364.20 | 340.13 | 367.56 | 347.12 |
|  |  | N | 16,832 | 14,527 | 13,653 | 15,692 | 14,216 | 10,768 | 13,242 | 8,339 |
|  | Unknown | Mean | 379.10 | 337.99 | 305.51 | 291.92 | 337.12 | 318.65 | 348.62 | 322.24 |
|  |  | N | 7,699 | 7,058 | 5,297 | 6,777 | 6,129 | 4,488 | 6,463 | 3,284 |
| 6-8 | Non-Hispanic Asian | Mean | 407.51 | 357.99 | 339.24 | 331.07 | 369.95 | 347.43 | 372.26 | 354.28 |
|  |  | N | 21,984 | 19,932 | 23,201 | 20,402 | 17,668 | 18,317 | 17,474 | 13,236 |
|  | Non-Hispanic Pacific Islander | Mean | 384.50 | 336.60 | 324.03 | 316.93 | 350.02 | 329.90 | 350.77 | 335.80 |
|  |  | N | 1,581 | 1,687 | 1,841 | 1,525 | 1,177 | 1,535 | 1,341 | 961 |
|  | Non-Hispanic Black | Mean | 388.56 | 337.99 | 319.04 | 321.25 | 354.68 | 326.93 | 352.30 | 334.43 |
|  |  | N | 12,155 | 11,711 | 13,434 | 12,021 | 9,674 | 10,454 | 9,581 | 7,029 |
|  | Hispanic (Of Any Race) | Mean | 392.28 | 343.53 | 326.74 | 320.86 | 357.19 | 333.97 | 357.88 | 341.32 |
|  |  | N | 131,525 | 125,542 | 141,608 | 120,031 | 102,642 | 115,855 | 108,899 | 80,157 |
|  | Non-Hispanic American Indian | Mean | 391.08 | 339.52 | 322.61 | 321.59 | 357.23 | 331.08 | 355.60 | 340.13 |
|  |  | N | 2,910 | 3,016 | 3,242 | 2,619 | 2,231 | 2,835 | 2,575 | 1,885 |
|  | Non-Hispanic Multi-racial | Mean | 414.76 | 352.83 | 334.91 | 332.83 | 374.68 | 342.86 | 371.54 | 352.81 |
|  |  | N | 590 | 547 | 628 | 536 | 472 | 508 | 480 | 367 |
|  | Non-Hispanic White | Mean | 405.76 | 351.17 | 332.55 | 330.56 | 368.92 | 340.56 | 367.08 | 349.51 |
|  |  | N | 16,765 | 15,655 | 18,277 | 16,034 | 13,294 | 14,237 | 13,299 | 9,844 |
|  | Unknown | Mean | 366.81 | 328.54 | 301.15 | 295.92 | 331.33 | 312.62 | 338.80 | 317.97 |
|  |  | N | 8,190 | 8,192 | 8,702 | 7,215 | 6,146 | 7,366 | 6,976 | 4,901 |


| Cluster | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 9-12 | Non-Hispanic Asian | Mean | 398.35 | 383.22 | 364.07 | 326.63 | 363.10 | 372.00 | 387.35 | 369.46 |
|  |  | N | 25,419 | 18,209 | 25,511 | 24,412 | 21,276 | 16,293 | 16,102 | 12,422 |
|  | Non-Hispanic Pacific Islander | Mean | 385.59 | 368.19 | 353.76 | 316.07 | 351.66 | 360.28 | 373.55 | 357.73 |
|  |  | N | 1,599 | 1,404 | 1,730 | 1,492 | 1,246 | 1,285 | 1,198 | 888 |
|  | Non-Hispanic Black | Mean | 380.20 | 367.32 | 348.51 | 318.31 | 348.75 | 355.39 | 370.02 | 352.13 |
|  |  | N | 15,913 | 11,670 | 16,762 | 15,650 | 12,900 | 10,300 | 9,744 | 7,196 |
|  | Hispanic <br> (Of Any Race) | Mean | 383.78 | 366.64 | 349.77 | 311.96 | 348.41 | 356.74 | 371.30 | 354.55 |
|  |  | N | 134,578 | 114,173 | 142,388 | 123,687 | 105,423 | 104,536 | 98,887 | 72,390 |
|  | Non-Hispanic American Indian | Mean | 394.57 | 371.79 | 357.73 | 323.33 | 359.60 | 364.08 | 378.28 | 362.91 |
|  |  | N | 2,580 | 2,325 | 2,799 | 2,452 | 2,103 | 2,187 | 2,011 | 1,547 |
|  | Non-Hispanic Multi-racial | Mean | 396.93 | 376.91 | 357.42 | 322.62 | 359.72 | 365.20 | 382.24 | 362.30 |
|  |  | N | 560 | 437 | 581 | 518 | 457 | 402 | 393 | 300 |
|  | Non-Hispanic White | Mean | 401.28 | 378.33 | 358.78 | 323.47 | 363.07 | 367.30 | 384.83 | 367.15 |
|  |  | N | 17,995 | 14,145 | 18,158 | 16,367 | 14,368 | 12,812 | 12,526 | 9,268 |
|  | Unknown | Mean | 366.48 | 354.39 | 330.34 | 293.36 | 330.46 | 338.87 | 356.37 | 335.99 |
|  |  | N | 9,730 | 8,271 | 10,069 | 8,577 | 7,267 | 7,324 | 7,042 | 4,762 |

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

### 3.2.2.2.1 By Grade

Table 3.2.2.2.1
Mean Scale Scores by Grade S401 Online

| Grade |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Mean | 328.49 | 290.64 | 260.96 | 271.80 | 300.96 | 275.63 | 301.76 | 283.21 |
|  | N | 155,365 | 155,144 | 170,761 | 135,531 | 125,541 | 155,093 | 144,040 | 116,652 |
| 2 | Mean | 328.84 | 311.66 | 297.30 | 264.26 | 297.64 | 304.22 | 316.59 | 302.49 |
|  | N | 154,676 | 151,219 | 174,393 | 141,170 | 127,608 | 151,153 | 137,371 | 113,409 |
| 3 | Mean | 352.39 | 329.23 | 316.61 | 275.21 | 314.62 | 323.00 | 336.25 | 320.80 |
|  | N | 163,623 | 158,690 | 182,548 | 152,475 | 138,893 | 158,650 | 145,486 | 123,749 |
| 4 | Mean | 401.72 | 343.69 | 324.97 | 311.23 | 357.50 | 333.11 | 360.44 | 340.21 |
|  | N | 110,019 | 97,309 | 80,224 | 100,539 | 91,410 | 64,908 | 89,088 | 49,613 |
| 5 | Mean | 406.82 | 347.74 | 333.35 | 312.63 | 360.63 | 339.36 | 364.79 | 345.50 |
|  | N | 81,552 | 71,156 | 60,268 | 74,734 | 68,450 | 48,594 | 65,554 | 37,622 |
| 6 | Mean | 384.07 | 337.19 | 319.93 | 318.48 | 351.88 | 327.56 | 351.05 | 335.22 |
|  | N | 64,820 | 62,850 | 69,202 | 60,103 | 51,264 | 57,241 | 54,115 | 39,883 |
| 7 | Mean | 395.25 | 344.69 | 327.17 | 321.61 | 359.34 | 334.64 | 359.43 | 342.55 |
|  | N | 65,713 | 62,539 | 71,312 | 59,887 | 50,778 | 57,760 | 53,919 | 39,356 |
| 8 | Mean | 402.12 | 352.18 | 333.80 | 325.65 | 364.24 | 341.58 | 366.71 | 348.51 |
|  | N | 65,167 | 60,893 | 70,419 | 60,393 | 51,262 | 56,106 | 52,591 | 39,141 |
| 9 | Mean | 384.38 | 364.10 | 345.20 | 311.46 | 349.00 | 352.58 | 369.34 | 352.28 |
|  | N | 83,661 | 70,625 | 87,734 | 78,155 | 66,533 | 63,997 | 61,056 | 44,742 |
| 10 | Mean | 385.51 | 368.92 | 351.52 | 314.14 | 350.18 | 358.75 | 373.37 | 356.18 |
|  | N | 56,588 | 46,083 | 59,166 | 52,350 | 44,777 | 41,909 | 39,938 | 29,366 |
| 11 | Mean | 387.49 | 373.75 | 357.17 | 317.14 | 352.61 | 364.04 | 377.54 | 360.44 |
|  | N | 40,338 | 31,489 | 42,079 | 36,469 | 31,294 | 28,668 | 27,298 | 19,798 |
| 12 | Mean | 390.99 | 377.45 | 361.17 | 321.97 | 356.41 | 368.21 | 381.42 | 364.73 |
|  | N | 27,787 | 22,437 | 29,019 | 26,181 | 22,436 | 20,565 | 19,611 | 14,867 |

### 3.2.2.2.2 By Grade by Gender

Table 3.2.2.2.2
Mean Scale Scores by Grade by Gender S401 Online

| Grade | Gender |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | F | Mean | 331.85 | 292.39 | 264.82 | 276.16 | 304.78 | 278.35 | 303.91 | 286.14 |
|  |  | N | 72,356 | 71,329 | 79,119 | 63,893 | 59,372 | 71,307 | 66,458 | 54,696 |
|  | M | Mean | 325.15 | 288.81 | 257.21 | 267.47 | 297.12 | 272.94 | 299.55 | 280.25 |
|  |  | N | 80,376 | 81,212 | 88,474 | 69,470 | 64,179 | 81,185 | 75,183 | 60,123 |
|  | Missing | Mean | 338.48 | 299.97 | 269.43 | 281.81 | 311.11 | 284.96 | 311.50 | 293.02 |
|  |  | N | 2,633 | 2,603 | 3,168 | 2,168 | 1,990 | 2,601 | 2,399 | 1,833 |
| 2 | F | Mean | 330.73 | 313.35 | 302.80 | 266.89 | 299.80 | 307.76 | 318.26 | 305.45 |
|  |  | N | 72,098 | 69,126 | 80,986 | 66,475 | 60,203 | 69,105 | 62,981 | 52,659 |
|  | M | Mean | 326.93 | 309.96 | 292.07 | 261.59 | 295.39 | 300.88 | 314.89 | 299.58 |
|  |  | N | 80,046 | 79,705 | 90,411 | 72,477 | 65,387 | 79,660 | 72,205 | 58,996 |
|  | Missing | Mean | 335.45 | 319.44 | 306.54 | 272.69 | 305.90 | 312.93 | 324.31 | 311.47 |
|  |  | N | 2,532 | 2,388 | 2,996 | 2,218 | 2,018 | 2,388 | 2,185 | 1,754 |
| 3 | F | Mean | 354.07 | 331.93 | 322.94 | 278.25 | 316.90 | 327.54 | 338.64 | 324.50 |
|  |  | N | 75,580 | 72,152 | 84,032 | 70,763 | 64,625 | 72,138 | 66,363 | 56,873 |
|  | M | Mean | 350.79 | 326.78 | 310.91 | 272.37 | 312.44 | 318.97 | 334.08 | 317.45 |
|  |  | N | 85,763 | 84,388 | 95,741 | 79,642 | 72,422 | 84,363 | 77,181 | 65,292 |
|  | Missing | Mean | 356.71 | 334.23 | 321.64 | 280.76 | 320.13 | 328.43 | 340.96 | 326.36 |
|  |  | N | 2,280 | 2,150 | 2,775 | 2,070 | 1,846 | 2,149 | 1,942 | 1,584 |
| 4 | F | Mean | 401.50 | 344.22 | 329.54 | 314.82 | 359.15 | 335.59 | 360.52 | 342.07 |
|  |  | N | 47,563 | 40,860 | 34,935 | 43,714 | 39,797 | 27,481 | 37,502 | 21,188 |
|  | M | Mean | 401.96 | 343.29 | 321.58 | 308.40 | 356.20 | 331.42 | 360.40 | 338.93 |
|  |  | N | 60,830 | 54,996 | 44,556 | 55,370 | 50,320 | 36,817 | 50,278 | 27,998 |
|  | Missing | Mean | 399.20 | 343.96 | 312.47 | 311.45 | 357.13 | 324.18 | 359.78 | 331.15 |
|  |  | N | 1,626 | 1,453 | 733 | 1,455 | 1,293 | 610 | 1,308 | 427 |
| 5 | F | Mean | 407.02 | 349.00 | 338.33 | 315.65 | 362.28 | 342.36 | 365.65 | 348.06 |
|  |  | N | 35,400 | 30,083 | 26,065 | 32,310 | 29,723 | 20,579 | 27,872 | 16,048 |
|  | M | Mean | 406.96 | 346.92 | 329.72 | 310.45 | 359.58 | 337.31 | 364.31 | 343.79 |
|  |  | N | 45,086 | 40,150 | 33,572 | 41,485 | 37,873 | 27,507 | 36,832 | 21,204 |
|  | Missing | Mean | 394.51 | 342.50 | 321.18 | 304.66 | 349.67 | 328.87 | 357.19 | 332.08 |
|  |  | N | 1,066 | 923 | 631 | 939 | 854 | 508 | 850 | 370 |


| Grade | Gender |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 6 | F | Mean | 383.96 | 340.35 | 326.25 | 320.70 | 352.84 | 332.30 | 353.11 | 338.64 |
|  |  | N | 28,304 | 26,503 | 29,812 | 25,888 | 22,329 | 24,159 | 23,195 | 17,094 |
|  | M | Mean | 384.40 | 335.04 | 315.31 | 317.05 | 351.39 | 324.26 | 349.67 | 332.82 |
|  |  | N | 35,748 | 35,571 | 38,602 | 33,498 | 28,346 | 32,454 | 30,312 | 22,377 |
|  | Missing | Mean | 373.21 | 327.88 | 307.47 | 305.00 | 339.64 | 316.02 | 341.58 | 323.32 |
|  |  | N | 768 | 776 | 788 | 717 | 589 | 628 | 608 | 412 |
| 7 | F | Mean | 395.66 | 347.97 | 332.87 | 324.21 | 360.86 | 339.07 | 361.72 | 345.92 |
|  |  | N | 28,898 | 26,674 | 30,901 | 25,960 | 22,284 | 24,671 | 23,333 | 17,066 |
|  | M | Mean | 395.40 | 342.47 | 323.06 | 319.87 | 358.49 | 331.55 | 357.98 | 340.21 |
|  |  | N | 36,036 | 35,119 | 39,610 | 33,244 | 27,935 | 32,449 | 29,978 | 21,884 |
|  | Missing | Mean | 373.38 | 332.18 | 310.75 | 307.39 | 341.25 | 320.48 | 343.13 | 326.93 |
|  |  | N | 779 | 746 | 801 | 683 | 559 | 640 | 608 | 406 |
| 8 | F | Mean | 402.32 | 354.99 | 338.73 | 327.86 | 365.28 | 345.41 | 368.65 | 351.22 |
|  |  | N | 28,737 | 25,878 | 30,542 | 26,224 | 22,644 | 23,881 | 22,714 | 16,939 |
|  | M | Mean | 402.49 | 350.38 | 330.28 | 324.23 | 363.82 | 339.03 | 365.58 | 346.75 |
|  |  | N | 35,637 | 34,251 | 39,011 | 33,447 | 28,048 | 31,573 | 29,257 | 21,780 |
|  | Missing | Mean | 378.46 | 338.12 | 318.15 | 311.15 | 343.84 | 324.74 | 349.14 | 330.35 |
|  |  | N | 793 | 764 | 866 | 722 | 570 | 652 | 620 | 422 |
| 9 | F | Mean | 387.46 | 367.58 | 351.21 | 314.72 | 352.16 | 357.31 | 372.58 | 356.40 |
|  |  | N | 35,295 | 28,578 | 36,271 | 32,496 | 28,069 | 25,876 | 25,036 | 18,363 |
|  | M | Mean | 382.74 | 362.14 | 341.45 | 309.74 | 347.23 | 349.83 | 367.54 | 349.85 |
|  |  | N | 46,798 | 40,743 | 49,824 | 44,260 | 37,335 | 37,007 | 34,987 | 25,713 |
|  | Missing | Mean | 363.71 | 349.00 | 326.33 | 290.10 | 328.50 | 333.93 | 351.91 | 332.42 |
|  |  | N | 1,568 | 1,304 | 1,639 | 1,399 | 1,129 | 1,114 | 1,033 | 666 |
| 10 | F | Mean | 387.01 | 371.43 | 355.79 | 315.82 | 351.85 | 362.16 | 375.57 | 358.92 |
|  |  | N | 24,545 | 19,098 | 25,125 | 22,274 | 19,295 | 17,308 | 16,796 | 12,193 |
|  | M | Mean | 384.58 | 367.25 | 348.42 | 313.05 | 349.11 | 356.46 | 371.90 | 354.38 |
|  |  | N | 31,171 | 26,315 | 33,159 | 29,311 | 24,847 | 24,027 | 22,594 | 16,815 |
|  | Missing | Mean | 376.22 | 363.21 | 346.80 | 307.18 | 341.69 | 352.25 | 366.30 | 347.11 |
|  |  | N | 872 | 670 | 882 | 765 | 635 | 574 | 548 | 358 |
| 11 | F | Mean | 388.30 | 375.97 | 360.69 | 318.43 | 353.70 | 366.97 | 379.52 | 362.87 |
|  |  | N | 18,012 | 13,456 | 18,405 | 16,058 | 13,948 | 12,196 | 11,764 | 8,512 |
|  | M | Mean | 386.79 | 372.04 | 354.37 | 316.15 | 351.75 | 361.83 | 376.00 | 358.63 |
|  |  | N | 21,703 | 17,607 | 23,053 | 19,891 | 16,923 | 16,115 | 15,188 | 11,072 |
|  | Missing | Mean | 388.84 | 374.22 | 357.01 | 314.98 | 351.32 | 363.64 | 377.64 | 357.76 |
|  |  | N | 623 | 426 | 621 | 520 | 423 | 357 | 346 | 214 |
| 12 | F | Mean | 392.23 | 380.73 | 365.54 | 323.66 | 358.11 | 372.34 | 384.38 | 368.25 |
|  |  | N | 12,818 | 9,915 | 13,072 | 11,825 | 10,300 | 9,065 | 8,817 | 6,654 |
|  | M | Mean | 389.97 | 374.81 | 357.51 | 320.63 | 355.01 | 364.92 | 378.98 | 361.90 |
|  |  | N | 14,621 | 12,256 | 15,562 | 14,017 | 11,884 | 11,271 | 10,586 | 8,076 |
|  | Missing | Mean | 388.25 | 377.17 | 360.67 | 318.56 | 352.95 | 366.34 | 380.42 | 360.72 |
|  |  | N | 348 | 266 | 385 | 339 | 252 | 229 | 208 | 137 |

### 3.2.2.2.3 By Grade by Ethnicity

Table 3.2.2.2.3
Mean Scale Scores by Grade by Ethnicity S401 Online

| Grade | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Non-Hispanic Asian | Mean | 338.38 | 305.96 | 273.71 | 278.38 | 309.17 | 289.68 | 315.43 | 295.57 |
|  |  | N | 21,490 | 21,252 | 23,367 | 19,050 | 17,745 | 21,249 | 19,872 | 16,454 |
|  | Non-Hispanic Pacific Islander | Mean | 308.48 | 284.09 | 259.62 | 263.94 | 286.42 | 271.98 | 291.41 | 276.15 |
|  |  | N | 1,206 | 1,228 | 1,326 | 1,034 | 963 | 1,228 | 1,141 | 914 |
|  | Non-Hispanic Black | Mean | 318.47 | 291.23 | 259.36 | 277.18 | 299.20 | 275.20 | 299.40 | 282.73 |
|  |  | N | 7,395 | 7,524 | 8,353 | 6,657 | 6,021 | 7,522 | 6,807 | 5,571 |
|  | Hispanic (Of Any Race) | Mean | 326.66 | 286.42 | 257.86 | 269.36 | 298.70 | 271.99 | 298.30 | 279.88 |
|  |  | N | 101,226 | 101,400 | 110,678 | 87,788 | 81,434 | 101,363 | 94,323 | 76,004 |
|  | Non-Hispanic American Indian | Mean | 319.82 | 282.46 | 254.27 | 262.72 | 292.81 | 268.42 | 293.70 | 276.16 |
|  |  | N | 1,532 | 1,602 | 1,710 | 1,336 | 1,210 | 1,600 | 1,452 | 1,151 |
|  | Non-Hispanic <br> Multi-racial | Mean | 345.33 | 299.38 | 267.79 | 283.29 | 314.98 | 283.15 | 312.59 | 292.58 |
|  |  | N | 786 | 778 | 887 | 711 | 645 | 778 | 709 | 580 |
|  | Non-Hispanic White | Mean | 336.66 | 297.18 | 266.73 | 278.75 | 308.55 | 281.64 | 308.49 | 289.59 |
|  |  | N | 15,375 | 15,050 | 16,951 | 13,645 | 12,596 | 15,045 | 13,879 | 11,412 |
|  | Unknown | Mean | 320.01 | 292.80 | 256.77 | 266.05 | 294.66 | 274.68 | 300.59 | 280.86 |
|  |  | N | 6,355 | 6,310 | 7,489 | 5,310 | 4,927 | 6,308 | 5,857 | 4,566 |
| 2 | Non-Hispanic Asian | Mean | 341.57 | 325.25 | 308.38 | 269.58 | 306.58 | 316.90 | 329.99 | 314.12 |
|  |  | N | 19,577 | 19,400 | 21,740 | 18,034 | 16,474 | 19,400 | 17,813 | 15,055 |
|  | Non-Hispanic Pacific Islander | Mean | 313.28 | 304.37 | 295.21 | 253.90 | 283.63 | 299.70 | 306.59 | 294.81 |
|  |  | N | 1,086 | 1,054 | 1,191 | 950 | 887 | 1,054 | 980 | 801 |
|  | Non-Hispanic Black | Mean | 322.56 | 308.37 | 290.45 | 266.48 | 296.09 | 299.12 | 312.41 | 298.67 |
|  |  | N | 7,232 | 7,267 | 8,462 | 6,820 | 5,968 | 7,262 | 6,363 | 5,275 |
|  | Hispanic (Of Any Race) | Mean | 325.85 | 308.39 | 295.36 | 262.58 | 295.20 | 301.56 | 313.40 | 299.74 |
|  |  | N | 103,460 | 100,979 | 116,165 | 93,988 | 85,072 | 100,929 | 91,856 | 75,487 |
|  | Non-Hispanic American Indian | Mean | 321.95 | 302.79 | 291.13 | 260.60 | 292.00 | 296.70 | 308.56 | 296.11 |
|  |  | N | 1,756 | 1,780 | 2,034 | 1,607 | 1,412 | 1,778 | 1,575 | 1,276 |
|  | Non-Hispanic Multi-racial | Mean | 346.84 | 320.88 | 304.05 | 273.98 | 311.00 | 312.43 | 328.02 | 311.76 |
|  |  | N | 765 | 726 | 857 | 706 | 641 | 725 | 666 | 563 |
|  | Non-Hispanic White | Mean | 340.58 | 318.97 | 302.59 | 270.54 | 306.61 | 310.31 | 325.16 | 309.51 |
|  |  | N | 14,830 | 14,265 | 16,852 | 13,831 | 12,385 | 14,260 | 12,844 | 10,732 |
|  | Unknown | Mean | 319.98 | 312.02 | 292.00 | 258.32 | 291.36 | 301.56 | 313.85 | 299.26 |
|  |  | N | 5,970 | 5,748 | 7,092 | 5,234 | 4,769 | 5,745 | 5,274 | 4,220 |


| Grade | Ethnicity |  | List | Read | Writ | S pek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 3 | Non-Hispanic Asian | Mean | 363.29 | 343.86 | 325.64 | 280.92 | 322.86 | 335.02 | 349.77 | 331.62 |
|  |  | N | 18,957 | 18,876 | 20,847 | 17,877 | 16,455 | 18,873 | 17,455 | 15,197 |
|  | Non-Hispanic <br> Pacific Islander | Mean | 335.63 | 319.70 | 315.11 | 265.48 | 300.52 | 317.64 | 324.76 | 312.44 |
|  |  | N | 1,176 | 1,165 | 1,333 | 1,074 | 974 | 1,165 | 1,052 | 870 |
|  | Non-Hispanic Black | Mean | 342.45 | 323.83 | 308.43 | 273.05 | 308.58 | 315.68 | 329.34 | 313.64 |
|  |  | N | 7,764 | 7,695 | 9,015 | 7,479 | 6,583 | 7,692 | 6,815 | 5,788 |
|  | Hispanic (Of Any Race) | Mean | 350.76 | 326.32 | 315.76 | 274.27 | 313.29 | 321.06 | 333.74 | 318.98 |
|  |  | N | 112,170 | 108,134 | 124,590 | 104,070 | 94,949 | 108,109 | 99,343 | 84,243 |
|  | Non-Hispanic American Indian | Mean | 347.97 | 318.79 | 311.15 | 271.52 | 310.66 | 314.99 | 327.59 | 314.23 |
|  |  | N | 2,012 | 1,974 | 2,298 | 1,867 | 1,644 | 1,972 | 1,763 | 1,441 |
|  | Non-Hispanic <br> Multi-racial | Mean | 366.02 | 337.59 | 318.05 | 282.11 | 324.82 | 328.15 | 346.18 | 327.62 |
|  |  | N | 741 | 719 | 813 | 695 | 645 | 719 | 669 | 588 |
|  | Non-Hispanic White | Mean | 362.57 | 338.27 | 321.06 | 281.64 | 322.93 | 329.78 | 345.49 | 328.08 |
|  |  | N | 15,280 | 14,756 | 17,067 | 14,372 | 13,042 | 14,751 | 13,472 | 11,506 |
|  | Unknown | Mean | 337.17 | 324.03 | 305.88 | 261.64 | 300.66 | 315.00 | 327.57 | 310.93 |
|  |  | N | 5,523 | 5,371 | 6,585 | 5,041 | 4,601 | 5,369 | 4,917 | 4,116 |
| 4 | Non-Hispanic Asian | Mean | 409.39 | 353.79 | 333.97 | 314.85 | 363.27 | 344.15 | 370.09 | 350.39 |
|  |  | N | 11,822 | 10,782 | 7,738 | 10,763 | 9,872 | 6,538 | 9,974 | 5,030 |
|  | Non-Hispanic Pacific Islander | Mean | 389.14 | 336.74 | 318.19 | 302.68 | 346.60 | 326.54 | 352.20 | 332.14 |
|  |  | N | 776 | 717 | 659 | 701 | 642 | 567 | 650 | 439 |
|  | Non-Hispanic Black | Mean | 391.10 | 336.90 | 313.20 | 310.42 | 351.80 | 322.73 | 352.11 | 330.83 |
|  |  | N | 5,810 | 5,297 | 3,416 | 5,538 | 4,890 | 2,727 | 4,706 | 2,042 |
|  | Hispanic (Of Any Race) | Mean | 402.02 | 342.48 | 325.47 | 311.17 | 357.47 | 332.72 | 359.70 | 339.83 |
|  |  | N | 75,854 | 66,551 | 56,194 | 69,175 | 63,092 | 45,250 | 61,091 | 34,744 |
|  | Non-Hispanic American Indian | Mean | 398.91 | 339.23 | 320.08 | 307.17 | 354.63 | 329.19 | 356.70 | 337.36 |
|  |  | N | 1,580 | 1,457 | 1,526 | 1,445 | 1,279 | 1,247 | 1,294 | 892 |
|  | Non-Hispanic Multi-racial | Mean | 416.19 | 352.08 | 333.15 | 319.04 | 368.56 | 342.59 | 371.23 | 351.45 |
|  |  | N | 438 | 386 | 277 | 397 | 374 | 232 | 364 | 189 |
|  | Non-Hispanic White | Mean | 407.34 | 348.39 | 327.53 | 316.30 | 362.92 | 336.60 | 365.37 | 344.14 |
|  |  | N | 9,379 | 8,104 | 7,572 | 8,686 | 7,827 | 5,952 | 7,353 | 4,551 |
|  | Unknown | Mean | 379.70 | 338.12 | 301.31 | 294.12 | 339.02 | 316.28 | 349.05 | 321.09 |
|  |  | N | 4,360 | 4,015 | 2,842 | 3,834 | 3,434 | 2,395 | 3,656 | 1,726 |


| Grade | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 5 | Non-Hispanic Asian | Mean | 411.71 | 358.23 | 341.09 | 316.08 | 364.65 | 349.28 | 373.65 | 354.07 |
|  |  | N | 8,839 | 7,860 | 5,820 | 8,065 | 7,496 | 4,865 | 7,367 | 3,890 |
|  | Non-Hispanic Pacific Islander | Mean | 391.51 | 340.55 | 329.78 | 306.55 | 348.77 | 334.17 | 355.61 | 337.98 |
|  |  | N | 660 | 622 | 550 | 588 | 534 | 485 | 569 | 375 |
|  | Non-Hispanic Black | Mean | 394.74 | 340.88 | 322.82 | 312.58 | 354.85 | 329.68 | 355.96 | 337.30 |
|  |  | N | 4,971 | 4,374 | 2,978 | 4,708 | 4,223 | 2,299 | 3,930 | 1,756 |
|  | Hispanic (Of Any Race) | Mean | 408.30 | 346.57 | 333.96 | 312.70 | 361.35 | 339.06 | 364.51 | 345.43 |
|  |  | N | 54,785 | 47,505 | 41,034 | 50,095 | 45,901 | 32,942 | 43,761 | 25,445 |
|  | Non-Hispanic American Indian | Mean | 408.80 | 344.99 | 331.30 | 312.99 | 361.70 | 337.78 | 364.08 | 344.89 |
|  |  | N | 1,214 | 1,067 | 1,171 | 1,063 | 967 | 945 | 992 | 700 |
|  | Non-Hispanic Multi-racial | Mean | 423.01 | 355.41 | 340.23 | 322.75 | 374.64 | 344.42 | 374.12 | 351.66 |
|  |  | N | 291 | 262 | 179 | 266 | 245 | 149 | 239 | 110 |
|  | Non-Hispanic White | Mean | 411.43 | 353.77 | 336.80 | 318.16 | 365.76 | 344.48 | 370.29 | 350.71 |
|  |  | N | 7,453 | 6,423 | 6,081 | 7,006 | 6,389 | 4,816 | 5,889 | 3,788 |
|  | Unknown | Mean | 378.31 | 337.81 | 310.38 | 289.05 | 334.69 | 321.35 | 348.06 | 323.50 |
|  |  | N | 3,339 | 3,043 | 2,455 | 2,943 | 2,695 | 2,093 | 2,807 | 1,558 |
| 6 | Non-Hispanic Asian | Mean | 395.32 | 348.38 | 330.00 | 324.93 | 360.88 | 338.30 | 362.15 | 345.46 |
|  |  | N | 7,007 | 6,559 | 7,358 | 6,530 | 5,678 | 5,985 | 5,777 | 4,375 |
|  | Non-Hispanic Pacific Islander | Mean | 372.87 | 328.99 | 316.27 | 311.88 | 341.12 | 322.21 | 341.23 | 326.98 |
|  |  | N | 561 | 575 | 636 | 520 | 417 | 523 | 476 | 345 |
|  | Non-Hispanic Black | Mean | 377.50 | 330.73 | 310.33 | 317.69 | 347.62 | 319.20 | 343.98 | 327.00 |
|  |  | N | 3,930 | 4,011 | 4,394 | 3,976 | 3,187 | 3,560 | 3,221 | 2,373 |
|  | Hispanic (Of Any Race) | Mean | 382.65 | 336.33 | 320.12 | 317.89 | 350.88 | 327.34 | 350.16 | 334.83 |
|  |  | N | 43,623 | 42,311 | 46,511 | 40,183 | 34,398 | 38,681 | 36,617 | 26,980 |
|  | Non-Hispanic American Indian | Mean | 379.58 | 332.10 | 314.91 | 316.70 | 349.08 | 323.78 | 346.55 | 332.35 |
|  |  | N | 962 | 993 | 1,045 | 870 | 745 | 910 | 841 | 602 |
|  | Non-Hispanic <br> Multi-racial | Mean | 400.04 | 340.44 | 323.63 | 323.33 | 363.26 | 330.64 | 357.33 | 341.37 |
|  |  | N | 224 | 205 | 236 | 198 | 180 | 191 | 183 | 139 |
|  | Non-Hispanic White | Mean | 395.78 | 342.91 | 325.64 | 326.71 | 361.88 | 332.85 | 358.21 | 341.76 |
|  |  | N | 5,795 | 5,411 | 6,152 | 5,414 | 4,586 | 4,901 | 4,653 | 3,403 |
|  | Unknown | Mean | 365.06 | 325.39 | 295.96 | 295.24 | 330.51 | 308.97 | 336.27 | 315.09 |
|  |  | N | 2,718 | 2,785 | 2,870 | 2,412 | 2,073 | 2,490 | 2,347 | 1,666 |


| Grade | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 7 | Non-Hispanic Asian | Mean | 408.91 | 358.77 | 339.87 | 331.14 | 370.72 | 348.23 | 372.93 | 355.07 |
|  |  | N | 7,430 | 6,692 | 7,834 | 6,782 | 5,888 | 6,186 | 5,893 | 4,414 |
|  | Non-Hispanic Pacific Islander | Mean | 386.29 | 336.14 | 322.75 | 315.59 | 351.19 | 329.02 | 351.05 | 335.88 |
|  |  | N | 540 | 580 | 618 | 534 | 407 | 529 | 462 | 337 |
|  | Non-Hispanic Black | Mean | 391.49 | 337.94 | 319.58 | 321.58 | 356.89 | 327.38 | 353.32 | 335.53 |
|  |  | N | 4,129 | 3,943 | 4,550 | 4,011 | 3,236 | 3,529 | 3,237 | 2,357 |
|  | Hispanic (Of Any Race) | Mean | 393.55 | 343.52 | 326.80 | 320.39 | 357.86 | 333.90 | 358.20 | 341.61 |
|  |  | N | 44,310 | 42,247 | 48,050 | 39,917 | 34,122 | 39,214 | 36,649 | 26,722 |
|  | Non-Hispanic American Indian | Mean | 394.47 | 340.37 | 324.40 | 322.12 | 359.38 | 331.91 | 357.35 | 341.54 |
|  |  | N | 989 | 1,026 | 1,131 | 896 | 757 | 980 | 872 | 650 |
|  | Non-Hispanic <br> Multi-racial | Mean | 417.42 | 355.39 | 337.00 | 332.01 | 375.03 | 345.86 | 373.79 | 354.75 |
|  |  | N | 190 | 178 | 201 | 168 | 150 | 166 | 157 | 120 |
|  | Non-Hispanic White | Mean | 408.23 | 351.90 | 332.87 | 330.41 | 370.64 | 341.24 | 368.62 | 351.08 |
|  |  | N | 5,400 | 5,176 | 6,069 | 5,253 | 4,228 | 4,722 | 4,330 | 3,183 |
|  | Unknown | Mean | 366.21 | 326.97 | 299.88 | 295.25 | 331.40 | 311.46 | 337.47 | 317.68 |
|  |  | N | 2,725 | 2,697 | 2,859 | 2,326 | 1,990 | 2,434 | 2,319 | 1,573 |
| 8 | Non-Hispanic Asian | Mean | 417.44 | 366.65 | 347.11 | 336.66 | 377.64 | 355.51 | 381.66 | 362.16 |
|  |  | N | 7,547 | 6,681 | 8,009 | 7,090 | 6,102 | 6,146 | 5,804 | 4,447 |
|  | Non-Hispanic Pacific Islander | Mean | 396.08 | 345.32 | 333.77 | 324.03 | 359.19 | 339.18 | 361.69 | 346.62 |
|  |  | N | 480 | 532 | 587 | 471 | 353 | 483 | 403 | 279 |
|  | Non-Hispanic Black | Mean | 396.23 | 345.79 | 327.01 | 324.42 | 359.40 | 334.64 | 359.83 | 340.99 |
|  |  | N | 4,096 | 3,757 | 4,490 | 4,034 | 3,251 | 3,365 | 3,123 | 2,299 |
|  | Hispanic (Of Any Race) | Mean | 400.64 | 350.97 | 333.22 | 324.31 | 362.89 | 340.79 | 365.49 | 347.66 |
|  |  | N | 43,592 | 40,984 | 47,047 | 39,931 | 34,122 | 37,960 | 35,633 | 26,455 |
|  | Non-Hispanic American Indian | Mean | 399.13 | 346.04 | 328.28 | 326.02 | 363.34 | 337.26 | 362.66 | 346.09 |
|  |  | N | 959 | 997 | 1,066 | 853 | 729 | 945 | 862 | 633 |
|  | Non-Hispanic <br> Multi-racial | Mean | 430.64 | 365.52 | 346.64 | 344.69 | 388.80 | 355.01 | 387.59 | 365.37 |
|  |  | N | 176 | 164 | 191 | 170 | 142 | 151 | 140 | 108 |
|  | Non-Hispanic White | Mean | 413.75 | 359.23 | 339.25 | 334.58 | 374.50 | 348.05 | 375.11 | 356.06 |
|  |  | N | 5,570 | 5,068 | 6,056 | 5,367 | 4,480 | 4,614 | 4,316 | 3,258 |
|  | Unknown | Mean | 369.13 | 333.32 | 307.38 | 297.22 | 332.07 | 317.49 | 342.71 | 321.15 |
|  |  | N | 2,747 | 2,710 | 2,973 | 2,477 | 2,083 | 2,442 | 2,310 | 1,662 |


| Grade | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 9 | Non-Hispanic Asian | Mean | 402.52 | 382.51 | 361.56 | 327.20 | 365.75 | 370.18 | 387.87 | 369.29 |
|  |  | N | 9,017 | 6,668 | 9,047 | 8,746 | 7,621 | 5,953 | 5,904 | 4,555 |
|  | Non-Hispanic Pacific Islander | Mean | 386.07 | 364.62 | 350.67 | 314.87 | 350.22 | 356.96 | 371.58 | 355.90 |
|  |  | N | 596 | 555 | 649 | 585 | 467 | 496 | 462 | 348 |
|  | Non-Hispanic Black | Mean | 379.02 | 362.00 | 341.24 | 314.68 | 346.82 | 348.88 | 365.95 | 347.79 |
|  |  | N | 5,793 | 4,519 | 6,103 | 5,723 | 4,705 | 3,970 | 3,762 | 2,753 |
|  | Hispanic (Of Any Race) | Mean | 382.65 | 362.47 | 344.45 | 309.56 | 347.09 | 351.61 | 367.73 | 350.89 |
|  |  | N | 55,520 | 48,169 | 58,936 | 51,454 | 43,733 | 44,037 | 41,698 | 30,519 |
|  | Non-Hispanic American Indian | Mean | 392.16 | 367.51 | 353.75 | 321.63 | 358.59 | 359.64 | 374.17 | 359.41 |
|  |  | N | 999 | 920 | 1,085 | 985 | 825 | 846 | 783 | 585 |
|  | Non-Hispanic <br> Multi-racial | Mean | 400.36 | 375.34 | 358.59 | 324.06 | 362.06 | 366.23 | 382.60 | 364.98 |
|  |  | N | 239 | 185 | 242 | 230 | 201 | 171 | 168 | 132 |
|  | Non-Hispanic White | Mean | 397.84 | 373.06 | 353.51 | 320.30 | 359.95 | 361.83 | 380.16 | 363.12 |
|  |  | N | 6,780 | 5,355 | 6,751 | 6,306 | 5,516 | 4,774 | 4,703 | 3,501 |
|  | Unknown | Mean | 354.50 | 343.31 | 314.30 | 280.09 | 318.52 | 325.36 | 344.93 | 323.48 |
|  |  | N | 4,717 | 4,254 | 4,921 | 4,126 | 3,465 | 3,750 | 3,576 | 2,349 |
| 10 | Non-Hispanic Asian | Mean | 396.11 | 381.58 | 362.39 | 325.47 | 361.30 | 370.28 | 385.18 | 367.27 |
|  |  | N | 6,524 | 4,616 | 6,559 | 6,215 | 5,449 | 4,156 | 4,075 | 3,150 |
|  | Non-Hispanic Pacific Islander | Mean | 382.56 | 368.15 | 350.70 | 314.63 | 350.55 | 359.16 | 372.32 | 355.50 |
|  |  | N | 414 | 358 | 442 | 385 | 328 | 325 | 311 | 231 |
|  | Non-Hispanic Black | Mean | 379.48 | 367.08 | 348.49 | 318.26 | 348.20 | 355.25 | 369.44 | 351.45 |
|  |  | N | 4,174 | 2,930 | 4,375 | 4,110 | 3,424 | 2,577 | 2,463 | 1,832 |
|  | Hispanic (Of Any Race) | Mean | 383.22 | 366.71 | 350.00 | 311.33 | 347.61 | 357.14 | 371.32 | 354.45 |
|  |  | N | 37,562 | 31,877 | 39,728 | 34,554 | 29,386 | 29,151 | 27,546 | 20,152 |
|  | Non-Hispanic American Indian | Mean | 395.56 | 372.37 | 358.44 | 324.50 | 360.77 | 364.88 | 379.75 | 364.48 |
|  |  | N | 683 | 615 | 755 | 638 | 551 | 588 | 530 | 419 |
|  | Non-Hispanic <br> Multi-racial | Mean | 394.21 | 379.29 | 355.79 | 321.53 | 357.75 | 365.84 | 383.58 | 360.38 |
|  |  | N | 131 | 103 | 140 | 120 | 104 | 95 | 89 | 66 |
|  | Non-Hispanic White | Mean | 400.06 | 377.74 | 358.08 | 322.26 | 361.75 | 366.88 | 383.73 | 365.69 |
|  |  | N | 4,668 | 3,589 | 4,703 | 4,209 | 3,700 | 3,255 | 3,193 | 2,335 |
|  | Unknown | Mean | 371.98 | 360.37 | 337.81 | 298.93 | 335.18 | 345.94 | 362.35 | 341.55 |
|  |  | N | 2,432 | 1,995 | 2,464 | 2,119 | 1,835 | 1,762 | 1,731 | 1,181 |


| Grade | Ethnicity |  | List | Read | Writ | Spek | Oral | Litr | Cphn | Over |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 11 | Non-Hispanic Asian | Mean | 397.21 | 385.43 | 367.27 | 327.07 | 362.86 | 374.86 | 389.09 | 371.94 |
|  |  | N | 5,638 | 3,881 | 5,631 | 5,248 | 4,591 | 3,465 | 3,433 | 2,590 |
|  | Non-Hispanic Pacific Islander | Mean | 387.40 | 370.69 | 356.63 | 315.47 | 352.18 | 362.33 | 374.17 | 357.55 |
|  |  | N | 321 | 277 | 367 | 288 | 245 | 267 | 232 | 169 |
|  | Non-Hispanic Black | Mean | 381.59 | 372.72 | 353.70 | 320.13 | 350.12 | 361.04 | 374.56 | 356.00 |
|  |  | N | 3,277 | 2,333 | 3,485 | 3,175 | 2,594 | 2,073 | 1,938 | 1,418 |
|  | Hispanic (Of Any Race) | Mean | 383.98 | 370.67 | 355.01 | 313.52 | 349.03 | 361.68 | 374.42 | 357.74 |
|  |  | N | 25,267 | 20,435 | 26,598 | 22,593 | 19,362 | 18,690 | 17,677 | 12,712 |
|  | Non-Hispanic American Indian | Mean | 395.86 | 376.16 | 359.97 | 321.32 | 358.28 | 367.67 | 381.34 | 364.24 |
|  |  | N | 545 | 460 | 569 | 483 | 429 | 436 | 408 | 305 |
|  | Non-Hispanic <br> Multi-racial | Mean | 393.05 | 377.02 | 353.67 | 323.94 | 358.27 | 362.13 | 380.38 | 360.07 |
|  |  | N | 110 | 89 | 117 | 100 | 90 | 82 | 81 | 61 |
|  | Non-Hispanic White | Mean | 403.39 | 382.47 | 363.43 | 326.68 | 365.47 | 371.39 | 388.26 | 369.87 |
|  |  | N | 3,550 | 2,750 | 3,638 | 3,144 | 2,753 | 2,519 | 2,433 | 1,767 |
|  | Unknown | Mean | 382.43 | 370.21 | 350.50 | 308.61 | 344.99 | 358.07 | 372.13 | 352.18 |
|  |  | N | 1,630 | 1,264 | 1,674 | 1,438 | 1,230 | 1,136 | 1,096 | 776 |
| 12 | Non-Hispanic Asian | Mean | 394.42 | 384.42 | 367.76 | 326.61 | 360.52 | 374.95 | 387.24 | 370.05 |
|  |  | N | 4,240 | 3,044 | 4,274 | 4,203 | 3,615 | 2,719 | 2,690 | 2,127 |
|  | Non-Hispanic Pacific Islander | Mean | 387.02 | 374.31 | 362.22 | 322.18 | 356.05 | 367.74 | 379.48 | 366.19 |
|  |  | N | 268 | 214 | 272 | 234 | 206 | 197 | 193 | 140 |
|  | Non-Hispanic Black | Mean | 382.17 | 373.78 | 357.94 | 324.03 | 352.17 | 364.04 | 375.03 | 358.55 |
|  |  | N | 2,669 | 1,888 | 2,799 | 2,642 | 2,177 | 1,680 | 1,581 | 1,193 |
|  | Hispanic (Of Any Race) | Mean | 388.59 | 375.11 | 359.36 | 319.20 | 353.73 | 366.37 | 379.13 | 362.69 |
|  |  | N | 16,229 | 13,692 | 17,126 | 15,086 | 12,942 | 12,658 | 11,966 | 9,007 |
|  | Non-Hispanic American Indian | Mean | 397.51 | 376.54 | 364.14 | 328.85 | 362.10 | 369.52 | 382.37 | 367.04 |
|  |  | N | 353 | 330 | 390 | 346 | 298 | 317 | 290 | 238 |
|  | Non-Hispanic Multi-racial | Mean | 396.48 | 377.48 | 362.11 | 317.74 | 357.52 | 365.48 | 381.73 | 360.12 |
|  |  | N | 80 | 60 | 82 | 68 | 62 | 54 | 55 | 41 |
|  | Non-Hispanic White | Mean | 408.47 | 386.04 | 365.96 | 329.03 | 369.57 | 374.89 | 392.66 | 374.76 |
|  |  | N | 2,997 | 2,451 | 3,066 | 2,708 | 2,399 | 2,264 | 2,197 | 1,665 |
|  | Unknown | Mean | 384.54 | 374.53 | 356.88 | 316.91 | 350.58 | 363.07 | 377.13 | 358.47 |
|  |  | N | 951 | 758 | 1,010 | 894 | 737 | 676 | 639 | 456 |

### 3.2.2.3 Correlations Among Scale Scores by Cluster

Table 3.2.2.3A
Correlations Among Scale Scores: 1 S401 Online

|  |  | Listening | Reading | Writing | Speaking |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Listening | Pearson Correlation | 1 | .405 | .520 | .558 |
|  | N | 155,365 | 144,040 | 155,320 | 125,541 |
| Reading | Pearson Correlation |  | 1 | .540 | .367 |
|  | N |  | 155,144 | 155,093 | 125,283 |
| Writing | Pearson Correlation |  |  | 1 | .456 |
|  | N |  |  | 170,761 | 135,492 |
| Speaking | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 135,531 |

Table 3.2.2.3B
Correlations Among Scale Scores: 2-3 S401 Online

| Listening |  | Listening | Reading | Writing | Speaking |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | .581 | .598 | .602 |
| Reading | Pearson Correlation |  | 1 | 1 | .637 |
|  | N |  | 309,909 | 309,803 | 258,985 |
| Writing | Pearson Correlation |  |  | 1 | .538 |
|  | N |  |  | 356,941 | 293,549 |
| Speaking | Pearson Correlation |  |  | 1 |  |
|  | N |  |  | 293,645 |  |

Table 3.2.2.3C
Correlations Among Scale Scores: 4-5 S401 Online

| Listening | Pearson Correlation | Listening | Reading | Writing | Speaking |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | N | 1 | .629 | .658 | .641 |
| Reading | Pearson Correlation |  | 154,642 | 127,567 | 159,860 |
|  | N |  | 1 | .624 | .531 |
| Writing | Pearson Correlation |  |  | 113,502 | 139,947 |
|  | N |  |  | 140,492 | 117,092 |
| Speaking | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 175,273 |

Table 3.2.2.3D
Correlations Among Scale Scores: 6-8 S401 Online

|  |  | Listening | Reading | Writing | Speaking |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Listening | Pearson Correlation | 1 | .716 | .693 | .672 |
|  | N | 195,700 | 160,625 | 178,705 | 153,304 |
| Reading | Pearson Correlation |  | 1 | .756 | .595 |
|  | N |  | 186,282 | 171,107 | 145,221 |
| Writing | Pearson Correlation |  |  | 1 | .664 |
|  | N |  |  | 210,933 | 164,267 |
| Speaking | Pearson Correlation |  |  |  | 1 |
|  | N |  |  |  | 180,383 |

Table 3.2.2.3E
Correlations Among Scale Scores: 9-12 S401 Online

|  |  | Listening | Reading | Writing | Speaking |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Listening | Pearson Correlation | 1 | .711 | .599 | .638 |
|  | N | 208,374 | 147,903 | 186,112 | 165,040 |
| Reading | Pearson Correlation |  | 1 | .680 | .635 |
|  | N |  | 170,634 | 155,139 | 134,311 |
| Writing | Pearson Correlation |  |  | 1 | .660 |
|  | N |  |  | 217,998 | 171,946 |
| Speaking | Pearson Correlation |  |  | 1 |  |
|  | N |  |  | 193,155 |  |

### 3.2.3 Proficiency Level Results

### 3.2.3.1 Domains

### 3.2.3.1.1 Listening

### 3.2.3.1.1.1 By Cluster

Table 3.2.3.1.1.1A
Proficiency Level by Cluster (Count): Listening S401 Online

| Cluster | Listening Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | Total |
| 1 | 14,848 | 6,303 | 18,398 | 7,375 | 15,320 | 93,121 | 155,365 |
| $2-3$ | 24,222 | 35,278 | 54,872 | 19,669 | 45,956 | 138,302 | 318,299 |
| $4-5$ | 9,110 | 4,980 | 17,995 | 12,174 | 16,090 | 131,222 | 191,571 |
| $6-8$ | 12,371 | 16,012 | 43,028 | 27,074 | 23,773 | 73,442 | 195,700 |
| $9-12$ | 29,600 | 31,734 | 56,018 | 32,949 | 18,683 | 39,390 | 208,374 |

Table 3.2.3.1.1.1B
Proficiency Level by Cluster (Percent): Listening S401 Online

| Cluster | Listening Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | Total |
| 1 | $9.6 \%$ | $4.1 \%$ | $11.8 \%$ | $4.7 \%$ | $9.9 \%$ | $59.9 \%$ | $100.0 \%$ |
| $2-3$ | $7.6 \%$ | $11.1 \%$ | $17.2 \%$ | $6.2 \%$ | $14.4 \%$ | $43.5 \%$ | $100.0 \%$ |
| $4-5$ | $4.8 \%$ | $2.6 \%$ | $9.4 \%$ | $6.4 \%$ | $8.4 \%$ | $68.5 \%$ | $100.0 \%$ |
| $6-8$ | $6.3 \%$ | $8.2 \%$ | $22.0 \%$ | $13.8 \%$ | $12.1 \%$ | $37.5 \%$ | $100.0 \%$ |
| $9-12$ | $14.2 \%$ | $15.2 \%$ | $26.9 \%$ | $15.8 \%$ | $9.0 \%$ | $18.9 \%$ | $100.0 \%$ |

### 3.2.3.1.1.2 By Grade

Table 3.2.3.1.1.2A
Proficiency Level by Grade (Count): Listening S401 Online

| Grade | Listening Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 14,848 | 6,303 | 18,398 | 7,375 | 15,320 | 93,121 | 155,365 |
| 2 | 11,676 | 18,369 | 31,542 | 9,072 | 19,280 | 64,737 | 154,676 |
| 3 | 12,546 | 16,909 | 23,330 | 10,597 | 26,676 | 73,565 | 163,623 |
| 4 | 4,283 | 2,745 | 8,242 | 6,807 | 10,564 | 77,378 | 110,019 |
| 5 | 4,827 | 2,235 | 9,753 | 5,367 | 5,526 | 53,844 | 81,552 |
| 6 | 2,970 | 4,956 | 16,003 | 9,655 | 10,069 | 21,167 | 64,820 |
| 7 | 4,072 | 5,190 | 14,059 | 9,647 | 8,383 | 24,362 | 65,713 |
| 8 | 5,329 | 5,866 | 12,966 | 7,772 | 5,321 | 27,913 | 65,167 |
| 9 | 9,317 | 14,482 | 20,420 | 14,120 | 7,263 | 18,059 | 83,661 |
| 10 | 8,310 | 7,585 | 16,594 | 8,533 | 5,365 | 10,201 | 56,588 |
| 11 | 6,775 | 5,486 | 11,268 | 6,669 | 3,238 | 6,902 | 40,338 |
| 12 | 5,198 | 4,181 | 7,736 | 3,627 | 2,817 | 4,228 | 27,787 |

Table 3.2.3.1.1.2B
Proficiency Level by Grade (Percent): Listening S401 Online

| Grade | Listening Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $9.6 \%$ | $4.1 \%$ | $11.8 \%$ | $4.7 \%$ | $9.9 \%$ | $59.9 \%$ | $100.0 \%$ |
| 2 | $7.5 \%$ | $11.9 \%$ | $20.4 \%$ | $5.9 \%$ | $12.5 \%$ | $41.9 \%$ | $100.0 \%$ |
| 3 | $7.7 \%$ | $10.3 \%$ | $14.3 \%$ | $6.5 \%$ | $16.3 \%$ | $45.0 \%$ | $100.0 \%$ |
| 4 | $3.9 \%$ | $2.5 \%$ | $7.5 \%$ | $6.2 \%$ | $9.6 \%$ | $70.3 \%$ | $100.0 \%$ |
| 5 | $5.9 \%$ | $2.7 \%$ | $12.0 \%$ | $6.6 \%$ | $6.8 \%$ | $66.0 \%$ | $100.0 \%$ |
| 6 | $4.6 \%$ | $7.6 \%$ | $24.7 \%$ | $14.9 \%$ | $15.5 \%$ | $32.7 \%$ | $100.0 \%$ |
| 7 | $6.2 \%$ | $7.9 \%$ | $21.4 \%$ | $14.7 \%$ | $12.8 \%$ | $37.1 \%$ | $100.0 \%$ |
| 8 | $8.2 \%$ | $9.0 \%$ | $19.9 \%$ | $11.9 \%$ | $8.2 \%$ | $42.8 \%$ | $100.0 \%$ |
| 9 | $11.1 \%$ | $17.3 \%$ | $24.4 \%$ | $16.9 \%$ | $8.7 \%$ | $21.6 \%$ | $100.0 \%$ |
| 10 | $14.7 \%$ | $13.4 \%$ | $29.3 \%$ | $15.1 \%$ | $9.5 \%$ | $18.0 \%$ | $100.0 \%$ |
| 11 | $16.8 \%$ | $13.6 \%$ | $27.9 \%$ | $16.5 \%$ | $8.0 \%$ | $17.1 \%$ | $100.0 \%$ |
| 12 | $18.7 \%$ | $15.0 \%$ | $27.8 \%$ | $13.1 \%$ | $10.1 \%$ | $15.2 \%$ | $100.0 \%$ |

### 3.2.3.1.2 Reading

3.2.3.1.2.1 By Cluster

Table 3.2.3.1.2.1A
Proficiency Level by Cluster (Count): Reading S401 Online

| Cluster | Reading Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 30,402 | 41,364 | 30,896 | 13,856 | 23,014 | 15,612 | 155,144 |
| $2-3$ | 59,494 | 89,846 | 53,280 | 23,423 | 37,396 | 46,470 | 309,909 |
| $4-5$ | 23,409 | 46,554 | 36,178 | 18,918 | 27,982 | 15,424 | 168,465 |
| $6-8$ | 66,900 | 49,195 | 33,131 | 9,636 | 15,085 | 12,335 | 186,282 |
| $9-12$ | 49,277 | 54,319 | 25,905 | 7,720 | 17,565 | 15,848 | 170,634 |

Table 3.2.3.1.2.1B
Proficiency Level by Cluster (Percent): Reading S401 Online

| Cluster | Reading Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
|  | $19.6 \%$ | $26.7 \%$ | $19.9 \%$ | $8.9 \%$ | $14.8 \%$ | $10.1 \%$ | $100.0 \%$ |
| $2-3$ | $19.2 \%$ | $29.0 \%$ | $17.2 \%$ | $7.6 \%$ | $12.1 \%$ | $15.0 \%$ | $100.0 \%$ |
| $4-5$ | $13.9 \%$ | $27.6 \%$ | $21.5 \%$ | $11.2 \%$ | $16.6 \%$ | $9.2 \%$ | $100.0 \%$ |
| $6-8$ | $35.9 \%$ | $26.4 \%$ | $17.8 \%$ | $5.2 \%$ | $8.1 \%$ | $6.6 \%$ | $100.0 \%$ |
| $9-12$ | $28.9 \%$ | $31.8 \%$ | $15.2 \%$ | $4.5 \%$ | $10.3 \%$ | $9.3 \%$ | $100.0 \%$ |

### 3.2.3.1.2.2 By Grade

Table 3.2.3.1.2.2A
Proficiency Level by Grade (Count): Reading S401 Online

| Grade | Reading Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | Total |
| 1 | 30,402 | 41,364 | 30,896 | 13,856 | 23,014 | 15,612 | 155,144 |
| 2 | 26,869 | 46,113 | 30,183 | 12,658 | 17,874 | 17,522 | 151,219 |
| 3 | 32,625 | 43,733 | 23,097 | 10,765 | 19,522 | 28,948 | 158,690 |
| 4 | 10,541 | 24,448 | 24,091 | 11,395 | 18,532 | 8,302 | 97,309 |
| 5 | 12,868 | 22,106 | 12,087 | 7,523 | 9,450 | 7,122 | 71,156 |
| 6 | 22,705 | 17,132 | 12,098 | 3,259 | 5,099 | 2,557 | 62,850 |
| 7 | 22,501 | 16,676 | 11,042 | 3,457 | 4,627 | 4,236 | 62,539 |
| 8 | 21,694 | 15,387 | 9,991 | 2,920 | 5,359 | 5,542 | 60,893 |
| 9 | 22,281 | 20,143 | 10,934 | 2,869 | 7,259 | 7,139 | 70,625 |
| 10 | 13,198 | 14,476 | 7,615 | 2,076 | 4,700 | 4,018 | 46,083 |
| 11 | 8,409 | 11,073 | 4,246 | 1,620 | 3,287 | 2,854 | 31,489 |
| 12 | 5,389 | 8,627 | 3,110 | 1,155 | 2,319 | 1,837 | 22,437 |

Table 3.2.3.1.2.2B
Proficiency Level by Grade (Percent): Reading S401 Online

| Grade | Reading Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $19.6 \%$ | $26.7 \%$ | $19.9 \%$ | $8.9 \%$ | $14.8 \%$ | $10.1 \%$ | $100.0 \%$ |
| 2 | $17.8 \%$ | $30.5 \%$ | $20.0 \%$ | $8.4 \%$ | $11.8 \%$ | $11.6 \%$ | $100.0 \%$ |
| 3 | $20.6 \%$ | $27.6 \%$ | $14.6 \%$ | $6.8 \%$ | $12.3 \%$ | $18.2 \%$ | $100.0 \%$ |
| 4 | $10.8 \%$ | $25.1 \%$ | $24.8 \%$ | $11.7 \%$ | $19.0 \%$ | $8.5 \%$ | $100.0 \%$ |
| 5 | $18.1 \%$ | $31.1 \%$ | $17.0 \%$ | $10.6 \%$ | $13.3 \%$ | $10.0 \%$ | $100.0 \%$ |
| 6 | $36.1 \%$ | $27.3 \%$ | $19.2 \%$ | $5.2 \%$ | $8.1 \%$ | $4.1 \%$ | $100.0 \%$ |
| 7 | $36.0 \%$ | $26.7 \%$ | $17.7 \%$ | $5.5 \%$ | $7.4 \%$ | $6.8 \%$ | $100.0 \%$ |
| 8 | $35.6 \%$ | $25.3 \%$ | $16.4 \%$ | $4.8 \%$ | $8.8 \%$ | $9.1 \%$ | $100.0 \%$ |
| 9 | $31.5 \%$ | $28.5 \%$ | $15.5 \%$ | $4.1 \%$ | $10.3 \%$ | $10.1 \%$ | $100.0 \%$ |
| 10 | $28.6 \%$ | $31.4 \%$ | $16.5 \%$ | $4.5 \%$ | $10.2 \%$ | $8.7 \%$ | $100.0 \%$ |
| 11 | $26.7 \%$ | $35.2 \%$ | $13.5 \%$ | $5.1 \%$ | $10.4 \%$ | $9.1 \%$ | $100.0 \%$ |
| 12 | $24.0 \%$ | $38.4 \%$ | $13.9 \%$ | $5.1 \%$ | $10.3 \%$ | $8.2 \%$ | $100.0 \%$ |

### 3.2.3.1.3 Writing

3.2.3.1.3.1 By Cluster

Table 3.2.3.1.3.1A
Proficiency Level by Cluster (Count): Writing S401 Online

| Cluster | Writing Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 29,580 | 100,575 | 37,390 | 3,166 | 48 | 2 | 170,761 |
| $2-3$ | 16,999 | 56,087 | 220,527 | 62,678 | 626 | 24 | 356,941 |
| $4-5$ | 8,617 | 10,669 | 85,241 | 34,560 | 1,320 | 85 | 140,492 |
| $6-8$ | 25,201 | 37,563 | 106,224 | 41,488 | 437 | 20 | 210,933 |
| $9-12$ | 23,399 | 38,285 | 107,066 | 47,910 | 1,322 | 16 | 217,998 |

Table 3.2.3.1.3.1B
Proficiency Level by Cluster (Percent): Writing S401 Online

| Cluster | Writing Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
|  | $17.3 \%$ | $58.9 \%$ | $21.9 \%$ | $1.9 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| $2-3$ | $4.8 \%$ | $15.7 \%$ | $61.8 \%$ | $17.6 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| $4-5$ | $6.1 \%$ | $7.6 \%$ | $60.7 \%$ | $24.6 \%$ | $0.9 \%$ | $0.1 \%$ | $100.0 \%$ |
| $6-8$ | $11.9 \%$ | $17.8 \%$ | $50.4 \%$ | $19.7 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| $9-12$ | $10.7 \%$ | $17.6 \%$ | $49.1 \%$ | $22.0 \%$ | $0.6 \%$ | $0.0 \%$ | $100.0 \%$ |

### 3.2.3.1.3.2 By Grade

Table 3.2.3.1.3.2A
Proficiency Level by Grade (Count): Writing S401 Online

| Grade | Writing Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 29,580 | 100,575 | 37,390 | 3,166 | 48 | 2 | 170,761 |
| 2 | 10,144 | 31,959 | 112,559 | 19,619 | 110 | 2 | 174,393 |
| 3 | 6,855 | 24,128 | 107,968 | 43,059 | 516 | 22 | 182,548 |
| 4 | 5,016 | 6,077 | 50,373 | 17,947 | 773 | 38 | 80,224 |
| 5 | 3,601 | 4,592 | 34,868 | 16,613 | 547 | 47 | 60,268 |
| 6 | 7,507 | 12,290 | 36,910 | 12,396 | 97 | 2 | 69,202 |
| 7 | 8,310 | 14,410 | 34,535 | 13,917 | 133 | 7 | 71,312 |
| 8 | 9,384 | 10,863 | 34,779 | 15,175 | 207 | 11 | 70,419 |
| 9 | 10,502 | 14,309 | 39,237 | 22,847 | 833 | 6 | 87,734 |
| 10 | 4,839 | 11,581 | 29,853 | 12,654 | 232 | 7 | 59,166 |
| 11 | 4,098 | 8,392 | 20,999 | 8,423 | 165 | 2 | 42,079 |
| 12 | 3,960 | 4,003 | 16,977 | 3,986 | 92 | 1 | 29,019 |

Table 3.2.3.1.3.2B
Proficiency Level by Grade (Percent): Writing S401 Online

| Grade | Writing Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $17.3 \%$ | $58.9 \%$ | $21.9 \%$ | $1.9 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| 2 | $5.8 \%$ | $18.3 \%$ | $64.5 \%$ | $11.2 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| 3 | $3.8 \%$ | $13.2 \%$ | $59.1 \%$ | $23.6 \%$ | $0.3 \%$ | $0.0 \%$ | $100.0 \%$ |
| 4 | $6.3 \%$ | $7.6 \%$ | $62.8 \%$ | $22.4 \%$ | $1.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| 5 | $6.0 \%$ | $7.6 \%$ | $57.9 \%$ | $27.6 \%$ | $0.9 \%$ | $0.1 \%$ | $100.0 \%$ |
| 6 | $10.8 \%$ | $17.8 \%$ | $53.3 \%$ | $17.9 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| 7 | $11.7 \%$ | $20.2 \%$ | $48.4 \%$ | $19.5 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| 8 | $13.3 \%$ | $15.4 \%$ | $49.4 \%$ | $21.5 \%$ | $0.3 \%$ | $0.0 \%$ | $100.0 \%$ |
| 9 | $12.0 \%$ | $16.3 \%$ | $44.7 \%$ | $26.0 \%$ | $0.9 \%$ | $0.0 \%$ | $100.0 \%$ |
| 10 | $8.2 \%$ | $19.6 \%$ | $50.5 \%$ | $21.4 \%$ | $0.4 \%$ | $0.0 \%$ | $100.0 \%$ |
| 11 | $9.7 \%$ | $19.9 \%$ | $49.9 \%$ | $20.0 \%$ | $0.4 \%$ | $0.0 \%$ | $100.0 \%$ |
| 12 | $13.6 \%$ | $13.8 \%$ | $58.5 \%$ | $13.7 \%$ | $0.3 \%$ | $0.0 \%$ | $100.0 \%$ |

### 3.2.3.1.4 Speaking

3.2.3.1.4.1 By Cluster

Table 3.2.3.1.4.1A
Proficiency Level by Cluster (Count): Speaking S401 Online

| Cluster | Speaking Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 10,249 | 40,120 | 62,464 | 22,404 | 291 | 3 | 135,531 |
| $2-3$ | 35,624 | 103,841 | 138,660 | 15,104 | 374 | 42 | 293,645 |
| $4-5$ | 13,729 | 35,789 | 82,428 | 42,134 | 1,158 | 35 | 175,273 |
| $6-8$ | 20,894 | 45,753 | 94,748 | 18,621 | 353 | 14 | 180,383 |
| $9-12$ | 53,615 | 62,630 | 74,024 | 2,825 | 55 | 6 | 193,155 |

Table 3.2.3.1.4.1B
Proficiency Level by Cluster (Percent): Speaking S401 Online

| Cluster | Speaking Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $7.6 \%$ | $29.6 \%$ | $46.1 \%$ | $16.5 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| $2-3$ | $12.1 \%$ | $35.4 \%$ | $47.2 \%$ | $5.1 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| $4-5$ | $7.8 \%$ | $20.4 \%$ | $47.0 \%$ | $24.0 \%$ | $0.7 \%$ | $0.0 \%$ | $100.0 \%$ |
| $6-8$ | $11.6 \%$ | $25.4 \%$ | $52.5 \%$ | $10.3 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| $9-12$ | $27.8 \%$ | $32.4 \%$ | $38.3 \%$ | $1.5 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |

### 3.2.3.1.4.2 By Grade

Table 3.2.3.1.4.2A
Proficiency Level by Grade (Count): Speaking S401 Online

| Grade | Speaking Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 10,249 | 40,120 | 62,464 | 22,404 | 291 | 3 | 135,531 |
| 2 | 15,225 | 51,119 | 68,014 | 6,606 | 200 | 6 | 141,170 |
| 3 | 20,399 | 52,722 | 70,646 | 8,498 | 174 | 36 | 152,475 |
| 4 | 5,736 | 20,247 | 46,894 | 26,862 | 776 | 24 | 100,539 |
| 5 | 7,993 | 15,542 | 35,534 | 15,272 | 382 | 11 | 74,734 |
| 6 | 5,346 | 17,855 | 31,603 | 5,199 | 100 | 0 | 60,103 |
| 7 | 6,095 | 14,481 | 31,805 | 7,424 | 76 | 6 | 59,887 |
| 8 | 9,453 | 13,417 | 31,340 | 5,998 | 177 | 8 | 60,393 |
| 9 | 24,507 | 22,236 | 29,971 | 1,401 | 37 | 3 | 78,155 |
| 10 | 14,581 | 15,892 | 20,960 | 911 | 6 | 0 | 52,350 |
| 11 | 8,448 | 12,224 | 15,422 | 363 | 10 | 2 | 36,469 |
| 12 | 6,079 | 12,278 | 7,671 | 150 | 2 | 1 | 26,181 |

Table 3.2.3.1.4.2B
Proficiency Level by Grade (Percent): Speaking S401 Online

| Grade | Speaking Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $7.6 \%$ | $29.6 \%$ | $46.1 \%$ | $16.5 \%$ | $0.2 \%$ | $00.0 \%$ |  |
| 2 | $10.8 \%$ | $36.2 \%$ | $48.2 \%$ | $4.7 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| 3 | $13.4 \%$ | $34.6 \%$ | $46.3 \%$ | $5.6 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| 4 | $5.7 \%$ | $20.1 \%$ | $46.6 \%$ | $26.7 \%$ | $0.8 \%$ | $0.0 \%$ | $100.0 \%$ |
| 5 | $10.7 \%$ | $20.8 \%$ | $47.5 \%$ | $20.4 \%$ | $0.5 \%$ | $0.0 \%$ | $100.0 \%$ |
| 6 | $8.9 \%$ | $29.7 \%$ | $52.6 \%$ | $8.7 \%$ | $0.2 \%$ | $0.0 \%$ | $100.0 \%$ |
| 7 | $10.2 \%$ | $24.2 \%$ | $53.1 \%$ | $12.4 \%$ | $0.1 \%$ | $0.0 \%$ | $100.0 \%$ |
| 8 | $15.7 \%$ | $22.2 \%$ | $51.9 \%$ | $9.9 \%$ | $0.3 \%$ | $0.0 \%$ | $100.0 \%$ |
| 9 | $31.4 \%$ | $28.5 \%$ | $38.3 \%$ | $1.8 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| 10 | $27.9 \%$ | $30.4 \%$ | $40.0 \%$ | $1.7 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| 11 | $23.2 \%$ | $33.5 \%$ | $42.3 \%$ | $1.0 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |
| 12 | $23.2 \%$ | $46.9 \%$ | $29.3 \%$ | $0.6 \%$ | $0.0 \%$ | $0.0 \%$ | $100.0 \%$ |

### 3.2.3.2 Composites

### 3.2.3.2.1 Oral Composite

### 3.2.3.2.1.1 By Cluster

Table 3.2.3.2.1.1A
Proficiency Level by Cluster (Count): Oral S401 Online

| Cluster | Oral Language Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 8,309 | 11,441 | 33,020 | 38,477 | 29,593 | 4,701 | 125,541 |
| $2-3$ | 19,974 | 46,583 | 93,975 | 85,814 | 19,228 | 927 | 266,501 |
| $4-5$ | 7,940 | 10,210 | 28,173 | 66,802 | 38,828 | 7,907 | 159,860 |
| $6-8$ | 11,689 | 19,318 | 53,017 | 51,824 | 14,889 | 2,567 | 153,304 |
| $9-12$ | 27,676 | 38,863 | 65,777 | 28,598 | 3,833 | 293 | 165,040 |

Table 3.2.3.2.1.1B
Proficiency Level by Cluster (Percent): Oral S401 Online

| Cluster | Oral Language Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $6.6 \%$ | $9.1 \%$ | $26.3 \%$ | $30.6 \%$ | $23.6 \%$ | $3.7 \%$ | $100.0 \%$ |
| $2-3$ | $7.5 \%$ | $17.5 \%$ | $35.3 \%$ | $32.2 \%$ | $7.2 \%$ | $0.3 \%$ | $100.0 \%$ |
| $4-5$ | $5.0 \%$ | $6.4 \%$ | $17.6 \%$ | $41.8 \%$ | $24.3 \%$ | $4.9 \%$ | $100.0 \%$ |
| $6-8$ | $7.6 \%$ | $12.6 \%$ | $34.6 \%$ | $33.8 \%$ | $9.7 \%$ | $1.7 \%$ | $100.0 \%$ |
| $9-12$ | $16.8 \%$ | $23.5 \%$ | $39.9 \%$ | $17.3 \%$ | $2.3 \%$ | $0.2 \%$ | $100.0 \%$ |

### 3.2.3.2.1.2 By Grade

Table 3.2.3.2.1.2A
Proficiency Level by Grade (Count): Oral S401 Online

| Grade | Oral Language Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 8,309 | 11,441 | 33,020 | 38,477 | 29,593 | 4,701 | 125,541 |
| 2 | 9,129 | 24,958 | 48,031 | 35,823 | 9,220 | 447 | 127,608 |
| 3 | 10,845 | 21,625 | 45,944 | 49,991 | 10,008 | 480 | 138,893 |
| 4 | 3,587 | 5,271 | 15,479 | 40,131 | 22,355 | 4,587 | 91,410 |
| 5 | 4,353 | 4,939 | 12,694 | 26,671 | 16,473 | 3,320 | 68,450 |
| 6 | 2,796 | 6,186 | 19,901 | 17,276 | 4,446 | 659 | 51,264 |
| 7 | 3,887 | 6,447 | 17,402 | 16,817 | 5,372 | 853 | 50,778 |
| 8 | 5,006 | 6,685 | 15,714 | 17,731 | 5,071 | 1,055 | 51,262 |
| 9 | 10,582 | 14,826 | 24,873 | 14,030 | 2,043 | 179 | 66,533 |
| 10 | 7,511 | 10,501 | 18,379 | 7,429 | 901 | 56 | 44,777 |
| 11 | 5,628 | 7,895 | 12,849 | 4,314 | 575 | 33 | 31,294 |
| 12 | 3,955 | 5,641 | 9,676 | 2,825 | 314 | 25 | 22,436 |

Table 3.2.3.2.1.2B
Proficiency Level by Grade (Percent): Oral S401 Online

| Grade | Oral Language Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $6.6 \%$ | $9.1 \%$ | $26.3 \%$ | $30.6 \%$ | $23.6 \%$ | $3.7 \%$ | $100.0 \%$ |
| 2 | $7.2 \%$ | $19.6 \%$ | $37.6 \%$ | $28.1 \%$ | $7.2 \%$ | $0.4 \%$ | $100.0 \%$ |
| 3 | $7.8 \%$ | $15.6 \%$ | $33.1 \%$ | $36.0 \%$ | $7.2 \%$ | $0.3 \%$ | $100.0 \%$ |
| 4 | $3.9 \%$ | $5.8 \%$ | $16.9 \%$ | $43.9 \%$ | $24.5 \%$ | $5.0 \%$ | $100.0 \%$ |
| 5 | $6.4 \%$ | $7.2 \%$ | $18.5 \%$ | $39.0 \%$ | $24.1 \%$ | $4.9 \%$ | $100.0 \%$ |
| 6 | $5.5 \%$ | $12.1 \%$ | $38.8 \%$ | $33.7 \%$ | $8.7 \%$ | $1.3 \%$ | $100.0 \%$ |
| 7 | $7.7 \%$ | $12.7 \%$ | $34.3 \%$ | $33.1 \%$ | $10.6 \%$ | $1.7 \%$ | $100.0 \%$ |
| 8 | $9.8 \%$ | $13.0 \%$ | $30.7 \%$ | $34.6 \%$ | $9.9 \%$ | $2.1 \%$ | $100.0 \%$ |
| 9 | $15.9 \%$ | $22.3 \%$ | $37.4 \%$ | $21.1 \%$ | $3.1 \%$ | $0.3 \%$ | $100.0 \%$ |
| 10 | $16.8 \%$ | $23.5 \%$ | $41.0 \%$ | $16.6 \%$ | $2.0 \%$ | $0.1 \%$ | $100.0 \%$ |
| 11 | $18.0 \%$ | $25.2 \%$ | $41.1 \%$ | $13.8 \%$ | $1.8 \%$ | $0.1 \%$ | $100.0 \%$ |
| 12 | $17.6 \%$ | $25.1 \%$ | $43.1 \%$ | $12.6 \%$ | $1.4 \%$ | $0.1 \%$ | $100.0 \%$ |

### 3.2.3.2.2 Literacy Composite

### 3.2.3.2.2.1 By Cluster

Table 3.2.3.2.2.1A
Proficiency Level by Cluster (Count): Literacy S401 Online

| Cluster | Literacy Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 27,564 | 70,291 | 43,314 | 11,308 | 2,186 | 430 | 155,093 |
| $2-3$ | 24,497 | 74,423 | 135,122 | 63,658 | 11,191 | 912 | 309,803 |
| $4-5$ | 10,452 | 15,532 | 55,071 | 27,576 | 4,069 | 802 | 113,502 |
| $6-8$ | 37,992 | 39,931 | 63,037 | 26,839 | 3,057 | 251 | 171,107 |
| $9-12$ | 26,494 | 40,000 | 57,945 | 24,760 | 5,522 | 418 | 155,139 |

Table 3.2.3.2.2.1B
Proficiency Level by Cluster (Percent): Literacy S401 Online

| Cluster | Literacy Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $17.8 \%$ | $45.3 \%$ | $27.9 \%$ | $7.3 \%$ | $1.4 \%$ | $0.3 \%$ | $100.0 \%$ |
| $2-3$ | $7.9 \%$ | $24.0 \%$ | $43.6 \%$ | $20.5 \%$ | $3.6 \%$ | $0.3 \%$ | $100.0 \%$ |
| $4-5$ | $9.2 \%$ | $13.7 \%$ | $48.5 \%$ | $24.3 \%$ | $3.6 \%$ | $0.7 \%$ | $100.0 \%$ |
| $6-8$ | $22.2 \%$ | $23.3 \%$ | $36.8 \%$ | $15.7 \%$ | $1.8 \%$ | $0.1 \%$ | $100.0 \%$ |
| $9-12$ | $17.1 \%$ | $25.8 \%$ | $37.4 \%$ | $16.0 \%$ | $3.6 \%$ | $0.3 \%$ | $100.0 \%$ |

### 3.2.3.2.2.2 By Grade

Table 3.2.3.2.2.2A
Proficiency Level by Grade (Count): Literacy S401 Online

| Grade | Literacy Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 27,564 | 70,291 | 43,314 | 11,308 | 2,186 | 430 | 155,093 |
| 2 | 12,286 | 43,248 | 67,839 | 24,235 | 3,255 | 290 | 151,153 |
| 3 | 12,211 | 31,175 | 67,283 | 39,423 | 7,936 | 622 | 158,650 |
| 4 | 5,322 | 8,333 | 33,546 | 15,455 | 1,851 | 401 | 64,908 |
| 5 | 5,130 | 7,199 | 21,525 | 12,121 | 2,218 | 401 | 48,594 |
| 6 | 12,373 | 13,345 | 23,379 | 7,478 | 616 | 50 | 57,241 |
| 7 | 12,836 | 13,694 | 21,205 | 8,863 | 1,063 | 99 | 57,760 |
| 8 | 12,783 | 12,892 | 18,453 | 10,498 | 1,378 | 102 | 56,106 |
| 9 | 12,601 | 14,283 | 22,700 | 11,486 | 2,669 | 258 | 63,997 |
| 10 | 6,465 | 11,008 | 16,451 | 6,565 | 1,327 | 93 | 41,909 |
| 11 | 4,280 | 8,371 | 10,835 | 4,233 | 897 | 52 | 28,668 |
| 12 | 3,148 | 6,338 | 7,959 | 2,476 | 629 | 15 | 20,565 |

Table 3.2.3.2.2.2B
Proficiency Level by Grade (Percent): Literacy S401 Online

| Grade | Literacy Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $17.8 \%$ | $45.3 \%$ | $27.9 \%$ | $7.3 \%$ | $1.4 \%$ | $0.0 \%$ |  |
| 2 | $8.1 \%$ | $28.6 \%$ | $44.9 \%$ | $16.0 \%$ | $2.2 \%$ | $0.2 \%$ | $100.0 \%$ |
| 3 | $7.7 \%$ | $19.7 \%$ | $42.4 \%$ | $24.8 \%$ | $5.0 \%$ | $0.4 \%$ | $100.0 \%$ |
| 4 | $8.2 \%$ | $12.8 \%$ | $51.7 \%$ | $23.8 \%$ | $2.9 \%$ | $0.6 \%$ | $100.0 \%$ |
| 5 | $10.6 \%$ | $14.8 \%$ | $44.3 \%$ | $24.9 \%$ | $4.6 \%$ | $0.8 \%$ | $100.0 \%$ |
| 6 | $21.6 \%$ | $23.3 \%$ | $40.8 \%$ | $13.1 \%$ | $1.1 \%$ | $0.1 \%$ | $100.0 \%$ |
| 7 | $22.2 \%$ | $23.7 \%$ | $36.7 \%$ | $15.3 \%$ | $1.8 \%$ | $0.2 \%$ | $100.0 \%$ |
| 8 | $22.8 \%$ | $23.0 \%$ | $32.9 \%$ | $18.7 \%$ | $2.5 \%$ | $0.2 \%$ | $100.0 \%$ |
| 9 | $19.7 \%$ | $22.3 \%$ | $35.5 \%$ | $17.9 \%$ | $4.2 \%$ | $0.4 \%$ | $100.0 \%$ |
| 10 | $15.4 \%$ | $26.3 \%$ | $39.3 \%$ | $15.7 \%$ | $3.2 \%$ | $0.2 \%$ | $100.0 \%$ |
| 11 | $14.9 \%$ | $29.2 \%$ | $37.8 \%$ | $14.8 \%$ | $3.1 \%$ | $0.2 \%$ | $100.0 \%$ |
| 12 | $15.3 \%$ | $30.8 \%$ | $38.7 \%$ | $12.0 \%$ | $3.1 \%$ | $0.1 \%$ | $100.0 \%$ |

### 3.2.3.2.3 Comprehension Composite

### 3.2.3.2.3.1 By Cluster

Table 3.2.3.2.3.1A
Proficiency Level by Cluster (Count): Comprehension S401 Online

| Cluster | Comprehension Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 14,004 | 22,852 | 30,988 | 19,230 | 27,932 | 29,034 | 144,040 |
| $2-3$ | 29,538 | 58,801 | 65,092 | 33,997 | 44,225 | 51,204 | 282,857 |
| $4-5$ | 9,916 | 19,473 | 29,182 | 23,798 | 34,903 | 37,370 | 154,642 |
| $6-8$ | 32,204 | 37,239 | 33,116 | 19,486 | 19,582 | 18,998 | 160,625 |
| $9-12$ | 31,162 | 42,685 | 30,206 | 13,955 | 14,402 | 15,493 | 147,903 |

Table 3.2.3.2.3.1B
Proficiency Level by Cluster (Percent): Comprehension S401 Online

| Cluster | Comprehension Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $9.7 \%$ | $15.9 \%$ | $21.5 \%$ | $13.4 \%$ | $19.4 \%$ | $20.2 \%$ | $100.0 \%$ |
| $2-3$ | $10.4 \%$ | $20.8 \%$ | $23.0 \%$ | $12.0 \%$ | $15.6 \%$ | $18.1 \%$ | $100.0 \%$ |
| $4-5$ | $6.4 \%$ | $12.6 \%$ | $18.9 \%$ | $15.4 \%$ | $22.6 \%$ | $24.2 \%$ | $100.0 \%$ |
| $6-8$ | $20.0 \%$ | $23.2 \%$ | $20.6 \%$ | $12.1 \%$ | $12.2 \%$ | $11.8 \%$ | $100.0 \%$ |
| $9-12$ | $21.1 \%$ | $28.9 \%$ | $20.4 \%$ | $9.4 \%$ | $9.7 \%$ | $10.5 \%$ | $100.0 \%$ |

### 3.2.3.2.3.2 By Grade

Table 3.2.3.2.3.2A
Proficiency Level by Grade (Count): Comprehension S401 Online

| Grade | Comprehension Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | Total |
| 1 | 14,004 | 22,852 | 30,988 | 19,230 | 27,932 | 29,034 | 144,040 |
| 2 | 13,013 | 31,566 | 33,782 | 17,960 | 20,608 | 20,442 | 137,371 |
| 3 | 16,525 | 27,235 | 31,310 | 16,037 | 23,617 | 30,762 | 145,486 |
| 4 | 3,981 | 10,076 | 17,143 | 14,387 | 21,942 | 21,559 | 89,088 |
| 5 | 5,935 | 9,397 | 12,039 | 9,411 | 12,961 | 15,811 | 65,554 |
| 6 | 9,904 | 13,817 | 12,648 | 6,844 | 6,187 | 4,715 | 54,115 |
| 7 | 11,049 | 12,248 | 11,044 | 6,647 | 6,399 | 6,532 | 53,919 |
| 8 | 11,251 | 11,174 | 9,424 | 5,995 | 6,996 | 7,751 | 52,591 |
| 9 | 13,174 | 16,502 | 12,023 | 5,774 | 6,365 | 7,218 | 61,056 |
| 10 | 8,195 | 11,497 | 8,649 | 3,952 | 3,732 | 3,913 | 39,938 |
| 11 | 5,764 | 8,372 | 5,547 | 2,399 | 2,658 | 2,558 | 27,298 |
| 12 | 4,029 | 6,314 | 3,987 | 1,830 | 1,647 | 1,804 | 19,611 |

Table 3.2.3.2.3.2B
Proficiency Level by Grade (Percent): Comprehension S401 Online

| Grade | Comprehension Proficiency Range |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $9.7 \%$ | $15.9 \%$ | $21.5 \%$ | $13.4 \%$ | $19.4 \%$ | $20.2 \%$ | $100.0 \%$ |
| 2 | $9.5 \%$ | $23.0 \%$ | $24.6 \%$ | $13.1 \%$ | $15.0 \%$ | $14.9 \%$ | $100.0 \%$ |
| 3 | $11.4 \%$ | $18.7 \%$ | $21.5 \%$ | $11.0 \%$ | $16.2 \%$ | $21.1 \%$ | $100.0 \%$ |
| 4 | $4.5 \%$ | $11.3 \%$ | $19.2 \%$ | $16.1 \%$ | $24.6 \%$ | $24.2 \%$ | $100.0 \%$ |
| 5 | $9.1 \%$ | $14.3 \%$ | $18.4 \%$ | $14.4 \%$ | $19.8 \%$ | $24.1 \%$ | $100.0 \%$ |
| 6 | $18.3 \%$ | $25.5 \%$ | $23.4 \%$ | $12.6 \%$ | $11.4 \%$ | $8.7 \%$ | $100.0 \%$ |
| 7 | $20.5 \%$ | $22.7 \%$ | $20.5 \%$ | $12.3 \%$ | $11.9 \%$ | $12.1 \%$ | $100.0 \%$ |
| 8 | $21.4 \%$ | $21.2 \%$ | $17.9 \%$ | $11.4 \%$ | $13.3 \%$ | $14.7 \%$ | $100.0 \%$ |
| 9 | $21.6 \%$ | $27.0 \%$ | $19.7 \%$ | $9.5 \%$ | $10.4 \%$ | $11.8 \%$ | $100.0 \%$ |
| 10 | $20.5 \%$ | $28.8 \%$ | $21.7 \%$ | $9.9 \%$ | $9.3 \%$ | $9.8 \%$ | $100.0 \%$ |
| 11 | $21.1 \%$ | $30.7 \%$ | $20.3 \%$ | $8.8 \%$ | $9.7 \%$ | $9.4 \%$ | $100.0 \%$ |
| 12 | $20.5 \%$ | $32.2 \%$ | $20.3 \%$ | $9.3 \%$ | $8.4 \%$ | $9.2 \%$ | $100.0 \%$ |

### 3.2.3.2.4 Overall Composite

### 3.2.3.2.4.1 By Cluster

Table 3.2.3.2.4.1A
Proficiency Level by Cluster (Count): Overall S401 Online

| Cluster | Overall Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 9,503 | 31,831 | 57,026 | 14,756 | 3,160 | 376 | 116,652 |
| $2-3$ | 16,730 | 49,774 | 105,962 | 56,572 | 7,932 | 188 | 237,158 |
| $4-5$ | 5,748 | 9,254 | 33,951 | 32,224 | 5,442 | 616 | 87,235 |
| $6-8$ | 16,582 | 27,000 | 44,390 | 26,854 | 3,321 | 233 | 118,380 |
| $9-12$ | 19,413 | 24,359 | 44,603 | 17,334 | 2,890 | 174 | 108,773 |

Table 3.2.3.2.4.1B
Proficiency Level by Cluster (Percent): Overall S401 Online

| Cluster | Overall Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $8.1 \%$ | $27.3 \%$ | $48.9 \%$ | $12.6 \%$ | $2.7 \%$ | $0.3 \%$ | $100.0 \%$ |
| $2-3$ | $7.1 \%$ | $21.0 \%$ | $44.7 \%$ | $23.9 \%$ | $3.3 \%$ | $0.1 \%$ | $100.0 \%$ |
| $4-5$ | $6.6 \%$ | $10.6 \%$ | $38.9 \%$ | $36.9 \%$ | $6.2 \%$ | $0.7 \%$ | $100.0 \%$ |
| $6-8$ | $14.0 \%$ | $22.8 \%$ | $37.5 \%$ | $22.7 \%$ | $2.8 \%$ | $0.2 \%$ | $100.0 \%$ |
| $9-12$ | $17.8 \%$ | $22.4 \%$ | $41.0 \%$ | $15.9 \%$ | $2.7 \%$ | $0.2 \%$ | $100.0 \%$ |

### 3.2.3.2.4.2 By Grade

Table 3.2.3.2.4.2A
Proficiency Level by Grade (Count): Overall S401 Online

| Grade | Overall Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | 9,503 | 31,831 | 57,026 | 14,756 | 3,160 | 376 | 116,652 |
| 2 | 7,449 | 29,005 | 52,867 | 21,413 | 2,594 | 81 | 113,409 |
| 3 | 9,281 | 20,769 | 53,095 | 35,159 | 5,338 | 107 | 123,749 |
| 4 | 2,682 | 4,845 | 20,216 | 18,854 | 2,672 | 344 | 49,613 |
| 5 | 3,066 | 4,409 | 13,735 | 13,370 | 2,770 | 272 | 37,622 |
| 6 | 4,672 | 9,621 | 16,932 | 7,868 | 739 | 51 | 39,883 |
| 7 | 5,549 | 9,009 | 14,579 | 8,991 | 1,132 | 96 | 39,356 |
| 8 | 6,361 | 8,370 | 12,879 | 9,995 | 1,450 | 86 | 39,141 |
| 9 | 8,354 | 8,692 | 17,684 | 8,396 | 1,499 | 117 | 44,742 |
| 10 | 5,110 | 6,570 | 12,512 | 4,518 | 626 | 30 | 29,366 |
| 11 | 3,489 | 4,924 | 8,241 | 2,668 | 456 | 20 | 19,798 |
| 12 | 2,460 | 4,173 | 6,166 | 1,752 | 309 | 7 | 14,867 |

Table 3.2.3.2.4.2B
Proficiency Level by Grade (Percent): Overall S401 Online

| Grade | Overall Proficiency Range |  |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| 1 | $8.1 \%$ | $27.3 \%$ | $48.9 \%$ | $12.6 \%$ | $2.7 \%$ | $0.0 \%$ |  |
| 2 | $6.6 \%$ | $25.6 \%$ | $46.6 \%$ | $18.9 \%$ | $2.3 \%$ | $0.1 \%$ | $100.0 \%$ |
| 3 | $7.5 \%$ | $16.8 \%$ | $42.9 \%$ | $28.4 \%$ | $4.3 \%$ | $0.1 \%$ | $100.0 \%$ |
| 4 | $5.4 \%$ | $9.8 \%$ | $40.7 \%$ | $38.0 \%$ | $5.4 \%$ | $0.7 \%$ | $100.0 \%$ |
| 5 | $8.1 \%$ | $11.7 \%$ | $36.5 \%$ | $35.5 \%$ | $7.4 \%$ | $0.7 \%$ | $100.0 \%$ |
| 6 | $11.7 \%$ | $24.1 \%$ | $42.5 \%$ | $19.7 \%$ | $1.9 \%$ | $0.1 \%$ | $100.0 \%$ |
| 7 | $14.1 \%$ | $22.9 \%$ | $37.0 \%$ | $22.8 \%$ | $2.9 \%$ | $0.2 \%$ | $100.0 \%$ |
| 8 | $16.3 \%$ | $21.4 \%$ | $32.9 \%$ | $25.5 \%$ | $3.7 \%$ | $0.2 \%$ | $100.0 \%$ |
| 9 | $18.7 \%$ | $19.4 \%$ | $39.5 \%$ | $18.8 \%$ | $3.4 \%$ | $0.3 \%$ | $100.0 \%$ |
| 10 | $17.4 \%$ | $22.4 \%$ | $42.6 \%$ | $15.4 \%$ | $2.1 \%$ | $0.1 \%$ | $100.0 \%$ |
| 11 | $17.6 \%$ | $24.9 \%$ | $41.6 \%$ | $13.5 \%$ | $2.3 \%$ | $0.1 \%$ | $100.0 \%$ |
| 12 | $16.5 \%$ | $28.1 \%$ | $41.5 \%$ | $11.8 \%$ | $2.1 \%$ | $0.0 \%$ | $100.0 \%$ |

### 3.3. Analyses of Domain Scores: Results

3.3.1 Grade: 1

### 3.3.1.1 Listening 1

Table 3.3.1.1A
Complete Item Analysis and Summary: List 1 S401 Online

| Item Type | Average Item Difficulty (in logits) | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | -1.40 | 54 | 0.67 | 0.98 | 0.98 |
|  | Item |  |  | Fit <br> Statistics |  |
| Name | Difficulty (in logits) | Anchored? | P-value | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.LO1A_SI_Go ingHome_P 100_Screen_2_12445 | -4.38 |  | 0.97 | 0.90 | 0.87 |
| 2.LO1A_SI_GoingHome_P 100_Screen_3_12446 | -2.54 |  | 0.87 | 1.22 | 1.84 |
| 3.LO1A_SI_Go ingHome_P 100_Screen_4_12447 | 1.15 |  | 0.38 | 1.27 | 2.29 |
| 4.L0IB_SI_Choos ingCenters_P 100_Screen_2_12403 | -2.48 |  | 0.87 | 1.00 | 0.91 |
| 5.L01B_SI_Choos ingCenters_P 100_Screen_3_12404 | -2.92 |  | 0.91 | 1.01 | 0.99 |
| 6.L0 IB_SI_Choos ingCenters_P 100_Screen_4_12405 | -2.44 |  | 0.88 | 1.14 | 1.33 |
| 7.L01A_LA_Outdo orFun_P 100_Screen_2_12690 | -3.39 | Yes | 0.67 | 0.87 | 0.81 |
| 8.L01A_LA_Outdo orFun_P 100_Screen_3_12691 | -2.96 |  | 0.56 | 0.99 | 0.98 |
| 9.L01A_LA_Outdo orFun_P 100_Screen_4_12692 | -3.17 | Yes | 0.67 | 0.91 | 0.84 |
| 10.L01A_MA_DrawingaRobot_P 100_A202_Screen_2_13889 | -3.27 |  | 0.66 | 0.92 | 0.86 |
| 11.L01A_MA_DrawingaRobot_P 100_A202_Screen_3_13890 | -2.23 |  | 0.44 | 1.00 | 1.00 |
| 12.L01A_MA_Dra wingaRobot_P 100_A202_Screen_4_13891 | -2.55 |  | 0.51 | 1.01 | 1.00 |
| 13.L01C_SS_CampingTrip_P 100_A301FT_Screen_2_13808 | -2.41 |  | 0.48 | 1.04 | 1.05 |
| 14.LO1C_SS_CampingTrip_P 100_A301FT_Screen_3_13809 | -1.90 |  | 0.37 | 0.95 | 0.97 |
| 15.L01C_SS_CampingTrip_P 100_A301FT_Screen_4_13810 | -2.97 |  | 0.60 | 0.93 | 0.90 |
| 16.L01A_SC_GymClas s_P 100_Screen_2_11063 | -2.72 |  | 0.53 | 0.92 | 0.90 |
| 17.L01A_SC_GymClas s_P 100_Screen_3_11064 | -2.45 |  | 0.49 | 0.97 | 0.96 |
| 18.L01A_SC_GymClas s_P 100_Screen_4_11065 | -3.58 | Yes | 0.77 | 0.94 | 0.87 |
| 19.L01B_LA_ASpecialDay_P 100_Screen_2_12693 | -2.63 | Yes | 0.83 | 0.92 | 0.92 |
| 20.LO1B_LA_ASpecialDay_P 100_Screen_3_12694 | -1.83 | Yes | 0.72 | 0.97 | 0.98 |
| 21.L01B_LA_ASpecialDay_P 100_Screen_4_12695 | -1.51 | Yes | 0.66 | 0.96 | 0.91 |
| 22.Lo1C_MA_RainyDay_P 100_A202_Screen_2_13898 | -0.94 |  | 0.55 | 1.06 | 1.07 |
| 23.LO1C_MA_RainyDay_P 100_A202_Screen_3_13899 | -1.15 | Yes | 0.60 | 0.97 | 0.95 |
| 24.LO1C_MA_RainyDay_P 100_A202_Screen_4_13900 | -1.82 | Yes | 0.71 | 0.96 | 0.93 |
| 25.L01B_SS_ConstructionWorker_P 100_A301FT_alt1_Screen_2_13802 | -2.40 |  | 0.81 | 0.92 | 0.84 |


| Name | $\begin{array}{\|c} \text { Item } \\ \text { Difficulty } \\ \text { (in logits) } \\ \hline \end{array}$ | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | $\begin{gathered} \hline \text { Outfit } \\ \text { Mnsq } \\ \hline \end{gathered}$ |
| 26.L01B_SS_Construction Worker_P 100_A301FT_alt I_Screen_3_13803 | -1.67 | Yes | 0.75 | 0.92 | 0.85 |
| 27.L01B_SS_ConstructionWorker_P 100_A301FT_alt1_Screen_4_13804 | -1.03 | Yes | 0.53 | 0.90 | 0.87 |
| 28.L01C_SC_ForestHabitat_P 100_Screen_2_11667 | 0.44 |  | 0.27 | 0.94 | 0.93 |
| 29.L01C_SC_ForestHabitat_P 100_Screen_3_11668 | -0.81 | Yes | 0.51 | 0.97 | 0.96 |
| 30.LO1C_SC_ForestHabitat_P 100_Screen_4_11671 | -2.02 |  | 0.77 | 0.96 | 0.92 |
| 31.L0 1B_LA_Laura Ingalls Wilder_P 100_A301FT_Screen_2_13883 | -2.09 |  | 0.77 | 0.95 | 0.90 |
| 32.L01B_LA_Lauralngalls Wilder_P 100_A 301 IFT_Screen_3_13884 | -1.78 | Yes | 0.69 | 0.94 | 0.90 |
| 33.L01B_LA_Laura Ingalls Wilder_P 100_A 301FT_Screen_4_13885 | -0.28 |  | 0.39 | 0.93 | 0.92 |
| 34.L01B_MA_Subtraction_P 100_Screen_2_12457 | -2.08 |  | 0.76 | 0.94 | 0.90 |
| 35.L01B_MA_Subtractio n_P 100_Screen_3_12458 | -1.86 |  | 0.73 | 0.94 | 0.90 |
| 36.L01B_MA_Subtractio n_P 100_Screen_4_12459 | -0.22 |  | 0.39 | 0.95 | 0.95 |
| 37.LO1B_LA_ASpecialDay_P 100_alt_Screen_2_13814 | -1.56 | Yes | 0.86 | 0.99 | 0.97 |
| 38.LO1B_LA_ASpecialDay_P 100_alt_Screen_3_13815 | -0.38 |  | 0.72 | 1.07 | 1.10 |
| 39.L01B_LA_ASpecialDay_P 100_alt_Screen_4_13816 | -1.89 | Yes | 0.92 | 0.95 | 0.85 |
| 40.L01C_MA_ShapeRiddles _P 100_Screen_2_12846 | 0.58 |  | 0.58 | 1.08 | 1.10 |
| 41.L01C_MA_ShapeRiddles_P 100_Screen_3_12847 | 0.02 |  | 0.68 | 1.01 | 1.01 |
| 42.L01C_MA_Shape Riddles _P 100_Screen_4_12848 | 0.10 |  | 0.67 | 0.98 | 0.96 |
| 43.L01B_SS_SchoolCommunity_P 100_Screen_2_12745 | -1.35 |  | 0.88 | 0.99 | 1.00 |
| 44.L0IB_SS_SchoolCommunity_P 100_Screen_3_12746 | -0.86 |  | 0.85 | 0.95 | 0.88 |
| 45.L01B_SS_SchoolCommunity_P 100_Screen_4_12754 | -0.47 |  | 0.79 | 0.94 | 0.91 |
| 46.LOIB_SC_Solids_P 100_alt_Screen_2_13877 | 1.34 |  | 0.47 | 0.96 | 0.95 |
| 47.L01B_SC_Solids_P 100_alt_Screen_3_13878 | 0.31 |  | 0.69 | 0.92 | 0.89 |
| 48.L01B_SC_Solids_P 100_alt_Screen_4_13879 | -0.64 |  | 0.84 | 0.94 | 0.90 |
| 49.L01B_LA_The Wind_P 100_Screen_2_12460 | 1.09 |  | 0.56 | 0.95 | 0.94 |
| 50.L01B_LA_The Wind_P 100_Screen_3_12461 | 0.79 |  | 0.63 | 0.94 | 0.92 |
| 51.L01B_LA_The Wind_P 100_Screen_4_12462 | -1.14 |  | 0.90 | 0.94 | 0.82 |
| 52.L01B_MA_Shape Riddles _P 100_Screen_2_12843 | 0.21 |  | 0.71 | 0.93 | 0.90 |
| 53.L01B_MA_Shape Riddles_P 100_Screen_3_12844 | 1.46 |  | 0.46 | 1.01 | 1.01 |
| 54.L01B_MA_ShapeRiddles _P 100_Screen_4_12845 | -0.16 |  | 0.79 | 1.01 | 1.04 |

Table 3.3.1.1B
DIF Analysis and Summary: List 1 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| A | 31 | 22 | 27 | 27 |
| B | 0 | 1 | 0 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.LO1A_SI_GoingHo me_P 100_Screen_2_12445 | A | M | A | O |
| 2.LO1A_SI_GoingHo me_P 100_Screen_3_12446 | A | M | A | O |
| 3.LO1A_SI_Go ingHo me_P 100_Screen_4_12447 | A | M | A | O |
| 4.LOIB_SI_Choos ingCenters_P 100_Screen_2_12403 | A | M | A | H |
| 5.LOIB_SI_Choos ingCenters_P 100_Screen_3_12404 | A | M | A | O |
| 6.L01B_SI_Choos ingCenters_P 100_Screen_4_12405 | A | M | A | O |
| 7.L01A_LA_Outdo orFun_P 100_Screen_2_12690 | A | M | A | H |
| 8.L01A_LA_Outdo orFun_P 100_Screen_3_12691 | A | M | A | H |
| 9.L01A_LA_Outdo orFun_P 100_Screen_4_12692 | A | M | A | O |
| 10.L01A_MA_DrawingaRobot_P 100_A202_Screen_2_13889 | A | M | A | O |
| 11.L01A_MA_DrawingaRobot_P 100_A202_Screen_3_13890 | A | F | A | H |
| 12.L01A_MA_DrawingaRobot_P 100_A202_Screen_4_13891 | A | F | A | H |
| 13.L01C_SS_CampingTrip_P 100_A301FT_Screen_2_13808 | A | M | A | O |
| 14.L01C_SS_CampingTrip_P 100_A301FT_Screen_3_13809 | A | M | A | H |
| 15.L01C_SS_CampingTrip_P 100_A301FT_Screen_4_13810 | A | F | A | O |
| 16.L01A_SC_GymClas s_P 100_Screen_2_11063 | A | F | A | H |
| 17.L01A_SC_GymClas s_P 100_Screen_3_11064 | A | F | A | H |
| 18.LO1A_SC_GymClass_P 100_Screen_4_11065 | A | F | A | O |
| 19.L01B_LA_ASpecialDay_P 100_Screen_2_12693 | A | M | A | O |
| 20.L01B_LA_ASpecialDay_P 100_Screen_3_12694 | A | M | A | O |
| 21.L0 1B_LA_ASpecialDay_P 100_Screen_4_12695 | A | M | A | H |
| 22.L01C_MA_RainyDay_P 100_A202_Screen_2_13898 | A | F | A | O |
| 23.L01C_MA_RainyDay_P 100_A202_Screen_3_13899 | A | F | A | H |
| 24.L01C_MA_RainyDay_P 100_A202_Screen_4_13900 | A | M | A | O |
| 25.L01B_SS_ConstructionWorker_P 100_A301FT_alt_Screen_2_13802 | A | M | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 26.L01B_SS_ConstructionWorker_P 100_A301FT_alt_Screen_3_13803 | A | M | A | O |
| 27.L01B_SS_ConstructionWorker_P 100_A301FT_alt_Screen_4_13804 | A | F | A | H |
| 28.LO1C_SC_ForestHabitat_P 100_Screen_2_11667 | A | M | A | H |
| 29.L01C_SC_ForestHabitat_P 100_Screen_3_11668 | A | M | A | H |
| 30.Lo1C_SC_ForestHabitat_P 100_Screen_4_11671 | A | F | A | O |
| 31.L01B_LA_Laura Ingalls Wilder_P 100_A301FT_Screen_2_13883 | A | M | A | O |
| 32.L01B_LA_Laura Ingalls Wilder_P 100_A301FT_Screen_3_13884 | A | F | A | H |
| 33.L01B_LA_Laura Ingalls Wilder_P 100_A301FT_Screen_4_13885 | A | F | A | O |
| 34.LO1B_MA_Subtraction_P 100_Screen_2_12457 | A | F | A | H |
| 35.LO1B_MA_Subtraction_P 100_Screen_3_12458 | A | M | A | H |
| 36.L01B_MA_Subtraction_P 100_Screen_4_12459 | A | F | A | H |
| 37.L01B_LA_ASpecialDay_P 100_alt_Screen_2_13814 | A | M | A | O |
| 38.L01B_LA_ASpecialDay_P 100_alt_Screen_3_13815 | A | F | A | O |
| 39.L01B_LA_ASpecialDay_P 100_alt_Screen_4_13816 | A | M | A | O |
| 40.L01C_MA_ShapeRiddles_P 100_Screen_2_12846 | A | F | A | H |
| 41.L01C_MA_ShapeRiddles_P 100_Screen_3_12847 | A | F | A | O |
| 42.L01C_MA_ShapeRiddles_P 100_Screen_4_12848 | A | M | A | H |
| 43.L01B_SS_SchoolCommunity_P 100_Screen_2_12745 | A | M | A | O |
| 44.LO1B_SS_SchoolCommunity_P 100_Screen_3_12746 | A | F | A | H |
| 45.LO1B_SS_SchoolCommunity_P 100_Screen_4_12754 | A | M | A | O |
| 46.L01B_SC_So lids_P 100_altI_Screen_2_13877 | A | F | A | H |
| 47.L01B_SC_So lids_P 100_altıScreen_3_13878 | A | M | A | H |
| 48.L01B_SC_So lids_P 100_altI_Screen_4_13879 | A | M | A | H |
| 49.L01B_LA_The Wind_P 100_Screen_2_12460 | A | F | A | H |
| 50.L0 1B_LA_The Wind_P 100_Screen_3_12461 | A | M | A | H |
| $51 . L 01 \mathrm{~B}$ _LA_The Wind_P 100_Screen_4_12462 | B | F | A | O |
| 52.L01B_MA_ShapeRiddles_P 100_Screen_2_12843 | A | M | A | H |
| 53.L01B_MA_ShapeRiddles_P 100_Screen_3_12844 | A | F | A | O |
| 54.L01B_MA_ShapeRiddles_P 100_Screen_4_12845 | A | F | A | H |

Figure 3.3.1.1C
Raw Scores: List 1 S401 Online
n/a
Table 3.3.1.1C
Raw Score Descriptive Statistics: List 1 S401 Online n/a



Table 3.3.1.1D
Scale Score Descriptive Statistics: List 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 146,126 | 104 | 434 | 328.03 | 58.63 |
| Total | 146,126 | 104 | 434 | 328.03 | 58.63 |

Table 3.3.1.1E
Proficiency Level Distribution: List 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 14,016 | $9.59 \%$ | 14,016 | $9.59 \%$ |
| $\mathbf{2}$ | 5,963 | $4.08 \%$ | 5,963 | $4.08 \%$ |
| $\mathbf{3}$ | 17,462 | $11.95 \%$ | 17,462 | $11.95 \%$ |
| $\mathbf{4}$ | 6,992 | $4.78 \%$ | 6,992 | $4.78 \%$ |
| $\mathbf{5}$ | 14,551 | $9.96 \%$ | 14,551 | $9.96 \%$ |
| $\mathbf{6}$ | 87,142 | $59.63 \%$ | 87,142 | $59.63 \%$ |
| Total | 146,126 | $100.00 \%$ | 146,126 | $100.00 \%$ |

Table 3.3.1.1F
Raw Score to Scale Score Conversion: List 1 S401 Online n/a

Table 3.3.1.1G
Equating Summary: List 1 S401 Online


[^5]Figure 3.3.1.1H
Test Characteristic Curve: List 1 S401 Online n/a


Table 3.3.1.1J
Reliability: List 1 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 146,126 | 54 | .86 |

Table 3.3.1.1K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 1 S401 Online

| Proficiency <br> Level | Grade | Cut Score | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $1 / 2$ | 1 | 236 | 184 | 16.84 | 19.39 | 19.21 | 0.66 |
| $2 / 3$ | 1 | 259 | 284 | 16.84 | 17.86 | 17.45 | 0.43 |
| $3 / 4$ | 1 | 291 | 576 | 17.86 | 20.41 | 18.61 | 0.79 |
| $4 / 5$ | 1 | 303 | 358 | 17.86 | 21.94 | 19.57 | 0.98 |
| $5 / 6$ | 1 | 327 | N/A | N/A | N/A | N/A | N/A |

Table 3.3.1.1L
Accuracy and Consistency of Classification Indices: List (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.719 | 0.659 |  | 0.447 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.853 |  | 0.717 |  |
|  | 2 | 0.254 |  | 0.183 |  |
|  | 3 | 0.500 |  | 0.375 |  |
|  | 4 | 0.179 |  | 0.128 |  |
|  | 5 | 0.306 |  | 0.223 |  |
|  | 6 | 0.919 |  | 0.886 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.958 | 0.011 | 0.030 | 0.942 |
|  | 2/3 | 0.947 | 0.023 | 0.030 | 0.923 |
|  | 3/4 | 0.919 | 0.031 | 0.050 | 0.887 |
|  | 4/5 | 0.913 | 0.036 | 0.051 | 0.876 |
|  | 5/6 | 0.895 | 0.057 | 0.048 | 0.850 |

### 3.3.1.2 Reading 1

Table 3.3.1.2A
Complete Item Analysis and Summary: Read 1 S401 Online

| Item Type | Average Item Difficulty (in logits) | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | -1.01 | 72 | 0.41 | 0.99 | 0.98 |
|  | Item |  |  | Fit <br> Statistics |  |
| Name | Difficulty (in logits) | Anchored? | P-value | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1R01A_SI_GettingReady_P 100_A203_Screen_2_13193 | -2.11 |  | 0.66 | 0.93 | 0.83 |
| 2.R01A_SI_GettingReady_P 100_A203_Screen_3_13194 | -2.06 |  | 0.65 | 0.99 | 0.89 |
| 3.R01A_SI_GettingReady_P 100_A203_Screen_4_13195 | -1.52 |  | 0.55 | 1.10 | 1.15 |
| 4.R01B_SI_MorningMes sage_P 100_A203_Screen_2_13196 | -1.11 |  | 0.47 | 1.06 | 1.12 |
| 5.R01B_SI_MorningMes sage_P 100_A203_Screen_3_13197 | -1.54 |  | 0.55 | 1.08 | 1.09 |
| 6.R01B_SI_MorningMes sage_P 100_A203_Screen_4_13198 | -0.64 |  | 0.40 | 1.31 | 1.50 |
| 7.R01A_LA_CatAdventure_203_P 100_A301_Screen_2_13211 | -2.30 |  | 0.36 | 0.97 | 0.97 |
| 8.R01A_LA_CatAdventure_203_P 100_A301_Screen_3_13212 | -1.67 | Yes | 0.29 | 0.97 | 0.97 |
| $9 . \mathrm{R} 01 \mathrm{~A}$ _LA_CatAdventure_203_P 100_A301_Screen_4_13213 | -2.28 |  | 0.37 | 0.97 | 0.97 |
| 10.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_2_13274 | -4.50 |  | 0.75 | 1.03 | 1.08 |
| 11.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_3_13275 | -1.10 | Yes | 0.14 | 0.98 | 1.00 |
| 12.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_4_13276 | -1.13 | Yes | 0.15 | 0.95 | 0.91 |
| 13.R01C_SS_PetStore_P 100_A202_Screen_2_13313 | -4.96 |  | 0.82 | 1.01 | 1.04 |
| 14-R01C_SS_PetStore_P 100_A 202_Screen_3_13314 | -2.42 |  | 0.35 | 0.99 | 0.98 |
| 15.R01C_SS_PetStore_P 100_A202_Screen_4_13315 | -2.58 |  | 0.38 | 0.98 | 0.98 |
| 16.R01A_SC_Birds_dode_P 100_A301_Screen_2_13208 | -4.30 |  | 0.71 | 0.99 | 0.98 |
| 17.R01A_SC_Birds_dode_P 100_A301_Screen_3_13209 | -1.65 |  | 0.18 | 0.99 | 0.98 |
| 18.R01A_SC_Birds_dode_P 100_A301_Screen_4_13210 | -0.40 | Yes | 0.09 | 0.99 | 0.99 |
| 19.R01A_LA_A ${ }^{\text {is it }}$ (To TheP ond_P 100_A203_Screen_2_13199 | -1.94 | Yes | 0.25 | 0.95 | 0.92 |
| 20.R01A_LA_AVisitTo TheP ond_P 100_A203_Screen_3_B200 | -2.58 |  | 0.35 | 0.92 | 0.90 |
| 21.R01A_LA_A Vis itTo TheP ond_P 100_A203_Screen_4_13201 | -1.52 |  | 0.16 | 0.96 | 0.91 |
| 22.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_2_13202 | -2.25 | Yes | 0.28 | 0.92 | 0.89 |
| 23.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_3_13203 | -2.13 |  | 0.26 | 0.98 | 0.97 |
| 24.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_4_13204 | -2.65 |  | 0.35 | 0.97 | 0.96 |
| 25.R01B_LA_IvanAtTheMarket_P 100_A202_Screen_2_13229 | -2.87 |  | 0.73 | 0.98 | 0.99 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \\ \hline \end{gathered}$ | Outfit <br> Mnsq |
| 26.R01B_LA_IvanAtTheMarket_P 100_A202_Screen_3_13230 | -0.56 |  | 0.26 | 0.99 | 0.99 |
| 27.R01B_LA_IvanAtTheMarket_P 100_A202_Screen_4_B231 | -0.29 | Yes | 0.27 | 0.98 | 0.97 |
| 28.R01B_MA_EstimatingMoney_ShLo GiPe_P 100_A 203 _Screen_2_13220 | -0.83 |  | 0.32 | 0.99 | 0.96 |
| 29.R01B_MA_EstimatingMoney_ShLoGiPe_P 100_A203_Screen_3_13221 | -0.62 | Yes | 0.33 | 0.99 | 0.97 |
| 30R01B_MA_EstimatingMoney_ShLo GiPe_P 100_A203_Screen_4_13222 | -1.04 |  | 0.35 | 0.98 | 0.95 |
| 31R01B_SS_SchoolStore_jc_P 100_A203_Screen_2_13235 | -2.38 |  | 0.67 | 0.95 | 0.93 |
| 32.R01B_SS_SchoolStore_jc_P 100_A203_Screen_3_13236 | -1.69 | Yes | 0.51 | 0.97 | 0.96 |
| 33.R01B_SS_SchoolStore_jc_P 100_A203_Screen_4_13237 | -1.29 |  | 0.42 | 0.96 | 0.95 |
| 34.R01B_SC_AnimalCoverings_dode_P 100_A301_Screen_2_13238 | -1.58 |  | 0.48 | 0.97 | 0.96 |
| 35.R01B_SC_AnimalCoverings_dode_P 100_A301_Screen_3_13239 | -0.60 |  | 0.26 | 0.99 | 0.98 |
| 36.R01B_SC_AnimalCoverings_dode_P 100_A301_Screen_4_13240 | -0.95 | Yes | 0.37 | 0.99 | 0.98 |
| 37.R01B_LA_Lunchtime_P 100_A202_Screen_2_13283 | -1.23 |  | 0.40 | 0.96 | 0.95 |
| 38.R01B_LA_Lunchtime_P 100_A202_Screen_3_13284 | -1.71 | Yes | 0.57 | 0.96 | 0.96 |
| 39.R01B_LA_Lunchtime_P 100_A202_Screen_4_13285 | -0.62 |  | 0.27 | 0.99 | 0.98 |
| 40.R01C_MA_IreCreamAtTheP ark_AmEtPaAg_P 100_A203_Screen_2_13244 | -0.80 | Yes | 0.31 | 0.99 | 0.98 |
| 41.R01C_MA_leeCreamAtTheP ark_AmEtP a Ag_P 100_A203_Screen_3_13245 | -0.81 |  | 0.34 | 0.99 | 0.99 |
| 42.R01C_MA_IreCreamAtTheP ark_AmEtPa ag_P 100_A203_Screen_4_13246 | -0.45 | Yes | 0.31 | 1.00 | 1.00 |
| 43.R01A_SS_Homes OfThePast_203_P 100_A301_Screen_2_13205 | -0.34 |  | 0.23 | 0.98 | 0.97 |
| 44.R01A_SS_Homes OfThePast_203_P 100_A301_Screen_3_13206 | -1.82 |  | 0.55 | 0.96 | 0.96 |
| 45.R01A_SS_Homes OfThePast_203_P 100_A301_Screen_4_13207 | -0.32 |  | 0.24 | 1.01 | 1.01 |
| 46.R01C_SC_Cotto n_P 100_A201_Screen_2_13316 | -0.61 | Yes | 0.33 | 0.99 | 0.99 |
| 47.R01C_SC_Cotto n_P 100_A201_Screen_3_13317 | -0.63 |  | 0.29 | 1.00 | 1.00 |
| 48.R01C_SC_Cotto n_P 100_A201_Screen_4_3318 | -0.81 |  | 0.34 | 1.00 | 0.99 |
| 49.R01B_LA_Zoo Trip_203_P 100_A301_Screen_2_13217 | -1.03 |  | 0.64 | 0.81 | 0.74 |
| $50 . \mathrm{R} 01 \mathrm{~B}$ _LA_Zoo Trip_203_P 100_A301_Screen_3_13218 | -0.45 | Yes | 0.49 | 0.97 | 0.93 |
| 51 R 01 B _LA_ZooTrip_203_P 100_A301_Screen_4_13219 | 0.05 | Yes | 0.38 | 0.99 | 0.97 |
| 52.R01C_M A_BalanceScale_401_V2_Screen_2_14619 | -0.44 |  | 0.46 | 1.05 | 1.06 |
| 53.R01C_MA_BalanceScale_401_V2_Screen_3_14620 | 0.03 |  | 0.38 | 1.18 | 1.24 |
| 54.R01C_MA_B alanceScale_401_V2_Screen_4_14621 | -0.16 |  | 0.41 | 1.12 | 1.15 |
| 55.R01B_SS_Farmers Market_401_V2_Screen_2_4721 | -0.64 |  | 0.49 | 1.07 | 1.09 |
| 56.R01B_SS_Farmers Market_401_V2_Screen_3_14722 | -0.54 |  | 0.46 | 0.94 | 0.92 |
| 57.R01B_SS_Farmers Market_401_V2_Screen_4_4723 | -0.42 |  | 0.45 | 0.99 | 0.99 |
| $58 . \mathrm{R} 01 \mathrm{~B}$ _SC_AnimalCo verings_dode_P 100_A 301_altl_Screen_2_13304 | 0.90 |  | 0.33 | 0.91 | 0.87 |
| $59 . \mathrm{R} 01 \mathrm{~B}$ _SC_AnimalCo verings _dode_P 100_A 301_alt_Screen_3_13305 | 0.80 |  | 0.32 | 1.05 | 1.05 |
| $60 . \mathrm{R} 018$ _SC_AnimalCo verings_dode_P 100_A 301 _alt 1 Screen_4_13306 | -0.09 |  | 0.51 | 0.96 | 0.96 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | $\begin{aligned} & \hline \text { Outfit } \\ & \text { Mnsq } \\ & \hline \end{aligned}$ |
| 61.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_2_13241 | 0.50 |  | 0.39 | 0.83 | 0.79 |
| 62.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_3_13242 | 0.57 |  | 0.38 | 0.84 | 0.80 |
| 63.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_4_13243 | 0.47 |  | 0.42 | 1.05 | 1.03 |
| 64.R01C_MA_P uppetSho w_dode_P 100_A 203 _Screen_2_13256 | 0.57 |  | 0.42 | 1.04 | 1.03 |
| 65.R01C_MA_P uppetSho w_dode_P 100_A203_Screen_3_13257 | 0.61 |  | 0.42 | 0.99 | 0.98 |
| 66.R01C_MA_P uppetSho w_dode_P 100_A203_Screen_4_13258 | 0.14 |  | 0.49 | 0.95 | 0.92 |
| 67.R01C_S _UrbanNeighborho od_203_P 100_A301_Screen_2_13259 | 0.79 |  | 0.39 | 1.05 | 1.04 |
| 68.R01C_SS_UrbanNeighborho od_203_P 100_A301_Screen_3_13260 | -0.12 |  | 0.58 | 0.93 | 0.90 |
| 69.R01C_SS_UrbanNeighborho od_203_P 100_A301_Screen_4_13261 | 0.24 |  | 0.48 | 0.86 | 0.82 |
| 70.R01C_SC_Leaves_mika_P 100_A301_Screen_2_13262 | 1.00 |  | 0.36 | 1.01 | 1.01 |
| 71.R01C_SC_Leaves_mika_P 100_A 301_Screen_3_13263 | 0.24 |  | 0.52 | 1.06 | 1.09 |
| 72.R01C_SC_Leaves_mika_P 100_A301_Screen_4_13264 | 0.73 |  | 0.39 | 0.99 | 0.99 |

Table 3.3.1.2B
DIF Analysis and Summary: Read 1 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (0) |
| A | 34 | 38 | 34 | 37 |
| B | 0 | 0 | 1 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \hline \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \hline \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1R01A_SI_GettingReady_P 100_A203_Screen_2_13193 | A | M | A | O |
| 2.R01A_SI_GettingReady_P 100_A203_Screen_3_B194 | A | M | A | O |
| 3.R01A_SI_GettingReady_P 100_A203_Screen_4_13195 | A | M | A | O |
| 4.R01B_SI_MorningMes sage_P 100_A203_Screen_2_13196 | A | M | A | H |
| 5 R01B_SI_MorningMes sage_P 100_A 203 _Screen_3_13197 | A | M | A | H |
| 6.R01B_SI_MorningMes sage_P 100_A203_Screen_4_13198 | A | M | A | O |
| 7.R01A_LA_CatAdventure_203_P 100_A301_Screen_2_13211 | A | F | A | O |
| 8.R01A_LA_CatAdventure_203_P 100_A301_Screen_3_13212 | A | F | A | H |
| 9.R01A_LA_CatAdventure_203_P 100_A301_Screen_4_13213 | A | M | A | O |
| 10.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_2_13274 | A | M | A | H |
| 11.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_3_13275 | A | F | A | O |
| 12.R01A_MA_CountBallo ons_301_P 100_A301FT_Screen_4_13276 | A | F | A | O |
| 13.R01C_SS_PetStore_P 100_A202_Screen_2_13313 | A | M | B | H |
| 14.R01C_SS_PetStore_P 100_A202_Screen_3_13314 | A | F | A | O |
| 15.R01C_SS_PetStore_P 100_A 202 _Screen_4_13315 | A | F | A | O |
| 16.R01A_SC_Birds_dode_P 100_A301_Screen_2_13208 | A | F | A | H |
| 17.R01A_SC_B irds_dode_P 100_A301_Screen_3_13209 | A | M | A | O |
| 18.R01A_SC_Birds_dode_P 100_A 301_Screen_4_B210 | A | M | A | O |
| 19.R01A_LA_AVis it ${ }^{\text {a TheP ond_P 100_A203_Screen_2_13199 }}$ | A | F | A | O |
| 20.R01A_LA_A VisitTo ThePond_P 100_A203_Screen_3_13200 | A | M | A | H |
| 21.R01A_LA_A Vis itTo TheP ond_P 100_A203_Screen_4_13201 | A | F | A | O |
| 22.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_2_B202 | A | M | A | H |
| 23.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_3_13203 | A | F | A | O |
| 24.R01A_MA_FamilyGatheringatBeach_jc_P 100_A203_Screen_4_13204 | A | F | A | H |
| 25.R01B_LA_IvanAtTheMarket_P 100_A202_Screen_2_13229 | A | M | A | H |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 26.R01B_LA_IvanAtTheMarket_P 100_A202_Screen_3_13230 | A | F | A | O |
| 27.R01B_LA_LvanAtTheMarket_P 100_A202_Screen_4_13231 | A | M | A | O |
| 28.R01B_MA_EstimatingMoney_ShLo GiPe_P 100_A203_Screen_2_13220 | A | F | A | O |
| 29.R01B_MA_EstimatingMoney_ShLo GiP e_P 100_A203_Screen_3_13221 | A | M | A | H |
| 30.R01B_MA_EstimatingMoney_ShLo GiPe_P 100_A203_Screen_4_13222 | A | F | A | O |
| 31R01B_SS_SchoolStore_ic_P 100_A203_Screen_2_13235 | A | M | A | H |
| 32.R01B_SS_SchoolStore_jc_P 100_A203_Screen_3_13236 | A | F | A | O |
| 33.R01B_SS_SchoolStore_jc_P 100_A203_Screen_4_13237 | A | M | A | H |
| 34.R01B_SC_AnimalCo verings_dode_P 100_A 301 _Screen_2_13238 | A | M | A | O |
| 35.R01B_SC_AnimalCoverings_dode_P 100_A301_Screen_3_13239 | A | M | A | O |
| 36.R01B_SC_AnimalCo verings_dode_P 100_A 301 _Screen_4_B240 | A | F | A | O |
| 37.R01B_LA_Lunchtime_P 100_A202_Screen_2_13283 | A | F | A | O |
| 38.R01B_LA_Lunchtime_P 100_A202_Screen_3_13284 | A | M | A | H |
| 39.R01B_LA_Lunchtime_P 100_A202_Screen_4_13285 | A | F | A | O |
| 40.R01C_MA_IreCreamAtTheP ark_AmEtP a Ag_P 100_A203_Screen_2_13244 | A | M | A | H |
| 41.R01C_MA_IreCream AtTheP ark_AmEtP a Ag_P 100_A203_Screen_3_13245 | A | F | A | H |
| 42.R01C_MA_IceCreamAtTheP ark_AmEtP a Ag_P 100_A 203_Screen_4_13246 | A | F | A | O |
| 43.R01A_SS_Homes OfTheP ast_203_P 100_A301_Screen_2_13205 | A | M | A | O |
| 44.R01A_SS_Homes OfTheP ast_203_P 100_A301_Screen_3_13206 | A | F | A | O |
| 45.R01A_SS_Homes OfTheP ast_203_P 100_A301_Screen_4_13207 | A | F | A | O |
| 46.R01C_SC_Cotto n_P 100_A201_Screen_2_13316 | A | F | A | H |
| 47.R01C_SC_Cotto n_P 100_A201_Screen_3_13317 | A | F | A | H |
| 48.R01C_SC_Cotto n_P 100_A201_Screen_4_13318 | A | M | A | H |
| 49.R01B_LA_ZooTrip_203_P 100_A301_Screen_2_13217 | A | M | A | O |
| $50 . \mathrm{R} 01 \mathrm{~B}$ _LA_ZooTrip_203_P 100_A301_Screen_3_13218 | A | M | A | H |
| $51 \mathrm{R01B}$ _LA_Zoo Trip_203_P 100_A 301 Screen_4_132 19 | A | M | A | H |
| 52.R01C_MA_BalanceScale_401_V2_Screen_2_14619 | A | F | A | H |
| 53.R01C_MA_BalanceScale_401_V2_Screen_3_14620 | A | M | A | H |
| 54.R01C_MA_BalanceScale_401_V2_Screen_4_14621 | A | F | A | H |
| 55.R01B_SS_Farmers Market_401_V2_Screen_2_14721 | A | F | A | H |
| $56 . \mathrm{R} 01 \mathrm{~B}$ _SS_Farmers Market_401_V2_Screen_3_14722 | A | F | A | O |
| 57.R01B_SS_Farmers Market_401_V2_Screen_4_4723 | A | M | A | H |
| $58 . \mathrm{R} 01 \mathrm{~B}$ _SC_AnimalCo verings_dode_P 100_A 301_altl_Screen_2_13304 | A | F | A | O |
| 59.R01B_SC_AnimalCo verings_dode_P 100_A301_alt I_Screen_3_13305 | A | M | A | H |
| 60.R01B_SC_AnimalCo verings_dode_P 100_A301_alt 1_Screen_4_13306 | A | F | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 61.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_2_13241 | A | F | A | H |
| 62.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_3_13242 | A | M | A | O |
| 63.R01C_LA_Gro wingTaller_203_P 100_A301_Screen_4_13243 | A | F | A | H |
| 64.R01C_MA_PuppetShow_dode_P 100_A203_Screen_2_13256 | A | F | A | O |
| 65.R01C_MA_P uppetShow_dode_P 100_A203_Screen_3_13257 | A | M | A | H |
| 66.R01C_MA_P uppetShow_dode_P 100_A203_Screen_4_13258 | A | F | A | H |
| 67.R01C_SS_UrbanNeighborho od_203_P 100_A301_Screen_2_13259 | A | M | A | H |
| 68.R01C_SS_UrbanNeighborhood_203_P 100_A301_Screen_3_B260 | A | F | A | O |
| 69.R01C_SS_UrbanNeighborho od_203_P 100_A301_Screen_4_13261 | A | F | A | H |
| 70.R01C_SC_Leaves_mika_P 100_A301_Screen_2_13262 | A | M | A | O |
| 71.R01C_SC_Leaves_mika_P 100_A301_Screen_3_13263 | A | F | A | H |
| 72.R01C_SC_Leaves_mika_P 100_A301_Screen_4_13264 | A | F | A | H |

Figure 3.3.1.2C
Raw Scores: Read 1 S401 Online n/a

Table 3.3.1.2C
Raw Score Descriptive Statistics: Read 1 S401 Online n/a


Figure 3.3.1.2E
Proficiency Level: Read 1 S401 Online


Table 3.3.1.2D
Scale Score Descriptive Statistics: Read 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 146,074 | 141 | 409 | 290.12 | 34.64 |
| Total | 146,074 | 141 | 409 | 290.12 | 34.64 |

Table 3.3.1.2E
Proficiency Level Distribution: Read 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 28,938 | $19.81 \%$ | 28,938 | $19.81 \%$ |
| $\mathbf{2}$ | 39,222 | $26.85 \%$ | 39,222 | $26.85 \%$ |
| $\mathbf{3}$ | 29,296 | $20.06 \%$ | 29,296 | $20.06 \%$ |
| $\mathbf{4}$ | 13,132 | $8.99 \%$ | 13,132 | $8.99 \%$ |
| $\mathbf{5}$ | 21,419 | $14.66 \%$ | 21,419 | $14.66 \%$ |
| $\mathbf{6}$ | 14,067 | $9.63 \%$ | 14,067 | $9.63 \%$ |
| Total | 146,074 | $100.00 \%$ | 146,074 | $100.00 \%$ |

Table 3.3.1.2F
Raw Score to Scale Score Conversion: Read 1 S401 Online n/a

Table 3.3.1.2G
Equating Summary: Read 1 S401 Online


[^6]Figure 3.3.1.2H
Test Characteristic Curve: Read 1 S401 Online n/a

Figure 3.3.1.2I
Test Information Function: Read 1 S401 Online


Table 3.3.1.2J
Reliability: Read 1 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 146,074 | 72 | .89 |

Table 3.3.1.2K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 1 S401 Online

| Proficiency <br> Level | Grade | Cut Score | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $1 / 2$ | 1 | 264 | 102 | 10.71 | 11.22 | 10.88 | 0.24 |
| $2 / 3$ | 1 | 286 | 1,237 | 9.69 | 10.71 | 9.97 | 0.26 |
| $3 / 4$ | 1 | 304 | 709 | 9.69 | 10.71 | 10.25 | 0.24 |
| $4 / 5$ | 1 | 315 | 5,235 | 10.20 | 10.71 | 10.20 | 0.01 |
| $5 / 6$ | 1 | 334 | 11 | 10.20 | 10.20 | 10.20 | 0.00 |

Table 3.3.1.2L
Accuracy and Consistency of Classification Indices: Read (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.614 | 0.508 |  | 0.396 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.763 |  | 0.658 |  |
|  | 2 | 0.632 |  | 0.525 |  |
|  | 3 | 0.517 |  | 0.405 |  |
|  | 4 | 0.312 |  | 0.229 |  |
|  | 5 | 0.613 |  | 0.480 |  |
|  | 6 | $0.791$ |  | 0.661 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.910 | 0.049 | 0.041 | 0.872 |
|  | 2/3 | 0.884 | 0.058 | 0.058 | 0.840 |
|  | 3/4 | 0.908 | 0.049 | 0.043 | 0.868 |
|  | 4/5 | 0.921 | 0.049 | 0.030 | 0.889 |
|  | 5/6 | 0.958 | 0.023 | 0.019 | 0.938 |

### 3.3.1.3 Writing 1

### 3.3.1.3i <br> Writing 1 A

Table 3.3.1.3Ai
Complete Task Analysis and Summary: Writ 1 A S401 Online

| Task Type |  | Average Task Difficulty (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | -1.80 | 4 | 0.69 | 0.86 |
| Name |  | Task Difficulty (in logits) |  | Fit S | istics |
|  |  | Anchored? | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.W01A_SI_PartAH8_202_P 100_A302_14245 |  |  | -5.53 | Yes | 1.36 | 2.10 |
| 2.W01A_SI_CentersPartB_202_P100_A302_14246 |  | -1.36 | Yes | 0.72 | 0.68 |
| 3.W01A_SI_CentersPartC_202_P100_A302_14247 |  | -0.03 | Yes | 0.60 | 0.59 |
| 4.W01A_SI_PartDSentencesAboutMe_P100_A302_14248 |  | -0.27 | Yes | 0.08 | 0.08 |
| Raw Score <br> Distribution by Task | RawScore | Task 1 | Task 2 | Task 3 | Task 4 |
|  | 0 | 3.40\% | 5.83\% | 8.20\% | 5.76\% |
|  | 1 | 96.60\% | 91.96\% | 7.61\% | 5.58\% |
|  | 2 | N/A | 1.11\% | 11.71\% | 13.19\% |
|  | 3 | N/A | 1.11\% | 35.98\% | 22.57\% |
|  | 4 | N/A | N/A | 35.03\% | 47.86\% |
|  | 5 | N/A | N/A | 1.39\% | 4.58\% |
|  | 6 | N/A | N/A | 0.08\% | 0.45\% |
|  | 7 | N/A | N/A | 0.00\% | 0.00\% |
|  | 8 | N/A | N/A | 0.00\% | 0.00\% |
|  | 9 | N/A | N/A | 0.00\% | 0.00\% |

Table 3.3.1.3Bi
DIF Analysis and Summary: Writ 1 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 2 | 2 | 2 | 2 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |
| 1.W01A_SI_PartAH8_202_P 100_A302_14245 | AA | F | AA | O |
| 2.W01A_SI_CentersPartB_202_P 100_A302_14246 | AA | M | AA | H |
| 3.W01A_SI_CentersPartC_202_P 100_A302_14247 | AA | F | AA | O |
| 4.W01A_SI_PartDSentencesAboutMe_P 100_A302_14248 | AA | M | AA | H |

Figure 3.3.1.3Ci
Raw Scores: Writ 1 A S401 Online


Table 3.3.1.3Ci
Raw Score Descriptive Statistics: Writ 1 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 126,315 | 0 | 27 | 14.31 | 4.73 |
| Total | 126,315 | 0 | 27 | 14.31 | 4.73 |

Table 3.3.1.3Di
Scale Score Descriptive Statistics: Writ 1 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 126,315 | 111 | 334 | 249.70 | 27.69 |
| Total | 126,315 | 111 | 334 | 249.70 | 27.69 |



Table 3.3.1.3EB
Proficiency Level Distribution: Writ 1 A S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 28,030 | $22.19 \%$ | 28,030 | $22.19 \%$ |
| $\mathbf{2}$ | 89,943 | $71.21 \%$ | 89,943 | $71.21 \%$ |
| $\mathbf{3}$ | 8,342 | $6.60 \%$ | 8,342 | $6.60 \%$ |
| $\mathbf{4}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| $\mathbf{5}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 126,315 | $100.00 \%$ | 126,315 | $100.00 \%$ |

Table 3.3.1.3Fi
Raw Score to Scale Score Conversion: Writ 1 A S401 Online

| $\begin{gathered} \text { Raw } \\ \text { Score } \end{gathered}$ | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | $\begin{gathered} \text { Raw } \\ \text { Score } \end{gathered}$ | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 111^ | 49.94 | $100.00^{\wedge}$ | 157.94 | 34 | 381 | 13.21 | 367.79 | 394.21 |
| 1 | 148 | 32.25 | 115.75 | 180.25 | 35 | 387 | 13.64 | 373.36 | 400.64 |
| 2 | 177 | 23.74 | 153.26 | 200.74 | 36 | 395 | 14.47 | 380.53 | 409.47 |
| 3 | 193 | 17.86 | 175.14 | 210.86 | 37 | 403 | 16.00 | 387.00 | 419.00 |
| 4 | 202 | 14.47 | 187.53 | 216.47 | 38 | 414 | 18.98 | 395.02 | 432.98 |
| 5 | 209 | 12.54 | 196.46 | 221.54 | 39 | 433 | 26.42 | 406.58 | 459.42 |
| 6 | 214 | 11.38 | 202.62 | 225.38 | 40 | 464 | 48.52 | 415.48 | 512.52 |
| 7 | 219 | 10.66 | 208.34 | 229.66 |  |  |  |  |  |
| 8 | 223 | 10.23 | 212.77 | 233.23 |  |  |  |  |  |
| 9 | 227 | 10.02 | 216.98 | 237.02 |  |  |  |  |  |
| 10 | 231 | 9.96 | 221.04 | 240.96 |  |  |  |  |  |
| 11 | 234 | 10.07 | 223.93 | 244.07 |  |  |  |  |  |
| 12 | 238 | 10.31 | 227.69 | 248.31 |  |  |  |  |  |
| 13 | 242 | 10.69 | 231.31 | 252.69 |  |  |  |  |  |
| 14 | 247 | 11.20 | 235.80 | 258.20 |  |  |  |  |  |
| 15 | 252 | 11.84 | 240.16 | 263.84 |  |  |  |  |  |
| 16 | 257 | 12.54 | 244.46 | 269.54 |  |  |  |  |  |
| 17 | 263 | 13.21 | 249.79 | 276.21 |  |  |  |  |  |
| 18 | 270 | 13.80 | 256.20 | 283.80 |  |  |  |  |  |
| 19 | 277 | 14.18 | 262.82 | 291.18 |  |  |  |  |  |
| 20 | 285 | 14.31 | 270.69 | 299.31 |  |  |  |  |  |
| 21 | 293 | 14.23 | 278.77 | 307.23 |  |  |  |  |  |
| 22 | 300 | 13.99 | 286.01 | 313.99 |  |  |  |  |  |
| 23 | 307 | 13.72 | 293.28 | 320.72 |  |  |  |  |  |
| 24 | 314 | 13.53 | 300.47 | 327.53 |  |  |  |  |  |
| 25 | 321 | 13.45 | 307.55 | 334.45 |  |  |  |  |  |
| 26 | 328 | 13.48 | 314.52 | 341.48 |  |  |  |  |  |
| 27 | 334 | 13.59 | 320.41 | 347.59 |  |  |  |  |  |
| 28 | 341 | 13.61 | 327.39 | 354.61 |  |  |  |  |  |
| 29 | 348 | 13.56 | 334.44 | 361.56 |  |  |  |  |  |
| 30 | 355 | 13.40 | 341.60 | 368.40 |  |  |  |  |  |
| 31 | 362 | 13.21 | 348.79 | 375.21 |  |  |  |  |  |
| 32 | 368 | 13.08 | 354.92 | 381.08 |  |  |  |  |  |
| 33 | 374 | 13.05 | 360.95 | 387.05 |  |  |  |  |  |

[^7]Table 3.3.1.3Gi
Equating Summary: Writ 1 A S401 Online


[^8]

Table 3.3.1.3Ji
Reliability: Writ 1 A S401 Online

| Reliability | No. of Students | No. of Tasks | Response <br> Mode | Cronbach's <br> Alpha | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 126,315 | 4 | Hand-written <br> (HW) | .860 | 1.771 |
| Interrater <br> Reliability | Task | No. in Sample | \% AG | $\%$ AD | $\%$ NA |
|  | 1 | 60,726 | 100 | 0 | 0 |
|  | 2 | 67,496 | 100 | 0 | 0 |
|  | 3 | 71,834 | 98 | 2 | 0 |

Table 3.3.1.3Ki
Conditional Standard Error of Measurement at Cut Scores: Writ 1 A S401 Online n/a

Table 3.3.1.3Li
Accuracy and Consistency of Classification Indices: Writ 1 A S401 Online n/a

Table 3.3.1.3Aii
Complete Task Analysis and Summary: Writ 1 B/C S401 Online

| Task Type |  | Average <br> Task <br> Difficulty <br> (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average Outfit Mean Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 0.35 | 3 | 0.57 | 0.57 |
| Name |  | Task <br> Difficulty <br> (in logits) <br> 左 | Anchored? | Fit Statistics |  |
|  |  | Infit <br> Mnsq |  | Outfit <br> Mnsq |
| 1.W01B_SI_Centers_P100_1241 |  |  | 0.52 | Yes | 0.77 | 0.80 |
| 2.W01B_MS_Frogs_sakr_P100_A203_14242 |  | 0.52 | Yes | 0.60 | 0.59 |
| 3.W01C_IT_BikeRide_V2_401_14595 |  | 0.01 |  | 0.33 | 0.32 |
| Raw Score Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 0.32\% | 1.06\% | 0.60\% |  |
|  | 1 | 0.69\% | 0.84\% | 0.40\% |  |
|  | 2 | 4.05\% | 3.27\% | 1.87\% |  |
|  | 3 | 12.44\% | 9.55\% | 6.45\% |  |
|  | 4 | 60.65\% | 38.22\% | 36.49\% |  |
|  | 5 | 17.58\% | 30.85\% | 35.36\% |  |
|  | 6 | 3.58\% | 14.00\% | 15.81\% |  |
|  | 7 | 0.61\% | 2.02\% | 2.50\% |  |
|  | 8 | 0.07\% | 0.17\% | 0.46\% |  |
|  | 9 | 0.00\% | 0.01\% | 0.06\% |  |

Table 3.3.1.3Bii
DIF Analysis and Summary: Writ 1 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.W01B_SI_Centers_P 100_14241 | AA | F | AA | O |  |
| 2.Wo1B_MS_Frogs_sakr_P 100_A203_14242 | AA | M | AA | O |  |
| 3.W01C_IT_BikeRide_V2_401_14595 | AA | F | AA | H |  |

Figure 3.3.1.3Cii
Raw Scores: Writ 1 B/C S401 Online


Table 3.3.1.3Cii
Raw Score Descriptive Statistics: Writ 1 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 34,590 | 0 | 50 | 26.77 | 5.18 |
| Total | 34,590 | 0 | 50 | 26.77 | 5.18 |

Figure 3.3.1.3Dii Scale Scores: Writ 1 B/C S401 Online


Table 3.3.1.3Dii
Scale Score Descriptive Statistics: Writ 1 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 34,590 | 111 | 420 | 300.04 | 26.27 |
| Total | 34,590 | 111 | 420 | 300.04 | 26.27 |



Table 3.3.1.3Eii
Proficiency Level Distribution: Writ 1 B/C S401 Online

|  | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
| Level | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 269 | $0.78 \%$ | 269 | $0.78 \%$ |
| $\mathbf{2}$ | 5,232 | $15.13 \%$ | 5,232 | $15.13 \%$ |
| $\mathbf{3}$ | 26,214 | $75.78 \%$ | 26,214 | $75.78 \%$ |
| $\mathbf{4}$ | 2,829 | $8.18 \%$ | 2,829 | $8.18 \%$ |
| $\mathbf{5}$ | 44 | $0.13 \%$ | 44 | $0.13 \%$ |
| $\mathbf{6}$ | 2 | $0.01 \%$ | 2 | $0.01 \%$ |
| Total | 34,590 | $100.00 \%$ | 34,590 | $100.00 \%$ |

Table 3.3.1.3Fii
Raw Score to Scale Score Conversion: Writ 1 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 111^ | 126.20 | $100.00^{\wedge}$ | 286.20 | 34 | 340 | 12.35 | 327.65 | 352.35 |
| 1 | 188 | 24.30 | 163.70 | 212.30 | 35 | 346 | 12.19 | 333.81 | 358.19 |
| 2 | 203 | 16.57 | 186.43 | 219.57 | 36 | 351 | 11.98 | 339.02 | 362.98 |
| 3 | 211 | 13.40 | 197.60 | 224.40 | 37 | 357 | 11.76 | 345.24 | 368.76 |
| 4 | 217 | 11.65 | 205.35 | 228.65 | 38 | 362 | 11.55 | 350.45 | 373.55 |
| 5 | 221 | 10.55 | 210.45 | 231.55 | 39 | 366 | 11.33 | 354.67 | 377.33 |
| 6 | 225 | 9.83 | 215.17 | 234.83 | 40 | 371 | 11.14 | 359.86 | 382.14 |
| 7 | 229 | 9.32 | 219.68 | 238.32 | 41 | 376 | 10.98 | 365.02 | 386.98 |
| 8 | 232 | 8.97 | 223.03 | 240.97 | 42 | 380 | 10.87 | 369.13 | 390.87 |
| 9 | 235 | 8.70 | 226.30 | 243.70 | 43 | 385 | 10.82 | 374.18 | 395.82 |
| 10 | 238 | 8.54 | 229.46 | 246.54 | 44 | 389 | 10.85 | 378.15 | 399.85 |
| 11 | 240 | 8.43 | 231.57 | 248.43 | 45 | 393 | 10.93 | 382.07 | 403.93 |
| 12 | 243 | 8.40 | 234.60 | 251.40 | 46 | 398 | 11.12 | 386.88 | 409.12 |
| 13 | 245 | 8.40 | 236.60 | 253.40 | 47 | 403 | 11.44 | 391.56 | 414.44 |
| 14 | 248 | 8.46 | 239.54 | 256.46 | 48 | 408 | 11.95 | 396.05 | 419.95 |
| 15 | 251 | 8.57 | 242.43 | 259.57 | 49 | 413 | 12.67 | 400.33 | 425.67 |
| 16 | 254 | 8.73 | 245.27 | 262.73 | 50 | 420 | 13.77 | 406.23 | 433.77 |
| 17 | 257 | 8.94 | 248.06 | 265.94 | 51 | 428 | 15.57 | 412.43 | 443.57 |
| 18 | 260 | 9.21 | 250.79 | 269.21 | 52 | 438 | 18.77 | 419.23 | 456.77 |
| 19 | 263 | 9.53 | 253.47 | 272.53 | 53 | 457 | 26.42 | 430.58 | 483.42 |
| 20 | 266 | 9.91 | 256.09 | 275.91 | 54 | 488 | 48.63 | 439.37 | 536.63 |
| 21 | 270 | 10.28 | 259.72 | 280.28 |  |  |  |  |  |
| 22 | 274 | 10.69 | 263.31 | 284.69 |  |  |  |  |  |
| 23 | 279 | 11.09 | 267.91 | 290.09 |  |  |  |  |  |
| 24 | 283 | 11.47 | 271.53 | 294.47 |  |  |  |  |  |
| 25 | 288 | 11.79 | 276.21 | 299.79 |  |  |  |  |  |
| 26 | 294 | 12.06 | 281.94 | 306.06 |  |  |  |  |  |
| 27 | 299 | 12.30 | 286.70 | 311.30 |  |  |  |  |  |
| 28 | 305 | 12.46 | 292.54 | 317.46 |  |  |  |  |  |
| 29 | 311 | 12.57 | 298.43 | 323.57 |  |  |  |  |  |
| 30 | 317 | 12.62 | 304.38 | 329.62 |  |  |  |  |  |
| 31 | 323 | 12.62 | 310.38 | 335.62 |  |  |  |  |  |
| 32 | 329 | 12.59 | 316.41 | 341.59 |  |  |  |  |  |
| 33 | 334 | 12.49 | 321.51 | 346.49 |  |  |  |  |  |

[^9]Table 3.3.1.3Gii
Equating Summary: Writ 1 B/C S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 0.35 (0.29) | 3 |  | -0.11 (0.41) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | 0.01 |  | 0.52 | -0.58 |  | 0.16 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | 0.52 (0.00) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | (Sta. Dev.) $0.52(0.00)$ |  |  |  |
|  | Percentage Anchors |  | Average Displacement |  |  |  |
|  | 67\% |  | 0.03 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Item ID | Task Difficulty | yDisplacement | Item ID | Task Difficulty | Displacement |
|  | 14242 | 0.52 | -0.06 | 14241 | 0.52 | 0.13 |
|  | 14241 | 0.52 | 0.13 | 14242 | 0.52 | -0.06 |

[^10]

Table 3.3.1.3Jii
Reliability: Writ 1 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Response Mode | Cronbach's <br> Alpha | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Hand-written <br> $(H W)$ | .885 | 1.758 |
| Interrater <br> Reliability | 34,590 | Task | No. in Sample | \% AG | \% AD |
|  | 1 | 15,816 | 96 | 4 | \% NA |
|  | 2 | 16,302 | 94 | 5 | 0 |
|  | 3 | 16,114 | 94 | 5 | 0 |

Table 3.3.1.3Kii
Conditional Standard Error of Measurement at Cut Scores: Writ 1 B/C S401 Online n/a

Table 3.3.1.3Lii
Accuracy and Consistency of Classification Indices: Writ 1 B/C S401 Online n/a

### 3.3.1.3iii Writing 1 Across Tiers

Table 3.3.1.3Aiii
Complete Task Analysis and Summary: Writ 1 S401 Online n/a

Table 3.3.1.3Biii
DIF Analysis and Summary: Writ 1 S401 Online n/a


Table 3.3.1.3Ciii
Raw Score Descriptive Statistics: Writ 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 160,905 | 0 | 50 | 16.99 | 7.04 |
| Total | 160,905 | 0 | 50 | 16.99 | 7.04 |



Table 3.3.1.3Diii
Scale Score Descriptive Statistics: Writ 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 160,905 | 111 | 420 | 260.52 | 34.32 |
| Total | 160,905 | 111 | 420 | 260.52 | 34.32 |



Table 3.3.1.3Eiii
Proficiency Level Distribution: Writ 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 28,299 | $17.59 \%$ | 28,299 | $17.59 \%$ |
| $\mathbf{2}$ | 95,175 | $59.15 \%$ | 95,175 | $59.15 \%$ |
| $\mathbf{3}$ | 34,556 | $21.48 \%$ | 34,556 | $21.48 \%$ |
| $\mathbf{4}$ | 2,829 | $1.76 \%$ | 2,829 | $1.76 \%$ |
| $\mathbf{5}$ | 44 | $0.03 \%$ | 44 | $0.03 \%$ |
| $\mathbf{6}$ | 2 | $0.00 \%$ | 2 | $0.00 \%$ |
| Total | 160,905 | $100.00 \%$ | 160,905 | $100.00 \%$ |

Table 3.3.1.3Fiii
Raw Score to Scale Score Conversion: Writ 1 S401 Online n/a

Table 3.3.1.3Giii
Equating Summary: Writ 1 S401 Online n/a


Table 3.3.1.3Jiii
Reliability: Writ 1 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| A | 126,315 | 0.860 | 0.865 |
| B/C | 34,590 | 0.885 |  |

Table 3.3.1.3Kiii
Conditional Standard Error of Measurement at Cut Scores: Writ 1 S401 Online

| Proficiency <br> Level |  |  | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Grade | Cut Score | Tier A | Tier B/C |
| $1 / 2$ | 1 | 238 | 10.31 | 8.54 |
| $2 / 3$ | 1 | 275 | 13.96 | 10.74 |
| $3 / 4$ | 1 | 337 | 13.69 | 12.35 |
| $4 / 5$ | 1 | 382 | 13.16 | 10.74 |
| $5 / 6$ | 1 | 405 | 16.38 | 11.81 |

Table 3.3.1.3L
Accuracy and Consistency of Classification Indices: Writ (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.768 | 0.676 |  | 0.479 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.734 |  | 0.609 |  |
|  | 2 | 0.868 |  | 0.813 |  |
|  | 3 | 0.616 |  | 0.527 |  |
|  | 4 | - |  | 0.396 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.915 | 0.052 | 0.033 | 0.873 |
|  | 2/3 | 0.871 | 0.035 | 0.095 | 0.820 |
|  | 3/4 | 0.982 | 0.018 | 0.000 | 0.982 |
|  | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.1.4 Speaking 1

### 3.3.1.4i Speaking 1 Pre-A

Table 3.3.1.4Ai
Complete Task Analysis and Summary: Spek 1 Pre-A S401 Online
n/a

Table 3.3.1.4Bi
DIF Analys is and Summary: Spek 1 Pre-A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 2 | 1 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level |  | Favored <br> Group | DIF <br> Level | Favored <br> Group |
| 1.S0IP_SI_ClassroomDesk_401_14494 | AA | F | AA | O |  |
| 2.S01P_LA_Librarian_P 100_A203_14683 | AA | M | AA | O |  |
| 3.S01P_MS_ParkAdventure_401_14609 | AA | M | AA | H |  |

Figure 3.3.1.4Ci
Raw Scores: Spek 1 Pre-A S401 Online


Table 3.3.1.4Ci
Raw Score Descriptive Statistics: Spek 1 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 4,432 | 0 | 6 | 4.96 | 1.57 |
| Total | 4,432 | 0 | 6 | 4.96 | 1.57 |

Figure 3.3.1.4Di
Scale Scores: Spek 1 Pre-A S401 Online

Table 3.3.1.4Di
Scale Score Descriptive Statistics: Spek 1 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 4,432 | 106 | 174 | 160.70 | 19.86 |
| Total | 4,432 | 106 | 174 | 160.70 | 19.86 |

Table 3.3.1.4E
Proficiency Level Distribution: Spek 1 Pre-A S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
|  | 4,432 | $100.00 \%$ | 4,432 | $100.00 \%$ |
| Total | 4,432 | $100.00 \%$ | 4,432 | $100.00 \%$ |

Table 3.3.1.4Fi
Raw Score to Scale Score Conversion: Spek 1 Pre-A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $106^{\wedge}$ | 26.03 | $100.00^{\wedge}$ | 102.03 |
| 1 | $106^{\wedge}$ | 26.03 | $100.00^{\wedge}$ | 130.03 |
| 2 | 122 | 20.47 | 101.53 | 142.47 |
| 3 | 135 | 19.30 | 115.70 | 154.30 |
| 4 | 148 | 20.47 | 127.53 | 168.47 |
| 5 | $161^{*}$ | 24.28 | 141.72 | 190.28 |
| 6 | $174^{*}$ | 31.00 | 164.00 | 226.00 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.1.4Gi
Equating Summary: Spek 1 Pre-A S401 Online n/a



Table 3.3.1.4Ji
Reliability: Spek 1 Pre-A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 4,432 | 3 | .770 |  | 0.751 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 3,380 | 98 | 2 | 0 |
|  | 2 | 3,468 | 97 | 3 | 0 |
|  | 3 | 3,754 | 98 | 2 | 0 |

Table 3.3.1.4Ki
Conditional Standard Error of Measurement at Cut Scores: Spek 1 Pre-A S401 Online n/a

Table 3.3.1.4Li
Accuracy and Consistency of Classification Indices: Spek 1 Pre-A S401 Online n/a

### 3.3.1.4ii $\quad$ Speaking 1 A

Table 3.3.1.4Aii
Complete Task Analysis and Summary: Spek 1 A S401 Online n/a

Table 3.3.1.4Bii
DIF Analys is and Summary: Spek 1 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 2 | 4 | 2 | 4 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S01A_SI_Clas sroomDesk_401_V1_14494 | AA | F | AA | O |
| 2.S01A_SI_Clas sroomDesk_401_V1_14495 | AA | F | AA | O |
| 3.S01A_LA_Librarian_P 100_A203_ALT_13979 | AA | F | AA | O |
| 4.S01A_LA_Librarian_P 100_A203_ALT_13980 | AA | M | AA | O |
| 5.S01A_MS_ParkAdventure_401_14609 | AA | F | AA | H |
| 6.SO1A_MS_P arkAdventure_401_14610 | AA | M | AA | H |

Figure 3.3.1.4Cii
Raw Scores: Spek 1 A S401 Online


Table 3.3.1.4Cii
Raw Score Descriptive Statistics: Spek 1 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 43,064 | 0 | 18 | 11.68 | 2.19 |
| Total | 43,064 | 0 | 18 | 11.68 | 2.19 |



Table 3.3.1.4Dii
Scale Score Descriptive Statistics: Spek 1 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 43,064 | 106 | 391 | 256.34 | 41.88 |
| Total | 43,064 | 106 | 391 | 256.34 | 41.88 |

Figure 3.3.1.4Eii
Proficiency Level: Spek 1 A S401 Online


Table 3.3.1.4Eii
Proficiency Level Distribution: Spek 1 A S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 5,067 | $11.77 \%$ | 5,067 | $11.77 \%$ |
| $\mathbf{2}$ | 22,126 | $51.38 \%$ | 22,126 | $51.38 \%$ |
| $\mathbf{3}$ | 13,976 | $32.45 \%$ | 13,976 | $32.45 \%$ |
| $\mathbf{4}$ | 1,840 | $4.27 \%$ | 1,840 | $4.27 \%$ |
| $\mathbf{5}$ | 55 | $0.13 \%$ | 55 | $0.13 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 43,064 | $100.00 \%$ | 43,064 | $100.00 \%$ |

Table 3.3.1.4Fii
Raw Score to Scale Score Conversion: Spek 1 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $106^{\wedge}$ | 23.98 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 1 | $106^{\wedge}$ | 23.98 | $100.00^{\wedge}$ | 123.98 |
| 2 | 118 | 19.89 | $100.00^{\wedge}$ | 137.89 |
| 3 | 130 | 17.84 | 112.16 | 147.84 |
| 4 | 140 | 17.26 | 122.74 | 157.26 |
| 5 | 151 | 17.55 | 133.45 | 168.55 |
| 6 | 162 | 18.43 | 143.57 | 180.43 |
| 7 | 174 | 19.30 | 154.70 | 193.30 |
| 8 | 187 | 19.89 | 167.11 | 206.89 |
| 9 | 201 | 20.47 | 180.53 | 221.47 |
| 10 | 216 | 21.94 | 194.06 | 237.94 |
| 11 | 235 | 25.15 | 209.85 | 260.15 |
| 12 | 259 | 28.37 | 230.63 | 287.37 |
| 13 | 286 | 26.91 | 259.09 | 312.91 |
| 14 | 308 | 24.57 | 283.43 | 332.57 |
| 15 | 328 | 23.98 | 304.02 | 351.98 |
| 16 | 349 | 25.74 | 323.26 | 374.74 |
| 17 | $370^{*}$ | 30.42 | 345.58 | 406.42 |
| 18 | $391^{*}$ | 38.61 | 376.39 | 453.61 |

${ }^{\wedge}$ Truncated

* Adjusted for end of scale effect

Table 3.3.1.4Gii
Equating Summary: Spek 1 A S401 Online n/a



Table 3.3.1.4Jii
Reliability: Spek 1 A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 43,064 | 6 | .688 |  | 1.223 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 26,276 | 99 | 1 | 0 |
|  | 2 | 26,286 | 88 | 12 | 0 |
|  | 3 | 26,348 | 98 | 2 | 0 |
|  | 4 | 26,346 | 88 | 12 | 0 |
|  | 5 | 26,936 | 98 | 2 | 0 |

Table 3.3.1.4Kii
Conditional Standard Error of Measurement at Cut Scores: Spek 1 A S401 Online n/a

Table 3.3.1.4Lii
Accuracy and Consistency of Classification Indices: Spek 1 A S401 Online n/a

### 3.3.1.4iii $\quad$ Speaking 1 B/C

Table 3.3.1.4Aiii
Complete Task Analysis and Summary: Spek 1 B/C S401 Online n/a

Table 3.3.1.4Biii
DIF Analys is and Summary: Spek 1 B/C S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{array}{c}\text { DIF } \\ \text { Level }\end{array}$ | $\begin{array}{c}\text { Favoring } \\ \text { Male (M) }\end{array}$ | $\begin{array}{c}\text { Favoring } \\ \text { Female (F) }\end{array}$ | $\begin{array}{c}\text { Favoring } \\ \text { Hispanic (H) }\end{array}$ | $\begin{array}{c}\text { Favoring } \\ \text { Other (O) }\end{array}$ |
| AA | 2 | 4 | 3 | 3 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | DIF | Favored | DIF | Favored |
| (Level | Group | Level | Group |  |$]$

Figure 3.3.1.4Ciii
Raw Scores: Spek 1 B/C S401 Online


Table 3.3.1.4Ciii
Raw Score Descriptive Statistics: Spek 1 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 80,186 | 6 | 30 | 19.62 | 2.62 |
| Total | 80,186 | 6 | 30 | 19.62 | 2.62 |

Figure 3.3.1.4Diii
Scale Scores: Spek 1 B/C S401 Online


Table 3.3.1.4Diii
Scale Score Descriptive Statistics: Spek 1 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 80,186 | 106 | 407 | 285.48 | 30.78 |
| Total | 80,186 | 106 | 407 | 285.48 | 30.78 |

Table 3.3.1.4Eiii


Proficiency Level Distribution: Spek 1 B/C S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 272 | $0.34 \%$ | 272 | $0.34 \%$ |
| $\mathbf{2}$ | 16,149 | $20.14 \%$ | 16,149 | $20.14 \%$ |
| $\mathbf{3}$ | 44,825 | $55.90 \%$ | 44,825 | $55.90 \%$ |
| $\mathbf{4}$ | 18,725 | $23.35 \%$ | 18,725 | $23.35 \%$ |
| $\mathbf{5}$ | 212 | $0.26 \%$ | 212 | $0.26 \%$ |
| $\mathbf{6}$ | 3 | $0.00 \%$ | 3 | $0.00 \%$ |
| Total | 80,186 | $100.00 \%$ | 80,186 | $100.00 \%$ |

Table 3.3.1.4Fiii
Raw Score to Scale Score Conversion: Spek 1 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 6 | $106^{\wedge}$ | 23.11 | 130.89 | 177.11 |
| 7 | 163 | 16.09 | 146.91 | 179.09 |
| 8 | 172 | 16.09 | 155.91 | 188.09 |
| 9 | 180 | 15.79 | 164.21 | 195.79 |
| 10 | 189 | 15.50 | 173.50 | 204.50 |
| 11 | 197 | 15.21 | 181.79 | 212.21 |
| 12 | 205 | 15.21 | 189.79 | 220.21 |
| 13 | 213 | 15.50 | 197.50 | 228.50 |
| 14 | 221 | 16.09 | 204.91 | 237.09 |
| 15 | 230 | 16.67 | 213.33 | 246.67 |
| 16 | 241 | 17.84 | 223.16 | 258.84 |
| 17 | 252 | 19.01 | 232.99 | 271.01 |
| 18 | 265 | 19.89 | 245.11 | 284.89 |
| 19 | 279 | 19.89 | 259.11 | 298.89 |
| 20 | 292 | 19.01 | 272.99 | 311.01 |
| 21 | 303 | 18.13 | 284.87 | 321.13 |
| 22 | 314 | 17.55 | 296.45 | 331.55 |
| 23 | 325 | 16.96 | 308.04 | 341.96 |
| 24 | 334 | 16.96 | 317.04 | 350.96 |
| 25 | 344 | 17.55 | 326.45 | 361.55 |
| 26 | 355 | 18.43 | 336.57 | 373.43 |
| 27 | 368 | 19.89 | 348.11 | 387.89 |
| 28 | $381^{*}$ | 22.52 | 360.48 | 405.52 |
| 29 | $394^{*}$ | 26.03 | 380.97 | 433.03 |
| 30 | $407^{*}$ | 31.00 | 413.00 | 475.00 |

${ }^{\wedge}$ Truncated

* Adjusted for end of scale effect

Table 3.3.1.4Giii
Equating Summary: Spek 1 B/C S401 Online n/a


Table 3.3.1.4Jiii
Reliability: Spek 1 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 80,186 | 6 | .651 |  | 1.549 |
|  | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 45,470 | 84 | 16 | 0 |
|  | 2 | 45,468 | 86 | 14 | 0 |
|  | 3 | 46,086 | 89 | 11 | 0 |
|  | 4 | 46,082 | 86 | 14 | 0 |
|  | 5 | 43,866 | 85 | 15 | 0 |

Table 3.3.1.4Kiii
Conditional Standard Error of Measurement at Cut Scores: Spek 1 B/C S401 Online n/a

Table 3.3.1.4Liii
Accuracy and Consistency of Classification Indices: Spek 1 B/C S401 Online n/a

### 3.3.1.4iv Speaking 1 Across Tiers

Table 3.3.1.4Aiv
Complete Task Analysis and Summary: Spek 1 S401 Online

| Task Type |  | Average Task Difficulty (in logits) |  | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | -1.37 |  | 15 | 0.75 | 0.64 |
| Name |  | Tier | Task Difficulty (in logits) | Anchored? | Fit Statistics |  |
|  |  | Infit <br> Mnsq |  |  | Outfit <br> Mnsq |
| 1.S01A_SI_ClassroomDesk_401_V1_14494 |  |  | A** | -4.43 |  | 0.69 | 0.35 |
| 2.S01A_SI_ClassroomDesk_401_V1_14495 |  | A* | -0.14 | Yes | 0.59 | 0.59 |
| 3.S01A_LA_Librarian_P100_A203_ALT_13979 |  | A** | -4.32 |  | 0.66 | 0.27 |
| 4.S01A_LA_Librarian_P100_A203_ALT_13980 |  | A* | 0.02 | Yes | 0.82 | 0.83 |
| 5.S01A_MS_ParkAdventure_401_14609 |  | A** | -4.50 |  | 0.72 | 0.34 |
| 6.S01A_MS_ParkAdventure_401_14610 |  | A* | -0.07 | Yes | 0.82 | 0.83 |
| 7.S01C_SI_ClassroomDesk_401_V2_14603 |  | B/C* | -0.14 | Yes | 0.59 | 0.59 |
| 8.S01C_SI_ClassroomDesk_401_V2_14604 |  | B/C | 0.11 | Yes | 0.50 | 0.48 |
| 9.S01C_LS_Librarian_401_14611 |  | B/C* | 0.02 | Yes | 0.82 | 0.83 |
| 10.S01C_LS_Librarian_401_14612 |  | B/C | 0.15 | Yes | 1.12 | 1.17 |
| 11.S01C_MS_ParkAdventure_401_V2_14607 |  | B/C* | -0.07 | Yes | 0.82 | 0.83 |
| 12.S01C_MS_ParkAdventure_401_V2_14608 |  | B/C | 0.81 | Yes | 0.84 | 0.86 |
| 13.S01P_SI_ClassroomDesk_401_14494 |  | Pre-A** | -4.43 |  | 0.69 | 0.35 |
| 14.S01P_LA_Librarian_P 100_A203_14683 |  | Pre-A** | -4.32 |  | 0.66 | 0.27 |
| 15.S01P_MS_ParkAdventure_401_14609 |  | Pre-A** | -4.50 |  | 0.72 | 0.34 |
| Raw Score Distribution by Task | Task | Raw Score |  |  |  |  |
|  |  | 0 | 1 | 2 | 3 | 4 |
|  | Task 1 | 2.20\% | 8.32\% | 89.48\% | N/A | N/A |
|  | Task 2 | 1.14\% | 10.27\% | 65.58\% | 21.99\% | 1.03\% |
|  | Task 3 | 2.33\% | 7.32\% | 90.35\% | N/A | N/A |
|  | Task 4 | 1.34\% | 10.16\% | 46.87\% | 36.78\% | 4.85\% |
|  | Task 5 | 2.87\% | 5.54\% | 91.60\% | N/A | N/A |
|  | Task 6 | 0.96\% | 13.65\% | 44.35\% | 35.59\% | 5.45\% |
|  | Task 7 | 1.14\% | 10.27\% | 65.58\% | 21.99\% | 1.03\% |
|  | Task 8 | 0.37\% | 5.43\% | 77.61\% | 15.33\% | 1.25\% |
|  | Task 9 | 1.34\% | 10.16\% | 46.87\% | 36.78\% | 4.85\% |
|  | Task 10 | 1.02\% | 26.15\% | 27.48\% | 42.36\% | 2.99\% |
|  | Task 11 | 0.96\% | 13.65\% | 44.35\% | 35.59\% | 5.45\% |
|  | Task 12 | 0.98\% | 20.93\% | 46.20\% | 29.47\% | 2.42\% |
|  | Task 13 | 2.20\% | 8.32\% | 89.48\% | N/A | N/A |
|  | Task 14 | 2.33\% | 7.32\% | 90.35\% | N/A | N/A |
|  | Task 15 | 2.87\% | 5.54\% | 91.60\% | N/A | N/A |

** This task is shared between Pre-A and A.

* This task is shared between A and B/C.

Table 3.3.1.4Biv
DIF Analysis and Summary: Spek 1 S401 Online n/a

Table 3.3.1.4Civ
Raw Score Descriptive Statistics: Spek 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 127,682 | 0 | 30 | 16.43 | 4.96 |
| Total | 127,682 | 0 | 30 | 16.43 | 4.96 |

Figure 3.3.1.4Div
Scale Scores: Spek 1 S401 Online


Table 3.3.1.4Div
Scale Score Descriptive Statistics: Spek 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 127,682 | 106 | 407 | 271.32 | 42.74 |
| Total | 127,682 | 106 | 407 | 271.32 | 42.74 |

Table 3.3.1.4Ev


Proficiency Level Distribution: Spek 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 9,771 | $7.65 \%$ | 9,771 | $7.65 \%$ |
| $\mathbf{2}$ | 38,275 | $29.98 \%$ | 38,275 | $29.98 \%$ |
| $\mathbf{3}$ | 58,801 | $46.05 \%$ | 58,801 | $46.05 \%$ |
| $\mathbf{4}$ | 20,565 | $16.11 \%$ | 20,565 | $16.11 \%$ |
| $\mathbf{5}$ | 267 | $0.21 \%$ | 267 | $0.21 \%$ |
| $\mathbf{6}$ | 3 | $0.00 \%$ | 3 | $0.00 \%$ |
| Total | 127,682 | $100.00 \%$ | 127,682 | $100.00 \%$ |

Table 3.3.1.4Fiv
Raw Score to Scale Score Conversion: Spek 1 S401 Online n/a

Table 3.3.1.4Giv
Equating Summary: Spek 1 S401 Online


[^11]

Table 3.3.1.4Jiv
Reliability: Spek 1 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| Pre-A | 4,432 | 0.770 | 0.668 |
| A | 43,064 | 0.688 |  |
| B/C | 80,186 | 0.651 |  |

Table 3.3.1.4Kiv
Conditional Standard Error of Measurement at Cut Scores: Spek 1 S401 Online

| Proficiency <br> Level |  |  |  | SEM |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 |  |  |  |  |
|  | Grade | Cut Score | Tier A | Tier B/C |  |
| $2 / 3$ | 1 | 205 | 20.77 | 15.21 |  |
| $3 / 4$ | 1 | 311 | 28.37 | 19.89 |  |
| $4 / 5$ | 1 | 361 | 28.28 | 17.55 |  |
| $5 / 6$ | 1 | 403 | 45.63 | 19.01 |  |

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficeny level higher than 2.

Table 3.3.1.4L
Accuracy and Consistency of Classification Indices: Spek (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.623 | 0.504 |  | 0.247 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.774 |  | 0.602 |  |
|  | 2 | 0.704 |  | 0.509 |  |
|  | 3 | 0.585 |  | 0.558 |  |
|  | 4 | - |  | 0.262 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.961 | 0.015 | 0.024 | 0.941 |
|  | 2/3 | 0.824 | 0.039 | 0.137 | 0.753 |
|  | 3/4 | 0.837 | 0.163 | 0.000 | 0.766 |
|  | 4/5 | 0.998 | 0.002 | 0.000 | 0.997 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.2 Grades: 2-3

### 3.3.2.1 Listening 2-3

Table 3.3.2.1A
Complete Item Analysis and Summary: List 2-3 S401 Online

| Item Type | Average Item Difficulty (in logits) | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | -1.16 | 54 | 0.68 | 0.98 | 1.03 |
|  | Item |  |  | Fit <br> Statistics |  |
| Name | Difficulty <br> (in logits) | Anchored? | P-value | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 1.L23B_SI_ArtClas s_P 100_Screen_2_12825 | -3.26 |  | 0.95 | 1.07 | 2.25 |
| 2.L23B_SI_ArtClas s_P 100_Screen_3_12828 | -2.29 |  | 0.90 | 1.10 | 1.94 |
| 3.L23B_SI_ArtClas s_P 100_Screen_4_12830 | -0.87 |  | 0.79 | 1.06 | 1.15 |
| 4.L23A_SI_ThankYouCards_P 100_Screen_2_13788 | -3.85 |  | 0.97 | 1.09 | 3.13 |
| 5.L23A_SI_Thank YouCards_P 100_Screen_3_13789 | -0.32 |  | 0.71 | 1.19 | 1.39 |
| 6.L23A_SI_Thank YouCards_P 100_Screen_4_13790 | -2.68 |  | 0.93 | 0.94 | 1.24 |
| 7.L23A_LA_Mina AndGinger_jc_P 100_A202_Screen_2_13797 | -3.68 |  | 0.76 | 0.84 | 0.70 |
| 8.L23A_LA_Mina AndGinger_jc_P 100_A202_Screen_3_13798 | -2.72 |  | 0.59 | 0.87 | 0.81 |
| 9.L23A_LA_MinaAndGinger_jc_P 100_A202_Screen_4_13799 | -2.87 |  | 0.64 | 0.91 | 0.84 |
| 10.L23B_MA_TellingTime_P 100_Screen_2_13904 | -3.24 |  | 0.71 | 0.93 | 0.85 |
| 11.L23B_MA_TellingTime_P 100_Screen_3_13905 | -3.26 |  | 0.69 | 0.93 | 0.85 |
| 12.L23B_MA_TellingTime_P 100_Screen_4_13906 | -2.24 |  | 0.52 | 0.91 | 0.88 |
| 13.L23A_SS_AMapOfOakValley_P 100_Screen_2_12730 | -2.07 |  | 0.43 | 0.99 | 1.01 |
| 14.L23A_SS_AMapOfOakValley_P 100_Screen_3_12731 | -3.89 |  | 0.80 | 0.92 | 0.81 |
| 15.L23A_SS_AMapOfOakValley_P 100_Screen_4_12732 | -2.36 |  | 0.49 | 0.98 | 0.98 |
| 16.L23C_SC_BirdAdaptations_P 100_Screen_2_11544 | -3.56 |  | 0.78 | 0.93 | 0.85 |
| 17.L23C_SC_B irdAdaptations_P 100_Screen_3_11545 | -2.32 |  | 0.48 | 1.00 | 1.01 |
| 18.L23C_SC_BirdAdaptations_P 100_Screen_4_11546 | -2.17 |  | 0.39 | 1.06 | 1.13 |
| 19.L23B_LA_PresentforGrandma_P 100_Screen_2_13440 | -3.06 |  | 0.94 | 0.95 | 0.81 |
| 20.L23B_LA_PresentforGrandma_P 100_Screen_3_13441 | -0.57 |  | 0.63 | 0.89 | 0.79 |
| 21.L23B_LA_PresentforGrandma_P 100_Screen_4_13442 | -1.50 |  | 0.77 | 0.96 | 0.88 |
| 22.L23C_MA_GettingTo School_P 100_Screen_2_12956 | -3.09 |  | 0.91 | 0.95 | 0.82 |
| 23.L23C_MA_GettingTo School_P 100_Screen_3_12957 | -1.12 |  | 0.64 | 0.96 | 0.92 |
| 24.L23C_MA_GettingTo School_P 100_Screen_4_12971 | 0.35 |  | 0.33 | 1.02 | 1.04 |
| 25.L23C_SS_AMapOfOakValley_P 100_Screen_2_12813 | -1.83 |  | 0.79 | 0.93 | 0.86 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 26.L23C_SS_AMapOfOakValley_P 100_Screen_3_12814 | -1.26 |  | 0.68 | 0.93 | 0.89 |
| 27.L23C_SS_AMapOfOakValley_P 100_Screen_4_12815 | -0.47 |  | 0.51 | 0.93 | 0.90 |
| 28.L23B_SC_WaterCycle_P 100_A202_Screen_2_13910 | -2.33 |  | 0.85 | 0.96 | 0.90 |
| 29.L23B_SC_WaterC ycle_P 100_A202_Screen_3_13911 | -0.87 |  | 0.55 | 0.99 | 0.98 |
| 30.L23B_SC_WaterCycle_P 100_A 202 _Screen_4_13912 | -0.24 |  | 0.45 | 0.94 | 0.93 |
| 31.L23B_LA_TheEmptyP ot_P 100_Screen_2_12580 | -0.96 |  | 0.55 | 1.01 | 1.00 |
| 32.L23B_LA_TheEmptyP ot_P 100_Screen_3_12582 | -1.01 |  | 0.63 | 0.97 | 0.95 |
| 33.L23B_LA_TheEmptyP ot_P 100_Screen_4_12584 | -1.52 |  | 0.74 | 0.94 | 0.90 |
| 34.L23B_MA_LinearMeas urement_P 100_Screen_2_12988 | -1.07 |  | 0.66 | 0.95 | 0.93 |
| 35.L23B_MA_Line arMeas urement_P 100_Screen_3_12990 | 0.29 |  | 0.31 | 0.99 | 0.99 |
| 36.L23B_MA_LinearMeasurement_P 100_Screen_4_12991 | -0.07 |  | 0.39 | 0.97 | 0.97 |
| 37.L23C_LA_Cro wnAndThe Coins_P 100_Screen_2_12705 | -0.17 |  | 0.77 | 1.02 | 1.03 |
| 38.L23C_LA_Cro wnAndTheCoins_P 100_Screen_3_12706 | -1.64 | Yes | 0.93 | 0.98 | 0.93 |
| 39.L23C_LA_Cro wnAndTheCo ins_P 100_Screen_4_12707 | -0.30 |  | 0.78 | 0.94 | 0.87 |
| 40.L23B_MA_GettingTo School_P 100_Screen_2_12953 | 0.13 | Yes | 0.75 | 1.10 | 1.21 |
| 41.L23B_MA_GettingTo School_P 100_Screen_3_12954 | -1.65 | Yes | 0.96 | 0.99 | 0.98 |
| 42.L23B_MA_GettingTo School_P 100_Screen_4_12955 | 1.23 | Yes | 0.54 | 1.05 | 1.06 |
| 43.L23B_SS_AMapOfOakValley_P 100_Screen_2_12733 | -1.69 |  | 0.94 | 0.95 | 0.79 |
| 44.L23B_SS_AMapOfOakValley_P 100_Screen_3_12734 | 0.26 | Yes | 0.78 | 1.00 | 1.00 |
| 45.L23B_SS_AMapOfOakValley_P 100_Screen_4_12735 | 1.28 | Yes | 0.52 | 1.02 | 1.03 |
| 46.L23B_SC_Habitats_P 100_Screen_2_12785 | 0.00 | Yes | 0.80 | 0.97 | 0.93 |
| 47.L23B_SC_Habitats_P 100_Screen_3_12786 | 1.56 | Yes | 0.51 | 0.92 | 0.90 |
| 48.L23B_SC_Habitats_P 100_Screen_4_12787 | -0.11 | Yes | 0.83 | 0.95 | 0.91 |
| 49.L23C_LA_TheEmptyP ot_P 100_Screen_2_12586 | -0.24 |  | 0.83 | 0.96 | 0.93 |
| 50.L23C_LA_TheEmptyP ot_P 100_Screen_3_12587 | 1.82 |  | 0.46 | 0.97 | 0.95 |
| 51.L23C_LA_TheEmptyP ot_P 100_Screen_4_12588 | 0.80 |  | 0.68 | 0.95 | 0.94 |
| 52.L23C_MA_Metric Measurement_P 100_Screen_2_12992 | 1.62 |  | 0.47 | 0.99 | 0.98 |
| 53.L23C_MA_Metric Measurement_P 100_Screen_3_12993 | 0.63 | Yes | 0.65 | 1.00 | 1.01 |
| 54.L23C_MA_Metric Measurement_P 100_Screen_4_12994 | 1.84 | Yes | 0.43 | 0.99 | 0.98 |

Table 3.3.2.1B
DIF Analysis and Summary: List 2-3 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (0) |
| A | 30 | 22 | 27 | 27 |
| B | 1 | 0 | 0 | 0 |
| C | 1 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.L23B_S I_ArtClas s_P 100_Screen_2_12825 | A | M | A | H |
| 2.L23B_SI_ArtClas s_P 100_Screen_3_12828 | A | M | A | O |
| 3.L23B_SI_ArtClass_P 100_Screen_4_12830 | A | M | A | O |
| 4.L23A_SI_Thank YouCards_P 100_Screen_2_13788 | C | M | A | O |
| 5.L23A_SI_Thank YouCards_P 100_Screen_3_13789 | A | F | A | H |
| 6.L23A_SI_ThankYouCards_P 100_Screen_4_13790 | A | M | A | O |
| 7.L23A_LA_M ina AndGinger_jc_P 100_A202_Screen_2_13797 | A | F | A | O |
| 8.L23A_LA_M ina AndGinger_jc_P 100_A202_Screen_3_13798 | A | M | A | H |
| 9.L23A_LA_Mina AndGinger_jc_P 100_A202_Screen_4_13799 | A | M | A | O |
| 10.L23B_MA_TellingTime_P 100_Screen_2_13904 | A | F | A | H |
| 11.L23B_M A_Te llingTime_P 100_Screen_3_13905 | A | M | A | H |
| 12.L23B_MA_TellingTime_P 100_Screen_4_13906 | A | F | A | H |
| 13.L23A_SS_AMapOfOakValley_P 100_Screen_2_12730 | A | F | A | H |
| 14.L23A_SS_AMapOfOakValley_P 100_Screen_3_12731 | A | F | A | H |
| 15.L23A_SS_AMapOfOakValley_P 100_Screen_4_12732 | A | M | A | O |
| 16.L23C_SC_B irdAdaptations_P 100_Screen_2_11544 | A | F | A | O |
| 17.L23C_SC_B irdAdaptations_P 100_Screen_3_11545 | A | M | A | O |
| 18.L23C_SC_B irdAdaptations_P 100_Screen_4_11546 | A | F | A | H |
| 19.L23B_LA_PresentforGrandma_P 100_Screen_2_13440 | A | M | A | H |
| 20.L23B_LA_PresentforGrandma_P 100_Screen_3_13441 | B | M | A | O |
| 21.L23B_LA_PresentforGrandma_P 100_Screen_4_13442 | A | M | A | H |
| 22.L23C_MA_GettingTo School_P 100_Screen_2_12956 | A | M | A | O |
| 23.L23C_MA_GettingTo School_P 100_Screen_3_12957 | A | M | A | O |
| 24.L23C_MA_GettingTo School_P 100_Screen_4_12971 | A | M | A | H |
| 25.L23C_SS_AMapOfOakValley_P 100_Screen_2_12813 | A | F | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 26.L23C_SS_AMapOfOakValley_P 100_Screen_3_12814 | A | F | A | H |
| 27.L23C_SS_AMapOfOakValle y_P 100_Screen_4_12815 | A | F | A | H |
| 28.L23B_SC_WaterCycle_P 100_A202_Screen_2_13910 | A | F | A | O |
| 29.L23B_SC_WaterCycle_P 100_A202_Screen_3_13911 | A | M | A | H |
| 30.L23B_SC_WaterCycle_P 100_A 202 _Screen_4_13912 | A | F | A | O |
| 31.L23B_LA_TheEmptyP ot_P 100_Screen_2_12580 | A | F | A | H |
| 32.L23B_LA_TheEmptyP ot_P 100_Screen_3_12582 | A | M | A | O |
| 33.L23B_LA_TheEmptyP ot_P 100_Screen_4_12584 | A | M | A | O |
| 34.L23B_MA_LinearMeasurement_P 100_Screen_2_12988 | A | F | A | H |
| 35.L23B_MA_LinearMe as urement_P 100_Screen_3_12990 | A | F | A | H |
| 36.L23B_MA_LinearMeasurement_P 100_Screen_4_12991 | A | M | A | O |
| 37.L23C_LA_Cro wnAndTheCo ins_P 100_Screen_2_12705 | A | M | A | H |
| 38.L23C_LA_Cro wnAndTheCo ins_P 100_Screen_3_12706 | A | M | A | O |
| 39.L23C_LA_Cro wnAndTheCo ins _P 100_Screen_4_12707 | A | M | A | H |
| 40.L23B_MA_GettingTo School_P 100_Screen_2_12953 | A | M | A | O |
| 41.L23B_MA_GettingTo School_P 100_Screen_3_12954 | A | M | A | H |
| 42.L23B_MA_GettingTo School_P 100_Screen_4_12955 | A | F | A | O |
| 43.L23B_SS_AMapOfOakValley_P 100_Screen_2_12733 | A | M | A | O |
| 44.L23B_SS_AMapOfOakValley_P 100_Screen_3_12734 | A | M | A | O |
| 45.L23B_SS_AMapOfOakValley_P 100_Screen_4_12735 | A | F | A | O |
| 46.L23B_SC_Habitats_P 100_Screen_2_12785 | A | F | A | H |
| 47.L23B_SC_Habitats_P 100_Screen_3_12786 | A | F | A | H |
| 48.L23B_SC_Habitats_P 100_Screen_4_12787 | A | F | A | O |
| 49.L23C_LA_TheEmptyP ot_P 100_Screen_2_12586 | A | M | A | O |
| 50.L23C_LA_TheEmptyP ot_P 100_Screen_3_12587 | A | M | A | H |
| 51.L23C_LA_TheEmptyP ot_P 100_Screen_4_12588 | A | M | A | H |
| 52.L23C_MA_Metric Meas urement_P 100_Screen_2_12992 | A | M | A | H |
| 53.L23C_MA_Metric Meas urement_P 100_Screen_3_12993 | A | F | A | O |
| 54.L23C_MA_Metric Meas urement_P 100_Screen_4_12994 | A | M | A | H |

Figure 3.3.2.1C
Raw Scores: List 2-3 S401 Online
n/a

Table 3.3.2.1C
Raw Score Descriptive Statistics: List 2-3 S401 Online n/a


Figure 3.3.2.1E
Proficiency Level: List 2-3 S401 Online


Table 3.3.2.1D
Scale Score Descriptive Statistics: List 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 145,452 | 108 | 446 | 328.06 | 54.92 |
| $\mathbf{3}$ | 153,970 | 112 | 446 | 351.63 | 55.54 |
| Total | 299,422 | 108 | 446 | 340.18 | 56.48 |

Table 3.3.2.1E
Proficiency Level Distribution: List 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 11,107 | $7.64 \%$ | 11,937 | $7.75 \%$ | 23,044 | $7.70 \%$ |
| $\mathbf{2}$ | 17,509 | $12.04 \%$ | 16,179 | $10.51 \%$ | 33,688 | $11.25 \%$ |
| $\mathbf{3}$ | 30,031 | $20.65 \%$ | 22,228 | $14.44 \%$ | 52,259 | $17.45 \%$ |
| $\mathbf{4}$ | 8,594 | $5.91 \%$ | 10,069 | $6.54 \%$ | 18,663 | $6.23 \%$ |
| $\mathbf{5}$ | 18,255 | $12.55 \%$ | 25,313 | $16.44 \%$ | 43,568 | $14.55 \%$ |
| $\mathbf{6}$ | 59,956 | $41.22 \%$ | 68,244 | $44.32 \%$ | 128,200 | $42.82 \%$ |
| Total | 145,452 | $100.00 \%$ | 153,970 | $100.00 \%$ | 299,422 | $100.00 \%$ |

Table 3.3.2.1F
Raw Score to Scale Score Conversion: List 2-3 S401 Online n/a

Table 3.3.2.1G
Equating Summary: List 2-3 S401 Online


[^12]Figure 3.3.2.1H
Test Characteristic Curve: List 2-3 S401 Online n/a

Figure 3.3.2.1I


Table 3.3.2.1J
Reliability: List 2-3 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 299,422 | 54 | .86 |

Table 3.3.2.1K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 2-3 S401 Online

| Proficiency <br> Level | Grade | Cut Score | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1} / \mathbf{2}$ | 2 | 245 | 860 | 19.90 | 20.41 | 19.96 | 0.17 |
|  | 3 | 262 | N/A | N/A | N/A | N/A | N/A |
| $\mathbf{2 / 3}$ | 2 | 283 | 90 | 17.86 | 17.86 | 17.86 | 0.00 |
|  | 3 | 300 | 428 | 17.35 | 17.35 | 17.35 | 0.00 |
| $\mathbf{3 / 4}$ | 2 | 314 | 185 | 17.86 | 18.88 | 18.29 | 0.39 |
|  | 3 | 331 | 216 | 19.90 | 19.90 | 19.90 | 0.00 |
| $\mathbf{4 / 5}$ | 2 | 330 | 336 | 18.37 | 19.90 | 19.06 | 0.67 |
|  | 3 | 349 | 82 | 20.41 | 22.45 | 22.03 | 0.44 |

Table 3.3.2.1Li
Accuracy and Consistency of Classification Indices: List (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.629 | 0.545 |  | 0.404 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.789 |  | 0.622 |  |
|  | 2 | 0.514 |  | 0.396 |  |
|  | 3 | 0.591 |  | 0.474 |  |
|  | 4 | 0.177 |  | 0.127 |  |
|  | 5 | 0.346 |  | 0.252 |  |
|  | 6 | 0.894 |  | 0.832 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.961 | 0.014 | 0.025 | 0.943 |
|  | 2/3 | 0.916 | 0.039 | 0.045 | 0.884 |
|  | 3/4 | 0.894 | 0.033 | 0.073 | 0.853 |
|  | 4/5 | 0.903 | 0.048 | 0.049 | 0.854 |
|  | 5/6 | 0.887 | 0.073 | 0.040 | 0.842 |

Table 3.3.2.1 Lii
Accuracy and Consistency of Classification Indices: List (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.639 | 0.557 |  | 0.402 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.800 |  | 0.651 |  |
|  | 2 | 0.518 |  | 0.399 |  |
|  | 3 | 0.499 |  | 0.381 |  |
|  | 4 | 0.208 |  | 0.148 |  |
|  | 5 | 0.422 |  | 0.317 |  |
|  | 6 | 0.871 |  | 0.814 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.963 | 0.013 | 0.024 | 0.946 |
|  | 2/3 | 0.929 | 0.029 | 0.043 | 0.902 |
|  | 3/4 | 0.909 | 0.033 | 0.058 | 0.870 |
|  | 4/5 | 0.902 | 0.052 | 0.046 | 0.856 |
|  | 5/6 | 0.875 | 0.070 | 0.055 | 0.827 |

### 3.3.2.2 Reading 2-3

Table 3.3.2.2A
Complete Item Analysis and Summary: Read 2-3 S401 Online

|  | Average <br> Item <br> Difficulty <br> (in logits) | No. of Items | Average <br> Infit <br> Merage <br> P-value | Average <br> Outfit <br> Mean <br> Square |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Square |  |  |  |  |$|$| Item Type |
| :--- |
| Selected Response |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 26.R23C_LA_CarverandFord_203_P 100_A203_Screen_3_13366 | -0.10 | Yes | 0.46 | 0.93 | 0.89 |
| 27.R23C_LA_CarverandFord_203_P 100_A203_Screen_4_13367 | -0.19 |  | 0.39 | 1.15 | 1.17 |
| 28.R23C_MA_BookFair_Screen_2_13368 | -0.86 |  | 0.54 | 0.99 | 0.99 |
| 29.R23C_MA_BookFair_Screen_3_13369 | -0.43 |  | 0.48 | 0.90 | 0.87 |
| 30.R23C_MA_BookFair_Screen_4_13370 | -0.95 |  | 0.56 | 0.98 | 0.95 |
| 31.R23C_SS_UrbanNeighborhood_P 100_A301_Screen_2_13371 | -0.21 |  | 0.36 | 1.03 | 1.02 |
| 32.R23C_SS_UrbanNeighborho od_P 100_A301_Screen_3_13372 | -1.27 |  | 0.65 | 0.90 | 0.85 |
| 33.R23C_SS_UrbanNeighborhood_P 100_A301_Screen_4_13373 | -0.63 |  | 0.52 | 0.85 | 0.82 |
| 34.R23B_SC_TempClimate_401_V1_Screen_2_13932 | -0.08 |  | 0.37 | 1.01 | 1.01 |
| 35.R23B_SC_TempClimate_401_V1_Screen_3_13936 | 0.45 |  | 0.24 | 1.04 | 1.04 |
| 36.R23B_SC_TempClimate_401_V1_Screen_4_13937 | 0.42 |  | 0.24 | 1.06 | 1.10 |
| 37.R23B_LA_ZebraStripes_203_P 100_A301_altıScreen_2_13410 | -0.34 | Yes | 0.44 | 0.94 | 0.93 |
| 38.R23B_LA_ZebraStripes_203_P 100_A301_altl_Screen_3_13411 | 0.21 |  | 0.31 | 1.00 | 0.99 |
| 39.R23B_LA_ZebraStripes_203_P 100_A301_alt1_Screen_4_13412 | -0.11 | Yes | 0.41 | 0.93 | 0.92 |
| 40.R23B_MA_Library_P 100_A201_altıScreen_2_13413 | -0.47 | Yes | 0.52 | 0.90 | 0.89 |
| 41.R23B_MA_Library_P 100_A201_alt1_Screen_3_13414 | -0.76 | Yes | 0.50 | 0.93 | 0.92 |
| 42.R23B_MA_Library_P 100_A201_altıScreen_4_13415 | -0.14 | Yes | 0.44 | 0.98 | 0.97 |
| 43.R23B_SS_OurNeighbo rho od_203_P 100_A203_Screen_2_13350 | -0.32 | Yes | 0.43 | 0.97 | 0.96 |
| 44.R23B_SS_OurNeighborhoo d_203_P 100_A203_Screen_3_13351 | -0.24 | Yes | 0.44 | 0.96 | 0.96 |
| 45.R23B_SS_OurNe ighborhoo d_203_P 100_A203_Screen_4_13352 | 0.05 |  | 0.34 | 1.02 | 1.03 |
| 46.R23A_SC_Birds_dode_P 100_A 301_Screen_2_13280 | -3.26 |  | 0.93 | 0.96 | 0.77 |
| 47.R23A_SC_Birds_dode_P 100_A301_Screen_3_13281 | -0.62 | Yes | 0.48 | 0.90 | 0.89 |
| 48.R23A_SC_Birds_dode_P 100_A 301_Screen_4_13282 | -0.23 |  | 0.42 | 0.85 | 0.84 |
| 49.R23A_LA_KittenFable_203_P 100_A203FT_Screen_2_13338 | 0.80 |  | 0.49 | 1.22 | 1.31 |
| 50.R23A_LA_KittenFable_203_P 100_A203FT_Screen_3_13339 | 0.38 | Yes | 0.55 | 0.87 | 0.82 |
| 51.R23A_LA_KittenFable_203_P 100_A203FT_Screen_4_13340 | -0.25 |  | 0.69 | 0.84 | 0.72 |
| 52.R23C_MA_P uppetSho w_dode_P 100_A203_Screen_2_13322 | -0.01 | Yes | 0.68 | 0.95 | 0.91 |
| 53.R23C_MA_P uppetSho w_dode_P 100_A203_Screen_3_13323 | 0.06 | Yes | 0.70 | 0.86 | 0.77 |
| 54.R23C_MA_P uppetSho w_dode_P 100_A203_Screen_4_13324 | -0.63 |  | 0.75 | 0.88 | 0.78 |
| 55.R23B_SS_Boats LongAgo_401_V2_Screen_2_14583 | 0.05 |  | 0.70 | 0.92 | 0.89 |
| 56.R23B_SS_Boats LongAgo_401_V2_Screen_3_14584 | 0.59 |  | 0.53 | 1.13 | 1.16 |
| 57.R23B_SS_Boats LongAgo_401_V2_Screen_4_4585 | 0.64 |  | 0.53 | 1.00 | 0.98 |
| 58.R23B_SC_AnimalMo vement_dode_P 100_A301_altl_Screen_2_13419 | -1.67 | Yes | 0.94 | 0.94 | 0.74 |
| 59.R23B_SC_AnimalMovement_dode_P 100_A301_altıScreen_3_13420 | -0.90 | Yes | 0.85 | 0.93 | 0.88 |
| 60.R23B_SC_AnimalMovement_dode_P 100_A301_altı_Screen_4_13421 | 0.70 |  | 0.63 | 0.96 | 0.93 |


| Name | Item (in logits) | Anchored? | P-value | FitStatistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | Outfit <br> Mnsq |
| 61R23B_LA_ZebraStripes_203_P 100_A 301_Screen_2_13344 | 0.33 |  | 0.70 | 0.94 | 0.92 |
| 62.R23B_LA_Zebrastripes_203_P 100_A 301 _Screen_3_13345 | 1.53 |  | 0.49 | 0.90 | 0.88 |
| 63 2R23B_LA_Zebrastripes_203_P 100_A 301 _Screen_4_13346 | 0.60 |  | 0.67 | 0.91 | 0.85 |
| 64.R23B_MA_LibraryPP 100_A20_Screen_2_13356 | 0.07 | Yes | 0.72 | 0.90 | 0.85 |
| 65. R23B_MA_Librar_P P 100_A201_Screen_3_B357 | -0.13 | Yes | 0.75 | 0.90 | 0.82 |
| 66. R23B_MA_Librar_P P 100_A20_Screen_4_-3358 | 0.54 |  | 0.66 | 0.97 | 0.95 |
| 67.R23B_SS_OurNeighborho od_203_P 100_A203_all_Scree __2_3416 | 0.46 |  | 0.68 | 0.89 | 0.84 |
| 68.R23B_SS_OurNeighborho od_203_P 100_A203_all_Screen_3_1347 | 0.67 |  | 0.66 | 0.90 | 0.84 |
| 69.R23B_SS_OurNeighborho od_203_P 100_A203_all_Screen_4_3418 | 1.24 |  | 0.55 | 0.95 | 0.93 |
| 70.R23C_SC_Insects_dode_P 100_A 301 _Screen_2_13374 | -0.30 |  | 0.81 | 0.94 | 0.88 |
| 71.R23C_SC_Insects_dode_P 100_A301_Screen_3_13375 | 1.34 |  | 0.54 | 0.96 | 0.94 |
| 72.R23C_SC_Insects_dode_P 100_A301_Screen_4_13376 | 0.66 |  | 0.65 | 1.05 | 1.06 |

Table 3.3.2.2B
DIF Analysis and Summary: Read 2-3 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| A | 33 | 39 | 33 | 39 |
| B | 0 | 0 | 0 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.R23A_SI_Getting_Ready_P 100_A203_Screen_2_13325 | A | M | A | H |
| 2.R23A_SI_Getting_Ready_P 100_A203_Screen_3_13326 | A | F | A | H |
| 3.R23A_SI_Getting_Ready_P 100_A203_Screen_4_13327 | A | M | A | O |
| 4.R23B_SI_RunningClub_401_V2_Screen_2_14589 | A | M | A | O |
| 5.R23B_SI_RunningClub_401_V2_Screen_3_14590 | A | M | A | O |
| 6.R23B_SI_RunningClub_401_V2_Screen_4_14591 | A | F | A | H |
| 7.R23A_LA_BenFranklin_203_P 100_A301_Screen_2_13331 | A | M | A | O |
| 8.R23A_LA_BenFranklin_203_P 100_A 301_Screen_3_13332 | A | F | A | O |
| 9.R23A_LA_BenFranklin_203_P 100_A 301_Screen_4_13333 | A | F | A | O |
| 10.R23A_MA_BuyingSchoo ISupplies_P 100_A201_Screen_2_13392 | A | M | A | O |
| 11.R23A_MA_B uyingSchoolS upplies_P 100_A201_Screen_3_13393 | A | F | A | O |
| 12.R23A_MA_BuyingSchoolSupplies_P 100_A201_Screen_4_13394 | A | M | A | O |
| B.R23A_SS_Homes oftheP ast_203_P 100_A 301 _Screen_2_13277 | A | M | A | O |
| 14.R23A_SS_Homes oftheP ast_203_P 100_A301_Screen_3_13278 | A | F | A | H |
| 15.R23A_SS_Homes oftheP ast_203_P 100_A 301 _Screen_4_13279 | A | F | A | O |
| 16.R23B_SC_AnimalMo vement_dode_P 100_A301_Screen_2_13292 | A | F | A | H |
| 17.R23B_SC_AnimalMo vement_dode_P 100_A301_Screen_3_13293 | A | F | A | H |
| 18.R23B_SC_AnimalMo vement_dode_P 100_A301_Screen_4_13294 | A | F | A | O |
| 19.R23B_LA_IvanAtTheMarket_P 100_A202_Screen_2_13353 | A | M | A | H |
| 20.R23B_LA_IvanAtTheMarket_P 100_A202_Screen_3_13354 | A | F | A | H |
| 21.R23B_LA_IvanAtTheMarket_P 100_A202_Screen_4_13355 | A | M | A | O |
| 22.R23B_MA_Cupcakes_P 100_A 201 _Screen_2_13401 | A | M | A | H |
| 23.R23B_MA_Cupcakes_P 100_A 201 _Screen_3_13402 | A | F | A | O |
| 24.R23B_MA_Cupcakes_P 100_A201_Screen_4_13403 | A | F | A | O |
| 25.R23C_LA_CarverandFord_203_P 100_A203_Screen_2_13365 | A | F | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | DIF Level | Favored Group | DIF Level | Favored Group |
| 26.R23C_LA_CarverandFord_203_P 100_A203_Screen_3_13366 | A | M | A | H |
| 27.R23C_LA_CarverandFord_203_P 100_A203_Screen_4_13367 | A | M | A | O |
| 28.R23C_MA_BookFair_Screen_2_13368 | A | F | A | H |
| 29.R23C_MA_BookFair_Screen_3_13369 | A | M | A | O |
| 30.R23C_MA_BookFair_Screen_4_13370 | A | F | A | H |
| 31.R23C_SS_UrbanNeighborhood_P 100_A301_Screen_2_13371 | A | F | A | H |
| 32.R23C_SS_UrbanNeighborhood_P 100_A301_Screen_3_13372 | A | M | A | O |
| 33.R23C_SS_UrbanNeighborhood_P 100_A301_Screen_4_13373 | A | F | A | H |
| 34.R23B_SC_TempClimate_401_V1_Screen_2_13932 | A | F | A | O |
| 35.R23B_SC_TempClimate_401_V1_Screen_3_13936 | A | F | A | H |
| 36.R23B_SC_TempClimate_401_V1_Screen_4_13937 | A | M | A | H |
| 37.R23B_LA_ZebraStripes_203_P 100_A 301_alt1_Screen_2_13410 | A | F | A | O |
| 38.R23B_LA_ZebraStripes_203_P 100_A301_alt 1_S creen_3_13411 | A | M | A | O |
| 39.R23B_LA_ZebraStripes_203_P 100_A301_alt1_Screen_4_13412 | A | M | A | H |
| 40.R23B_MA_Library_P 100_A201_alt1_Screen_2_13413 | A | M | A | H |
| 41.R23B_MA_Library_P 100_A201_alt1_Screen_3_13414 | A | F | A | O |
| 42.R23B_MA_Library_P 100_A201_alt1_Screen_4_13415 | A | M | A | O |
| 43.R23B_SS_OurNeighborhood_203_P 100_A203_Screen_2_13350 | A | F | A | H |
| 44.R23B_SS_OurNeighborhood_203_P 100_A203_Screen_3_13351 | A | F | A | O |
| 45.R23B_SS_OurNeighborhood_203_P 100_A203_Screen_4_13352 | A | F | A | H |
| 46.R23A_SC_Birds_dode_P 100_A301_Screen_2_13280 | A | M | A | H |
| 47.R23A_SC_B irds_dode_P 100_A301_Screen_3_13281 | A | M | A | O |
| 48.R23A_SC_B irds_dode_P 100_A301_Screen_4_13282 | A | M | A | H |
| 49.R23A_LA_KittenFable_203_P 100_A203FT_Screen_2_13338 | A | F | A | O |
| 50.R23A_LA_KittenFable_203_P 100_A203FT_Screen_3_13339 | A | M | A | H |
| 51.R23A_LA_KittenFable_203_P 100_A203FT_Screen_4_13340 | A | M | A | H |
| 52.R23C_MA_P uppetShow_dode_P 100_A203_Screen_2_13322 | A | M | A | H |
| 53.R23C_MA_P uppetShow_dode_P 100_A203_Screen_3_13323 | A | M | A | H |
| 54.R23C_MA_P uppetShow_dode_P 100_A203_Screen_4_13324 | A | M | A | O |
| 55.R23B_SS_BoatsLongAgo_401_V2_Screen_2_14583 | A | F | A | H |
| 56.R23B_SS_Boats LongAgo_401_V2_Screen_3_14584 | A | F | A | H |
| 57.R23B_SS_Boats LongAgo_401_V2_Screen_4_14585 | A | F | A | O |
| 58.R23B_SC_AnimalMovement_dode_P 100_A301_alt 1_Screen_2_13419 | A | F | A | O |
| 59.R23B_SC_AnimalMovement_dode_P 100_A301_alt 1_Screen_3_13420 | A | F | A | O |
| 60.R23B_SC_AnimalMovement_dode_P 100_A301_alt 1_Screen_4_13421 | A | F | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | DIF Level | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 61.R23B_LA_ZebraStripes_203_P 100_A301_Screen_2_13344 | A | M | A | H |
| 62.R23B_LA_ZebraStripes_203_P 100_A301_Screen_3_13345 | A | F | A | O |
| 63.R23B_LA_ZebraStripes_203_P 100_A301_Screen_4_13346 | A | M | A | O |
| 64.R23B_MA_Library_P 100_A201_Screen_2_13356 | A | F | A | O |
| 65.R23B_MA_Library_P 100_A201_Screen_3_13357 | A | F | A | O |
| 66.R23B_MA_Library_P 100_A201_Screen_4_13358 | A | M | A | H |
| 67.R23B_SS_OurNeighborhood_203_P 100_A203_alt1_Screen_2_13416 | A | F | A | O |
| 68.R23B_SS_OurNeighborhood_203_P 100_A203_alt1_Screen_3_13417 | A | F | A | O |
| 69.R23B_SS_OurNeighborhood_203_P 100_A203_alt1_Screen_4_13418 | A | M | A | H |
| 70.R23C_SC_Insects_dode_P 100_A301_Screen_2_13374 | A | F | A | H |
| 71.R23C_SC_Insects_dode_P 100_A301_Screen_3_13375 | A | M | A | O |
| 72.R23C_SC_Insects_dode_P 100_A301_Screen_4_13376 | A | F | A | H |

Figure 3.3.2.2C
Raw Scores: Read 2-3 S401 Online n/a

Table 3.3.2.2C
Raw Score Descriptive Statistics: Read 2-3 S401 Online n/a


Figure 3.3.2.2E
Proficiency Level: Read 2-3 S401 Online


Table 3.3.2.2D
Scale Score Descriptive Statistics: Read 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 142,362 | 194 | 415 | 311.00 | 32.51 |
| $\mathbf{3}$ | 149,280 | 158 | 415 | 328.49 | 38.39 |
| Total | 291,642 | 158 | 415 | 319.95 | 36.70 |

Table 3.3.2.2E
Proficiency Level Distribution: Read 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 25,704 | $18.06 \%$ | 31,198 | $20.90 \%$ | 56,902 | $19.51 \%$ |
| $\mathbf{2}$ | 43,882 | $30.82 \%$ | 41,585 | $27.86 \%$ | 85,467 | $29.31 \%$ |
| $\mathbf{3}$ | 28,595 | $20.09 \%$ | 21,938 | $14.70 \%$ | 50,533 | $17.33 \%$ |
| $\mathbf{4}$ | 11,938 | $8.39 \%$ | 10,179 | $6.82 \%$ | 22,117 | $7.58 \%$ |
| $\mathbf{5}$ | 16,671 | $11.71 \%$ | 18,250 | $12.23 \%$ | 34,921 | $11.97 \%$ |
| $\mathbf{6}$ | 15,572 | $10.94 \%$ | 26,130 | $17.50 \%$ | 41,702 | $14.30 \%$ |
| Total | 142,362 | $100.00 \%$ | 149,280 | $100.00 \%$ | 291,642 | $100.00 \%$ |

Table 3.3.2.2F
Raw Score to Scale Score Conversion: Read 2-3 S401 Online n/a

Table 3.3.2.2G
Equating Summary: Read 2-3 S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Items |  | Average Difficulty (Std. Dev.) | No. of Items |  | Average Difficulty (Std. Dev.) |
|  | 72 |  | -0.55 (1.06) | 72 |  | -0.50 (0.94) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | -3.39 |  | 1.53 | -3.63 |  | 1.51 |
| Anchoring Items | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 63 |  | -0.64 (1.09) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 20 |  | -0.50 (0.63) |  |  |  |
|  | Percentage Anchors |  | Average Displacement |  |  |  |
|  | 28\% |  | -0.02 |  |  |  |
| Displacement of Anchor Items | Anchor Items by Displacement |  |  | Anchor Items by Item Difficulty |  |  |
|  | Item ID | Item Difficulty | Displacement | Item ID | Item Difficulty | Displacement |
|  | 13323 | 0.06 | -0.29 | 13392 | -1.78 | -0.02 |
|  | 13331 | -1.45 | -0.28 | 13419 | -1.67 | -0.10 |
|  | 13277 | -1.42 | -0.21 | 13331 | -1.45 | -0.28 |
|  | 13415 | -0.14 | -0.20 | 13277 | -1.42 | -0.21 |
|  | 13322 | -0.01 | -0.16 | 13420 | -0.90 | 0.20 |
|  | 13366 | -0.10 | -0.15 | 13414 | -0.76 | 0.24 |
|  | 13412 | -0.11 | -0.12 | 13281 | -0.62 | 0.19 |
|  | 13419 | -1.67 | -0.10 | 13413 | -0.47 | -0.09 |
|  | 13413 | -0.47 | -0.09 | 13410 | -0.34 | 0.06 |
|  | 13351 | -0.24 | -0.06 | 13350 | -0.32 | -0.01 |
|  | 13392 | -1.78 | -0.02 | 13351 | -0.24 | -0.06 |
|  | 13350 | -0.32 | -0.01 | 13415 | -0.14 | -0.20 |
|  | 13410 | -0.34 | 0.06 | 13357 | -0.13 | 0.20 |
|  | 13339 | 0.38 | 0.13 | 13412 | -0.11 | -0.12 |
|  | 13356 | 0.07 | 0.18 | 13366 | -0.10 | -0.15 |
|  | 13281 | -0.62 | 0.19 | 13322 | -0.01 | -0.16 |
|  | 13365 | $0.00$ | 0.19 | 13365 | 0.00 | 0.19 |
|  | 13420 | -0.90 | 0.20 | 13323 | 0.06 | -0.29 |
|  | 13357 | -0.13 | 0.20 | 13356 | 0.07 | 0.18 |
|  | 13414 | -0.76 | 0.24 | 13339 | 0.38 | 0.13 |

[^13]Figure 3.3.2.2H
Test Characteristic Curve: Read 2-3 S401 Online n/a

Figure 3.3.2.2I


Table 3.3.2.2J
Reliability: Read 2-3 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 291,642 | 72 | .90 |

Table 3.3.2.2K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 2-3 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 2 | 283 | 906 | 10.20 | 12.24 | 11.00 | 0.46 |
|  | 3 | 297 | 1,267 | 10.20 | 11.22 | 10.68 | 0.51 |
| 2/3 | 2 | 307 | 3,171 | 9.69 | 10.20 | 9.73 | 0.13 |
|  | 3 | 323 | 464 | 9.69 | 11.22 | 9.96 | 0.29 |
| 3/4 | 2 | 326 | 578 | 9.69 | 10.20 | 10.01 | 0.25 |
|  | 3 | 342 | 182 | 10.20 | 11.22 | 10.85 | 0.25 |
| 4/5 | 2 | 337 | 3,730 | 10.20 | 10.71 | 10.20 | 0.01 |
|  | 3 | 352 | 861 | 10.71 | 12.24 | 11.30 | 0.21 |
| 5/6 | 2 | 355 | 2,937 | 11.22 | 11.73 | 11.23 | 0.03 |
|  | 3 | 370 | 593 | 13.27 | 13.27 | 13.27 | 0.00 |

Table 3.3.2.2Li
Accuracy and Consistency of Classification Indices: Read (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.651 | 0.544 |  | 0.433 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.754 |  | 0.645 |  |
|  | 2 | 0.698 |  | 0.599 |  |
|  | 3 | 0.564 |  | 0.446 |  |
|  | 4 | 0.353 |  | 0.256 |  |
|  | 5 | 0.583 |  | 0.455 |  |
|  | 6 | 0.851 |  | 0.747 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.914 | 0.046 | 0.041 | 0.877 |
|  | 2/3 | 0.895 | 0.051 | 0.054 | 0.854 |
|  | 3/4 | 0.924 | 0.040 | 0.036 | 0.890 |
|  | 4/5 | 0.938 | 0.038 | 0.025 | 0.911 |
|  | 5/6 | 0.960 | 0.025 | 0.015 | 0.944 |

Table 3.3.2.2Lii
Accuracy and Consistency of Classification Indices: Read (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.648 | 0.550 |  | 0.446 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.756 |  | 0.665 |  |
|  | 2 | 0.675 |  | 0.568 |  |
|  | 3 | 0.470 |  | 0.356 |  |
|  | 4 | 0.293 |  | 0.208 |  |
|  | 5 | 0.542 |  | 0.420 |  |
|  | 6 | 0.877 |  | 0.796 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.904 | 0.054 | 0.042 | 0.867 |
|  | 2/3 | 0.904 | 0.044 | 0.052 | 0.865 |
|  | 3/4 | 0.927 | 0.040 | 0.033 | 0.894 |
|  | 4/5 | 0.933 | 0.040 | 0.027 | 0.905 |
|  | 5/6 | 0.948 | 0.032 | 0.020 | 0.927 |

### 3.3.2.3 Writing 2-3

3.3.2.3i Writing 2-3 A

Table 3.3.2.3Ai
Complete Task Analysis and Summary: Writ 2-3 A S401 Online
$\left.\begin{array}{|l|c|c|c|c|c|}\hline \text { Task Type } & \begin{array}{c}\text { Average } \\ \text { Task } \\ \text { Difficulty } \\ \text { (in logits) }\end{array} & \begin{array}{c}\text { Average } \\ \text { No. of } \\ \text { Tasks }\end{array} & \begin{array}{c}\text { Average } \\ \text { Infit } \\ \text { Mean } \\ \text { Square }\end{array} & \begin{array}{c}\text { Outfit } \\ \text { Mean } \\ \text { Square }\end{array} \\ \hline \text { Constructed Response } & 0.11 & 3 & 0.41 & 0.40 \\ \hline & & & & \text { Task } \\ \text { Difficulty } \\ \text { (in logits) }\end{array}\right)$

Table 3.3.2.3Bi
DIF Analysis and Summary: Writ 2-3 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 2 | 1 | 1 | 2 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |
| 1.W23A_SI_WritingTime_P 100_14257 | AA | M | AA | H |
| 2.W23A_LA_Umbrella_203_P 100_A302_14258 | AA | F | AA | O |
| 3.W23A_MS_GiraffeCheetah_P 100_14259 | AA | M | AA | O |

Figure 3.3.2.3Ci
Raw Scores: Writ 2-3 A S401 Online


Table 3.3.2.3Ci
Raw Score Descriptive Statistics: Writ 2-3 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 61,602 | 0 | 19 | 10.33 | 3.62 |
| $\mathbf{3}$ | 42,063 | 0 | 20 | 10.92 | 3.64 |
| Total | 103,665 | 0 | 20 | 10.57 | 3.64 |

Table 3.3.2.3Di
Scale Score Descriptive Statistics: Writ 2-3 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 61,602 | 122 | 358 | 267.14 | 32.90 |
| $\mathbf{3}$ | 42,063 | 133 | 367 | 273.01 | 32.36 |
| Total | 103,665 | 122 | 367 | 269.52 | 32.81 |



Table 3.3.2.3Ei
Proficiency Level Distribution: Writ 2-3 A S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
|  | 8,943 | $14.52 \%$ | 6,192 | $14.72 \%$ | 15,135 | $14.60 \%$ |
| $\mathbf{2}$ | 23,378 | $37.95 \%$ | 19,478 | $46.31 \%$ | 42,856 | $41.34 \%$ |
| $\mathbf{3}$ | 29,254 | $47.49 \%$ | 16,327 | $38.82 \%$ | 45,581 | $43.97 \%$ |
| $\mathbf{4}$ | 27 | $0.04 \%$ | 66 | $0.16 \%$ | 93 | $0.09 \%$ |
| $\mathbf{5}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 61,602 | $100.00 \%$ | 42,063 | $100.00 \%$ | 103,665 | $100.00 \%$ |

Table 3.3.2.3Fi
Raw Score to Scale Score Conversion: Writ 2-3 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 133^ | 108.75 | $100.00^{\wedge}$ | 281.75 |
| 1 | 199 | 23.41 | 175.59 | 222.41 |
| 2 | 213 | 16.41 | 196.59 | 229.41 |
| 3 | 222 | 13.77 | 208.23 | 235.77 |
| 4 | 228 | 12.51 | 215.49 | 240.51 |
| 5 | 234 | 11.90 | 222.10 | 245.90 |
| 6 | 239 | 11.71 | 227.29 | 250.71 |
| 7 | 244 | 11.84 | 232.16 | 255.84 |
| 8 | 249 | 12.27 | 236.73 | 261.27 |
| 9 | 255 | 13.02 | 241.98 | 268.02 |
| 10 | 262 | 14.02 | 247.98 | 276.02 |
| 11 | 270 | 15.20 | 254.80 | 285.20 |
| 12 | 279 | 16.27 | 262.73 | 295.27 |
| 13 | 290 | 17.10 | 272.90 | 307.10 |
| 14 | 301 | 17.64 | 283.36 | 318.64 |
| 15 | 313 | 17.88 | 295.12 | 330.88 |
| 16 | 325 | 17.83 | 307.17 | 342.83 |
| 17 | 336 | 17.51 | 318.49 | 353.51 |
| 18 | 347 | 16.97 | 330.03 | 363.97 |
| 19 | 358 | 16.33 | 341.67 | 374.33 |
| 20 | 367 | 15.73 | 351.27 | 382.73 |
| 21 | 376 | 15.33 | 360.67 | 391.33 |
| 22 | 385 | 15.25 | 369.75 | 400.25 |
| 23 | 394 | 15.65 | 378.35 | 409.65 |
| 24 | 403 | 16.81 | 386.19 | 419.81 |
| 25 | 415 | 19.41 | 395.59 | 434.41 |
| 26 | 434 | 26.50 | 407.50 | 460.50 |
| 27 | 465 | 48.41 | 416.59 | 513.41 |

$\wedge$ Truncated

Table 3.3.2.3Gi
Equating Summary: Writ 2-3 A S401 Online


[^14]

Table 3.3.2.3Ji
Reliability: Writ 2-3 A S401 Online

| Reliability | No. of Students | No. of Tasks | Response <br> Mode | Cronbach's <br> Alpha | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 103,665 | 3 | Hand-written <br> $(H W)$ | .863 | 1.347 |
| Interrater <br> Reliability | Task | No. in Sample | $\%$ AG | $\%$ AD | $\%$ NA |
|  | 1 | 52,575 | 94 | 5 | 1 |
|  | 2 | 52,446 | 96 | 3 | 0 |

Table 3.3.2.3Ki
Conditional Standard Error of Measurement at Cut Scores: Writ 2-3 A S401 Online n/a

Table 3.3.2.3Li
Accuracy and Consistency of Classification Indices: Writ 2-3 A S401 Online n/a

Table 3.3.2.3Aii
Complete Task Analysis and Summary: Writ 2-3 B/C S401 Online
$\left.\begin{array}{|l|c|c|c|c|c|}\hline \text { Task Type } & \begin{array}{c}\text { Average } \\ \text { Task } \\ \text { Difficulty } \\ \text { (in logits) }\end{array} & \begin{array}{c}\text { Average } \\ \text { No. of } \\ \text { Tasks }\end{array} & \begin{array}{c}\text { Average } \\ \text { Infit } \\ \text { Mean } \\ \text { Square }\end{array} & \begin{array}{c}\text { Outfit } \\ \text { Mean } \\ \text { Square }\end{array} \\ \hline \text { Constructed Response } & 0.47 & 3 & 0.54 & 0.53 \\ \hline & & & & \text { Task } \\ \text { Tifficulty } \\ \text { (in logits) }\end{array}\right)$

Table 3.3.2.3Bii
DIF Analys is and Summary: Writ 2-3 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 2 | 1 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF | Favored |  |  |  |
| Group | DIF | Level | Favored <br> Group |  |  |
| 1.W23B_SI_Gro wnUpJobs_203_P 100_A302_14260 | AA | F | AA | O |  |
| 2.W23C_MS_Plants_MaBaPaLe_P 100_A203_14261 | AA | F | AA | H |  |
| 3.W23C_IT_LibrarySale_401_14480 | AA | M | AA | H |  |

Figure 3.3.2.3Cii
Raw Scores: Writ 2-3 B/C S401 Online


Table 3.3.2.3Cii
Raw Score Descriptive Statistics: Writ 2-3 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 102,601 | 0 | 52 | 29.36 | 5.10 |
| $\mathbf{3}$ | 129,812 | 0 | 53 | 32.25 | 4.68 |
| Total | 232,413 | 0 | 53 | 30.97 | 5.08 |

Table 3.3.2.3Dii
Scale Score Descriptive Statistics: Writ 2-3 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 102,601 | 122 | 441 | 314.83 | 26.31 |
| $\mathbf{3}$ | 129,812 | 133 | 459 | 330.37 | 25.12 |
| Total | 232,413 | 122 | 459 | 323.51 | 26.78 |



Table 3.3.2.3Eii
Proficiency Level Distribution: Writ 2-3 B/C S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 803 | $0.78 \%$ | 379 | $0.29 \%$ | 1,182 | $0.51 \%$ |
| $\mathbf{2}$ | 7,130 | $6.95 \%$ | 3,560 | $2.74 \%$ | 10,690 | $4.60 \%$ |
| $\mathbf{3}$ | 76,364 | $74.43 \%$ | 85,198 | $65.63 \%$ | 161,562 | $69.52 \%$ |
| $\mathbf{4}$ | 18,198 | $17.74 \%$ | 40,170 | $30.94 \%$ | 58,368 | $25.11 \%$ |
| $\mathbf{5}$ | 104 | $0.10 \%$ | 484 | $0.37 \%$ | 588 | $0.25 \%$ |
| $\mathbf{6}$ | 2 | $0.00 \%$ | 21 | $0.02 \%$ | 23 | $0.01 \%$ |
| Total | 102,601 | $100.00 \%$ | 129,812 | $100.00 \%$ | 232,413 | $100.00 \%$ |

Table 3.3.2.3Fii
Raw Score to Scale Score Conversion: Writ 2-3 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 133^ | 80.82 | $100.00^{\wedge}$ | 239.82 | 34 | 341 | 12.30 | 328.70 | 353.30 |
| 1 | 187 | 24.19 | 162.81 | 211.19 | 35 | 346 | 12.14 | 333.86 | 358.14 |
| 2 | 201 | 16.54 | 184.46 | 217.54 | 36 | 352 | 11.95 | 340.05 | 363.95 |
| 3 | 210 | 13.45 | 196.55 | 223.45 | 37 | 357 | 11.73 | 345.27 | 368.73 |
| 4 | 216 | 11.79 | 204.21 | 227.79 | 38 | 362 | 11.52 | 350.48 | 373.52 |
| 5 | 220 | 10.77 | 209.23 | 230.77 | 39 | 367 | 11.33 | 355.67 | 378.33 |
| 6 | 224 | 10.10 | 213.90 | 234.10 | 40 | 371 | 11.17 | 359.83 | 382.17 |
| 7 | 228 | 9.64 | 218.36 | 237.64 | 41 | 376 | 11.06 | 364.94 | 387.06 |
| 8 | 231 | 9.32 | 221.68 | 240.32 | 42 | 381 | 10.98 | 370.02 | 391.98 |
| 9 | 234 | 9.10 | 224.90 | 243.10 | 43 | 385 | 10.98 | 374.02 | 395.98 |
| 10 | 237 | 8.94 | 228.06 | 245.94 | 44 | 390 | 11.04 | 378.96 | 401.04 |
| 11 | 240 | 8.86 | 231.14 | 248.86 | 45 | 394 | 11.14 | 382.86 | 405.14 |
| 12 | 243 | 8.81 | 234.19 | 251.81 | 46 | 399 | 11.38 | 387.62 | 410.38 |
| 13 | 246 | 8.78 | 237.22 | 254.78 | 47 | 404 | 11.71 | 392.29 | 415.71 |
| 14 | 249 | 8.81 | 240.19 | 257.81 | 48 | 409 | 12.19 | 396.81 | 421.19 |
| 15 | 252 | 8.89 | 243.11 | 260.89 | 49 | 415 | 12.92 | 402.08 | 427.92 |
| 16 | 255 | 9.00 | 246.00 | 264.00 | 50 | 422 | 13.99 | 408.01 | 435.99 |
| 17 | 258 | 9.13 | 248.87 | 267.13 | 51 | 430 | 15.73 | 414.27 | 445.73 |
| 18 | 261 | 9.32 | 251.68 | 270.32 | 52 | 441 | 18.88 | 422.12 | 459.88 |
| 19 | 264 | 9.56 | 254.44 | 273.56 | 53 | 459 | 26.45 | 432.55 | 485.45 |
| 20 | 268 | 9.85 | 258.15 | 277.85 | 54 | 490 | 48.60 | 441.40 | 538.60 |
| 21 | 272 | 10.20 | 261.80 | 282.20 |  |  |  |  |  |
| 22 | 276 | 10.55 | 265.45 | 286.55 |  |  |  |  |  |
| 23 | 280 | 10.96 | 269.04 | 290.96 |  |  |  |  |  |
| 24 | 285 | 11.33 | 273.67 | 296.33 |  |  |  |  |  |
| 25 | 290 | 11.65 | 278.35 | 301.65 |  |  |  |  |  |
| 26 | 295 | 11.95 | 283.05 | 306.95 |  |  |  |  |  |
| 27 | 300 | 12.19 | 287.81 | 312.19 |  |  |  |  |  |
| 28 | 306 | 12.38 | 293.62 | 318.38 |  |  |  |  |  |
| 29 | 312 | 12.49 | 299.51 | 324.49 |  |  |  |  |  |
| 30 | 317 | 12.57 | 304.43 | 329.57 |  |  |  |  |  |
| 31 | 323 | 12.57 | 310.43 | 335.57 |  |  |  |  |  |
| 32 | 329 | 12.51 | 316.49 | 341.51 |  |  |  |  |  |
| 33 | 335 | 12.43 | 322.57 | 347.43 |  |  |  |  |  |

[^15]Table 3.3.2.3Gii
Equating Summary: Writ 2-3 B/C S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 0.47 (0.55) | 3 |  | 0.16 (0.26) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | -0.14 |  | 0.95 | -0.11 |  | 0.41 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | 0.77 (0.26) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 1 |  | 0.58 (N/A) |  |  |  |
|  | Percentage <br> Anchors |  | Average <br> Displacement |  |  |  |
|  | 33\% |  | 0.00 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Task ID ${ }^{\text {D }}$ | Task Difficulty | y y Displacement | Task ID | Task Difficulty | Displacement |
|  | 14261 | 0.58 | 0.00 | 14261 | 0.58 | 0.00 |

[^16]

Table 3.3.2.3Jii
Reliability: Writ 2-3 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Response Mode | Cronbach's <br> Alpha | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Hand-written <br> $(H W)$ | .896 | 1.636 |
| Interrater <br> Reliability | 232,413 | 3 | Task | No. in Sample | \% AG |
|  | 1 | 101,738 | 95 | \% AD | \% NA |
|  | 2 | 101,640 | 93 | 5 | 0 |
|  | 3 | 102,426 | 95 | 6 | 0 |

Table 3.3.2.3Kii
Conditional Standard Error of Measurement at Cut Scores: Writ 2-3 B/C S401 Online n/a

Table 3.3.2.3Lii
Accuracy and Consistency of Classification Indices: Writ 2-3 B/C S401 Online n/a

### 3.3.2.3iii Writing 2-3 Across Tiers

Table 3.3.2.3Aiii
Complete Task Analysis and Summary: Writ 2-3 S401 Online n/a

Table 3.3.2.3Biii
DIF Analysis and Summary: Writ 2-3 S401 Online n/a


Table 3.3.2.3Ciii
Raw Score Descriptive Statistics: Writ 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 164,203 | 0 | 52 | 22.22 | 10.30 |
| $\mathbf{3}$ | 171,875 | 0 | 53 | 27.03 | 10.19 |
| Total | 336,078 | 0 | 53 | 24.68 | 10.52 |



Table 3.3.2.3Diii
Scale Score Descriptive Statistics: Writ 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 164,203 | 122 | 441 | 296.94 | 37.03 |
| $\mathbf{3}$ | 171,875 | 133 | 459 | 316.33 | 36.62 |
| Total | 336,078 | 122 | 459 | 306.85 | 38.08 |



Table 3.3.2.3Eiii
Proficiency Level Distribution: Writ 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 9,746 | $5.94 \%$ | 6,571 | $3.82 \%$ | 16,317 | $4.86 \%$ |
| $\mathbf{2}$ | 30,508 | $18.58 \%$ | 23,038 | $13.40 \%$ | 53,546 | $15.93 \%$ |
| $\mathbf{3}$ | 105,618 | $64.32 \%$ | 101,525 | $59.07 \%$ | 207,143 | $61.64 \%$ |
| $\mathbf{4}$ | 18,225 | $11.10 \%$ | 40,236 | $23.41 \%$ | 58,461 | $17.40 \%$ |
| $\mathbf{5}$ | 104 | $0.06 \%$ | 484 | $0.28 \%$ | 588 | $0.17 \%$ |
| $\mathbf{6}$ | 2 | $0.00 \%$ | 21 | $0.01 \%$ | 23 | $0.01 \%$ |
| Total | 164,203 | $100.00 \%$ | 171,875 | $100.00 \%$ | 336,078 | $100.00 \%$ |

Table 3.3.2.3Fiii
Raw Score to Scale Score Conversion: Writ 2-3 S401 Online n/a

Table 3.3.2.3Giii
Equating Summary: Writ 2-3 S401 Online n/a



Table 3.3.2.3Jiii
Reliability: Writ 2-3 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| A | 103,665 | 0.863 | 0.886 |
| B/C | 232,413 | 0.896 |  |

Table 3.3.2.3Kiii
Conditional Standard Error of Measurement at Cut Scores: Writ 2-3 S401 Online

| Proficiency <br> Level |  | SEM |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Tier A | Tier B/C |
| $1 / 2$ | 2 | 242 | 11.81 | 8.86 |
|  | 3 | 247 | 12.08 | 8.86 |
| $2 / 3$ | 2 | 279 | 16.27 | 11.01 |
|  | 3 | 283 | 16.65 | 11.28 |
| $3 / 4$ | 2 | 341 | 17.18 | 12.30 |
|  | 3 | 346 | 17.18 | 12.14 |
| $4 / 5$ | 2 | 388 | 15.31 | 11.01 |
|  | 3 | 394 | 15.65 | 11.14 |
| $5 / 6$ | 2 | 411 | 18.26 | 12.35 |
|  | 3 | 418 | 20.41 | 13.43 |

Table 3.3.2.3Li
Accuracy and Consistency of Classification Indices: Writ (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.795 | 0.732 |  | 0.476 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.794 |  | 0.673 |  |
|  | 2 | 0.776 |  | 0.661 |  |
|  | 3 | 0.799 |  | 0.795 |  |
|  | 4 | - |  | 0.324 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
|  | 1/2 | 0.976 | 0.013 | 0.011 | 0.965 |
|  | 2/3 | 0.930 | 0.027 | 0.043 | 0.902 |
|  | 3/4 | 0.888 | 0.112 | 0.000 | 0.865 |
|  | 4/5 | 0.999 | 0.001 | 0.000 | 0.999 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.2.3 Lii
Accuracy and Consistency of Classification Indices: Writ (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.775 | 0.697 |  | 0.468 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.781 |  | 0.650 |  |
|  | 2 | 0.772 |  | 0.653 |  |
|  | 3 | 0.811 |  | 0.758 |  |
|  | 4 | 0.678 |  | 0.566 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.983 | 0.008 | 0.008 | 0.975 |
|  | 2/3 | 0.948 | 0.019 | 0.033 | 0.927 |
|  | 3/4 | 0.847 | 0.085 | 0.068 | 0.796 |
|  | 4/5 | 0.997 | 0.003 | 0.000 | 0.997 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.2.4 Speaking 2-3

3.3.2.4i Speaking 2-3 Pre-A

Table 3.3.2.4Ai
Complete Task Analysis and Summary: Spek 2-3 Pre-A S401 Online n/a

Table 3.3.2.4Bi
DIF Analysis and Summary: Spek 2-3 Pre-A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 2 | 1 | 2 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S23P_SI_P laygro und_401_14579 | AA | M | AA | O |
| 2.S23P_LA_Clas s Garden_P 100_New_14680 | AA | F | AA | H |
| 3.S23P_MS_WaterCycle_401_4575 | AA | F | AA | O |

Figure 3.3.2.4Ci
Raw Scores: Spek 2-3 Pre-A S401 Online


Table 3.3.2.4Ci
Raw Score Descriptive Statistics: Spek 2-3 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 5,414 | 0 | 6 | 5.27 | 1.36 |
| $\mathbf{3}$ | 8,618 | 0 | 6 | 5.39 | 1.26 |
| Total | 14,032 | 0 | 6 | 5.35 | 1.30 |



Table 3.3.2.4Ei
Proficiency Level Distribution: Spek 2-3 Pre-A S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
|  | 5,414 | $100.00 \%$ | 8,618 | $100.00 \%$ | 14,032 | $100.00 \%$ |
| Total | 5,414 | $100.00 \%$ | 8,618 | $100.00 \%$ | 14,032 | $100.00 \%$ |

Table 3.3.2.4Fi
Raw Score to Scale Score Conversion: Spek 2-3 Pre-A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $118^{\wedge}$ | 26.62 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 1 | $118^{\wedge}$ | 26.62 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 2 | $118^{\wedge}$ | 26.62 | $100.00^{\wedge}$ | 100.62 |
| 3 | $118^{\wedge}$ | 26.62 | $100.00^{\wedge}$ | 113.62 |
| 4 | $118^{\wedge}$ | 26.62 | $100.00^{\wedge}$ | 126.62 |
| 5 | $118^{*}$ | 26.62 | $100.00^{\wedge}$ | 144.62 |
| 6 | $126^{*}$ | 31.00 | 116.00 | 178.00 |

[^17]Table 3.3.2.4Gi
Equating Summary: Spek 2-3 Pre-A S401 Online n/a

Figure 3.3.2.4Hi


Figure 3.3.2.4ii


Table 3.3.2.4Ji
Reliability: Spek 2-3 Pre-A S401 Online

| Reliability |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
|  | 14,032 | 3 | .786 |  | 0.600 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 9,290 | 98 | 2 | 0 |
|  | 2 | 9,956 | 98 | 2 | 0 |
|  | 3 | 9,410 | 98 | 2 | 0 |

Table 3.3.2.4Ki
Conditional Standard Error of Measurement at Cut Scores: Spek 2-3 Pre-A S401 Online n/a

Table 3.3.2.4Li
Accuracy and Consistency of Classification Indices: Spek 2-3 Pre-A S401 Online n/a

### 3.3.2.4ii $\quad$ Speaking 2-3 A

Table 3.3.2.4Aii
Complete Task Analysis and Summary: Spek 2-3 A S401 Online
n/a

Table 3.3.2.4Bii
DIF Analysis and Summary: Spek 2-3 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 5 | 2 | 4 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S23A_SI_P layground_401_14579 | AA | F | AA | O |
| 2.S23A_SI_P laygro und_401_14580 | AA | M | AA | H |
| 3.S23A_LA_Clas s Garden_P 100_New_ALT_14201 | AA | F | AA | O |
| 4.S23A_LA_Clas sGarden_P 100_New_ALT_14202 | AA | F | AA | H |
| 5.S23A_MS_WaterCycle_401_V1_14575 | AA | F | AA | O |
| 6.S23A_MS_WaterCycle_401_V1_14576 | AA | F | AA | O |

Figure 3.3.2.4Cii
Raw Scores: Spek 2-3 A S401 Online


Table 3.3.2.4Cii
Raw Score Descriptive Statistics: Spek 2-3 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 34,384 | 0 | 18 | 11.96 | 1.90 |
| $\mathbf{3}$ | 32,253 | 0 | 18 | 12.49 | 1.67 |
| Total | 66,637 | 0 | 18 | 12.21 | 1.81 |



Table 3.3.2.4Dii
Scale Score Descriptive Statistics: Spek 2-3 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 34,384 | 112 | 383 | 249.37 | 40.12 |
| $\mathbf{3}$ | 32,253 | 118 | 383 | 260.92 | 37.04 |
| Total | 66,637 | 112 | 383 | 254.96 | 39.09 |

Table 3.3.2.4Eii
Proficiency Level Distribution: Spek 2-3 A S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
|  | 5,663 | $16.47 \%$ | 7,542 | $23.38 \%$ | 13,205 | $19.82 \%$ |
| $\mathbf{2}$ | 15,345 | $44.63 \%$ | 16,926 | $52.48 \%$ | 32,271 | $48.43 \%$ |
| $\mathbf{3}$ | 12,729 | $37.02 \%$ | 6,822 | $21.15 \%$ | 19,551 | $29.34 \%$ |
| $\mathbf{4}$ | 618 | $1.80 \%$ | 963 | $2.99 \%$ | 1,581 | $2.37 \%$ |
| $\mathbf{5}$ | 29 | $0.08 \%$ | 0 | $0.00 \%$ | 29 | $0.04 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 34,384 | $100.00 \%$ | 32,253 | $100.00 \%$ | 66,637 | $100.00 \%$ |

Table 3.3.2.4Fii
Raw Score to Scale Score Conversion: Spek 2-3 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 1 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 2 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | $100.00^{\wedge}$ |
| 3 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | 107.81 |
| 4 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | 119.81 |
| 5 | $118^{\wedge}$ | 22.81 | $100.00^{\wedge}$ | 132.81 |
| 6 | 128 | 24.57 | 103.43 | 152.57 |
| 7 | 150 | 24.86 | 125.14 | 174.86 |
| 8 | 170 | 23.11 | 146.89 | 193.11 |
| 9 | 187 | 22.23 | 164.77 | 209.23 |
| 10 | 204 | 23.11 | 180.89 | 227.11 |
| 11 | 224 | 25.45 | 198.55 | 249.45 |
| 12 | 248 | 27.79 | 220.21 | 275.79 |
| 13 | 274 | 26.62 | 247.38 | 300.62 |
| 14 | 297 | 24.86 | 272.14 | 321.86 |
| 15 | 317 | 24.57 | 292.43 | 341.57 |
| 16 | 339 | 26.03 | 312.97 | 365.03 |
| 17 | $361^{*}$ | 31.00 | 336.00 | 398.00 |
| 18 | $383^{*}$ | 39.48 | 368.52 | 447.48 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.2.4Gii
Equating Summary: Spek 2-3 A S401 Online n/a



Table 3.3.2.4Jii
Reliability: Spek 2-3 A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 66,637 | 6 | .585 |  | 1.169 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 38,804 | 99 | 1 | 0 |
|  | 2 | 38,805 | 83 | 16 | 0 |
|  | 3 | 39,156 | 99 | 1 | 0 |
|  | 4 | 39,152 | 88 | 11 | 0 |
|  | 5 | 38,506 | 99 | 1 | 0 |

Table 3.3.2.4Kii
Conditional Standard Error of Measurement at Cut Scores: Spek 2-3 A S401 Online n/a

Table 3.3.2.4Lii
Accuracy and Consistency of Classification Indices: Spek 2-3 A S401 Online n/a

### 3.3.2.4iii Speaking 2-3 B/C

Table 3.3.2.4Aiii
Complete Task Analysis and Summary: Spek 2-3 B/C S401 Online n/a

Table 3.3.2.4Biii
DIF Analys is and Summary: Spek 2-3 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 2 | 4 | 4 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.S23C_SI_Playground_401_V2_14569 | AA | M | AA | H |  |
| 2.S23C_SI_Playground_401_V2_14570 | AA | F | AA | O |  |
| 3.S23C_LS_ClassGarden_401_14581 | AA | F | AA | H |  |
| 4.S23C_LS_ClassGarden_401_14582 | AA | F | AA | H |  |
| 5.S23C_MS_WaterCycle_401_V1_14571 | AA | F | AA | H |  |
| 6.S23C_MS_WaterCycle_401_V1_14572 | AA | M | AA | O |  |

Figure 3.3.2.4Ciii
Raw Scores: Spek 2-3 B/C S401 Online


Table 3.3.2.4Ciii
Raw Score Descriptive Statistics: Spek 2-3 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 93,004 | 6 | 30 | 19.26 | 2.73 |
| $\mathbf{3}$ | 102,739 | 6 | 30 | 20.50 | 2.73 |
| Total | 195,743 | 6 | 30 | 19.91 | 2.80 |

Figure 3.3.2.4Diii
Scale Scores: Spek 2-3 B/C S401 Online

Table 3.3.2.4Diii
Scale Score Descriptive Statistics: Spek 2-3 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 93,004 | 112 | 415 | 277.46 | 31.63 |
| $\mathbf{3}$ | 102,739 | 118 | 425 | 291.88 | 31.49 |
| Total | 195,743 | 112 | 425 | 285.03 | 32.37 |

Figure 3.3.2.4Eiii
Proficiency Level: Spek 2-3 B/C S401 Online


Table 3.3.2.4Eiii
Proficiency Level Distribution: Spek 2-3 B/C S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 3,494 | $3.76 \%$ | 3,367 | $3.28 \%$ | 6,861 | $3.51 \%$ |
| $\mathbf{2}$ | 33,072 | $35.56 \%$ | 32,941 | $32.06 \%$ | 66,013 | $33.72 \%$ |
| $\mathbf{3}$ | 50,743 | $54.56 \%$ | 59,221 | $57.64 \%$ | 109,964 | $56.18 \%$ |
| $\mathbf{4}$ | 5,525 | $5.94 \%$ | 7,006 | $6.82 \%$ | 12,531 | $6.40 \%$ |
| $\mathbf{5}$ | 164 | $0.18 \%$ | 168 | $0.16 \%$ | 332 | $0.17 \%$ |
| $\mathbf{6}$ | 6 | $0.01 \%$ | 36 | $0.04 \%$ | 42 | $0.02 \%$ |
| Total | 93,004 | $100.00 \%$ | 102,739 | $100.00 \%$ | 195,743 | $100.00 \%$ |

Table 3.3.2.4Fiii
Raw Score to Scale Score Conversion: Spek 2-3 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 6 | $118^{\wedge}$ | 21.35 | 102.65 | 145.35 |
| 7 | 141 | 22.23 | 118.77 | 163.23 |
| 8 | 156 | 20.18 | 135.82 | 176.18 |
| 9 | 169 | 18.43 | 150.57 | 187.43 |
| 10 | 180 | 17.26 | 162.74 | 197.26 |
| 11 | 190 | 16.67 | 173.33 | 206.67 |
| 12 | 199 | 16.38 | 182.62 | 215.38 |
| 13 | 209 | 16.38 | 192.62 | 225.38 |
| 14 | 218 | 16.67 | 201.33 | 234.67 |
| 15 | 228 | 17.26 | 210.74 | 245.26 |
| 16 | 238 | 17.84 | 220.16 | 255.84 |
| 17 | 250 | 18.72 | 231.28 | 268.72 |
| 18 | 262 | 19.30 | 242.70 | 281.30 |
| 19 | 275 | 19.30 | 255.70 | 294.30 |
| 20 | 287 | 18.72 | 268.28 | 305.72 |
| 21 | 299 | 18.13 | 280.87 | 317.13 |
| 22 | 310 | 17.84 | 292.16 | 327.84 |
| 23 | 320 | 17.55 | 302.45 | 337.55 |
| 24 | 331 | 17.55 | 313.45 | 348.55 |
| 25 | 342 | 18.13 | 323.87 | 360.13 |
| 26 | 353 | 19.01 | 333.99 | 372.01 |
| 27 | 367 | 20.47 | 346.53 | 387.47 |
| 28 | $381^{*}$ | 23.11 | 359.89 | 406.11 |
| 29 | $395^{*}$ | 26.91 | 381.09 | 434.91 |
| 30 | $425^{*}$ | 40.07 | 405.93 | 486.07 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.2.4Giii
Equating Summary: Spek 2-3 B/C S401 Online n/a


Table 3.3.2.4Jiii
Reliability: Spek 2-3 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 195,743 | 6 | .688 |  | 1.561 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 102,167 | 74 | 25 | 1 |
|  | 2 | 102,167 | 71 | 28 | 1 |
|  | 3 | 105,128 | 78 | 21 | 1 |
|  | 4 | 105,129 | 74 | 26 | 1 |
|  | 5 | 105,894 | 78 | 22 | 1 |

Table 3.3.2.4Kiii
Conditional Standard Error of Measurement at Cut Scores: Spek 2-3 B/C S401 Online n/a

Table 3.3.2.4Liii
Accuracy and Consistency of Classification Indices: Spek 2-3 B/C S401 Online n/a

### 3.3.2.4iv Speaking 2-3 Across Tiers

Table 3.3.2.4Aiv
Complete Task Analys is and Summary: Spek 2-3 S401 Online

| Task Type |  | Average Task Difficulty (in logits) |  | No. of Tasks | Average Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | -1.78 |  | 15 | 0.73 | 0.62 |
| Name |  | Tier | Task Difficulty (in logits) |  | Fit S | stics |
|  |  | Anchored? |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 1.S23A_SI_Playground_401_14579 |  |  | A** | -6.11 |  | 0.63 | 0.21 |
| 2.S23A_SI_Playground_401_14580 |  | A* | -0.93 | Yes | 0.74 | 0.77 |
| 3.S23A_LA_ClassGarden_P 100_New_ALT_14201 |  | A** | -5.98 |  | 0.63 | 0.17 |
| 4.S23A_LA_ClassGarden_P 100_New_ALT_14202 |  | A | 0.20 | Yes | 0.61 | 0.63 |
| 5.S23A_MS_WaterCycle_401_V1_14575 |  | A** | -6.10 |  | 0.61 | 0.19 |
| 6.S23A_MS_WaterCycle_401_V1_14576 |  | A* | -0.58 | Yes | 0.72 | 0.73 |
| 7.S23C_SI_Playground_401_V2_14569 |  | B/C* | -0.93 | Yes | 0.74 | 0.77 |
| 8.S23C_SI_Playground_401_V2_14570 |  | B/C | 0.48 | Yes | 0.71 | 0.72 |
| 9.S23C_LS_ClassGarden_401_14581 |  | B/C | 0.38 | Yes | 0.91 | 0.93 |
| 10.S23C_LS_ClassGarden_401_14582 |  | B/C | -0.26 | Yes | 0.75 | 0.77 |
| 11.S23C_MS_WaterCycle_401_V1_14571 |  | B/C* | -0.58 | Yes | 0.72 | 0.73 |
| 12.S23C_MS_WaterCycle_401_V1_14572 |  | B/C | 1.07 | Yes | 0.99 | 1.04 |
| 13.S23P_SI_Playground_401_14579 |  | Pre-A** | -6.11 |  | 0.63 | 0.21 |
| 14.S23P_LA_ClassGarden_P 100_New_14680 |  | Pre-A** | -5.98 |  | 0.63 | 0.17 |
| 15.S23P_MS_WaterCycle_401_14575 |  | Pre-A** | -6.10 |  | 0.61 | 0.19 |
| Raw Score Distribution by Task | Task | Raw Score |  |  |  |  |
|  |  | 0 | 1 | 2 | 3 | 4 |
|  | Task 1 | 1.21\% | 4.01\% | 94.78\% | N/A | N/A |
|  | Task 2 | 0.16\% | 2.64\% | 35.87\% | 53.70\% | 7.64\% |
|  | Task 3 | 1.55\% | 3.81\% | 94.63\% | N/A | N/A |
|  | Task 4 | 1.09\% | 26.52\% | 51.58\% | 18.70\% | 2.11\% |
|  | Task 5 | 1.29\% | 4.58\% | 94.13\% | N/A | N/A |
|  | Task 6 | 0.47\% | 10.26\% | 48.17\% | 35.32\% | 5.77\% |
|  | Task 7 | 0.16\% | 2.64\% | 35.87\% | 53.70\% | 7.64\% |
|  | Task 8 | 0.06\% | 15.35\% | 58.97\% | 23.02\% | 2.61\% |
|  | Task 9 | 0.29\% | 21.50\% | 51.39\% | 21.18\% | 5.65\% |
|  | Task 10 | 0.48\% | 8.59\% | 39.62\% | 45.37\% | 5.94\% |
|  | Task 11 | 0.47\% | 10.26\% | 48.17\% | 35.32\% | 5.77\% |
|  | Task 12 | 0.67\% | 30.91\% | 43.65\% | 20.63\% | 4.14\% |
|  | Task 13 | 1.21\% | 4.01\% | 94.78\% | N/A | N/A |
|  | Task 14 | 1.55\% | 3.81\% | 94.63\% | N/A | N/A |
|  | Task 15 | 1.29\% | 4.58\% | 94.13\% | N/A | N/A |

** This task is shared between Pre-A and A.

* This task is shared between A and B/C.

Table 3.3.2.4Biv
DIF Analysis and Summary: Spek 2-3 S401 Online n/a


Figure 3.3.2.4Civ Raw Scores: Spek 2-3 S401 Online

Table 3.3.2.4Civ
Raw Score Descriptive Statistics: Spek 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 132,802 | 0 | 30 | 16.80 | 4.68 |
| $\mathbf{3}$ | 143,610 | 0 | 30 | 17.79 | 5.18 |
| Total | 276,412 | 0 | 30 | 17.31 | 4.97 |



Table 3.3.2.4Div
Scale Score Descriptive Statistics: Spek 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 132,802 | 112 | 415 | 263.85 | 46.09 |
| $\mathbf{3}$ | 143,610 | 118 | 425 | 274.85 | 51.34 |
| Total | 276,412 | 112 | 425 | 269.56 | 49.19 |



Table 3.3.2.4Eiv
Proficiency Level Distribution: Spek 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 14,571 | $10.97 \%$ | 19,527 | $13.60 \%$ | 34,098 | $12.34 \%$ |
| $\mathbf{2}$ | 48,417 | $36.46 \%$ | 49,867 | $34.72 \%$ | 98,284 | $35.56 \%$ |
| $\mathbf{3}$ | 63,472 | $47.79 \%$ | 66,043 | $45.99 \%$ | 129,515 | $46.86 \%$ |
| $\mathbf{4}$ | 6,143 | $4.63 \%$ | 7,969 | $5.55 \%$ | 14,112 | $5.11 \%$ |
| $\mathbf{5}$ | 193 | $0.15 \%$ | 168 | $0.12 \%$ | 361 | $0.13 \%$ |
| $\mathbf{6}$ | 6 | $0.00 \%$ | 36 | $0.03 \%$ | 42 | $0.02 \%$ |
| Total | 132,802 | $100.00 \%$ | 143,610 | $100.00 \%$ | 276,412 | $100.00 \%$ |

Table 3.3.2.4Fiv
Raw Score to Scale Score Conversion: Spek 2-3 S401 Online n/a

Table 3.3.2.4Giv
Equating Summary: Spek 2-3 S401 Online


[^18]


Table 3.3.2.4Jiv
Reliability: Spek 2-3 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| Pre-A | 14,032 | 0.786 | 0.668 |
| A | 66,637 | 0.585 |  |
| B/C | 195,743 | 0.688 |  |

Table 3.3.2.4Kiv
Conditional Standard Error of Measurement at Cut Scores: Spek 2-3 S401 Online

| Proficiency Level | Grade | Cut Score | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Tier A | Tier B/C |
| 1/2 | 2 | 220 | 24.86 | 16.67 |
|  | 3 | 234 | 26.62 | 17.55 |
| 2/3 | 2 | 273 | 26.91 | 19.30 |
|  | 3 | 283 | 26.03 | 19.01 |
| 3/4 | 2 | 322 | 24.57 | 17.55 |
|  | 3 | 332 | 25.45 | 17.55 |
| 4/5 | 2 | 374 | 35.39 | 21.64 |
|  | 3 | 386 | 40.95 | 24.28 |
| 5/6 | 2 | 415 | 62.30 | 34.80 |
|  | 3 | 425 | 72.83 | 40.07 |

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficeny level higher than 2.

Table 3.3.2.4Li
Accuracy and Consistency of Classification Indices: Spek (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.617 | 0.497 |  | 0.203 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.865 |  | 0.675 |  |
|  | 2 | 0.806 |  | 0.415 |  |
|  | 3 | 0.554 |  | 0.568 |  |
|  | 4 | - |  | 0.061 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.960 | 0.013 | 0.027 | 0.932 |
|  | $2 / 3$ | 0.746 | 0.003 | 0.251 | 0.606 |
|  | 3/4 | 0.952 | 0.048 | 0.000 | 0.906 |
|  | 4/5 | 0.999 | 0.002 | 0.000 | 0.998 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.2.4Lii
Accuracy and Consistency of Classification Indices: Spek (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.572 | 0.465 |  | 0.182 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.931 |  | 0.701 |  |
|  | 2 | 0.568 |  | 0.383 |  |
|  | 3 | 0.528 |  | 0.523 |  |
|  | 4 | - |  | 0.065 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.953 | 0.007 | 0.039 | 0.919 |
|  | 2/3 | 0.654 | 0.016 | 0.330 | 0.576 |
|  | 3/4 | 0.943 | 0.057 | 0.000 | 0.881 |
|  | 4/5 | 0.999 | 0.001 | 0.000 | 0.998 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.3 Grades: 4-5

### 3.3.3.1 Listening 4-5

Table 3.3.3.1A
Complete Item Analysis and Summary: List 4-5 S401 Online

| Item Type | Average Item Difficulty (in logits) | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | 0.25 | 54 | 0.63 | 0.98 | 1.02 |
|  | Item |  |  | Fit <br> Statistics |  |
| Name | Difficulty <br> (in logits) | Anchored? | P-value | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.LA5A_SL_MorningArrivalAtClas s_P 100_Screen_2_12413 | -1.91 |  | 0.96 | 1.22 | 2.58 |
| 2.L45A_SI_MorningArrivalAtClas s_P 100_Screen_3_12415 | -2.76 |  | 0.98 | 0.95 | 1.45 |
| 3.L45A_SI_MorningArrivalAtClas s_P 100_Screen_4_12416 | -1.40 |  | 0.93 | 0.93 | 1.04 |
| 4.L45B_SI_SchoolClubs_P 100_Screen_2_12386 | -1.45 |  | 0.93 | 0.98 | 1.07 |
| 5.L45B_SI_SchoolClubs_P 100_Screen_3_12394 | -0.89 |  | 0.91 | 1.20 | 1.84 |
| 6.L45B_SI_SchoolClubs_P 100_Screen_4_12396 | 1.26 |  | 0.68 | 1.01 | 1.07 |
| 7.L45A_LA_Mina AndGinger_P 100_Screen_2_14205 | -2.34 |  | 0.68 | 0.95 | 0.91 |
| 8.L45A_LA_Mina AndGinger_P 100_Screen_3_14206 | -1.29 | Yes | 0.47 | 0.92 | 0.89 |
| 9.L45A_LA_MinaAndGinger_P 100_Screen_4_14207 | -0.82 |  | 0.39 | 0.97 | 0.96 |
| 10.L45B_MA_TheSchoolMascot_P 100_Screen_2_12917 | -0.78 |  | 0.30 | 0.89 | 0.87 |
| 11.L45B_MA_TheSchoolMas cot_P 100_Screen_3_12918 | -0.47 |  | 0.28 | 1.00 | 1.02 |
| 12.L45B_MA_TheSchoolMascot_P 100_Screen_4_12919 | -2.65 |  | 0.76 | 0.96 | 0.92 |
| 13.L45A_SS_RomanToo ls Artifacts_P 100_Screen_2_13024 | -1.78 |  | 0.58 | 0.99 | 0.99 |
| 14.L45A_SS_RomanTools Artifacts_P 100_Screen_3_13025 | -0.38 |  | 0.25 | 0.96 | 0.97 |
| 15.L45A_SS_RomanTools Artifacts_P 100_Screen_4_13026 | -0.76 |  | 0.33 | 0.96 | 0.96 |
| 16.L45A_SC_WindSpeed_P 100_Screen_2_12792 | -2.32 |  | 0.63 | 0.96 | 0.93 |
| 17.L45A_SC_WindSpeed_P 100_Screen_3_12793 | -2.36 |  | 0.68 | 0.99 | 0.98 |
| 18.L45A_SC_WindSpeed_P 100_Screen_4_12794 | -0.22 |  | 0.24 | 1.00 | 1.03 |
| 19.L45B_LA_BookDiscus sion_P 100_Screen_2_12571 | -1.61 |  | 0.84 | 0.84 | 0.69 |
| 20.L45B_LA_BookDiscus sion_P 100_Screen_3_12572 | -1.25 |  | 0.84 | 0.87 | 0.80 |
| 21.L45B_LA_B ookDiscus sion_P 100_Screen_4_12573 | 0.54 |  | 0.51 | 1.01 | 1.00 |
| 22.L45B_MA_Garden_P 100_Screen_2_12596 | -1.12 |  | 0.84 | 0.94 | 0.86 |
| 23.L45B_MA_Garden_P 100_Screen_3_12597 | 0.43 | Yes | 0.57 | 0.97 | 0.94 |
| 24.L45B_MA_Garden_P 100_Screen_4_12598 | 1.98 | Yes | 0.34 | 1.04 | 1.07 |
| 25.L45C_SS_RomanTools Artifacts_P 100_Screen_2_13030 | 0.75 | Yes | 0.53 | 0.94 | 0.91 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \\ \hline \end{gathered}$ | Outfit <br> Mnsq |
| 26.L45C_SS_RomanToo is Artifacts_P 100_Screen_3_13031 | 1.00 | Yes | 0.45 | 0.94 | 0.92 |
| 27.L45C_SS_RomanToo is Artifacts_P 100_Screen_4_13032 | 1.52 |  | 0.39 | 1.05 | 1.06 |
| 28.L45B_SC_Ice_P 100_Screen_2_12795 | 0.44 | Yes | 0.60 | 0.96 | 0.94 |
| 29.L45B_SC_Ice_P 100_Screen_3_12796 | 0.75 | Yes | 0.59 | 0.96 | 0.94 |
| 30.L45B_SC_Ice_P 100_Screen_4_12797 | 1.21 |  | 0.48 | 0.99 | 0.99 |
| 31.L45B_LA_EarlyLife_P 100_Alt1_Screen_2_14224 | -0.77 | Yes | 0.78 | 0.95 | 0.89 |
| 32.L45B_LA_EarlyLife_P 100_Alt I_Screen_3_14225 | -0.27 |  | 0.75 | 0.98 | 0.95 |
| 33.L45B_LA_EarlyLife_P 100_Alt I_Screen_4_14226 | -1.35 |  | 0.85 | 0.93 | 0.81 |
| 34.L45C_MA_3DShapes_P 100_Screen_2_14215 | -1.42 |  | 0.87 | 0.96 | 0.89 |
| 35.L45C_MA_3DShapes_P 100_Screen_3_14216 | 1.40 | Yes | 0.29 | 0.94 | 0.92 |
| 36.L45C_MA_3DShapes_P 100_Screen_4_14217 | 0.98 |  | 0.50 | 0.99 | 0.99 |
| 37.L45C_LA_BookDiscus sion_P 100_Screen_2_12581 | -1.90 |  | 0.98 | 1.00 | 1.22 |
| 38.L45C_LA_B ookDiscuss ion_P 100_Screen_3_12583 | 1.38 |  | 0.71 | 0.98 | 0.95 |
| 39.L45C_LA_B ookDiscuss ion_P 100_Screen_4_12585 | 0.89 |  | 0.78 | 0.97 | 0.91 |
| 40.L45C_MA_TheSchoolMascot_P 100_Screen_2_12942 | 1.44 |  | 0.74 | 1.01 | 1.03 |
| 41.L45C_MA_TheSchoolMascot_P 100_Screen_3_12943 | 3.60 |  | 0.33 | 0.95 | 0.93 |
| 42.L45C_MA_TheSchoolMascot_P 100_Screen_4_12946 | 1.46 |  | 0.75 | 0.95 | 0.92 |
| 43.L45B_S _Ro manToo ls Artifacts_P 100_Screen_2_13027 | 3.33 |  | 0.45 | 1.05 | 1.06 |
| 44.L45B_SS_RomanToo is Artifacts_P 100_Screen_3_13028 | 3.20 |  | 0.48 | 1.11 | 1.14 |
| 45.L45B_S _RomanToo is Artifacts_P 100_Screen_4_13029 | 1.83 |  | 0.70 | 0.96 | 0.94 |
| 46.L45C_SC_FromRocks To So il_P 100_Screen_2_14212 | 2.49 |  | 0.60 | 1.01 | 1.00 |
| 47.L45C_SC_FromRocks To Soil_P 100_Screen_3_14213 | 1.13 |  | 0.82 | 0.95 | 0.87 |
| 48.L45C_SC_FromRocks To Soil_P 100_Screen_4_14214 | 4.50 |  | 0.24 | 1.05 | 1.14 |
| 49.L45B_LA_EarlyLife_P 100_Screen_2_12923 | 0.31 |  | 0.92 | 0.97 | 0.89 |
| 50.L45B_LA_EarlyLife_P 100_Screen_3_12925 | 1.42 |  | 0.81 | 1.00 | 1.04 |
| 51.L45B_LA_EarlyLife_P 100_Screen_4_12928 | 1.69 |  | 0.75 | 0.96 | 0.94 |
| 52.L45B_MA_RunTime_P 100_Alti_Screen_2_13069 | 2.71 |  | 0.56 | 0.98 | 0.97 |
| 53.L45B_MA_RunTime_P 100_Alt1_Screen_3_13070 | 3.09 |  | 0.49 | 1.00 | 1.00 |
| 54.L45B_MA_RunTime_P 100_Alt_Screen_4_13071 | 0.77 |  | 0.87 | 0.97 | 0.89 |

Table 3.3.3.1B
DIF Analysis and Summary: List 4-5 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (0) |
| A | 30 | 21 | 30 | 23 |
| B | 1 | 2 | 0 | 1 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.L45A_SI_MorningArrivalAtClas s_P 100_Screen_2_12413 | A | M | A | O |
| 2.L45A_SI_MorningArrivalAtClas _PP 100_Screen_3_12415 | A | M | A | O |
| 3.L45A_SI_MorningArrivalAtClas _PP 100_Screen_4_12416 | A | F | A | O |
| 4.L45B_SI_SchoolClubs_P 100_Screen_2_12386 | A | M | A | O |
| 5.L45B_SI_SchoolClubs_P 100_Screen_3_12394 | A | F | A | O |
| 6.L45B_SI_SchoolClubs_P 100_Screen_4_12396 | A | F | A | H |
| 7.L45A_LA_Mina AndGinger_P 100_Screen_2_14205 | B | M | A | O |
| 8.L45A_LA_M ina AndGinger_P 100_Screen_3_14206 | A | F | A | H |
| 9.L45A_LA_M ${ }^{\text {ana AndGinger_P 100_Screen_4_14207 }}$ | A | F | A | H |
| 10.L45B_MA_TheSchoolMascot_P 100_Screen_2_12917 | A | M | A | H |
| 11.L45B_MA_TheSchoolMascot_P 100_Screen_3_12918 | A | M | A | H |
| 12.L45B_MA_TheSchoolMascot_P 100_Screen_4_12919 | A | F | B | O |
| 13.L45A_SS_RomanTools Artifacts_P 100_Screen_2_13024 | A | M | A | H |
| 14.L45A_SS_RomanToo ls Artifacts_P 100_Screen_3_13025 | A | M | A | H |
| 15.L45A_SS_RomanToo ls Artifacts_P 100_Screen_4_13026 | A | F | A | O |
| 16.LA5A_SC_WindSpeed_P 100_Screen_2_12792 | A | F | A | H |
| 17.L45A_SC_WindSpeed_P 100_Screen_3_12793 | A | F | A | O |
| 18.L45A_SC_WindSpeed_P 100_Screen_4_12794 | A | F | A | H |
| 19.L45B_LA_BookDis cus sion_P P 100_Screen_2_12571 | A | M | A | H |
| 20.L45B_LA_BookDiscuss ion_P 100_Screen_3_12572 | A | M | A | H |
| 21.L45B_LA_BookDiscussion_P 100_Screen_4_12573 | A | M | A | O |
| 22.L45B_MA_Garden_P 100_Screen_2_12596 | A | M | A | H |
| 23.L45B_MA_Garden_P 100_Screen_3_12597 | A | F | A | H |
| 24.L45B_MA_Garden_P 100_Screen_4_12598 | A | M | A | H |
| 25.L45C_SS_RomanTools Artifacts_P 100_Screen_2_13030 | B | F | A | H |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 26.L45C_S _RomanToo ls Artifacts_P 100_Screen_3_13031 | A | F | A | O |
| 27.L45C_SS_RomanToo ls Artifacts_P 100_Screen_4_13032 | A | M | A | H |
| 28.L45B_SC_Ice_P 100_Screen_2_12795 | A | F | A | H |
| 29.L45B_SC_Ice_P 100_Screen_3_12796 | A | F | A | O |
| 30.L45B_SC_Ice_P 100_Screen_4_12797 | A | F | A | H |
| 31.L45B_LA_EarlyLife_P 100_Alt l_Screen_2_14224 | A | M | A | O |
| 32.L45B_LA_EarlyLife_P 100_AltI_Screen_3_14225 | A | F | A | H |
| 33.L45B_LA_EarlyLife_P 100_AltI_Screen_4_14226 | A | M | A | O |
| 34.L45C_MA_3DShapes_P 100_Screen_2_14215 | A | M | A | O |
| 35.L45C_MA_3DShapes_P 100_Screen_3_14216 | A | M | A | O |
| 36.L45C_MA_3DShapes_P 100_Screen_4_14217 | A | M | A | H |
| 37.L45C_LA_BookDiscus sion_P 100_Screen_2_12581 | A | M | A | O |
| 38.L45C_LA_B ookDis cuss ion_P 100_Screen_3_12583 | A | M | A | O |
| 39.L45C_LA_B ookDis cuss ion_P 100_Screen_4_12585 | A | M | A | O |
| 40.L45C_MA_TheSchoolMascot_P 100_Screen_2_12942 | A | M | A | H |
| 41.L45C_MA_TheSchoolMascot_P 100_Screen_3_12943 | A | M | A | O |
| 42.L45C_MA_TheSchoolMascot_P 100_Screen_4_12946 | A | F | A | H |
| 43.L45B_SS_RomanToo is Artifacts_P 100_Screen_2_13027 | A | M | A | H |
| 44.L45B_S _RomanToo is Artifacts_P 100_Screen_3_13028 | A | M | A | H |
| 45.L45B_SS_RomanToo ls Artifacts_P 100_Screen_4_13029 | A | F | A | H |
| 46.L45C_SC_FromRocks To Soil_P 100_Screen_2_14212 | B | F | A | O |
| 47.L45C_SC_FromRocksTo Soil_P 100_Screen_3_14213 | A | F | A | H |
| 48.L45C_SC_FromRocksToSoil_P 100_Screen_4_14214 | A | F | A | O |
| 49.L45B_LA_EarlyLife_P 100_Screen_2_12923 | A | M | A | H |
| 50.L45B_LA_EarlyLife_P 100_Screen_3_12925 | A | M | A | O |
| 51.L45B_LA_EarlyLife_P 100_Screen_4_12928 | A | M | A | H |
| 52.L45B_MA_RunTime_P 100_AltI_Screen_2_13069 | A | F | A | O |
| 53.L45B_MA_RunTime_P 100_AltI_Screen_3_13070 | A | M | A | H |
| 54.L45B_MA_RunTime_P 100_AltıScreen_4_13071 | A | M | A | H |

Figure 3.3.3.1C
Raw Scores: List 4-5 S401 Online
n/a
Table 3.3.3.1C
Raw Score Descriptive Statistics: List 4-5 S401 Online n/a


Figure 3.3.3.1E
Proficiency Level: List 4-5 S401 Online


Table 3.3.3.1D
Scale Score Descriptive Statistics: List 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 102,781 | 116 | 516 | 400.80 | 53.57 |
| $\mathbf{5}$ | 74,998 | 120 | 516 | 405.35 | 60.33 |
| Total | 177,779 | 116 | 516 | 402.72 | 56.56 |

Table 3.3.3.1E
Proficiency Level Distribution: List 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 4,070 | $3.96 \%$ | 4,534 | $6.05 \%$ | 8,604 | $4.84 \%$ |
| $\mathbf{2}$ | 2,587 | $2.52 \%$ | 2,114 | $2.82 \%$ | 4,701 | $2.64 \%$ |
| $\mathbf{3}$ | 7,858 | $7.65 \%$ | 9,281 | $12.37 \%$ | 17,139 | $9.64 \%$ |
| $\mathbf{4}$ | 6,510 | $6.33 \%$ | 5,124 | $6.83 \%$ | 11,634 | $6.54 \%$ |
| $\mathbf{5}$ | 10,085 | $9.81 \%$ | 5,225 | $6.97 \%$ | 15,310 | $8.61 \%$ |
| $\mathbf{6}$ | 71,671 | $69.73 \%$ | 48,720 | $64.96 \%$ | 120,391 | $67.72 \%$ |
| Total | 102,781 | $100.00 \%$ | 74,998 | $100.00 \%$ | 177,779 | $100.00 \%$ |

Table 3.3.3.1F
Raw Score to Scale Score Conversion: List 4-5 S401 Online n/a

Table 3.3.3.1G
Equating Summary: List 4-5 S401 Online


[^19]Figure 3.3.3.1H
Test Characteristic Curve: List 4-5 S401 Online n/a


Table 3.3.3.1J
Reliability: List 4-5 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 177,779 | 54 | .85 |

Table 3.3.3.1K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 4-5 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 4 | 275 | 9 | 17.35 | 18.88 | 18.71 | 0.51 |
|  | 5 | 285 | N/A | N/A | N/A | N/A | N/A |
| 2/3 | 4 | 313 | 67 | 17.86 | 18.37 | 17.86 | 0.06 |
|  | 5 | 323 | N/A | N/A | N/A | N/A | N/A |
| 3/4 | 4 | 343 | 1,684 | 18.37 | 19.39 | 18.37 | 0.04 |
|  | 5 | 354 | 35 | 17.86 | 19.39 | 19.08 | 0.62 |
| 4/5 | 4 | 363 | 879 | 17.86 | 19.39 | 19.33 | 0.27 |
|  | 5 | 375 | 6 | 19.39 | 20.92 | 19.90 | 0.79 |
| 5/6 | 4 | 388 | N/A | N/A | N/A | N/A | N/A |
|  | 5 | 401 | N/A | N/A | N/A | N/A | N/A |

Table 3.3.3.1 Li
Accuracy and Consistency of Classification Indices: List (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.780 | 0.724 |  | 0.451 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.907 |  | 0.781 |  |
|  | 2 | 0.331 |  | 0.230 |  |
|  | 3 | 0.540 |  | 0.393 |  |
|  | 4 | 0.314 |  | 0.226 |  |
|  | 5 | 0.347 |  | 0.249 |  |
|  | 6 | 0.932 |  | 0.911 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.983 | 0.003 | 0.014 | 0.978 |
|  | 2/3 | 0.975 | 0.013 | 0.012 | 0.961 |
|  | 3/4 | 0.946 | 0.019 | 0.035 | 0.926 |
|  | 4/5 | 0.932 | 0.026 | 0.042 | 0.902 |
|  | 5/6 | 0.906 | 0.047 | 0.048 | 0.863 |

Table 3.3.3.1 Lii
Accuracy and Consistency of Classification Indices: List (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.741 | 0.681 |  | 0.442 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.900 |  | 0.753 |  |
|  | 2 | 0.241 |  | 0.167 |  |
|  | 3 | 0.602 |  | 0.463 |  |
|  | 4 | 0.272 |  | 0.200 |  |
|  | 5 | 0.231 |  | 0.161 |  |
|  | 6 | 0.939 |  | 0.910 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.975 | 0.004 | 0.021 | 0.965 |
|  | 2/3 | 0.961 | 0.023 | 0.016 | 0.940 |
|  | 3/4 | 0.921 | 0.023 | 0.057 | 0.897 |
|  | 4/5 | 0.922 | 0.026 | 0.053 | 0.886 |
|  | 5/6 | 0.905 | 0.057 | 0.038 | 0.857 |

### 3.3.3.2 Reading 4-5

Table 3.3.3.2A
Complete Item Analysis and Summary: Read 4-5 S401 Online

|  | Average <br> Item <br> Difficulty <br> (in logits) | No. of Items | Average <br> Infit <br> Merage <br> P-value | Average <br> Outfit <br> Mean <br> Square |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Square |  |  |  |  |$|$| Item Type |
| :--- |
| Selected Response |


| Name | Item <br> Difficulty <br> (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 26.R45C_LA_BrunelDavinci_203_P 100_A301_Screen_3_13489 | 1.33 |  | 0.44 | 1.04 | 1.05 |
| 27.R45C_LA_BrunelDavinci_203_P 100_A301_Screen_4_13490 | 2.11 | Yes | 0.29 | 1.14 | 1.28 |
| 28.R45A_MA_Clas sroomCleanup_401_V2_Screen_2_14625 | -0.08 |  | 0.51 | 1.00 | 1.00 |
| 29.R45A_MA_Clas sroomCleanup_401_V2_Screen_3_14626 | -0.26 |  | 0.57 | 0.94 | 0.93 |
| 30.R45A_MA_Clas sroomCleanup_401_V2_Screen_4_4627 | 1.59 |  | 0.20 | 0.98 | 0.97 |
| 31.R45B_SS_Antarctica_jc_P 100_A203_Alt_Screen_2_13536 | -1.15 |  | 0.81 | 0.94 | 0.88 |
| 32.R45B_SS_Antarctica_jc_P 100_A 203 _Altt_Screen_3_13537 | 1.83 | Yes | 0.23 | 0.96 | 0.95 |
| 33.R45B_SS_Antarctic a_ic_P 100_A 203 _Altı_Screen_4_13538 | 1.78 |  | 0.22 | 0.99 | 0.98 |
| 34.R45B_SC_Rocks_mika_P 100_A301_Screen_2_13485 | 0.30 | Yes | 0.50 | 0.94 | 0.93 |
|  | 1.18 | Yes | 0.36 | 0.96 | 0.96 |
| 36.R45B_SC_Rocks_mika_P 100_A301_Screen_4_13487 | 1.36 |  | 0.28 | 0.98 | 0.97 |
| 37.R45B_MA_Class Schedule_jc_P 100_A201_Screen_2_13479 | -0.69 |  | 0.75 | 0.93 | 0.88 |
| 38.R45B_MA_Clas s Schedule_jc_P 100_A201_Screen_3_13480 | 0.75 | Yes | 0.45 | 0.93 | 0.92 |
| 39.R45B_MA_Clas S Schedule_jc_P 100_A201_Screen_4_13481 | 1.68 |  | 0.23 | 0.98 | 0.98 |
| 40.R45B_SS_Declaratio n_P 100_A203FT_Screen_2_13527 | 1.08 | Yes | 0.38 | 0.99 | 0.99 |
| 41.R45B_SS_Declaration_P 100_A 203FT_Screen_3_13528 | 1.64 | Yes | 0.26 | 1.01 | 1.01 |
| 42.R45B_SS_Declaratio n_P 100_A203FT_Screen_4_13529 | 0.98 | Yes | 0.34 | 1.00 | 1.00 |
| 43.R45C_SC_Tides_P 100_A301_Screen_2_13497 | 1.99 | Yes | 0.19 | 1.00 | 1.00 |
| 44.R45C_SC_Tides_P 100_A301_Screen_3_13498 | -0.05 | Yes | 0.62 | 0.98 | 0.98 |
| 45.R45C_SC_Tides_P 100_A301_Screen_4_13499 | 1.17 |  | 0.33 | 1.01 | 1.01 |
| 46.R45C_MA_So occerlnTheP ark_401_V1_Screen_2_13926 | 1.50 |  | 0.47 | 1.02 | 1.02 |
| 47.R45C_MA_SoccerlnTheP ark_401_V1_Screen_3_13927 | 2.35 |  | 0.32 | 1.25 | 1.34 |
| 48.R45C_MA_SoccerlnTheP ark_401_V1_Screen_4_13928 | 2.86 |  | 0.23 | 1.22 | 1.34 |
| 49.R45B_SS_Native Americ anShelters_namo_P 100_A202_Screen_2_13482 | 2.46 |  | 0.34 | 0.95 | 0.94 |
| 50.R45B_SS_Native Americ anShelters_namo_P 100_A 202_Screen_3_13483 | 2.13 | Yes | 0.37 | 0.97 | 0.96 |
| 51.R45B_SS_Native AmericanShelters_namo_P 100_A202_Screen_4_13484 | 2.41 |  | 0.36 | 1.01 | 1.00 |
| 52.R45B_SC_Geysers_P 100_A301_Alt _S Screen_2_13539 | -0.63 | Yes | 0.87 | 0.87 | 0.66 |
| 53.R45B_SC_Geysers_P 100_A301_Alt I_Screen_3_13540 | 1.92 |  | 0.47 | 0.93 | 0.91 |
| 54.R45B_SC_Geysers_P 100_A301_Alt I_Screen_4_13541 | 2.13 | Yes | 0.39 | 0.97 | 0.95 |
| 55.R45B_LA_Onthe Train_P 100_A202_Alt I_Screen_2_13530 | 1.35 |  | 0.60 | 0.97 | 0.98 |
| 56.R45B_LA_Onthe Train_P 100_A202_AltI_Screen_3_13531 | 2.11 |  | 0.42 | 0.92 | 0.90 |
| 57.R45B_LA_OntheTrain_P 100_A202_Alt1_Screen_4_13532 | 1.51 | Yes | 0.52 | 0.92 | 0.90 |
| 58.R45B_MA_Clas s Schedule_jc_P 100_A201_Alt _Screen_2_13533 | 1.43 |  | 0.60 | 0.93 | 0.91 |
| 59.R45B_MA_Clas s Schedule_jc_P 100_A201_Alt | 0.99 |  | 0.65 | 0.94 | 0.95 |
| 60.R45B_MA_Clas s Schedule_jc_P 100_A201_AltıScreen_4_13535 | 2.38 |  | 0.38 | 0.93 | 0.91 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 61.R45C_SS_EleanorRoosevelt_203_P 100_A301_Screen_2_13503 | 2.15 | Yes | 0.42 | 0.94 | 0.92 |
| 62.R45C_SS_EleanorRoosevelt_203_P 100_A301_Screen_3_13504 | 2.57 |  | 0.39 | 0.97 | 0.96 |
| 63.R45C_SS_EleanorRoosevelt_203_P 100_A301_Screen_4_13505 | 1.96 |  | 0.49 | 0.92 | 0.90 |
| 64.R45C_SC_Glaciers_Pas z_P 100_A 301_Screen_2_13506 | 1.88 |  | 0.53 | 0.97 | 0.95 |
| 65.R45C_SC_Glaciers_pas z_P 100_A 301_Screen_3_13507 | 1.55 |  | 0.58 | 0.93 | 0.91 |
| 66.R45C_SC_Glaciers_pas z_P 100_A 301 _Screen_4_13508 | 3.18 |  | 0.28 | 0.99 | 1.00 |

Table 3.3.3.2B
DIF Analysis and Summary: Read 4-5 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (0) |
| A | 37 | 29 | 31 | 35 |
| B | 0 | 0 | 0 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.R45A_SI_WatchingTV_P 100_A301FT_Screen_2_13509 | A | F | A | H |
| 2.R45A_SI_WatchingTV_P 100_A 301 FT _Screen_3_13510 | A | F | A | H |
| 3.R45A_SI_WatchingTV_P 100_A 301 FT _Screen_4_13511 | A | M | A | H |
| 4.R45B_SI_P hoto Contest_P 100_A301FT_Screen_2_13449 | A | F | A | O |
| 5.R45B_SI_P hoto Contest_P 100_A301FT_Screen_3_13450 | A | F | A | O |
| 6.R45B_SI_P hoto Contest_P 100_A301FT_Screen_4_13451 | A | F | A | H |
| 7.R45A_LA_Jo sephStraus s_203_P 100_A301_Screen_2_13461 | A | M | A | O |
| 8.R45A_LA_Jo sephStraus s_203_P 100_A301_Screen_3_13462 | A | F | A | O |
| 9.R45A_LA_Jo sephStraus s_203_P 100_A301_Screen_4_13463 | A | M | A | O |
| 10.R45A_MA_Supermarket_pasz_P 100_A203_Screen_2_13464 | A | F | A | H |
| 11.R45A_MA_Supermarket_pasz_P 100_A203_Screen_3_13465 | A | M | A | O |
| 12.R45A_MA_Supermarket_pasz_P 100_A203_Screen_4_13466 | A | F | A | H |
| 13.R45A_SS_ShoppingCart_401_V2_Screen_2_1774 | A | F | A | O |
| 14.R45A_SS_ShoppingCart_401_V2_Screen_3_14715 | A | M | A | H |
| 15.R45A_SS_ShoppingCart_401_V2_Screen_4_4716 | A | M | A | O |
| 16.R45B_SC_Geysers_P 100_A 301_45_Screen_2_13407 | A | F | A | H |
| 17.R45B_SC_Geysers_P 100_A 301_45_Screen_3_13408 | A | M | A | H |
| 18.R45B_SC_Geysers_P 100_A301_45_Screen_4_13409 | A | M | A | O |
| 19.R45B_LA_Onthe Train_P 100_A202_Screen_2_13467 | A | M | A | H |
| 20.R45B_LA_OntheTrain_P 100_A 202_Screen_3_13468 | A | M | A | O |
| 21.R 45B_LA_OntheTrain_P 100_A202_Screen_4_13469 | A | F | A | O |
| 22.R45A_MA_BuyingSchoolSupplies_P 100_A201_Screen_2_13518 | A | M | A | H |
| 23.R45A_MA_BuyingSchoolSupplies_P 100_A201_Screen_3_13519 | A | M | A | H |
| 24.R45A_MA_BuyingSchoo ISupplies_P 100_A201_Screen_4_13520 | A | M | A | H |
| 25.R45C_LA_BrunelDavinci_203_P 100_A301_Screen_2_13488 | A | F | A | H |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | DIF Level | Favored Group |  | Favored Group |
| 26.R45C_LA_BrunelDavinci_203_P 100_A301_Screen_3_13489 | A | F | A | O |
| 27.R45C_LA_BrunelDavinci_203_P 100_A301_Screen_4_13490 | A | F | A | O |
| 28.R45A_MA_Clas sroomCle anup_401_V2_Screen_2_14625 | A | F | A | O |
| 29.R45A_MA_Clas sroomCleanup_401_V2_Screen_3_14626 | A | F | A | H |
| 30.R45A_MA_Clas sroomCle anup_401_V2_Screen_4_14627 | A | M | A | H |
| 31.R45B_SS_Antarctica_jc_P 100_A203_Alt 1_Screen_2_13536 | A | M | A | H |
| 32.R45B_SS_Antarctica_jc_P 100_A203_Alt 1_Screen_3_13537 | A | F | A | O |
| 33.R45B_SS_Antarctica_jc_P 100_A203_Alt 1_Screen_4_13538 | A | F | A | O |
| 34.R45B_SC_Rocks_mika_P 100_A301_Screen_2_13485 | A | M | A | H |
| 35.R45B_SC_Rocks_mika_P 100_A301_Screen_3_13486 | A | F | A | O |
| 36.R45B_SC_Rocks_mika_P 100_A301_Screen_4_13487 | A | M | A | H |
| 37.R45B_MA_Class Schedule_jc_P 100_A201_Screen_2_13479 | A | M | A | H |
| 38.R45B_MA_Class Schedule_jc_P 100_A201_Screen_3_13480 | A | M | A | O |
| 39.R45B_MA_Class Schedule_jc_P 100_A201_Screen_4_13481 | A | F | A | O |
| 40.R45B_SS_Declaration_P 100_A203FT_Screen_2_13527 | A | M | A | O |
| 41.R45B_SS_Declaration_P 100_A203FT_Screen_3_13528 | A | M | A | O |
| 42.R45B_SS_Declaration_P 100_A203FT_Screen_4_13529 | A | M | A | O |
| 43.R45C_SC_Tides_P 100_A301_Screen_2_13497 | A | F | A | O |
| 44.R45C_SC_Tides_P 100_A301_Screen_3_13498 | A | M | A | H |
| 45.R45C_SC_Tides_P 100_A301_Screen_4_13499 | A | F | A | O |
| 46.R45C_M A_SoccerinTheP ark_401_V1_Screen_2_13926 | A | M | A | H |
| 47.R45C_MA_SoccerinTheP ark_401_V1_Screen_3_13927 | A | M | A | H |
| 48.R45C_MA_SoccerlnTheP ark_401_V1_Screen_4_13928 | A | M | A | O |
| 49.R45B_SS_Native AmericanShelters_namo_P 100_A202_Screen_2_13482 | A | F | A | H |
| 50.R45B_SS_Native AmericanShelters_namo_P 100_A202_Screen_3_13483 | A | M | A | H |
| 51.R45B_SS_Native AmericanShelters _namo_P 100_A202_Screen_4_13484 | A | F | A | O |
| 52.R45B_SC_Geysers _P 100_A301_Alt 1_Screen_2_13539 | A | M | A | H |
| 53.R45B_SC_Geysers _P 100_A301_Alt 1_Screen_3_13540 | A | F | A | H |
| 54.R45B_SC_Geysers _P 100_A301_Alt 1_Screen_4_13541 | A | F | A | O |
| 55.R45B_LA_Onthe Train_P 100_A202_Alt1_Screen_2_13530 | A | M | A | H |
| 56.R45B_LA_OntheTrain_P 100_A202_Alt1_Screen_3_13531 | A | M | A | H |
| 57.R45B_LA_Onthe Train_P 100_A202_Alt1_Screen_4_13532 | A | M | A | O |
| 58.R45B_MA_Class Schedule_jc_P 100_A201_Alt 1_Screen_2_13533 | A | F | A | H |
| 59.R45B_MA_Clas SSchedule_jc_P 100_A201_Alt 1_Screen_3_13534 | A | M | A | O |
| 60.R45B_MA_Class Schedule_jc_P 100_A201_Alt 1_Screen_4_13535 | A | M | A | O |


|  | Male/Female |  | Hispanic/Other |  |
| :--- | :---: | :---: | :---: | :---: |
|  | Name | $\begin{array}{c}\text { DIF } \\ \text { Level }\end{array}$ | $\begin{array}{c}\text { Favored } \\ \text { Group }\end{array}$ | $\begin{array}{c}\text { DIF } \\ \text { Level }\end{array}$ |
| 6avored |  |  |  |  |
| Group |  |  |  |  |$]$

Figure 3.3.3.2C
Raw Scores: Read 4-5 S401 Online n/a

Table 3.3.3.2C
Raw Score Descriptive Statistics: Read 4-5 S401 Online n/a


Figure 3.3.3.2E
Proficiency Level: Read 4-5 S401 Online


Table 3.3.3.2D
Scale Score Descriptive Statistics: Read 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 91,056 | 220 | 456 | 342.98 | 30.34 |
| $\mathbf{5}$ | 65,391 | 220 | 456 | 346.52 | 35.02 |
| Total | 156,447 | 220 | 456 | 344.46 | 32.43 |

Table 3.3.3.2E
Proficiency Level Distribution: Read 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 10,044 | $11.03 \%$ | 12,235 | $18.71 \%$ | 22,279 | $14.24 \%$ |
| $\mathbf{2}$ | 23,320 | $25.61 \%$ | 20,799 | $31.81 \%$ | 44,119 | $28.20 \%$ |
| $\mathbf{3}$ | 22,740 | $24.97 \%$ | 11,189 | $17.11 \%$ | 33,929 | $21.69 \%$ |
| $\mathbf{4}$ | 10,713 | $11.77 \%$ | 6,850 | $10.48 \%$ | 17,563 | $11.23 \%$ |
| $\mathbf{5}$ | 17,089 | $18.77 \%$ | 8,353 | $12.77 \%$ | 25,442 | $16.26 \%$ |
| $\mathbf{6}$ | 7,150 | $7.85 \%$ | 5,965 | $9.12 \%$ | 13,115 | $8.38 \%$ |
| Total | 91,056 | $100.00 \%$ | 65,391 | $100.00 \%$ | 156,447 | $100.00 \%$ |

Table 3.3.3.2F
Raw Score to Scale Score Conversion: Read 4-5 S401 Online n/a

Table 3.3.3.2G
Equating Summary: Read 4-5 S401 Online


[^20]Figure 3.3.3.2H
Test Characteristic Curve: Read 4-5 S401 Online n/a

Figure 3.3.3.2I


Table 3.3.3.2J
Reliability: Read 4-5 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 156,447 | 66 | .88 |

Table 3.3.3.2K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 4-5 S401 Online

| Proficiency <br> Level | Grade | Cut Score | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1 / 2}$ | 4 | 307 | 111 | 11.22 | 13.27 | 11.82 | 0.48 |
|  | 5 | 316 | 413 | 11.22 | 11.22 | 11.22 | 0.00 |
| $\mathbf{2 / 3}$ | 4 | 335 | 5,439 | 10.20 | 11.22 | 10.54 | 0.39 |
|  | 5 | 345 | 304 | 10.20 | 10.71 | 10.56 | 0.23 |
| $\mathbf{3 / 4}$ | 4 | 354 | 284 | 10.20 | 10.71 | 10.30 | 0.20 |
|  | 5 | 5 | 364 | 1,448 | 10.20 | 10.71 | 10.22 |
| $\mathbf{4 / 5}$ | 4 | 364 | 4,630 | 10.20 | 10.71 | 10.21 | 0.09 |
|  | 5 | 373 | 99 | 10.20 | 10.71 | 10.56 | 0.23 |

Table 3.3.3.2Li
Accuracy and Consistency of Classification Indices: Read (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.614 | 0.504 |  | 0.386 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.804 |  | 0.666 |  |
|  | 2 | 0.686 |  | 0.578 |  |
|  | 3 | 0.574 |  | 0.465 |  |
|  | 4 | 0.328 |  | 0.250 |  |
|  | 5 | 0.629 |  | 0.502 |  |
|  | 6 | 0.714 |  | 0.543 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.949 | 0.019 | 0.032 | 0.927 |
|  | 2/3 | 0.894 | 0.053 | 0.053 | 0.850 |
|  | 3/4 | 0.887 | 0.062 | 0.051 | 0.841 |
|  | 4/5 | 0.897 | 0.063 | 0.040 | 0.859 |
|  | 5/6 | 0.953 | 0.025 | 0.021 | 0.930 |

Table 3.3.3.2Lii
Accuracy and Consistency of Classification Indices: Read (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.616 | 0.513 |  | 0.393 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.790 |  | 0.677 |  |
|  | 2 | 0.702 |  | 0.601 |  |
|  | 3 | 0.442 |  | 0.344 |  |
|  | 4 | 0.347 |  | 0.260 |  |
|  | 5 | $0.518$ |  | 0.398 |  |
|  | 6 | 0.779 |  | 0.624 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.918 | 0.038 | 0.044 | 0.883 |
|  | 2/3 | 0.889 | 0.050 | 0.061 | 0.843 |
|  | 3/4 | 0.898 | 0.064 | 0.038 | 0.860 |
|  | 4/5 | 0.913 | 0.048 | 0.039 | 0.880 |
|  | 5/6 | 0.955 | 0.026 | 0.018 | 0.934 |

### 3.3.3.3 Writing 4-5

### 3.3.3.3i <br> Writing 4-5 A

Table 3.3.3.3Ai
Complete Task Analysis and Summary: Writ 4-5 A S401 Online

| Task Type |  | Average Task Difficulty (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 1.26 | 3 | 0.39 | 0.41 |
| Name |  | Task Difficulty (in logits) |  | Fit Statistics |  |
|  |  | Anchored? | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.W45A_SI_DismissalRules_kawo_P100_A301_14266 |  |  | 1.41 | Yes | 0.43 | 0.45 |
| 2.W45A_LA_Umbrella_203_P100_A 301 _14267 |  | 1.15 | Yes | 0.35 | 0.38 |
| 3.W45A_MS_PatternsInNature_pale_P100_A 203 _14270 |  | 1.22 | Yes | 0.39 | 0.40 |
| Raw Score Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 17.44\% | 9.89\% | 15.54\% |  |
|  | 1 | 10.67\% | 9.61\% | 6.44\% |  |
|  | 2 | 16.31\% | 17.42\% | 15.95\% |  |
|  | 3 | 22.70\% | 27.18\% | 29.29\% |  |
|  | 4 | 22.90\% | 26.47\% | 26.78\% |  |
|  | 5 | 8.08\% | 8.04\% | 5.06\% |  |
|  | 6 | 1.75\% | 1.31\% | 0.87\% |  |
|  | 7 | 0.14\% | 0.07\% | 0.06\% |  |
|  | 8 | 0.01\% | 0.00\% | 0.01\% |  |
|  | 9 | 0.00\% | 0.00\% | 0.00\% |  |

Table 3.3.3.3Bi
DIF Analysis and Summary: Writ 4-5 A S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.W45A_SI_DismissalRules_Kawo_P 100_A301_14266 | AA | F | AA | H |  |
| 2.W45A_LA_Umbrella_203_P 100_A301_14267 | AA | F | AA | O |  |
| 3.W45A_MS_Patterns InNature_pale_P 100_A203_14270 | AA | M | AA | O |  |

Figure 3.3.3.3Ci
Raw Scores: Writ 4-5 A S401 Online


Table 3.3.3.3Ci
Raw Score Descriptive Statistics: Writ 4-5 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 14,443 | 0 | 19 | 7.47 | 4.10 |
| $\mathbf{5}$ | 14,205 | 0 | 20 | 8.53 | 4.11 |
| Total | 28,648 | 0 | 20 | 8.00 | 4.14 |

Table 3.3.3.3Di
Scale Score Descriptive Statistics: Writ 4-5 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 14,443 | 144 | 388 | 272.23 | 44.40 |
| $\mathbf{5}$ | 14,205 | 155 | 398 | 282.60 | 39.98 |
| Total | 28,648 | 144 | 398 | 277.37 | 42.58 |



Table 3.3.3.3Ei
Proficiency Level Distribution: Writ 4-5 A S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 4,389 | $30.39 \%$ | 3,287 | $23.14 \%$ | 7,676 | $26.79 \%$ |
| $\mathbf{2}$ | 4,624 | $32.02 \%$ | 3,885 | $27.35 \%$ | 8,509 | $29.70 \%$ |
| $\mathbf{3}$ | 5,359 | $37.10 \%$ | 6,970 | $49.07 \%$ | 12,329 | $43.04 \%$ |
| $\mathbf{4}$ | 71 | $0.49 \%$ | 63 | $0.44 \%$ | 134 | $0.47 \%$ |
| $\mathbf{5}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 14,443 | $100.00 \%$ | 14,205 | $100.00 \%$ | 28,648 | $100.00 \%$ |

Table 3.3.3.3Fi
Raw Score to Scale Score Conversion: Writ 4-5 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $155^{\wedge}$ | 129.15 | $100.00^{\wedge}$ | 333.15 |
| 1 | 231 | 23.41 | 207.59 | 254.41 |
| 2 | 245 | 16.35 | 228.65 | 261.35 |
| 3 | 253 | 13.69 | 239.31 | 266.69 |
| 4 | 259 | 12.41 | 246.59 | 271.41 |
| 5 | 264 | 11.79 | 252.21 | 275.79 |
| 6 | 270 | 11.63 | 258.37 | 281.63 |
| 7 | 275 | 11.76 | 263.24 | 286.76 |
| 8 | 280 | 12.24 | 267.76 | 292.24 |
| 9 | 286 | 13.00 | 273.00 | 299.00 |
| 10 | 293 | 14.04 | 278.96 | 307.04 |
| 11 | 301 | 15.22 | 285.78 | 316.22 |
| 12 | 310 | 16.30 | 293.70 | 326.30 |
| 13 | 320 | 17.13 | 302.87 | 337.13 |
| 14 | 332 | 17.64 | 314.36 | 349.64 |
| 15 | 343 | 17.88 | 325.12 | 360.88 |
| 16 | 355 | 17.83 | 337.17 | 372.83 |
| 17 | 367 | 17.51 | 349.49 | 384.51 |
| 18 | 378 | 16.97 | 361.03 | 394.97 |
| 19 | 388 | 16.33 | 371.67 | 404.33 |
| 20 | 398 | 15.71 | 382.29 | 413.71 |
| 21 | 407 | 15.31 | 391.69 | 422.31 |
| 22 | 416 | 15.20 | 400.80 | 431.20 |
| 23 | 424 | 15.60 | 408.40 | 439.60 |
| 24 | 434 | 16.76 | 417.24 | 450.76 |
| 25 | 446 | 19.39 | 426.61 | 465.39 |
| 26 | 464 | 26.50 | 437.50 | 490.50 |
| 27 | 496 | 48.41 | 447.59 | 544.41 |

$\wedge$ Truncated

Table 3.3.3.3Gi
Equating Summary: Writ 4-5 A S401 Online


[^21]

Table 3.3.3.3Ji
Reliability: Writ 4-5 A S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's <br> Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 28,648 | 3 | Hand-written (HW) | Keyboarded (KB) | . 881 | 1.429 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 9,116 | 98 | 2 | 0 |
|  |  | KB | 14,122 | 95 | 5 | 0 |
|  | 2 | HW | 7,186 | 97 | 3 | 0 |
|  |  | KB | 13,974 | 96 | 4 | 0 |
|  | 3 | HW | 8,108 | 98 | 2 | 0 |
|  |  | KB | 14,000 | 96 | 4 | 0 |

Table 3.3.3.3Ki
Conditional Standard Error of Measurement at Cut Scores: Writ 4-5 A S401 Online n/a

Table 3.3.3.3Li
Accuracy and Consistency of Classification Indices: Writ 4-5 A S401 Online n/a

Table 3.3.3.3Aii
Complete Task Analysis and Summary: Writ 4-5 B/C S401 Online

| Task Type |  | Average Task Difficulty (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average Outfit Mean Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 2.00 | 3 | 0.59 | 0.59 |
| Name |  | Task Difficulty (in logits) | Anchored? | Fit Statistics |  |
|  |  | Infit <br> Mnsq |  | Outfit Mnsq |
| 1.W45B_SI_FieldTripRules_MaEs SaNa_P100_A301_1272 |  |  | 1.51 | Yes | 0.96 | 0.94 |
| 2.W45B_MS_ThePlanets_JeCaDaKi_P100_A203_14273 |  | 2.39 | Yes | 0.50 | 0.52 |
| 3.W45C_IT_AlgonquinChildren_401_HW_15757 |  | 2.10 |  | 0.31 | 0.30 |
| Raw Score <br> Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 0.62\% | 0.84\% | 1.53\% |  |
|  | 1 | 0.86\% | 0.85\% | 0.83\% |  |
|  | 2 | 3.16\% | 4.10\% | 4.79\% |  |
|  | 3 | 8.32\% | 27.05\% | 13.54\% |  |
|  | 4 | 25.22\% | 41.26\% | 38.09\% |  |
|  | 5 | 36.63\% | 19.95\% | 29.45\% |  |
|  | 6 | 21.32\% | 5.40\% | 10.59\% |  |
|  | 7 | 3.54\% | 0.52\% | 1.04\% |  |
|  | 8 | 0.31\% | 0.03\% | 0.13\% |  |
|  | 9 | 0.03\% | 0.00\% | 0.02\% |  |

Table 3.3.3.3Bii
DIF Analys is and Summary: Writ 4-5 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF | Favored | DIF | Favored |  |
| Level | Group | Level | Group |  |  |
| L.W45B_S_FieldTripRules_MaEsSaNa_P 100_A301_14272 | AA | F | AA | O |  |
| 2.W45B_MS_TheP lanets_JeCaDaKi_P 100_A203_14273 | AA | F | AA | H |  |
| 3.W45C_IT_AlgonquinChildren_401_HW_15757 | AA | M | AA | O |  |

Figure 3.3.3.3Cii Raw Scores: Writ 4-5 B/C S401 Online


Table 3.3.3.3Cii
Raw Score Descriptive Statistics: Writ 4-5 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 58,798 | 0 | 47 | 24.15 | 5.72 |
| $\mathbf{5}$ | 39,842 | 0 | 49 | 26.79 | 5.19 |
| Total | 98,640 | 0 | 49 | 25.22 | 5.66 |

Figure 3.3.3.3Dii


Table 3.3.3.3Dii
Scale Score Descriptive Statistics: Writ 4-5 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 58,798 | 144 | 452 | 336.84 | 27.03 |
| $\mathbf{5}$ | 39,842 | 155 | 463 | 349.67 | 26.19 |
| Total | 98,640 | 144 | 463 | 342.02 | 27.43 |



Table 3.3.3.3Eii
Proficiency Level Distribution: Writ 4-5 B/C S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 344 | $0.59 \%$ | 80 | $0.20 \%$ | 424 | $0.43 \%$ |
| $\mathbf{2}$ | 1,137 | $1.93 \%$ | 431 | $1.08 \%$ | 1,568 | $1.59 \%$ |
| $\mathbf{3}$ | 40,885 | $69.53 \%$ | 24,675 | $61.93 \%$ | 65,560 | $66.46 \%$ |
| $\mathbf{4}$ | 15,721 | $26.74 \%$ | 14,157 | $35.53 \%$ | 29,878 | $30.29 \%$ |
| $\mathbf{5}$ | 676 | $1.15 \%$ | 463 | $1.16 \%$ | 1,139 | $1.15 \%$ |
| $\mathbf{6}$ | 35 | $0.06 \%$ | 36 | $0.09 \%$ | 71 | $0.07 \%$ |
| Total | 58,798 | $100.00 \%$ | 39,842 | $100.00 \%$ | 98,640 | $100.00 \%$ |

Table 3.3.3.3Fii
Raw Score to Scale Score Conversion: Writ 4-5 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 155^ | 138.82 | $100.00^{\wedge}$ | 347.82 | 34 | 389 | 12.35 | 376.65 | 401.35 |
| 1 | 237 | 24.25 | 212.75 | 261.25 | 35 | 395 | 12.19 | 382.81 | 407.19 |
| 2 | 252 | 16.59 | 235.41 | 268.59 | 36 | 401 | 11.98 | 389.02 | 412.98 |
| 3 | 260 | 13.51 | 246.49 | 273.51 | 37 | 406 | 11.76 | 394.24 | 417.76 |
| 4 | 266 | 11.79 | 254.21 | 277.79 | 38 | 411 | 11.55 | 399.45 | 422.55 |
| 5 | 270 | 10.71 | 259.29 | 280.71 | 39 | 416 | 11.36 | 404.64 | 427.36 |
| 6 | 274 | 9.96 | 264.04 | 283.96 | 40 | 420 | 11.17 | 408.83 | 431.17 |
| 7 | 278 | 9.45 | 268.55 | 287.45 | 41 | 425 | 11.01 | 413.99 | 436.01 |
| 8 | 281 | 9.08 | 271.92 | 290.08 | 42 | 429 | 10.90 | 418.10 | 439.90 |
| 9 | 284 | 8.78 | 275.22 | 292.78 | 43 | 434 | 10.85 | 423.15 | 444.85 |
| 10 | 287 | 8.59 | 278.41 | 295.59 | 44 | 438 | 10.87 | 427.13 | 448.87 |
| 11 | 290 | 8.49 | 281.51 | 298.49 | 45 | 443 | 10.96 | 432.04 | 453.96 |
| 12 | 292 | 8.40 | 283.60 | 300.40 | 46 | 447 | 11.14 | 435.86 | 458.14 |
| 13 | 295 | 8.40 | 286.60 | 303.40 | 47 | 452 | 11.47 | 440.53 | 463.47 |
| 14 | 298 | 8.46 | 289.54 | 306.46 | 48 | 457 | 11.95 | 445.05 | 468.95 |
| 15 | 300 | 8.57 | 291.43 | 308.57 | 49 | 463 | 12.67 | 450.33 | 475.67 |
| 16 | 303 | 8.73 | 294.27 | 311.73 | 50 | 469 | 13.77 | 455.23 | 482.77 |
| 17 | 306 | 8.94 | 297.06 | 314.94 | 51 | 477 | 15.55 | 461.45 | 492.55 |
| 18 | 309 | 9.21 | 299.79 | 318.21 | 52 | 488 | 18.77 | 469.23 | 506.77 |
| 19 | 312 | 9.51 | 302.49 | 321.51 | 53 | 506 | 26.42 | 479.58 | 532.42 |
| 20 | 316 | 9.88 | 306.12 | 325.88 | 54 | 538 | 48.63 | 489.37 | 586.63 |
| 21 | 320 | 10.28 | 309.72 | 330.28 |  |  |  |  |  |
| 22 | 324 | 10.69 | 313.31 | 334.69 |  |  |  |  |  |
| 23 | 328 | 11.09 | 316.91 | 339.09 |  |  |  |  |  |
| 24 | 333 | 11.47 | 321.53 | 344.47 |  |  |  |  |  |
| 25 | 338 | 11.79 | 326.21 | 349.79 |  |  |  |  |  |
| 26 | 343 | 12.06 | 330.94 | 355.06 |  |  |  |  |  |
| 27 | 349 | 12.27 | 336.73 | 361.27 |  |  |  |  |  |
| 28 | 354 | 12.43 | 341.57 | 366.43 |  |  |  |  |  |
| 29 | 360 | 12.57 | 347.43 | 372.57 |  |  |  |  |  |
| 30 | 366 | 12.62 | 353.38 | 378.62 |  |  |  |  |  |
| 31 | 372 | 12.62 | 359.38 | 384.62 |  |  |  |  |  |
| 32 | 378 | 12.57 | 365.43 | 390.57 |  |  |  |  |  |
| 33 | 384 | 12.49 | 371.51 | 396.49 |  |  |  |  |  |

[^22]Table 3.3.3.3Gii
Equating Summary: Writ 4-5 B/C S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 2.00 (0.45) | 3 |  | 2.34 (0.48) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | 1.51 |  | 2.39 | 2.02 |  | 2.90 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | 1.95 (0.62) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | (Std. Dev.) |  |  |  |
|  | Percentage Anchors |  | Average <br> Displacement |  |  |  |
|  | 67\% |  | 0.01 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Task ID ${ }^{\text {D }}$ | Task Difficulty | y ${ }^{\text {Displacement }}$ | Task ID | Task Difficulty | Displacement |
|  | 14273 | 2.39 | -0.01 | 14272 | 1.51 | 0.02 |
|  | 14272 | 1.51 | 0.02 | 14273 | 2.39 | -0.01 |

[^23]

Table 3.3.3.3Jii
Reliability: Writ 4-5 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 98,640 | 3 | Hand-written (HW) | Keyboarded (KB) | . 901 | 1.778 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 17,654 | 94 | 6 | 0 |
|  |  | KB | 52,566 | 94 | 6 | 0 |
|  | 2 | HW | 17,510 | 95 | 5 | 0 |
|  |  | KB | 52,858 | 95 | 5 | 0 |
|  | 3 | HW | 18,104 | 95 | 5 | 0 |
|  |  | KB | 54,126 | 96 | 4 | 0 |

Table 3.3.3.3Kii
Conditional Standard Error of Measurement at Cut Scores: Writ 4-5 B/C S401 Online n/a

Table 3.3.3.3Lii
Accuracy and Consistency of Classification Indices: Writ 4-5 B/C S401 Online n/a

### 3.3.3.3iii Writing 4-5 Across Tiers

Table 3.3.3.3Aiii
Complete Task Analysis and Summary: Writ 4-5 S401 Online n/a

Table 3.3.3.3Biii
DIF Analysis and Summary: Writ 4-5 S401 Online n/a


Table 3.3.3.3Ciii
Raw Score Descriptive Statistics: Writ 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 73,241 | 0 | 47 | 20.86 | 8.58 |
| $\mathbf{5}$ | 54,047 | 0 | 49 | 22.00 | 9.43 |
| Total | 127,288 | 0 | 49 | 21.34 | 8.97 |



Table 3.3.3.3Diii
Scale Score Descriptive Statistics: Writ 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 73,241 | 144 | 452 | 324.10 | 40.45 |
| $\mathbf{5}$ | 54,047 | 155 | 463 | 332.05 | 42.39 |
| Total | 127,288 | 144 | 463 | 327.47 | 41.47 |



Table 3.3.3.3Eiii
Proficiency Level Distribution: Writ 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 4,733 | $6.46 \%$ | 3,367 | $6.23 \%$ | 8,100 | $6.36 \%$ |
| $\mathbf{2}$ | 5,761 | $7.87 \%$ | 4,316 | $7.99 \%$ | 10,077 | $7.92 \%$ |
| $\mathbf{3}$ | 46,244 | $63.14 \%$ | 31,645 | $58.55 \%$ | 77,889 | $61.19 \%$ |
| $\mathbf{4}$ | 15,792 | $21.56 \%$ | 14,220 | $26.31 \%$ | 30,012 | $23.58 \%$ |
| $\mathbf{5}$ | 676 | $0.92 \%$ | 463 | $0.86 \%$ | 1,139 | $0.89 \%$ |
| $\mathbf{6}$ | 35 | $0.05 \%$ | 36 | $0.07 \%$ | 71 | $0.06 \%$ |
| Total | 73,241 | $100.00 \%$ | 54,047 | $100.00 \%$ | 127,288 | $100.00 \%$ |

Table 3.3.3.3Fiii
Raw Score to Scale Score Conversion: Writ 4-5 S401 Online n/a

Table 3.3.3.3Giii
Equating Summary: Writ 4-5 S401 Online n/a



Table 3.3.3.3Jiii
Reliability: Writ 4-5 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| A | 28,648 | 0.881 | 0.897 |
| B/C | 98,640 | 0.901 |  |

Table 3.3.3.3Kiii
Conditional Standard Error of Measurement at Cut Scores: Writ 4-5 S401 Online

| Proficiency <br> Level |  | SEM |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Cut Score | Tier A | Tier B/C |
| $1 / 2$ | 4 | 266 | 11.81 | 11.79 |
|  | 5 | 267 | 11.55 | 11.55 |
| $2 / 3$ | 4 | 288 | 13.43 | 8.59 |
|  | 5 | 293 | 14.04 | 8.32 |
| $3 / 4$ | 4 | 351 | 17.99 | 12.35 |
|  | 5 | 356 | 17.72 | 12.35 |
| $4 / 5$ | 4 | 401 | 15.57 | 11.98 |
|  | 5 | 407 | 15.31 | 11.81 |
| $5 / 6$ | 4 | 425 | 15.57 | 11.01 |
|  | 5 | 433 | 16.65 | 10.74 |

Table 3.3.3.3Li
Accuracy and Consistency of Classification Indices: Writ (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.714 | 0.635 |  | 0.334 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.811 |  | 0.722 |  |
|  | 2 | 0.658 |  | 0.516 |  |
|  | 3 | 0.710 |  | 0.740 |  |
|  | 4 | - |  | 0.372 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.979 | 0.013 | 0.008 | 0.969 |
|  | 2/3 | 0.959 | 0.015 | 0.026 | 0.944 |
|  | 3/4 | 0.775 | 0.225 | 0.000 | 0.721 |
|  | 4/5 | 0.990 | 0.010 | 0.000 | 0.991 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.3.3Lii
Accuracy and Consistency of Classification Indices: Writ (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.748 | 0.669 |  | 0.444 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.823 |  | 0.726 |  |
|  | 2 | 0.604 |  | 0.462 |  |
|  | 3 | 0.861 |  | 0.772 |  |
|  | 4 | 0.604 |  | 0.539 |  |
|  | 5 | - |  | 0.056 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.979 | 0.011 | 0.010 | 0.969 |
|  | 2/3 | 0.954 | 0.021 | 0.025 | 0.936 |
|  | 3/4 | 0.824 | 0.046 | 0.129 | 0.767 |
|  | 4/5 | 0.991 | 0.009 | 0.000 | 0.991 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

### 3.3.3.4 Speaking 4-5

### 3.3.3.4i Speaking 4-5 Pre-A

Table 3.3.3.4Ai
Complete Task Analysis and Summary: Spek 4-5 Pre-A S401 Online n/a

Table 3.3.3.4Bi
DIF Analys is and Summary: Spek 4-5 Pre-A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 2 | 2 | 1 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | DIF | Favored | DIF | Favored |
| ( | Level | Group | Level | Group |
| 1.S45P_SL_NewStudent_401_14500 | AA | M | AA | H |
| 2.S45P_LS_ElishaOtis_401_14731 | AA | F | AA | O |
| 3.S45P_MS_GrowingPlants_401_4735 | AA | F | AA | H |

Figure 3.3.3.4Ci
Raw Scores: Spek 4-5 Pre-A S401 Online


Table 3.3.3.4Ci
Raw Score Descriptive Statistics: Spek 4-5 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 1,859 | 0 | 6 | 4.69 | 1.60 |
| $\mathbf{5}$ | 3,065 | 0 | 6 | 4.91 | 1.51 |
| Total | 4,924 | 0 | 6 | 4.83 | 1.55 |



Table 3.3.3.4Ei
Proficiency Level Distribution: Spek 4-5 Pre-A S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
|  | 1,859 | $100.00 \%$ | 3,065 | $100.00 \%$ | 4,924 | $100.00 \%$ |
| Total | 1,859 | $100.00 \%$ | 3,065 | $100.00 \%$ | 4,924 | $100.00 \%$ |

Table 3.3.3.4Fi
Raw Score to Scale Score Conversion: Spek 4-5 Pre-A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $130^{\wedge}$ | 21.64 | $100.00^{\wedge}$ | 109.64 |
| 1 | $130^{\wedge}$ | 21.64 | $100.00^{\wedge}$ | 138.64 |
| 2 | 135 | 20.47 | 114.53 | 155.47 |
| 3 | 148 | 19.30 | 128.70 | 167.30 |
| 4 | 161 | 20.47 | 140.53 | 181.47 |
| 5 | $174^{*}$ | 24.57 | 154.43 | 203.57 |
| 6 | $187^{*}$ | 31.00 | 177.00 | 239.00 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.3.4Gi
Equating Summary: Spek 4-5 Pre-A S401 Online n/a


Table 3.3.3.4Ji
Reliability: Spek 4-5 Pre-A S401 Online

| Reliability |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
|  | 4,924 | 3 | .745 |  | 0.782 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 3,608 | 96 | 4 | 0 |
|  | 2 | 4,116 | 96 | 4 | 0 |
|  | 3 | 3,870 | 95 | 5 | 0 |

Table 3.3.3.4Ki
Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 Pre-A S401 Online n/a

Table 3.3.3.4Li
Accuracy and Consistency of Classification Indices: Spek 4-5 Pre-A S401 Online n/a

### 3.3.3.4ii Speaking 4-5 A

Table 3.3.3.4Aii
Complete Task Analysis and Summary: Spek 4-5 A S401 Online n/a

Table 3.3.3.4Bii
DIF Analysis and Summary: Spek 4-5 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 5 | 3 | 3 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \\ \hline \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S45A_SL_NewS Sudent_401_14500 | AA | F | AA | O |
| 2.S45A_SL_NewStudent_401_4501 | AA | F | AA | O |
| 3.S45A_LS_Elis haotis_401_Vl_4731 | AA | F | AA | O |
| 4.S45A_LS_Elis haOtis_401_Vl_4732 | AA | M | AA | H |
| 5.S45A_MS_Gro wingP lants_401_14735 | AA | F | AA | H |
| 6.S45A_MS_Gro wingP lants_401_4736 | AA | F | AA | H |

Figure 3.3.3.4Cii
Raw Scores: Spek 4-5 A S401 Online


Table 3.3.3.4Cii
Raw Score Descriptive Statistics: Spek 4-5 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 12,552 | 0 | 18 | 10.79 | 2.23 |
| $\mathbf{5}$ | 9,791 | 0 | 18 | 11.13 | 2.08 |
| Total | 22,343 | 0 | 18 | 10.94 | 2.17 |

Figure 3.3.3.4Dii Scale Scores: Spek 4-5 A S401 Online


Table 3.3.3.4Dii
Scale Score Descriptive Statistics: Spek 4-5 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 12,552 | 124 | 423 | 267.82 | 43.33 |
| $\mathbf{5}$ | 9,791 | 130 | 423 | 274.63 | 42.27 |
| Total | 22,343 | 124 | 423 | 270.80 | 43.00 |

Figure 3.3.3.4Eii
Proficiency Level: Spek 4-5 A S401 Online


Table 3.3.3.4Eii
Proficiency Level Distribution: Spek 4-5 A S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,960 | $23.58 \%$ | 3,510 | $35.85 \%$ | 6,470 | $28.96 \%$ |
| $\mathbf{2}$ | 7,029 | $56.00 \%$ | 3,839 | $39.21 \%$ | 10,868 | $48.64 \%$ |
| $\mathbf{3}$ | 2,235 | $17.81 \%$ | 2,080 | $21.24 \%$ | 4,315 | $19.31 \%$ |
| $\mathbf{4}$ | 318 | $2.53 \%$ | 359 | $3.67 \%$ | 677 | $3.03 \%$ |
| $\mathbf{5}$ | 10 | $0.08 \%$ | 3 | $0.03 \%$ | 13 | $0.06 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 12,552 | $100.00 \%$ | 9,791 | $100.00 \%$ | 22,343 | $100.00 \%$ |

Table 3.3.3.4Fii
Raw Score to Scale Score Conversion: Spek 4-5 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $130^{\wedge}$ | 20.77 | $100.00^{\wedge}$ | 105.77 |
| 1 | $130^{\wedge}$ | 20.77 | $100.00^{\wedge}$ | 134.77 |
| 2 | 132 | 20.18 | 111.82 | 152.18 |
| 3 | 145 | 18.43 | 126.57 | 163.43 |
| 4 | 156 | 18.13 | 137.87 | 174.13 |
| 5 | 168 | 19.30 | 148.70 | 187.30 |
| 6 | 182 | 21.06 | 160.94 | 203.06 |
| 7 | 198 | 21.94 | 176.06 | 219.94 |
| 8 | 215 | 21.64 | 193.36 | 236.64 |
| 9 | 231 | 21.64 | 209.36 | 252.64 |
| 10 | 247 | 22.52 | 224.48 | 269.52 |
| 11 | 266 | 25.15 | 240.85 | 291.15 |
| 12 | 291 | 28.08 | 262.92 | 319.08 |
| 13 | 317 | 26.91 | 290.09 | 343.91 |
| 14 | 340 | 24.57 | 315.43 | 364.57 |
| 15 | 360 | 24.28 | 335.72 | 384.28 |
| 16 | 381 | 26.03 | 354.97 | 407.03 |
| 17 | $402^{*}$ | 30.42 | 378.58 | 439.42 |
| 18 | $423^{*}$ | 38.31 | 410.69 | 487.31 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.3.4Gii
Equating Summary: Spek 4-5 A S401 Online n/a



Table 3.3.3.4Jii
Reliability: Spek 4-5 A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 22,343 | 6 | .700 |  | 1.187 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 13,062 | 98 | 2 | 0 |
|  | 2 | 13,062 | 84 | 16 | 0 |
|  | 3 | 13,600 | 98 | 2 | 0 |
|  | 4 | 13,600 | 84 | 16 | 0 |
|  | 5 | 13,444 | 97 | 3 | 0 |

Table 3.3.3.4Kii
Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 A S401 Online n/a

Table 3.3.3.4Lii
Accuracy and Consistency of Classification Indices: Spek 4-5 A S401 Online n/a

### 3.3.3.4iii Speaking 4-5 B/C

Table 3.3.3.4Aiii
Complete Task Analysis and Summary: Spek 4-5 B/C S401 Online n/a

Table 3.3.3.4Biii
DIF Analysis and Summary: Spek 4-5 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 3 | 3 | 3 | 3 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.S45C_SI_NewStudent_401_14502 | AA | F | AA | O |  |
| 2.S45C_SI_NewStudent_401_14503 | AA | F | AA | O |  |
| 3.S45C_LS_ElishaOtis_401_V1_14737 | AA | M | AA | H |  |
| 4.S45C_LS_ElishaOtis_401_V1_14738 | AA | M | AA | O |  |
| 5.S45C_MS_Gro wingP lants_401_V2_14743 | AA | F | AA | H |  |
| 6.S45C_MS_Gro wingPlants_401_V2_14744 | AA | M | AA | H |  |

Figure 3.3.3.4Ciii
Raw Scores: Spek 4-5 B/C S401 Online


Table 3.3.3.4Ciii
Raw Score Descriptive Statistics: Spek 4-5 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 79,554 | 6 | 30 | 19.16 | 2.74 |
| $\mathbf{5}$ | 55,891 | 6 | 30 | 19.61 | 2.79 |
| Total | 135,445 | 6 | 30 | 19.35 | 2.77 |



Figure 3.3.3.4Eiii
Proficiency Level: Spek 4-5 B/C S401 Online

| $80.0 \%$ |
| :--- |

Table 3.3.3.4Eiii
Proficiency Level Distribution: Spek 4-5 B/C S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 640 | $0.80 \%$ | 989 | $1.77 \%$ | 1,629 | $1.20 \%$ |
| $\mathbf{2}$ | 12,268 | $15.42 \%$ | 10,832 | $19.38 \%$ | 23,100 | $17.05 \%$ |
| $\mathbf{3}$ | 41,699 | $52.42 \%$ | 30,465 | $54.51 \%$ | 72,164 | $53.28 \%$ |
| $\mathbf{4}$ | 24,245 | $30.48 \%$ | 13,260 | $23.72 \%$ | 37,505 | $27.69 \%$ |
| $\mathbf{5}$ | 679 | $0.85 \%$ | 337 | $0.60 \%$ | 1,016 | $0.75 \%$ |
| $\mathbf{6}$ | 23 | $0.03 \%$ | 8 | $0.01 \%$ | 31 | $0.02 \%$ |
| Total | 79,554 | $100.00 \%$ | 55,891 | $100.00 \%$ | 135,445 | $100.00 \%$ |

Table 3.3.3.4Fiii
Raw Score to Scale Score Conversion: Spek 4-5 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 6 | $130^{\wedge}$ | 20.47 | 156.53 | 197.47 |
| 7 | 190 | 19.60 | 170.40 | 209.60 |
| 8 | 203 | 18.72 | 184.28 | 221.72 |
| 9 | 214 | 17.84 | 196.16 | 231.84 |
| 10 | 225 | 16.96 | 208.04 | 241.96 |
| 11 | 234 | 16.67 | 217.33 | 250.67 |
| 12 | 244 | 16.38 | 227.62 | 260.38 |
| 13 | 253 | 16.38 | 236.62 | 269.38 |
| 14 | 262 | 16.67 | 245.33 | 278.67 |
| 15 | 272 | 17.26 | 254.74 | 289.26 |
| 16 | 283 | 17.84 | 265.16 | 300.84 |
| 17 | 294 | 18.72 | 275.28 | 312.72 |
| 18 | 306 | 19.30 | 286.70 | 325.30 |
| 19 | 319 | 19.01 | 299.99 | 338.01 |
| 20 | 331 | 18.72 | 312.28 | 349.72 |
| 21 | 343 | 18.13 | 324.87 | 361.13 |
| 22 | 354 | 17.84 | 336.16 | 371.84 |
| 23 | 365 | 17.55 | 347.45 | 382.55 |
| 24 | 375 | 17.55 | 357.45 | 392.55 |
| 25 | 386 | 18.13 | 367.87 | 404.13 |
| 26 | 398 | 19.01 | 378.99 | 417.01 |
| 27 | 411 | 20.47 | 390.53 | 431.47 |
| 28 | $424^{*}$ | 22.81 | 404.19 | 449.81 |
| 29 | $437 *$ | 26.32 | 425.68 | 478.32 |
| 30 | $450^{*}$ | 30.71 | 459.29 | 520.71 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.3.4Giii
Equating Summary: Spek 4-5 B/C S401 Online
n/a



Table 3.3.3.4Jiii
Reliability: Spek 4-5 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 135,445 | 6 | .726 |  | 1.448 |
|  | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 74,322 | 82 | 17 | 0 |
|  | 2 | 74,322 | 77 | 23 | 1 |
|  | 3 | 74,390 | 72 | 27 | 1 |
|  | 4 | 74,393 | 74 | 25 | 1 |
|  | 5 | 73,178 | 72 | 27 | 1 |

Table 3.3.3.4Kiii
Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 B/C S401 Online n/a

Table 3.3.3.4Liii
Accuracy and Consistency of Classification Indices: Spek 4-5 B/C S401 Online n/a

### 3.3.3.4iv Speaking 4-5 Across Tiers

Table 3.3.3.4Aiv
Complete Task Analysis and Summary: Spek 4-5 S401 Online

| Task Type |  | Average <br> Task Difficulty (in logits) |  | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | -0.30 |  | 15 | 0.66 | 0.52 |
| Name |  | Tier | Task Difficulty (in logits) | Anchored? | Fit Statistics |  |
|  |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \end{gathered}$ |  |  | Outfit <br> Mnsq |
| 1.S45A_SI_NewStudent_401_14500 |  |  | A** | -4.04 |  | 0.66 | 0.18 |
| 2.S45A_SI_NewStudent_401_14501 |  | A* | 0.75 | Yes | 0.69 | 0.72 |
| 3.S45A_LS_ElishaOtis_401_V1_14731 |  | A** | -4.15 |  | 0.60 | 0.10 |
| 4.S45A_LS_ElishaOtis_401_V1_14732 |  | A* | 1.55 | Yes | 0.71 | 0.73 |
| 5.S45A_MS_GrowingPlants_401_14735 |  | A** | -3.76 |  | 0.61 | 0.18 |
| 6.S45A_MS_GrowingPlants_401_14736 |  | A* | 0.77 | Yes | 0.73 | 0.74 |
| 7.S45C_SI_NewStudent_401_14502 |  | B/C* | 0.75 | Yes | 0.69 | 0.72 |
| 8.S45C_SI_NewStudent_401_14503 |  | B/C | 2.12 | Yes | 0.72 | 0.74 |
| 9.S45C_LS_ElishaOtis_401_V1_14737 |  | B/C* | 1.55 | Yes | 0.71 | 0.73 |
| 10.S45C_LS_ElishaOtis_401_V1_14738 |  | B/C | 2.68 | Yes | 0.66 | 0.68 |
| 11.S45C_MS_GrowingPlants_401_V2_14743 |  | B/C* | 0.77 | Yes | 0.73 | 0.74 |
| 12.S45C_MS_GrowingPlants_401_V2_14744 |  | B/C | 1.35 | Yes | 0.58 | 0.57 |
| 13.S45P_SI_NewStudent_401_14500 |  | Pre-A** | -4.04 |  | 0.66 | 0.18 |
| 14.S45P_LS_ElishaOtis_401_14731 |  | Pre-A** | -4.15 |  | 0.60 | 0.10 |
| 15.S45P_MS_GrowingPlants_401_14735 |  | Pre-A** | -3.76 |  | 0.61 | 0.18 |
| Raw Score <br> Distribution by Task | Task | Raw Score |  |  |  |  |
|  |  | 0 | 1 | 2 | 3 | 4 |
|  | Task 1 | 2.15\% | 8.09\% | 89.76\% | N/A | N/A |
|  | Task 2 | 0.46\% | 12.56\% | 41.56\% | 43.09\% | 2.33\% |
|  | Task 3 | 3.33\% | 7.53\% | 89.14\% | N/A | N/A |
|  | Task 4 | 0.73\% | 20.53\% | 47.00\% | 28.38\% | 3.36\% |
|  | Task 5 | 2.56\% | 8.81\% | 88.63\% | N/A | N/A |
|  | Task 6 | 0.39\% | 8.76\% | 46.97\% | 37.52\% | 6.36\% |
|  | Task 7 | 0.46\% | 12.56\% | 41.56\% | 43.09\% | 2.33\% |
|  | Task 8 | 0.40\% | 21.57\% | 53.88\% | 21.70\% | 2.45\% |
|  | Task 9 | 0.73\% | 20.53\% | 47.00\% | 28.38\% | 3.36\% |
|  | Task 10 | 0.42\% | 27.13\% | 53.86\% | 16.38\% | 2.21\% |
|  | Task 11 | 0.39\% | 8.76\% | 46.97\% | 37.52\% | 6.36\% |
|  | Task 12 | 0.20\% | 7.12\% | 65.61\% | 25.20\% | 1.87\% |
|  | Task 13 | 2.15\% | 8.09\% | 89.76\% | N/A | N/A |
|  | Task 14 | 3.33\% | 7.53\% | 89.14\% | N/A | N/A |
|  | Task 15 | 2.56\% | 8.81\% | 88.63\% | N/A | N/A |

** This task is shared between Pre-A and A.

* This task is shared between A and B/C.

Table 3.3.3.4Biv
DIF Analysis and Summary: Spek 4-5 S401 Online n/a

Table 3.3.3.4Civ
Raw Score Descriptive Statistics: Spek 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 93,965 | 0 | 30 | 17.76 | 4.31 |
| $\mathbf{5}$ | 68,747 | 0 | 30 | 17.75 | 4.84 |
| Total | 162,712 | 0 | 30 | 17.75 | 4.54 |

Figure 3.3.3.4Div
Scale Scores: Spek 4-5 S401 Online


Table 3.3.3.4Div
Scale Score Descriptive Statistics: Spek 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 93,965 | 124 | 450 | 310.59 | 42.74 |
| $\mathbf{5}$ | 68,747 | 130 | 450 | 311.68 | 48.21 |
| Total | 162,712 | 124 | 450 | 311.05 | 45.13 |



Table 3.3.3.4Eiv
Proficiency Level Distribution: Spek 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 5,459 | $5.81 \%$ | 7,564 | $11.00 \%$ | 13,023 | $8.00 \%$ |
| $\mathbf{2}$ | 19,297 | $20.54 \%$ | 14,671 | $21.34 \%$ | 33,968 | $20.88 \%$ |
| $\mathbf{3}$ | 43,934 | $46.76 \%$ | 32,545 | $47.34 \%$ | 76,479 | $47.00 \%$ |
| $\mathbf{4}$ | 24,563 | $26.14 \%$ | 13,619 | $19.81 \%$ | 38,182 | $23.47 \%$ |
| $\mathbf{5}$ | 689 | $0.73 \%$ | 340 | $0.49 \%$ | 1,029 | $0.63 \%$ |
| $\mathbf{6}$ | 23 | $0.02 \%$ | 8 | $0.01 \%$ | 31 | $0.02 \%$ |
| Total | 93,965 | $100.00 \%$ | 68,747 | $100.00 \%$ | 162,712 | $100.00 \%$ |

Table 3.3.3.4Fiv
Raw Score to Scale Score Conversion: Spek 4-5 S401 Online n/a

Table 3.3.3.4Giv
Equating Summary: Spek 4-5 S401 Online


[^24]


Table 3.3.3.4Jiv
Reliability: Spek 4-5 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| Pre-A | 4,924 | 0.745 | 0.723 |
| A | 22,343 | 0.700 |  |
| B/C | 135,445 | 0.726 |  |

Table 3.3.3.4Kiv
Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 S401 Online

| Proficiency Level | Grade | Cut Score | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Tier A | Tier B/C |
| 1/2 | 4 | 246 | 22.52 | 16.38 |
|  | 5 | 258 | 23.98 | 16.67 |
| 2/3 | 4 | 293 | 28.08 | 18.72 |
|  | 5 | 302 | 28.08 | 19.01 |
| 3/4 | 4 | 342 | 24.57 | 18.13 |
|  | 5 | 350 | 24.28 | 17.84 |
| 4/5 | 4 | 397 | 29.25 | 19.01 |
|  | 5 | 407 | 31.88 | 20.18 |
| 5/6 | 4 | 435 | 45.04 | 25.74 |
|  | 5 | 443 | 50.60 | 28.08 |

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficeny level higher than 2.

Table 3.3.3.4Li
Accuracy and Consistency of Classification Indices: Spek (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.582 | 0.503 |  | 0.247 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.733 |  | 0.561 |  |
|  | 2 | 0.679 |  | 0.493 |  |
|  | 3 | 0.552 |  | 0.542 |  |
|  | 4 | - |  | 0.416 |  |
|  | 5 | - |  | 0.015 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.968 | 0.015 | 0.017 | 0.950 |
|  | $2 / 3$ | 0.881 | 0.035 | 0.084 | 0.829 |
|  | 3/4 | 0.731 | 0.269 | 0.000 | 0.697 |
|  | 4/5 | 0.992 | 0.008 | 0.000 | 0.991 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.3.4Lii
Accuracy and Consistency of Classification Indices: Spek (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.605 | 0.484 |  | 0.236 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.816 |  | 0.678 |  |
|  | 2 | 0.586 |  | 0.373 |  |
|  | 3 | 0.579 |  | 0.560 |  |
|  | 4 | - |  | 0.292 |  |
|  | 5 | - |  | 0.008 |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.955 | 0.019 | 0.026 | 0.932 |
|  | 2/3 | 0.848 | 0.036 | 0.117 | 0.775 |
|  | 3/4 | 0.797 | 0.203 | 0.000 | 0.717 |
|  | 4/5 | 0.995 | 0.005 | 0.000 | 0.994 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.4 Grades: 6-8

### 3.3.4.1 Listening 6-8

Table 3.3.4.1A
Complete Item Analysis and Summary: List 6-8 S401 Online
$\left.\begin{array}{|l|c|c|c|c|c|}\hline & \begin{array}{c}\text { Average } \\ \text { Item } \\ \text { Difficulty } \\ \text { (in logits) }\end{array} & \text { No. of Items } & \begin{array}{c}\text { Average } \\ \text { Infit } \\ \text { P-value }\end{array} & \begin{array}{c}\text { Average } \\ \text { Outfit } \\ \text { Mean } \\ \text { Square }\end{array} \\ \text { Square }\end{array}\right]$

| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 26.L68B_SS_Types OfGo vernment_P 100_alt1_Screen_3_13830 | 0.60 |  | 0.62 | 0.93 | 0.90 |
| 27.L68B_SS_Types OfGo vernment_P 100_altı_Screen_4_13831 | 1.94 | Yes | 0.37 | 1.01 | 1.01 |
| 28.L68B_SC_Lea aClas sific atio n_P 100_Screen_2_11467 | 0.84 | Yes | 0.55 | 0.98 | 0.97 |
| 29.L68B_SC_LeafClas sification_P 100_Screen_3_11472 | 1.61 | Yes | 0.39 | 0.97 | 0.96 |
| 30.L68B_SC_Lea fClas sification_P 100_Screen_4_11474 | 1.71 | Yes | 0.45 | 1.02 | 1.02 |
| 31.L68B_LA_WritingAboutReading_P 100_Screen_2_12526 | -0.95 | Yes | 0.89 | 0.92 | 0.77 |
| 32.L68B_LA_WritingAboutReading_P 100_Screen_3_12528 | 0.23 | Yes | 0.67 | 0.91 | 0.87 |
| 33.L68B_LA_WritingAboutReading_P 100_Screen_4_12529 | 1.42 | Yes | 0.40 | 0.96 | 0.96 |
| 34.L68B_MA_SchoolGarden_P 100_Screen_2_12481 | -0.65 | Yes | 0.82 | 0.98 | 0.95 |
| 35.L68B_MA_SchoolGarden_P 100_Screen_3_12482 | -0.38 |  | 0.80 | 0.96 | 0.92 |
| 36.L68B_MA_SchoolGarden_P 100_Screen_4_12483 | 1.17 | Yes | 0.49 | 1.02 | 1.02 |
| 37.L68C_LA_MaryAnders on_P 100_Screen_2_12717 | 0.37 |  | 0.88 | 0.90 | 0.70 |
| 38.L68C_LA_MaryAnders on_P 100_Screen_3_12718 | 3.36 |  | 0.41 | 1.06 | 1.12 |
| 39.L68C_LA_MaryAnders on_P 100_Screen_4_12719 | 2.49 |  | 0.57 | 0.94 | 0.93 |
| 40.L68C_MA_Class Pres identSurve y_P 100_Screen_2_13042 | 0.58 |  | 0.86 | 0.94 | 0.90 |
| 41.L68C_MA_Clas sPres identSurve y_P 100_Screen_3_13043 | 0.59 |  | 0.90 | 0.99 | 1.00 |
| 42.L68C_MA_Clas sPres identSurve y_P 100_Screen_4_13044 | 2.44 |  | 0.62 | 0.97 | 0.96 |
| 43.L68C_SS_ThailandMap_P 100_A301FT_Screen_2_13838 | 2.51 |  | 0.65 | 1.04 | 1.05 |
| 44.L68C_SS_ThailandMap_P 100_A301FT_Screen_3_13839 | 2.15 |  | 0.71 | 1.03 | 1.09 |
| 45.L68C_SS_ThailandMap_P 100_A301FT_Screen_4_13840 | 2.02 |  | 0.72 | 0.94 | 0.91 |
| 46.L68C_SC_Nature Trail_P 100_Screen_2_12899 | 2.09 |  | 0.73 | 0.96 | 0.96 |
| 47.L68C_SC_Nature Trail_P 100_Screen_3_12901 | 2.25 |  | 0.70 | 0.94 | 0.92 |
| 48.L68C_SC_Nature Trail_P 100_Screen_4_12905 | 2.50 |  | 0.66 | 0.94 | 0.92 |
| 49.L68C_LA_WritingAboutReading_P 100_Screen_2_12531 | 2.46 |  | 0.68 | 0.94 | 0.92 |
| 50.L68C_LA_WritingAboutReading_P 100_Screen_3_12532 | 3.95 |  | 0.38 | 0.97 | 0.96 |
| 51.L68C_LA_WritingAboutReading_P 100_Screen_4_12533 | 1.47 |  | 0.83 | 0.94 | 0.89 |
| 52.L68C_MA_SchoolGarden_P 100_Screen_2_12484 | 0.33 | Yes | 0.92 | 0.97 | 0.87 |
| 53.L68C_MA_SchoolGarden_P 100_Screen_3_12485 | 3.36 |  | 0.56 | 1.03 | 1.04 |
| 54.L68C_MA_SchoolGarden_P 100_Screen_4_12486 | 2.93 |  | 0.61 | 0.95 | 0.94 |

Table 3.3.4.1B
DIF Analysis and Summary: List 6-8 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| A | 27 | 26 | 28 | 25 |
| B | 1 | 0 | 0 | 1 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.L68A_SI_TalentS ho w_P 100_Screen_2_12454 | A | M | A | H |
| 2.L68A_SI_TalentShow_P 100_Screen_3_12455 | A | M | A | O |
| 3.L68A_SI_TalentShow_P 100_Screen_4_12456 | A | M | A | O |
| 4.L68B_SI_Autho rP roject_P 100_Screen_2_12380 | A | M | A | O |
| 5.L68B_SI_AuthorP roject_P 100_Screen_3_12381 | A | M | A | H |
| 6.L68B_SI_AuthorP ro ject_P 100_Screen_4_12382 | A | M | A | H |
| 7.L68A_LA_RemoteContro I_P 100_Screen_2_12849 | A | F | A | H |
| 8.L68A_LA_RemoteControl_P 100_Screen_3_12850 | A | F | A | H |
| 9.L68A_LA_RemoteContro L_P 100_Screen_4_12922 | A | F | B | O |
| 10.L68A_MA_Clas sPres identSurvey_P 100_Screen_2_13033 | A | M | A | O |
| 11.L68A_MA_Clas sPres identSurvey_P 100_Screen_3_13034 | A | M | A | O |
| 12.L68A_MA_Clas sPres identSurvey_P 100_Screen_4_13035 | A | M | A | H |
| 13.L68A_SS_Types OfGo vernment_P 100_Screen_2_13060 | A | M | A | H |
| 14.L68A_SS_Types OfGo vernment_P 100_Screen_3_13061 | A | M | A | H |
| 15.L68A_SS_Types OfGo vernment_P 100_Screen_4_13062 | A | F | A | H |
| 16.L68A_SC_Nature Trail_P 100_Screen_2_12884 | A | F | A | O |
| 17.L68A_SC_NatureTrail_P 100_Screen_3_12885 | A | F | A | H |
| 18.L68A_SC_NatureTrail_P 100_Screen_4_12886 | A | F | A | O |
| 19.L68B_LA_MaryAnders on_P 100_Screen_2_12687 | A | F | A | O |
| 20.L68B_LA_MaryAnders on_P 100_Screen_3_12688 | A | F | A | O |
| 21.L68B_LA_MaryAnders on_P 100_Screen_4_12689 | A | F | A | O |
| 22.L68B_MA_Clas sPresidentSurvey_P 100_Screen_2_13039 | A | M | A | H |
| 23.L68B_MA_Clas sPres identSurvey_P 100_Screen_3_13040 | A | M | A | O |
| 24.L68B_MA_Clas sPres identSurvey_P 100_Screen_4_13041 | A | M | A | H |
| 25.L68B_SS_Types OfGo vernment_P 100_altıScreen_2_13829 | A | M | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 26.L68B_SS_Types OfGo vernment_P 100_alt1_Screen_3_13830 | A | F | A | O |
| 27.L68B_SS_Types OfGo vernment_P 100_altıScreen_4_13831 | A | F | A | O |
| 28.L68B_SC_LeafClas sification_P 100_Screen_2_11467 | A | M | A | O |
| 29.L68B_SC_LeafClas sificatio n_P 100_Screen_3_1472 | A | F | A | H |
| 30.L68B_SC_Lea aClas sification_P 100_Screen_4_11474 | A | F | A | O |
| 31.L68B_LA_WritingAboutReading_P 100_Screen_2_12526 | A | M | A | H |
| 32.L68B_LA_WritingAboutReading_P 100_Screen_3_12528 | A | F | A | O |
| 33.L68B_LA_WritingAboutReading_P 100_Screen_4_12529 | A | M | A | H |
| 34.L68B_MA_SchoolGarden_P 100_Screen_2_12481 | A | M | A | H |
| 35.L68B_MA_SchoolGarden_P 100_Screen_3_12482 | A | M | A | H |
| 36.L68B_MA_SchoolGarden_P 100_Screen_4_12483 | A | F | A | H |
| 37.L68C_LA_MaryAnders on_P 100_Screen_2_12717 | A | F | A | O |
| 38.L68C_LA_MaryAnders on_P 100_Screen_3_127 18 | A | M | A | H |
| 39.L68C_LA_MaryAnders on_P 100_Screen_4_127 19 | A | F | A | O |
| 40.L68C_MA_Clas sPres identSurvey_P 100_Screen_2_13042 | A | M | A | O |
| 41.L68C_MA_Clas sP res identS urvey_P 100_Screen_3_13043 | A | M | A | H |
| 42.L68C_MA_Clas sPres identSurvey_P 100_Screen_4_13044 | A | M | A | O |
| 43.L68C_SS_ThailandMap_P 100_A301FT_Screen_2_13838 | A | F | A | H |
| 44.L68C_SS_ThailandMap_P 100_A301FT_Screen_3_13839 | A | F | A | H |
| 45.L68C_SS_ThailandMap_P 100_A301FT_Screen_4_13840 | A | F | A | O |
| 46.L68C_SC_Nature Trail_P 100_Screen_2_12899 | A | F | A | O |
| 47.L68C_SC_Nature Trail_P 100_Screen_3_12901 | A | F | A | H |
| 48.L68C_SC_Nature Trail_P 100_Screen_4_12905 | A | F | A | H |
| 49.L68C_LA_WritingAboutReading_P 100_Screen_2_12531 | A | F | A | O |
| 50.L68C_LA_WritingAboutReading_P 100_Screen_3_12532 | A | F | A | H |
| 51.L68C_LA_WritingAboutReading_P 100_Screen_4_12533 | B | M | A | O |
| 52.L68C_MA_SchoolGarden_P 100_Screen_2_12484 | A | M | A | H |
| 53.L68C_MA_SchoolGarden_P 100_Screen_3_12485 | A | M | A | H |
| 54.L68C_MA_SchoolGarden_P 100_Screen_4_12486 | A | M | A | H |

Figure 3.3.4.1C
Raw Scores: List 6-8 S401 Online
n/a

Table 3.3.4.1C
Raw Score Descriptive Statistics: List 6-8 S401 Online n/a



Table 3.3.4.1D
Scale Score Descriptive Statistics: List 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 59,759 | 124 | 515 | 382.49 | 50.81 |
| $\mathbf{7}$ | 60,913 | 128 | 515 | 393.56 | 56.64 |
| $\mathbf{8}$ | 60,312 | 132 | 515 | 400.37 | 61.06 |
| Total | 180,984 | 124 | 515 | 392.17 | 56.82 |

Table 3.3.4.1E
Proficiency Level Distribution: List 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,823 | $4.72 \%$ | 3,905 | $6.41 \%$ | 5,043 | $8.36 \%$ | 11,771 | $6.50 \%$ |
| $\mathbf{2}$ | 4,700 | $7.86 \%$ | 4,968 | $8.16 \%$ | 5,623 | $9.32 \%$ | 15,291 | $8.45 \%$ |
| $\mathbf{3}$ | 15,178 | $25.40 \%$ | 13,354 | $21.92 \%$ | 12,325 | $20.44 \%$ | 40,857 | $22.57 \%$ |
| $\mathbf{4}$ | 9,057 | $15.16 \%$ | 9,080 | $14.91 \%$ | 7,310 | $12.12 \%$ | 25,447 | $14.06 \%$ |
| $\mathbf{5}$ | 9,310 | $15.58 \%$ | 7,829 | $12.85 \%$ | 4,953 | $8.21 \%$ | 22,092 | $12.21 \%$ |
| $\mathbf{6}$ | 18,691 | $31.28 \%$ | 21,777 | $35.75 \%$ | 25,058 | $41.55 \%$ | 65,526 | $36.21 \%$ |
| Total | 59,759 | $100.00 \%$ | 60,913 | $100.00 \%$ | 60,312 | $100.00 \%$ | 180,984 | $100.00 \%$ |

Table 3.3.4.1F
Raw Score to Scale Score Conversion: List 6-8 S401 Online n/a

Table 3.3.4.1G
Equating Summary: List 6-8 S401 Online


[^25]Figure 3.3.4.1H
Test Characteristic Curve: List 6-8 S401 Online n/a

Figure 3.3.4.1I


Ability Measure

Table 3.3.4.1J
Reliability: List 6-8 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 180,984 | 54 | .87 |

Table 3.3.4.1K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 6-8 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 6 | 294 | 2 | 17.35 | 17.86 | 17.60 | 0.36 |
|  | 7 | 302 | 12 | 16.84 | 21.43 | 18.15 | 1.99 |
|  | 8 | 308 | 24 | 17.86 | 17.86 | 17.86 | 0.00 |
| 2/3 | 6 | 332 | 13 | 17.35 | 17.35 | 17.35 | 0.00 |
|  | 7 | 340 | 10 | 17.86 | 17.86 | 17.86 | 0.00 |
|  | 8 | 347 | 1 | 17.35 | 17.35 | 17.35 | 0.00 |
| 3/4 | 6 | 363 | 3 | 17.86 | 17.86 | 17.86 | 0.00 |
|  | 7 | 370 | 52 | 17.86 | 17.86 | 17.86 | 0.00 |
|  | 8 | 377 | 74 | 17.35 | 17.86 | 17.37 | 0.12 |
| 4/5 | 6 | 385 | 145 | 17.35 | 18.88 | 17.87 | 0.19 |
|  | 7 | 394 | 43 | 17.86 | 18.88 | 18.20 | 0.33 |
|  | 8 | 402 | 653 | 17.86 | 19.90 | 19.55 | 0.61 |
| 5/6 | 6 | 411 | N/A | N/A | N/A | N/A | N/A |
|  | 7 | 420 | 5 | 19.90 | 19.90 | 19.90 | 0.00 |
|  | 8 | 427 | 2,681 | 18.37 | 19.39 | 18.37 | 0.06 |

Table 3.3.4.1 Li
Accuracy and Consistency of Classification Indices: List (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.611 | 0.514 |  | 0.387 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.809 |  | 0.581 |  |
|  | 2 | 0.413 |  | 0.298 |  |
|  | 3 | 0.655 |  | 0.543 |  |
|  | 4 | 0.382 |  | 0.296 |  |
|  | 5 | 0.437 |  | 0.328 |  |
|  | 6 | 0.879 |  | 0.801 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.970 | 0.005 | 0.024 | 0.959 |
|  | 2/3 | 0.928 | 0.043 | 0.029 | 0.893 |
|  | 3/4 | 0.875 | 0.049 | 0.076 | 0.833 |
|  | 4/5 | 0.894 | 0.050 | 0.056 | 0.846 |
|  | 5/6 | 0.908 | 0.056 | 0.035 | 0.870 |

Table 3.3.4.1 Lii
Accuracy and Consistency of Classification Indices: List (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.609 | 0.521 |  | 0.389 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.793 |  | 0.578 |  |
|  | 2 | 0.376 |  | 0.275 |  |
|  | 3 | 0.596 |  | 0.484 |  |
|  | 4 | 0.388 |  | 0.301 |  |
|  | 5 | 0.374 |  | 0.276 |  |
|  | 6 | 0.895 |  | 0.826 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.961 | 0.009 | 0.031 | 0.945 |
|  | 2/3 | 0.921 | 0.048 | 0.031 | 0.885 |
|  | 3/4 | 0.879 | 0.049 | 0.072 | 0.839 |
|  | 4/5 | 0.897 | 0.047 | 0.057 | 0.850 |
|  | 5/6 | 0.908 | 0.057 | 0.035 | 0.869 |

Table 3.3.4.1 Liii
Accuracy and Consistency of Classification Indices: List (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.627 | 0.546 |  | 0.408 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.784 |  | 0.594 |  |
|  | 2 | 0.380 |  | 0.287 |  |
|  | 3 | 0.580 |  | 0.471 |  |
|  | 4 | 0.355 |  | 0.268 |  |
|  | 5 | 0.267 |  | 0.189 |  |
|  | 6 | 0.927 |  | 0.870 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.951 | 0.013 | 0.036 | 0.931 |
|  | 2/3 | 0.913 | 0.051 | 0.036 | 0.878 |
|  | 3/4 | 0.888 | 0.042 | 0.071 | 0.851 |
|  | 4/5 | 0.912 | 0.038 | 0.050 | 0.867 |
|  | 5/6 | 0.909 | 0.063 | 0.028 | 0.871 |

### 3.3.4.2 Reading 6-8

Table 3.3.4.2A
Complete Item Analysis and Summary: Read 6-8 S401 Online

|  | Average <br> Item <br> Difficulty <br> (in logits) | No. of Items | Average <br> Infit <br> Merage <br> P-value | Average <br> Outfit <br> Mean <br> Square |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Square |  |  |  |  |$|$| Item Type |
| :--- |
| Selected Response |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | Outfit <br> Mnsq |
| 26.R68C_LA_MyPenPal_P 100_A202_Screen_3_13660 | 2.23 | Yes | 0.42 | 1.04 | 1.07 |
| 27.R68C_LA_MyP enP al_P 100_A202_Screen_4_13661 | 1.22 | Yes | 0.59 | 1.10 | 1.21 |
| 28.R68C_MA_Ratios_401_V1_Screen_2_13962 | 1.39 |  | 0.44 | 0.97 | 0.96 |
| 29.R68C_MA_Ratio s_401_V1_Screen_3_13963 | 1.94 |  | 0.35 | 1.02 | 1.03 |
| 30.R68C_MA_Ratios_401_V1_Screen_4_13964 | 2.35 |  | 0.28 | 1.07 | 1.13 |
| 31.R688__SS_RockAndRoll_203_P 100_A 301 SScreen_2_13584 | 1.26 |  | 0.44 | 1.03 | 1.02 |
| 32.R68B_SS_RockAndRoll_203_P 100_A 301_Screen_3_13585 | 0.57 | Yes | 0.62 | 0.90 | 0.87 |
| 33.R68B_SS_RockAndRoll_203_P 100_A 301_Screen_4_13586 | 0.61 | Yes | 0.57 | 0.96 | 0.94 |
| 34.R68B_SC_CellDivis io n_diab_P 100_A 301_Screen_2_13587 | 0.64 | Yes | 0.69 | 0.94 | 0.92 |
| 35.R68B_SC_CellDivis io n_diab_P 100_A 301_Screen_3_13588 | 0.78 |  | 0.60 | 0.95 | 0.93 |
| 36.R68B_SC_CellDivis io n_diab_P 100_A301_Screen_4_13589 | 1.93 | Yes | 0.35 | 1.02 | 1.02 |
| 37.R68C_LA_SolarOven_203_P 100_A203_Screen_2_13602 | 1.09 | Yes | 0.48 | 0.96 | 0.95 |
| 38.R68C_LA_So larOven_203_P 100_A203_Screen_3_13603 | 1.01 | Yes | 0.51 | 0.92 | 0.91 |
| 39.R68C_LA_SolarOven_203_P 100_A203_Screen_4_13604 | 1.87 | Yes | 0.37 | 0.92 | 0.90 |
| 40.R68B_MA_ne wro o m_diab_P 100_A203_alt1_Screen_2_13650 | 1.09 | Yes | 0.53 | 0.95 | 0.94 |
| 41.R68B_MA_ne wroom_diab_P 100_A203_alt I_Screen_3_13651 | 2.53 |  | 0.22 | 0.99 | 0.98 |
| 42.R68B_MA_ne wroom_diab_P 100_A 203 _alt | 2.41 | Yes | 0.25 | 1.06 | 1.10 |
| 43.R68B_SS_Greathventions_203_P 100_A301_Screen_2_13596 | 0.44 | Yes | 0.69 | 0.91 | 0.88 |
| 44.R68B_SS_Greatliventions_203_P 100_A301_Screen_3_13597 | 1.01 | Yes | 0.54 | 0.89 | 0.88 |
| 45.R68B_SS_Greathventions_203_P 100_A301_Screen_4_13598 | 1.80 | Yes | 0.33 | 0.95 | 0.94 |
| 46.R68C_SC_Butterflies_mabr_P 100_A301_Screen_2_13611 | 1.15 |  | 0.50 | 1.02 | 1.02 |
| 47.R68C_SC_Butterflies_mabr_P 100_A301_Screen_3_13612 | 1.23 | Yes | 0.48 | 0.99 | 0.99 |
| 48.R68C_SC_Butterflies_mabr_P 100_A301_Screen_4_13613 | 1.15 | Yes | 0.51 | 0.96 | 0.96 |
| 49.R68C_MA_Ratio s_401_V2_Screen_2_14616 | 2.52 |  | 0.43 | 0.96 | 0.94 |
| 50.R68C_MA_Ratios_401_V2_Screen_3_14617 | 3.34 |  | 0.25 | 1.01 | 1.04 |
| $51 . \mathrm{R} 68 \mathrm{C}$ _MA_Ratio s_401_V2_Screen_4_14618 | 2.44 |  | 0.41 | 0.98 | 0.98 |
| 52.R68C_SS_Nile River_401_V1_Screen_2_14491 | 1.72 |  | 0.67 | 0.94 | 0.92 |
| 53.R68C_SS_NileRiver_401_V1_Screen_3_14492 | 1.78 |  | 0.66 | 1.01 | 1.06 |
| 54.R68C_SS_Nile River_401_V1_Screen_4_14493 | 2.61 |  | 0.47 | 0.98 | 0.97 |
| $55 . \mathrm{R} 68 \mathrm{~B}$ _SC_RockCycle_DeRoSaJ a_P 100_A301_altıScreen_2_13656 | 2.35 |  | 0.59 | 0.97 | 0.96 |
| 56.R68B_SC_RockCycle_DeRoSaJa_P 100_A301_alt_Screen_3_13657 | 1.62 | Yes | 0.74 | 0.99 | 1.02 |
| 57.R68B_SC_RockCycle_DeRoSaJa_P 100_A301_altıScreen_4_13658 | 3.59 |  | 0.32 | 1.06 | 1.08 |
| 58.R68C_LA_Prodigy_203_P 100_A301_Screen_2_13614 | 1.95 | Yes | 0.65 | 0.85 | 0.80 |
| 59.R68C_LA_Prodigy_203_P 100_A301_Screen_3_13615 | 2.21 | Yes | 0.60 | 0.86 | 0.82 |
| 60.R68C_LA_Prodigy_203_P 100_A301_Screen_4_13616 | 2.67 |  | 0.52 | 0.92 | 0.91 |


| Name | ItemDifficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | Outfit <br> Mnsq |
| 61.R68C_MA_FoodCollection_RaObAmSc_P 100_A203_Screen_2_13617 | 3.33 |  | 0.41 | 1.04 | 1.04 |
| 62.R68C_MA_FoodCollection_RaObAmSc_P 100_A203_Screen_3_13618 | 3.13 |  | 0.45 | 0.98 | 0.97 |
| 63.R68C_MA_FoodCo llection_RaObAmSc_P 100_A203_Screen_4_13619 | 3.03 | Yes | 0.42 | 1.00 | 0.99 |
| 64.R68C_SS_StudentCouncil_203_P P 100_A 301_Screen_2_13620 | 1.82 | Yes | 0.71 | 0.89 | 0.83 |
| 65.R68C_SS_StudentCouncil_203_P 100_A ${ }^{\text {a }}$ L_S Sree __3_13621 | 3.33 |  | 0.43 | 0.98 | 0.98 |
| 66.R68C_SS_StudentCouncil_203_P 100_A 301_Screen_4_13622 | 3.78 |  | 0.32 | 1.01 | 1.00 |
| 67.R68C_SC_Vaporization_RaObAmSc_P 100_A 301_Screen_2_B623 | 2.40 |  | 0.62 | 0.97 | 0.98 |
| 68.R68C_SC_Vaporization_RaObAmSc_P P100_A301_Screen_3_13624 | 3.54 |  | 0.37 | 0.97 | 0.96 |
| 69.R68C_SC_Vaporization_RaObAmSc_P P100_A301_Screen_-_13625 | 3.31 |  | 0.42 | 0.98 | 0.98 |

Table 3.3.4.2B
DIF Analysis and Summary: Read 6-8 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (0) |
| A | 41 | 28 | 38 | 31 |
| B | 0 | 0 | 0 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.R68A_SI_ProperConduct_P 100_A203_Screen_2_13554 | A | M | A | H |
| 2.R68A_SI_ProperConduct_P 100_A203_Screen_3_13555 | A | M | A | H |
| 3.R68A_SI_ProperConduct_P 100_A203_Screen_4_13556 | A | M | A | H |
| 4.R68B_SI_Sports Equipment_301_P 100_A301FT_Screen_2_13629 | A | M | A | O |
| 5.R68B_SI_SportsEquipment_301_P 100_A301FT_Screen_3_13630 | A | M | A | O |
| 6.R68B_SI_Sports Equipment_301_P 100_A301FT_Screen_4_13631 | A | F | A | H |
| 7.R68A_LA_ScavengerHunt_401_V2_Screen_2_14640 | A | M | A | O |
| 8.R68A_LA_ScavengerHunt_401_V2_Screen_3_14641 | A | M | A | O |
| 9.R68A_LA_ScavengerHunt_401_V2_Screen_4_14642 | A | M | A | H |
| 10.R68A_MA_Clo set_rize_P 100_A203_Screen_2_13563 | A | F | A | H |
| 11.R68A_MA_Closet_rize_P 100_A203_Screen_3_13564 | A | M | A | H |
| 12.R68A_MA_Closet_rize_P 100_A203_Screen_4_13565 | A | M | A | O |
| 13.R68A_SS_Uruguay_203_P 100_A301_Screen_2_13566 | A | F | A | H |
| 14.R68A_SS_Uruguay_203_P 100_A301_Screen_3_13567 | A | M | A | O |
| 15.R68A_SS_Uruguay_203_P 100_A 301 _Screen_4_13568 | A | M | A | O |
| 16.R68A_SC_Ho wP lants MakeFood_mabr_P 100_A 301_Screen_2_13569 | A | M | A | H |
| 17.R68A_SC_Ho wP lants MakeFo od_mabr_P 100_A301_Screen_3_13570 | A | M | A | O |
| 18.R68A_SC_Ho wP lants MakeFoo d_m abr_P 100_A301_Screen_4_13571 | A | F | A | O |
| 19.R68A_LA_Collage_203_P 100_A301_Screen_2_13572 | A | M | A | O |
| 20.R68A_LA_Co llage_203_P 100_A301_Screen_3_13573 | A | M | A | O |
| 21.R68A_LA_Collage_203_P 100_A301_Screen_4_13574 | A | M | A | H |
| 22.R68A_MA_soccerfield_DeRoSaJ a_P 100_A203_Screen_2_13575 | A | F | A | H |
| 23.R68A_MA_soccerfield_DeRo SaJ a_P 100_A203_Screen_3_13576 | A | M | A | O |
| 24.R68A_MA_soccerfield_DeRoSaJ a_P 100_A203_Screen_4_13577 | A | F | A | O |
| 25.R68C_LA_MyPenPal_P 100_A202_Screen_2_13659 | A | M | A | H |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | DIF Level | Favored Group |  | Favored Group |
| 26.R68C_LA_MyPenP al_P 100_A202_Screen_3_13660 | A | M | A | H |
| 27.R68C_LA_MyPenP al_P 100_A202_Screen_4_13661 | A | M | A | H |
| 28.R68C_MA_Ratios_401_V1_Screen_2_13962 | A | F | A | H |
| 29.R68C_MA_Ratios_401_V1_Screen_3_13963 | A | F | A | H |
| 30.R68C_MA_Ratio s_401_V1_Screen_4_13964 | A | M | A | H |
| 31.R68B_SS_RockAndRoll_203_P 100_A301_Screen_2_13584 | A | M | A | H |
| 32.R68B_SS_RockAndRoll_203_P 100_A301_Screen_3_13585 | A | F | A | H |
| 33.R68B_SS_RockAndRo ll_203_P 100_A301_Screen_4_13586 | A | M | A | O |
| 34.R68B_SC_Celldivis ion_diab_P 100_A301_Screen_2_13587 | A | M | A | H |
| 35.R68B_SC_CellDivision_diab_P 100_A301_Screen_3_13588 | A | M | A | H |
| 36.R68B_SC_CellDivision_diab_P 100_A301_Screen_4_13589 | A | M | A | H |
| 37.R68C_LA_SolarOven_203_P 100_A203_Screen_2_13602 | A | M | A | H |
| 38.R68C_LA_SolarOven_203_P 100_A203_Screen_3_13603 | A | M | A | O |
| 39.R68C_LA_SolarOven_203_P 100_A203_Screen_4_13604 | A | F | A | H |
| 40.R68B_MA_ne wroom_diab_P 100_A203_alt 1_Screen_2_13650 | A | M | A | O |
| 41.R68B_MA_newroom_diab_P 100_A203_alt1_Screen_3_13651 | A | M | A | O |
| 42.R68B_MA_ne wroom_diab_P 100_A203_alt 1_Screen_4_13652 | A | F | A | H |
| 43.R68B_SS_GreatInventions_203_P 100_A301_Screen_2_13596 | A | M | A | O |
| 44.R68B_SS_GreatInventions_203_P 100_A301_Screen_3_13597 | A | F | A | O |
| 45.R68B_SS_Greatinventions_203_P 100_A301_Screen_4_13598 | A | F | A | O |
| 46.R68C_SC_Butterflies _mabr_P 100_A301_Screen_2_13611 | A | F | A | H |
| 47.R68C_SC_Butterflies _mabr_P 100_A301_Screen_3_13612 | A | F | A | H |
| 48.R68C_SC_Butterflies _mabr_P 100_A301_Screen_4_13613 | A | F | A | O |
| 49.R68C_MA_Ratios_401_V2_Screen_2_14616 | A | F | A | H |
| $50 . \mathrm{R} 68 \mathrm{C}$ _MA_Ratio s_401_V2_Screen_3_14617 | A | M | A | H |
| 51.R68C_MA_Ratio s_401_V2_Screen_4_14618 | A | M | A | H |
| 52.R68C_SS_NileRiver_401_V1_Screen_2_14491 | A | F | A | H |
| 53.R68C_SS_NileRiver_401_V1_Screen_3_14492 | A | F | A | H |
| 54.R68C_SS_NileRiver_401_V1_Screen_4_14493 | A | F | A | O |
| 55.R68B_SC_RockCycle_DeRoSaJa_P 100_A301_alt1_Screen_2_13656 | A | F | A | O |
| 56.R68B_SC_RockCycle_DeRoSaJa_P 100_A301_alt 1_Screen_3_13657 | A | F | A | H |
| 57.R68B_SC_RockCycle_DeRoSaJa_P 100_A301_alt 1_Screen_4_13658 | A | M | A | O |
| 58.R68C_LA_Prodigy_203_P 100_A301_Screen_2_13614 | A | M | A | O |
| 59.R68C_LA_Prodigy_203_P 100_A301_Screen_3_13615 | A | M | A | O |
| 60.R68C_LA_Prodigy_203_P 100_A301_Screen_4_13616 | A | M | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 61.R68C_MA_FoodCollection_RaObAmSc_P 100_A203_Screen_2_13617 | A | F | A | H |
| 62.R68C_MA_FoodCollection_RaObAmSc_P 100_A203_Screen_3_13618 | A | F | A | O |
| 63.R68C_MA_FoodCollection_RaObAmSc_P 100_A203_Screen_4_13619 | A | M | A | H |
| 64.R68C_SS_StudentCouncil_203_P 100_A 301_Screen_2_13620 | A | F | A | H |
| 65.R68C_SS_StudentCouncil_203_P 100_A 301_Screen_3_13621 | A | M | A | H |
| 66.R68C_SS_StudentCouncil_203_P 100_A301_Screen_4_13622 | A | M | A | O |
| 67.R68C_SC_Vaporization_RaObAm Sc_P 100_A301_Screen_2_13623 | A | F | A | O |
| 68.R68C_SC_Vaporization_RaObAmSc_P 100_A301_Screen_3_13624 | A | F | A | H |
| 69.R68C_SC_Vaporization_RaObAmSc_P 100_A301_Screen_4_13625 | A | F | A | O |

Figure 3.3.4.2C
Raw Scores: Read 6-8 S401 Online
n/a

Table 3.3.4.2C
Raw Score Descriptive Statistics: Read 6-8 S401 Online n/a


Figure 3.3.4.2E
Proficiency Level: Read 6-8 S401 Online


Table 3.3.4.2D
Scale Score Descriptive Statistics: Read 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 58,039 | 217 | 473 | 336.25 | 36.60 |
| $\mathbf{7}$ | 57,821 | 191 | 473 | 343.59 | 39.79 |
| $\mathbf{8}$ | 56,091 | 200 | 473 | 350.97 | 42.11 |
| Total | 171,951 | 191 | 473 | 343.52 | 39.99 |

Table 3.3.4.2E
Proficiency Level Distribution: Read 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
|  | 21,462 | $36.98 \%$ | 21,347 | $36.92 \%$ | 20,506 | $36.56 \%$ | 63,315 | $36.82 \%$ |
| $\mathbf{2}$ | 16,000 | $27.57 \%$ | 15,531 | $26.86 \%$ | 14,347 | $25.58 \%$ | 45,878 | $26.68 \%$ |
| $\mathbf{3}$ | 11,014 | $18.98 \%$ | 10,127 | $17.51 \%$ | 9,103 | $16.23 \%$ | 30,244 | $17.59 \%$ |
| $\mathbf{4}$ | 2,928 | $5.04 \%$ | 3,127 | $5.41 \%$ | 2,684 | $4.79 \%$ | 8,739 | $5.08 \%$ |
| $\mathbf{5}$ | 4,469 | $7.70 \%$ | 4,096 | $7.08 \%$ | 4,754 | $8.48 \%$ | 13,319 | $7.75 \%$ |
| $\mathbf{6}$ | 2,166 | $3.73 \%$ | 3,593 | $6.21 \%$ | 4,697 | $8.37 \%$ | 10,456 | $6.08 \%$ |
| Total | 58,039 | $100.00 \%$ | 57,821 | $100.00 \%$ | 56,091 | $100.00 \%$ | 171,951 | $100.00 \%$ |

Table 3.3.4.2F
Raw Score to Scale Score Conversion: Read 6-8 S401 Online n/a

Table 3.3.4.2G
Equating Summary: Read 6-8 S401 Online


[^26]Figure 3.3.4.2H
Test Characteristic Curve: Read 6-8 S401 Online n/a

Figure 3.3.4.2I


Table 3.3.4.2J
Reliability: Read 6-8 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 171,951 | 69 | .92 |

Table 3.3.4.2K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 6-8 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 6 | 323 | 471 | 10.71 | 12.24 | 12.01 | 0.51 |
|  | 7 | 329 | 629 | 10.71 | 12.24 | 12.02 | 0.42 |
|  | 8 | 335 | 477 | 10.71 | 12.24 | 12.09 | 0.46 |
| 2/3 | 6 | 353 | 2,514 | 10.20 | 10.71 | 10.20 | 0.02 |
|  | 7 | 360 | 26 | 10.20 | 12.24 | 11.24 | 0.55 |
|  | 8 | 366 | 79 | 10.20 | 11.73 | 10.49 | 0.53 |
| 3/4 | 6 | 373 | 107 | 10.20 | 10.71 | 10.32 | 0.21 |
|  | 7 | 380 | 410 | 10.20 | 12.76 | 10.69 | 0.15 |
|  | 8 | 386 | 327 | 10.71 | 12.24 | 11.09 | 0.24 |
| 4/5 | 6 | 382 | 101 | 10.71 | 10.71 | 10.71 | 0.00 |
|  | 7 | 389 | 76 | 10.71 | 11.22 | 10.81 | 0.20 |
|  | 8 | 395 | 221 | 10.71 | 11.73 | 10.92 | 0.38 |
| 5/6 | 6 | 399 | 86 | 10.71 | 11.22 | 10.72 | 0.06 |
|  | 7 | 406 | 16 | 11.73 | 12.24 | 11.89 | 0.24 |
|  | 8 | 412 | 1,108 | 11.22 | 12.24 | 11.23 | 0.04 |

Table 3.3.4.2Li
Accuracy and Consistency of Classification Indices: Read (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.726 | 0.639 |  | 0.514 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.899 |  | 0.850 |  |
|  | 2 | 0.687 |  | 0.589 |  |
|  | 3 | 0.636 |  | 0.520 |  |
|  | 4 | 0.281 |  | 0.204 |  |
|  | 5 | 0.586 |  | 0.443 |  |
|  | 6 | 0.748 |  | 0.573 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.920 | 0.037 | 0.044 | 0.887 |
|  | 2/3 | 0.915 | 0.050 | 0.035 | 0.882 |
|  | 3/4 | 0.940 | 0.032 | 0.028 | 0.914 |
|  | 4/5 | 0.953 | 0.029 | 0.017 | 0.933 |
|  | 5/6 | 0.979 | 0.013 | 0.008 | 0.969 |

Table 3.3.4.2 Lii
Accuracy and Consistency of Classification Indices: Read (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.721 | 0.634 |  | 0.512 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.893 |  | 0.844 |  |
|  | 2 | 0.682 |  | 0.581 |  |
|  | 3 | 0.608 |  | 0.491 |  |
|  | 4 | 0.301 |  | 0.217 |  |
|  | 5 | 0.512 |  | 0.381 |  |
|  | 6 | 0.828 |  | 0.694 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.918 | 0.039 | 0.043 | 0.885 |
|  | 2/3 | 0.917 | 0.048 | 0.035 | 0.884 |
|  | 3/4 | 0.939 | 0.033 | 0.028 | 0.914 |
|  | 4/5 | 0.953 | 0.026 | 0.020 | 0.932 |
|  | 5/6 | 0.973 | 0.018 | 0.009 | 0.962 |

Table 3.3.4.2Liii
Accuracy and Consistency of Classification Indices: Read (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.713 | 0.627 |  | 0.509 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.890 |  | 0.841 |  |
|  | 2 | 0.671 |  | 0.567 |  |
|  | 3 | 0.579 |  | 0.458 |  |
|  | 4 | 0.253 |  | 0.182 |  |
|  | 5 | 0.527 |  | 0.400 |  |
|  | 6 | 0.833 |  | 0.714 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.918 | 0.040 | 0.043 | 0.884 |
|  | 2/3 | 0.917 | 0.046 | 0.037 | 0.885 |
|  | 3/4 | 0.938 | 0.032 | 0.029 | 0.912 |
|  | 4/5 | 0.948 | 0.032 | 0.020 | 0.926 |
|  | 5/6 | 0.966 | 0.022 | 0.012 | 0.952 |

### 3.3.4.3 Writing 6-8

### 3.3.4.3i <br> Writing 6-8 A

Table 3.3.4.3Ai
Complete Task Analysis and Summary: Writ 6-8 A S401 Online

| Task Type |  | Average <br> Task Difficulty (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 0.92 | 3 | 0.46 | 0.49 |
| Name |  | Task Difficulty (in logits) |  | Fit Statistics |  |
|  |  | Anchored? | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.W68A_SI_Clubs_P100_A301_HW_14284 |  |  | 0.85 | Yes | 0.45 | 0.47 |
| 2.W68A_LA_Cat_P100_A301_HW_14285 |  | 0.60 | Yes | 0.45 | 0.47 |
| 3.W68A_MS_Ecosystems_MaMaNiVi_P100_A203_HW_14287 |  | 1.32 | Yes | 0.48 | 0.54 |
| Raw Score Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 6.53\% | 4.18\% | 7.99\% |  |
|  | 1 | 5.97\% | 7.68\% | 10.24\% |  |
|  | 2 | 9.30\% | 8.04\% | 15.07\% |  |
|  | 3 | 17.57\% | 17.37\% | 39.18\% |  |
|  | 4 | 29.76\% | 29.66\% | 21.36\% |  |
|  | 5 | 24.07\% | 23.94\% | 5.34\% |  |
|  | 6 | 5.92\% | 7.71\% | 0.73\% |  |
|  | 7 | 0.81\% | 1.25\% | 0.07\% |  |
|  | 8 | 0.07\% | 0.16\% | 0.01\% |  |
|  | 9 | 0.01\% | 0.02\% | 0.00\% |  |

Table 3.3.4.3Bi
DIF Analysis and Summary: Writ 6-8 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 2 | 1 | 2 | 1 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | DIF <br> Level | Favored Group |
| 1.W68A_SI_Clubs_P 100_A301_HW_14284 | AA | F | AA | H |
| 2.W68A_LA_Cat_P 100_A301_HW_14285 | AA | M | AA | O |
| 3.W68A_MS_Ecos ystems_MaMaNiVi_P 100_A203_HW_14287 | AA | M | AA | H |

Figure 3.3.4.3Ci
Raw Scores: Writ 6-8 A S401 Online


Figure 3.3.4.3Di
Scale Scores: Writ 6-8 A S401 Online


Table 3.3.4.3Ci
Raw Score Descriptive Statistics: Writ 6-8 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 24,619 | 0 | 21 | 9.33 | 4.01 |
| $\mathbf{7}$ | 28,520 | 0 | 23 | 10.09 | 3.95 |
| $\mathbf{8}$ | 28,652 | 0 | 23 | 10.60 | 3.94 |
| Total | 81,791 | 0 | 23 | 10.04 | 4.00 |

Table 3.3.4.3Di
Scale Score Descriptive Statistics: Writ 6-8 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 24,619 | 166 | 398 | 281.95 | 34.25 |
| $\mathbf{7}$ | 28,520 | 177 | 416 | 288.75 | 33.16 |
| $\mathbf{8}$ | 28,652 | 188 | 416 | 293.61 | 32.66 |
| Total | 81,791 | 166 | 416 | 288.41 | 33.65 |



Table 3.3.4.3Ei
Proficiency Level Distribution: Writ 6-8 A S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 6,769 | $27.50 \%$ | 7,776 | $27.27 \%$ | 8,738 | $30.50 \%$ | 23,283 | $28.47 \%$ |
| $\mathbf{2}$ | 9,399 | $38.18 \%$ | 12,299 | $43.12 \%$ | 9,488 | $33.11 \%$ | 31,186 | $38.13 \%$ |
| $\mathbf{3}$ | 8,409 | $34.16 \%$ | 8,299 | $29.10 \%$ | 10,356 | $36.14 \%$ | 27,064 | $33.09 \%$ |
| $\mathbf{4}$ | 42 | $0.17 \%$ | 146 | $0.51 \%$ | 70 | $0.24 \%$ | 258 | $0.32 \%$ |
| $\mathbf{5}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 24,619 | $100.00 \%$ | 28,520 | $100.00 \%$ | 28,652 | $100.00 \%$ | 81,791 | $100.00 \%$ |

Table 3.3.4.3Fi
Raw Score to Scale Score Conversion: Writ 6-8 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 188^ | 52.90 | 141.10 | 246.90 |
| 1 | 220 | 23.41 | 196.59 | 243.41 |
| 2 | 234 | 16.49 | 217.51 | 250.49 |
| 3 | 243 | 13.94 | 229.06 | 256.94 |
| 4 | 249 | 12.75 | 236.25 | 261.75 |
| 5 | 255 | 12.16 | 242.84 | 267.16 |
| 6 | 260 | 11.98 | 248.02 | 271.98 |
| 7 | 266 | 12.08 | 253.92 | 278.08 |
| 8 | 271 | 12.46 | 258.54 | 283.46 |
| 9 | 277 | 13.08 | 263.92 | 290.08 |
| 10 | 284 | 13.99 | 270.01 | 297.99 |
| 11 | 292 | 15.09 | 276.91 | 307.09 |
| 12 | 301 | 16.16 | 284.84 | 317.16 |
| 13 | 312 | 17.02 | 294.98 | 329.02 |
| 14 | 323 | 17.59 | 305.41 | 340.59 |
| 15 | 334 | 17.83 | 316.17 | 351.83 |
| 16 | 346 | 17.78 | 328.22 | 363.78 |
| 17 | 358 | 17.45 | 340.55 | 375.45 |
| 18 | 369 | 16.94 | 352.06 | 385.94 |
| 19 | 379 | 16.33 | 362.67 | 395.33 |
| 20 | 389 | 15.76 | 373.24 | 404.76 |
| 21 | 398 | 15.41 | 382.59 | 413.41 |
| 22 | 407 | 15.36 | 391.64 | 422.36 |
| 23 | 416 | 15.82 | 400.18 | 431.82 |
| 24 | 425 | 16.97 | 408.03 | 441.97 |
| 25 | 438 | 19.57 | 418.43 | 457.57 |
| 26 | 457 | 26.58 | 430.42 | 483.58 |
| 27 | 488 | 48.39 | 439.61 | 536.39 |

$\wedge$ Truncated

Table 3.3.4.3Gi
Equating Summary: Writ 6-8 A S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 0.92 (0.36) | 3 |  | 2.83 (0.38) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | 0.60 |  | 1.32 | 2.45 |  | 3.20 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 3 |  | 0.92 (0.36) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 3 |  | $0.92(0.36)$ |  |  |  |
|  | Percentage Anchors |  | Average <br> Displacement |  |  |  |
|  | 100\% |  | 0.00 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Task ID | Task Difficulty | y y ( Displacement | Task ID | Task Difficulty | Displacement |
|  | 14284 | 0.85 | -0.12 | 14285 | 0.60 | -0.02 |
|  | 14285 | 0.60 | -0.02 | 14284 | 0.85 | -0.12 |
|  | 14287 | 1.32 | 0.12 | 14287 | 1.32 | 0.12 |

[^27]

Table 3.3.4.3Ji
Reliability: Writ 6-8 A S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's <br> Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 81,791 | 3 | Hand-written (HW) | Keyboarded (KB) | . 870 | 1.439 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 636 | 98 | 2 | 0 |
|  |  | KB | 39,638 | 95 | 5 | 0 |
|  | 2 | HW | 590 | 97 | 3 | 0 |
|  |  | KB | 39,956 | 94 | 5 | 0 |
|  | 3 | HW | 628 | 97 | 2 | 0 |
|  |  | KB | 39,790 | 96 | 4 | 0 |

Table 3.3.4.3Ki
Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 A S401 Online n/a

Table 3.3.4.3Li
Accuracy and Consistency of Classification Indices: Writ 6-8 A S401 Online n/a

Table 3.3.4.3Aii
Complete Task Analysis and Summary: Writ 6-8 B/C S401 Online

| Task Type |  | $\begin{array}{\|c} \hline \text { Average } \\ \text { Task } \\ \text { Difficulty } \\ \text { (in logits) } \end{array}$ | No. of Tasks | $\begin{gathered} \hline \text { Average } \\ \text { Infit } \\ \text { Mean } \\ \text { Square } \end{gathered}$ | Average Outfit Mean Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 1.30 | 3 | 0.54 | 0.52 |
| Name |  | Task Difficulty (in logits) | Anchored? | Fit Statistics |  |
|  |  | Infit <br> Mnsq |  | Outfit <br> Mnsq |
| 1.W68B_SI_SchoolElectronics_P100_A 301 _HW_14655 |  |  | 1.10 | Yes | 0.76 | 0.73 |
| 2.W68B_MS_BodySystems_kiow_P100_A 203 _HW_14654 |  | 1.42 | Yes | 0.54 | 0.51 |
| 3.W68C_IT_ServiceProject_401_HW_15764 |  | 1.37 |  | 0.34 | 0.32 |
| Raw Score <br> Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 0.14\% | 0.30\% | 0.77\% |  |
|  | 1 | 0.22\% | 0.25\% | 0.32\% |  |
|  | 2 | 0.66\% | 1.00\% | 1.52\% |  |
|  | 3 | 2.42\% | 3.42\% | 3.67\% |  |
|  | 4 | 11.84\% | 15.28\% | 15.21\% |  |
|  | 5 | 41.38\% | 46.75\% | 36.29\% |  |
|  | 6 | 35.92\% | 28.56\% | 32.55\% |  |
|  | 7 | 6.70\% | 3.95\% | 7.80\% |  |
|  | 8 | 0.65\% | 0.43\% | 1.50\% |  |
|  | 9 | 0.07\% | 0.07\% | 0.36\% |  |

Table 3.3.4.3Bii
DIF Analysis and Summary: Writ 6-8 B/C S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 2 | 2 | 1 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level |  |
|  | Favored <br> Group |  |  |  |
| 1.W68B_SI_SchooIElectronics_P 100_A301_HW_14655 | AA | F | AA | H |
| 2.W68B_MS_BodySystems_kiow_P 100_A203_HW_14654 | AA | M | AA | H |
| 3.W68C_IT_ServiceProject_401_HW_15764 | AA | F | AA | O |

Figure 3.3.4.3Cii
Raw Scores: Writ 6-8 B/C S401 Online


Figure 3.3.4.3Dii
Scale Scores: Writ 6-8 B/C S401 Online


Table 3.3.4.3Cii
Raw Score Descriptive Statistics: Writ 6-8 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 39,209 | 0 | 50 | 29.05 | 5.63 |
| $\mathbf{7}$ | 37,403 | 0 | 52 | 31.46 | 4.97 |
| $\mathbf{8}$ | 36,268 | 0 | 54 | 33.17 | 4.81 |
| Total | 112,880 | 0 | 54 | 31.18 | 5.43 |

Table 3.3.4.3Dii
Scale Score Descriptive Statistics: Writ 6-8 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 39,209 | 166 | 448 | 341.89 | 28.70 |
| $\mathbf{7}$ | 37,403 | 177 | 467 | 354.75 | 26.40 |
| $\mathbf{8}$ | 36,268 | 188 | 516 | 364.06 | 25.65 |
| Total | 112,880 | 166 | 516 | 353.28 | 28.49 |

Figure 3.3.4.3Eii


Table 3.3.4.3Eii
Proficiency Level Distribution: Writ 6-8 B/C S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
|  | 408 | $1.04 \%$ | 190 | $0.51 \%$ | 203 | $0.56 \%$ | 801 | $0.71 \%$ |
| $\mathbf{2}$ | 2,321 | $5.92 \%$ | 1,446 | $3.87 \%$ | 790 | $2.18 \%$ | 4,557 | $4.04 \%$ |
| $\mathbf{3}$ | 25,530 | $65.11 \%$ | 23,194 | $62.01 \%$ | 21,364 | $58.91 \%$ | 70,088 | $62.09 \%$ |
| $\mathbf{4}$ | 10,862 | $27.70 \%$ | 12,444 | $33.27 \%$ | 13,710 | $37.80 \%$ | 37,016 | $32.79 \%$ |
| $\mathbf{5}$ | 86 | $0.22 \%$ | 124 | $0.33 \%$ | 192 | $0.53 \%$ | 402 | $0.36 \%$ |
| $\mathbf{6}$ | 2 | $0.01 \%$ | 5 | $0.01 \%$ | 9 | $0.02 \%$ | 16 | $0.01 \%$ |
| Total | 39,209 | $100.00 \%$ | 37,403 | $100.00 \%$ | 36,268 | $100.00 \%$ | 112,880 | $100.00 \%$ |

Table 3.3.4.3Fii
Raw Score to Scale Score Conversion: Writ 6-8 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | Raw Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 188^ | 49.14 | 140.86 | 239.14 | 34 | 369 | 12.38 | 356.62 | 381.38 |
| 1 | 218 | 24.38 | 193.62 | 242.38 | 35 | 375 | 12.22 | 362.78 | 387.22 |
| 2 | 233 | 16.57 | 216.43 | 249.57 | 36 | 380 | 12.00 | 368.00 | 392.00 |
| 3 | 241 | 13.35 | 227.65 | 254.35 | 37 | 386 | 11.79 | 374.21 | 397.79 |
| 4 | 247 | 11.57 | 235.43 | 258.57 | 38 | 391 | 11.55 | 379.45 | 402.55 |
| 5 | 251 | 10.45 | 240.55 | 261.45 | 39 | 396 | 11.33 | 384.67 | 407.33 |
| 6 | 255 | 9.69 | 245.31 | 264.69 | 40 | 400 | 11.12 | 388.88 | 411.12 |
| 7 | 258 | 9.16 | 248.84 | 267.16 | 41 | 405 | 10.96 | 394.04 | 415.96 |
| 8 | 261 | 8.78 | 252.22 | 269.78 | 42 | 409 | 10.82 | 398.18 | 419.82 |
| 9 | 264 | 8.51 | 255.49 | 272.51 | 43 | 414 | 10.74 | 403.26 | 424.74 |
| 10 | 267 | 8.35 | 258.65 | 275.35 | 44 | 418 | 10.74 | 407.26 | 428.74 |
| 11 | 269 | 8.24 | 260.76 | 277.24 | 45 | 422 | 10.85 | 411.15 | 432.85 |
| 12 | 272 | 8.22 | 263.78 | 280.22 | 46 | 427 | 11.04 | 415.96 | 438.04 |
| 13 | 274 | 8.24 | 265.76 | 282.24 | 47 | 431 | 11.33 | 419.67 | 442.33 |
| 14 | 277 | 8.32 | 268.68 | 285.32 | 48 | 436 | 11.84 | 424.16 | 447.84 |
| 15 | 279 | 8.46 | 270.54 | 287.46 | 49 | 442 | 12.57 | 429.43 | 454.57 |
| 16 | 282 | 8.65 | 273.35 | 290.65 | 50 | 448 | 13.69 | 434.31 | 461.69 |
| 17 | 285 | 8.89 | 276.11 | 293.89 | 51 | 456 | 15.49 | 440.51 | 471.49 |
| 18 | 288 | 9.18 | 278.82 | 297.18 | 52 | 467 | 18.74 | 448.26 | 485.74 |
| 19 | 291 | 9.53 | 281.47 | 300.53 | 53 | 485 | 26.42 | 458.58 | 511.42 |
| 20 | 295 | 9.93 | 285.07 | 304.93 | 54 | 516 | 48.65 | 467.35 | 564.65 |
| 21 | 299 | 10.34 | 288.66 | 309.34 |  |  |  |  |  |
| 22 | 303 | 10.77 | 292.23 | 313.77 |  |  |  |  |  |
| 23 | 307 | 11.17 | 295.83 | 318.17 |  |  |  |  |  |
| 24 | 312 | 11.52 | 300.48 | 323.52 |  |  |  |  |  |
| 25 | 317 | 11.84 | 305.16 | 328.84 |  |  |  |  |  |
| 26 | 323 | 12.11 | 310.89 | 335.11 |  |  |  |  |  |
| 27 | 328 | 12.32 | 315.68 | 340.32 |  |  |  |  |  |
| 28 | 334 | 12.49 | 321.51 | 346.49 |  |  |  |  |  |
| 29 | 340 | 12.59 | 327.41 | 352.59 |  |  |  |  |  |
| 30 | 346 | 12.65 | 333.35 | 358.65 |  |  |  |  |  |
| 31 | 352 | 12.65 | 339.35 | 364.65 |  |  |  |  |  |
| 32 | 358 | 12.62 | 345.38 | 370.62 |  |  |  |  |  |
| 33 | 364 | 12.51 | 351.49 | 376.51 |  |  |  |  |  |

[^28]Table 3.3.4.3Gii
Equating Summary: Writ 6-8 B/C S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 1.30 (0.17) | 3 |  | 1.59 (0.26) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | 1.10 |  | 1.42 | 1.35 |  | 1.87 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | 1.26 (0.23) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | (Std. Dev.) |  |  |  |
|  | Percentage Anchors |  | Average Displacement |  |  |  |
|  | 67\% |  | -0.02 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Task ID ${ }^{\text {D }}$ | Task Difficulty | y y Displacement | Task ID | Task Difficulty | Displacement |
|  | 14655 | 1.10 | -0.06 | 14655 | 1.10 | -0.06 |
|  | 14654 | 1.42 | 0.03 | 14654 | 1.42 | 0.03 |

[^29]

Table 3.3.4.3Jii
Reliability: Writ 6-8 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's <br> Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 112,880 | 3 | Hand-written (HW) | Keyboarded (KB) | . 904 | 1.686 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 378 | 97 | 3 | 0 |
|  |  | KB | 57,860 | 96 | 3 | 0 |
|  | 2 | HW | 382 | 96 | 4 | 0 |
|  |  | KB | 58,514 | 96 | 4 | 0 |
|  | 3 | HW | 402 | 94 | 6 | 0 |
|  |  | KB | 59,824 | 95 | 5 | 0 |

Table 3.3.4.3Kii
Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 B/C S401 Online n/a

Table 3.3.4.3Lii
Accuracy and Consistency of Classification Indices: Writ 6-8 B/C S401 Online n/a

### 3.3.4.3iii Writing 6-8 Across Tiers

Table 3.3.4.3Aiii
Complete Task Analysis and Summary: Writ 6-8 S401 Online n/a

Table 3.3.4.3Biii
DIF Analysis and Summary: Writ 6-8 S401 Online
n/a



Table 3.3.4.3Ciii
Raw Score Descriptive Statistics: Writ 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 63,828 | 0 | 50 | 21.45 | 10.85 |
| $\mathbf{7}$ | 65,923 | 0 | 52 | 22.22 | 11.53 |
| $\mathbf{8}$ | 64,920 | 0 | 54 | 23.21 | 12.06 |
| Total | 194,671 | 0 | 54 | 22.30 | 11.52 |

Table 3.3.4.3Diii
Scale Score Descriptive Statistics: Writ 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 63,828 | 166 | 448 | 318.77 | 42.54 |
| $\mathbf{7}$ | 65,923 | 177 | 467 | 326.20 | 44.05 |
| $\mathbf{8}$ | 64,920 | 188 | 516 | 332.97 | 45.41 |
| Total | 194,671 | 166 | 516 | 326.02 | 44.40 |



Table 3.3.4.3Eiii
Proficiency Level Distribution: Writ 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 7,177 | $11.24 \%$ | 7,966 | $12.08 \%$ | 8,941 | $13.77 \%$ | 24,084 | $12.37 \%$ |
| $\mathbf{2}$ | 11,720 | $18.36 \%$ | 13,745 | $20.85 \%$ | 10,278 | $15.83 \%$ | 35,743 | $18.36 \%$ |
| $\mathbf{3}$ | 33,939 | $53.17 \%$ | 31,493 | $47.77 \%$ | 31,720 | $48.86 \%$ | 97,152 | $49.91 \%$ |
| $\mathbf{4}$ | 10,904 | $17.08 \%$ | 12,590 | $19.10 \%$ | 13,780 | $21.23 \%$ | 37,274 | $19.15 \%$ |
| $\mathbf{5}$ | 86 | $0.13 \%$ | 124 | $0.19 \%$ | 192 | $0.30 \%$ | 402 | $0.21 \%$ |
| $\mathbf{6}$ | 2 | $0.00 \%$ | 5 | $0.01 \%$ | 9 | $0.01 \%$ | 16 | $0.01 \%$ |
| Total | 63,828 | $100.00 \%$ | 65,923 | $100.00 \%$ | 64,920 | $100.00 \%$ | 194,671 | $100.00 \%$ |

Table 3.3.4.3Fiii
Raw Score to Scale Score Conversion: Writ 6-8 S401 Online n/a

Table 3.3.4.3Giii
Equating Summary: Writ 6-8 S401 Online n/a



Table 3.3.4.3Jiii
Reliability: Writ 6-8 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| A | 81,791 | 0.870 | 0.890 |
| B/C | 112,880 | 0.904 |  |

Table 3.3.4.3Kiii
Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 S401 Online

| Proficiency <br> Level |  |  | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Cut Score | Tier A | Tier B/C |
| $1 / 2$ | 6 | 268 | 12.35 | 8.32 |
|  | 7 | 273 | 12.62 | 8.32 |
|  | 8 | 281 | 13.69 | 8.59 |
| $2 / 3$ | 6 | 298 | 15.84 | 10.20 |
|  | 7 | 305 | 16.65 | 11.01 |
|  | 8 | 311 | 16.92 | 11.55 |
| $3 / 4$ | 6 | 361 | 17.45 | 12.62 |
|  | 7 | 367 | 17.18 | 12.35 |
|  | 8 | 372 | 16.92 | 12.35 |
| $4 / 5$ | 6 | 413 | 15.57 | 10.74 |
|  | 7 | 419 | 16.11 | 10.74 |
|  | 8 | 424 | 16.92 | 11.01 |
| $5 / 6$ | 6 | 441 | 20.68 | 12.35 |
|  | 7 | 450 | 23.90 | 14.23 |
|  | 8 | 459 | 27.93 | 16.38 |

Table 3.3.4.3Li
Accuracy and Consistency of Classification Indices: Writ (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.761 | 0.676 |  | 0.490 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.811 |  | 0.701 |  |
|  | 2 | 0.688 |  | 0.565 |  |
|  | 3 | 0.790 |  | 0.740 |  |
|  | 4 | 0.687 |  | 0.550 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.957 | 0.021 | 0.022 | 0.937 |
|  | $2 / 3$ | 0.919 | 0.030 | 0.051 | 0.888 |
|  | 3/4 | 0.885 | 0.070 | 0.045 | 0.846 |
|  | 4/5 | 0.999 | 0.001 | 0.000 | 0.999 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.4.3Lii
Accuracy and Consistency of Classification Indices: Writ (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.762 | 0.673 |  | 0.514 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.810 |  | 0.699 |  |
|  | 2 | 0.706 |  | 0.590 |  |
|  | 3 | 0.778 |  | 0.716 |  |
|  | 4 | 0.739 |  | 0.623 |  |
|  | 5 | - |  | 0.032 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.953 | 0.023 | 0.024 | 0.932 |
|  | 2/3 | 0.913 | 0.031 | 0.056 | 0.881 |
|  | 3/4 | 0.896 | 0.059 | 0.045 | 0.857 |
|  | 4/5 | 0.998 | 0.002 | 0.000 | 0.998 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.4.3Liii
Accuracy and Consistency of Classification Indices: Writ (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.767 | 0.678 |  | 0.522 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.855 |  | 0.760 |  |
|  | 2 | 0.623 |  | 0.497 |  |
|  | 3 | 0.800 |  | 0.736 |  |
|  | 4 | 0.738 |  | 0.633 |  |
|  | 5 | - |  | 0.034 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.955 | 0.019 | 0.026 | 0.935 |
|  | 2/3 | 0.922 | 0.033 | 0.044 | 0.891 |
|  | 3/4 | 0.891 | 0.056 | 0.052 | 0.849 |
|  | 4/5 | 0.997 | 0.003 | 0.000 | 0.995 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.4.4 Speaking 6-8

### 3.3.4.4i Speaking 6-8 Pre-A

Table 3.3.4.4Ai
Complete Task Analysis and Summary: Spek 6-8 Pre-A S401 Online n/a

Table 3.3.4.4Bi
DIF Analysis and Summary: Spek 6-8 Pre-A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 2 | 2 | 1 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S68P_SI_PeerRevie wing_401_14510 | AA | M | AA | H |
| 2.S68P _LS_Luis Rivera_401_14745 | AA | F | AA | H |
| 3.S68P_MS_Pedometers_401_14747 | AA | F | AA | O |



Figure 3.3.4.4Di
Scale Scores: Spek 6-8 Pre-A S401 Online


Table 3.3.4.4Ci
Raw Score Descriptive Statistics: Spek 6-8 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 2,538 | 0 | 6 | 5.37 | 1.20 |
| $\mathbf{7}$ | 3,549 | 0 | 6 | 5.48 | 1.09 |
| $\mathbf{8}$ | 4,621 | 0 | 6 | 5.55 | 1.02 |
| Total | 10,708 | 0 | 6 | 5.48 | 1.09 |

Table 3.3.4.4Di
Scale Score Descriptive Statistics: Spek 6-8 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 2,538 | 136 | 210 | 201.77 | 15.72 |
| $\mathbf{7}$ | 3,549 | 142 | 210 | 203.21 | 14.07 |
| $\mathbf{8}$ | 4,621 | 148 | 210 | 204.23 | 12.91 |
| Total | 10,708 | 136 | 210 | 203.31 | 14.03 |

Figure 3.3.4.4Ei
Proficiency Level: Spek 6-8 Pre-A S401 Online


Table 3.3.4.4Ei
Proficiency Level Distribution: Spek 6-8 Pre-A S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,538 | $100.00 \%$ | 3,549 | $100.00 \%$ | 4,621 | $100.00 \%$ | 10,708 | $100.00 \%$ |
| Total | 2,538 | $100.00 \%$ | 3,549 | $100.00 \%$ | 4,621 | $100.00 \%$ | 10,708 | $100.00 \%$ |

Table 3.3.4.4Fi
Raw Score to Scale Score Conversion: Spek 6-8 Pre-A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $148^{\wedge}$ | 23.40 | $100.00^{\wedge}$ | 133.40 |
| 1 | $148^{\wedge}$ | 23.40 | 115.60 | 162.40 |
| 2 | 157 | 20.77 | 136.23 | 177.77 |
| 3 | 171 | 19.60 | 151.40 | 190.60 |
| 4 | 184 | 20.77 | 163.23 | 204.77 |
| 5 | $197^{*}$ | 24.57 | 177.43 | 226.57 |
| 6 | $210^{*}$ | 31.29 | 199.71 | 262.29 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.4.4Gi
Equating Summary: Spek 6-8 Pre-A S401 Online n/a



Table 3.3.4.4Ji
Reliability: Spek 6-8 Pre-A S401 Online

| Reliability |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | No. of Students | No. of Tasks | Cronbach's Alpha | SEM |  |
|  | 10,708 | 3 | .719 |  | 0.577 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 6,114 | 97 | 3 | 0 |
|  | 2 | 6,100 | 97 | 3 | 0 |

Table 3.3.4.4Ki
Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 Pre-A S401 Online n/a

Table 3.3.4.4Li
Accuracy and Consistency of Classification Indices: Spek 6-8 Pre-A S401 Online n/a

### 3.3.4.4ii $\quad$ Speaking 6-8 A

Table 3.3.4.4Aii
Complete Task Analysis and Summary: Spek 6-8 A S401 Online n/a

Table 3.3.4.4Bii
DIF Analysis and Summary: Spek 6-8 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\underset{\text { DiF }}{\text { DIFel }}$ | Favoring Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 1 | 5 | 2 | 4 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S68A_SİP eerRe vie wing_401_4510 | AA | F | AA | H |
| 2.S68A_SLPP eerRevie wing_401_14511 | AA | F | AA | O |
| 3.S68A_LS_Luis Rivera_401_14745 | AA | F | AA | O |
| 4.S68A_LS_Luis Rivera_401_14746 | AA | M | AA | H |
| 5.S68A_MS_Pedometers_401_Vl_4747 | AA | F | AA | O |
| 6.S68A_MS_Pedometers_401_Vl_14748 | AA | F | AA | O |

Figure 3.3.4.4Cii Raw Scores: Spek 6-8 A S401 Online


Figure 3.3.4.4Dii
Scale Scores: Spek 6-8 A S401 Online


Table 3.3.4.4Cii
Raw Score Descriptive Statistics: Spek 6-8 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 13,195 | 0 | 17 | 10.79 | 1.70 |
| $\mathbf{7}$ | 9,301 | 0 | 18 | 10.73 | 1.66 |
| $\mathbf{8}$ | 16,134 | 0 | 18 | 11.41 | 1.62 |
| Total | 38,630 | 0 | 18 | 11.04 | 1.69 |

Table 3.3.4.4Dii
Scale Score Descriptive Statistics: Spek 6-8 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 13,195 | 136 | 438 | 301.53 | 34.46 |
| $\mathbf{7}$ | 9,301 | 142 | 459 | 300.17 | 33.91 |
| $\mathbf{8}$ | 16,134 | 148 | 459 | 314.85 | 35.27 |
| Total | 38,630 | 136 | 459 | 306.77 | 35.34 |

Figure 3.3.4.4Eii
Proficiency Level: Spek 6-8 A S401 Online


Table 3.3.4.4Eii
Proficiency Level Distribution: Spek 6-8 A S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,402 | $18.20 \%$ | 1,832 | $19.70 \%$ | 3,938 | $24.41 \%$ | 8,172 | $21.15 \%$ |
| $\mathbf{2}$ | 6,375 | $48.31 \%$ | 4,535 | $48.76 \%$ | 4,320 | $26.78 \%$ | 15,230 | $39.43 \%$ |
| $\mathbf{3}$ | 4,006 | $30.36 \%$ | 2,675 | $28.76 \%$ | 7,486 | $46.40 \%$ | 14,167 | $36.67 \%$ |
| $\mathbf{4}$ | 394 | $2.99 \%$ | 257 | $2.76 \%$ | 371 | $2.30 \%$ | 1,022 | $2.65 \%$ |
| $\mathbf{5}$ | 18 | $0.14 \%$ | 1 | $0.01 \%$ | 19 | $0.12 \%$ | 38 | $0.10 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 1 | $0.01 \%$ | 0 | $0.00 \%$ | 1 | $0.00 \%$ |
| Total | 13,195 | $100.00 \%$ | 9,301 | $100.00 \%$ | 16,134 | $100.00 \%$ | 38,630 | $100.00 \%$ |

Table 3.3.4.4Fii
Raw Score to Scale Score Conversion: Spek 6-8 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $148^{\wedge}$ | 22.52 | $100.00^{\wedge}$ | 130.52 |
| 1 | $148^{\wedge}$ | 22.52 | 114.48 | 159.52 |
| 2 | 155 | 20.47 | 134.53 | 175.47 |
| 3 | 168 | 18.72 | 149.28 | 186.72 |
| 4 | 181 | 19.01 | 161.99 | 200.01 |
| 5 | 194 | 20.77 | 173.23 | 214.77 |
| 6 | 211 | 23.69 | 187.31 | 234.69 |
| 7 | 231 | 23.98 | 207.02 | 254.98 |
| 8 | 250 | 22.23 | 227.77 | 272.23 |
| 9 | 266 | 21.64 | 244.36 | 287.64 |
| 10 | 282 | 22.52 | 259.48 | 304.52 |
| 11 | 302 | 25.45 | 276.55 | 327.45 |
| 12 | 327 | 28.37 | 298.63 | 355.37 |
| 13 | 353 | 26.91 | 326.09 | 379.91 |
| 14 | 376 | 24.57 | 351.43 | 400.57 |
| 15 | 396 | 23.98 | 372.02 | 419.98 |
| 16 | 417 | 25.74 | 391.26 | 442.74 |
| 17 | $438^{*}$ | 30.42 | 414.58 | 475.42 |
| 18 | $459^{*}$ | 38.61 | 446.39 | 523.61 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.4.4Gii
Equating Summary: Spek 6-8 A S401 Online n/a



Table 3.3.4.4Jii
Reliability: Spek 6-8 A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 38,630 | 6 | .622 |  | 1.039 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 24,040 | 99 | 1 | 0 |
|  | 2 | 24,045 | 83 | 17 | 0 |
|  | 3 | 24,051 | 98 | 2 | 0 |
|  | 4 | 24,044 | 83 | 17 | 0 |
|  | 5 | 23,800 | 99 | 1 | 0 |

Table 3.3.4.4Kii
Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 A S401 Online n/a

Table 3.3.4.4Lii
Accuracy and Consistency of Classification Indices: Spek 6-8 A S401 Online n/a

### 3.3.4.4iii Speaking 6-8 B/C

Table 3.3.4.4Aiii
Complete Task Analysis and Summary: Spek 6-8 B/C S401 Online n/a

Table 3.3.4.4Biii
DIF Analysis and Summary: Spek 6-8 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 3 | 3 | 4 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.S68C_SI_PeerReviewing_401_14504 | AA | F | AA | H |  |
| 2.S68C_SI_PeerReviewing_401_14505 | AA | F | AA | H |  |
| 3.S68C_LS_LuisRivera_401_V1_14751 | AA | M | AA | H |  |
| 4.S68C_LS_LuisRivera_401_V1_14752 | AA | F | AA | H |  |
| 5.S68C_MS_Pedometers_401_V2_14757 | AA | M | AA | O |  |
| 6.S68C_MS_Pedometers_401_V2_14758 | AA | M | AA | O |  |

Figure 3.3.4.4Ciii


Figure 3.3.4.4Diii Scale Scores: Spek 6-8 B/C S401 Online


Table 3.3.4.4Ciii
Raw Score Descriptive Statistics: Spek 6-8 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 39,608 | 6 | 28 | 17.77 | 2.33 |
| $\mathbf{7}$ | 42,380 | 6 | 29 | 18.15 | 2.43 |
| $\mathbf{8}$ | 34,814 | 6 | 30 | 19.02 | 2.46 |
| Total | 116,802 | 6 | 30 | 18.28 | 2.46 |

Table 3.3.4.4Diii
Scale Score Descriptive Statistics: Spek 6-8 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 39,608 | 136 | 447 | 330.32 | 27.76 |
| $\mathbf{7}$ | 42,380 | 142 | 459 | 334.96 | 28.98 |
| $\mathbf{8}$ | 34,814 | 148 | 471 | 345.34 | 29.41 |
| Total | 116,802 | 136 | 471 | 336.48 | 29.34 |

Figure 3.3.4.4Eiii
Proficiency Level: Spek 6-8 B/C S401 Online


Table 3.3.4.4Eiii
Proficiency Level Distribution: Spek 6-8 B/C S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 157 | $0.40 \%$ | 433 | $1.02 \%$ | 416 | $1.19 \%$ | 1,006 | $0.86 \%$ |
| $\mathbf{2}$ | 10,503 | $26.52 \%$ | 9,158 | $21.61 \%$ | 8,189 | $23.52 \%$ | 27,850 | $23.84 \%$ |
| $\mathbf{3}$ | 24,680 | $62.31 \%$ | 26,375 | $62.23 \%$ | 21,095 | $60.59 \%$ | 72,150 | $61.77 \%$ |
| $\mathbf{4}$ | 4,192 | $10.58 \%$ | 6,342 | $14.96 \%$ | 4,979 | $14.30 \%$ | 15,513 | $13.28 \%$ |
| $\mathbf{5}$ | 76 | $0.19 \%$ | 68 | $0.16 \%$ | 130 | $0.37 \%$ | 274 | $0.23 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 4 | $0.01 \%$ | 5 | $0.01 \%$ | 9 | $0.01 \%$ |
| Total | 39,608 | $100.00 \%$ | 42,380 | $100.00 \%$ | 34,814 | $100.00 \%$ | 116,802 | $100.00 \%$ |

Table 3.3.4.4Fiii
Raw Score to Scale Score Conversion: Spek 6-8 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 6 | $148^{\wedge}$ | 22.23 | 181.77 | 226.23 |
| 7 | 219 | 20.47 | 198.53 | 239.47 |
| 8 | 232 | 19.01 | 212.99 | 251.01 |
| 9 | 244 | 17.55 | 226.45 | 261.55 |
| 10 | 254 | 16.38 | 237.62 | 270.38 |
| 11 | 263 | 15.79 | 247.21 | 278.79 |
| 12 | 271 | 15.50 | 255.50 | 286.50 |
| 13 | 279 | 15.79 | 263.21 | 294.79 |
| 14 | 288 | 16.09 | 271.91 | 304.09 |
| 15 | 297 | 16.96 | 280.04 | 313.96 |
| 16 | 308 | 17.84 | 290.16 | 325.84 |
| 17 | 319 | 19.30 | 299.70 | 338.30 |
| 18 | 332 | 20.18 | 311.82 | 352.18 |
| 19 | 346 | 19.89 | 326.11 | 365.89 |
| 20 | 359 | 19.01 | 339.99 | 378.01 |
| 21 | 371 | 18.13 | 352.87 | 389.13 |
| 22 | 382 | 17.55 | 364.45 | 399.55 |
| 23 | 392 | 16.96 | 375.04 | 408.96 |
| 24 | 402 | 16.96 | 385.04 | 418.96 |
| 25 | 412 | 17.55 | 394.45 | 429.55 |
| 26 | 423 | 18.13 | 404.87 | 441.13 |
| 27 | 435 | 19.89 | 415.11 | 454.89 |
| 28 | $447^{*}$ | 22.23 | 428.77 | 473.23 |
| 29 | $459^{*}$ | 25.45 | 448.55 | 499.45 |
| 30 | $471^{*}$ | 29.54 | 481.46 | 540.54 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.4.4Giii
Equating Summary: Spek 6-8 B/C S401 Online n/a



Table 3.3.4.4Jiii
Reliability: Spek 6-8 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 116,802 | 6 | .731 |  | 1.277 |
|  | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 70,442 | 78 | 22 | 1 |
|  | 2 | 70,444 | 77 | 23 | 0 |
|  | 3 | 70,708 | 80 | 20 | 0 |
|  | 4 | 70,722 | 77 | 23 | 1 |
|  | 5 | 70,984 | 79 | 20 | 1 |

Table 3.3.4.4Kiii
Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 B/C S401 Online n/a

Table 3.3.4.4Liii
Accuracy and Consistency of Classification Indices: Spek 6-8 B/C S401 Online n/a

### 3.3.4.4iv $\quad$ Speaking 6-8 Across Tiers

Table 3.3.4.4Aiv
Complete Task Analysis and Summary: Spek 6-8 S401 Online

| Task Type |  | Average Task Difficulty (in logits) |  | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 0.57 |  | 15 | 0.57 | 0.41 |
| Name |  | Tier | Task Difficulty (in logits) |  | Fit S | stics |
|  |  | Anchored? |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 1.S68A_SI_PeerReviewing_401_14510 |  |  | A** | -3.54 |  | 0.66 | 0.17 |
| 2.S68A_SI_PeerReviewing_401_14511 |  | A* | 2.10 | Yes | 0.60 | 0.60 |
| 3.S68A_LS_LuisRivera_401_14745 |  | A** | -2.97 |  | 0.63 | 0.15 |
| 4.S68A_LS_LuisRivera_401_14746 |  | A* | 2.03 | Yes | 0.46 | 0.45 |
| 5.S68A_MS_Pedometers_401_V1_14747 |  | A** | -3.13 |  | 0.64 | 0.14 |
| 6.S68A_MS_Pedometers_401_V1_14748 |  | A* | 2.64 | Yes | 0.51 | 0.53 |
| 7.S68C_SI_PeerReviewing_401_14504 |  | B/C* | 2.10 | Yes | 0.60 | 0.60 |
| 8.S68C_SI_PeerReviewing_401_14505 |  | B/C | 2.84 | Yes | 0.51 | 0.51 |
| 9.S68C_LS_LuisRivera_401_V1_14751 |  | B/C* | 2.03 | Yes | 0.46 | 0.45 |
| 10.S68C_LS_LuisRivera_401_V1_14752 |  | B/C | 2.54 | Yes | 0.60 | 0.60 |
| 11.S68C_MS_Pedometers_401_V2_14757 |  | B/C* | 2.64 | Yes | 0.51 | 0.53 |
| 12.S68C_MS_Pedometers_401_V2_14758 |  | B/C | 2.58 | Yes | 0.52 | 0.52 |
| 13.S68P_SI_PeerReviewing_401_14510 |  | Pre-A** | -3.54 |  | 0.66 | 0.17 |
| 14.S68P_LS_LuisRivera_401_14745 |  | Pre-A** | -2.97 |  | 0.63 | 0.15 |
| 15.S68P_MS_Pedometers_401_14747 |  | Pre-A** | -3.13 |  | 0.64 | 0.14 |
| Raw Score Distribution by Task | Task | Raw Score |  |  |  |  |
|  |  | 0 | 1 | 2 | 3 | 4 |
|  | Task 1 | 1.17\% | 4.10\% | 94.73\% | N/A | N/A |
|  | Task 2 | 0.48\% | 13.22\% | 63.05\% | 20.97\% | 2.28\% |
|  | Task 3 | 1.22\% | 5.20\% | 93.58\% | N/A | N/A |
|  | Task 4 | 0.61\% | 8.84\% | 68.32\% | 20.50\% | 1.74\% |
|  | Task 5 | 1.01\% | 3.37\% | 95.62\% | N/A | N/A |
|  | Task 6 | 0.65\% | 28.21\% | 59.92\% | 10.00\% | 1.22\% |
|  | Task 7 | 0.48\% | 13.22\% | 63.05\% | 20.97\% | 2.28\% |
|  | Task 8 | 0.47\% | 23.24\% | 64.17\% | 11.19\% | 0.93\% |
|  | Task 9 | 0.61\% | 8.84\% | 68.32\% | 20.50\% | 1.74\% |
|  | Task 10 | 0.42\% | 18.43\% | 58.49\% | 20.42\% | 2.24\% |
|  | Task 11 | 0.65\% | 28.21\% | 59.92\% | 10.00\% | 1.22\% |
|  | Task 12 | 0.38\% | 20.10\% | 65.99\% | 12.21\% | 1.31\% |
|  | Task 13 | 1.17\% | 4.10\% | 94.73\% | N/A | N/A |
|  | Task 14 | 1.22\% | 5.20\% | 93.58\% | N/A | N/A |
|  | Task 15 | 1.01\% | 3.37\% | 95.62\% | N/A | N/A |

[^30]Table 3.3.4.4Biv
DIF Analysis and Summary: Spek 6-8 S401 Online n/a



Table 3.3.4.4Civ
Raw Score Descriptive Statistics: Spek 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 55,341 | 0 | 28 | 15.54 | 4.28 |
| $\mathbf{7}$ | 55,230 | 0 | 29 | 16.09 | 4.52 |
| $\mathbf{8}$ | 55,569 | 0 | 30 | 15.69 | 5.04 |
| Total | 166,140 | 0 | 30 | 15.77 | 4.63 |

Table 3.3.4.4Div
Scale Score Descriptive Statistics: Spek 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 55,341 | 136 | 447 | 317.56 | 40.48 |
| $\mathbf{7}$ | 55,230 | 142 | 459 | 320.63 | 44.33 |
| $\mathbf{8}$ | 55,569 | 148 | 471 | 324.75 | 49.18 |
| Total | 166,140 | 136 | 471 | 320.99 | 44.91 |



Table 3.3.4.4Eiv
Proficiency Level Distribution: Spek 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 5,097 | $9.21 \%$ | 5,814 | $10.53 \%$ | 8,975 | $16.15 \%$ | 19,886 | $11.97 \%$ |
| $\mathbf{2}$ | 16,878 | $30.50 \%$ | 13,693 | $24.79 \%$ | 12,509 | $22.51 \%$ | 43,080 | $25.93 \%$ |
| $\mathbf{3}$ | 28,686 | $51.83 \%$ | 29,050 | $52.60 \%$ | 28,581 | $51.43 \%$ | 86,317 | $51.95 \%$ |
| $\mathbf{4}$ | 4,586 | $8.29 \%$ | 6,599 | $11.95 \%$ | 5,350 | $9.63 \%$ | 16,535 | $9.95 \%$ |
| $\mathbf{5}$ | 94 | $0.17 \%$ | 69 | $0.12 \%$ | 149 | $0.27 \%$ | 312 | $0.19 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 5 | $0.01 \%$ | 5 | $0.01 \%$ | 10 | $0.01 \%$ |
| Total | 55,341 | $100.00 \%$ | 55,230 | $100.00 \%$ | 55,569 | $100.00 \%$ | 166,140 | $100.00 \%$ |

Table 3.3.4.4Fiv
Raw Score to Scale Score Conversion: Spek 6-8 S401 Online n/a

Table 3.3.4.4Giv
Equating Summary: Spek 6-8 S401 Online


[^31]

Table 3.3.4.4Jiv
Reliability: Spek 6-8 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| Pre-A | 10,708 | 0.719 | 0.705 |
| A | 38,630 | 0.622 |  |
| B/C | 116,802 | 0.731 |  |

Table 3.3.4.4Kiv
Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 S401 Online

| Proficiency Level | Grade | Cut Score | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Tier A | Tier B/C |
| 1/2 | 6 | 268 | 21.64 | 15.79 |
|  | 7 | 277 | 22.23 | 15.50 |
|  | 8 | 284 | 22.81 | 15.79 |
| 2/3 | 6 | 310 | 26.62 | 18.13 |
|  | 7 | 317 | 27.49 | 19.01 |
|  | 8 | 323 | 28.08 | 19.60 |
| 3/4 | 6 | 360 | 26.03 | 19.01 |
|  | 7 | 369 | 25.15 | 18.43 |
|  | 8 | 377 | 24.57 | 17.84 |
| 4/5 | 6 | 417 | 25.74 | 17.84 |
|  | 7 | 425 | 27.20 | 18.43 |
|  | 8 | 433 | 28.96 | 19.60 |
| 5/6 | 6 | 451 | 35.10 | 23.11 |
|  | 7 | 457 | 37.73 | 24.86 |
|  | 8 | 463 | 40.95 | 26.62 |

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficeny level higher than 2.

Table 3.3.4.4Li
Accuracy and Consistency of Classification Indices: Spek (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.681 | 0.546 |  | 0.276 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.737 |  | 0.569 |  |
|  | 2 | 0.676 |  | 0.496 |  |
|  | 3 | 0.676 |  | 0.641 |  |
|  | 4 | - |  | 0.149 |  |
|  | 5 | - |  | 0.870 |  |
|  | 6 | N/A |  | N/A |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.950 | 0.023 | 0.027 | 0.923 |
|  | 2/3 | 0.813 | 0.043 | 0.144 | 0.746 |
|  | 3/4 | 0.915 | 0.085 | 0.000 | 0.843 |
|  | 4/5 | 0.998 | 0.002 | 0.000 | 0.998 |
|  | 5/6 | N/A | N/A | N/A | N/A |

Table 3.3.4.4Lii
Accuracy and Consistency of Classification Indices: Spek (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.652 | 0.523 |  | 0.243 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.771 |  | 0.602 |  |
|  | 2 | 0.576 |  | 0.376 |  |
|  | 3 | 0.654 |  | 0.629 |  |
|  | 4 | - |  | 0.189 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.948 | 0.023 | 0.029 | 0.920 |
|  | 2/3 | 0.818 | 0.048 | 0.134 | 0.735 |
|  | 3/4 | 0.879 | 0.121 | 0.000 | 0.817 |
|  | 4/5 | 0.999 | 0.001 | 0.000 | 0.999 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.4.4Liii
Accuracy and Consistency of Classification Indices: Spek (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.663 | 0.525 |  | 0.272 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.837 |  | 0.690 |  |
|  | 2 | 0.495 |  | 0.318 |  |
|  | 3 | 0.665 |  | 0.638 |  |
|  | 4 | - |  | 0.151 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.935 | 0.023 | 0.042 | 0.902 |
|  | 2/3 | 0.815 | 0.047 | 0.138 | 0.729 |
|  | 3/4 | 0.901 | 0.099 | 0.000 | 0.829 |
|  | 4/5 | 0.997 | 0.003 | 0.000 | 0.997 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.3.5 Grades: 9-12

### 3.3.5.1 Listening 9-12

Table 3.3.5.1A
Complete Item Analysis and Summary: List 9-12 S401 Online

| Item Type | Average <br> Item <br> Difficulty <br> (in <br> logits) <br> in | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | 1.30 | 54 | 0.61 | 0.97 | 0.96 |
|  | Item Difficulty |  |  | Fit <br> Statistics |  |
| Name | (in <br> logits) | Anchored? | $\mathbf{P}$-value | Infit <br> Mnsq | Outfit <br> Mnsq |
| 1.L91A_SI_ArtContest_P100_Screen_2_12360 | -0.42 |  | 0.85 | 0.99 | 1.07 |
| 2.L91A_SI_ArtContest_P100_Screen_3_12361 | 2.15 |  | 0.43 | 1.17 | 1.42 |
| 3.L91A_SI_ArtContest_P100_Screen_4_12441 | 1.47 |  | 0.55 | 1.16 | 1.31 |
| 4.L91B_SI_RenewableEnergy_P100_Screen_2_12442 | 1.10 |  | 0.62 | 1.18 | 1.34 |
| 5.L91B_SI_RenewableEnergy_P 100_Screen_3_12443 | 2.22 |  | 0.45 | 1.29 | 1.54 |
| 6.L91B_SI_RenewableEnergy_P 100_Screen_4_12444 | 0.11 |  | 0.77 | 0.97 | 0.99 |
| 7.L91B_LA_Graduation_301_P100_A301_Screen_2_13856 | -0.50 |  | 0.63 | 0.86 | 0.80 |
| 8.L91B_LA_Graduation_301_P100_A301_Screen_3_13857 | -0.38 |  | 0.61 | 0.94 | 0.90 |
| 9.L91B_LA_Graduation_301_P100_A301_Screen_4_13858 | -0.41 |  | 0.63 | 0.98 | 0.95 |
| 10.L91A_MA_WaterSale_P100_Screen_2_12711 | 0.08 |  | 0.46 | 0.94 | 0.93 |
| 11.L91A_MA_WaterSale_P100_Screen_3_12712 | 0.65 |  | 0.34 | 0.94 | 0.93 |
| 12.L91A_MA_WaterSale_P100_Screen_4_12713 | 0.25 | Yes | 0.50 | 1.01 | 1.00 |
| 13.L91B_SS_GlobalCarProduction_P100_Screen_2_12869 | -0.48 |  | 0.61 | 0.94 | 0.92 |
| 14.L91B_SS_GlobalCarProduction_P100_Screen_3_12870 | 0.37 |  | 0.42 | 0.95 | 0.94 |
| 15.L91B_SS_GlobalCarProduction_P100_Screen_4_12871 | 1.19 |  | 0.27 | 1.02 | 1.06 |
| 16.L91A_SC_DesertPlants_P100_Screen_2_11034 | -0.40 |  | 0.59 | 0.94 | 0.92 |
| 17.L91A_SC_DesertPlants_P100_Screen_3_11035 | -0.52 |  | 0.61 | 0.93 | 0.91 |
| 18.L91A_SC_DesertPlants_P100_Screen_4_11036 | 0.51 |  | 0.39 | 0.96 | 0.95 |
| 19.L91B_LA_FreeVerse_P100_Screen_2_12887 | -0.46 | Yes | 0.84 | 0.84 | 0.66 |
| 20.L91B_LA_FreeVerse_P100_Screen_3_12889 | 0.57 | Yes | 0.66 | 0.92 | 0.88 |
| 21.L91B_LA_FreeVerse_P100_Screen_4_12890 | 1.50 | Yes | 0.54 | 0.98 | 0.96 |
| 22.L91B_MA_WaterSale_P100_Screen_2_12714 | 1.14 |  | 0.61 | 0.98 | 0.96 |
| 23.L91B_MA_WaterSale_P100_Screen_3_12715 | 1.89 | Yes | 0.41 | 0.95 | 0.93 |
| 24.L91B_MA_WaterSale_P100_Screen_4_12716 | 0.41 | Yes | 0.69 | 0.98 | 0.96 |
| 25.L91C_SS_USCities_301_P100_A301_Screen_2_13868 | 1.74 | Yes | 0.48 | 0.97 | 0.96 |


| Name | $\begin{array}{\|c} \text { Item } \\ \text { Difficulty } \\ \text { (in } \\ \text { logits) } \\ \hline \end{array}$ | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | Outfit <br> Mnsq |
| 26.L91C_SS_USCities_301_P100_A301_Screen_3_13869 | 1.45 | Yes | 0.56 | 0.91 | 0.89 |
| 27.L91C_SS_USCities_301_P100_A301_Screen_4_13870 | 2.21 | Yes | 0.40 | 0.97 | 0.96 |
| 28.L91B_SC_GrasshopperLifeCycle_P100_Screen_2_11198 | 1.98 | Yes | 0.45 | 1.04 | 1.04 |
| 29.L91B_SC_GrasshopperLifeCycle_P100_Screen_3_11199 | 0.40 | Yes | 0.76 | 0.87 | 0.79 |
| 30.L91B_SC_GrasshopperLifeCycle_P100_Screen_4_11215 | 1.44 | Yes | 0.52 | 0.95 | 0.94 |
| 31.L91B_LA_CharacterDev_P 100_alt 1_Screen_2_13844 | 0.10 |  | 0.80 | 0.95 | 0.92 |
| 32.L91B_LA_CharacterDev_P 100_alt 1_Screen_3_13845 | 1.12 | Yes | 0.65 | 0.91 | 0.89 |
| 33.L91B_LA_CharacterDev_P 100_alt 1_Screen_4_13846 | 1.21 | Yes | 0.56 | 0.94 | 0.93 |
| 34.L91B_MA_RunnersVelocity_P100_Screen_2_12525 | 0.40 | Yes | 0.72 | 0.98 | 0.96 |
| 35.L91B_MA_RunnersVelocity_P100_Screen_3_12527 | 0.26 | Yes | 0.74 | 0.95 | 0.92 |
| 36.L91B_MA_RunnersVelocity_P100_Screen_4_12530 | 2.24 | Yes | 0.36 | 1.01 | 1.02 |
| 37.L91C_LA_Poetry_P100_Screen_2_12893 | 0.07 |  | 0.93 | 0.98 | 0.97 |
| 38.L91C_LA_Poetry_P100_Screen_3_12894 | 2.82 | Yes | 0.55 | 0.96 | 0.95 |
| 39.L91C_LA_Poetry_P100_Screen_4_12895 | 2.83 |  | 0.57 | 1.05 | 1.07 |
| 40.L91C_MA_WaterSale_P100_Screen_2_12720 | 1.60 |  | 0.78 | 0.95 | 0.89 |
| 41.L91C_MA_WaterSale_P100_Screen_3_12903 | 1.40 |  | 0.81 | 0.92 | 0.84 |
| 42.L91C_MA_WaterSale_P100_Screen_4_12907 | 2.18 |  | 0.69 | 0.94 | 0.90 |
| 43.L91C_SS_GlobalCoffeeProduction_P100_Screen_2_13036 | 1.07 | Yes | 0.87 | 0.95 | 0.87 |
| 44.L91C_SS_GlobalCoffeeProduction_P100_Screen_3_13037 | 2.66 |  | 0.67 | 0.96 | 0.94 |
| 45.L91C_SS_GlobalCoffeeProduction_P100_Screen_4_13038 | 3.60 |  | 0.48 | 1.03 | 1.04 |
| 46.L91B_SC_PlantLifeCycle_P100_Screen_2_11245 | 1.21 |  | 0.88 | 0.94 | 0.84 |
| 47.L91B_SC_PlantLifeCycle_P100_Screen_3_11343 | 2.65 |  | 0.67 | 0.92 | 0.88 |
| 48.L91B_SC_PlantLifeCycle_P100_Screen_4_11344 | 3.91 |  | 0.44 | 0.99 | 0.98 |
| 49.L91C_LA_EagleFlies_301_P100_A301_Screen_2_13862 | 2.18 |  | 0.78 | 0.91 | 0.84 |
| 50.L91C_LA_EagleFlies_301_P100_A301_Screen_3_13863 | 3.10 |  | 0.64 | 0.96 | 0.96 |
| 51.L91C_LA_EagleFlies_301_P100_A301_Screen_4_13864 | 1.80 |  | 0.83 | 0.91 | 0.82 |
| 52.L91C_MA_angles_joga_P100_A202_Screen_2_13865 | 2.68 |  | 0.72 | 0.97 | 0.99 |
| 53.L91C_MA_angles_joga_P100_A202_Screen_3_13866 | 3.83 |  | 0.51 | 0.97 | 0.97 |
| 54.L91C_M A_angles_joga_P100_A202_Screen_4_13867 | 3.88 |  | 0.47 | 0.94 | 0.93 |

Table 3.3.5.1B
DIF Analysis and Summary: List 9-12 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| A | 25 | 29 | 29 | 21 |
| B | 0 | 0 | 1 | 3 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \hline \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.L91A_SI_ArtContest_P 100_Screen_2_12360 | A | M | A | O |
| 2.L91A_SI_ArtContest_P 100_Screen_3_12361 | A | M | A | O |
| 3.L91A_SI_ArtContest_P 100_Screen_4_12441 | A | F | A | H |
| 4.L9 IB_SI_Rene wable Energy_P 100_Screen_2_12442 | A | M | A | O |
| 5.L91B_SI_Rene wable Energy_P 100_Scree n_3_12443 | A | F | A | O |
| 6.L9 IB_SI_Rene wable Energy_P 100_Screen_4_12444 | A | M | A | H |
| 7.L9 1B_LA_Graduation_301_P 100_A301_Screen_2_13856 | A | F | A | H |
| 8.L91B_LA_Graduation_301_P 100_A301_Screen_3_13857 | A | M | A | H |
| 9.L91B_LA_Graduation_301_P 100_A301_Screen_4_13858 | A | M | B | O |
| 10.L91A_MA_WaterSale_P 100_Screen_2_12711 | A | F | A | H |
| 11.L91A_MA_WaterS ale_P 100_Screen_3_12712 | A | M | A | H |
| 12.L91A_MA_WaterSale_P 100_Screen_4_12713 | A | F | A | O |
| 13.L9 1B_SS_Glo balCarP roduction_P 100_Screen_2_12869 | A | F | A | H |
| 14.L9 1B_SS_Glo balCarP roduction_P 100_Screen_3_12870 | A | F | A | O |
| 15.L9 IB_SS_Glo balCarP roduction_P 100_Screen_4_12871 | A | F | A | O |
| 16.L91A_SC_DesertP lants_P 100_Screen_2_11034 | A | F | A | H |
| 17.L91A_SC_DesertP lants_P 100_Screen_3_11035 | A | M | A | O |
| 18.L91A_SC_DesertP lants_P 100_Screen_4_11036 | A | F | A | O |
| 19.L9 1B_LA_Free Verse_P 100_Screen_2_12887 | A | M | A | H |
| 20.1918 _LA_Free Verse_P 100_Screen_3_12889 | A | M | A | H |
| 21.L91B_LA_Free Verse_P 100_Screen_4_12890 | A | M | A | H |
| $22 . L 91 B$ _MA_WaterSale_P 100_Screen_2_12714 | A | M | A | O |
| 23.L9 1B_MA_WaterSale_P 100_Screen_3_12715 | A | F | A | H |
| 24.L9 1B_MA_WaterS ale_P 100_Screen_4_12716 | A | M | A | O |
| 25.L91C_SS_USCities_301_P 100_A301_Screen_2_13868 | A | F | A | O |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored <br> Group |
| 26.L91C_SS_USCitie s_301_P 100_A301_Screen_3_13869 | A | F | A | H |
| 27.L91C_SS_USCities_301_P 100_A301_Screen_4_13870 | A | M | A | O |
| 28.L9 1B_SC_Gras sho pperLife Cycle_P 100_Screen_2_1198 | A | F | A | O |
| 29.L9 1B_SC_Gras sho pperLife Cycle_P 100_Screen_3_11199 | A | F | A | H |
| 30.L9 1B_SC_Gras sho pperLife Cycle_P 100_Screen_4_112 15 | A | F | A | H |
| 31.L9 1B_LA_CharacterDev_P 100_alt I_Screen_2_13844 | A | M | A | O |
| 32.L9 1B_LA_CharacterDev_P 100_alt_Screen_3_13845 | A | M | A | O |
| 33.L9 1B_LA_CharacterDev_P 100_alt_Screen_4_13846 | A | M | A | H |
| 34.L9 IB_MA_Runners Velocity_P 100_Screen_2_12525 | A | F | A | H |
| 35.L9 1B_MA_Runners Velocity_P 100_Screen_3_12527 | A | F | A | H |
| 36.L9 1B_MA_Runners Velocity_P 100_Screen_4_12530 | A | M | A | H |
| 37.L91C_LA_P oetry_P 100_Screen_2_12893 | A | F | B | O |
| 38.L91C_LA_P oetry_P 100_Screen_3_12894 | A | M | A | H |
| 39.L91C_LA_P oetry_P 100_Screen_4_12895 | A | F | A | O |
| 40.L91C_MA_WaterSale_P 100_Screen_2_12720 | A | F | A | O |
| 41.L91C_MA_WaterSale_P 100_Screen_3_12903 | A | M | A | H |
| 42.L91C_MA_WaterSale_P 100_Screen_4_12907 | A | F | A | H |
| 43.L91C_SS_Glo balCoffeeP roduction_P 100_Screen_2_13036 | A | F | A | H |
| 44.L91C_SS_Glo balCoffeeP roduction_P 100_Screen_3_13037 | A | M | A | H |
| 45.L91C_SS_Glo balCoffeeP roduction_P 100_Screen_4_13038 | A | M | A | H |
| 46.L9 IB_SC_P lantLifeCycle_P 100_Screen_2_11245 | A | F | A | O |
| 47.L9 1B_SC_P lantLifeCycle_P 100_Screen_3_11343 | A | F | A | H |
| 48.L9 IB_SC_P lantLife Cycle_P 100_Screen_4_11344 | A | F | A | O |
| 49.L91C_LA_EagleFlies_301_P 100_A301_Screen_2_13862 | A | F | A | H |
| 50.L91C_LA_EagleFlies_301_P 100_A301_Screen_3_13863 | A | M | A | H |
| 51.L91C_LA_EagleFlies_301_P 100_A301_Screen_4_13864 | A | M | A | O |
| 52.L9 1C_MA_angles _joga_P 100_A202_Screen_2_13865 | A | F | B | O |
| 53.L91C_MA_angles _joga_P 100_A202_Screen_3_13866 | A | M | A | H |
| 54.L9 1C_MA_a agle _joga_P 100_A202_Screen_4_13867 | A | F | B | H |

Figure 3.3.5.1C
Raw Scores: List 9-12 S401 Online
n/a

Table 3.3.5.1C
Raw Score Descriptive Statistics: List 9-12 S401 Online n/a



Table 3.3.5.1D
Scale Score Descriptive Statistics: List 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 77,726 | 136 | 527 | 382.96 | 54.82 |
| $\mathbf{1 0}$ | 52,056 | 140 | 527 | 383.51 | 53.02 |
| $\mathbf{1 1}$ | 37,047 | 214 | 527 | 385.78 | 52.25 |
| $\mathbf{1 2}$ | 24,783 | 214 | 527 | 388.18 | 51.19 |
| Total | 191,612 | 136 | 527 | 384.33 | 53.41 |

Table 3.3.5.1E
Proficiency Level Distribution: List 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 8,924 | $11.48 \%$ | 7,916 | $15.21 \%$ | 6,472 | $17.47 \%$ | 4,834 | $19.51 \%$ | 28,146 | $14.69 \%$ |
| $\mathbf{2}$ | 13,792 | $17.74 \%$ | 7,215 | $13.86 \%$ | 5,155 | $13.91 \%$ | 3,872 | $15.62 \%$ | 30,034 | $15.67 \%$ |
| $\mathbf{3}$ | 19,215 | $24.72 \%$ | 15,500 | $29.78 \%$ | 10,448 | $28.20 \%$ | 7,051 | $28.45 \%$ | 52,214 | $27.25 \%$ |
| $\mathbf{4}$ | 13,037 | $16.77 \%$ | 7,841 | $15.06 \%$ | 6,065 | $16.37 \%$ | 3,269 | $13.19 \%$ | 30,212 | $15.77 \%$ |
| $\mathbf{5}$ | 6,678 | $8.59 \%$ | 4,887 | $9.39 \%$ | 2,953 | $7.97 \%$ | 2,461 | $9.93 \%$ | 16,979 | $8.86 \%$ |
| $\mathbf{6}$ | 16,080 | $20.69 \%$ | 8,697 | $16.71 \%$ | 5,954 | $16.07 \%$ | 3,296 | $13.30 \%$ | 34,027 | $17.76 \%$ |
| Total | 77,726 | $100.00 \%$ | 52,056 | $100.00 \%$ | 37,047 | $100.00 \%$ | 24,783 | $100.00 \%$ | 191,612 | $100.00 \%$ |

Table 3.3.5.1F
Raw Score to Scale Score Conversion: List 9-12 S401 Online n/a

Table 3.3.5.1G
Equating Summary: List 9-12 S401 Online


[^32]Figure 3.3.5.1H
Test Characteristic Curve: List 9-12 S401 Online n/a

Figure 3.3.5.1I


Ability Measure

Table 3.3.5.1J
Reliability: List 9-12 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 191,612 | 54 | .87 |

Table 3.3.5.1 K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 9-12 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 9 | 314 | N/A | N/A | N/A | N/A | N/A |
|  | 10 | 325 | 74 | 17.86 | 19.39 | 19.33 | 0.30 |
|  | 11 | 335 | 248 | 16.33 | 19.90 | 19.44 | 0.26 |
|  | 12 | 342 | 545 | 16.33 | 19.39 | 19.37 | 0.22 |
| 2/3 | 9 | 353 | N/A | N/A | N/A | N/A | N/A |
|  | 10 | 358 | 1,532 | 16.33 | 16.33 | 16.33 | 0.00 |
|  | 11 | 364 | 207 | 16.33 | 16.84 | 16.46 | 0.22 |
|  | 12 | 368 | N/A | N/A | N/A | N/A | N/A |
| 3/4 | 9 | 383 | 352 | 16.33 | 16.84 | 16.83 | 0.04 |
|  | 10 | 389 | 39 | 16.84 | 17.86 | 17.67 | 0.32 |
|  | 11 | 394 | 2 | 18.37 | 18.37 | 18.37 | 0.00 |
|  | 12 | 398 | 158 | 17.35 | 18.88 | 17.58 | 0.54 |
| 4/5 | 9 | 409 | 240 | 17.86 | 17.86 | 17.86 | 0.00 |
|  | 10 | 415 | 285 | 18.37 | 19.39 | 19.22 | 0.38 |
|  | 11 | 420 | 59 | 18.88 | 20.41 | 19.45 | 0.75 |
|  | 12 | 426 | 171 | 18.37 | 20.92 | 19.50 | 1.07 |
| 5/6 | 9 | 434 | 108 | 18.37 | 19.39 | 18.41 | 0.19 |
|  | 10 | 441 | 91 | 19.39 | 22.45 | 20.05 | 0.72 |
|  | 11 | 447 | 111 | 20.92 | 20.92 | 20.92 | 0.00 |
|  | 12 | 452 | 128 | 20.92 | 20.92 | 20.92 | 0.00 |

Table 3.3.5.1 Li
Accuracy and Consistency of Classification Indices: List (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.585 | 0.483 |  | 0.372 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.742 |  | 0.567 |  |
|  | 2 | 0.495 |  | 0.396 |  |
|  | 3 | 0.569 |  | 0.464 |  |
|  | 4 | 0.471 |  | 0.360 |  |
|  | 5 | 0.345 |  | 0.245 |  |
|  | 6 | 0.897 |  | 0.807 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.927 | 0.022 | 0.050 | 0.899 |
|  | 2/3 | 0.886 | 0.067 | 0.047 | 0.841 |
|  | 3/4 | 0.880 | 0.058 | 0.062 | 0.837 |
|  | 4/5 | 0.921 | 0.037 | 0.042 | 0.882 |
|  | 5/6 | 0.938 | 0.043 | 0.019 | 0.912 |

Table 3.3.5.1 Lii
Accuracy and Consistency of Classification Indices: List (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.586 | 0.481 |  | 0.370 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.788 |  | 0.635 |  |
|  | 2 | 0.387 |  | 0.298 |  |
|  | 3 | 0.651 |  | 0.543 |  |
|  | 4 | 0.440 |  | 0.333 |  |
|  | 5 | 0.412 |  | 0.295 |  |
|  | 6 | 0.886 |  | 0.791 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.918 | 0.026 | 0.056 | 0.885 |
|  | 2/3 | 0.883 | 0.078 | 0.039 | 0.837 |
|  | 3/4 | 0.882 | 0.053 | 0.065 | 0.839 |
|  | 4/5 | 0.928 | 0.037 | 0.035 | 0.891 |
|  | 5/6 | 0.946 | 0.037 | 0.017 | 0.924 |

Table 3.3.5.1 Liii
Accuracy and Consistency of Classification Indices: List (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.586 | 0.482 |  | 0.372 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.786 |  | 0.648 |  |
|  | 2 | 0.380 |  | 0.289 |  |
|  | 3 | 0.621 |  | 0.515 |  |
|  | 4 | 0.479 |  | 0.366 |  |
|  | 5 | 0.365 |  | 0.258 |  |
|  | 6 | 0.899 |  | 0.801 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.913 | 0.033 | 0.054 | 0.875 |
|  | 2/3 | 0.881 | 0.077 | 0.041 | 0.835 |
|  | 3/4 | 0.881 | 0.057 | 0.062 | 0.839 |
|  | 4/5 | 0.931 | 0.032 | 0.037 | 0.894 |
|  | 5/6 | 0.948 | 0.038 | 0.014 | 0.928 |

Table 3.3.5.1 Liv
Accuracy and Consistency of Classification Indices: List (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.587 | 0.481 |  | 0.368 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.758 |  | 0.629 |  |
|  | 2 | 0.391 |  | 0.303 |  |
|  | 3 | 0.637 |  | 0.523 |  |
|  | 4 | 0.424 |  | 0.316 |  |
|  | 5 | 0.481 |  | 0.353 |  |
|  | 6 | 0.864 |  | 0.760 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.900 | 0.044 | 0.056 | 0.856 |
|  | 2/3 | 0.874 | 0.079 | 0.047 | 0.827 |
|  | 3/4 | 0.895 | 0.045 | 0.060 | 0.853 |
|  | 4/5 | 0.935 | 0.039 | 0.027 | 0.903 |
|  | 5/6 | 0.952 | 0.032 | 0.016 | 0.934 |

### 3.3.5.2 Reading 9-12

Table 3.3.5.2A
Complete Item Analysis and Summary: Read 9-12 S401 Online

| Item Type | Average Item Difficulty (in logits) | No. of Items | Average P-value | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Selected Response | 1.84 | 72 | 0.53 | 0.98 | 0.98 |
|  | Item Difficulty |  |  | Fit <br> Statistics |  |
| Name | $\begin{gathered} \text { (in } \\ \text { logits) } \end{gathered}$ | Anchored? | P-value | $\begin{gathered} \hline \text { Infit } \\ \text { Mnsq } \end{gathered}$ | Outfit <br> Mnsq |
| 1.R91A_SI_JobSearch_301_P100_A301FT_Screen_2_13743 | 0.26 |  | 0.74 | 1.12 | 1.46 |
| 2.R91A_SI_JobSearch_301_P100_A301FT_Screen_3_13744 | 0.60 |  | 0.69 | 0.87 | 0.80 |
| 3.R91A_SI_JobSearch_301_P100_A301FT_Screen_4_13745 | 1.04 |  | 0.62 | 1.14 | 1.29 |
| 4.R91B_SI_ChoosingCollege_401_V1_Screen_2_13950 | 1.41 |  | 0.57 | 1.14 | 1.35 |
| 5.R91B_SI_ChoosingCollege_401_V1_Screen_3_13951 | 2.69 |  | 0.33 | 1.19 | 1.53 |
| 6.R91B_SI_ChoosingCollege_401_V1_Screen_4_13952 | 1.92 |  | 0.47 | 1.28 | 1.46 |
| 7.R91A_LA_CharlesSchulz_203_P100_A301_Screen_2_13674 | -1.77 |  | 0.87 | 0.90 | 0.67 |
| 8.R91A_LA_CharlesSchulz_203_P100_A301_Screen_3_13675 | 0.12 |  | 0.56 | 1.00 | 1.00 |
| 9.R91A_LA_CharlesSchulz_203_P100_A301_Screen_4_13676 | 1.32 | Yes | 0.35 | 1.01 | 1.04 |
| 10.R91A_MA_DrawingShapes_kaje_P100_A203_Screen_2_13677 | -0.32 | Yes | 0.65 | 0.95 | 0.93 |
| 11.R91A_MA_DrawingShapes_kaje_P100_A203_Screen_3_13678 | 0.16 |  | 0.53 | 1.00 | 1.01 |
| 12.R91A_MA_DrawingShapes_kaje_P100_A203_Screen_4_13679 | 0.86 | Yes | 0.40 | 0.89 | 0.88 |
| 13.R91A_SS_AviationHistory_P100_A203_Screen_2_13680 | -0.46 | Yes | 0.72 | 0.96 | 0.92 |
| 14.R91A_SS_AviationHistory_P100_A203_Screen_3_13681 | 0.55 | Yes | 0.48 | 1.00 | 0.99 |
| 15.R91A_SS_AviationHistory_P100_A203_Screen_4_13682 | 0.77 | Yes | 0.40 | 1.00 | 0.99 |
| 16.R91B_SC_FindingthepHlevel_keto_P100_A301_L_Screen_2_13701 | 0.31 | Yes | 0.59 | 0.99 | 0.98 |
| 17.R91B_SC_FindingthepHlevel_keto_P100_A301_L_Screen_3_13702 | 1.24 |  | 0.31 | 1.06 | 1.10 |
| 18.R91B_SC_FindingthepHlevel_keto_P100_A301_L_Screen_4_13703 | 1.51 | Yes | 0.29 | 0.97 | 0.98 |
| 19.R91B_LA_Auntie_203_P100_A 301 _Screen_2_13692 | 0.12 | Yes | 0.58 | 0.91 | 0.88 |
| 20.R91B_LA_Auntie_203_P100_A 301_Screen_3_13693 | 1.19 | Yes | 0.37 | 0.96 | 0.96 |
| 21.R91B_LA_Auntie_203_P100_A 301_Screen_4_13694 | 1.14 |  | 0.32 | 0.94 | 0.93 |
| 22.R91A_MA_RoadSigns_P100_A203_Screen_2_13689 | 0.26 | Yes | 0.58 | 0.94 | 0.92 |
| 23.R91A_MA_RoadSigns_P100_A 203 _Screen_3_13690 | 0.10 |  | 0.57 | 0.93 | 0.91 |
| 24.R91A_MA_RoadSigns_P100_A203_Screen_4_13691 | 1.49 |  | 0.27 | 1.07 | 1.14 |
| 25.R91A_LA_M iltonHershey_203_P100_A301_Screen_2_13686 | -1.14 |  | 0.92 | 0.83 | 0.45 |


| Name | ItemDifficulty(inlogits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Infit <br> Mnsq | Outfit <br> Mnsq |
| 26.R91A_LA_MiltonHershey_203_P10__A301_Screen_3_13687 | 1.08 |  | 0.62 | 0.93 | 0.90 |
| 27.R91A_LA_MiltonHershey_203_P100_A 301 _Screen_4_13688 | 1.19 |  | 0.62 | 0.85 | 0.79 |
| 28.R91B_MA_CakeM aking_JaGeKaKo_P 100_A203_alt __Sreen_2_13767 | -0.89 |  | 0.91 | 0.93 | 0.80 |
| 29.R91B_MA_CakeMaking_JaGeKaKo_P100_A203_alt I_Screen_3_13768 | 2.02 | Yes | 0.42 | 1.00 | 0.99 |
| 30.R91B_MA_CakeM aking_JaGeKaKo_P100_A203_altı_Screen_4_13769 | 2.14 | Yes | 0.41 | 0.98 | 0.97 |
| 31.R91B_SS_Reading_203_P100_A 301 _Screen_2_13698 | 1.20 |  | 0.62 | 1.02 | 1.03 |
| 32.R91B_SS_Reading_203_P100_A 301_Screen_3_13699 | 1.48 | Yes | 0.58 | 0.94 | 0.92 |
| 33.R91B_SS_Reading_203_P100_A 301 _Scren_4_13700 | 2.04 | Yes | 0.42 | 0.98 | 0.97 |
| 34.R911_SC_SpiceEffects_kaje_P100_A 301 _Screen_2_13713 | 0.98 |  | 0.69 | 0.92 | 0.90 |
| 35.R91B_SC_Spiceffect_kaje_P100_A301_Srree_3_13714 | 1.89 | Yes | 0.54 | 0.92 | 0.91 |
| 36.R91B_SC_SpiceEffects_kaje_P100_A301_Screen_4_13715 | 2.86 | Yes | 0.23 | 1.00 | 1.00 |
| 37.R91B_LA_Mattie_203_P10__A301_Screen_2_13704 | 0.80 | Yes | 0.70 | 0.89 | 0.84 |
| 38.R91B_LA_Mattie_203_P100_A301_Screen_3_13705 | 1.29 | Yes | 0.59 | 0.91 | 0.89 |
| 39.R91B_LA_Mattie_203_P100_A301_Screen_4_13706 | 1.94 | Yes | 0.45 | 0.96 | 0.95 |
| 40.R91B_MA_CircumferenceandDiameter_ryroP100_A203_Screen_2_13707 | 1.60 | Yes | 0.54 | 0.93 | 0.92 |
| 41. R91B_MA_CircumferenceandDiameter_ryroP100_A203_Screen_3_13708 | 2.16 |  | 0.45 | 0.95 | 0.94 |
| 42.R91B_MA_CircumferenceandDiameter_ryo_P100_A203_Screen_4_13709 | 2.72 |  | 0.33 | 0.99 | 0.99 |
| 43.R91B_SS_Recycling_203_P100_A301_alt _Screen_2_13770 | 2.46 |  | 0.38 | 0.98 | 0.98 |
| 44.R91B_SS_Recycling_203_P100_A301_alt _Screen_3_13771 | 2.32 | Yes | 0.41 | 0.94 | 0.93 |
| 45.R91B_SS_Recycling_203_P100_A301_alt _Screen_4_13772 | 1.83 |  | 0.49 | 0.94 | 0.93 |
| 46.R91C_SC_Diffusion_elle_P100_A301_Screen_2_13728 | 1.59 | Yes | 0.57 | 0.95 | 0.95 |
| 47-R91C_SC_Diffusion_elleP100_A301_Screen_3_13729 | 2.30 | Yes | 0.34 | 0.99 | 0.99 |
| 48.R91C_SC_Diffusion_elle_P100_A301_Screen_4_13730 | 2.86 |  | 0.29 | 1.03 | 1.04 |
| 49.R91C_LA_EdmundH_203_P100_A301_Screen_2_13719 | 1.46 |  | 0.76 | 0.92 | 0.86 |
| 50 R91C_LA_EdmundH_203_P100_A301_Screen_3_13720 | 3.13 | Yes | 0.51 | 0.99 | 0.98 |
| 51.R91C_LA_EdmundH_203_P100_A301_Screen_4_13721 | 3.18 | Yes | 0.50 | 1.14 | 1.19 |
| $52 . \mathrm{R} 91 \mathrm{C}$ _MA_Functions_401_V1_Screen_2_13968 | 2.61 |  | 0.61 | 1.03 | 1.07 |
| 53.R91C_MA_Functions_401_V1_Screen_3_13969 | 3.47 |  | 0.45 | 1.03 | 1.04 |
| 54.R91C_MA_Functions_401_V1_Screen_4_13970 | 2.85 |  | 0.57 | 1.06 | 1.09 |
| 55.R91C_SS_Mesopotami__401_V2_Screen_2_14634 | 2.09 |  | 0.72 | 0.94 | 0.94 |
| 56.R91C_SS_Mesopotami_-401_V2_Screen_3_14635 | 3.35 |  | 0.49 | 0.93 | 0.91 |
| 57.R91C_SS_Mesopotami__401_V2_Screen_4_14636 | 3.11 |  | 0.56 | 0.98 | 0.97 |
| $58 . \mathrm{R91C}$ _SC_pHScale_P100_A201_Scren_2_13785 | 2.54 | Yes | 0.64 | 0.93 | 0.90 |
| 59.R91C_SC_pHScale_P100_A201_Screen_3_13786 | 4.30 |  | 0.34 | 0.99 | 1.00 |
| $60 . \mathrm{R91C}$ _SC_pHScale_P100_A201_Screen_4_13787 | 3.69 |  | 0.46 | 1.01 | 1.01 |


| Name | Item Difficulty (in logits) | Anchored? | P-value | Fit <br> Statistics |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \\ \hline \end{gathered}$ | Outfit <br> Mnsq |
| 61.R91C_LA_Malta_203_P100_A 301_L_Screen_2_13731 | 3.95 |  | 0.41 | 1.04 | 1.04 |
| 62.R91C_LA_Malta_203_P100_A 301_L_Screen_3_13732 | 2.92 | Yes | 0.58 | 0.90 | 0.87 |
| 63.R91C_LA_Malta_203_P100_A 301_L_Screen_4_13733 | 2.87 |  | 0.65 | 0.97 | 0.96 |
| 64.R91C_MA_CircumferenceandDiameter_ry ro_P100_A 203_Screen_2_13722 | 2.50 |  | 0.73 | 0.92 | 0.87 |
| 65.R91C_MA_CircumferenceandDiameter_ry ro_P100_A 203_Screen_3_13723 | 2.82 |  | 0.67 | 0.91 | 0.87 |
| 66.R91C_MA_CircumferenceandDiameter_ry ro_P100_A203_Screen_4_13724 | 3.96 |  | 0.42 | 0.99 | 0.97 |
| 67.R91C_SS_SongDy nasty_401_V1_Screen_2_13956 | 3.92 |  | 0.44 | 0.97 | 0.96 |
| 68.R91C_SS_SongDy nasty_401_V1_Screen_3_13957 | 4.58 |  | 0.32 | 1.03 | 1.02 |
| 69.R91C_SS_SongDy nasty_401_V1_Screen_4_13958 | 4.25 |  | 0.38 | 1.02 | 1.01 |
| 70.R91C_SC_PolymerBall_kaje_P100_A301_Screen_2_13740 | 2.76 |  | 0.70 | 0.93 | 0.89 |
| 71.R91C_SC_Poly merBall_kaje_P100_A301_Screen_3_13741 | 3.30 |  | 0.59 | 0.92 | 0.89 |
| 72.R91C_SC_Poly merBall_kaje_P100_A301_Screen_4_13742 | 3.87 |  | 0.48 | 0.94 | 0.93 |

Table 3.3.5.2B
DIF Analysis and Summary: Read 9-12 S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (0) |
| A | 33 | 38 | 37 | 35 |
| B | 1 | 0 | 0 | 0 |
| C | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.R91A_SI_JobSearch_301_P 100_A301FT_Screen_2_13743 | A | M | A | H |
| 2.R91A_SI_JobSearch_301_P 100_A301FT_Screen_3_13744 | A | M | A | H |
| 3.R91A_SI_JobSearch_301_P 100_A301FT_Screen_4_13745 | A | M | A | H |
| 4.R9 1B_SI_Choos singCo llege_401_V1_Screen_2_13950 | A | M | A | H |
| $5 . \mathrm{R} 9$ 1B_SI_Choos ingCo llege_401_V1_Screen_3_13951 | A | F | A | H |
| 6.R9 1B_SI_Choos ingCollege_401_V1_Screen_4_13952 | A | F | A | O |
| 7.R91A_LA_Charle S Schulz_203_P 100_A301_Screen_2_13674 | A | M | A | O |
| 8.R91A_LA_Charles Schulz_203_P 100_A 301_Screen_3_13675 | A | M | A | H |
| 9.R91A_LA_Charle S Schulz_203_P 100_A301_Screen_4_13676 | A | F | A | H |
| 10.R91A_MA_DrawingShapes_kaje_P 100_A203_Screen_2_13677 | A | M | A | H |
| 11.R91A_MA_DrawingShapes_kaje_P 100_A203_Screen_3_13678 | A | F | A | H |
| 12.R91A_MA_DrawingShapes_kaje_P 100_A203_Screen_4_13679 | A | F | A | O |
| 13.R91A_SS_AviationHistory_P 100_A203_Screen_2_13680 | A | M | A | H |
| 14.R91A_SS_AviationHis tory_P 100_A203_Screen_3_13681 | A | F | A | H |
| 15.R91A_SS_AviationHistory_P 100_A 203_Screen_4_13682 | A | M | A | O |
| 16.R9 1B_SC_FindingthepHle vel_keto_P 100_A301_L_Screen_2_13701 | A | M | A | H |
| 17.R91B_SC_FindingthepHle vel_keto_P 100_A301_L_Screen_3_13702 | A | F | A | H |
| 18.R91B_SC_FindingthepHle vel_keto_P 100_A301_L_Screen_4_13703 | A | F | A | O |
| 19.R91B_LA_Auntie_203_P 100_A301_Screen_2_13692 | A | M | A | O |
| 20.R91B_LA_Auntie_203_P 100_A301_Screen_3_13693 | A | M | A | O |
| 21.R91B_LA_Auntie_203_P 100_A301_Screen_4_13694 | A | M | A | O |
| 22.R91A_MA_RoadS igns_P 100_A203_Screen_2_13689 | A | F | A | H |
| 23.R91A_MA_RoadSigns_P 100_A 203 _Screen_3_13690 | A | F | A | H |
| 24.R91A_MA_RoadS igns_P 100_A203_Screen_4_13691 | A | F | A | H |
| 25.R91A_LA_MiltonHers hey_203_P 100_A 301 _Screen_2_13686 | A | M | A | H |


| Name | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 26.R91A_LA_Milto nHers hey_203_P 100_A301_Screen_3_13687 | A | F | A | O |
| 27.R91A_LA_MiltonHers hey_203_P 100_A301_Screen_4_13688 | A | F | A | H |
| 28.R91B_MA_CakeMaking_JaGeKaKo_P 100_A203_alt1_Screen_2_13767 | A | F | A | H |
| 29.R91B_MA_CakeMaking_J a Ge KaKo_P 100_A203_altI_Screen_3_13768 | A | F | A | H |
| 30.R91B_MA_CakeMaking_J a Ge KaKo_P 100_A 203_altI_Screen_4_13769 | A | M | A | H |
| 31.R91B_SS_Reading_203_P 100_A 301 _Screen_2_13698 | A | F | A | H |
| 32.R91B_SS_Reading_203_P 100_A301_Screen_3_13699 | A | M | A | H |
| 33.R91B_SS_Reading_203_P 100_A301_Screen_4_13700 | A | F | A | O |
| $34 . \mathrm{R9} 9 \mathrm{~B}$ _SC_Spice Effects_kaje_P 100_A 301_Screen_2_13713 | A | M | A | O |
| 35.R91B_SC_SpiceEffects_kaje_P 100_A301_Screen_3_13714 | A | M | A | O |
| 36.R91B_SC_SpiceEffects_kaje_P 100_A 301_Screen_4_13715 | A | M | A | O |
| 37.R91B_LA_Mattie_203_P 100_A301_Screen_2_13704 | A | M | A | O |
| 38.R91B_LA_Mattie_203_P 100_A 301_Screen_3_13705 | A | M | A | O |
| 39.R91B_LA_Mattie_203_P 100_A 301_Screen_4_13706 | A | F | A | O |
| 40.R91B_MA_Circum ferenceandDiameter_ryro_P 100_A203_Screen_2_13707 | A | F | A | O |
| 41.R91B_MA_CircumferenceandDiameter_ryro_P 100_A203_Screen_3_13708 | A | M | A | O |
| 42.R91B_MA_Circum ferenceandDiameter_ryro_P 100_A203_Screen_4_13709 | A | F | A | H |
| 43.R91B_SS_Recycling_203_P 100_A301_altl_Screen_2_13770 | A | M | A | O |
| 44.R91B_SS_Recycling_203_P 100_A301_alt_Screen_3_13771 | A | F | A | H |
| 45.R91B_SS_Recycling_203_P 100_A301_altıScreen_4_13772 | A | M | A | O |
| 46.R91C_SC_Diffus ion_elle_P 100_A301_Screen_2_13728 | A | M | A | O |
| 47.R91C_SC_Diffus ion_elle_P 100_A301_Screen_3_13729 | A | M | A | O |
| 48.R91C_SC_Diffus io n_elle_P 100_A301_Screen_4_13730 | A | F | A | H |
| 49.R91C_LA_EdmundH_203_P 100_A 301_Screen_2_13719 | A | F | A | H |
| 50.R91C_LA_EdmundH_203_P 100_A 301 _Screen_3_13720 | A | F | A | O |
| $51 . \mathrm{P} 91 \mathrm{C}$ _LA_EdmundH_203_P 100_A301_Screen_4_13721 | A | M | A | O |
| $52 . \mathrm{R91C}$ _MA_Functions_401_V1_Screen_2_13968 | B | M | A | H |
| $53 . \mathrm{R91C}$ _MA_Functions_401_V1_Screen_3_13969 | A | F | A | O |
| 54.R91C_MA_Functions_401_V1_Screen_4_13970 | A | M | A | H |
| 55.R91C_SS_Mesopotamia_401_V2_Screen_2_14634 | A | F | A | O |
| 56.R91C_SS_Mesopotamia_401_V2_Screen_3_14635 | A | F | A | H |
| 57.R91C_SS_Mesopotamia_401_V2_Screen_4_4636 | A | F | A | O |
| 58.R91C_SC_pHScale_P 100_A201_Screen_2_13785 | A | F | A | O |
| 59.R91C_SC_pHScale_P 100_A201_Screen_3_13786 | A | F | A | O |
| $60 . \mathrm{R91C}$ _SC_pHScale_P 100_A201_Screen_4_13787 | A | F | A | H |


|  | Male/Female |  | Hispanic/Other |  |
| :--- | :---: | :---: | :---: | :---: |
|  | Name | $\begin{array}{c}\text { DIF } \\ \text { Level }\end{array}$ | $\begin{array}{c}\text { Favored } \\ \text { Group }\end{array}$ | $\begin{array}{c}\text { DIF } \\ \text { Level }\end{array}$ |
| Favored |  |  |  |  |
| Group |  |  |  |  |$]$

Figure 3.3.5.2C
Raw Scores: Read 9-12 S401 Online n/a

Table 3.3.5.2C
Raw Score Descriptive Statistics: Read 9-12 S401 Online n/a


Figure 3.3.5.2E
Proficiency Level: Read 9-12 S401 Online


Table 3.3.5.2D
Scale Score Descriptive Statistics: Read 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 65,870 | 208 | 491 | 363.10 | 38.59 |
| $\mathbf{1 0}$ | 42,507 | 250 | 491 | 367.67 | 36.39 |
| $\mathbf{1 1}$ | 28,833 | 250 | 491 | 372.55 | 36.93 |
| $\mathbf{1 2}$ | 19,870 | 250 | 491 | 375.59 | 35.58 |
| Total | 157,080 | 208 | 491 | 367.65 | 37.61 |

Table 3.3.5.2E
Proficiency Level Distribution: Read 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 21,306 | $32.35 \%$ | 12,531 | $29.48 \%$ | 7,962 | $27.61 \%$ | 4,925 | $24.79 \%$ | 46,724 | $29.75 \%$ |
| $\mathbf{2}$ | 18,917 | $28.72 \%$ | 13,538 | $31.85 \%$ | 10,221 | $35.45 \%$ | 7,855 | $39.53 \%$ | 50,531 | $32.17 \%$ |
| $\mathbf{3}$ | 10,155 | $15.42 \%$ | 7,001 | $16.47 \%$ | 3,873 | $13.43 \%$ | 2,794 | $14.06 \%$ | 23,823 | $15.17 \%$ |
| $\mathbf{4}$ | 2,627 | $3.99 \%$ | 1,898 | $4.47 \%$ | 1,474 | $5.11 \%$ | 1,007 | $5.07 \%$ | 7,006 | $4.46 \%$ |
| $\mathbf{5}$ | 6,614 | $10.04 \%$ | 4,219 | $9.93 \%$ | 2,892 | $10.03 \%$ | 1,949 | $9.81 \%$ | 15,674 | $9.98 \%$ |
| $\mathbf{6}$ | 6,251 | $9.49 \%$ | 3,320 | $7.81 \%$ | 2,411 | $8.36 \%$ | 1,340 | $6.74 \%$ | 13,322 | $8.48 \%$ |
| Total | 65,870 | $100.00 \%$ | 42,507 | $100.00 \%$ | 28,833 | $100.00 \%$ | 19,870 | $100.00 \%$ | 157,080 | $100.00 \%$ |

Table 3.3.5.2F
Raw Score to Scale Score Conversion: Read 9-12 S401 Online n/a

Table 3.3.5.2G
Equating Summary: Read 9-12 S401 Online


[^33]Figure 3.3.5.2H
Test Characteristic Curve: Read 9-12 S401 Online n/a

Figure 3.3.5.2I


Ability Measure

Table 3.3.5.2J
Reliability: Read 9-12 S401 Online

| No. of Students | No. of Items | Rasch <br> Reliability <br> Estimate |
| :---: | :---: | :---: |
| 157,080 | 72 | .91 |

Table 3.3.5.2K
Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 9-12 S401 Online

| Proficiency Level | Grade | Cut Score | No. of Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1/2 | 9 | 340 | 455 | 11.22 | 12.24 | 11.24 | 0.12 |
|  | 10 | 344 | 75 | 11.22 | 12.24 | 11.36 | 0.29 |
|  | 11 | 348 | 475 | 11.22 | 12.24 | 11.23 | 0.07 |
|  | 12 | 352 | 674 | 10.71 | 11.73 | 11.23 | 0.04 |
| 2/3 | 9 | 372 | 184 | 10.20 | 10.71 | 10.25 | 0.15 |
|  | 10 | 377 | 1,168 | 10.20 | 11.22 | 10.21 | 0.07 |
|  | 11 | 382 | 580 | 10.20 | 11.22 | 10.22 | 0.14 |
|  | 12 | 386 | 591 | 10.20 | 11.73 | 10.24 | 0.16 |
| 3/4 | 9 | 392 | 49 | 10.20 | 11.73 | 10.94 | 0.38 |
|  | 10 | 397 | 27 | 10.20 | 10.71 | 10.34 | 0.23 |
|  | 11 | 402 | 71 | 10.20 | 11.73 | 10.91 | 0.35 |
|  | 12 | 407 | 339 | 10.20 | 12.24 | 10.25 | 0.19 |
| 4/5 | 9 | 401 | 99 | 10.20 | 11.73 | 10.95 | 0.62 |
|  | 10 | 406 | 26 | 11.22 | 12.24 | 11.75 | 0.51 |
|  | 11 | 410 | 46 | 10.20 | 12.76 | 11.46 | 0.58 |
|  | 12 | 414 | 73 | 11.22 | 12.24 | 11.32 | 0.21 |
| 5/6 | 9 | 418 | 2 | 11.22 | 11.73 | 11.48 | 0.36 |
|  | 10 | 423 | 4 | 11.73 | 13.27 | 12.12 | 0.77 |
|  | 11 | 427 | 81 | 11.22 | 12.76 | 11.79 | 0.22 |
|  | 12 | 432 | N/A | N/A | N/A | N/A | N/A |

Table 3.3.5.2 Li
Accuracy and Consistency of Classification Indices: Read (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.706 | 0.614 |  | 0.498 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.858 |  | 0.798 |  |
|  | 2 | 0.687 |  | 0.585 |  |
|  | 3 | 0.578 |  | 0.450 |  |
|  | 4 | 0.218 |  | 0.157 |  |
|  | 5 | 0.602 |  | 0.472 |  |
|  | 6 | 0.831 |  | 0.730 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.905 | 0.045 | 0.050 | 0.866 |
|  | 2/3 | 0.920 | 0.045 | 0.035 | 0.888 |
|  | 3/4 | 0.943 | 0.031 | 0.026 | 0.917 |
|  | 4/5 | 0.947 | 0.036 | 0.017 | 0.926 |
|  | 5/6 | 0.965 | 0.019 | 0.015 | 0.951 |

Table 3.3.5.2Lii
Accuracy and Consistency of Classification Indices: Read (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.706 | 0.613 |  | 0.496 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.853 |  | 0.786 |  |
|  | 2 | 0.718 |  | 0.624 |  |
|  | 3 | 0.589 |  | 0.463 |  |
|  | 4 | 0.239 |  | 0.173 |  |
|  | 5 | 0.608 |  | 0.477 |  |
|  | 6 | 0.812 |  | 0.696 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.909 | 0.042 | 0.048 | 0.873 |
|  | 2/3 | 0.917 | 0.046 | 0.036 | 0.884 |
|  | 3/4 | 0.941 | 0.032 | 0.027 | 0.915 |
|  | 4/5 | 0.947 | 0.036 | 0.018 | 0.926 |
|  | 5/6 | 0.968 | 0.018 | 0.014 | 0.954 |

Table 3.3.5.2Liii
Accuracy and Consistency of Classification Indices: Read (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.705 | 0.611 |  | 0.491 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.824 |  | 0.750 |  |
|  | 2 | 0.759 |  | 0.669 |  |
|  | 3 | 0.509 |  | 0.389 |  |
|  | 4 | 0.273 |  | 0.200 |  |
|  | 5 | 0.592 |  | 0.468 |  |
|  | 6 | 0.816 |  | 0.702 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.905 | 0.049 | 0.046 | 0.866 |
|  | 2/3 | 0.922 | 0.039 | 0.039 | 0.888 |
|  | 3/4 | 0.940 | 0.038 | 0.022 | 0.915 |
|  | 4/5 | 0.944 | 0.036 | 0.020 | 0.924 |
|  | 5/6 | 0.967 | 0.018 | 0.015 | 0.953 |

Table 3.3.5.2Liv
Accuracy and Consistency of Classification Indices: Read (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.713 | 0.618 |  | 0.492 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.797 |  | 0.716 |  |
|  | 2 | 0.792 |  | 0.711 |  |
|  | 3 | 0.523 |  | 0.403 |  |
|  | 4 | 0.285 |  | 0.209 |  |
|  | 5 | 0.614 |  | 0.489 |  |
|  | 6 | 0.811 |  | 0.688 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.906 | 0.052 | 0.041 | 0.868 |
|  | 2/3 | 0.921 | 0.038 | 0.041 | 0.886 |
|  | 3/4 | 0.942 | 0.038 | 0.020 | 0.919 |
|  | 4/5 | 0.947 | 0.033 | 0.020 | 0.928 |
|  | 5/6 | 0.973 | 0.015 | 0.012 | 0.960 |

### 3.3.5.3 Writing 9-12

### 3.3.5.3i <br> Writing 9-12 A

Table 3.3.5.3Ai
Complete Task Analysis and Summary: Writ 9-12 A S401 Online
$\begin{array}{|l|c|c|c|c|c|c|}\hline \text { Task Type } & \begin{array}{c}\text { Average } \\ \text { Task } \\ \text { Difficulty } \\ \text { (in logits) }\end{array} & \begin{array}{c}\text { Average } \\ \text { No. of } \\ \text { Tasks }\end{array} & \begin{array}{c}\text { Average } \\ \text { Infit } \\ \text { Mean } \\ \text { Square }\end{array} & \begin{array}{c}\text { Outfit } \\ \text { Mean } \\ \text { Square }\end{array} \\ \hline \text { Constructed Response } & 2.14 & 3 & 0.58 & 0.61 \\ \hline & & & & \\$\cline { 3 - 6 } \& \& Task <br> Difficulty <br> (in logits)\end{array}$)$

Table 3.3.5.3Bi
DIF Analysis and Summary: Writ 9-12 A S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.W91A_SL_SchoolDance_P 100_A301_HW_14265 | AA | F | AA | O |  |
| 2.W91A_LA_Toaster_P 100_A301_HW_14290 | AA | F | AA | O |  |
| 3.W91A_MS_BouncingBalls_P 100_A203_alt1_HW_14294 | AA | M | AA | H |  |

Figure 3.3.5.3Ci
Raw Scores: Writ 9-12 A S401 Online


Figure 3.3.5.3Di
Scale Scores: Writ 9-12 A S401 Online


Table 3.3.5.3Ci
Raw Score Descriptive Statistics: Writ 9-12 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 38,220 | 0 | 24 | 9.01 | 4.62 |
| $\mathbf{1 0}$ | 23,565 | 0 | 21 | 10.39 | 4.02 |
| $\mathbf{1 1}$ | 15,956 | 0 | 22 | 11.23 | 3.74 |
| $\mathbf{1 2}$ | 9,142 | 0 | 21 | 11.70 | 3.63 |
| Total | 86,883 | 0 | 24 | 10.08 | 4.33 |

Table 3.3.5.3Di
Scale Score Descriptive Statistics: Writ 9-12 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 38,220 | 199 | 458 | 314.20 | 39.38 |
| $\mathbf{1 0}$ | 23,565 | 210 | 430 | 325.30 | 33.54 |
| $\mathbf{1 1}$ | 15,956 | 221 | 439 | 332.05 | 32.00 |
| $\mathbf{1 2}$ | 9,142 | 232 | 430 | 336.06 | 31.63 |
| Total | 86,883 | 199 | 458 | 322.79 | 36.73 |



Table 3.3.5.3Ei
Proficiency Level Distribution: Writ 9-12 A S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 9,907 | $25.92 \%$ | 4,212 | $17.87 \%$ | 3,443 | $21.58 \%$ | 2,915 | $31.89 \%$ | 20,477 | $23.57 \%$ |
| $\mathbf{2}$ | 12,043 | $31.51 \%$ | 9,009 | $38.23 \%$ | 5,833 | $36.56 \%$ | 1,944 | $21.26 \%$ | 28,829 | $33.18 \%$ |
| $\mathbf{3}$ | 13,887 | $36.33 \%$ | 9,464 | $40.16 \%$ | 6,383 | $40.00 \%$ | 4,061 | $44.42 \%$ | 33,795 | $38.90 \%$ |
| $\mathbf{4}$ | 2,372 | $6.21 \%$ | 880 | $3.73 \%$ | 297 | $1.86 \%$ | 222 | $2.43 \%$ | 3,771 | $4.34 \%$ |
| $\mathbf{5}$ | 11 | $0.03 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 11 | $0.01 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 38,220 | $100.00 \%$ | 23,565 | $100.00 \%$ | 15,956 | $100.00 \%$ | 9,142 | $100.00 \%$ | 86,883 | $100.00 \%$ |

Table 3.3.5.3Fi
Raw Score to Scale Score Conversion: Writ 9-12 A S401 Online

| Raw Score | Scale <br> Score | CSEM | Low Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $232^{\wedge}$ | 39.47 | 186.53 | 265.47 |
| 1 | 252 | 23.44 | 228.56 | 275.44 |
| 2 | 266 | 16.67 | 249.33 | 282.67 |
| 3 | 275 | 14.18 | 260.82 | 289.18 |
| 4 | 282 | 12.94 | 269.06 | 294.94 |
| 5 | 288 | 12.30 | 275.70 | 300.30 |
| 6 | 293 | 12.03 | 280.97 | 305.03 |
| 7 | 299 | 12.06 | 286.94 | 311.06 |
| 8 | 304 | 12.38 | 291.62 | 316.38 |
| 9 | 310 | 13.02 | 296.98 | 323.02 |
| 10 | 317 | 13.94 | 303.06 | 330.94 |
| 11 | 325 | 15.06 | 309.94 | 340.06 |
| 12 | 334 | 16.16 | 317.84 | 350.16 |
| 13 | 344 | 17.02 | 326.98 | 361.02 |
| 14 | 355 | 17.59 | 337.41 | 372.59 |
| 15 | 367 | 17.83 | 349.17 | 384.83 |
| 16 | 379 | 17.78 | 361.22 | 396.78 |
| 17 | 390 | 17.45 | 372.55 | 407.45 |
| 18 | 402 | 16.94 | 385.06 | 418.94 |
| 19 | 412 | 16.33 | 395.67 | 428.33 |
| 20 | 421 | 15.79 | 405.21 | 436.79 |
| 21 | 430 | 15.47 | 414.53 | 445.47 |
| 22 | 439 | 15.44 | 423.56 | 454.44 |
| 23 | 448 | 15.84 | 432.16 | 463.84 |
| 24 | 458 | 16.97 | 441.03 | 474.97 |
| 25 | 471 | 19.55 | 451.45 | 490.55 |
| 26 | 489 | 26.56 | 462.44 | 515.56 |
| 27 | 521 | 48.39 | 472.61 | 569.39 |

$\wedge$ Truncated

Table 3.3.5.3Gi
Equating Summary: Writ 9-12 A S401 Online


[^34]

Table 3.3.5.3Ji
Reliability: Writ 9-12 A S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 86,883 | 3 | Hand-written (HW) | Keyboarded (KB) | . 848 | 1.688 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 802 | 98 | 2 | 0 |
|  |  | KB | 43,556 | 93 | 7 | 0 |
|  | 2 | HW | 630 | 100 | 0 | 0 |
|  |  | KB | 43,410 | 95 | 5 | 0 |
|  | 3 | HW | 826 | 99 | 1 | 0 |
|  |  | KB | 43,321 | 94 | 6 | 0 |

Table 3.3.5.3Ki
Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 A S401 Online n/a

Table 3.3.5.3Li
Accuracy and Consistency of Classification Indices: Writ 9-12 A S401 Online n/a

Table 3.3.5.3Aii
Complete Task Analysis and Summary: Writ 9-12 B/C S401 Online

| Task Type |  | Average Task Difficulty (in logits) | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 1.91 | 3 | 0.76 | 0.74 |
| Name |  |  | Anchored? | Fit Statistics |  |
|  |  | Infit <br> Mnsq |  | Outfit <br> Mnsq |
| 1.W91B_SI_BestTeacher_P100_A 301_HW_14653 |  |  | 1.84 | Yes | 1.15 | 1.14 |
| 2.W91B_MS_Viscosity_P100_A203_HW_14652 |  | 1.83 | Yes | 0.73 | 0.71 |
| 3.W91C_IT_BoiardiChild_401_HW_15770 |  | 2.07 |  | 0.40 | 0.38 |
| Raw Score Distribution by Task | Raw Score | Task 1 | Task 2 | Task 3 |  |
|  | 0 | 0.12\% | 0.64\% | 0.54\% |  |
|  | 1 | 0.17\% | 0.20\% | 0.23\% |  |
|  | 2 | 0.99\% | 0.85\% | 1.45\% |  |
|  | 3 | 4.96\% | 2.55\% | 4.79\% |  |
|  | 4 | 16.14\% | 12.16\% | 19.62\% |  |
|  | 5 | 31.84\% | 34.68\% | 36.86\% |  |
|  | 6 | 30.17\% | 35.34\% | 26.90\% |  |
|  | 7 | 12.06\% | 11.52\% | 7.58\% |  |
|  | 8 | 2.92\% | 1.82\% | 1.67\% |  |
|  | 9 | 0.65\% | 0.24\% | 0.37\% |  |

Table 3.3.5.3Bii
DIF Analys is and Summary: Writ 9-12 B/C S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 1 | 2 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.W91B_SI_BestTeacher_P 100_A301_HW_14653 | AA | F | AA | O |  |
| 2.W91B_MS_Viscosity_P 100_A203_HW_14652 | AA | M | AA | H |  |
| 3.W91C_IT_BoiardiChild_401_HW_15770 | AA | F | AA | O |  |



Figure 3.3.5.3Dii
Scale Scores: Writ 9-12 B/C S401 Online


Table 3.3.5.3Cii
Raw Score Descriptive Statistics: Writ 9-12 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 43,685 | 0 | 53 | 31.36 | 5.61 |
| $\mathbf{1 0}$ | 30,962 | 0 | 53 | 31.25 | 5.67 |
| $\mathbf{1 1}$ | 22,723 | 0 | 54 | 31.93 | 5.70 |
| $\mathbf{1 2}$ | 16,772 | 0 | 52 | 31.87 | 5.84 |
| Total | 114,142 | 0 | 54 | 31.52 | 5.69 |

Table 3.3.5.3Dii
Scale Score Descriptive Statistics: Writ 9-12 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 43,685 | 199 | 501 | 370.30 | 29.74 |
| $\mathbf{1 0}$ | 30,962 | 210 | 501 | 369.75 | 29.92 |
| $\mathbf{1 1}$ | 22,723 | 221 | 533 | 373.44 | 30.11 |
| $\mathbf{1 2}$ | 16,772 | 232 | 483 | 373.20 | 30.60 |
| Total | 114,142 | 199 | 533 | 371.20 | 30.03 |



Table 3.3.5.3Eii
Proficiency Level Distribution: Writ 9-12 B/C S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 239 | $0.55 \%$ | 378 | $1.22 \%$ | 454 | $2.00 \%$ | 741 | $4.42 \%$ | 1,812 | $1.59 \%$ |
| $\mathbf{2}$ | 1,709 | $3.91 \%$ | 2,023 | $6.53 \%$ | 2,087 | $9.18 \%$ | 1,740 | $10.37 \%$ | 7,559 | $6.62 \%$ |
| $\mathbf{3}$ | 22,749 | $52.08 \%$ | 17,999 | $58.13 \%$ | 12,797 | $56.32 \%$ | 11,062 | $65.96 \%$ | 64,607 | $56.60 \%$ |
| $\mathbf{4}$ | 18,257 | $41.79 \%$ | 10,349 | $33.42 \%$ | 7,239 | $31.86 \%$ | 3,153 | $18.80 \%$ | 38,998 | $34.17 \%$ |
| $\mathbf{5}$ | 725 | $1.66 \%$ | 208 | $0.67 \%$ | 144 | $0.63 \%$ | 76 | $0.45 \%$ | 1,153 | $1.01 \%$ |
| $\mathbf{6}$ | 6 | $0.01 \%$ | 5 | $0.02 \%$ | 2 | $0.01 \%$ | 0 | $0.00 \%$ | 13 | $0.01 \%$ |
| Total | 43,685 | $100.00 \%$ | 30,962 | $100.00 \%$ | 22,723 | $100.00 \%$ | 16,772 | $100.00 \%$ | 114,142 | $100.00 \%$ |

Table 3.3.5.3Fii
Raw Score to Scale Score Conversion: Writ 9-12 B/C S401 Online

| $\begin{gathered} \text { Raw } \\ \text { Score } \end{gathered}$ | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound | $\begin{gathered} \text { Raw } \\ \text { Score } \end{gathered}$ | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | $232^{\wedge}$ | 26.05 | 179.95 | 232.05 | 34 | 386 | 12.38 | 373.62 | 398.38 |
| 1 | 234 | 24.38 | 209.62 | 258.38 | 35 | 391 | 12.22 | 378.78 | 403.22 |
| 2 | 249 | 16.57 | 232.43 | 265.57 | 36 | 397 | 12.00 | 385.00 | 409.00 |
| 3 | 257 | 13.35 | 243.65 | 270.35 | 37 | 402 | 11.79 | 390.21 | 413.79 |
| 4 | 263 | 11.57 | 251.43 | 274.57 | 38 | 407 | 11.55 | 395.45 | 418.55 |
| 5 | 268 | 10.45 | 257.55 | 278.45 | 39 | 412 | 11.33 | 400.67 | 423.33 |
| 6 | 271 | 9.69 | 261.31 | 280.69 | 40 | 417 | 11.12 | 405.88 | 428.12 |
| 7 | 275 | 9.16 | 265.84 | 284.16 | 41 | 421 | 10.96 | 410.04 | 431.96 |
| 8 | 278 | 8.78 | 269.22 | 286.78 | 42 | 426 | 10.82 | 415.18 | 436.82 |
| 9 | 280 | 8.51 | 271.49 | 288.51 | 43 | 430 | 10.74 | 419.26 | 440.74 |
| 10 | 283 | 8.35 | 274.65 | 291.35 | 44 | 434 | 10.74 | 423.26 | 444.74 |
| 11 | 286 | 8.24 | 277.76 | 294.24 | 45 | 438 | 10.85 | 427.15 | 448.85 |
| 12 | 288 | 8.22 | 279.78 | 296.22 | 46 | 443 | 11.04 | 431.96 | 454.04 |
| 13 | 291 | 8.24 | 282.76 | 299.24 | 47 | 448 | 11.36 | 436.64 | 459.36 |
| 14 | 293 | 8.32 | 284.68 | 301.32 | 48 | 453 | 11.84 | 441.16 | 464.84 |
| 15 | 296 | 8.46 | 287.54 | 304.46 | 49 | 458 | 12.57 | 445.43 | 470.57 |
| 16 | 299 | 8.65 | 290.35 | 307.65 | 50 | 464 | 13.69 | 450.31 | 477.69 |
| 17 | 301 | 8.89 | 292.11 | 309.89 | 51 | 472 | 15.49 | 456.51 | 487.49 |
| 18 | 304 | 9.18 | 294.82 | 313.18 | 52 | 483 | 18.74 | 464.26 | 501.74 |
| 19 | 308 | 9.53 | 298.47 | 317.53 | 53 | 501 | 26.42 | 474.58 | 527.42 |
| 20 | 311 | 9.93 | 301.07 | 320.93 | 54 | 533 | 48.65 | 484.35 | 581.65 |
| 21 | 315 | 10.34 | 304.66 | 325.34 |  |  |  |  |  |
| 22 | 319 | 10.77 | 308.23 | 329.77 |  |  |  |  |  |
| 23 | 324 | 11.17 | 312.83 | 335.17 |  |  |  |  |  |
| 24 | 328 | 11.52 | 316.48 | 339.52 |  |  |  |  |  |
| 25 | 334 | 11.84 | 322.16 | 345.84 |  |  |  |  |  |
| 26 | 339 | 12.11 | 326.89 | 351.11 |  |  |  |  |  |
| 27 | 344 | 12.32 | 331.68 | 356.32 |  |  |  |  |  |
| 28 | 350 | 12.49 | 337.51 | 362.49 |  |  |  |  |  |
| 29 | 356 | 12.59 | 343.41 | 368.59 |  |  |  |  |  |
| 30 | 362 | 12.65 | 349.35 | 374.65 |  |  |  |  |  |
| 31 | 368 | 12.65 | 355.35 | 380.65 |  |  |  |  |  |
| 32 | 374 | 12.62 | 361.38 | 386.62 |  |  |  |  |  |
| 33 | 380 | 12.51 | 367.49 | 392.51 |  |  |  |  |  |

[^35]Table 3.3.5.3Gii
Equating Summary: Writ 9-12 B/C S401 Online

| Comparison of Forms* | Form 401 |  |  | Form 400 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. of Tasks |  | Average Difficulty (Std. Dev.) | No. of Tasks |  | Average Difficulty (Std. Dev.) |
|  | 3 |  | 1.91 (0.13) | 3 |  | 3.16 (0.16) |
|  | Easiest |  | Hardest | Easiest |  | Hardest |
|  | 1.83 |  | 2.07 | 3.00 |  | 3.32 |
| Anchoring Tasks | No. of Possible Anchors |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | 1.84 (0.01) |  |  |  |
|  | No. of Anchors Used |  | Average Difficulty (Std. Dev.) |  |  |  |
|  | 2 |  | (Std. Dev.) |  |  |  |
|  | Percentage Anchors |  | Average Displacement |  |  |  |
|  | 67\% |  | 0.00 |  |  |  |
| Common <br> Rating Scale <br> Step <br> Measures | Anchored Scale Steps |  |  |  |  |  |
|  | Step |  | Measure |  |  |  |
|  | 1 |  | -2.47 |  |  |  |
|  | 2 |  | -2.78 |  |  |  |
|  | 3 |  | -2.61 |  |  |  |
|  | 4 |  | -1.68 |  |  |  |
|  | 5 |  | -0.48 |  |  |  |
|  | 6 |  | 0.97 |  |  |  |
|  | 7 |  | 2.25 |  |  |  |
|  | 8 |  | 3.21 |  |  |  |
|  | 9 |  | 3.59 |  |  |  |
| Displacement of Anchor Tasks | Anchor Tasks by Displacement |  |  | Anchor Tasks by Task Difficulty |  |  |
|  | Task ID ${ }^{\text {D }}$ | Task Difficulty | y y Displacement | Task ID | Task Difficulty | Displacement |
|  | 14653 | 1.84 | -0.01 | 14652 | 1.83 | 0.01 |
|  | 14652 | 1.83 | 0.01 | 14653 | 1.84 | -0.01 |

[^36]

Table 3.3.5.3Jii
Reliability: Writ 9-12 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Response Modes |  | Cronbach's Alpha | SEM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 114,142 | 3 | Hand-written (HW) | Keyboarded (KB) | . 881 | 1.959 |
| Interrater Reliability | Task | Mode of Response | No. in Sample | \% AG | \% AD | \% NA |
|  | 1 | HW | 178 | 99 | 1 | 0 |
|  |  | KB | 59,618 | 95 | 5 | 0 |
|  | 2 | HW | 198 | 98 | 2 | 0 |
|  |  | KB | 60,552 | 95 | 5 | 0 |
|  | 3 | HW | 208 | 99 | 1 | 0 |
|  |  | KB | 60,474 | 95 | 5 | 0 |

Table 3.3.5.3Kii
Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 B/C S401 Online n/a

Table 3.3.5.3Lii
Accuracy and Consistency of Classification Indices: Writ 9-12 B/C S401 Online n/a

### 3.3.5.3iii Writing 9-12 Across Tiers

Table 3.3.5.3Aiii
Complete Task Analysis and Summary: Writ 9-12 S401 Online n/a

Table 3.3.5.3Biii
DIF Analysis and Summary: Writ 9-12 S401 Online
n/a



Table 3.3.5.3Ciii
Raw Score Descriptive Statistics: Writ 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 81,905 | 0 | 53 | 20.93 | 12.29 |
| $\mathbf{1 0}$ | 54,527 | 0 | 53 | 22.23 | 11.49 |
| $\mathbf{1 1}$ | 38,679 | 0 | 54 | 23.39 | 11.34 |
| $\mathbf{1 2}$ | 25,914 | 0 | 52 | 24.75 | 10.94 |
| Total | 201,025 | 0 | 54 | 22.25 | 11.80 |

Table 3.3.5.3Diii
Scale Score Descriptive Statistics: Writ 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 81,905 | 199 | 501 | 344.12 | 44.48 |
| $\mathbf{1 0}$ | 54,527 | 210 | 501 | 350.54 | 38.46 |
| $\mathbf{1 1}$ | 38,679 | 221 | 533 | 356.37 | 37.02 |
| $\mathbf{1 2}$ | 25,914 | 232 | 483 | 360.10 | 35.70 |
| Total | 201,025 | 199 | 533 | 350.28 | 40.87 |



Table 3.3.5.3Eiii
Proficiency Level Distribution: Writ 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
|  | 10,146 | $12.39 \%$ | 4,590 | $8.42 \%$ | 3,897 | $10.08 \%$ | 3,656 | $14.11 \%$ | 22,289 | $11.09 \%$ |
| $\mathbf{2}$ | 13,752 | $16.79 \%$ | 11,032 | $20.23 \%$ | 7,920 | $20.48 \%$ | 3,684 | $14.22 \%$ | 36,388 | $18.10 \%$ |
| $\mathbf{3}$ | 36,636 | $44.73 \%$ | 27,463 | $50.37 \%$ | 19,180 | $49.59 \%$ | 15,123 | $58.36 \%$ | 98,402 | $48.95 \%$ |
| $\mathbf{4}$ | 20,629 | $25.19 \%$ | 11,229 | $20.59 \%$ | 7,536 | $19.48 \%$ | 3,375 | $13.02 \%$ | 42,769 | $21.28 \%$ |
| $\mathbf{5}$ | 736 | $0.90 \%$ | 208 | $0.38 \%$ | 144 | $0.37 \%$ | 76 | $0.29 \%$ | 1,164 | $0.58 \%$ |
| $\mathbf{6}$ | 6 | $0.01 \%$ | 5 | $0.01 \%$ | 2 | $0.01 \%$ | 0 | $0.00 \%$ | 13 | $0.01 \%$ |
| Total | 81,905 | $100.00 \%$ | 54,527 | $100.00 \%$ | 38,679 | $100.00 \%$ | 25,914 | $100.00 \%$ | 201,025 | $100.00 \%$ |

Table 3.3.5.3Fiii
Raw Score to Scale Score Conversion: Writ 9-12 S401 Online
n/a

Table 3.3.5.3Giii
Equating Summary: Writ 9-12 S401 Online n/a



Table 3.3.5.3Jiii
Reliability: Writ 9-12 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| A | 86,883 | 0.848 | 0.867 |
| B/C | 114,142 | 0.881 |  |

Table 3.3.5.3Kiii
Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 S401 Online

| Proficiency <br> Level |  | SEM |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Cut Score | Tier A | Tier B/C |
| $1 / 2$ |  | 289 | 12.35 | 8.32 |
|  | 10 | 298 | 12.08 | 8.59 |
|  | 11 | 308 | 12.89 | 9.53 |
|  | 12 | 318 | 14.23 | 10.74 |
|  | 9 | 319 | 14.23 | 10.77 |
|  | 10 | 326 | 15.31 | 11.28 |
|  | 11 | 335 | 16.38 | 11.81 |
|  | 12 | 344 | 17.02 | 12.32 |
| $3 / 4$ | 9 | 378 | 17.72 | 12.62 |
|  | 10 | 385 | 17.72 | 12.35 |
|  | 11 | 391 | 17.45 | 12.22 |
|  | 12 | 398 | 17.18 | 12.08 |
|  | 9 | 430 | 15.47 | 10.74 |
|  | 10 | 436 | 15.31 | 10.74 |
|  | 11 | 441 | 15.57 | 11.01 |
|  | 12 | 447 | 15.84 | 11.28 |
| $5 / 6$ | 9 | 469 | 19.33 | 14.77 |
|  | 10 | 479 | 22.29 | 17.45 |
|  | 11 | 490 | 27.12 | 21.48 |
|  | 12 | 501 | 33.03 | 26.42 |

Table 3.3.5.3Li
Accuracy and Consistency of Classification Indices: Writ (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.739 | 0.644 |  | 0.483 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.825 |  | 0.711 |  |
|  | 2 | 0.627 |  | 0.501 |  |
|  | 3 | 0.754 |  | 0.676 |  |
|  | 4 | 0.740 |  | 0.646 |  |
|  | 5 | - |  | 0.073 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.952 | 0.020 | 0.028 | 0.931 |
|  | 2/3 | 0.915 | 0.032 | 0.053 | 0.883 |
|  | 3/4 | 0.878 | 0.064 | 0.057 | 0.831 |
|  | 4/5 | 0.991 | 0.009 | 0.000 | 0.988 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.5.3Lii
Accuracy and Consistency of Classification Indices: Writ (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.764 | 0.675 |  | 0.501 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.759 |  | 0.608 |  |
|  | 2 | 0.688 |  | 0.566 |  |
|  | 3 | 0.783 |  | 0.729 |  |
|  | 4 | 0.781 |  | 0.668 |  |
|  | 5 | - |  | 0.102 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.957 | 0.019 | 0.024 | 0.936 |
|  | 2/3 | 0.905 | 0.033 | 0.061 | 0.870 |
|  | 3/4 | 0.904 | 0.058 | 0.038 | 0.867 |
|  | 4/5 | 0.996 | 0.004 | 0.000 | 0.995 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.5.3Liii
Accuracy and Consistency of Classification Indices: Writ (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.761 | 0.673 |  | 0.503 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.789 |  | 0.652 |  |
|  | 2 | 0.679 |  | 0.559 |  |
|  | 3 | 0.781 |  | 0.726 |  |
|  | 4 | 0.771 |  | 0.654 |  |
|  | 5 | - |  | 0.097 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.954 | 0.020 | 0.027 | 0.932 |
|  | 2/3 | 0.906 | 0.034 | 0.060 | 0.870 |
|  | 3/4 | 0.904 | 0.059 | 0.037 | 0.868 |
|  | 4/5 | 0.996 | 0.004 | 0.000 | 0.995 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.5.3Liv
Accuracy and Consistency of Classification Indices: Writ (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.779 | 0.692 |  | 0.502 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.868 |  | 0.760 |  |
|  | 2 | 0.519 |  | 0.396 |  |
|  | 3 | 0.852 |  | 0.811 |  |
|  | 4 | 0.691 |  | 0.540 |  |
|  | 5 | - |  | 0.371 |  |
|  | 6 | N/A |  | N/A |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
|  | 1/2 | 0.948 | 0.016 | 0.036 | 0.927 |
|  | 2/3 | 0.912 | 0.046 | 0.042 | 0.873 |
|  | 3/4 | 0.920 | 0.045 | 0.035 | 0.884 |
|  | 4/5 | 0.997 | 0.003 | 0.000 | 0.997 |
|  | 5/6 | N/A | N/A | N/A | N/A |

### 3.3.5.4 Speaking 9-12

3.3.5.4i Speaking 9-12 Pre-A

Table 3.3.5.4Ai
Complete Task Analysis and Summary: Spek 9-12 Pre-A S401 Online n/a

Table 3.3.5.4Bi
DIF Analys is and Summary: Spek 9-12 Pre-A S401 Online

| DIF Summary |  | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DIF <br> Level | Favoring <br> Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |  |
| AA | 2 | 1 | 1 | 2 |  |
| BB | 0 | 0 | 0 | 0 |  |
| CC | 0 | 0 | 0 | 0 |  |
|  | Male/Female |  | Hispanic/Other |  |  |
| Name | DIF <br> Level | Favored <br> Group | DIF <br> Level | Favored <br> Group |  |
| 1.S91P_SI_StudyPreferences_401_14543 | AA | M | AA | O |  |
| 2.S91P_SS_PahonaCity_P 100_A201_14692 | AA | F | AA | H |  |
| 3.S91P_SC_MonarchButterfly_P 100_A203_14693 | AA | M | AA | O |  |

Figure 3.3.5.4Ci
Raw Scores: Spek 9-12 Pre-A S401 Online


Figure 3.3.5.4Di
Scale Scores: Spek 9-12 Pre-A S401 Online


Table 3.3.5.4Ci
Raw Score Descriptive Statistics: Spek 9-12 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 6,708 | 0 | 6 | 5.30 | 1.23 |
| $\mathbf{1 0}$ | 5,520 | 0 | 6 | 5.66 | 0.87 |
| $\mathbf{1 1}$ | 3,987 | 0 | 6 | 5.75 | 0.75 |
| $\mathbf{1 2}$ | 2,548 | 0 | 6 | 5.82 | 0.65 |
| Total | 18,763 | 0 | 6 | 5.57 | 0.99 |

Table 3.3.5.4Di
Scale Score Descriptive Statistics: Spek 9-12 Pre-A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 6,708 | 154 | 225 | 215.86 | 15.91 |
| $\mathbf{1 0}$ | 5,520 | 160 | 225 | 220.55 | 11.15 |
| $\mathbf{1 1}$ | 3,987 | 166 | 225 | 221.85 | 9.44 |
| $\mathbf{1 2}$ | 2,548 | 172 | 225 | 222.69 | 7.90 |
| Total | 18,763 | 154 | 225 | 219.44 | 12.73 |



Table 3.3.5.4Ei
Proficiency Level Distribution: Spek 9-12 Pre-A S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 6,708 | $100.00 \%$ | 5,520 | $100.00 \%$ | 3,987 | $100.00 \%$ | 2,548 | $100.00 \%$ | 18,763 | $100.00 \%$ |
| Total | 6,708 | $100.00 \%$ | 5,520 | $100.00 \%$ | 3,987 | $100.00 \%$ | 2,548 | $100.00 \%$ | 18,763 | $100.00 \%$ |

Table 3.3.5.4Fi
Raw Score to Scale Score Conversion: Spek 9-12 Pre-A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $172^{\wedge}$ | 20.77 | 105.23 | 146.77 |
| 1 | $172^{\wedge}$ | 20.77 | 134.23 | 175.77 |
| 2 | 172 | 20.47 | 151.53 | 192.47 |
| 3 | 186 | 19.30 | 166.70 | 205.30 |
| 4 | 199 | 20.47 | 178.53 | 219.47 |
| 5 | $212^{*}$ | 24.57 | 192.43 | 241.57 |
| 6 | $225^{*}$ | 31.59 | 213.41 | 276.59 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.5.4Gi
Equating Summary: Spek 9-12 Pre-A S401 Online n/a


Table 3.3.5.4Ji
Reliability: Spek 9-12 Pre-A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 18,763 | 3 | .652 |  | 0.584 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 11,806 | 97 | 3 | 0 |
|  | 2 | 11,474 | 98 | 2 | 0 |
|  | 3 | 11,402 | 98 | 2 | 0 |

Table 3.3.5.4Ki
Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 Pre-A S401 Online n/a

Table 3.3.5.4Li
Accuracy and Consistency of Classification Indices: Spek 9-12 Pre-A S401 Online n/a

### 3.3.5.4ii $\quad$ Speaking 9-12 A

Table 3.3.5.4Aii
Complete Task Analysis and Summary: Spek 9-12 A S401 Online n/a

Table 3.3.4.4Bii
DIF Analys is and Summary: Spek 9-12 A S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\underset{\substack{\text { DIF } \\ \text { Level }}}{ }$ | Favoring Male (M) | Favoring <br> Female (F) | Favoring <br> Hispanic (H) | Favoring <br> Other (O) |
| AA | 3 | 3 | 1 | 5 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S91A_SI_StudyP references_40 L_V2_4543 | AA | M | AA | O |
| 2.S91A_SI_StudyP references_401_V2_14544 | AA | F | AA | O |
| 3.591A_SS_P aho naCity_P 100_A201_4163 | AA | F | AA | O |
| 4.S91A_SS_P aho naCity_P 100_A201_4164 | AA | M | AA | O |
| 5.S91A_SC_MonarchButterfly_P 100_A 203 _14159 | AA | F | AA | O |
| 6.S91A_SC_Monarch Butterfly_P 100_A 203 _14160 | AA | M | AA | H |

Figure 3.3.5.4Cii


Figure 3.3.5.4Dii
Scale Scores: Spek 9-12 A S401 Online


Table 3.3.5.4Cii
Raw Score Descriptive Statistics: Spek 9-12 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 38,370 | 0 | 18 | 11.55 | 1.98 |
| $\mathbf{1 0}$ | 21,147 | 0 | 18 | 11.89 | 1.66 |
| $\mathbf{1 1}$ | 8,122 | 0 | 17 | 11.80 | 1.55 |
| $\mathbf{1 2}$ | 12,300 | 0 | 18 | 12.49 | 1.50 |
| Total | 79,939 | 0 | 18 | 11.81 | 1.82 |

Table 3.3.5.4Dii
Scale Score Descriptive Statistics: Spek 9-12 A S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 38,370 | 154 | 445 | 306.16 | 39.67 |
| $\mathbf{1 0}$ | 21,147 | 160 | 445 | 312.46 | 35.28 |
| $\mathbf{1 1}$ | 8,122 | 166 | 424 | 310.26 | 33.16 |
| $\mathbf{1 2}$ | 12,300 | 172 | 445 | 325.62 | 33.43 |
| Total | 79,939 | 154 | 445 | 311.24 | 37.59 |

Figure 3.3.5.4Eii
Proficiency Level: Spek 9-12 A S401 Online


Table 3.3.5.4Eii
Proficiency Level Distribution: Spek 9-12 A S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 16,247 | $42.34 \%$ | 7,556 | $35.73 \%$ | 3,070 | $37.80 \%$ | 2,623 | $21.33 \%$ | 29,496 | $36.90 \%$ |
| $\mathbf{2}$ | 10,563 | $27.53 \%$ | 6,592 | $31.17 \%$ | 2,677 | $32.96 \%$ | 6,919 | $56.25 \%$ | 26,751 | $33.46 \%$ |
| $\mathbf{3}$ | 11,151 | $29.06 \%$ | 6,740 | $31.87 \%$ | 2,302 | $28.34 \%$ | 2,708 | $22.02 \%$ | 22,901 | $28.65 \%$ |
| $\mathbf{4}$ | 397 | $1.03 \%$ | 259 | $1.22 \%$ | 73 | $0.90 \%$ | 50 | $0.41 \%$ | 779 | $0.97 \%$ |
| $\mathbf{5}$ | 12 | $0.03 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 12 | $0.02 \%$ |
| $\mathbf{6}$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ | 0 | $0.00 \%$ |
| Total | 38,370 | $100.00 \%$ | 21,147 | $100.00 \%$ | 8,122 | $100.00 \%$ | 12,300 | $100.00 \%$ | 79,939 | $100.00 \%$ |

Table 3.3.5.4Fii
Raw Score to Scale Score Conversion: Spek 9-12 A S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 0 | $172^{\wedge}$ | 19.01 | 101.99 | 140.01 |
| 1 | $172^{\wedge}$ | 19.01 | 131.99 | 170.01 |
| 2 | $172^{\wedge}$ | 19.01 | 148.99 | 187.01 |
| 3 | 180 | 17.84 | 162.16 | 197.84 |
| 4 | 191 | 17.26 | 173.74 | 208.26 |
| 5 | 201 | 17.84 | 183.16 | 218.84 |
| 6 | 213 | 18.72 | 194.28 | 231.72 |
| 7 | 225 | 19.89 | 205.11 | 244.89 |
| 8 | 239 | 20.47 | 218.53 | 259.47 |
| 9 | 254 | 21.06 | 232.94 | 275.06 |
| 10 | 270 | 22.23 | 247.77 | 292.23 |
| 11 | 289 | 24.86 | 264.14 | 313.86 |
| 12 | 313 | 27.79 | 285.21 | 340.79 |
| 13 | 339 | 26.91 | 312.09 | 365.91 |
| 14 | 361 | 24.86 | 336.14 | 385.86 |
| 15 | 382 | 24.28 | 357.72 | 406.28 |
| 16 | 403 | 26.03 | 376.97 | 429.03 |
| 17 | $424^{*}$ | 30.42 | 400.58 | 461.42 |
| 18 | $445^{*}$ | 38.61 | 432.39 | 509.61 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.5.4Gii
Equating Summary: Spek 9-12 A S401 Online n/a



Table 3.3.5.4Jii
Reliability: Spek 9-12 A S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 79,939 | 6 | .632 |  | 1.103 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 50,202 | 99 | 1 | 0 |
|  | 2 | 50,202 | 82 | 17 | 0 |
|  | 3 | 49,277 | 99 | 1 | 0 |
|  | 4 | 49,276 | 82 | 17 | 1 |
|  | 5 | 49,694 | 99 | 1 | 0 |

Table 3.3.5.4Kii
Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 A S401 Online n/a

Table 3.3.5.4Lii
Accuracy and Consistency of Classification Indices: Spek 9-12 A S401 Online n/a

### 3.3.5.4iii Speaking 9-12 B/C

Table 3.3.5.4Aii
Complete Task Analysis and Summary: Spek 9-12 B/C S401 Online n/a

Table 3.3.5.4Biii
DIF Analysis and Summary: Spek 9-12 B/C S401 Online

| DIF Summary | Male/Female |  | Hispanic/Other |  |
| :---: | :---: | :---: | :---: | :---: |
| $\underset{\text { DiF }}{\text { DIFel }}$ | Favoring Male (M) | Favoring <br> Female (F) | Favoring Hispanic (H) | Favoring <br> Other (O) |
| AA | 3 | 3 | 4 | 2 |
| BB | 0 | 0 | 0 | 0 |
| CC | 0 | 0 | 0 | 0 |
|  | Male/Female |  | Hispanic/Other |  |
| Name | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group | $\begin{gathered} \text { DIF } \\ \text { Level } \end{gathered}$ | Favored Group |
| 1.S91C_SL_StudyP references _ 401_V2_14545 | AA | F | AA | H |
| 2.S91C_SL_StudyP references _40 L_V2_14546 | AA | F | AA | O |
| 3.S91C_LS_P aho naCity_401_VI_14761 | AA | M | AA | H |
| 4.S91C_LS_P aho naC City_401_Vl_14762 | AA | F | AA | H |
| 5.S91c_MS_MonarchButterflies_401_Vl_14508 | AA | M | AA | H |
| 6.S91C_MS_MonarchButterfies_40 _Vl_14509 | AA | M | AA | O |

Figure 3.3.5.4Ciii
Raw Scores: Spek 9-12 B/C S401 Online


Figure 3.3.5.4Diii Scale Scores: Spek 9-12 B/C S401 Online


Table 3.3.5.4Ciii
Raw Score Descriptive Statistics: Spek 9-12 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 27,601 | 6 | 30 | 19.26 | 2.20 |
| $\mathbf{1 0}$ | 21,530 | 6 | 29 | 19.08 | 2.16 |
| $\mathbf{1 1}$ | 21,445 | 6 | 30 | 18.94 | 2.23 |
| $\mathbf{1 2}$ | 8,571 | 8 | 30 | 19.60 | 2.25 |
| Total | 79,147 | 6 | 30 | 19.16 | 2.21 |

Table 3.3.5.4Diii
Scale Score Descriptive Statistics: Spek 9-12 B/C S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 27,601 | 154 | 468 | 339.93 | 26.03 |
| $\mathbf{1 0}$ | 21,530 | 160 | 454 | 337.77 | 25.62 |
| $\mathbf{1 1}$ | 21,445 | 166 | 474 | 336.05 | 26.42 |
| $\mathbf{1 2}$ | 8,571 | 227 | 476 | 343.96 | 26.45 |
| Total | 79,147 | 154 | 476 | 338.73 | 26.18 |



Table 3.3.5.4Eiii
Proficiency Level Distribution: Spek 9-12 B/C S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 431 | $1.56 \%$ | 811 | $3.77 \%$ | 1,007 | $4.70 \%$ | 465 | $5.43 \%$ | 2,714 | $3.43 \%$ |
| $\mathbf{2}$ | 9,966 | $36.11 \%$ | 7,973 | $37.03 \%$ | 8,528 | $39.77 \%$ | 4,029 | $47.01 \%$ | 30,496 | $38.53 \%$ |
| $\mathbf{3}$ | 16,261 | $58.91 \%$ | 12,144 | $56.41 \%$ | 11,625 | $54.21 \%$ | 3,982 | $46.46 \%$ | 44,012 | $55.61 \%$ |
| $\mathbf{4}$ | 919 | $3.33 \%$ | 597 | $2.77 \%$ | 273 | $1.27 \%$ | 92 | $1.07 \%$ | 1,881 | $2.38 \%$ |
| $\mathbf{5}$ | 22 | $0.08 \%$ | 5 | $0.02 \%$ | 10 | $0.05 \%$ | 2 | $0.02 \%$ | 39 | $0.05 \%$ |
| $\mathbf{6}$ | 2 | $0.01 \%$ | 0 | $0.00 \%$ | 2 | $0.01 \%$ | 1 | $0.01 \%$ | 5 | $0.01 \%$ |
| Total | 27,601 | $100.00 \%$ | 21,530 | $100.00 \%$ | 21,445 | $100.00 \%$ | 8,571 | $100.00 \%$ | 79,147 | $100.00 \%$ |

Table 3.3.5.4Fiii
Raw Score to Scale Score Conversion: Spek 9-12 B/C S401 Online

| Raw <br> Score | Scale <br> Score | CSEM | Low <br> Bound | High <br> Bound |
| :---: | :---: | :---: | :---: | :---: |
| 6 | $172^{\wedge}$ | 18.72 | 188.28 | 225.72 |
| 7 | 217 | 16.96 | 200.04 | 233.96 |
| 8 | 227 | 16.96 | 210.04 | 243.96 |
| 9 | 236 | 16.67 | 219.33 | 252.67 |
| 10 | 246 | 16.09 | 229.91 | 262.09 |
| 11 | 254 | 15.79 | 238.21 | 269.79 |
| 12 | 263 | 15.79 | 247.21 | 278.79 |
| 13 | 271 | 15.79 | 255.21 | 286.79 |
| 14 | 280 | 16.38 | 263.62 | 296.38 |
| 15 | 290 | 16.96 | 273.04 | 306.96 |
| 16 | 300 | 17.84 | 282.16 | 317.84 |
| 17 | 312 | 19.01 | 292.99 | 331.01 |
| 18 | 324 | 19.60 | 304.40 | 343.60 |
| 19 | 337 | 19.60 | 317.40 | 356.60 |
| 20 | 350 | 19.01 | 330.99 | 369.01 |
| 21 | 362 | 18.13 | 343.87 | 380.13 |
| 22 | 373 | 17.55 | 355.45 | 390.55 |
| 23 | 383 | 17.26 | 365.74 | 400.26 |
| 24 | 393 | 17.26 | 375.74 | 410.26 |
| 25 | 404 | 17.84 | 386.16 | 421.84 |
| 26 | 415 | 18.43 | 396.57 | 433.43 |
| 27 | 428 | 20.18 | 407.82 | 448.18 |
| 28 | $441^{*}$ | 22.81 | 421.19 | 466.81 |
| 29 | $454^{*}$ | 26.03 | 441.97 | 494.03 |
| 30 | $476^{*}$ | 35.10 | 469.90 | 540.10 |

$\wedge$ Truncated

* Adjusted for end of scale effect

Table 3.3.5.4Giii
Equating Summary: Spek 9-12 B/C S401 Online
n/a


Table 3.3.5.4Jiii
Reliability: Spek 9-12 B/C S401 Online

| Reliability | No. of Students | No. of Tasks | Cronbach's Alpha |  | SEM |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | 79,147 | 6 | .686 |  | 1.240 |
| Interrater <br> Reliability | Task | No. in Sample | \% EX | \% AD | \% NA |
|  | 1 | 49,278 | 82 | 17 | 0 |
|  | 2 | 49,280 | 82 | 18 | 0 |
|  | 3 | 49,070 | 77 | 22 | 1 |
|  | 4 | 49,070 | 78 | 22 | 0 |
|  | 5 | 49,012 | 75 | 25 | 0 |

Table 3.3.5.4Kiii
Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 B/C S401 Online n/a

Table 3.3.5.4Liii
Accuracy and Consistency of Classification Indices: Spek 9-12 B/C S401 Online n/a

### 3.3.5.4iv Speaking 9-12 Across Tiers

Table 3.3.5.4Aiv
Complete Task Analysis and Summary: Spek 9-12 S401 Online

| Task Type |  | Average Task Difficulty (in logits) |  | No. of Tasks | Average <br> Infit <br> Mean <br> Square | Average <br> Outfit <br> Mean <br> Square |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Constructed Response |  | 0.76 |  | 15 | 0.57 | 0.43 |
| Name |  | Tier | Task Difficulty (in logits) | Anchored? | Fit Statistics |  |
|  |  | $\begin{gathered} \text { Infit } \\ \text { Mnsq } \end{gathered}$ |  |  | Outfit <br> Mnsq |
| 1.S91A_SI_StudyPreferences_401_V2_14543 |  |  | A** | -2.72 |  | 0.72 | 0.20 |
| 2.S91A_SI_StudyPreferences_401_V2_14544 |  | A* | 1.83 | Yes | 0.49 | 0.49 |
| 3.S91A_SS_PahonaCity_P100_A201_14163 |  | A** | -2.77 |  | 0.67 | 0.22 |
| 4.S91A_SS_PahonaCity_P 100_A201_14164 |  | A | 1.24 | Yes | 0.44 | 0.43 |
| 5.S91A_SC_MonarchButterfly_P 100_A203_14159 |  | A** | -2.61 |  | 0.70 | 0.22 |
| 6.S91A_SC_MonarchButterfly_P 100_A203_14160 |  | A | 2.22 | Yes | 0.57 | 0.58 |
| 7.S91C_SI_StudyPreferences_401_V2_14545 |  | B/C* | 1.83 | Yes | 0.49 | 0.49 |
| 8.S91C_SI_StudyPreferences_401_V2_14546 |  | B/C | 2.31 | Yes | 0.48 | 0.47 |
| 9.S91C_LS_PahonaCity_401_V1_14761 |  | B/C | 2.01 | Yes | 0.47 | 0.46 |
| 10.S91C_LS_PahonaCity_401_V1_14762 |  | B/C | 2.39 | Yes | 0.44 | 0.43 |
| 11.S91C_MS_MonarchButterflies_401_V1_14508 |  | B/C | 1.37 | Yes | 0.55 | 0.54 |
| 12.S91C_MS_MonarchButterflies_401_V1_14509 |  | B/C | 3.04 | Yes | 0.74 | 0.74 |
| 13.S91P_SI_StudyPreferences_401_14543 |  | Pre-A** | -2.72 |  | 0.72 | 0.20 |
| 14.S91P_SS_PahonaCity_P 100_A201_14692 |  | Pre-A** | -2.77 |  | 0.67 | 0.22 |
| 15.S91P_SC_MonarchButterfly_P 100_A203_14693 |  | Pre-A** | -2.61 |  | 0.70 | 0.22 |
| Raw Score Distribution by Task | Task | Raw Score |  |  |  |  |
|  |  | 0 | 1 | 2 | 3 | 4 |
|  | Task 1 | 1.21\% | 3.08\% | 95.71\% | N/A | N/A |
|  | Task 2 | 0.53\% | 16.24\% | 65.33\% | 16.68\% | 1.21\% |
|  | Task 3 | 0.94\% | 3.08\% | 95.99\% | N/A | N/A |
|  | Task 4 | 0.86\% | 12.65\% | 68.04\% | 16.57\% | 1.88\% |
|  | Task 5 | 1.18\% | 3.52\% | 95.29\% | N/A | N/A |
|  | Task 6 | 1.03\% | 24.47\% | 50.81\% | 21.35\% | 2.33\% |
|  | Task 7 | 0.53\% | 16.24\% | 65.33\% | 16.68\% | 1.21\% |
|  | Task 8 | 0.29\% | 9.28\% | 73.31\% | 15.51\% | 1.62\% |
|  | Task 9 | 0.22\% | 4.60\% | 70.24\% | 23.74\% | 1.20\% |
|  | Task 10 | 0.29\% | 4.91\% | 70.84\% | 22.72\% | 1.24\% |
|  | Task 11 | 0.21\% | 2.35\% | 54.99\% | 39.58\% | 2.87\% |
|  | Task 12 | 0.43\% | 19.48\% | 56.10\% | 21.74\% | 2.26\% |
|  | Task 13 | 1.21\% | 3.08\% | 95.71\% | N/A | N/A |
|  | Task 14 | 0.94\% | 3.08\% | 95.99\% | N/A | N/A |
|  | Task 15 | 1.18\% | 3.52\% | 95.29\% | N/A | N/A |

[^37]* This task is shared between A and B/C.

Table 3.3.5.4Biv
DIF Analysis and Summary: Spek 9-12 S401 Online n/a



Table 3.3.5.4Civ
Raw Score Descriptive Statistics: Spek 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 72,679 | 0 | 30 | 13.90 | 4.97 |
| $\mathbf{1 0}$ | 48,197 | 0 | 29 | 14.39 | 4.97 |
| $\mathbf{1 1}$ | 33,554 | 0 | 30 | 15.64 | 5.10 |
| $\mathbf{1 2}$ | 23,419 | 0 | 30 | 14.36 | 4.79 |
| Total | 177,849 | 0 | 30 | 14.42 | 5.01 |

Table 3.3.5.4Div
Scale Score Descriptive Statistics: Spek 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 72,679 | 154 | 468 | 310.65 | 47.72 |
| $\mathbf{1 0}$ | 48,197 | 160 | 454 | 313.24 | 45.90 |
| $\mathbf{1 1}$ | 33,554 | 166 | 474 | 316.24 | 45.18 |
| $\mathbf{1 2}$ | 23,419 | 172 | 476 | 321.13 | 45.89 |
| Total | 177,849 | 154 | 476 | 313.79 | 46.65 |



Table 3.3.5.4Eiv
Proficiency Level Distribution: Spek 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 23,386 | $32.18 \%$ | 13,887 | $28.81 \%$ | 8,064 | $24.03 \%$ | 5,636 | $24.07 \%$ | 50,973 | $28.66 \%$ |
| $\mathbf{2}$ | 20,529 | $28.25 \%$ | 14,565 | $30.22 \%$ | 11,205 | $33.39 \%$ | 10,948 | $46.75 \%$ | 57,247 | $32.19 \%$ |
| $\mathbf{3}$ | 27,412 | $37.72 \%$ | 18,884 | $39.18 \%$ | 13,927 | $41.51 \%$ | 6,690 | $28.57 \%$ | 66,913 | $37.62 \%$ |
| $\mathbf{4}$ | 1,316 | $1.81 \%$ | 856 | $1.78 \%$ | 346 | $1.03 \%$ | 142 | $0.61 \%$ | 2,660 | $1.50 \%$ |
| $\mathbf{5}$ | 34 | $0.05 \%$ | 5 | $0.01 \%$ | 10 | $0.03 \%$ | 2 | $0.01 \%$ | 51 | $0.03 \%$ |
| $\mathbf{6}$ | 2 | $0.00 \%$ | 0 | $0.00 \%$ | 2 | $0.01 \%$ | 1 | $0.00 \%$ | 5 | $0.00 \%$ |
| Total | 72,679 | $100.00 \%$ | 48,197 | $100.00 \%$ | 33,554 | $100.00 \%$ | 23,419 | $100.00 \%$ | 177,849 | $100.00 \%$ |

Table 3.3.5.4Fiv
Raw Score to Scale Score Conversion: Spek 9-12 S401 Online n/a

Table 3.3.5.4Giv
Equating Summary: Spek 9-12 S401 Online


[^38]

Table 3.3.5.4Jiv
Reliability: Spek 9-12 Weighted Reliability S401 Online

| Tiers | No. of Students | Reliability | Weighted <br> Reliability |
| :---: | :---: | :---: | :---: |
| Pre-A | 18,763 | 0.652 | 0.658 |
| A | 79,939 | 0.632 |  |
| B/C | 79,147 | 0.686 |  |

Table 3.3.5.4Kiv
Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 S401 Online

| Proficiency Level | Grade | Cut Score | SEM |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Tier A | Tier B/C |
| 1/2 | 9 | 290 | 25.15 | 16.96 |
|  | 10 | 295 | 26.03 | 17.55 |
|  | 11 | 299 | 26.62 | 17.84 |
|  | 12 | 302 | 26.91 | 18.13 |
| 2/3 | 9 | 328 | 27.79 | 19.60 |
|  | 10 | 333 | 27.20 | 19.60 |
|  | 11 | 337 | 26.91 | 19.60 |
|  | 12 | 340 | 26.62 | 19.30 |
| 3/4 | 9 | 385 | 24.28 | 17.26 |
|  | 10 | 393 | 24.86 | 17.26 |
|  | 11 | 400 | 25.74 | 17.55 |
|  | 12 | 406 | 26.32 | 17.84 |
| 4/5 | 9 | 440 | 36.27 | 22.52 |
|  | 10 | 446 | 38.90 | 23.98 |
|  | 11 | 451 | 41.82 | 25.15 |
|  | 12 | 455 | 43.87 | 26.62 |
| 5/6 | 9 | 468 | 52.94 | 31.29 |
|  | 10 | 471 | 55.57 | 32.46 |
|  | 11 | 474 | 57.91 | 33.93 |
|  | 12 | 476 | 59.67 | 35.10 |

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficeny level higher than 2.

Table 3.3.5.4Li
Accuracy and Consistency of Classification Indices: Spek (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.632 | 0.526 |  | 0.301 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.866 |  | 0.716 |  |
|  | 2 | 0.432 |  | 0.342 |  |
|  | 3 | 0.610 |  | 0.554 |  |
|  | 4 | - |  | 0.034 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.869 | 0.035 | 0.096 | 0.810 |
|  | 2/3 | 0.752 | 0.078 | 0.171 | 0.674 |
|  | 3/4 | 0.981 | 0.019 | 0.000 | 0.956 |
|  | 4/5 | 1.000 | 0.001 | 0.000 | 0.999 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.5.4Lii
Accuracy and Consistency of Classification Indices: Spek (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.616 | 0.516 |  | 0.281 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.863 |  | 0.685 |  |
|  | 2 | 0.437 |  | 0.356 |  |
|  | 3 | 0.593 |  | 0.545 |  |
|  | 4 | - |  | 0.030 |  |
|  | 5 | - |  | - |  |
|  | 6 | N/A |  | N/A |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.881 | 0.032 | 0.087 | 0.818 |
|  | 2/3 | 0.719 | 0.091 | 0.190 | 0.646 |
|  | 3/4 | 0.982 | 0.018 | 0.000 | 0.971 |
|  | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |
|  | 5/6 | N/A | N/A | N/A | N/A |

Table 3.3.5.4Liii
Accuracy and Consistency of Classification Indices: Spek (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.581 | 0.496 |  | 0.240 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.917 |  | 0.628 |  |
|  | 2 | - |  | 0.382 |  |
|  | 3 | 0.036 |  | 0.528 |  |
|  | 4 | - |  | 0.015 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.906 | 0.015 | 0.079 | 0.825 |
|  | 2/3 | - | - | - | - |
|  | 3/4 | 0.989 | 0.011 | 0.000 | 0.983 |
|  | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

Table 3.3.5.4Liv
Accuracy and Consistency of Classification Indices: Spek (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.548 | 0.491 |  | 0.236 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.820 |  | 0.609 |  |
|  | 2 | 0.964 |  | 0.533 |  |
|  | 3 | 0.352 |  | 0.383 |  |
|  | 4 | - |  | 0.010 |  |
|  | 5 | - |  | - |  |
|  | 6 | - |  | - |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.887 | 0.036 | 0.077 | 0.817 |
|  | 2/3 | - | - | - | - |
|  | 3/4 | 0.994 | 0.006 | 0.000 | 0.987 |
|  | 4/5 | 1.000 | 0.000 | 0.000 | 1.000 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.4. Analyses of Composite Scores: Results

### 3.4.1 Grade: 1

### 3.4.1.1 Oral Language Composite 1




Table 3.4.1.1A
Scale Score Descriptive Statistics: Oral 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 118,149 | 118 | 414 | 300.49 | 44.56 |
| Total | 118,149 | 118 | 414 | 300.49 | 44.56 |

Table 3.4.1.1B
Proficiency Level Distribution: Oral 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 7,882 | $6.67 \%$ | 7,882 | $6.67 \%$ |
| $\mathbf{2}$ | 10,862 | $9.19 \%$ | 10,862 | $9.19 \%$ |
| $\mathbf{3}$ | 31,456 | $26.62 \%$ | 31,456 | $26.62 \%$ |
| $\mathbf{4}$ | 36,343 | $30.76 \%$ | 36,343 | $30.76 \%$ |
| $\mathbf{5}$ | 27,382 | $23.18 \%$ | 27,382 | $23.18 \%$ |
| $\mathbf{6}$ | 4,224 | $3.58 \%$ | 4,224 | $3.58 \%$ |
| Total | 118,149 | $100.00 \%$ | 118,149 | $100.00 \%$ |

Table 3.4.1.1C
Reliability: Oral 1 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.50 | 3369.627 | 0.860 |
| Speaking | 0.50 | 1837.482 | 0.668 |
| Oral | 1994.085 | 0.864 |  |

* Variances from students who had results in all four domains

Table 3.4.1.1D
Accuracy and Consistency of Classification Indices: Oral (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.629 | 0.515 |  | 0.366 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.867 |  | 0.756 |  |
|  | 2 | 0.578 |  | 0.441 |  |
|  | 3 | 0.704 |  | 0.580 |  |
|  | 4 | 0.579 |  | 0.469 |  |
|  | 5 | 0.589 |  | 0.505 |  |
|  | 6 | - |  | 0.140 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.976 | 0.008 | 0.017 | 0.966 |
|  | 2/3 | 0.947 | 0.026 | 0.027 | 0.922 |
|  | 3/4 | 0.886 | 0.046 | 0.069 | 0.840 |
|  | 4/5 | 0.849 | 0.070 | 0.081 | 0.790 |
|  | 5/6 | 0.964 | 0.036 | 0.000 | 0.952 |

### 3.4.1.2 Literacy Language Composite 1




Table 3.4.1.2A
Scale Score Descriptive Statistics: Litr 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 146,025 | 140 | 415 | 275.16 | 30.06 |
| Total | 146,025 | 140 | 415 | 275.16 | 30.06 |

Table 3.4.1.2B
Proficiency Level Distribution: Litr 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 26,314 | $18.02 \%$ | 26,314 | $18.02 \%$ |
| $\mathbf{2}$ | 66,589 | $45.60 \%$ | 66,589 | $45.60 \%$ |
| $\mathbf{3}$ | 40,608 | $27.81 \%$ | 40,608 | $27.81 \%$ |
| $\mathbf{4}$ | 10,268 | $7.03 \%$ | 10,268 | $7.03 \%$ |
| $\mathbf{5}$ | 1,884 | $1.29 \%$ | 1,884 | $1.29 \%$ |
| $\mathbf{6}$ | 362 | $0.25 \%$ | 362 | $0.25 \%$ |
| Total | 146,025 | $100.00 \%$ | 146,025 | $100.00 \%$ |

Table 3.4.1.2C
Reliability: Litr 1 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Reading | 0.50 | 1226.883 | 0.890 |
| Writing | 0.50 | 1122.100 | 0.865 |
| Literacy |  | 909.410 | 0.921 |

[^39]Table 3.4.1.2D
Accuracy and Consistency of Classification Indices: Litr (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.805 | 0.723 |  | 0.597 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.780 |  | 0.678 |  |
|  | 2 | 0.839 |  | 0.779 |  |
|  | 3 | 0.778 |  | 0.696 |  |
|  | 4 | 0.784 |  | 0.674 |  |
|  | 5 | 0.725 |  | 0.577 |  |
|  | 6 | 0.850 |  | 0.660 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.925 | 0.041 | 0.033 | 0.892 |
|  | 2/3 | 0.915 | 0.037 | 0.048 | 0.879 |
|  | 3/4 | 0.973 | 0.018 | 0.009 | 0.962 |
|  | 4/5 | 0.993 | 0.005 | 0.002 | 0.991 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 0.998 |

### 3.4.1.3 Comprehension Language Composite 1




Table 3.4.1.3A
Scale Score Descriptive Statistics: Cphn 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 135,505 | 130 | 417 | 301.27 | 35.20 |
| Total | 135,505 | 130 | 417 | 301.27 | 35.20 |

Table 3.4.1.3B
Proficiency Level Distribution: Cphn 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 13,313 | $9.82 \%$ | 13,313 | $9.82 \%$ |
| $\mathbf{2}$ | 21,693 | $16.01 \%$ | 21,693 | $16.01 \%$ |
| $\mathbf{3}$ | 29,379 | $21.68 \%$ | 29,379 | $21.68 \%$ |
| $\mathbf{4}$ | 18,213 | $13.44 \%$ | 18,213 | $13.44 \%$ |
| $\mathbf{5}$ | 26,346 | $19.44 \%$ | 26,346 | $19.44 \%$ |
| $\mathbf{6}$ | 26,561 | $19.60 \%$ | 26,561 | $19.60 \%$ |
| Total | 135,505 | $100.00 \%$ | 135,505 | $100.00 \%$ |

Table 3.4.1.3C
Reliability: Cphn 1 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.30 | 3369.627 | 0.860 |
| Reading | 0.70 | 1226.883 | 0.890 |
| Comprehension | 1253.142 | 0.913 |  |

*Variances from students who had results in all four domains

Table 3.4.1.3D
Accuracy and Consistency of Classification Indices: Cphn (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.646 | 0.543 |  | 0.445 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.790 |  | 0.654 |  |
|  | 2 | 0.606 |  | 0.489 |  |
|  | 3 | 0.596 |  | 0.486 |  |
|  | 4 | 0.420 |  | 0.321 |  |
|  | 5 | 0.614 |  | 0.499 |  |
|  | 6 | 0.861 |  | 0.777 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.952 | 0.018 | 0.030 | 0.932 |
|  | 2/3 | 0.921 | 0.037 | 0.042 | 0.888 |
|  | 3/4 | 0.903 | 0.054 | 0.043 | 0.866 |
|  | 4/5 | 0.908 | 0.047 | 0.045 | 0.872 |
|  | 5/6 | 0.940 | 0.034 | 0.026 | 0.915 |

### 3.4.1.4 Overall Language Composite 1




Table 3.4.1.4A
Scale Score Descriptive Statistics: Over 1 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{1}$ | 109,820 | 153 | 405 | 282.73 | 30.64 |
| Total | 109,820 | 153 | 405 | 282.73 | 30.64 |

Table 3.4.1.4B
Proficiency Level Distribution: Over 1 S401 Online

| Level | Grade 1 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 9,061 | $8.25 \%$ | 9,061 | $8.25 \%$ |
| $\mathbf{2}$ | 30,321 | $27.61 \%$ | 30,321 | $27.61 \%$ |
| $\mathbf{3}$ | 53,835 | $49.02 \%$ | 53,835 | $49.02 \%$ |
| $\mathbf{4}$ | 13,518 | $12.31 \%$ | 13,518 | $12.31 \%$ |
| $\mathbf{5}$ | 2,782 | $2.53 \%$ | 2,782 | $2.53 \%$ |
| $\mathbf{6}$ | 303 | $0.28 \%$ | 303 | $0.28 \%$ |
| Total | 109,820 | $100.00 \%$ | 109,820 | $100.00 \%$ |

Table 3.4.1.4C
Reliability: Over 1 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.15 | 3369.627 | 0.860 |
| Reading | 0.35 | 1226.883 | 0.890 |
| Speaking | 0.15 | 1837.482 | 0.668 |
| Writing | 0.35 | 1122.100 | 0.865 |
| Overall Composite |  | 939.072 | 0.937 |

*Variances from students who had results in all four domains

Table 3.4.1.4D
Accuracy and Consistency of Classification Indices: Over (Grade 1) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.831 | 0.761 |  | 0.642 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.836 |  | 0.730 |  |
|  | 2 | 0.801 |  | 0.723 |  |
|  | 3 | 0.871 |  | 0.828 |  |
|  | 4 | 0.752 |  | 0.646 |  |
|  | 5 | 0.775 |  | 0.649 |  |
|  | 6 | 0.892 |  | 0.685 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.970 | 0.013 | 0.017 | 0.957 |
|  | 2/3 | 0.923 | 0.039 | 0.038 | 0.892 |
|  | 3/4 | 0.952 | 0.026 | 0.022 | 0.932 |
|  | 4/5 | 0.988 | 0.010 | 0.003 | 0.984 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.998 |

### 3.4.2 Grades: 2-3

### 3.4.2.1 Oral Language Composite 2-3




Table 3.4.2.1A
Scale Score Descriptive Statistics: Oral 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 119,993 | 110 | 421 | 297.01 | 44.26 |
| $\mathbf{3}$ | 130,821 | 115 | 436 | 314.05 | 48.10 |
| Total | 250,814 | 110 | 436 | 305.90 | 47.08 |

Table 3.4.2.1B
Proficiency Level Distribution: Oral 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 8,713 | $7.26 \%$ | 10,339 | $7.90 \%$ | 19,052 | $7.60 \%$ |
| $\mathbf{2}$ | 23,839 | $19.87 \%$ | 20,719 | $15.84 \%$ | 44,558 | $17.77 \%$ |
| $\mathbf{3}$ | 45,482 | $37.90 \%$ | 43,627 | $33.35 \%$ | 89,109 | $35.53 \%$ |
| $\mathbf{4}$ | 33,340 | $27.78 \%$ | 46,689 | $35.69 \%$ | 80,029 | $31.91 \%$ |
| $\mathbf{5}$ | 8,217 | $6.85 \%$ | 9,008 | $6.89 \%$ | 17,225 | $6.87 \%$ |
| $\mathbf{6}$ | 402 | $0.34 \%$ | 439 | $0.34 \%$ | 841 | $0.34 \%$ |
| Total | 119,993 | $100.00 \%$ | 130,821 | $100.00 \%$ | 250,814 | $100.00 \%$ |

Table 3.4.2.1C
Reliability: Oral 2-3 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.50 | 3230.825 | 0.860 |
| Speaking | 0.50 | 2449.825 | 0.668 |
| Oral | 2279.670 | 0.861 |  |

*Variances from students who had results in all four domains

Table 3.4.2.1Di
Accuracy and Consistency of Classification Indices: Oral (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.652 | 0.543 |  | 0.372 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.791 |  | 0.665 |  |
|  | 2 | 0.726 |  | 0.597 |  |
|  | 3 | 0.694 |  | 0.567 |  |
|  | 4 | 0.558 |  | 0.500 |  |
|  | 5 | - |  | 0.205 |  |
|  | 6 | - |  | 0.667 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.970 | 0.015 | 0.015 | 0.955 |
|  | 2/3 | 0.913 | 0.035 | 0.052 | 0.878 |
|  | 3/4 | 0.837 | 0.054 | 0.109 | 0.774 |
|  | 4/5 | 0.928 | 0.072 | 0.000 | 0.905 |
|  | 5/6 | 0.997 | 0.003 | 0.000 | 0.997 |

Table 3.4.2.1Dii
Accuracy and Consistency of Classification Indices: Oral (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.636 | 0.522 |  | 0.333 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.814 |  | 0.706 |  |
|  | 2 | 0.712 |  | 0.571 |  |
|  | 3 | 0.634 |  | 0.473 |  |
|  | 4 | 0.590 |  | 0.548 |  |
|  | 5 | - |  | 0.142 |  |
|  | 6 | - |  | 0.053 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.972 | 0.015 | 0.013 | 0.958 |
|  | 2/3 | 0.927 | 0.026 | 0.047 | 0.897 |
|  | 3/4 | 0.805 | 0.052 | 0.143 | 0.729 |
|  | 4/5 | 0.928 | 0.072 | 0.000 | 0.895 |
|  | 5/6 | 0.997 | 0.003 | 0.000 | 0.996 |

### 3.4.2.2 Literacy Language Composite 2-3



Figure 3.4.2.2B
Proficiency Level: Litr 2-3 S401 Online


Table 3.4.2.2A
Scale Score Descriptive Statistics: Litr 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 142,299 | 158 | 428 | 303.70 | 31.07 |
| $\mathbf{3}$ | 149,242 | 180 | 437 | 322.48 | 34.09 |
| Total | 291,541 | 158 | 437 | 313.32 | 33.97 |

Table 3.4.2.2B
Proficiency Level Distribution: Litr 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 11,781 | $8.28 \%$ | 11,676 | $7.82 \%$ | 23,457 | $8.05 \%$ |
| $\mathbf{2}$ | 41,196 | $28.95 \%$ | 29,736 | $19.92 \%$ | 70,932 | $24.33 \%$ |
| $\mathbf{3}$ | 64,002 | $44.98 \%$ | 63,691 | $42.68 \%$ | 127,693 | $43.80 \%$ |
| $\mathbf{4}$ | 22,277 | $15.66 \%$ | 36,564 | $24.50 \%$ | 58,841 | $20.18 \%$ |
| $\mathbf{5}$ | 2,804 | $1.97 \%$ | 7,024 | $4.71 \%$ | 9,828 | $3.37 \%$ |
| $\mathbf{6}$ | 239 | $0.17 \%$ | 551 | $0.37 \%$ | 790 | $0.27 \%$ |
| Total | 142,299 | $100.00 \%$ | 149,242 | $100.00 \%$ | 291,541 | $100.00 \%$ |

Table 3.4.2.2C
Reliability: Litr 2-3 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Reading | 0.50 | 1359.713 | 0.900 |
| Writing | 0.50 | 1422.881 | 0.886 |
| Literacy |  | 1140.740 | 0.935 |

[^40]Table 3.4.2.2Di
Accuracy and Consistency of Classification Indices: Litr (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.821 | 0.747 |  | 0.629 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.808 |  | 0.695 |  |
|  | 2 | 0.824 |  | 0.748 |  |
|  | 3 | 0.836 |  | 0.785 |  |
|  | 4 | 0.788 |  | 0.695 |  |
|  | 5 | 0.694 |  | 0.520 |  |
|  | 6 | 0.867 |  | 0.457 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.968 | 0.016 | 0.016 | 0.953 |
|  | 2/3 | 0.923 | 0.033 | 0.045 | 0.892 |
|  | 3/4 | 0.944 | 0.033 | 0.023 | 0.922 |
|  | 4/5 | 0.987 | 0.010 | 0.003 | 0.983 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.998 |

Table 3.4.2.2Dii
Accuracy and Consistency of Classification Indices: Litr (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.799 | 0.718 |  | 0.603 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.847 |  | 0.749 |  |
|  | 2 | 0.783 |  | 0.690 |  |
|  | 3 | 0.829 |  | 0.771 |  |
|  | 4 | 0.767 |  | 0.687 |  |
|  | 5 | 0.650 |  | 0.485 |  |
|  | 6 | - |  | 0.233 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.974 | 0.011 | 0.015 | 0.962 |
|  | 2/3 | 0.937 | 0.027 | 0.035 | 0.912 |
|  | 3/4 | 0.924 | 0.040 | 0.036 | 0.893 |
|  | 4/5 | 0.967 | 0.023 | 0.010 | 0.954 |
|  | 5/6 | 0.996 | 0.004 | 0.000 | 0.996 |

### 3.4.2.3 Comprehension Language Composite 2-3




Table 3.4.2.3A
Scale Score Descriptive Statistics: Cphn 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 129,302 | 195 | 424 | 315.88 | 34.31 |
| $\mathbf{3}$ | 136,880 | 144 | 424 | 335.49 | 39.11 |
| Total | 266,182 | 144 | 424 | 325.97 | 38.14 |

Table 3.4.2.3B
Proficiency Level Distribution: Cphn 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 12,428 | $9.61 \%$ | 15,801 | $11.54 \%$ | 28,229 | $10.61 \%$ |
| $\mathbf{2}$ | 30,102 | $23.28 \%$ | 26,030 | $19.02 \%$ | 56,132 | $21.09 \%$ |
| $\mathbf{3}$ | 32,165 | $24.88 \%$ | 29,790 | $21.76 \%$ | 61,955 | $23.28 \%$ |
| $\mathbf{4}$ | 17,019 | $13.16 \%$ | 15,188 | $11.10 \%$ | 32,207 | $12.10 \%$ |
| $\mathbf{5}$ | 19,311 | $14.93 \%$ | 22,268 | $16.27 \%$ | 41,579 | $15.62 \%$ |
| $\mathbf{6}$ | 18,277 | $14.14 \%$ | 27,803 | $20.31 \%$ | 46,080 | $17.31 \%$ |
| Total | 129,302 | $100.00 \%$ | 136,880 | $100.00 \%$ | 266,182 | $100.00 \%$ |

Table 3.4.2.3C
Reliability: Cphn 2-3 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.30 | 3230.825 | 0.860 |
| Reading | 0.70 | 1359.713 | 0.900 |
| Comprehension | 1468.833 | 0.927 |  |

[^41]Table 3.4.2.3Di
Accuracy and Consistency of Classification Indices: Cphn (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.686 | 0.583 |  | 0.489 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.803 |  | 0.678 |  |
|  | 2 | 0.725 |  | 0.627 |  |
|  | 3 | $0.667$ |  | 0.560 |  |
|  | 4 | $0.474$ |  | 0.362 |  |
|  | 5 | $0.620$ |  | 0.498 |  |
|  | 6 | $0.868$ |  | 0.781 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.957 | 0.017 | 0.025 | 0.939 |
|  | 2/3 | 0.918 | 0.041 | 0.041 | 0.885 |
|  | 3/4 | 0.915 | 0.043 | 0.041 | 0.882 |
|  | 4/5 | 0.930 | 0.037 | 0.033 | 0.901 |
|  | 5/6 | 0.956 | 0.027 | 0.017 | 0.937 |

Table 3.4.2.3Dii
Accuracy and Consistency of Classification Indices: Cphn (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.677 | 0.577 |  | 0.487 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.819 |  | 0.712 |  |
|  | 2 | 0.670 |  | 0.559 |  |
|  | 3 | 0.635 |  | 0.524 |  |
|  | 4 | 0.403 |  | 0.303 |  |
|  | 5 | 0.601 |  | 0.480 |  |
|  | 6 | 0.881 |  | 0.806 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.954 | 0.020 | 0.027 | 0.934 |
|  | $2 / 3$ | 0.922 | 0.039 | 0.039 | 0.891 |
|  | 3/4 | 0.916 | 0.042 | 0.042 | 0.882 |
|  | 4/5 | 0.925 | 0.040 | 0.036 | 0.893 |
|  | 5/6 | 0.944 | 0.033 | 0.023 | 0.921 |

### 3.4.2.4 Overall Language Composite 2-3



Figure 3.4.2.4B
Proficiency Level: Over 2-3 S401 Online


Table 3.4.2.4A
Scale Score Descriptive Statistics: Over 2-3 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{2}$ | 106,741 | 157 | 398 | 301.91 | 31.90 |
| $\mathbf{3}$ | 116,535 | 171 | 422 | 320.25 | 35.44 |
| Total | 223,276 | 157 | 422 | 311.48 | 35.01 |

Table 3.4.2.4B
Proficiency Level Distribution: Over 2-3 S401 Online

| Level | Grade 2 |  | Grade 3 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 7,127 | $6.68 \%$ | 8,857 | $7.60 \%$ | 15,984 | $7.16 \%$ |
| $\mathbf{2}$ | 27,687 | $25.94 \%$ | 19,909 | $17.08 \%$ | 47,596 | $21.32 \%$ |
| $\mathbf{3}$ | 50,067 | $46.91 \%$ | 50,320 | $43.18 \%$ | 100,387 | $44.96 \%$ |
| $\mathbf{4}$ | 19,608 | $18.37 \%$ | 32,718 | $28.08 \%$ | 52,326 | $23.44 \%$ |
| $\mathbf{5}$ | 2,190 | $2.05 \%$ | 4,637 | $3.98 \%$ | 6,827 | $3.06 \%$ |
| $\mathbf{6}$ | 62 | $0.06 \%$ | 94 | $0.08 \%$ | 156 | $0.07 \%$ |
| Total | 106,741 | $100.00 \%$ | 116,535 | $100.00 \%$ | 223,276 | $100.00 \%$ |

Table 3.4.2.4C
Reliability: Over 2-3 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.15 | 3230.825 | 0.860 |
| Reading | 0.35 | 1359.713 | 0.900 |
| Speaking | 0.15 | 2449.825 | 0.668 |
| Writing | 0.35 | 1422.881 | 0.886 |
| Overall Composite |  |  |  | $1225.721 \quad 0.947 \quad$.

*Variances from students who had results in all four domains

Table 3.4.2.4Di
Accuracy and Consistency of Classification Indices: Over (Grade 2) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.833 | 0.772 |  | 0.662 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.828 |  | 0.738 |  |
|  | 2 | 0.845 |  | 0.776 |  |
|  | 3 | 0.862 |  | 0.816 |  |
|  | 4 | 0.751 |  | 0.689 |  |
|  | 5 | - |  | 0.499 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.977 | 0.012 | 0.011 | 0.968 |
|  | 2/3 | 0.936 | 0.028 | 0.035 | 0.911 |
|  | 3/4 | 0.939 | 0.031 | 0.030 | 0.914 |
|  | 4/5 | 0.979 | 0.021 | 0.000 | 0.979 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

Table 3.4.2.4Dii
Accuracy and Consistency of Classification Indices: Over (Grade 3) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.810 | 0.744 |  | 0.632 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.871 |  | 0.807 |  |
|  | 2 | 0.796 |  | 0.706 |  |
|  | 3 | 0.860 |  | 0.803 |  |
|  | 4 | 0.741 |  | 0.693 |  |
|  | 5 | - |  | 0.326 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.980 | 0.010 | 0.011 | 0.972 |
|  | 2/3 | 0.949 | 0.025 | 0.026 | 0.929 |
|  | 3/4 | 0.920 | 0.033 | 0.047 | 0.888 |
|  | 4/5 | 0.959 | 0.041 | 0.000 | 0.955 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

### 3.4.3 Grades: 4-5

### 3.4.3.1 Oral Language Composite 4-5



Figure 3.4.3.1B
Proficiency Level: Oral 4-5 S401 Online


Table 3.4.3.1 A
Scale Score Descriptive Statistics: Oral 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 85,468 | 152 | 467 | 356.74 | 43.00 |
| $\mathbf{5}$ | 62,948 | 134 | 475 | 359.43 | 49.58 |
| Total | 148,416 | 134 | 475 | 357.89 | 45.93 |

Table 3.4.3.1B
Proficiency Level Distribution: Oral 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 3,409 | $3.99 \%$ | 4,097 | $6.51 \%$ | 7,506 | $5.06 \%$ |
| $\mathbf{2}$ | 5,026 | $5.88 \%$ | 4,714 | $7.49 \%$ | 9,740 | $6.56 \%$ |
| $\mathbf{3}$ | 14,796 | $17.31 \%$ | 12,041 | $19.13 \%$ | 26,837 | $18.08 \%$ |
| $\mathbf{4}$ | 37,822 | $44.25 \%$ | 24,663 | $39.18 \%$ | 62,485 | $42.10 \%$ |
| $\mathbf{5}$ | 20,380 | $23.85 \%$ | 14,607 | $23.20 \%$ | 34,987 | $23.57 \%$ |
| $\mathbf{6}$ | 4,035 | $4.72 \%$ | 2,826 | $4.49 \%$ | 6,861 | $4.62 \%$ |
| Total | 85,468 | $100.00 \%$ | 62,948 | $100.00 \%$ | 148,416 | $100.00 \%$ |

Table 3.4.3.1C
Reliability: Oral 4-5 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.50 | 3292.827 | 0.850 |
| Speaking | 0.50 | 2190.505 | 0.723 |
| Oral | 2252.993 | 0.878 |  |

*Variances from students who had results in all four domains

Table 3.4.3.1Di
Accuracy and Consistency of Classification Indices: Oral (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.640 | 0.537 |  | 0.356 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.881 |  | 0.787 |  |
|  | 2 | 0.623 |  | 0.482 |  |
|  | 3 | 0.685 |  | 0.548 |  |
|  | 4 | 0.735 |  | 0.614 |  |
|  | 5 | 0.507 |  | 0.445 |  |
|  | 6 | - |  | 0.138 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.987 | 0.004 | 0.009 | 0.982 |
|  | 2/3 | 0.969 | 0.015 | 0.016 | 0.955 |
|  | 3/4 | 0.918 | 0.036 | 0.046 | 0.883 |
|  | 4/5 | 0.810 | 0.052 | 0.138 | 0.749 |
|  | 5/6 | 0.953 | 0.047 | 0.000 | 0.940 |

Table 3.4.3.1Dii
Accuracy and Consistency of Classification Indices: Oral (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.614 | 0.508 |  | 0.337 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.873 |  | 0.779 |  |
|  | 2 | 0.577 |  | 0.437 |  |
|  | 3 | 0.665 |  | 0.526 |  |
|  | 4 | 0.649 |  | 0.540 |  |
|  | 5 | 0.507 |  | 0.439 |  |
|  | 6 | - |  | 0.122 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.980 | 0.008 | 0.013 | 0.971 |
|  | 2/3 | 0.957 | 0.021 | 0.022 | 0.937 |
|  | 3/4 | 0.906 | 0.036 | 0.058 | 0.868 |
|  | 4/5 | 0.808 | 0.073 | 0.120 | 0.747 |
|  | 5/6 | 0.955 | 0.045 | 0.000 | 0.940 |

### 3.4.3.2 Literacy Language Composite 4-5



Figure 3.4.3.2B
Proficiency Level: Litr 4-5 S401 Online


Table 3.4.3.2A
Scale Score Descriptive Statistics: Litr 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 59,348 | 182 | 443 | 332.12 | 31.95 |
| $\mathbf{5}$ | 43,490 | 188 | 457 | 337.83 | 35.74 |
| Total | 102,838 | 182 | 457 | 334.54 | 33.72 |

Table 3.4.3.2B
Proficiency Level Distribution: Litr 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 5,016 | $8.45 \%$ | 4,797 | $11.03 \%$ | 9,813 | $9.54 \%$ |
| $\mathbf{2}$ | 7,862 | $13.25 \%$ | 6,757 | $15.54 \%$ | 14,619 | $14.22 \%$ |
| $\mathbf{3}$ | 31,137 | $52.47 \%$ | 19,619 | $45.11 \%$ | 50,756 | $49.36 \%$ |
| $\mathbf{4}$ | 13,569 | $22.86 \%$ | 10,281 | $23.64 \%$ | 23,850 | $23.19 \%$ |
| $\mathbf{5}$ | 1,463 | $2.47 \%$ | 1,717 | $3.95 \%$ | 3,180 | $3.09 \%$ |
| $\mathbf{6}$ | 301 | $0.51 \%$ | 319 | $0.73 \%$ | 620 | $0.60 \%$ |
| Total | 59,348 | $100.00 \%$ | 43,490 | $100.00 \%$ | 102,838 | $100.00 \%$ |

Table 3.4.3.2C
Reliability: Litr 4-5 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Reading | 0.50 | 1025.399 | 0.880 |
| Writing | 0.50 | 1733.272 | 0.897 |
| Literacy |  | 1106.535 | 0.932 |

[^42]Table 3.4.3.2Di
Accuracy and Consistency of Classification Indices: Litr (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.799 | 0.731 |  | 0.589 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.825 |  | 0.792 |  |
|  | 2 | 0.711 |  | 0.592 |  |
|  | 3 | 0.874 |  | 0.821 |  |
|  | 4 | 0.695 |  | 0.626 |  |
|  | 5 | - |  | 0.231 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.972 | 0.015 | 0.013 | 0.966 |
|  | 2/3 | 0.942 | 0.030 | 0.027 | 0.924 |
|  | 3/4 | 0.900 | 0.041 | 0.059 | 0.866 |
|  | 4/5 | 0.970 | 0.030 | 0.000 | 0.973 |
|  | 5/6 | 0.995 | 0.005 | 0.000 | 1.000 |

Table 3.4.3.2Dii
Accuracy and Consistency of Classification Indices: Litr (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.779 | 0.700 |  | 0.575 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.870 |  | 0.798 |  |
|  | 2 | 0.703 |  | 0.587 |  |
|  | 3 | 0.846 |  | 0.787 |  |
|  | 4 | 0.692 |  | 0.630 |  |
|  | 5 | 0.501 |  | 0.348 |  |
|  | 6 | - |  | 0.894 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.969 | 0.014 | 0.018 | 0.956 |
|  | 2/3 | 0.938 | 0.030 | 0.032 | 0.913 |
|  | 3/4 | 0.915 | 0.037 | 0.047 | 0.881 |
|  | 4/5 | 0.957 | 0.041 | 0.002 | 0.950 |
|  | 5/6 | 0.993 | 0.007 | 0.000 | 0.993 |

### 3.4.3.3 Comprehension Language Composite 4-5




Table 3.4.3.3A
Scale Score Descriptive Statistics: Cphn 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 83,406 | 207 | 474 | 359.65 | 33.63 |
| $\mathbf{5}$ | 60,264 | 205 | 474 | 363.42 | 39.35 |
| Total | 143,670 | 205 | 474 | 361.23 | 36.18 |

Table 3.4.3.3B
Proficiency Level Distribution: Cphn 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 3,775 | $4.53 \%$ | 5,616 | $9.32 \%$ | 9,391 | $6.54 \%$ |
| $\mathbf{2}$ | 9,644 | $11.56 \%$ | 8,965 | $14.88 \%$ | 18,609 | $12.95 \%$ |
| $\mathbf{3}$ | 16,391 | $19.65 \%$ | 11,394 | $18.91 \%$ | 27,785 | $19.34 \%$ |
| $\mathbf{4}$ | 13,658 | $16.38 \%$ | 8,791 | $14.59 \%$ | 22,449 | $15.63 \%$ |
| $\mathbf{5}$ | 20,586 | $24.68 \%$ | 11,876 | $19.71 \%$ | 32,462 | $22.59 \%$ |
| $\mathbf{6}$ | 19,352 | $23.20 \%$ | 13,622 | $22.60 \%$ | 32,974 | $22.95 \%$ |
| Total | 83,406 | $100.00 \%$ | 60,264 | $100.00 \%$ | 143,670 | $100.00 \%$ |

Table 3.4.3.3C
Reliability: Cphn 4-5 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.30 | 3292.827 | 0.850 |
| Reading | 0.70 | 1025.399 | 0.880 |
| Comprehension |  | 1278.180 | 0.918 |

*Variances from students who had results in all four domains

Table 3.4.3.4Di
Accuracy and Consistency of Classification Indices: Cphn (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kарра (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.676 | 0.574 |  | 0.472 |  |
| Conditional on | Level | Accuracy |  | Consistency |  |
| Level | 1 | 0.861 |  | 0.744 |  |
|  | 2 | $0.703$ |  | 0.583 |  |
|  | 3 | $0.650$ |  | 0.532 |  |
|  | 4 | $0.482$ |  | 0.376 |  |
|  | 5 | $0.639$ |  | 0.531 |  |
|  | 6 | $0.830$ |  | 0.744 |  |
| Indices at Cut <br> Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.982 | 0.005 | 0.013 | 0.975 |
|  | 2/3 | 0.951 | 0.023 | 0.027 | 0.929 |
|  | 3/4 | 0.912 | 0.045 | 0.043 | 0.876 |
|  | 4/5 | 0.898 | 0.050 | 0.052 | 0.860 |
|  | 5/6 | 0.921 | 0.040 | 0.039 | 0.887 |

Table 3.4.3.4Dii
Accuracy and Consistency of Classification Indices: Cphn (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.654 | 0.554 |  | 0.458 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.862 |  | 0.760 |  |
|  | 2 | 0.663 |  | 0.544 |  |
|  | 3 | 0.586 |  | 0.470 |  |
|  | 4 | 0.439 |  | 0.339 |  |
|  | 5 | 0.572 |  | 0.461 |  |
|  | 6 | 0.841 |  | 0.754 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.966 | 0.011 | 0.022 | 0.953 |
|  | 2/3 | 0.935 | 0.031 | 0.034 | 0.907 |
|  | 3/4 | 0.906 | 0.050 | 0.044 | 0.869 |
|  | 4/5 | 0.901 | 0.049 | 0.050 | 0.863 |
|  | 5/6 | 0.925 | 0.040 | 0.035 | 0.893 |

### 3.4.3.4 Overall Language Composite 4-5



Figure 3.4.3.4B
Proficiency Level: Over 4-5S401 Online


Table 3.4.3.4A
Scale Score Descriptive Statistics: Over 4-5 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{4}$ | 45,375 | 176 | 445 | 339.17 | 33.02 |
| $\mathbf{5}$ | 33,626 | 189 | 450 | 343.86 | 37.97 |
| Total | 79,001 | 176 | 450 | 341.16 | 35.29 |

Table 3.4.3.4B
Proficiency Level Distribution: Over 4-5 S401 Online

| Level | Grade 4 |  | Grade 5 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,504 | $5.52 \%$ | 2,842 | $8.45 \%$ | 5,346 | $6.77 \%$ |
| $\mathbf{2}$ | 4,597 | $10.13 \%$ | 4,163 | $12.38 \%$ | 8,760 | $11.09 \%$ |
| $\mathbf{3}$ | 18,992 | $41.86 \%$ | 12,699 | $37.77 \%$ | 31,691 | $40.11 \%$ |
| $\mathbf{4}$ | 16,885 | $37.21 \%$ | 11,581 | $34.44 \%$ | 28,466 | $36.03 \%$ |
| $\mathbf{5}$ | 2,142 | $4.72 \%$ | 2,131 | $6.34 \%$ | 4,273 | $5.41 \%$ |
| $\mathbf{6}$ | 255 | $0.56 \%$ | 210 | $0.62 \%$ | 465 | $0.59 \%$ |
| Total | 45,375 | $100.00 \%$ | 33,626 | $100.00 \%$ | 79,001 | $100.00 \%$ |

Table 3.4.3.4C
Reliability: Over 4-5 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.15 | 3292.827 | 0.850 |
| Reading | 0.35 | 1025.399 | 0.880 |
| Speaking | 0.15 | 2190.505 | 0.723 |
| Writing | 0.35 | 1733.272 | 0.897 |
| Overall Composite |  |  |  | $1245.101 \quad 0.950 \quad$.

*Variances from students who had results in all four domains

Table 3.4.3.4Di
Accuracy and Consistency of Classification Indices: Over (Grade 4) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.822 | 0.769 |  | 0.653 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.812 |  | 0.831 |  |
|  | 2 | 0.768 |  | 0.662 |  |
|  | 3 | 0.886 |  | 0.833 |  |
|  | 4 | 0.779 |  | 0.747 |  |
|  | 5 | - |  | 0.279 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.982 | 0.011 | 0.007 | 0.982 |
|  | 2/3 | 0.961 | 0.022 | 0.017 | 0.952 |
|  | 3/4 | 0.917 | 0.034 | 0.049 | 0.891 |
|  | 4/5 | 0.947 | 0.053 | 0.000 | 0.945 |
|  | 5/6 | 0.994 | 0.006 | 0.000 | 1.000 |

Table 3.4.3.4Dii
Accuracy and Consistency of Classification Indices: Over (Grade 5) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.794 | 0.728 |  | 0.616 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.864 |  | 0.836 |  |
|  | 2 | 0.754 |  | 0.648 |  |
|  | 3 | 0.861 |  | 0.797 |  |
|  | 4 | 0.735 |  | 0.708 |  |
|  | 5 | - |  | 0.340 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.978 | 0.012 | 0.011 | 0.973 |
|  | 2/3 | 0.956 | 0.023 | 0.021 | 0.941 |
|  | 3/4 | 0.921 | 0.032 | 0.046 | 0.893 |
|  | 4/5 | 0.930 | 0.070 | 0.000 | 0.922 |
|  | 5/6 | 0.994 | 0.006 | 0.000 | 0.997 |

### 3.4.4 Grades: 6-8

### 3.4.4.1 Oral Language Composite 6-8




Table 3.4.4.1 A
Scale Score Descriptive Statistics: Oral 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 47,283 | 161 | 475 | 350.63 | 41.68 |
| $\mathbf{7}$ | 47,065 | 164 | 487 | 357.98 | 46.97 |
| $\mathbf{8}$ | 47,355 | 171 | 493 | 362.92 | 51.56 |
| Total | 141,703 | 161 | 493 | 357.18 | 47.18 |

Table 3.4.4.1B
Proficiency Level Distribution: Oral 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 2,658 | $5.62 \%$ | 3,733 | $7.93 \%$ | 4,738 | $10.01 \%$ | 11,129 | $7.85 \%$ |
| $\mathbf{2}$ | 5,921 | $12.52 \%$ | 6,193 | $13.16 \%$ | 6,408 | $13.53 \%$ | 18,522 | $13.07 \%$ |
| $\mathbf{3}$ | 18,769 | $39.70 \%$ | 16,407 | $34.86 \%$ | 14,774 | $31.20 \%$ | 49,950 | $35.25 \%$ |
| $\mathbf{4}$ | 15,592 | $32.98 \%$ | 15,346 | $32.61 \%$ | 16,123 | $34.05 \%$ | 47,061 | $33.21 \%$ |
| $\mathbf{5}$ | 3,782 | $8.00 \%$ | 4,674 | $9.93 \%$ | 4,435 | $9.37 \%$ | 12,891 | $9.10 \%$ |
| $\mathbf{6}$ | 561 | $1.19 \%$ | 712 | $1.51 \%$ | 877 | $1.85 \%$ | 2,150 | $1.52 \%$ |
| Total | 47,283 | $100.00 \%$ | 47,065 | $100.00 \%$ | 47,355 | $100.00 \%$ | 141,703 | $100.00 \%$ |

Table 3.4.4.1C
Reliability: Oral 6-8 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.50 | 3391.687 | 0.870 |
| Speaking | 0.50 | 2114.953 | 0.705 |
| Oral | 2290.219 | 0.884 |  |

*Variances from students who had results in all four domains

Table 3.4.4.1Di
Accuracy and Consistency of Classification Indices: Oral (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.704 | 0.601 |  | 0.440 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.833 |  | 0.712 |  |
|  | 2 | 0.637 |  | 0.499 |  |
|  | 3 | 0.787 |  | 0.696 |  |
|  | 4 | 0.654 |  | 0.588 |  |
|  | 5 | 0.506 |  | 0.356 |  |
|  | 6 | - |  | 0.315 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.978 | 0.009 | 0.014 | 0.968 |
|  | 2/3 | 0.933 | 0.035 | 0.031 | 0.903 |
|  | 3/4 | 0.878 | 0.046 | 0.076 | 0.830 |
|  | 4/5 | 0.920 | 0.068 | 0.012 | 0.897 |
|  | 5/6 | 0.988 | 0.012 | 0.000 | 0.988 |

Table 3.4.4.1 Dii
Accuracy and Consistency of Classification Indices: Oral (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.673 | 0.565 |  | 0.413 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.837 |  | 0.721 |  |
|  | 2 | 0.611 |  | 0.475 |  |
|  | 3 | 0.745 |  | 0.641 |  |
|  | 4 | 0.633 |  | 0.562 |  |
|  | 5 | 0.505 |  | 0.368 |  |
|  | 6 | - |  | 0.117 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.971 | 0.012 | 0.017 | 0.957 |
|  | 2/3 | 0.928 | 0.037 | 0.036 | 0.896 |
|  | 3/4 | 0.878 | 0.046 | 0.076 | 0.831 |
|  | 4/5 | 0.903 | 0.074 | 0.022 | 0.874 |
|  | 5/6 | 0.985 | 0.015 | 0.000 | 0.983 |

Table 3.4.4.1Diii
Accuracy and Consistency of Classification Indices: Oral (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.658 | 0.546 |  | 0.395 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.842 |  | 0.735 |  |
|  | 2 | 0.600 |  | 0.466 |  |
|  | 3 | 0.703 |  | 0.587 |  |
|  | 4 | 0.629 |  | 0.570 |  |
|  | 5 | 0.428 |  | 0.304 |  |
|  | 6 | - |  | 0.099 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.965 | 0.015 | 0.020 | 0.949 |
|  | 2/3 | 0.925 | 0.036 | 0.038 | 0.894 |
|  | 3/4 | 0.877 | 0.048 | 0.076 | 0.828 |
|  | 4/5 | 0.894 | 0.087 | 0.019 | 0.861 |
|  | 5/6 | 0.981 | 0.019 | 0.000 | 0.977 |

### 3.4.4.2 Literacy Language Composite 6-8



Figure 3.4.4.2B
Proficiency Level: Litr 6-8 S401 Online


Table 3.4.4.2 A
Scale Score Descriptive Statistics: Litr 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 52,892 | 202 | 452 | 326.49 | 36.69 |
| $\mathbf{7}$ | 53,499 | 214 | 453 | 333.57 | 39.63 |
| $\mathbf{8}$ | 51,693 | 213 | 490 | 340.48 | 41.66 |
| Total | 158,084 | 202 | 490 | 333.46 | 39.77 |

Table 3.4.4.2B
Proficiency Level Distribution: Litr 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 11,793 | $22.30 \%$ | 12,305 | $23.00 \%$ | 12,163 | $23.53 \%$ | 36,261 | $22.94 \%$ |
| $\mathbf{2}$ | 12,616 | $23.85 \%$ | 12,900 | $24.11 \%$ | 12,118 | $23.44 \%$ | 37,634 | $23.81 \%$ |
| $\mathbf{3}$ | 21,427 | $40.51 \%$ | 19,436 | $36.33 \%$ | 16,843 | $32.58 \%$ | 57,706 | $36.50 \%$ |
| $\mathbf{4}$ | 6,476 | $12.24 \%$ | 7,853 | $14.68 \%$ | 9,274 | $17.94 \%$ | 23,603 | $14.93 \%$ |
| $\mathbf{5}$ | 533 | $1.01 \%$ | 922 | $1.72 \%$ | 1,206 | $2.33 \%$ | 2,661 | $1.68 \%$ |
| $\mathbf{6}$ | 47 | $0.09 \%$ | 83 | $0.16 \%$ | 89 | $0.17 \%$ | 219 | $0.14 \%$ |
| Total | 52,892 | $100.00 \%$ | 53,499 | $100.00 \%$ | 51,693 | $100.00 \%$ | 158,084 | $100.00 \%$ |

Table 3.4.4.2 C
Reliability: Litr 6-8 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Reading | 0.50 | 1625.254 | 0.920 |
| Writing | 0.50 | 2000.288 | 0.890 |
| Literacy |  | 1597.394 | 0.945 |

*Variances from students who had results in all four domains

Table 3.4.4.2Di
Accuracy and Consistency of Classification Indices: Litr (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.823 | 0.752 |  | 0.654 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.906 |  | 0.855 |  |
|  | 2 | 0.743 |  | 0.650 |  |
|  | 3 | 0.852 |  | 0.800 |  |
|  | 4 | 0.748 |  | 0.652 |  |
|  | 5 | 0.628 |  | 0.372 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.952 | 0.020 | 0.028 | 0.933 |
|  | 2/3 | 0.930 | 0.036 | 0.034 | 0.902 |
|  | 3/4 | 0.950 | 0.027 | 0.023 | 0.929 |
|  | 4/5 | 0.990 | 0.009 | 0.001 | 0.988 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

Table 3.4.4.2Dii
Accuracy and Consistency of Classification Indices: Litr (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.811 | 0.734 |  | 0.640 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.905 |  | 0.852 |  |
|  | 2 | 0.745 |  | 0.652 |  |
|  | 3 | 0.826 |  | 0.765 |  |
|  | 4 | 0.761 |  | 0.666 |  |
|  | 5 | 0.585 |  | 0.389 |  |
|  | 6 | - |  | 0.864 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.950 | 0.021 | 0.029 | 0.930 |
|  | 2/3 | 0.931 | 0.035 | 0.034 | 0.903 |
|  | 3/4 | 0.945 | 0.030 | 0.025 | 0.923 |
|  | 4/5 | 0.986 | 0.011 | 0.003 | 0.979 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.999 |

Table 3.4.4.2Diii
Accuracy and Consistency of Classification Indices: Litr (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.798 | 0.717 |  | 0.624 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.900 |  | 0.847 |  |
|  | 2 | 0.750 |  | 0.654 |  |
|  | 3 | 0.796 |  | 0.722 |  |
|  | 4 | 0.756 |  | 0.675 |  |
|  | 5 | 0.530 |  | 0.334 |  |
|  | 6 | - |  | 0.300 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.950 | 0.023 | 0.027 | 0.929 |
|  | 2/3 | 0.933 | 0.032 | 0.035 | 0.906 |
|  | 3/4 | 0.938 | 0.033 | 0.029 | 0.913 |
|  | 4/5 | 0.978 | 0.017 | 0.005 | 0.968 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.998 |

### 3.4.4.3 Comprehension Language Composite 6-8



Figure 3.4.4.3B
ProficiencyLevel: Cphn 6-8 S401 Online


Table 3.4.4.3A
Scale Score Descriptive Statistics: Cphn 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 50,026 | 238 | 486 | 349.91 | 37.32 |
| $\mathbf{7}$ | 49,974 | 220 | 486 | 358.12 | 41.86 |
| $\mathbf{8}$ | 48,606 | 224 | 486 | 365.31 | 45.29 |
| Total | 148,606 | 220 | 486 | 357.71 | 42.05 |

Table 3.4.4.3B
Proficiency Level Distribution: Cphn 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 9,378 | $18.75 \%$ | 10,570 | $21.15 \%$ | 10,727 | $22.07 \%$ | 30,675 | $20.64 \%$ |
| $\mathbf{2}$ | 13,109 | $26.20 \%$ | 11,593 | $23.20 \%$ | 10,552 | $21.71 \%$ | 35,254 | $23.72 \%$ |
| $\mathbf{3}$ | 11,828 | $23.64 \%$ | 10,338 | $20.69 \%$ | 8,835 | $18.18 \%$ | 31,001 | $20.86 \%$ |
| $\mathbf{4}$ | 6,224 | $12.44 \%$ | 6,094 | $12.19 \%$ | 5,491 | $11.30 \%$ | 17,809 | $11.98 \%$ |
| $\mathbf{5}$ | 5,478 | $10.95 \%$ | 5,741 | $11.49 \%$ | 6,338 | $13.04 \%$ | 17,557 | $11.81 \%$ |
| $\mathbf{6}$ | 4,009 | $8.01 \%$ | 5,638 | $11.28 \%$ | 6,663 | $13.71 \%$ | 16,310 | $10.98 \%$ |
| Total | 50,026 | $100.00 \%$ | 49,974 | $100.00 \%$ | 48,606 | $100.00 \%$ | 148,606 | $100.00 \%$ |

Table 3.4.4.3C
Reliability: Cphn 6-8 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.30 | 3391.687 | 0.870 |
| Reading | 0.70 | 1625.254 | 0.920 |
| Comprehension |  | 1813.255 | 0.943 |

*Variances from students who had results in all four domains

Table 3.4.4.3Di
Accuracy and Consistency of Classification Indices: Cphn (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.725 | 0.632 |  | 0.543 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.894 |  | 0.828 |  |
|  | 2 | 0.752 |  | 0.665 |  |
|  | 3 | 0.688 |  | 0.585 |  |
|  | 4 | 0.521 |  | 0.406 |  |
|  | 5 | 0.611 |  | 0.485 |  |
|  | 6 | 0.847 |  | 0.740 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.951 | 0.019 | 0.030 | 0.932 |
|  | 2/3 | 0.927 | 0.040 | 0.034 | 0.897 |
|  | 3/4 | 0.926 | 0.042 | 0.033 | 0.897 |
|  | 4/5 | 0.944 | 0.029 | 0.028 | 0.920 |
|  | 5/6 | 0.971 | 0.018 | 0.011 | 0.959 |

Table 3.4.4.3Dii
Accuracy and Consistency of Classification Indices: Cphn (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.713 | 0.619 |  | 0.534 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.894 |  | 0.833 |  |
|  | 2 | 0.717 |  | 0.620 |  |
|  | 3 | 0.652 |  | 0.542 |  |
|  | 4 | 0.501 |  | 0.388 |  |
|  | 5 | 0.585 |  | 0.461 |  |
|  | 6 | 0.856 |  | 0.762 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.947 | 0.021 | 0.032 | 0.926 |
|  | 2/3 | 0.929 | 0.038 | 0.033 | 0.899 |
|  | 3/4 | 0.926 | 0.041 | 0.033 | 0.897 |
|  | 4/5 | 0.940 | 0.032 | 0.029 | 0.915 |
|  | 5/6 | 0.963 | 0.022 | 0.015 | 0.947 |

Table 3.4.4.3Diii
Accuracy and Consistency of Classification Indices: Cphn (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.705 | 0.612 |  | 0.529 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.892 |  | 0.832 |  |
|  | 2 | 0.704 |  | 0.602 |  |
|  | 3 | 0.612 |  | 0.497 |  |
|  | 4 | 0.462 |  | 0.354 |  |
|  | 5 | 0.592 |  | 0.472 |  |
|  | 6 | 0.854 |  | 0.765 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.946 | 0.023 | 0.031 | 0.924 |
|  | 2/3 | 0.930 | 0.036 | 0.034 | 0.901 |
|  | 3/4 | 0.927 | 0.040 | 0.033 | 0.898 |
|  | 4/5 | 0.935 | 0.035 | 0.030 | 0.909 |
|  | 5/6 | 0.956 | 0.025 | 0.019 | 0.937 |

### 3.4.4.4 Overall Language Composite 6-8



Figure 3.4.4.4B
Proficiency Level: Over 6-8S401 Online


Table 3.4.4.4A
Scale Score Descriptive Statistics: Over 6-8 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{6}$ | 36,879 | 207 | 454 | 334.09 | 36.03 |
| $\mathbf{7}$ | 36,536 | 209 | 455 | 341.35 | 40.14 |
| $\mathbf{8}$ | 36,097 | 216 | 481 | 347.29 | 43.32 |
| Total | 109,512 | 207 | 481 | 340.86 | 40.28 |

Table 3.4.4.4B
Proficiency Level Distribution: Over 6-8 S401 Online

| Level | Grade 6 |  | Grade 7 |  | Grade 8 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
|  | 4,452 | $12.07 \%$ | 5,346 | $14.63 \%$ | 6,063 | $16.80 \%$ | 15,861 | $14.48 \%$ |
| $\mathbf{2}$ | 9,175 | $24.88 \%$ | 8,592 | $23.52 \%$ | 7,943 | $22.00 \%$ | 25,710 | $23.48 \%$ |
| $\mathbf{3}$ | 15,730 | $42.65 \%$ | 13,531 | $37.03 \%$ | 11,892 | $32.94 \%$ | 41,153 | $37.58 \%$ |
| $\mathbf{4}$ | 6,847 | $18.57 \%$ | 8,021 | $21.95 \%$ | 8,879 | $24.60 \%$ | 23,747 | $21.68 \%$ |
| $\mathbf{5}$ | 627 | $1.70 \%$ | 967 | $2.65 \%$ | 1,247 | $3.45 \%$ | 2,841 | $2.59 \%$ |
| $\mathbf{6}$ | 48 | $0.13 \%$ | 79 | $0.22 \%$ | 73 | $0.20 \%$ | 200 | $0.18 \%$ |
| Total | 36,879 | $100.00 \%$ | 36,536 | $100.00 \%$ | 36,097 | $100.00 \%$ | 109,512 | $100.00 \%$ |

Table 3.4.4.4C
Reliability: Over 6-8 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.15 | 3391.687 | 0.870 |
| Reading | 0.35 | 1625.254 | 0.920 |
| Speaking | 0.15 | 2114.953 | 0.705 |
| Writing | 0.35 | 2000.288 | 0.890 |
| Overall Composite |  |  |  | $1622.385 \quad 0.959 \quad$.

*Variances from students who had results in all four domains

Table 3.4.4.4Di
Accuracy and Consistency of Classification Indices: Over (Grade 6) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.841 | 0.782 |  | 0.693 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.877 |  | 0.821 |  |
|  | 2 | 0.832 |  | 0.761 |  |
|  | 3 | 0.866 |  | 0.819 |  |
|  | 4 | 0.779 |  | 0.722 |  |
|  | 5 | - |  | 0.435 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.970 | 0.015 | 0.015 | 0.959 |
|  | 2/3 | 0.942 | 0.028 | 0.031 | 0.919 |
|  | 3/4 | 0.944 | 0.028 | 0.028 | 0.922 |
|  | 4/5 | 0.982 | 0.018 | 0.000 | 0.982 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

Table 3.4.4.4Dii
Accuracy and Consistency of Classification Indices: Over (Grade 7) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.826 | 0.760 |  | 0.675 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.881 |  | 0.829 |  |
|  | 2 | 0.812 |  | 0.734 |  |
|  | 3 | 0.839 |  | 0.781 |  |
|  | 4 | 0.790 |  | 0.737 |  |
|  | 5 | 0.742 |  | 0.439 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.966 | 0.017 | 0.017 | 0.953 |
|  | 2/3 | 0.941 | 0.028 | 0.031 | 0.919 |
|  | 3/4 | 0.941 | 0.030 | 0.029 | 0.919 |
|  | 4/5 | 0.974 | 0.023 | 0.003 | 0.971 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.999 |

Table 3.4.4.4Diii
Accuracy and Consistency of Classification Indices: Over (Grade 8) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.816 | 0.742 |  | 0.658 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.899 |  | 0.845 |  |
|  | 2 | 0.802 |  | 0.719 |  |
|  | 3 | 0.816 |  | 0.747 |  |
|  | 4 | 0.789 |  | 0.735 |  |
|  | 5 | 0.582 |  | 0.385 |  |
|  | 6 | - |  | 0.889 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.965 | 0.017 | 0.018 | 0.951 |
|  | 2/3 | 0.944 | 0.025 | 0.031 | 0.922 |
|  | 3/4 | 0.938 | 0.031 | 0.030 | 0.913 |
|  | 4/5 | 0.968 | 0.025 | 0.006 | 0.958 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.998 |

### 3.4.5 Grades: 9-12

### 3.4.5.1 Oral Language Composite 9-12



Figure 3.4.5.1B
Proficiency Level: Oral 9-12 S401 Online


Table 3.4.5.1A
Scale Score Descriptive Statistics: Oral 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 61,666 | 154 | 491 | 347.85 | 47.25 |
| $\mathbf{1 0}$ | 41,084 | 207 | 484 | 348.68 | 45.25 |
| $\mathbf{1 1}$ | 28,692 | 215 | 501 | 351.22 | 44.70 |
| $\mathbf{1 2}$ | 19,971 | 217 | 477 | 354.46 | 43.09 |
| Total | 151,413 | 154 | 501 | 349.58 | 45.75 |

Table 3.4.5.1B
Proficiency Level Distribution: Oral 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 10,143 | $16.45 \%$ | 7,150 | $17.40 \%$ | 5,382 | $18.76 \%$ | 3,674 | $18.40 \%$ | 26,349 | $17.40 \%$ |
| $\mathbf{2}$ | 13,979 | $22.67 \%$ | 9,884 | $24.06 \%$ | 7,372 | $25.69 \%$ | 5,194 | $26.01 \%$ | 36,429 | $24.06 \%$ |
| $\mathbf{3}$ | 23,042 | $37.37 \%$ | 16,894 | $41.12 \%$ | 11,703 | $40.79 \%$ | 8,645 | $43.29 \%$ | 60,284 | $39.81 \%$ |
| $\mathbf{4}$ | 12,577 | $20.40 \%$ | 6,369 | $15.50 \%$ | 3,708 | $12.92 \%$ | 2,214 | $11.09 \%$ | 24,868 | $16.42 \%$ |
| $\mathbf{5}$ | 1,772 | $2.87 \%$ | 738 | $1.80 \%$ | 494 | $1.72 \%$ | 224 | $1.12 \%$ | 3,228 | $2.13 \%$ |
| $\mathbf{6}$ | 153 | $0.25 \%$ | 49 | $0.12 \%$ | 33 | $0.12 \%$ | 20 | $0.10 \%$ | 255 | $0.17 \%$ |
| Total | 61,666 | $100.00 \%$ | 41,084 | $100.00 \%$ | 28,692 | $100.00 \%$ | 19,971 | $100.00 \%$ | 151,413 | $100.00 \%$ |

Table 3.4.5.1 C
Reliability: Oral 9-12 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.50 | 2992.442 | 0.870 |
| Speaking | 0.50 | 2358.334 | 0.658 |
| Oral | 2210.830 | 0.865 |  |

*Variances from students who had results in all four domains

Table 3.4.5.1 Di
Accuracy and Consistency of Classification Indices: Oral (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.692 | 0.582 |  | 0.438 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.833 |  | 0.724 |  |
|  | 2 | 0.619 |  | 0.502 |  |
|  | 3 | 0.709 |  | 0.617 |  |
|  | 4 | 0.645 |  | 0.545 |  |
|  | 5 | 0.515 |  | 0.263 |  |
|  | 6 | - |  | 0.109 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.937 | 0.026 | 0.037 | 0.910 |
|  | 2/3 | 0.890 | 0.054 | 0.056 | 0.846 |
|  | 3/4 | 0.893 | 0.055 | 0.052 | 0.849 |
|  | 4/5 | 0.969 | 0.030 | 0.001 | 0.959 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.997 |

Table 3.4.5.1Dii
Accuracy and Consistency of Classification Indices: Oral (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.699 | 0.592 |  | 0.437 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.827 |  | 0.721 |  |
|  | 2 | 0.637 |  | 0.517 |  |
|  | 3 | 0.728 |  | 0.643 |  |
|  | 4 | 0.591 |  | 0.474 |  |
|  | 5 | - |  | 0.156 |  |
|  | 6 | - |  | 0.667 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
|  | 1/2 | 0.936 | 0.029 | 0.035 | 0.908 |
|  | 2/3 | 0.886 | 0.054 | 0.061 | 0.840 |
|  | 3/4 | 0.894 | 0.054 | 0.052 | 0.849 |
|  | 4/5 | 0.981 | 0.019 | 0.000 | 0.976 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 0.999 |

Table 3.4.5.1 Diii
Accuracy and Consistency of Classification Indices: Oral (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.699 | 0.593 |  | 0.435 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.819 |  | 0.714 |  |
|  | 2 | 0.647 |  | 0.527 |  |
|  | 3 | 0.724 |  | 0.641 |  |
|  | 4 | 0.556 |  | 0.433 |  |
|  | 5 | - |  | 0.159 |  |
|  | 6 | - |  | 0.600 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.931 | 0.034 | 0.035 | 0.900 |
|  | 2/3 | 0.881 | 0.053 | 0.066 | 0.834 |
|  | 3/4 | 0.902 | 0.052 | 0.045 | 0.860 |
|  | 4/5 | 0.982 | 0.018 | 0.000 | 0.978 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 0.999 |

Table 3.4.5.1Div
Accuracy and Consistency of Classification Indices: Oral (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.713 | 0.610 |  | 0.445 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.802 |  | 0.696 |  |
|  | 2 | 0.645 |  | 0.527 |  |
|  | 3 | 0.746 |  | 0.673 |  |
|  | 4 | 0.573 |  | 0.431 |  |
|  | 5 | - |  | 0.145 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False Negatives | Consistency |
|  | 1/2 | 0.928 | 0.037 | 0.036 | 0.896 |
|  | 2/3 | 0.878 | 0.055 | 0.067 | 0.832 |
|  | 3/4 | 0.914 | 0.050 | 0.036 | 0.877 |
|  | 4/5 | 0.988 | 0.012 | 0.000 | 0.987 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

### 3.4.5.2 Literacy Language Composite 9-12



Figure 3.4.5.2B
Proficiency Level: Litr 9-12 S401 Online


Table 3.4.5.2A
Scale Score Descriptive Statistics: Litr 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 59,767 | 235 | 477 | 351.47 | 38.83 |
| $\mathbf{1 0}$ | 38,625 | 230 | 474 | 357.56 | 34.27 |
| $\mathbf{1 1}$ | 26,179 | 252 | 487 | 362.91 | 33.67 |
| $\mathbf{1 2}$ | 18,165 | 261 | 482 | 366.57 | 31.97 |
| Total | 142,736 | 230 | 487 | 357.14 | 36.31 |

Table 3.4.5.2B
Proficiency Level Distribution: Litr 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 12,166 | $20.36 \%$ | 6,178 | $15.99 \%$ | 4,064 | $15.52 \%$ | 2,882 | $15.87 \%$ | 25,290 | $17.72 \%$ |
| $\mathbf{2}$ | 13,606 | $22.77 \%$ | 10,399 | $26.92 \%$ | 7,803 | $29.81 \%$ | 5,778 | $31.81 \%$ | 37,586 | $26.33 \%$ |
| $\mathbf{3}$ | 21,157 | $35.40 \%$ | 15,177 | $39.29 \%$ | 9,853 | $37.64 \%$ | 7,090 | $39.03 \%$ | 53,277 | $37.33 \%$ |
| $\mathbf{4}$ | 10,324 | $17.27 \%$ | 5,732 | $14.84 \%$ | 3,676 | $14.04 \%$ | 1,988 | $10.94 \%$ | 21,720 | $15.22 \%$ |
| $\mathbf{5}$ | 2,301 | $3.85 \%$ | 1,069 | $2.77 \%$ | 741 | $2.83 \%$ | 421 | $2.32 \%$ | 4,532 | $3.18 \%$ |
| $\mathbf{6}$ | 213 | $0.36 \%$ | 70 | $0.18 \%$ | 42 | $0.16 \%$ | 6 | $0.03 \%$ | 331 | $0.23 \%$ |
| Total | 59,767 | $100.00 \%$ | 38,625 | $100.00 \%$ | 26,179 | $100.00 \%$ | 18,165 | $100.00 \%$ | 142,736 | $100.00 \%$ |

Table 3.4.5.2 C
Reliability: Litr 9-12 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Reading | 0.50 | 1432.799 | 0.910 |
| Writing | 0.50 | 1677.917 | 0.867 |
| Literacy |  | 1309.167 | 0.933 |

*Variances from students who had results in all four domains

Table 3.4.5.2Di
Accuracy and Consistency of Classification Indices: Litr (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.782 | 0.695 |  | 0.594 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.893 |  | 0.826 |  |
|  | 2 | 0.701 |  | 0.601 |  |
|  | 3 | 0.798 |  | 0.728 |  |
|  | 4 | 0.757 |  | 0.654 |  |
|  | 5 | 0.692 |  | 0.537 |  |
|  | 6 | 0.778 |  | 0.449 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.946 | 0.020 | 0.034 | 0.925 |
|  | 2/3 | 0.924 | 0.039 | 0.036 | 0.893 |
|  | 3/4 | 0.936 | 0.036 | 0.028 | 0.910 |
|  | 4/5 | 0.978 | 0.014 | 0.008 | 0.968 |
|  | 5/6 | 0.997 | 0.003 | 0.000 | 0.996 |

Table 3.4.5.2Dii
Accuracy and Consistency of Classification Indices: Litr (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.802 | 0.721 |  | 0.616 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.870 |  | 0.788 |  |
|  | 2 | 0.748 |  | 0.661 |  |
|  | 3 | 0.825 |  | 0.765 |  |
|  | 4 | 0.780 |  | 0.678 |  |
|  | 5 | 0.757 |  | 0.603 |  |
|  | 6 | 0.870 |  | 0.658 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.950 | 0.019 | 0.031 | 0.930 |
|  | 2/3 | 0.921 | 0.040 | 0.039 | 0.889 |
|  | 3/4 | 0.946 | 0.031 | 0.023 | 0.924 |
|  | 4/5 | 0.986 | 0.009 | 0.005 | 0.980 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 0.999 |

Table 3.4.5.2 Diii
Accuracy and Consistency of Classification Indices: Litr (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.801 | 0.720 |  | 0.614 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.860 |  | 0.773 |  |
|  | 2 | 0.774 |  | 0.693 |  |
|  | 3 | 0.813 |  | 0.749 |  |
|  | 4 | 0.771 |  | 0.665 |  |
|  | 5 | 0.759 |  | 0.604 |  |
|  | 6 | 0.871 |  | 0.663 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.951 | 0.021 | 0.029 | 0.930 |
|  | 2/3 | 0.919 | 0.040 | 0.041 | 0.886 |
|  | 3/4 | 0.947 | 0.031 | 0.022 | 0.926 |
|  | 4/5 | 0.986 | 0.009 | 0.005 | 0.979 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 0.999 |

Table 3.4.5.2Div
Accuracy and Consistency of Classification Indices: Litr (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.809 | 0.730 |  | 0.621 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.853 |  | 0.765 |  |
|  | 2 | 0.783 |  | 0.704 |  |
|  | 3 | 0.830 |  | 0.770 |  |
|  | 4 | 0.750 |  | 0.635 |  |
|  | 5 | 0.819 |  | 0.665 |  |
|  | 6 | 1.000 |  | 0.737 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.949 | 0.022 | 0.029 | 0.927 |
|  | 2/3 | 0.916 | 0.042 | 0.042 | 0.882 |
|  | 3/4 | 0.956 | 0.026 | 0.019 | 0.937 |
|  | 4/5 | 0.989 | 0.008 | 0.003 | 0.984 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

### 3.4.5.3 Comprehension Language Composite 9-12




Table 3.4.5.3A
Scale Score Descriptive Statistics: Cphn 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 56,813 | 246 | 502 | 368.17 | 40.84 |
| $\mathbf{1 0}$ | 36,742 | 269 | 502 | 371.84 | 38.44 |
| $\mathbf{1 1}$ | 24,906 | 268 | 502 | 376.07 | 38.52 |
| $\mathbf{1 2}$ | 17,298 | 277 | 502 | 379.12 | 36.99 |
| Total | 135,759 | 246 | 502 | 372.01 | 39.50 |

Table 3.4.5.3B
Proficiency Level Distribution: Cphn 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 12,623 | $22.22 \%$ | 7,811 | $21.26 \%$ | 5,487 | $22.03 \%$ | 3,686 | $21.31 \%$ | 29,607 | $21.81 \%$ |
| $\mathbf{2}$ | 15,611 | $27.48 \%$ | 10,783 | $29.35 \%$ | 7,748 | $31.11 \%$ | 5,782 | $33.43 \%$ | 39,924 | $29.41 \%$ |
| $\mathbf{3}$ | 11,163 | $19.65 \%$ | 8,037 | $21.87 \%$ | 5,069 | $20.35 \%$ | 3,568 | $20.63 \%$ | 27,837 | $20.50 \%$ |
| $\mathbf{4}$ | 5,356 | $9.43 \%$ | 3,589 | $9.77 \%$ | 2,150 | $8.63 \%$ | 1,596 | $9.23 \%$ | 12,691 | $9.35 \%$ |
| $\mathbf{5}$ | 5,744 | $10.11 \%$ | 3,318 | $9.03 \%$ | 2,333 | $9.37 \%$ | 1,381 | $7.98 \%$ | 12,776 | $9.41 \%$ |
| $\mathbf{6}$ | 6,316 | $11.12 \%$ | 3,204 | $8.72 \%$ | 2,119 | $8.51 \%$ | 1,285 | $7.43 \%$ | 12,924 | $9.52 \%$ |
| Total | 56,813 | $100.00 \%$ | 36,742 | $100.00 \%$ | 24,906 | $100.00 \%$ | 17,298 | $100.00 \%$ | 135,759 | $100.00 \%$ |

Table 3.4.5.3C
Reliability: Cphn 9-12 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.30 | 2992.442 | 0.870 |
| Reading | 0.70 | 1432.799 | 0.910 |
| Comprehension | 1591.969 | 0.938 |  |

*Variances from students who had results in all four domains

Table 3.4.5.3Di
Accuracy and Consistency of Classification Indices: Cphn (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.721 | 0.628 |  | 0.538 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.868 |  | 0.799 |  |
|  | 2 | 0.729 |  | 0.639 |  |
|  | 3 | 0.658 |  | 0.544 |  |
|  | 4 | 0.461 |  | 0.349 |  |
|  | 5 | 0.600 |  | 0.474 |  |
|  | 6 | 0.875 |  | 0.794 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives | Consistency |
|  | 1/2 | 0.933 | 0.028 | 0.039 | 0.907 |
|  | 2/3 | 0.926 | 0.041 | 0.033 | 0.896 |
|  | 3/4 | 0.936 | 0.034 | 0.030 | 0.911 |
|  | 4/5 | 0.950 | 0.028 | 0.022 | 0.928 |
|  | 5/6 | 0.968 | 0.019 | 0.013 | 0.954 |

Table 3.4.5.3Dii
Accuracy and Consistency of Classification Indices: Cphn (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.728 | 0.634 |  | 0.540 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.867 |  | 0.794 |  |
|  | 2 | 0.749 |  | 0.664 |  |
|  | 3 | 0.683 |  | 0.575 |  |
|  | 4 | 0.484 |  | 0.368 |  |
|  | 5 | 0.590 |  | 0.462 |  |
|  | 6 | 0.862 |  | 0.768 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False <br> Negatives |  |
|  | 1/2 | 0.937 | 0.027 | 0.036 | 0.911 |
|  | 2/3 | 0.924 | 0.041 | 0.035 | 0.892 |
|  | 3/4 | 0.936 | 0.034 | 0.029 | 0.911 |
|  | 4/5 | 0.953 | 0.026 | 0.021 | 0.932 |
|  | 5/6 | 0.972 | 0.017 | 0.011 | 0.960 |

Table 3.4.5.3Diii
Accuracy and Consistency of Classification Indices: Cphn (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.729 | 0.635 |  | 0.539 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.851 |  | 0.778 |  |
|  | 2 | 0.758 |  | 0.673 |  |
|  | 3 | 0.667 |  | 0.555 |  |
|  | 4 | 0.454 |  | 0.341 |  |
|  | 5 | 0.619 |  | 0.490 |  |
|  | 6 | 0.864 |  | 0.773 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.932 | 0.032 | 0.036 | 0.904 |
|  | 2/3 | 0.923 | 0.041 | 0.035 | 0.892 |
|  | 3/4 | 0.939 | 0.032 | 0.028 | 0.915 |
|  | 4/5 | 0.954 | 0.026 | 0.020 | 0.935 |
|  | 5/6 | 0.973 | 0.016 | 0.011 | 0.962 |

Table 3.4.5.3Div
Accuracy and Consistency of Classification Indices: Cphn (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.737 | 0.643 |  | 0.543 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.838 |  | 0.761 |  |
|  | 2 | 0.778 |  | 0.697 |  |
|  | 3 | 0.673 |  | 0.561 |  |
|  | 4 | 0.509 |  | 0.387 |  |
|  | 5 | 0.606 |  | 0.476 |  |
|  | 6 | 0.877 |  | 0.788 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False Negatives |  |
|  | 1/2 | 0.931 | 0.035 | 0.034 | 0.902 |
|  | 2/3 | 0.922 | 0.041 | 0.037 | 0.891 |
|  | 3/4 | 0.942 | 0.031 | 0.027 | 0.918 |
|  | 4/5 | 0.960 | 0.022 | 0.019 | 0.942 |
|  | 5/6 | 0.977 | 0.015 | 0.008 | 0.968 |

### 3.4.5.4 Overall Language Composite 9-12



Figure 3.4.5.3B
Proficiency Level: Cphn 9-12S401 Online


Table 3.4.5.4A
Scale Score Descriptive Statistics: Over 9-12 S401 Online

| Grade | No. of <br> Students | Min. | Max. | Mean | Std. Dev. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{9}$ | 41,580 | 229 | 469 | 351.06 | 39.70 |
| $\mathbf{1 0}$ | 26,964 | 244 | 474 | 354.75 | 36.06 |
| $\mathbf{1 1}$ | 18,004 | 248 | 477 | 359.02 | 35.73 |
| $\mathbf{1 2}$ | 13,070 | 257 | 472 | 362.71 | 33.53 |
| Total | 99,618 | 229 | 477 | 355.03 | 37.48 |

Table 3.4.5.4B
Proficiency Level Distribution: Over 9-12 S401 Online

| Level | Grade 9 |  | Grade 10 |  | Grade 11 |  | Grade 12 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Count | Percent | Count | Percent | Count | Percent | Count | Percent | Count | Percent |
| $\mathbf{1}$ | 8,065 | $19.40 \%$ | 4,877 | $18.09 \%$ | 3,341 | $18.56 \%$ | 2,272 | $17.38 \%$ | 18,555 | $18.63 \%$ |
| $\mathbf{2}$ | 8,236 | $19.81 \%$ | 6,205 | $23.01 \%$ | 4,562 | $25.34 \%$ | 3,809 | $29.14 \%$ | 22,812 | $22.90 \%$ |
| $\mathbf{3}$ | 16,441 | $39.54 \%$ | 11,527 | $42.75 \%$ | 7,454 | $41.40 \%$ | 5,442 | $41.64 \%$ | 40,864 | $41.02 \%$ |
| $\mathbf{4}$ | 7,463 | $17.95 \%$ | 3,861 | $14.32 \%$ | 2,269 | $12.60 \%$ | 1,358 | $10.39 \%$ | 14,951 | $15.01 \%$ |
| $\mathbf{5}$ | 1,284 | $3.09 \%$ | 470 | $1.74 \%$ | 362 | $2.01 \%$ | 185 | $1.42 \%$ | 2,301 | $2.31 \%$ |
| $\mathbf{6}$ | 91 | $0.22 \%$ | 24 | $0.09 \%$ | 16 | $0.09 \%$ | 4 | $0.03 \%$ | 135 | $0.14 \%$ |
| Total | 41,580 | $100.00 \%$ | 26,964 | $100.00 \%$ | 18,004 | $100.00 \%$ | 13,070 | $100.00 \%$ | 99,618 | $100.00 \%$ |

Table 3.4.5.4C
Reliability: Over 9-12 S401 Online

| Component | Weight | Variance | Reliability |
| :---: | :---: | :---: | :---: |
| Listening | 0.15 | 2992.442 | 0.870 |
| Reading | 0.35 | 1432.799 | 0.910 |
| Speaking | 0.15 | 2358.334 | 0.658 |
| Writing | 0.35 | 1677.917 | 0.867 |
| Overall Composite |  | 1405.064 | 0.950 |

*Variances from students who had results in all four domains

Table 3.4.5.4Di
Accuracy and Consistency of Classification Indices: Over (Grade 9) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.818 | 0.745 |  | 0.654 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.918 |  | 0.868 |  |
|  | 2 | 0.718 |  | 0.622 |  |
|  | 3 | 0.856 |  | 0.802 |  |
|  | 4 | 0.777 |  | 0.691 |  |
|  | 5 | 0.693 |  | 0.536 |  |
|  | 6 | - |  | 0.688 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False Positives | False Negatives |  |
|  | 1/2 | 0.958 | 0.015 | 0.027 | 0.942 |
|  | 2/3 | 0.938 | 0.035 | 0.027 | 0.913 |
|  | 3/4 | 0.942 | 0.029 | 0.028 | 0.918 |
|  | 4/5 | 0.981 | 0.013 | 0.006 | 0.973 |
|  | 5/6 | 0.998 | 0.002 | 0.000 | 0.998 |

Table 3.4.5.4Dii
Accuracy and Consistency of Classification Indices: Over (Grade 10) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.832 | 0.765 |  | 0.672 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.908 |  | 0.856 |  |
|  | 2 | 0.754 |  | 0.667 |  |
|  | 3 | 0.870 |  | 0.824 |  |
|  | 4 | 0.771 |  | 0.682 |  |
|  | 5 | 0.766 |  | 0.554 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.959 | 0.016 | 0.025 | 0.944 |
|  | 2/3 | 0.935 | 0.036 | 0.029 | 0.908 |
|  | 3/4 | 0.949 | 0.026 | 0.025 | 0.929 |
|  | 4/5 | 0.987 | 0.010 | 0.002 | 0.984 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

Table 3.4.5.4Diii
Accuracy and Consistency of Classification Indices: Over (Grade 11) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.832 | 0.763 |  | 0.670 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.897 |  | 0.839 |  |
|  | 2 | 0.767 |  | 0.683 |  |
|  | 3 | 0.866 |  | 0.818 |  |
|  | 4 | 0.770 |  | 0.674 |  |
|  | 5 | 0.794 |  | 0.635 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  | Consistency |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives |  |
|  | 1/2 | 0.956 | 0.018 | 0.026 | 0.939 |
|  | 2/3 | 0.932 | 0.036 | 0.032 | 0.904 |
|  | 3/4 | 0.955 | 0.024 | 0.021 | 0.936 |
|  | 4/5 | 0.988 | 0.009 | 0.003 | 0.985 |
|  | 5/6 | 0.999 | 0.001 | 0.000 | 1.000 |

Table 3.4.5.4Div
Accuracy and Consistency of Classification Indices: Over (Grade 12) S401 Online

| Overall Indices | Accuracy | Consistency |  | Kappa (k) |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.838 | 0.772 |  | 0.677 |  |
| Conditional on Level | Level | Accuracy |  | Consistency |  |
|  | 1 | 0.884 |  | 0.820 |  |
|  | 2 | 0.797 |  | 0.722 |  |
|  | 3 | 0.869 |  | 0.822 |  |
|  | 4 | 0.759 |  | 0.662 |  |
|  | 5 | 0.854 |  | 0.682 |  |
|  | 6 | - |  | 1.000 |  |
| Indices at Cut Points |  | Accuracy |  |  |  |
|  | Cut Point | Accuracy | False <br> Positives | False <br> Negatives | Consistency |
|  | 1/2 | 0.956 | 0.020 | 0.024 | 0.939 |
|  | 2/3 | 0.929 | 0.037 | 0.034 | 0.901 |
|  | 3/4 | 0.961 | 0.021 | 0.019 | 0.944 |
|  | 4/5 | 0.991 | 0.008 | 0.001 | 0.989 |
|  | 5/6 | 1.000 | 0.000 | 0.000 | 1.000 |

## References

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

American Educational Research Association, American Psychological Association, \& National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, D.C.: American Psychological Association.
Bachman, L. F., \& Palmer, A. S. (2010). Language assessment in practice. Oxford: Oxford University Press.

Brennan, R. (2004). Linking with equivalent group or single group design (LEGS) (version 2.0). [Computer software]. University of Iowa: Center for Advanced Studies in Measurement and Assessment (CASMA).

Center for Applied Linguistics (2016). ACCESS for ELLs ${ }^{\circledR}$ Series 400 Listening and Reading Scale Maintenance: Technical Brief. Washington, D.C.: Author.

Center for Applied Linguistics (2017). ACCESS for ELLs ${ }^{\circledR}$ 2.0 Speaking and Writing Score Scale Reconstruction: Technical Brief. Washington, D.C.: Author.

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

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

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 ELLss ${ }^{\circledR}$ Assessment. Madison, WI: Author.

Cook, H. G. and MacGregor, D. (2017). The ACCESS for ELLs 2.0 2016 Standard Setting Study (Technical Report). Madison, WI: Board of Regents of the University of Wisconsin System.

Data Recognition Corporation (2016a). DRC INSIGHT Online Learning System Technology User Guide. Maple Grove, MN: DRC/WIDA.

Data Recognition Corporation (2016b). WIDA Assessment Management System (WIDA AMS). Maple Grove, MN: DRC/WIDA.

Gottlieb, M. (2004). English Language Proficiency Standards for English Language Learners in Kindergarten through Grade 12: Framework for Large-Scale State and Classroom Assessment. Madison, WI: WIDA Consortium.

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

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

Kane, M. (2013). The argument-based approach to validation. School Psychology Review, 42(4), 448-457.

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

Kenyon, D. M. (2006). Development and Field Test of ACCESS for ELLs® (WIDA Consortium Technical Report No. 1). Center for Applied Linguistics: Washington, DC.

Kenyon, D. M., Ryu, J. R., \& MacGregor, D. (2013). Setting Grade Level Cut Scores for ACCESS for ELLs® (WIDA Consortium Technical Report No. 4). Center for Applied Linguistics: Washington, DC.

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. (1999, Autumn). Relating Cronbach and Rasch Reliabilities Rasch Measurement Transactions, 1999, 13:2 p. 696. Retrieved from http://www.rasch.org/rmt/rmt132i.htm

Linacre, J. M. (2002, Autumn). What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions, 16(2), 878. Retrieved from http://www.rasch.org/rmt/rmt162f.htm
Linacre, J. M. (2006). Winsteps Rasch analysis (Version 3.60.1) [Computer software]. Retrieved from http://www.winsteps.com

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

Llosa, L. (2008). Building and supporting a validity argument for a standards-based classroom assessment of English proficiency based on teacher judgments. Educational Measurement: Issues and Practice, 27(3), 32-42.

Mantel, N. \& Haenszel, W. (1959). Statistical aspect of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719-748.

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

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.

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

Thissen, D. (2000). Reliability and measurement precision. In H. Wainer, N. Dorans, D. Eignor, R. Flaugher, B. Green, R. Mislevy, L. Steinberg \& D. Thissen (Eds.), Computerized adaptive testing: A primer (Second Edition). Hillsdale, NJ: Lawrence Erlbaum Associates, 159-184.

Toulmin, S. E. (2003). The uses of argument (Updated ed.). Cambridge: Cambridge University Press.

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

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

WIDA Consortium. (2016). WIDA ACCESS for ELLs 2.0 2016-2017 Test Administrator Manual. Madison, WI: Author.

Wright, B. D., \& Stone, M. H. (1979). Best test design: Rasch measurement. Chicago, IL: MESA Press.

Young, M.J., \& Yoon, B. (1998, April). Estimating the consistency and accuracy of classifications in a standards-referenced assessment (CSE Tech. Rep. 475). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education and Information Studies.

Zwick, R., \& Bridgeman, B. (2014). Evaluating validity, fairness, and differential item functioning in multistage testing. In Y. Duanli, A. A. von Davier, \& C. Lewis (Eds.), Computer Multistage Testing: Theory and Applications (pp. 271-284). Hoboken, NJ: CRC Press.

Zwick, R., Donoghue, J. R., \& Grima, A. (1993). Assessment of differential item functioning for performance tasks. Journal of Educational Measurement, 30, 233-251.
Zwick R., Thayer D. T., \& Wingersky, M. (1993). A simulation study of methods for addressing differential item functioning in computer-adaptive tests. ETS Research Report, 93-11.

## Acknowledgements

We would like to extend our appreciation to the many CAL staff members who have supported this work, including:

Melissa Amos, M.S.
Keira Ballantyne, Ph.D.
Tanya Bitterman, M.A.
Michele Kawood, M.S.Ed.
Justin Kelly, Ph.D.
Dorry M. Kenyon, Ph.D.
Tristan Kirkman, M.A.
Daniel Lee, M.S.
Chen Li., M.S. Ed
Ashley Lipps, M.A., M.P.A.
Mohammed Louguit, Ph.D.
Meg Montee, Ph.D.
Jennifer Norton, Ed.D.
Lin Pan, M.S.
Mary Spanarkel, M.A.
Jing Wei, Ph.D.
Yu-Chia Wu, M.Ed.
Shu Jing Yen, Ph.D.
Xin Yu, M.A.
Jiaqi Zheng, M.S.


[^0]:    ${ }^{1}$ Note that although students' final scores for Series 401 were interpreted in terms of the ACCESS 2.0 proficiency level cuts set at the 2016 standard setting, the settings to route students through the adaptive test engine are keyed to ACCESS 1.0 proficiency level cuts.

[^1]:    ${ }^{2}$ Note, however, that on the Series 401 test there are a few exceptions where the PL3 task differs on Tiers A and $B / C$.

[^2]:    ${ }^{1}$ For WIDA webinars, see: https://www.wida.us/assessment/ACCESS\%202.0/WebinarRecordings.aspx. For webinars focusing on technological issues, see: https://www.wida.us/assessment/video/DuringTestingTechnologyTroubleshooting.aspx. For technology FAQs, see: https://www.wida.us/assessment/ACCESS\%202.0/technology.aspx\#1.
    ${ }^{2}$ The WIDA System Status Dashboard is located at: http://status.drcedirect.com/WIDA.

[^3]:    ${ }^{3}$ WIDA state pages can be found at: https://www.wida.us/membership/states/index.aspx
    ${ }^{4}$ ACCESS for ELLs 2.0 FAQs can be found at: https://www.wida.us/assessment/ACCESS\%202.0/administration.aspx\#8

[^4]:    ${ }^{5}$ In the dataset, Hispanic ethnicity, as well as each of the race categories, are coded as a binary variable (Y/N). Ethnicity information is counted as "Unknown" in cases where the student is recorded as N for Hispanic ethnicity and also N for every race category.

[^5]:    * No equating was performed for the previous series (S400).

[^6]:    * No equating was performed for the previous series (S400).

[^7]:    ${ }^{\wedge}$ Truncated

[^8]:    * No equating was performed for the previous series (S400).

[^9]:    ${ }^{\wedge}$ Truncated

[^10]:    * No equating was performed for the previous series (S400).

[^11]:    * No equating was performed for the previous series (S400).

[^12]:    * No equating was performed for the previous series (S400).

[^13]:    * No equating was performed for the previous series (S400).

[^14]:    * No equating was performed for the previous series (S400).

[^15]:    ${ }^{\wedge}$ Truncated

[^16]:    * No equating was performed for the previous series (S400).

[^17]:    $\wedge$ Truncated

    * Adjusted for end of scale effect

[^18]:    * No equating was performed for the previous series (S400).

[^19]:    * No equating was performed for the previous series (S400).

[^20]:    * No equating was performed for the previous series (S400).

[^21]:    * No equating was performed for the previous series (S400).

[^22]:    ${ }^{\wedge}$ Truncated

[^23]:    * No equating was performed for the previous series (S400).

[^24]:    * No equating was performed for the previous series (S400).

[^25]:    * No equating was performed for the previous series (S400).

[^26]:    * No equating was performed for the previous series (S400).

[^27]:    * No equating was performed for the previous series (S400).

[^28]:    ${ }^{\wedge}$ Truncated

[^29]:    * No equating was performed for the previous series (S400).

[^30]:    ** This task is shared between Pre-A and A.

    * This task is shared between A and B/C.

[^31]:    * No equating was performed for the previous series (S400).

[^32]:    * No equating was performed for the previous series (S400).

[^33]:    * No equating was performed for the previous series (S400).

[^34]:    * No equating was performed for the previous series (S400).

[^35]:    $\wedge$ Truncated

[^36]:    * No equating was performed for the previous series (S400).

[^37]:    ** This task is shared between Pre-A and A.

[^38]:    * No equating was performed for the previous series (S400).

[^39]:    * Variances from students who had results in all four domains

[^40]:    *Variances from students who had results in all four domains

[^41]:    *Variances from students who had results in all four domains

[^42]:    *Variances from students who had results in all four domains

