



California Department of Education Assessment Development and Administration Division



California Assessment of
Student Performance and Progress

California Spanish Assessment 2018–19 Field Test Technical Report

Contract #CN150012

Prepared for the California Department of Education by
Educational Testing Service

Final Submitted June 14, 2019

Table of Contents

Chapter 1: Introduction	1
1.1. Overview	1
1.2. Purpose of the Field Test	2
1.3. Participation of the Field Test	2
1.3.1. Recruitment of Local Educational Agencies (LEAs)	2
1.3.2. Intended Population	2
1.4. Testing Window and Times	2
1.5. Groups and Organizations Involved with the CAASPP System	3
1.5.1. State Board of Education (SBE)	3
1.5.2. California Department of Education (CDE)	3
1.5.3. California Educators	3
1.5.4. Contractors	4
1.6. Systems Overview and Functionality	4
1.6.1. Test Operations Management System (TOMS)	4
1.6.2. Test Delivery System (TDS)	5
1.6.3. Training Test	5
1.6.4. Constructed Response (CR) Scoring Systems	5
1.7. Overview of the Technical Report	6
References	7
Chapter 2: Item Development and Test Assembly	8
2.1. Overview	8
2.2. Test Blueprint	8
2.3. Item Development Plan	8
2.3.1. Selection of Item Writers	10
2.3.2. Item Writer Training	10
2.4. Item Types and Features	11
2.5. Item Review Process	12
2.5.1. ETS Content Review	13
2.5.2. ETS Editorial Review	13
2.5.3. ETS Sensitivity and Fairness Review	13
2.5.4. California Educator Review	14
2.6. Test Assembly and Length	15
Reference	17
Appendix 2.A: CSA Blueprint Overview—Operational Forms	18
Grade Span Three Through Five	19
Grade Span Six Through Eight	20
Grade Span Nine Through Twelve (High School)	21
Chapter 3: Test Administration	22
3.1. Field Test Administration	22
3.2. Test Security and Confidentiality	22
3.2.1. ETS' Office of Testing Integrity (OTI)	23
3.2.2. Procedures to Maintain Standardization of Test Security	23
3.2.3. Security of Electronic Files Using a Firewall	24
3.2.4. Transfer of Scores via Secure Data Exchange	24
3.2.5. Data Management in the Secure Database	25

3.2.6. Statistical Analysis on Secure Servers	25
3.2.7. Student Confidentiality	25
3.2.8. Security and Test Administration Incident Reporting System (STAIRS) Process	25
3.3. Accessibility Resources	26
3.3.1. Universal Tools	26
3.3.2. Designated Supports	27
3.3.3. Accommodations	27
3.3.4. Resources for Selection of Accessibility Resources	28
3.3.5. Delivery of Accessibility Resources	28
3.3.6. Unlisted Resources	28
3.4. Participation	28
3.5. Demographic Summaries	29
References	31
Appendix 3.A Demographic Summaries	32
Chapter 4: Summary Statistics for the 2018–19 Field Test Administration	44
4.1. Overview	44
4.2. Sample Used for the Analyses	44
4.3. Raw Score Distributions	45
4.4. Reliability and Standard Error of Measurement (SEM) Estimation	46
4.4.1. Raw Score Reliability Estimation	46
4.4.2. Standard Error of Measurement (SEM)	47
4.4.3. Results for the Field-Test Forms	47
4.5. Classical Item Analyses	48
4.5.1. Item Difficulty (Overall and by Item Type)	49
4.5.2. Item Discrimination (Overall and by Item Type)	49
4.5.3. Distractor Analyses (Overall and by Item Type)	50
4.5.4. Summary of Classical Item Analyses Flagging Criteria	50
4.5.5. Distribution of Item Scores	51
4.5.6. Classical Item Analyses Results Summaries	51
4.6. Response Time Analyses	53
4.7. Differential Item Functioning (DIF) Analyses	54
4.7.1. DIF Procedure for Dichotomous Items	54
4.7.2. DIF Procedure for Polytomous Items	55
4.7.3. Classification	56
4.7.4. Items Exhibiting Significant DIF	56
References	57
Appendix 4.A Test Scores Distribution	58
Appendix 4.B Classical Item Analyses	79
Appendix 4.C Response Time Analysis	139
Appendix 4.D Differential Item Functioning (DIF) Analyses	148
Chapter 5: Quality Control	150
5.1. Quality Control of Item Development	150
5.2. Quality Control of Test Form Development	150
5.3. Quality Control of Test Materials	151
5.3.1. Developing Test Administration Instructions	151
5.3.2. Processing Test Materials	151
5.4. Quality Control of Test Administration	151

5.5. Quality Control of Machine Scoring	152
5.6. Quality Control of Psychometric Processes	152
5.6.1. Development of Psychometric Specifications	152
5.6.2. Development of Psychometric Procedures	153
Reference	154
Chapter 6: Continuous Improvement	155
6.1. Feedback from the Data Review Meeting	155
6.2. Test Delivery and Administration	155
6.2.1. Survey Results	155
6.2.2. Training and Communication	156
6.3. Psychometric Analyses	156
6.4. Accessibility	157
Reference	158

List of Tables

<i>Acronyms and Initialisms Used in the CSA Field Test Technical Report</i>	vii
Table 2.1 Number of Field Test Items to Administer per Form	8
Table 2.2 Number of Items Developed per Grade Level for the CSA Field Test	9
Table 2.3 Item Types for the CSA Field Test	11
Table 2.4 ETS Field Test Forms Review Process	15
Table 2.5 Overview of Field Test Forms	16
Table 3.1 Participation Rates by Region	29
Table 3.2 Demographic Student Groups to Be Reported	29
Table 3.A.1 Demographic Summary—Grade Three	32
Table 3.A.2 Demographic Summary—Grade Four	34
Table 3.A.3 Demographic Summary—Grade Five	35
Table 3.A.4 Demographic Summary—Grade Six	36
Table 3.A.5 Demographic Summary—Grade Seven	38
Table 3.A.6 Demographic Summary—Grade Eight	40
Table 3.A.7 Demographic Summary—High School	42
Table 4.1 Summary of Completion of the Field Test by Grade Level	44
Table 4.2 Summary Statistics of the Raw Scores	45
Table 4.3 Test Reliability of the Total Scores	48
Table 4.4 Item Difficulty and Item-Total Correlation in Each Form by Grade Level	52
Table 4.5 Flagged Items Summary in Each Form by Grade Level	52
Table 4.6 DIF Categories for Dichotomous Items	56
Table 4.7 DIF Categories for Polytomous Items	56
Table 4.A.1 Raw Score Frequency Distribution—Grade Three	58
Table 4.A.2 Raw Score Frequency Distribution—Grade Four	61
Table 4.A.3 Raw Score Frequency Distribution—Grade Five	64
Table 4.A.4 Raw Score Frequency Distribution—Grade Six	67
Table 4.A.5 Raw Score Frequency Distribution—Grade Seven	70
Table 4.A.6 Raw Score Frequency Distribution—Grade Eight	73
Table 4.A.7 Raw Score Frequency Distribution—High School	76
Table 4.B.1 Dichotomous Item Statistics for Form 1—Grade Three	79
Table 4.B.2 Dichotomous Item Statistics for Form 2—Grade Three	81

Table 4.B.3 Dichotomous Item Statistics for the Accommodated Form—Grade Three	83
Table 4.B.4 Dichotomous Item Statistics for Form 1—Grade Four	85
Table 4.B.5 Dichotomous Item Statistics for Form 2—Grade Four	87
Table 4.B.6 Dichotomous Item Statistics for the Accommodated Form—Grade Four	89
Table 4.B.7 Dichotomous Item Statistics for Form 1—Grade Five	91
Table 4.B.8 Dichotomous Item Statistics for Form 2—Grade Five	93
Table 4.B.9 Dichotomous Item Statistics for the Accommodated Form—Grade Five	95
Table 4.B.10 Dichotomous Item Statistics for Form 1—Grade Six	97
Table 4.B.11 Dichotomous Item Statistics for Form 2—Grade Six	99
Table 4.B.12 Dichotomous Item Statistics for the Accommodated Form—Grade Six	101
Table 4.B.13 Dichotomous Item Statistics for Form 1—Grade Seven	103
Table 4.B.14 Dichotomous Item Statistics for Form 2—Grade Seven	105
Table 4.B.15 Dichotomous Item Statistics for the Accommodated Form—Grade Seven	107
Table 4.B.16 Dichotomous Item Statistics for Form 1—Grade Eight	109
Table 4.B.17 Dichotomous Item Statistics for Form 2—Grade Eight	111
Table 4.B.18 Dichotomous Item Statistics for the Accommodated Form—Grade Eight	113
Table 4.B.19 Dichotomous Item Statistics for Form 1—High School	115
Table 4.B.20 Dichotomous Item Statistics for Form 2—High School	117
Table 4.B.21 Dichotomous Item Statistics for the Accommodated Form—High School	119
Table 4.B.22 Polytomous Items Statistics—Grade Three	121
Table 4.B.23 Polytomous Items Statistics—Grade Four	122
Table 4.B.24 Polytomous Items Statistics—Grade Five	124
Table 4.B.25 Polytomous Items Statistics—Grade Six	126
Table 4.B.26 Polytomous Items Statistics—Grade Seven	128
Table 4.B.27 Polytomous Items Statistics—Grade Eight	130
Table 4.B.28 Polytomous Items Statistics—High School	131
Table 4.B.29 Summary of the Classical Item Statistics by Item Type—Grade Three	132
Table 4.B.30 Summary of the Classical Item Statistics by Item Type—Grade Four	132
Table 4.B.31 Summary of the Classical Item Statistics by Item Type—Grade Five	133
Table 4.B.32 Summary of the Classical Item Statistics by Item Type—Grade Six	133
Table 4.B.33 Summary of the Classical Item Statistics by Item Type—Grade Seven	133
Table 4.B.34 Summary of the Classical Item Statistics by Item Type—Grade Eight	134
Table 4.B.35 Summary of the Classical Item Statistics by Item Type—High School	134
Table 4.B.36 Summary of the Classical Item Statistics by Item and Response Types— Grade Three	134
Table 4.B.37 Summary of the Classical Item Statistics by Item and Response Types— Grade Four	135
Table 4.B.38 Summary of the Classical Item Statistics by Item and Response Types— Grade Five	136
Table 4.B.39 Summary of the Classical Item Statistics by Item and Response Types— Grade Six	136
Table 4.B.40 Summary of the Classical Item Statistics by Item Type and Response Type— Grade Seven	137
Table 4.B.41 Summary of the Classical Item Statistics by Item and Response Types— Grade Eight	137
Table 4.B.42 Summary of the Classical Item Statistics by Item and Response Types— High School	138

Table 4.C.1 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Three	139
Table 4.C.2 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Four	140
Table 4.C.3 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Five	141
Table 4.C.4 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Six	142
Table 4.C.5 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Seven	143
Table 4.C.6 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Eight	144
Table 4.C.7 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Nine	145
Table 4.C.8 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Ten	146
Table 4.C.9 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Eleven	147
Table 4.D.1 Gender DIF Classifications Summary by Grade Level	148
Table 4.D.2 Items Exhibiting Significant DIF by Student Group	149
Table 6.1 Summary of Flagged Items	155

Acronyms and Initialisms Used in the *CSA Field Test Technical Report*

Abbreviation	Term
AD	Assessment Development
AERA	American Educational Research Association
AIR	American Institutes for Research
AIS	average item score
APA	American Psychological Association
CAA	California Alternate Assessment
CAASPP	California Assessment of Student Performance and Progress
CaCCSSeE	California Common Core State Standards en Español
CALPADS	California Longitudinal Pupil Achievement Data System
CCR	<i>California Code of Regulations</i>
CDE	California Department of Education
CR	constructed response
CSA	California Spanish Assessment
DIF	differential item functioning
EL	English learner
ELA	English language arts/literacy
ETS	Educational Testing Service
FIA	final item analyses
IBIS	Item Banking Information System
IFEP	initial fluent English proficient
IEP	individualized education program
LEA	local educational agency
MC	multiple choice
MH DIF	Mantel-Haenszel Differential Item Functioning
NCME	National Council on Measurement in Education
OTI	Office of Testing Integrity
PAR	Psychometric Analysis and Research
PIA	preliminary item analyses
QA	quality assurance
RFEP	reclassified fluent English proficient
SBE	State Board of Education
SD	standard deviation
SEM	standard error of measurement
SFTP	secure file transfer protocol
SMD	standardized mean difference
SS	single select
STAIRS	Security and Test Administration Incident Reporting System
STS	Standards-based Tests in Spanish
TDS	test delivery system

CAASPP System

Abbreviation	Term
TE	technology enhanced
TEI	technology-enhanced items
TOMS	Test Operations Management System
UAT	user acceptance testing
USC	United States Code

Chapter 1: Introduction

1.1. Overview

In October 2013, Assembly Bill 484 established the California Assessment of Student Performance and Progress (CAASPP) as the new student assessment system that replaced the Standardized Testing and Reporting program. The primary purpose of the CAASPP System of assessments is to assist teachers, administrators, and students and their parents/guardians by promoting high-quality teaching and learning through the use of a variety of item types and assessment approaches. These tests provide the foundation for the state's school accountability system.

During the spring 2017–18 administration, the CAASPP System comprised the following assessments:

- Smarter Balanced assessments and tools:
 - Summative Assessments—Online assessments for English language arts/literacy (ELA) and mathematics in grades three through eight and grade eleven
 - Interim Assessments—Optional resources developed for grades three through eight and grade eleven designed to inform and promote teaching and learning by providing information that can be used to monitor student progress toward mastery of the Common Core State Standards and that may be administered to students at any grade level
 - Digital Library—Tools and practices designed to help teachers utilize formative assessment processes for improved teaching and learning in all grades
- California Alternate Assessments (CAAs) for ELA and mathematics in grades three through eight and grade eleven
- Science assessments in grades five, eight, and high school (grades ten, eleven, or twelve; these are the California Science Test and the CAA for Science)
- The final year of a primary language assessment, the Standards-based Tests in Spanish (STS) for Reading/Language Arts, in grades two through eleven (optional for eligible students)

As part of the CAASPP System of assessments, the California Spanish Assessment (CSA) is being developed as an optional assessment that will replace the STS. This new computer-based assessment for students in grades three through eight and high school is designed to measure a student's Spanish skills in reading, writing mechanics, and listening for the purposes of

- providing student-level data in Spanish competency,
- providing aggregate data that may be used for evaluating the implementation of Spanish language arts programs at the local level, and
- providing a high school measure suitable to be used, in part, for the California State Seal of Biliteracy.

Development of the CSA started in September 2016 with the State Board of Education's (SBE's) approval of the high-level test design. Following item development and reviews with

California educators, the CSA field test was administered to voluntarily participating LEAs during fall 2018.

1.2. Purpose of the Field Test

The CSA field test evaluated the new computer-based CSA reading, writing mechanics, and listening items and created a pool of high-quality, statistically viable items to be used for the operational administration. The focus of the CSA field test was on student aggregate performance on each item in the online environment (refer to subsection [2.4](#) for a list of item types). Student responses to each item were collected, aggregated, and reviewed to examine whether the item performed as expected. After the items were evaluated, feedback from the field test was used to update and inform item writing guidelines and future item development. Approved items were deployed for the operational item pool.

The CSA field test was intended to assess item performance and not student performance. Although data was collected at the item level, student scores were not reported for the CSA field test.

1.3. Participation in the Field Test

1.3.1. Recruitment of Local Educational Agencies (LEAs)

Because the CSA will be a voluntary assessment once operational, the goal of the field test recruitment was to have as many eligible students and LEAs as possible participate. All LEAs in California were invited to participate in the field test. Educational Testing Service (ETS) developed recruitment packets for LEAs to indicate their intent to participate. The LEAs were asked to identify the number and grade level of participating students as well as the number of students who have received or are receiving Spanish instruction.

ETS, with the input of the California Department of Education (CDE), developed participation guidelines that specifically outlined which students LEAs should encourage to take the CSA field test (i.e., students who have received at least one year of formal Spanish instruction).

LEA outreach efforts and communication began in the winter of 2017. Additionally, grade-band training tests were released in spring 2018 to assist LEA staff in preparing students and familiarizing themselves with the new assessment.

1.3.2. Intended Population

The CSA field test was administered to 8,049 students from the participating LEAs. The assessed population for the CSA field test was all students in grades four through twelve who receive instruction in Spanish in California and who seek a measure that recognizes their Spanish-specific reading, writing mechanics, and listening skills. Because the CSA field test was administered in fall 2018 as a summative assessment, participating students took the grade-level test that reflected the grade they were enrolled in for the 2017–18 school year. For example, students in grade four in fall 2018 took the grade three assessment; students in grade nine in fall 2018 took the grade eight assessment.

1.4. Testing Window and Times

The CSA field test was delivered online using the fixed testing forms during a testing window from September 17 to October 19, 2018. Three field test forms were used per grade level, including two general forms and one form with accessibility features. Similar to other

CAASPP assessments, the CSA field test was untimed for test takers. A student could take the CSA field test within the testing window over as many days as required to meet a student's needs (*California Code of Regulations*, Title 5, Education, Division 1, Chapter 2, Subchapter 3.75, Article 2, Section 855[a]).

1.5. Groups and Organizations Involved with the CAASPP System

1.5.1. State Board of Education (SBE)

The SBE is the state agency that establishes educational policy for kindergarten through grade twelve in the areas of standards, instructional materials, assessment, and accountability. The SBE adopts textbooks for kindergarten through grade eight, adopts regulations to implement legislation, and has the authority to grant waivers of the *Education Code*.

In addition to adopting the rules and regulations for itself, its appointees, and California's public schools, the SBE is also the state educational agency responsible for overseeing California's compliance with programs that meet the requirements of the federal Every Student Succeeds Act and the state's Public School Accountability Act, which measure the academic performance and progress of schools on a variety of academic metrics (CDE, 2017).

1.5.2. California Department of Education (CDE)

The CDE oversees California's public school system, which is responsible for the education of more than 6,200,000 children and young adults in more than 10,450¹ schools. California aims to provide a world-class education for all students, from early childhood to adulthood. The CDE serves the state by innovating and collaborating with educators, school staff, parents/guardians, and community partners which together, as a team, prepares students to live, work, and thrive in a highly connected world.

Within the CDE, it is the Performance, Planning, & Technology Branch that oversees programs promoting innovation and improved student achievement. Programs include oversight of statewide assessments and the collection and reporting of educational data (CDE, 2018b). Within the Performance, Planning, & Technology Branch, the Assessment Development & Administration Division manages the development and administration for all statewide assessments.

1.5.3. California Educators

A variety of California educators, including teachers and school administrators, who were selected based on their qualifications, experiences, demographics and geographic locations, were invited to participate in the entire process of the assessments, including defining the purpose and scopes, test design, item development, and scoring the constructed-response items.

¹ Retrieved from the CDE Fingertip Facts on Education in California – *CalEdFacts* web page at <https://www.cde.ca.gov/ds/sd/cb/ceffingertipfacts.asp>

1.5.4. Contractors

1.5.4.1 Educational Testing Service

The CDE and the SBE contract with ETS to develop and administer the CSA field test. As the prime contractor, ETS has the overall responsibility for working with the CDE to implement and maintain an effective assessment system and to coordinate the work of its subcontractors. Activities directly conducted by ETS include, but are not limited to, the following:

- Providing management of the program activities
- Supporting and training counties, LEAs, and direct funded charter schools
- Providing tiered help desk support to LEAs
- Hosting and maintaining a website with resources for LEA CAASPP coordinators
- Developing, hosting, and providing support for TOMS
- Developing all CSA test items
- Constructing, producing, and controlling the quality of CSA test forms and related test materials, including *Directions for Administration*
- Processing student test assignments
- Completing all psychometric procedures

1.5.4.2 American Institutes for Research (AIR)

ETS also monitors and manages the work of AIR, subcontractor to ETS for the CAASPP System of online assessments. Activities conducted by AIR include

- providing the AIR proprietary TDS, including the Student Testing Interface, Test Administrator Interface, secure browser, and practice and training tests;
- hosting and providing support for its TDS and Online Reporting System (ORS), a component of the overall CAASPP Assessment Delivery System;
- scoring machine-scorable items; and
- providing Level 3 technology help desk support to LEAs for technology issues directly related to the TDS.

1.6. Systems Overview and Functionality

1.6.1. Test Operations Management System (TOMS)

TOMS is the password-protected, web-based system that LEAs use to manage all aspects of CAASPP testing. TOMS serves various functions, which, for the CSA field test, included but were not limited to the following:

- Managing test administration windows
- Assigning and managing CSA online user roles
- Managing student test assignments and accessibility resources

- Providing a platform for authorized user access to secure materials such as user information and access to the Security and Test Administration Incident Reporting System form and the Appeals module

TOMS receives student enrollment data and LEA and school hierarchy data from CALPADS via a daily feed. CALPADS is “a longitudinal data system used to maintain individual-level data including student demographics, course data, discipline, assessments, staff assignments, and other data for state and federal reporting.”² LEA staff involved in the administration of the CSA, such as LEA coordinators, test site coordinators, test administrators, and test examiners are assigned varying levels of access to TOMS. For example, only an LEA coordinator has permission to set up the LEA’s test administration window; a test administrator cannot download student reports. A description of user roles is explained more extensively in the *2017–18 Online Test Administration Manual* (CDE, 2018a).

1.6.2. Test Delivery System (TDS)

The TDS is the means by which the statewide online assessments are delivered to students. Components of TDS include

- The Test Administrator Interface, the web browser–based application that allows test administrators to activate student tests and monitor student testing;
- The Student Testing Interface, on which students take the test using the secure browser; and
- The secure browser, the online application through which the Student Testing Interface may be accessed. The secure browser prevents students from accessing other applications during testing.

1.6.3. Training Test

The publicly available training tests, offered by grade band (grades three through five, grades six through eight, and high school), is provided to prepare students for the summative assessment. These tests simulate the experience of the CSA online assessments. Students may access them using a web browser.

The purposes of the training test are to

- allow students and administrators to quickly become familiar with the user interface and components of TDS and the process of starting and completing a testing session, and
- introduce students and administrators to new item types similar to those on the field test.

1.6.4. Constructed Response (CR) Scoring Systems

CR items from the TDS are routed to ETS’s CR scoring systems. CR items are scored by certified human raters. The hired human raters are provided in-depth training and certified before starting the scoring process. Human raters are organized under a scoring leader and are provided CSA scoring materials such as anchor sets, scoring rubrics, validity samples, qualifying sets, and condition codes for unscorable responses within the interface. Because

² From the CDE California Longitudinal Pupil Achievement Data System (CALPADS) web page at <http://www.cde.ca.gov/ds/sp/cl/>.

the scoring of CR items will take place in May through July of 2019, the evaluation results of CR scoring will be included in a separate CR analysis report.

1.7. Overview of the Technical Report

This technical report addresses the characteristics of the CSA field test administered in fall 2018. It contains seven additional chapters as follows:

- [Chapter 2](#) discusses the detailed procedures of item development, item review, and field test assembly for the fall 2018 field test administration.
- [Chapter 3](#) describes the details of administering the CSA field test forms, LEA participation, and demographic summaries.
- [Chapter 4](#) summarizes the results of the psychometric analyses for the fall 2018 CSA field test, including classical item analyses, response time analyses, test completion analyses, and differential item functioning (DIF) analyses.
- [Chapter 5](#) highlights the quality control processes used at various stages of administration such as item and test form development, scoring, and psychometric processes.
- [Chapter 6](#) summarizes the findings from the fall 2018 CSA field test administration and discusses the continuous improvement steps and implications for the first operational assessment, to be administered in spring 2019.

References

California Code of Regulations, Title 5, Education, Division 1, Chapter 2, Subchapter 3.75, Article 2, § 855. Retrieved from <https://bit.ly/326oEEs>

California Department of Education. (2017, October). *State Board of Education responsibilities.* Retrieved from <http://www.cde.ca.gov/be/ms/po/sberesponsibilities.asp>

California Department of Education. (2018a). *CSA online field test administration manual.* Sacramento, CA: California Department of Education. Retrieved from <http://www.caaspp.org/rsc/pdfs/CSA.ft-otam.2018-19.pdf>

California Department of Education. (2018b, August). *Organization.* Retrieved from <http://www.cde.ca.gov/re/di/or/>

Chapter 2: Item Development and Test Assembly

This chapter discusses the detailed procedures of item development and field test assembly for the California Spanish Assessment (CSA) field test administration. In particular, new item types and features that differ from traditional item types are described.

2.1. Overview

Educational Testing Service (ETS) developed 757 field test items across the seven grade levels (i.e., grades three through eight and high school) and delivered them to the California Department of Education (CDE) via the ETS Item Banking Information System (IBIS). The items developed were designed to be engaging to the student population and represented a wide variety of item types. All items for the CSA field tests were developed in accordance with the *ETS Standards for Quality and Fairness* (2014) across all phases of item and test development. While under initial development, the assessment materials, including items, passages, constructed-response (CR) prompts, and listening stimuli, were kept on password-protected ETS computers and secure internal network drives. Audio recordings were produced as electronic audio files and delivered to the CDE for review. All secure documents needed for CDE review that were not available in IBIS were delivered to the CDE via the Tumbleweed secure file transfer protocol server.

2.2. Test Blueprint

Each field test form contains items that approximate the proportions in the test blueprint. The proposed test blueprint for the CSA provides the proposed numbers of items to be included in an operational assessment for each language-arts domain assessed in grades three through eight and high school. [Table 2.1](#) shows the distribution of the field test items by domain and grade level. [Appendix 2.A](#) presents the overview of the CSA blueprint by grade span.

Table 2.1 Number of Field Test Items to Administer per Form

Domain	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	High School
Listening	8–10	8–10	8–10	8–10	8–10	8–10	8–10
Reading	27–31	27–31	27–31	27–31	27–31	27–31	27–31
Writing Mechanics	20–24	20–24	20–24	20–24	20–24	20–24	20–24
Total Number of Items	60 items	60 items	60 items	60 items	60 items	60 items	60 items

2.3. Item Development Plan

The initial item development plan for the CSA field test was to develop sufficient items across a variety of item types to eventually build an operational form to blueprint. The items developed for the CSA field test closely reflected the distribution of domains in the blueprint. [Table 2.2](#) shows the number of items developed in each of the domains of reading, writing mechanics, and listening for the CSA field test.

The total number of machine-scorable items developed and field-tested (757) was greater than the number to be administered operationally (364) because overage was built in. ETS develops overage to account for the potential rejection of items during item review and data

review meetings. If item reviewers at the item review meeting determine that certain items are not appropriate for field testing, the overage ensures that the minimum item counts for the field test forms will be satisfied. Similarly, if item reviewers at the data review meeting determine that certain items are not performing well enough for operational use, the overage ensures that the blueprint for the operational test forms will still be satisfied. For the general forms in the field test, there was substantial overage built in. However, for the accommodated forms in the field test, there was little overage when accommodations were applied to some of the items. In the future, there is a need to develop more items for the accommodated forms.

Table 2.2 Number of Items Developed per Grade Level for the CSA Field Test

Domain	Number
Listening	16–17
Reading	37–43
Writing Mechanics	46–54
Total Number of Items	106–110

All items created for the CSA adhere to the *ETS Standards for Quality and Fairness* (2014) across all phases of item and test development. Each CSA item was developed through a comprehensive development cycle and designed to conform to the principles of quality item writing as defined by ETS.

Throughout the item-writing process, ETS adhered to its foundational guidelines for quality item writing. According to these guidelines, item developers conformed to the following list of attributes for each item:

1. The question is clearly and concisely presented.
2. There is an absence of clueing in the item stem and supporting stimuli.
3. The supporting stimulus and stimuli are presented clearly and are construct-relevant.
4. There is a single correct answer (for selected-response items only).
5. Distractors are plausible, but incorrect (for selected-response only).
6. The answer key is correct.
7. The scoring rubric and annotations are accurate, precise, and complete.
8. Item format and content adhere to the principles of universal design.

ETS maintains item specifications for the CSA. These specifications describe the characteristics of items written to measure the Common Core State Standards en Español that, in turn, provide evidence for the CSA’s reading, writing mechanics, and listening domains. Using the item specifications helps ensure that all items developed for the CSA measure standards consistently. Item writing assignments are guided by the CSA blueprints, developed in consultation with the CDE.

The specifications include

- a description of best practices for item writing:
 - universal design,
 - bias and sensitivity avoidance,
 - cognitive level,
 - anatomy of an item,
 - item types and characteristics,

- a general list of elements to avoid, and
- stand-alone items;
- information about passages used to assess CSA domains;
- a description of standards used for items associated with reading passages, writing mechanics passages, and listening passages;
- a full statement of each standard featured on the CSA blueprint; and
- sample item stems at each grade level for some standards.

2.3.1. Selection of Item Writers

Senior ETS content staff screened applications for item writers for the CSA field test, and ETS approved only those with strong content and teaching backgrounds for the item writing training program. ETS selected item writers after the training, but not all recipients of the training became an item writer.

Because some of the participants were current or former California educators, they were particularly knowledgeable about the standards assessed by the CSA. All item writers met the following minimum qualifications:

- Possession of a bachelor's degree in a relevant field of education; an advanced degree in the relevant content was desirable
- Previous experience or training in writing items for standards-based assessments, including knowledge of the many considerations that are important when developing items for special student populations

2.3.2. Item Writer Training

ETS assessment specialists provided item-writer training to California educators and ETS contractors. The in-person meeting trained California educators on how to write items for the computer-based CSA. ETS led educators through the Common Core State Standards en Español, detailed how to write a strong item, and described the functionality of the internet-delivered item types used on this new assessment.

ETS held item-writer training workshops in September 2016 in Sacramento, California, and in January 2017 in Santa Ana, California, to provide prospective item writers with professional development in several areas. A review of the general assessment development process gave trainees a sense of the total lifecycle of an item.

Participants learned best practices in item writing to provide clarity within the item and avoid bias or sensitivity concerns; how to review a passage for item opportunities; and an introduction to how the new, innovative item types work.

Given that the trainees were California educators and educational leaders, ETS also emphasized incorporation of current effective teaching practices and instructional activities. Small-group and individual work generated sample items that the ETS facilitators then used in a large-group discussion to analyze and ascertain overall item quality. The ETS team also provided post hoc feedback via email and phone calls to train item writers on further item samples and ideas submitted ahead of contractual item submissions.

The primary goals for the training were:

1. to provide teachers with knowledge, via professional development on writing items, that they can use to help develop or refine their own classroom teaching and assessments;
2. to ensure that teachers who successfully completed the training were ready to develop high-quality items for the CSA field test; and
3. to leverage the experiences, perspectives, and expertise of the teachers in writing items for the CSA field test.

2.4. Item Types and Features

ETS developed a variety of technology-enhanced (TE) item types that required the student to respond to a question in different ways from typical selected-response items. In addition to TE items, CR items were also used in the assessment. Items may contain a stimulus (e.g., a passage, audio, or image).

Students responded to TE items by typing an answer, completing a graph, dragging a response to a designated area, using a drop-down list selection, or selecting multiple areas in a graphic (also known as “hot spots”). All TE item types were designed to be machine-scorable.

The following item types were included in the 2018–19 field test:

- Multiple choice (MC) (single select and multiple select)
- Zone (single select and multiple select)
- Inline choice list (single select and multiple select)
- Text choices (single select and multiple select)
- Numeric
- Grid (multiple select)
- Match (single select and multiple select)
- Composite
- Extended text

[Table 2.3](#) lists item types used in the CSA field test. Response types marked with an asterisk (*) are technology-enhanced items.

Table 2.3 Item Types for the CSA Field Test

Item Type	Response Type	Description
MC	Multiple choice single select	The item generally consists of a stem and list of choices; test taker can select only one choice to respond. May also include a stimulus.
MC	Multiple choice multiple select	The item generally consists of a stem and list of choices; test taker can select two or more choices to respond. May also include a stimulus.
Hot Spot	Zones single select *	An item where the answer choices are predefined “hotspots” on an image. When the test taker selects (clicks) on the spot, the selection is highlighted, shaded, or outlined in red. The test taker selects one zone to respond.

Item Type	Response Type	Description
Hot Spot	Zone multiple select *	An item where the answer choices are predefined “hotspots” on an image. When the test taker selects (clicks) on the spot, the selection is highlighted, shaded, or outlined in red. The test taker selects two or more zones to respond.
MC	Inline choice list single select *	The stem contains a single blank, and the test taker must fill the blank by selecting a choice from its corresponding choice list.
MC	Inline choice list multiple select *	The stem contains two or more blanks, and the test taker must fill each blank by selecting a choice from the corresponding choice lists.
MC	Text choices single select *	The test taker responds by selecting only one of several underlined words or phrases embedded in a larger section of text.
MC	Text choices multiple select *	The test taker responds by selecting two or more underlined words or phrases embedded in a larger section of text.
Numeric CR	Numeric	The test taker responds by filling in a blank entry box with a numeric value.
MC	Grid multiple select *	The test taker responds by marking two or more cells in a table grid.
Drag & Drop	Match single select *	The test taker responds by dragging and dropping a single choice (“source”) into the appropriate location (“target”).
Drag & Drop	Match multiple select *	The test taker responds by dragging and dropping one or more choices (“sources”) into the appropriate locations (“targets”).
All item types except CR	Composite *	The test taker completes multiple tasks based on a combination of machine-scored items.
Short CR and Extended CR	Extended Text *	An item that consists of a stem in which the test taker must provide a written response, usually in the form of a paragraph or an essay.

2.5. Item Review Process

After items were drafted, they went through ETS reviews, a review by CDE staff, a Content Review, as well as a Bias and Sensitivity Review.

All items were entered into IBIS with corresponding artwork and metadata. Within IBIS, items received ETS internal content, fairness, and edit reviews.

The CDE reviewed proposed changes to items in response to reviews by the participants of the Item and Passage Review meetings to ensure the quality of the item pool. The CDE then gained access to CSA field test items and conducted reviews in IBIS. ETS revised items in response to comments from the CDE prior to using them in the field test forms.

2.5.1. ETS Content Review

On all items ETS develops, content-area assessment specialists conduct three reviews on items and stimuli. These assessment specialists verify that the items and stimuli are in compliance with ETS's written guidelines for clarity, style, accuracy, and appropriateness for California students as well as in compliance with the approved item specifications.

Assessment specialists review each item in terms of the following characteristics:

- Relevance of each item to the purpose of the test
- Match of each item to the item specifications, including the tier of item complexity
- Match of each item to the principles of quality item writing
- Match of each item to the identified standard or standards
- Difficulty of the item
- Accuracy of the content of the item
- Readability of the item or passage
- Grade-level appropriateness of the item
- Appropriateness of any illustrations, graphs, or figures

Assessment specialists check each item against its classification codes, both to evaluate the correctness of the classification and to confirm that the task posed by the item is relevant to the outcome it is intended to measure. The reviewers can accept the item and classification as written, suggest revisions, or recommend that the item be discarded. These steps occur prior to the CDE's review.

2.5.2. ETS Editorial Review

After content-area assessment specialists and researchers review each item, a group of specially trained editors also review each item in preparation for consideration by the CDE and participants at the item review meeting. The editors check items for clarity, correctness of language, appropriateness of language for the grade level assessed, adherence to the style guidelines, and conformity with accepted item-writing practices.

2.5.3. ETS Sensitivity and Fairness Review

ETS assessment specialists who are specially trained to identify and eliminate questions that contain content or wording that could be construed to be offensive to or biased against members of specific ethnic, racial, or gender groups conduct the next level of review (ETS, 2014, 2016). These trained staff members review every item before the CDE and item-review meeting reviews.

The review process promotes a general awareness of and responsiveness to the following:

- Cultural diversity
- Diversity of background, cultural tradition, and viewpoints to be found in the test-taking populations
- Changing roles and attitudes toward various groups
- Role of language in setting and changing attitudes toward various groups

- Contributions of diverse groups (including ethnic and minority groups, individuals with disabilities, and women) to the history and culture of the United States and the achievements of individuals within these groups
- Item accessibility for English learners

2.5.4. California Educator Review

In preparation for the fall 2018 CSA field test administration, ETS convened a four-day meeting with California educators in Sacramento, California, to

- review Spanish passages and items for the fall 2018 field test for grade-level appropriateness, content, bias and sensitivity, readability, and overall interest for the test taker; and
- obtain feedback from California educators about the passages and items in order to inform ETS on the appropriateness of their use on future test forms for the CSA.

2.5.4.1 Item Review

Upon the completion of an introductory presentation, an ETS assessment specialist led the participants through a thorough training for reviewing items. This training included the structure of an item, the best practices for item reviewing, an explanation of item types and functionality, and a discussion of the metadata accompanying items. These metadata—alignment with the California Common Core State Standards en Español, depth of knowledge levels, difficulty levels, etc.—were available for each item on a comment sheet.

The group discussed each item together, reviewing for grade-level appropriateness, content, bias and sensitivity, depth of knowledge, standard alignment, and the correct answer or answers (as indicated in the metadata). ETS summarized comments, captured any recommended edits, and reached consensus from the group before moving forward to the next item. The group continued in this manner until all items were reviewed. The CDE made decisions separately from the group, as needed, and gave the final approval after requested edits had been applied. Items were then placed on the field test forms.

The educators reviewed grade six items as a group and then, upon completion of the grade six review, were divided into two groups to continue the review process: One group focused on grades three through five, and the other on grade seven, grade eight, and high school.

Following the training, ETS specialists facilitated the review of items by projecting the items on-screen with printed copies of passages associated with the items. The participants were asked to read a passage. When all participants finished, the facilitators projected each item associated with that passage one at a time. The facilitators read each item aloud and displayed any technology-enabled functions.

2.5.4.2 Passage Review

Participants were similarly trained to review passages. An ETS assessment specialist led the participants through a training that highlighted what to look for in a strong passage and present more detailed information on content and bias and sensitivity issues. Each participant received a grade-level comment sheet, a bias and sensitivity reference document, and a binder containing the passages for review.

Educators began by reviewing grade six passages. Grade six was chosen as a starting point to train participants because it is a grade in the middle of the range of grades, and it

requires neither the extra training in foundational reading for grades three through five, nor the secondary consideration of the State Seal of Biliteracy.

Once complete, the ETS specialists brought the full group together to discuss each grade six passage for grade-level appropriateness, content, bias and sensitivity, readability, and overall interest for the test taker. The CDE made decisions separate from the group, as needed, and gave the final approval after requested edits had been applied.

Upon completion of the grade six review, ETS divided the participants into two groups: One group focused on grades three through five and the other on grade seven, grade eight, and high school.

2.6. Test Assembly and Length

Following the item review meeting, the CDE conducted one round of item review. ETS assessment specialists worked closely with the CDE to select items and assemble field-test forms once the item review was complete.

The field-test forms were assembled so that they covered a variety of item types, item difficulties, cognitive levels, and key distributions. The forms were evaluated prior to CDE review—via the ETS review process shown in [Table 2.4](#)—and reviewed and approved by the CDE.

Table 2.4 ETS Field Test Forms Review Process

Step	Task
1. Test Assembly	Assessment specialists select test items that meet the specifications, are fair, and reflect appropriate content coverage. These items are collected in the item bank so they can be tracked as a unit.
2. Senior Review	An assessment specialist with content-area expertise, who did not assemble the test, reviews all of the items and checks for content-related issues (e.g., incorrect keys, overlapping content, cueing of one item by another) and other concerns (e.g., confirming that the items match the test framework). The assessment specialist also verifies that the test meets content and statistical specifications.
3. Senior Fresh-Eyes Review	Every new test form goes through a “fresh-eyes” review. During this review, a senior-level content expert, who has never seen the form, reviews it carefully for any content errors that may have been missed during earlier stages of review.
4. Certification	Once these reviews are completed and the test form is judged to be free from errors, ETS certifies the test form and sends it to be packaged for device delivery.

ETS developed three field-test forms per grade. Each grade level had two general forms with 60 items per form. Each grade level had one form with accessibility features. It included 60 items that were identical to or close variants of selected items on the two general field-test forms; this form was assigned to students with an individualized education program or Section 504 plan. The other two general forms were assigned to students randomly.

[Table 2.5](#) shows the total number of items needed per grade level to accommodate the field test. The estimated duration for the field test was 215 minutes in grades three through eight to 250 minutes in high school.

Table 2.5 Overview of Field Test Forms

Criteria	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	High School
Number of Forms	3	3	3	3	3	3	3
Number of Listening Passages per Form	2	2	2	2	2	2	2
Number of Reading Passages per Form	3	3	3	3	3	3	3
Number of Writing Mechanics Passages per Form	2	2	2	2	2	2	2
Listening Passage- based Items	16	17	16	17	17	16	16
Reading Passage- based Items	52	51	51	43	50	52	50
Writing Mechanics Passage-based Items	32	33	32	32	32	29	30
Number of Stand- alone Items	8	6	8	12	8	8	9
Number of Writing Mechanics Constructed Response (CR) Items	2	2	2	2	2	2	2
Estimated Testing Time	215 minutes	215 minutes	215 minutes	215 minutes	215 minutes	215 minutes	250 minutes

Reference

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>

Appendix 2.A: CSA Blueprint Overview—Operational Forms

The proposed test blueprint for the California Spanish Assessment (CSA) provides the proposed numbers of items and points to be included in an operational assessment for each Spanish reading/language arts domain assessed in grades three through eight and high school. Note, however, that the numbers of items and points are subject to revision in response to a qualitative evaluation of the items after the first pilot assessment and in response to statistical analyses of the first field test and first operational use. Note also that the blueprint does not include constructed-response writing items.

All items are aligned with the translated and linguistically augmented version of the Common Core English Language Arts (ELA) & Literacy standards (i.e., “California Common Core State Standards en Español” [CaCCSSeE]).

Each grade has more than 50 testable standards at its disposal. There are three overview tables provided—grades three through five, six through eight, and high school (grades nine through twelve)—to clarify the overall proportions of the blueprint. After the overview tables, specifics for each tested grade level are given, enumerating further subdivisions of the content and specific groups of testing standards. Note that high school grades are tested together in one level using the CaCCSSeE designated as “9–10” and “11–12” and uses the designation “high school” (HS).

The proposed blueprint is represented in tables. Each overview table is organized by the three domains assessed: reading, writing (mechanics), and listening—referred to as claim/score reporting category—and are provided in the first column. Other columns in the proposed blueprint are as follows:

- *Second column:* Content Category
- *Third column:* Number of items representing the content category on an operational assessment
- *Fourth column:* Number of points for the given content category
- *Remaining columns:* Aggregated item counts, points, and percentages by claim

Item counts and point values may be adjusted further during future stages of the CSA design and development effort to take into consideration the evaluation of pilot test results as well as the analyses of statistics of both the first field test and the first operational administration of the CSA.

Finally, for all tables, some items are anticipated to be polytomously scored (maximum of two points), so the number of items is smaller than the number of score points.

Grade Span Three Through Five

Claim/Score Reporting Category	Content Category	Total Items by Content Category	Total Score Points by Content Category*	Total Items by Claim	Percentage of Items by Claim	Total Score Points by Claim*	Percentage of Score Points by Claim
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Literary	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Informational	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Vocabulary and Meaning	8–10	10–13	24	46%	27–35	40–58%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Foundational Mechanics and Conventions	8–10	10–13	16	31%	19–22	28–37%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Revising and Editing	5–7	6–9	16	31%	19–22	28–37%
Listening Claim: Students can comprehend spoken Spanish in a range of contexts.	Listening Comprehension	12	15–17	12	23%	15–17	22–28%
NA	NA	NA	TOTALS:	52	100%	61–66	100%

Grade Span Six Through Eight

Claim/Score Reporting Category	Content Category	Total Items by Content Category	Total Score Points by Content Category*	Total Items by Claim	Percentage of Items by Claim	Total Score Points by Claim*	Percentage of Score Points by Claim
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Literary	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Informational	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Vocabulary and Meaning	8–10	10–13	24	46%	27–35	40–58%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Mechanics and Conventions	7–9	8–11	16	31%	19–22	28–37%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Revising and Editing	7–9	8–11	16	31%	19–22	28–37%
Listening Claim: Students can comprehend spoken Spanish in a range of contexts.	Listening Comprehension	12	15–17	12	23%	15–17	22–28%
NA	NA	NA	TOTALS:	52	100%	61–66	100%

Grade Span Nine Through Twelve (High School)

Claim/Score Reporting Category	Content Category	Total Items by Content Category	Total Score Points by Content Category*	Total Items by Claim	Percentage of Items by Claim	Total Score Points by Claim*	Percentage of Score Points by Claim
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Literary	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Informational	6–9	7–11	24	46%	27–35	40–58%
Reading Claim: Students can read, analyze, and interpret a variety of texts and genres through Spanish.	Vocabulary and Meaning	8–10	10–13	24	46%	27–35	40–58%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Mechanics and Conventions	7–9	8–11	16	31%	19–22	28–37%
Writing Claim: Students can revise writing products that accurately and convincingly present, describe, and explain ideas for a range of purposes and audiences through Spanish.	Revising and Editing	7–9	8–11	16	31%	19–22	28–37%
Listening Claim: Students can comprehend spoken Spanish in a range of contexts.	Listening Comprehension	12	15–17	12	23%	15–17	22–28%
NA	NA	NA	TOTALS:	52	100%	61–66	100%

Chapter 3: Test Administration

This chapter describes the details of administering the California Spanish Assessment (CSA) field-test forms, as well as local educational agency (LEA) participation and demographic summaries. It describes the efforts and measures to ensure test security, and summarizes demographics and procedures for implementation of test accommodations based on the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014, Chapter 6).

3.1. Field Test Administration

The CSA field test was administered to all eligible students in grades four through twelve in fall 2018. Because the CSA is a summative assessment, participating students took the grade-level test that reflected the grade they were enrolled in for the 2017–18 school year. For example, students enrolled in grade four in fall 2018 took the grade three assessment and students enrolled in grade nine in fall 2018 took the grade eight assessment. Students in the current grade three who were in grade two during the previous spring were not tested.

In accordance with the procedures for all online California Assessment of Student Performance and Progress (CAASPP) assessments, LEAs identified test administrators to administer the CSA field test and entered them into the Test Operations Management System (TOMS). Educational Testing Service (ETS) provided LEA staff with the appropriate training materials, such as test administration manuals, videos, and webcasts, to ensure that the LEA staff and test administrators understood how to administer the computer-based CSA field test.

The window for fall 2018 administration of the CSA field test was September 17 through October 19, 2018. Once the field test administration window opened, each participating LEA and school locally determined administration dates. Students reported to the testing classroom or center and were provided a computer or testing device on which to take the assessment. The field test utilized the same secure browser and online testing platform as all of the CAASPP assessments. The students received initial direction in Spanish from the test administrator as well as item-level directions, as needed. At the beginning of each field test, there were three additional questions, administered to collect information on whether the student received instruction in Spanish, the Spanish-language program type, and the percentage of instruction in Spanish.

3.2. Test Security and Confidentiality

For the CSA field test, every person who worked with the assessments or received testing information was responsible for maintaining the security and confidentiality of the tests, including California Department of Education (CDE) staff, Educational Testing Service (ETS) staff, ETS subcontractors, LEA assessment coordinators, school assessment coordinators, students, parents/guardians, teachers, and cooperative educational service agency staff. ETS' Code of Ethics requires that all test information, including tangible materials (e.g., test items), confidential files (e.g., those containing personally identifiable student information), and processes related to test administration (e.g., the configurations of secure servers) be kept secure. ETS has systems in place that maintained tight security for test items and test results, as well as for student data. To ensure security for all tests that

ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI), which is described in the next subsection.

All tests within the CAASPP System, as well as the confidentiality of student information, should be protected to ensure the validity, reliability, and fairness of the results. As stated in *Standard 7.9* (AERA, APA, & NCME, 2014), “The documentation should explain the steps necessary to protect test materials and to prevent inappropriate exchange of information during the test administration session” (p. 128).

This section of the *CSA 2018–19 Field Test Technical Report* describes the measures intended to prevent potential test security incidents prior to testing and the actions that were taken to handle security incidents occurring during or after the testing window using the Security and Test Administration Incident Reporting System (STAIRS) process.

There were zero test security incidents that were identified for the CSA field test.

3.2.1. ETS’ Office of Testing Integrity (OTI)

The OTI is a division of ETS that provides quality assurance services for all testing programs managed by ETS. This division resides in the ETS legal department. The Office of Professional Standards Compliance at ETS publishes and maintains the *ETS Standards for Quality and Fairness* (2014), which supports the OTI’s goals and activities. The *ETS Standards for Quality and Fairness* provides guidelines to help ETS staff design, develop, and deliver technically sound, fair, and beneficial products and services and help the public and auditors evaluate those products and services.

The OTI’s mission is to

- minimize any testing security violations that can impact the fairness of testing,
- minimize and investigate any security breach that threatens the validity of the interpretation of test scores, and
- report on security activities.

The OTI helps prevent misconduct on the part of students and administrators, detects potential misconduct through empirically established indicators, and resolves situations involving misconduct in a fair and balanced way that reflects the laws and professional standards governing the integrity of testing. In its pursuit of enforcing secure testing practices, the OTI strives to safeguard the various processes involved in a test development and administration cycle.

3.2.2. Procedures to Maintain Standardization of Test Security

Test security requires the accounting of all secure materials—including online summative test items and student data—before, during, and after each test administration. The LEA CAASPP coordinator is responsible for keeping all electronic test materials secure, keeping student information confidential, and making sure the CAASPP test site coordinators and test administrators are properly trained regarding security policies and procedures.

The CAASPP test site coordinator is responsible for mitigating test security incidents at the test site and for reporting incidents to the LEA CAASPP coordinator.

The test administrator is responsible for reporting testing incidents to the CAASPP test site coordinator and securely destroying printed and digital media for items and passages

generated by the print-on-demand feature of the test delivery system (CDE, 2018a and 2018b).

The following measures ensured the security of the CAASPP:

- LEA CAASPP coordinators and test site coordinators must have signed and submitted a “CAASPP Test Security Agreement for LEA CAASPP coordinators and CAASPP test site coordinators” form to the California Technical Assistance Center before ETS can grant the coordinators access to TOMS. (California Code of Regulations, Title 5 [5 CCR], Education, Division 1, Chapter 2, Subchapter 3.75, Article 1, Section 859[a])
- Anyone having access to the testing materials must have signed and submitted a “Test Security Affidavit for Test Examiners, Test Administrators, Proctors, Translators, Scribes, and Any Other Person Having Access to CAASPP Tests” form to the CAASPP test site coordinator before receiving access to any testing materials. (5 CCR, Section 859[c])

In addition, it was the responsibility of every participant in the CAASPP System to report immediately any violation or suspected violation of test security or confidentiality. The test site coordinator reported to the LEA CAASPP coordinator. The LEA CAASPP coordinator reported to the CDE within 24 hours of the incident. (5 CCR, Section 859[e])

3.2.3. Security of Electronic Files Using a Firewall

A firewall is software that prevents unauthorized entry to files, email, and other organization-specific information. All ETS data exchanges and internal email remain within the ETS firewall at all ETS locations, ranging from Princeton, New Jersey, to San Antonio, Texas, to Concord and Sacramento, California.

All electronic applications that are included in TOMS remain protected by the ETS firewall software at all times. Due to the sensitive nature of the student information processed by TOMS, the firewall plays a significant role in maintaining assurance of confidentiality among the users of this information.

Refer to the subsection [1.7 Systems Overview and Functionality](#) in [Chapter 1: Introduction](#) for more information on TOMS.

3.2.4. Transfer of Scores via Secure Data Exchange

Due to the confidential nature of test results, ETS uses secure file transfer protocol (SFTP) and encryption for all data file transfers, including student data files. SFTP is a method for reliable and exclusive routing of files. Files reside on a password-protected server that only authorized users can access. ETS shares an SFTP server with the CDE. On that site, ETS posts Microsoft Word and Excel files, Adobe Acrobat PDFs, or other document files for the CDE to review; the CDE returns reviewed materials in the same manner. Files are deleted upon retrieval.

The SFTP server is used as a conduit for the transfer of files; secure test data is stored only temporarily on the shared SFTP server. Industry-standard secure protocols are used to transfer test content and student data from the ETS internal data center to any external systems.

ETS enters information about the files posted to the SFTP server in a web form on a SharePoint website; a CDE staff member monitors this log throughout the day to check the

status of deliverables and downloads the file from the SFTP server when its status shows it has been posted.

3.2.5. Data Management in the Secure Database

ETS currently maintains a secure database to house all student demographic data and assessment results. Information associated with each student has a database relationship to the LEA, school, and grade codes as the data is collected during operational testing. Only individuals with the appropriate credentials can access the data. ETS builds all interfaces with the most stringent security considerations, including interfaces with data encryption for databases that store test items and student data. ETS applies best and up-to-date security practices, including system-to-system authentication and authorization, in all solution designs.

All stored test content and student data are encrypted. Industry-standard secure protocols are used to transfer test content and student data from the ETS internal data center to any external systems. ETS complies with the Family Educational Rights and Privacy Act (20 *United States Code [USC]* § 1232g; 34 *Code of Federal Regulations* Part 99) and the Children’s Online Privacy Protection Act (15 USC §§ 6501-6506, P.L. No. 105–277, 112 Stat. 2681–1728).

In TOMS, staff at LEAs and test sites have different levels of access appropriate to the role assigned to them.

3.2.6. Statistical Analysis on Secure Servers

During CAASPP testing, the information technology staff at ETS retrieves data files from the American Institutes for Research and loads them into a database. The ETS Data Quality Services staff extracts the data from the database and performs quality control procedures (e.g., the values of all variables are as expected) before passing files to the ETS statistical analysis group. The statistical analysis staff store the files on secure servers. All staff members involved with the data adhere to the ETS Code of Ethics and the ETS Information Protection Policies to prevent any unauthorized access to data.

3.2.7. Student Confidentiality

To meet requirements of the Every Student Succeeds Act, as well as state requirements, LEAs must collect demographic data about students’ ethnicity, disabilities, parent/guardian education, and so forth during the school year. ETS takes every precaution to prevent any of this information from becoming public or being used for anything other than for testing and score-reporting purposes. These procedures are applied to all documents in which student demographic data appears, such as technical reports.

3.2.8. Security and Test Administration Incident Reporting System (STAIRS) Process

Test security incidents, such as improprieties, irregularities, and breaches, are prohibited behaviors that give a student an unfair advantage or compromise the secure administration of the tests, which, in turn, compromise the reliability and validity of test results (CDE, 2018b). Whether intentional or unintentional, failure by staff or students to comply with security rules constitutes a test security incident. Test security incidents have impacts on scoring and affect students’ performance on the test.

LEA CAASPP coordinators and CAASPP test site coordinators must verify that all test security and summative administration incidents are documented by filling out the secure STAIRS form for reporting, which contains selectable options to guide coordinators in their submittal. Incidents are then resolved when the LEA CAASPP coordinator or CAASPP test site coordinator either files an appeal to reset, re-open, invalidate, or restore a student's test, or by following other instructions in a system-generated email in response to the STAIRS form submittal.

Prior to the field test administration, ETS and the CDE agreed that the following test security incidents would apply to the CSA field test; reports about these incidents were forwarded to the CDE.

- Administration error
- Cheating or accessing unauthorized devices
- Disruption or technical issue
- Exposing secure materials
- Incorrect Statewide Student Identifier used
- Student disruption

There were zero test security incidents identified during the CSA field test.

3.3. Accessibility Resources

The purpose of universal tools, designated supports, and accommodations in testing is to allow *all* students the opportunity to demonstrate what they know and what they are able to do, rather than giving students who use these resources an advantage over other students or artificially inflating their scores. Universal tools, designated supports, and accommodations minimize or remove barriers that could otherwise prevent students from demonstrating their knowledge, skills, and achievement in a specific content area.

The CSA field test offered commonly used accessibility resources available through the CAASPP online testing platform, where applicable for the tested construct. Some of these features could include a highlighter, the ability to mark an item for future review, and the ability to visually zoom the computer display in (making the display larger) or out (making the display smaller).

3.3.1. Universal Tools

Universal tools are available to all students by default, although they can be disabled if a student finds them distracting. Each universal tool falls into one of two categories: embedded and non-embedded. Embedded universal tools are provided through the student testing interface (through the CAASPP secure browser), although they can be turned off by a test administrator.

The following embedded universal tools were available to students during the CSA field test:

- Breaks
- Digital notepad
- Expandable passages
- Expandable items
- Highlighter
- Keyboard navigation
- Line reader

- Mark for review
- Strikethrough
- Writing tools (e.g., bold, italic, bullets, undo/redo) (full-write items)
- Zoom (in/out)

The following non-embedded universal tools were available for testing:

- Breaks
- Scratch paper
- Spanish dictionary (full-write items)
- Spanish thesaurus (full-write items)

3.3.2. Designated Supports

Designated supports are available to all students through the test settings in TOMS. The designated supports each fall into one of two categories: embedded and non-embedded. Embedded designated supports are provided through the student testing interface (through the CAASPP secure browser).

The following embedded designated supports were available during the CSA field test:

- Color contrast
- Masking
- Mouse pointer (size and color)
- Permissive mode
- Streamline
- Text-to-speech (items)
- Turn on/Turn off universal tool(s)

The following non-embedded designated supports were available during the CSA field test:

- Amplification
- Bilingual dictionary (full write items)
- Color contrast
- Color overlay
- Magnification
- Medical device
- Noise buffers
- Read aloud (items)
- Scribe (nonwriting items)
- Separate setting (special lighting/acoustics, adaptive furniture, time of day)
- Simplified test directions

3.3.3. Accommodations

Accommodations are changes in procedures or materials that increase equitable access during the CAASPP assessments. Assessment accommodations generate valid assessment results for students who need them; they allow these students to show what they know and can do. Accommodations do not compromise the learning expectations, construct, grade-level standard, or intended outcome of the assessments.

The following embedded accommodations were available during the CSA field test:

- Braille (embossed and refreshable)
- Closed captioning
- Text-to-speech (reading passages)

The following non-embedded accommodations were available during the CSA field test:

- Alternate response options
- Print on demand
- Read aloud (reading passages)
- Scribe (writing items)

3.3.4. Resources for Selection of Accessibility Resources

The full list of the universal tools, designated supports, and accommodations that are used in CAASPP online assessments are documented in Matrix One (CDE, 2018c). Part 1 of Matrix One lists the embedded universal tools, designated supports, and accommodations available for CAASPP online testing. Part 2 of Matrix One includes the non-embedded universal tools, designated supports, accommodations that are available. School-level personnel, individualized education program teams, and Section 504 teams use Matrix One when deciding how best to support the student's test-taking experience.

The Smarter Balanced Assessment Consortium's *Usability, Accessibility, and Accommodations Guidelines* ("Guidelines") (Smarter Balanced, 2018) aids in the selection of universal tools, designated supports, and accommodations deemed necessary for individual students. The *Guidelines* apply to all students and promote an individualized approach to the implementation of assessment practices. The *Guidelines* are intended to provide policy regarding universal tools, designated supports, and accommodations. Another manual, the *Smarter Balanced Usability, Accessibility, and Accommodations Implementation Guide* (Smarter Balanced, 2014), provides suggestions for implementation of these resources.

3.3.5. Delivery of Accessibility Resources

Universal tools, designated supports, and accommodations can be delivered as either embedded or non-embedded resources. Embedded resources are digitally delivered features or settings available as part of the technology platform for the online CAASPP assessments. Examples of embedded resources include the braille language resource, color contrast, and closed captioning for ELA listening items.

Non-embedded resources are not part of the technology platform for the computer-administered CAASPP tests. Examples of non-embedded resources include magnification, noise buffers, and the use of a scribe.

3.3.6. Unlisted Resources

Unlisted resources were not available for the CSA field test.

3.4. Participation

Because student participation in the operational CSA will be voluntary, the goal of the CSA field test recruitment was to have as many eligible students and LEAs as possible participate. All LEAs in California were invited to participate in the field test.

LEAs were given the following guidelines to determine if a student should take the CSA field test when either of these conditions apply:

- The student is receiving instruction in Spanish in the state of California.
- The student is seeking a measure that recognizes his or her Spanish reading, writing mechanics, and listening language arts skills.

As a result, a total of 52 LEAs participated in the CSA field test. [Table 3.1](#) presents the participation rates of each region.

Table 3.1 Participation Rates by Region

Region	# of LEAs	Total Students Registered	Total Students Completed	Mean Completion Rate	Minimum Completion Rate	Maximum Completion Rate
North	12	1,497	1,377	89.69	0	100
Central	18	4,742	3,761	70.50	0	100
South	22	4,111	2,884	72.66	0	100

3.5. Demographic Summaries

The number and the percent of students for selected groups with completed test scores are provided for grades three through eight and high school in [Table 3.A.1](#) through [Table 3.A.7](#) of [appendix 3.A](#). Grade levels reflect students' enrolled grade levels during the 2017–18 school year.

In the tables, students are grouped by demographic characteristics, including gender, ethnicity, English-language fluency, economic status (disadvantaged or not), special education services status, length of enrollment in U.S. schools, Spanish-language program type, and percentage of instruction in Spanish, as shown in [Table 3.2](#).

Note that data collected for program types and percentage of the school day instruction comes from the student demographic survey that was part of the field test. Note, too, that Spanish as a foreign language programs are only available for students in grades six through high school.

Table 3.2 Demographic Student Groups to Be Reported

Student Group	Definition
Gender	<ul style="list-style-type: none"> • Male • Female
Ethnicity	<ul style="list-style-type: none"> • American Indian or Alaska Native • Asian • Native Hawaiian or Other Pacific Islander • Filipino • Hispanic or Latino • Black or African American • White • Two or more races

Student Group	Definition
English-language Fluency	<ul style="list-style-type: none"> • English only • Initial fluent English proficient (IFEP) • English learner (EL) • Reclassified fluent English proficient (RFEP) • Ever-ELs (EL or RFEP) • To be determined • English proficiency unknown
Economic Status	<ul style="list-style-type: none"> • Not economically disadvantaged • Economically disadvantaged
Special Education Services Status	<ul style="list-style-type: none"> • No special education services • Special education services
Enrollment in U.S. Schools	<ul style="list-style-type: none"> • Less than 12 months • 12 months or more
Received instruction in Spanish in the 2017–18 school year—program type	<ul style="list-style-type: none"> • One-Way Immersion • Dual-Language Immersion • Developmental Bilingual • Heritage Language or Indigenous Language • Spanish as a Foreign Language³
Percentage of school day instruction provided in Spanish	<ul style="list-style-type: none"> • 0–25% • 26–50% • 51–75% • 76–100%

³ For students in grades six through high school

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- California Department of Education. (2018a). *CSA online field test administration manual*. Sacramento, CA: California Department of Education. Retrieved from <http://www.caaspp.org/rsc/pdfs/CSA.ft-otam.2018-19.pdf>
- California Department of Education. (2018c). *Matrix one: Universal tools, designated supports, and accommodations for the California Assessment of Student Performance and Progress for 2018–19*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/ta/tg/ai/caasppmatrix1.asp>
- California Department of Education. (2018b). *Security incidents and appeals procedure guide, 2017–18 administration*. Sacramento, CA: California Department of Education. Retrieved from <http://www.caaspp.org/rsc/pdfs/CAASPP.stairs-appeals-guide.2017-18.pdf>
- Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>
- Smarter Balanced Assessment Consortium. (2018). *Smarter Balanced Assessment Consortium: Usability, accessibility, and accommodations guidelines*. Los Angeles, CA: Smarter Balanced Assessment Consortium and National Center on Educational Outcomes. Retrieved from <https://portal.smarterbalanced.org/library/en/usability-accessibility-and-accommodations-guidelines.pdf>
- Smarter Balanced Assessment Consortium. (2014). *Smarter Balanced Assessment Consortium: Usability, accessibility, and accommodations implementation guide*. Los Angeles, CA: Smarter Balanced Assessment Consortium and National Center on Educational Outcomes. Retrieved from <https://portal.smarterbalanced.org/library/en/usability-accessibility-and-accommodations-implementation-guide.pdf>

Appendix 3.A Demographic Summaries

Notes:

- Data collected for Spanish instruction program types and percentage of school day instruction in Spanish comes from the student survey as part of the field test.
- Grade levels reflect students' enrolled grade levels during the 2017–18 school year.

Table 3.A.1 Demographic Summary—Grade Three

Demographic	Number Tested	Percent
Students completed the test	1,999	100.00
Male	939	46.97
Female	1,060	53.03
American Indian or Alaska Native	6	0.30
Asian	22	1.10
Native Hawaiian or Other Pacific Islander	0	0.00
Filipino	5	0.25
Hispanic or Latino	1,676	83.84
Black or African American	27	1.35
White	213	10.66
Two or more races	34	1.70
English only	729	36.47
Initial fluent English proficient (IFEP)	85	4.25
English learner (EL)	878	43.92
Reclassified fluent English proficient (RFEP)	286	14.31
Ever-ELs (EL or RFEP)	1,164	58.23
To be determined	18	0.90
English proficiency unknown	3	0.15
No special education services	1,864	93.25
Special education services	135	6.75
Not economically disadvantaged	1,243	62.18
Economically disadvantaged	756	37.82
In U.S. schools < 12 months	18	0.90
In U.S. schools ≥ 12 months	1,981	99.10
Received instruction in Spanish in the 2017–18 School Year—Total	1,926	96.35
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	39	2.02
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	1,728	89.72
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	74	3.84

Demographic	Number Tested	Percent
Received instruction in Spanish in the 2017–18 school year— Heritage Language or Indigenous Language Program	9	0.47
Received instruction in Spanish in the 2017–18 school year—None of the above	70	3.63
Percentage of school-day instruction provided in Spanish—0–25%	34	1.77
Percentage of school-day instruction provided in Spanish—26–50%	858	44.55
Percentage of school-day instruction provided in Spanish—51–75%	863	44.81
Percentage of school-day instruction provided in Spanish—76–100%	162	8.41

Table 3.A.2 Demographic Summary—Grade Four

Demographic	Number Tested	Percent
Students completed the test	1,899	100.00
Male	913	48.08
Female	986	51.92
American Indian or Alaska Native	3	0.16
Asian	22	1.16
Native Hawaiian or Other Pacific Islander	1	0.05
Filipino	10	0.53
Hispanic or Latino	1,490	78.46
Black or African American	36	1.90
White	277	14.59
Two or more races	44	2.32
English only	822	43.29
IFEP	54	2.84
EL	603	31.75
RFEP	417	21.96
Ever-ELs (EL or RFEP)	1,020	53.71
To be determined	3	0.16
English proficiency unknown	0	0.00
No special education services	1,768	93.10
Special education services	131	6.90
Not economically disadvantaged	1,223	64.40
Economically disadvantaged	676	35.60
In U.S. schools < 12 months	10	0.53
In U.S. schools ≥ 12 months	1,889	99.47
Received instruction in Spanish in the 2017–18 School Year—Total	1,852	97.53
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	40	2.16
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	1,556	84.02
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	49	2.65
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	10	0.54
Received instruction in Spanish in the 2017–18 school year—None of the above	195	10.53
Percentage of school-day instruction provided in Spanish—0–25%	178	9.61
Percentage of school-day instruction provided in Spanish—26–50%	900	48.60
Percentage of school-day instruction provided in Spanish—51–75%	688	37.15
Percentage of school-day instruction provided in Spanish—76–100%	84	4.54

Table 3.A.3 Demographic Summary—Grade Five

Demographic	Number Tested	Percent
Students completed the test	1,339	100.00
Male	646	48.24
Female	693	51.76
American Indian or Alaska Native	3	0.22
Asian	13	0.97
Native Hawaiian or Other Pacific Islander	0	0.00
Filipino	3	0.22
Hispanic or Latino	1,111	82.97
Black or African American	24	1.79
White	159	11.87
Two or more races	15	1.12
English only	500	37.34
IFEP	53	3.96
EL	374	27.93
RFEP	406	30.32
Ever-ELs (EL or RFEP)	780	58.25
To be determined	5	0.37
English proficiency unknown	1	0.07
No special education services	1,257	93.88
Special education services	82	6.12
Not economically disadvantaged	840	62.73
Economically disadvantaged	499	37.27
In U.S. schools < 12 months	15	1.12
In U.S. schools ≥ 12 months	1,324	98.88
Received instruction in Spanish in the 2017–18 School Year—Total	1,280	95.59
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	24	1.88
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	1,213	94.77
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	23	1.80
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	3	0.23
Received instruction in Spanish in the 2017–18 school year—None of the above	16	1.25
Percentage of school-day instruction provided in Spanish—0–25%	21	1.64
Percentage of school-day instruction provided in Spanish—26–50%	743	58.05
Percentage of school-day instruction provided in Spanish—51–75%	444	34.69
Percentage of school-day instruction provided in Spanish—76–100%	71	5.55

Table 3.A.4 Demographic Summary—Grade Six

Demographic	Number Tested	Percent
Students completed the test	965	100.00
Male	446	46.22
Female	519	53.78
American Indian or Alaska Native	5	0.52
Asian	6	0.62
Native Hawaiian or Other Pacific Islander	1	0.10
Filipino	2	0.21
Hispanic or Latino	838	86.84
Black or African American	14	1.45
White	85	8.81
Two or more races	10	1.04
English only	295	30.57
IFEP	44	4.56
EL	232	24.04
RFEP	389	40.31
Ever-ELs (EL or RFEP)	621	64.35
To be determined	5	0.52
English proficiency unknown	0	0.00
No special education services	923	95.65
Special education services	42	4.35
Not economically disadvantaged	532	55.13
Economically disadvantaged	433	44.87
In U.S. schools < 12 months	18	1.87
In U.S. schools ≥ 12 months	947	98.13
Received instruction in Spanish in the 2017–18 School Year—Total	906	93.89
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	7	0.77
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	837	92.38
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	21	2.32
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	3	0.33
Received instruction in Spanish in the 2017–18 school year—Spanish as a Foreign Language Program	17	1.88
Received instruction in Spanish in the 2017–18 school year—None of the above	21	2.32

Demographic	Number Tested	Percent
Percentage of school-day instruction provided in Spanish—0–25%	94	10.38
Percentage of school-day instruction provided in Spanish—26–50%	685	75.61
Percentage of school-day instruction provided in Spanish—51–75%	85	9.38
Percentage of school-day instruction provided in Spanish—76–100%	42	4.64

Table 3.A.5 Demographic Summary—Grade Seven

Demographic	Number Tested	Percent
Students completed the test	892	100.00
Male	409	45.85
Female	483	54.15
American Indian or Alaska Native	1	0.11
Asian	8	0.90
Native Hawaiian or Other Pacific Islander	1	0.11
Filipino	0	0.00
Hispanic or Latino	725	81.28
Black or African American	7	0.78
White	126	14.13
Two or more races	16	1.79
English only	317	35.54
IFEP	36	4.04
EL	134	15.02
RFEP	394	44.17
Ever-ELs (EL or RFEP)	528	59.19
To be determined	11	1.23
English proficiency unknown	0	0.00
No special education services	848	95.07
Special education services	44	4.93
Not economically disadvantaged	547	61.32
Economically disadvantaged	345	38.68
In U.S. schools < 12 months	13	1.46
In U.S. schools ≥ 12 months	879	98.54
Received instruction in Spanish in the 2017–18 School Year—Total	861	96.52
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	10	1.16
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	755	87.69
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	11	1.28
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	6	0.70
Received instruction in Spanish in the 2017–18 school year—Spanish as a Foreign Language Program	54	6.27
Received instruction in Spanish in the 2017–18 school year—None of the above	24	2.79

Demographic	Number Tested	Percent
Percentage of school-day instruction provided in Spanish—0–25%	251	29.15
Percentage of school-day instruction provided in Spanish—26–50%	498	57.84
Percentage of school-day instruction provided in Spanish—51–75%	56	6.50
Percentage of school-day instruction provided in Spanish—76–100%	56	6.50

Table 3.A.6 Demographic Summary—Grade Eight

Demographic	Number Tested	Percent
Students completed the test	279	100.00
Male	134	48.03
Female	145	51.97
American Indian or Alaska Native	2	0.72
Asian	2	0.72
Native Hawaiian or Other Pacific Islander	1	0.36
Filipino	1	0.36
Hispanic or Latino	231	82.80
Black or African American	2	0.72
White	11	3.94
Two or more races	1	0.36
English only	37	13.26
IFEP	12	4.30
EL	46	16.49
RFEP	158	56.63
Ever-ELs (EL or RFEP)	204	73.12
To be determined	25	8.96
English proficiency unknown	1	0.36
No special education services	275	98.57
Special education services	4	1.43
Not economically disadvantaged	150	53.76
Economically disadvantaged	129	46.24
In U.S. schools < 12 months	10	3.58
In U.S. schools ≥ 12 months	269	96.42
Received instruction in Spanish in the 2017–18 School Year—Total	215	77.06
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	16	7.44
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	125	58.14
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	12	5.58
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	20	9.30
Received instruction in Spanish in the 2017–18 school year—Spanish as a Foreign Language Program	20	9.30
Received instruction in Spanish in the 2017–18 school year—None of the above	22	10.23

Demographic	Number Tested	Percent
Percentage of school-day instruction provided in Spanish—0–25%	51	23.72
Percentage of school-day instruction provided in Spanish—26–50%	94	43.72
Percentage of school-day instruction provided in Spanish—51–75%	30	13.95
Percentage of school-day instruction provided in Spanish—76–100%	40	18.60

Table 3.A.7 Demographic Summary—High School

Demographic	Number Tested	Percent
Students completed the test	649	100.00
Male	266	40.99
Female	383	59.01
American Indian or Alaska Native	2	0.31
Asian	14	2.16
Native Hawaiian or Other Pacific Islander	1	0.15
Filipino	2	0.31
Hispanic or Latino	490	75.50
Black or African American	6	0.92
White	75	11.56
Two or more races	0	0.00
English only	147	22.65
IFEP	10	1.54
EL	135	20.80
RFEP	281	43.30
Ever-ELs (EL or RFEP)	416	64.10
To be determined	67	10.32
English proficiency unknown	9	1.39
No special education services	640	98.61
Special education services	9	1.39
Not economically disadvantaged	377	58.09
Economically disadvantaged	272	41.91
In U.S. schools < 12 months	35	5.39
In U.S. schools ≥ 12 months	614	94.61
Received instruction in Spanish in the 2017–18 School Year—Total	559	86.13
Received instruction in Spanish in the 2017–18 school year—One-Way Immersion program	27	4.83
Received instruction in Spanish in the 2017–18 school year—Dual-Language Immersion program	45	8.05
Received instruction in Spanish in the 2017–18 school year—Developmental Bilingual Program	16	2.86
Received instruction in Spanish in the 2017–18 school year—Heritage Language or Indigenous Language Program	124	22.18
Received instruction in Spanish in the 2017–18 school year—Spanish as a Foreign Language Program	261	46.69
Received instruction in Spanish in the 2017–18 school year—None of the above	86	15.38

Demographic	Number Tested	Percent
Percentage of school-day instruction provided in Spanish—0–25%	351	62.79
Percentage of school-day instruction provided in Spanish—26–50%	57	10.20
Percentage of school-day instruction provided in Spanish—51–75%	34	6.08
Percentage of school-day instruction provided in Spanish—76–100%	117	20.93

Chapter 4: Summary Statistics for the 2018–19 Field Test Administration

4.1. Overview

This chapter summarizes the results of the item- and test-level analyses for the 2018–19 California Spanish Assessment (CSA) field test, including the following:

- Test completion analyses
- Test form reliability
- Classical item analyses
- Response time analyses
- Differential item functioning (DIF) analyses

Note that the analyses results of the scores are not intended for reporting but are used only for research and future test development.

4.2. Sample Used for the Analyses

The field test was administered to eligible students in grades four through twelve in fall 2018 to accommodate the test development schedule. Because fourth-graders in the fall are generally equivalent to the third-graders in spring, administration of the CSA field test started with grade four instead of grade three.

Two item analyses were run for the CSA field test: the preliminary item analyses (PIA) and the final item analyses (FIA). The PIA identifies potentially problematic items for further evaluation. PIA is typically conducted as soon as sufficient volume of data is collected to obtain reliable and valid estimates. The FIA was conducted after the administration was completed.

All students' responses that met the completion rule—that a student has a test logon and answered at least 10 items—were included in the analyses. All analyses in this report were performed using the final file for the FIA. All results presented in this chapter use the enrolled grades from the 2017–18 school year.

[Table 4.1](#) presents, for each grade level tested, the total number of students registered for the CSA field test as well as the number and percent of students who completed the test.

Table 4.1 Summary of Completion of the Field Test by Grade Level

Grade Level Tested ⁴	Total Number of Students Completing the Test	Percent Completion	Total Number of Registered Test Takers
3	1,999	84.96	2,353
4	1,899	84.51	2,247
5	1,339	85.89	1,559

⁴ The grade levels tested are included in the table. As the field test was administered in the fall, the students' actual grade level was one grade higher than the grade level in the *Grade Level Tested* column. For example, 1,999 fourth-graders completed the grade three test.

Grade Level Tested ⁴	Total Number of Students Completing the Test	Percent Completion	Total Number of Registered Test Takers
6	965	64.29	1,501
7	892	69.74	1,279
8	279	69.23	403
High school	649	64.38	1,008

4.3. Raw Score Distributions

For all of the CSA field tests, the total test raw score is defined as total points obtained for machine-scorable items. The summary statistics of the total test raw scores for grades three through eight and high school assessments are presented in [Table 4.2](#), which contains data for each test form. The summary statistics presented include the mean, mean as percent of total score points, and standard deviation (SD). For the grade eight and high school assessments, there were no students taking the accommodated forms. Note that only machine-scorable items are included in the analysis. One constructed-response (CR) item for each form is not included in the analysis.⁵

In addition, [Table 4.A.1](#) through [Table 4.A.7](#) present the distributions of the total test raw score of each form for grades three through eight and high school.

Table 4.2 Summary Statistics of the Raw Scores

Grade Level Tested	Form	No. of Items	N Points	N Students	Mean	Mean as % of Total	SD
3	1	59	70	934	33.26	47.52	9.56
3	2	59	69	941	32.91	47.70	10.55
3	A	59	67	124	28.06	41.88	8.79
4	1	59	73	894	34.32	47.01	10.75
4	2	59	76	883	35.45	46.65	11.07
4	A	59	70	122	36.19	51.70	10.34
5	1	59	71	648	31.87	44.89	9.50
5	2	59	70	652	32.89	46.99	9.26
5	A	59	71	39	32.74	46.12	7.25
6	1	59	74	458	37.28	50.38	10.92
6	2	59	74	467	38.55	52.10	10.99
6	A	59	72	40	36.45	50.63	10.00

⁵ CR items will be human scored in May through July 2019. The purpose of developing those CR items is to provide an opportunity for LEAs and schools to self-score those CRs in future administrations. Educational Testing Service will create sets of benchmark (anchor) and training papers, with sample-specific annotations, as well as sets of validity and calibration samples for each CR item for the local scoring. Those CR scores are not included in CSA summative score reports.

Grade Level Tested	Form	No. of Items	N Points	N Students	Mean	Mean as % of Total	SD
7	1	59	71	421	30.60	43.10	10.10
7	2	59	71	436	31.11	43.81	9.40
7	A	59	70	35	30.57	43.67	6.56
8	1	59	73	138	33.31	45.63	8.17
8	2	59	74	141	34.57	46.72	9.07
8	A	59	66	0	NA	NA	NA
High school	1	59	73	327	33.36	45.70	8.31
High school	2	59	71	322	33.65	47.40	8.93
High school	A	59	72	0	NA	NA	NA

4.4. Reliability and Standard Error of Measurement (SEM) Estimation

Reliability of the test scores is the extent to which differences in test scores reflect true differences in the knowledge, ability, or skill being tested, rather than fluctuations due to measurement error. Thus, reliability is the consistency of the scores across conditions that can be assumed to differ at random, especially which form of the test the student is administered. In statistical terms, the variance in the distributions of test scores—essentially, the differences among individuals—is due partly to real differences in the knowledge, skill, or ability being tested (true variance) and partly to random errors in the measurement process (error variance). The reliability coefficient is an estimate of the proportion of the total variance that is true variance.

There are several different ways of estimating reliability. The type of reliability estimate reported here is an internal-consistency measure, which is derived from analysis of the consistency of the performance of individuals across items within a test.

Reliability coefficients range from 0 to 1. The higher the reliability coefficient for a set of scores, the more likely individuals are to obtain very similar scores upon repeated testing occasions, if the students do not change in their level of the knowledge or skills measured by the test.

SEM quantifies the amount of error in the test scores. SEM is the extent by which students' scores tend to differ from the scores they would receive if the assessment were perfectly reliable. As the SEM increases, the variability of students' observed scores is likely to increase across repeated testing. Observed scores with large SEM pose a challenge to the valid interpretation of a single test score.

For the CSA field test, reliability and SEM estimates were calculated at the test-form level.

4.4.1. Raw Score Reliability Estimation

Coefficient alpha (Cronbach, 1951), which measures internal consistency reliability, is the most commonly used measure of reliability. Coefficient alpha is estimated by substituting sample estimates for the parameters and is defined as follows:

$$\alpha = \frac{K}{K - 1} \left(1 - \frac{\sum_{i=1}^K S_{X_i}^2}{S_X^2} \right), \tag{4.1}$$

where,

K is the number of items in the test,

$S_{X_i}^2$ is the observed variance of item i in the test, and

S_X^2 is the observed variance of the total test score.

Since CSA field test forms have mixed item types (dichotomous and polytomous items), it is more appropriate to report stratified alpha (Feldt & Brennan, 1989). Stratified alpha is a weighted average of coefficient alphas for item sets with different maximum score points or “strata.” It is a reliability estimate computed by dividing the test into parts (strata), computing coefficient alpha separately for each part, and using the results to estimate a reliability coefficient for the total score. The formula for the stratified alpha is:

$$\rho_{strata} = 1 - \frac{\sum S_{X_j}^2 (1 - \alpha_j)}{S_X^2}, \quad (4.2)$$

where,

$\sigma_{X_j}^2$ is the variance for strata j of the test,

σ_X^2 is the total variance of the test, and

α_j is the Cronbach’s alpha for strata j of the test.

Estimates of stratified alpha are computed by substituting sample estimates for the parameters in the formula.

4.4.2. Standard Error of Measurement (SEM)

The SEM provides a measure of score instability in a different metric. The formula for the SEM is:

$$SEM = S_X \sqrt{1 - \rho_{strata}} \quad (4.3)$$

where,

ρ_{strata} is the reliability estimated in equation 4.2, and

S_X is the standard deviation of the total score.

4.4.3. Results for the Field-Test Forms

[Table 4.3](#) provides the reliability estimates and SEM for each test form per grade level. Note that only machine-scorable items are included in the analysis.

Table 4.3 Test Reliability of the Total Scores

Grade Level Tested	Form	No. of Items	N Points	N Students	Mean	SD	Reliability	SEM
3	1	59	70	934	33.26	9.56	0.820	4.056
3	2	59	69	941	32.91	10.55	0.859	3.962
3	A	59	67	124	28.06	8.79	0.826	3.667
4	1	59	73	894	34.32	10.75	0.854	4.108
4	2	59	76	883	35.45	11.07	0.866	4.052
4	A	59	70	122	36.19	10.34	0.869	3.742
5	1	59	71	648	31.87	9.50	0.823	3.997
5	2	59	70	652	32.89	9.26	0.813	4.004
5	A	59	71	39	32.74	7.25	0.732	3.753
6	1	59	74	458	37.28	10.92	0.859	4.100
6	2	59	74	467	38.55	10.99	0.863	4.068
6	A	59	72	40	36.45	10.00	0.860	3.742
7	1	59	71	421	30.60	10.10	0.841	4.027
7	2	59	71	436	31.11	9.40	0.801	4.193
7	A	59	70	35	30.57	6.56	0.684	3.688
8	1	59	73	138	33.31	8.17	0.792	3.726
8	2	59	74	141	34.57	9.07	0.805	4.005
8	A	59	66	0	NA	NA	NA	NA
High school	1	59	73	327	33.36	8.31	0.752	4.138
High school	2	59	71	322	33.65	8.93	0.794	4.053
High school	A	59	72	0	NA	NA	NA	NA

4.5. Classical Item Analyses

For all field-tested items that are machine-scorable, classical item analyses are used to evaluate the item performance with respect to item difficulty, item discrimination, and distractor analysis. In addition, the distributions of score categories on key-based selected-response items and rule-based machine-scored items are also included in the classical item analyses results. Lastly, the associated flagging rules of these statistics are used to identify items that are not performing as expected. Items scored as one (correct) or zero (incorrect) are referred to as dichotomous items. Items with maximum score greater than one are called polytomous items. [Table 4.4](#) presents the summary results of item difficulty and item-total correlation in each form by grade level. [Table 4.5](#) presents the summary results of flagged items in each form by grade level. In addition, [appendix 4.B](#) presents results of the classical item analyses for dichotomous and polytomous items by grade level, as well as the summary statistics by item type and item and response types.

4.5.1. Item Difficulty (Overall and by Item Type)

For dichotomous items, item difficulty is indicated by its p -value, which is the proportion of students who answer the item correctly. The range of p -values is from 0.00 to 1.00. Items with high p -values are easier items; those with low p -values are more difficult items. Dichotomous items are flagged for review if they have p -values above 0.95 (i.e., too easy) or below 0.10 (i.e., too difficult).

The formula for the p -value for a dichotomous item is:

$$p\text{-value}_{dich} = \frac{\sum X_{ic}}{N_i} \quad (4.4)$$

where,

X_{ic} is the number of students that answered item i correctly, and

N_i is the total number of students who were presented with item i .

For polytomous items, the difficulty is indicated by the average item score (AIS). The AIS can range from 0.00 to the maximum total possible points for an item. Desired AIS values for polytomous items generally fall within the range of 30 percent to 70 percent of the maximum obtainable item score; items with values outside this range are flagged for review. To facilitate the interpretation, the AIS values for polytomous items are often expressed as the proportion of the maximum possible score, which are equivalent to the p -values of dichotomous items.

The formula for the p -value for a polytomous item is:

$$p\text{-value}_{poly} = \frac{\sum_j X_{ij}}{N_i \times \text{Max}(X_i)} \quad (4.5)$$

where,

X_{ij} is the score assigned for a given polytomous item i and student j ,

N_i is the total number of students who were presented with item i , and

$\text{Max}(X_i)$ is the maximum possible score for item i .

4.5.2. Item Discrimination (Overall and by Item Type)

The item-total correlation statistic describes the relationship between students' performance on a specific item and their performance on the total test. It is calculated as the correlation coefficient between the item score and total score. In general, item-total correlation ranges from -1.0 (for a perfect negative relationship) to 1.0 (for a perfect positive relationship).

A relatively high positive item-total correlation coefficient value is desired, as it indicates that students with higher scores on the overall test tend to perform better on the item. A negative item-total correlation typically signifies a problem with the item, as the students with higher scores on the overall test are more likely to get the item wrong or receive a low score, and the students with lower scores on the overall test are more likely to get the item correct or a high score.

For the CSA, the polyserial correlation is used for both polytomous and dichotomous items. Statistically, polyserial correlations are based on a polyserial regression model (Olsson, 1979; Drasgow, 1988), which assumes that performance on an item is determined by the examinee’s position on an underlying latent variable that is normally distributed at a given criterion score level. Based on this approach, the polyserial correlation can be estimated as:

$$r_{polyreg} = \frac{\hat{\beta} s_{tot}}{\sqrt{\hat{\beta}^2 s_{tot}^2 + 1}} \quad (4.6)$$

where,

s_{tot} is the standard deviation of the students’ total test scores as a criterion score, and

β is the item parameter to be estimated from the data, with the estimate denoted as $\hat{\beta}$, using maximum likelihood estimation. It is a regression coefficient (slope) for predicting the continuous version of an item score onto the continuous version of the total score.

There are as many regressions as there are boundaries between scores with all sharing a common slope, β . For a polytomous item, there are $m-1$ regressions, where m is the number of score points on the item. Beta (β) is the common slope for all $m-1$ regressions. Desired polyserial correlation values of items are positive and larger than 0.20.

4.5.3. Distractor Analyses (Overall and by Item Type)

The quality of distractors is an important component of an item’s overall quality. Distractors should be clearly incorrect, but at the same time be plausible and attractive to students who do not understand the content or skills being assessed. For the CSA field test, the following distractor analyses were conducted to evaluate the quality of distractors.

4.5.3.1 The Proportion of Students Choosing Each Distractor

The percentage of students at each response option is calculated for the highest-performing 20 percent of students. If the percentage of students who selected a distractor is greater than the percentage of students who selected the correct answer for the high-performing group, the item is flagged and examined to determine if it has multiple correct answers or the wrong key (i.e., the item is miskeyed).

4.5.3.2 Polyserial Correlation

The polyserial correlation is calculated for each response option. While the key should have a positive polyserial correlation with the criterion score, the distractors should exhibit negative polyserial correlations (i.e., lower-ability students would likely choose the distractors, while higher-ability students would not). An item with a positive distractor-total correlation is flagged for review, as this item may have multiple correct answers, be miskeyed, or have other content issues.

4.5.4. Summary of Classical Item Analyses Flagging Criteria

In summary, an item is flagged for review if the item analysis yields any of the following results. One item could have multiple flags if the statistics meet the flagging criteria:

- A **difficulty flag** indicates extreme values of the proportion-correct (for dichotomous items) or the proportion of the possible maximum points earned (for polytomous items).
 - A-flag: A value less than 0.10 for dichotomous items or 0.30 for polytomous items suggests that the item might be too difficult.
 - H-flag: A value greater than 0.95 for dichotomous items or 0.70 for polytomous items suggests that the item might be too easy.
- A **discrimination flag** (R-flag) indicates that the item does not discriminate effectively between high- and low-ability students. Items with a polyserial correlation less than 0.20 are flagged.
- An **omit flag** (O-flag) is set for dichotomous items with nonresponse rates greater than five percent and polytomous items with nonresponse rates greater than 10 percent.
- A **distractor flag** (P-flag) is used for any distractors having positive correlation with the criterion score.
- A **miskey flag** (D-flag) is used for multiple-choice (MC) items when more of the high-ability examinee group—the top 20 percent of examinees on the total assessment—choose any distractor rather than choosing the response keyed as correct.

ETS’s psychometric staff and assessment development staff carefully reviewed each of the flagged items during and at the end of the item analyses. All confirmed flagged items were also reviewed by a panel of educators and then summarized for the California Department of Education with recommendations for subsequent analyses.

4.5.5. Distribution of Item Scores

For polytomous items, examination of the distribution of scores helps evaluate how well the item functions. If no students achieved the highest possible score, the item may not be functioning as expected. The item may be confusing, poorly worded, unexpectedly difficult, or students may not have had an opportunity to learn the content.

4.5.6. Classical Item Analyses Results Summaries

The summary statistics of the classical item analyses, which include the means and ranges of overall item difficulty and item-total correlation for all machine-scorable items, is presented in [Table 4.4](#) for each test form in all grade levels tested. For all field tests within each grade level—except for the accommodated forms in grade eight and high school, which were not taken by any students in those grades—the average item difficulties and discriminations were reasonable, as shown in [Table 4.4](#).

Table 4.4 Item Difficulty and Item-Total Correlation in Each Form by Grade Level

Grade Level Tested	Form	No. of Items	No. of Examinees	Mean Difficulty	Mean Item-Total Corr.	Min. Difficulty	Min. Item-Total Corr.	Max. Difficulty	Max. Item-Total Corr.
3	1	59	934	0.47	0.36	0.15	-0.06	0.91	0.80
3	2	59	941	0.48	0.40	0.20	-0.09	0.81	0.71
3	A	59	124	0.43	0.37	0.19	-0.19	0.81	0.80
4	1	59	894	0.47	0.40	0.15	-0.00	0.78	0.71
4	2	59	883	0.46	0.40	0.16	-0.07	0.78	0.73
4	A	59	122	0.49	0.42	0.15	-0.03	0.78	0.73
5	1	59	648	0.45	0.36	0.07	0.01	0.84	0.64
5	2	59	652	0.47	0.35	0.12	-0.06	0.78	0.74
5	A	59	39	0.46	0.33	0.07	-0.20	0.78	0.74
6	1	59	458	0.49	0.40	0.06	-0.02	0.81	0.70
6	2	59	467	0.51	0.40	0.18	-0.05	0.87	0.71
6	A	59	40	0.49	0.39	0.03	-0.27	0.87	0.88
7	1	59	421	0.42	0.38	0.03	-0.07	0.84	0.69
7	2	59	436	0.43	0.34	0.03	-0.19	0.89	0.62
7	A	59	35	0.42	0.32	0.00	-0.39	0.89	0.88
8	1	59	138	0.43	0.34	0.04	-0.28	0.88	0.78
8	2	59	141	0.46	0.35	0.13	-0.08	0.91	0.66
8	A	59	0	NA	NA	NA	NA	NA	NA
High school	1	59	327	0.44	0.31	0.19	-0.19	0.86	0.61
High school	2	59	322	0.46	0.33	0.17	-0.12	0.78	0.53
High school	A	59	0	NA	NA	NA	NA	NA	NA

The summary of flagged items across the test forms by grade level is presented in [Table 4.5](#).

Table 4.5 Flagged Items Summary in Each Form by Grade Level

Grade Level Tested	Form	No. of Items	No. of Flag A Items	Percent of Flag A Items	No. of Flag H Items	Percent of Flag H Items	No. of Flag R Items	Percent of Flag R Items	No. of Flag D Items	Percent of Flag D Items	No. of Flag P Items	Percent of Flag P Items
3	1	59	0	0.00	2	3.39	10	16.95	3	5.08	17	28.81
3	2	59	0	0.00	2	3.39	7	11.86	5	8.47	10	16.95
3	A	59	0	0.00	1	1.69	11	18.64	8	13.56	18	30.51

Grade Level Tested	Form	No. of Items	No. of Flag A Items	Percent of Flag A Items	No. of Flag H Items	Percent of Flag H Items	No. of Flag R Items	Percent of Flag R Items	No. of Flag D Items	Percent of Flag D Items	No. of Flag P Items	Percent of Flag P Items
4	1	59	0	0.00	2	3.39	6	10.17	4	6.78	9	15.25
4	2	59	0	0.00	3	5.08	9	15.25	6	10.17	15	25.42
4	A	59	0	0.00	3	5.08	7	11.86	5	8.47	17	28.81
5	1	59	1	1.69	0	0.00	11	18.64	6	10.17	20	33.90
5	2	59	0	0.00	0	0.00	13	22.03	8	13.56	13	22.03
5	A	59	1	1.69	1	1.69	14	23.73	9	15.25	21	35.59
6	1	59	2	3.39	6	10.17	8	13.56	4	6.78	14	23.73
6	2	59	1	1.69	6	10.17	9	15.25	7	11.86	14	23.73
6	A	59	2	3.39	4	6.78	8	13.56	7	11.86	16	27.12
7	1	59	5	8.47	4	6.78	10	16.95	8	13.56	20	33.90
7	2	59	4	6.78	0	0.00	12	20.34	11	18.64	19	32.20
7	A	59	3	5.08	2	3.39	15	25.42	10	16.95	28	47.46
8	1	59	2	3.39	3	5.08	13	22.03	9	15.25	23	38.98
8	2	59	1	1.69	2	3.39	13	22.03	4	6.78	18	30.51
8	A	59	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
HS	1	59	1	1.69	1	1.69	11	18.64	7	11.86	20	33.90
HS	2	59	1	1.69	0	0.00	9	15.25	8	13.56	20	33.90
HS	A	59	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

In addition, the item statistics, including the p -value, polyserial correlation, omit rates, and not-reached rates for dichotomous items, are presented in [Table 4.B.1](#) through [Table 4.B.21](#). The item statistics, including AIS, polyserial correlation, not-reached rates, and the distribution of score points on each polytomous item are listed in [Table 4.B.22](#) through [Table 4.B.28](#).

[Table 4.B.29](#) through [Table 4.B.35](#) present the summary of classical item statistics by item type. Data is shown for all item types, with the exception of constructed-response (CR) items, for grades three through eight and high school.

[Table 4.B.36](#) through [Table 4.B.42](#) present the summary of statistics by item and response types. Data is shown for all item types, with the exception of CR items, for grades three through eight and high school.

4.6. Response Time Analyses

The length of time it takes students to complete a test is recorded and analyzed to build a profile describing what a typical testing event looks like for each test. In addition, variability in testing time is investigated to determine whether a student's testing time should be viewed as unusual or irregular for further investigation. It should be noted that the CSA assessments are untimed.

In these analyses, all students who completed the test are considered. The testing population is partitioned into quartiles based on machine-scorable items. The descriptive statistics—for example, the number of students, mean, standard deviation, minimum and maximum, percentiles—of the time required to complete the total test for the machine-scorable items are computed for each of the four quartile groups for each grade level (i.e., grades three through eight and high school grade levels).

[Appendix 4.C](#) summarizes results of testing time analysis. [Table 4.C.1](#) through [Table 4.C.9](#) provide descriptive statistics of total testing time for the full student population at each ability level for each grade level. The unit of testing time is in minutes; for example, in [Table 4.C.1](#), the median (i.e., 50th percentile) of the testing time for grade three Q1 group is 65.02 minutes. Overall, there is no consistent pattern in testing times across the grade levels, given that the CSA field tests were untimed.

4.7. Differential Item Functioning (DIF) Analyses

Analyses of DIF can provide evidence of the degree to which an item score interpretation or use is valid for individuals who differ in their demographic characteristics. An item may be biased if it contains content or language that is differentially familiar to student groups. It is important, however, to recognize that item performance differences flagged for DIF might be related to actual difference in relevant knowledge or skills (group impact) or statistical Type I error, which might falsely assert DIF exists for an item. As a result, DIF statistics are used to identify potential item bias. Subsequent reviews by content experts and bias and sensitivity experts are required to determine the source and meaning of item performance differences.

DIF analyses were performed on all machine-scorable items. In examining the DIF between groups, the reference group is often designated as the group assumed to have an advantage, while the focal group refers to the group anticipated to be disadvantaged by the test. The sample size requirements for the DIF analyses were 100 in the smaller of either group and 400 in the combined focal and reference groups. These sample size requirements are based on standard operating procedures with respect to DIF analyses at ETS to ensure reliable DIF results can be obtained.

4.7.1. DIF Procedure for Dichotomous Items

The DIF analyses for dichotomous items utilized the Mantel-Haenszel (MH) DIF statistic (Mantel & Haenszel, 1959; Holland & Thayer, 1988). For this method, students are classified to relevant student groups of interest (e.g., gender or ethnicity). Students at each total-score level in the focal group (e.g., females) are compared with students at each total-score level in the reference group (e.g., males). The common odds ratio—that is, the proportion of correct response over the proportion of incorrect response—is estimated across all levels of matched student ability using the formula in equation 4.1 (Dorans & Holland, 1993). The resulting estimate is interpreted as the relative probability of success on a particular item for members of two groups when matched on ability.

$$\alpha_{MH} = \frac{\left(\sum_m R_{rm} \frac{W_{fm}}{N_{tm}} \right)}{\left(\sum_m R_{fm} \frac{W_{rm}}{N_{tm}} \right)} \quad (4.7)$$

where,

m is the number of score categories of the total test,

R_{rm} is the number of students in the reference group who answer the item correctly at score level m ,

W_{fm} is the number of students in the focal group who answer the item incorrectly at score level m ,

N_m is the total number of students at score level m ,

R_{fm} is the number of students in the focal group who answer the item correctly at score level m , and

W_{rm} is the number of students in the reference group who answer the item incorrectly at score level m .

To facilitate the interpretation of MH results, the common odds ratio is frequently transformed to the delta scale using the following formula (Holland & Thayer, 1988):

$$MH\ D - DIF = -2.35 \ln[\alpha_{MH}] \quad (4.8)$$

Positive values indicate DIF in favor of the focal group (i.e., positive DIF items are differentially easier for the focal group), whereas negative values indicate DIF in favor of the reference group (i.e., negative DIF items are differentially harder for the focal group).

4.7.2. DIF Procedure for Polytomous Items

The standardization DIF (Dorans & Schmitt, 1993; Zwick, Thayer, & Mazzeo, 1997; Dorans, 2013) in conjunction with the Mantel chi-square statistic (Mantel, 1963; Mantel & Haenszel, 1959) is calculated for polytomous items. The standardized mean difference (SMD) compares the item means of the two groups after adjusting for differences in the distribution of students across all items and is calculated using the following formula:

$$SMD = \frac{\sum_{m=1}^M N_{fm} \times E_f(Y | X = m)}{\sum_{m=1}^M N_{fm}} - \frac{\sum_{m=1}^M N_{rm} \times E_r(Y | X = m)}{\sum_{m=1}^M N_{rm}} \quad (4.9)$$

where,

X is the criterion score (total raw score),

Y is the item score,

M is the number of score levels on X ,

N_{rm} is the number of students in the reference group at score level m ,

N_{fm} is the number of students in the focal group at score level m ,

E_r is the expected item score for the reference group, and

E_f is the expected item score for the focal group.

A positive SMD value means that, conditional on the criterion score, the focal group has a higher mean item score than the reference group (i.e., the item is differentially easier for the focal group). In contrast, a negative SMD value means that, conditional upon the criterion score, the focal group has a lower mean item score than the reference group (i.e., the item is differentially harder for the focal group).

4.7.3. Classification

Based on the DIF statistics and significance tests, items are classified into three categories and assigned values of A, B, or C. Category A items contain negligible DIF, Category B items exhibit slight to moderate DIF, and Category C items possess moderate to large DIF values.

The flagging criteria for dichotomous items are presented in [Table 4.6](#); the flagging criteria for polytomous items are provided in [Table 4.7](#).

Table 4.6 DIF Categories for Dichotomous Items

DIF Category	Criteria
A (negligible)	<ul style="list-style-type: none"> Absolute value of MH D-DIF is not significantly different from zero at the 0.05 level, or is less than one. Positive values are classified as “A+” and negative values as “A-.”
B (moderate)	<ul style="list-style-type: none"> Absolute value of MH D-DIF is significantly different from zero but not from one at the 0.05 level, and is at least one; OR Absolute value of MH D-DIF is significantly different from one, but is less than 1.5. Positive values are classified as “B+” and negative values as “B-.”
C (large)	<ul style="list-style-type: none"> Absolute value of MH D-DIF is significantly greater than one at the 0.05 level, and is at least 1.5. Positive values are classified as “C+” and negative values as “C-.”

Table 4.7 DIF Categories for Polytomous Items

DIF Category	Criteria
A (negligible)	Mantel Chi-square p -value > 0.05 level or $ SMD/SD \leq 0.17$
B (moderate)	Mantel Chi-square p -value < 0.05 level and $0.17 < SMD/SD \leq 0.25$
C (large)	Mantel Chi-square p -value < 0.05 level and $ SMD/SD > 0.25$

Note: SMD = standardized mean DIF; SD = total group standard deviation of item score

4.7.4. Items Exhibiting Significant DIF

DIF analyses for the gender group were conducted for the CSA field test for each grade level. [Appendix 4.D](#) provides detailed DIF results. [Table 4.D.1](#) shows the distributions of items across the DIF category classifications for each grade level. In addition, “Small N” indicates that the DIF analysis was not performed due to insufficient sample size in [Table 4.D.1](#). [Table 4.D.2](#) lists the items exhibiting significant DIF by the gender groups.

There is one positive C-DIF item that favors the female students in the grade four test and one negative C-DIF item that favors the male students in the grade five test. Formal DIF panel reviews revealed the items did not show any content flaw. Therefore, the C-DIF items were not removed from the item bank.

References

- Dorans, N. J. (2013). ETS contributions to the quantitative assessment of item, test, and score fairness. *ETS Research Report Series*, i–38.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Dorans, N. J., & Schmitt, A. P. (1993). Constructed response and differential item functioning: A pragmatic approach. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 135–65). Hillsdale, NH: Lawrence Erlbaum Associates, Inc.
- Dragow F. (1988). Polychoric and polyserial correlations. In L. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 7, pp. 69–74). New York: Wiley.
- Feldt, L. S., & Brennan, R. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105–146). New York: Macmillan.
- Holland, P. W., & Thayer, D. T. (1988). *An alternate definition of the ETS delta scale of stem difficulty* (Research Report 85–43). Princeton, NJ: Educational Testing Service. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2330-8516.1985.tb00128.x>
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58, 690–700.
- Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analyses of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–48.
- Olsson, U. (1979) Maximum likelihood estimation of the polychoric correlation coefficient *Psychometika* 44, 443–60.
- Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education*, 10(4), 321–44.

Appendix 4.A Test Scores Distribution

Grade levels reflect students' enrolled grade levels during the 2017–18 school year.

Table 4.A.1 Raw Score Frequency Distribution—Grade Three

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
1	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
2	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
3	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
4	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
5	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
6	2	0.21	0.21	2	0.21	0.21	0	0.00	0.00
7	2	0.21	0.43	4	0.43	0.64	0	0.00	0.00
8	3	0.32	0.75	0	0.00	0.64	0	0.00	0.00
9	1	0.11	0.86	1	0.11	0.74	0	0.00	0.00
10	1	0.11	0.96	1	0.11	0.85	0	0.00	0.00
11	2	0.21	1.18	6	0.64	1.49	0	0.00	0.00
12	1	0.11	1.28	3	0.32	1.81	0	0.00	0.00
13	3	0.32	1.61	4	0.43	2.23	0	0.00	0.00
14	7	0.75	2.36	2	0.21	2.44	0	0.00	0.00
15	6	0.64	3.00	5	0.53	2.98	4	3.23	3.23
16	3	0.32	3.32	7	0.74	3.72	1	0.81	4.03
17	7	0.75	4.07	6	0.64	4.36	2	1.61	5.65
18	6	0.64	4.71	16	1.70	6.06	8	6.45	12.10
19	10	1.07	5.78	18	1.91	7.97	6	4.84	16.94
20	17	1.82	7.60	27	2.87	10.84	7	5.65	22.58
21	20	2.14	9.74	24	2.55	13.39	9	7.26	29.84
22	27	2.89	12.63	27	2.87	16.26	6	4.84	34.68
23	38	4.07	16.70	38	4.04	20.30	4	3.23	37.90
24	29	3.10	19.81	36	3.83	24.12	3	2.42	40.32
25	26	2.78	22.59	40	4.25	28.37	9	7.26	47.58
26	20	2.14	24.73	39	4.14	32.52	6	4.84	52.42
27	36	3.85	28.59	28	2.98	35.49	5	4.03	56.45
28	33	3.53	32.12	28	2.98	38.47	3	2.42	58.87

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
29	35	3.75	35.87	35	3.72	42.19	3	2.42	61.29
30	40	4.28	40.15	31	3.29	45.48	5	4.03	65.32
31	41	4.39	44.54	27	2.87	48.35	2	1.61	66.94
32	38	4.07	48.61	19	2.02	50.37	2	1.61	68.55
33	43	4.60	53.21	28	2.98	53.35	5	4.03	72.58
34	30	3.21	56.42	37	3.93	57.28	2	1.61	74.19
35	37	3.96	60.39	27	2.87	60.15	5	4.03	78.23
36	33	3.53	63.92	31	3.29	63.44	4	3.23	81.45
37	31	3.32	67.24	20	2.13	65.57	0	0.00	81.45
38	26	2.78	70.02	34	3.61	69.18	2	1.61	83.06
39	30	3.21	73.23	25	2.66	71.84	2	1.61	84.68
40	27	2.89	76.12	28	2.98	74.81	8	6.45	91.13
41	20	2.14	78.27	29	3.08	77.90	3	2.42	93.55
42	27	2.89	81.16	23	2.44	80.34	0	0.00	93.55
43	22	2.36	83.51	23	2.44	82.78	1	0.81	94.35
44	20	2.14	85.65	18	1.91	84.70	1	0.81	95.16
45	27	2.89	88.54	16	1.70	86.40	3	2.42	97.58
46	21	2.25	90.79	16	1.70	88.10	1	0.81	98.39
47	18	1.93	92.72	21	2.23	90.33	0	0.00	98.39
48	10	1.07	93.79	12	1.28	91.60	0	0.00	98.39
49	10	1.07	94.86	8	0.85	92.45	0	0.00	98.39
50	13	1.39	96.25	17	1.81	94.26	0	0.00	98.39
51	11	1.18	97.43	9	0.96	95.22	1	0.81	99.19
52	8	0.86	98.29	9	0.96	96.17	0	0.00	99.19
53	5	0.54	98.82	8	0.85	97.02	0	0.00	99.19
54	6	0.64	99.46	3	0.32	97.34	0	0.00	99.19
55	2	0.21	99.68	9	0.96	98.30	1	0.81	100.00
56	1	0.11	99.79	6	0.64	98.94	0	0.00	100.00
57	1	0.11	99.89	5	0.53	99.47	0	0.00	100.00
58	0	0.00	99.89	3	0.32	99.79	0	0.00	100.00
59	1	0.11	100.00	0	0.00	99.79	0	0.00	100.00
60	0	0.00	100.00	2	0.21	100.00	0	0.00	100.00
61	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
62	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
63	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
64	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
65	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
66	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
67	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
68	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
69	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
70	0	0.00	100.00	NA	NA	NA	NA	NA	NA

Table 4.A.2 Raw Score Frequency Distribution—Grade Four

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
1	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
2	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
3	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
4	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
5	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
6	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
7	3	0.34	0.34	1	0.11	0.11	0	0.00	0.00
8	1	0.11	0.45	2	0.23	0.34	0	0.00	0.00
9	2	0.22	0.67	4	0.45	0.79	0	0.00	0.00
10	2	0.22	0.89	2	0.23	1.02	0	0.00	0.00
11	4	0.45	1.34	2	0.23	1.25	0	0.00	0.00
12	4	0.45	1.79	2	0.23	1.47	1	0.82	0.82
13	4	0.45	2.24	1	0.11	1.59	0	0.00	0.82
14	1	0.11	2.35	2	0.23	1.81	0	0.00	0.82
15	3	0.34	2.68	0	0.00	1.81	3	2.46	3.28
16	3	0.34	3.02	6	0.68	2.49	0	0.00	3.28
17	6	0.67	3.69	6	0.68	3.17	1	0.82	4.10
18	11	1.23	4.92	3	0.34	3.51	0	0.00	4.10
19	18	2.01	6.94	6	0.68	4.19	1	0.82	4.92
20	9	1.01	7.94	10	1.13	5.32	2	1.64	6.56
21	13	1.45	9.40	15	1.70	7.02	3	2.46	9.02
22	23	2.57	11.97	16	1.81	8.83	1	0.82	9.84
23	23	2.57	14.54	19	2.15	10.99	3	2.46	12.30
24	34	3.80	18.34	33	3.74	14.72	1	0.82	13.11
25	25	2.80	21.14	24	2.72	17.44	5	4.10	17.21
26	40	4.47	25.62	37	4.19	21.63	3	2.46	19.67
27	27	3.02	28.64	30	3.40	25.03	3	2.46	22.13
28	30	3.36	31.99	40	4.53	29.56	1	0.82	22.95
29	38	4.25	36.24	38	4.30	33.86	4	3.28	26.23
30	41	4.59	40.83	40	4.53	38.39	5	4.10	30.33
31	27	3.02	43.85	26	2.94	41.34	2	1.64	31.97

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	32	3.58	47.43	28	3.17	44.51	6	4.92	36.89
33	30	3.36	50.78	28	3.17	47.68	5	4.10	40.98
34	29	3.24	54.03	37	4.19	51.87	2	1.64	42.62
35	33	3.69	57.72	26	2.94	54.81	7	5.74	48.36
36	22	2.46	60.18	33	3.74	58.55	1	0.82	49.18
37	35	3.91	64.09	38	4.30	62.85	7	5.74	54.92
38	21	2.35	66.44	28	3.17	66.02	1	0.82	55.74
39	24	2.68	69.13	19	2.15	68.18	5	4.10	59.84
40	24	2.68	71.81	20	2.27	70.44	7	5.74	65.57
41	30	3.36	75.17	15	1.70	72.14	2	1.64	67.21
42	22	2.46	77.63	21	2.38	74.52	7	5.74	72.95
43	15	1.68	79.31	14	1.59	76.10	4	3.28	76.23
44	16	1.79	81.10	16	1.81	77.92	4	3.28	79.51
45	19	2.13	83.22	14	1.59	79.50	3	2.46	81.97
46	17	1.90	85.12	23	2.60	82.11	4	3.28	85.25
47	14	1.57	86.69	19	2.15	84.26	0	0.00	85.25
48	15	1.68	88.37	10	1.13	85.39	3	2.46	87.70
49	14	1.57	89.93	13	1.47	86.86	1	0.82	88.52
50	14	1.57	91.50	14	1.59	88.45	2	1.64	90.16
51	12	1.34	92.84	11	1.25	89.69	1	0.82	90.98
52	8	0.89	93.74	10	1.13	90.83	2	1.64	92.62
53	8	0.89	94.63	15	1.70	92.53	4	3.28	95.90
54	10	1.12	95.75	8	0.91	93.43	0	0.00	95.90
55	12	1.34	97.09	16	1.81	95.24	0	0.00	95.90
56	8	0.89	97.99	8	0.91	96.15	3	2.46	98.36
57	4	0.45	98.43	6	0.68	96.83	1	0.82	99.18
58	4	0.45	98.88	3	0.34	97.17	1	0.82	100.00
59	3	0.34	99.22	7	0.79	97.96	0	0.00	100.00
60	3	0.34	99.55	6	0.68	98.64	0	0.00	100.00
61	1	0.11	99.66	2	0.23	98.87	0	0.00	100.00
62	1	0.11	99.78	2	0.23	99.09	0	0.00	100.00
63	0	0.00	99.78	5	0.57	99.66	0	0.00	100.00
64	1	0.11	99.89	0	0.00	99.66	0	0.00	100.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	99.89	1	0.11	99.77	0	0.00	100.00
66	0	0.00	99.89	0	0.00	99.77	0	0.00	100.00
67	1	0.11	100.00	1	0.11	99.89	0	0.00	100.00
68	0	0.00	100.00	0	0.00	99.89	0	0.00	100.00
69	0	0.00	100.00	0	0.00	99.89	0	0.00	100.00
70	0	0.00	100.00	0	0.00	99.89	0	0.00	100.00
71	0	0.00	100.00	1	0.11	100.00	NA	NA	NA
72	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
73	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
74	NA	NA	NA	0	0.00	100.00	NA	NA	NA
75	NA	NA	NA	0	0.00	100.00	NA	NA	NA
76	NA	NA	NA	0	0.00	100.00	NA	NA	NA

Table 4.A.3 Raw Score Frequency Distribution—Grade Five

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
1	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
2	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
3	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
4	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
5	0	0.00	0.00	1	0.15	0.15	0	0.00	0.00
6	1	0.15	0.15	0	0.00	0.15	0	0.00	0.00
7	0	0.00	0.15	0	0.00	0.15	0	0.00	0.00
8	1	0.15	0.31	1	0.15	0.31	0	0.00	0.00
9	1	0.15	0.46	2	0.31	0.61	0	0.00	0.00
10	4	0.62	1.08	3	0.46	1.07	0	0.00	0.00
11	2	0.31	1.39	1	0.15	1.23	0	0.00	0.00
12	3	0.46	1.85	3	0.46	1.69	0	0.00	0.00
13	0	0.00	1.85	2	0.31	1.99	0	0.00	0.00
14	2	0.31	2.16	5	0.77	2.76	0	0.00	0.00
15	8	1.23	3.40	3	0.46	3.22	0	0.00	0.00
16	2	0.31	3.70	2	0.31	3.53	0	0.00	0.00
17	10	1.54	5.25	4	0.61	4.14	0	0.00	0.00
18	7	1.08	6.33	7	1.07	5.21	0	0.00	0.00
19	9	1.39	7.72	5	0.77	5.98	0	0.00	0.00
20	20	3.09	10.80	16	2.45	8.44	1	2.56	2.56
21	14	2.16	12.96	16	2.45	10.89	1	2.56	5.13
22	15	2.31	15.28	11	1.69	12.58	1	2.56	7.69
23	18	2.78	18.06	25	3.83	16.41	2	5.13	12.82
24	24	3.70	21.76	20	3.07	19.48	0	0.00	12.82
25	35	5.40	27.16	21	3.22	22.70	2	5.13	17.95
26	27	4.17	31.33	21	3.22	25.92	1	2.56	20.51
27	23	3.55	34.88	24	3.68	29.60	1	2.56	23.08
28	29	4.48	39.35	25	3.83	33.44	1	2.56	25.64
29	21	3.24	42.59	21	3.22	36.66	3	7.69	33.33
30	35	5.40	47.99	28	4.29	40.95	1	2.56	35.90
31	24	3.70	51.70	25	3.83	44.79	5	12.82	48.72

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	25	3.86	55.56	30	4.60	49.39	3	7.69	56.41
33	24	3.70	59.26	19	2.91	52.30	1	2.56	58.97
34	25	3.86	63.12	23	3.53	55.83	1	2.56	61.54
35	20	3.09	66.20	25	3.83	59.66	3	7.69	69.23
36	27	4.17	70.37	27	4.14	63.80	1	2.56	71.79
37	20	3.09	73.46	17	2.61	66.41	2	5.13	76.92
38	14	2.16	75.62	29	4.45	70.86	0	0.00	76.92
39	16	2.47	78.09	25	3.83	74.69	1	2.56	79.49
40	21	3.24	81.33	29	4.45	79.14	2	5.13	84.62
41	8	1.23	82.56	17	2.61	81.75	0	0.00	84.62
42	17	2.62	85.19	17	2.61	84.36	2	5.13	89.74
43	11	1.70	86.88	15	2.30	86.66	1	2.56	92.31
44	11	1.70	88.58	18	2.76	89.42	0	0.00	92.31
45	16	2.47	91.05	13	1.99	91.41	1	2.56	94.87
46	7	1.08	92.13	10	1.53	92.94	0	0.00	94.87
47	12	1.85	93.98	9	1.38	94.33	1	2.56	97.44
48	4	0.62	94.60	7	1.07	95.40	0	0.00	97.44
49	8	1.23	95.83	7	1.07	96.47	1	2.56	100.00
50	8	1.23	97.07	5	0.77	97.24	0	0.00	100.00
51	5	0.77	97.84	6	0.92	98.16	0	0.00	100.00
52	4	0.62	98.46	3	0.46	98.62	0	0.00	100.00
53	4	0.62	99.07	3	0.46	99.08	0	0.00	100.00
54	3	0.46	99.54	1	0.15	99.23	0	0.00	100.00
55	0	0.00	99.54	3	0.46	99.69	0	0.00	100.00
56	0	0.00	99.54	2	0.31	100.00	0	0.00	100.00
57	1	0.15	99.69	0	0.00	100.00	0	0.00	100.00
58	0	0.00	99.69	0	0.00	100.00	0	0.00	100.00
59	1	0.15	99.85	0	0.00	100.00	0	0.00	100.00
60	0	0.00	99.85	0	0.00	100.00	0	0.00	100.00
61	0	0.00	99.85	0	0.00	100.00	0	0.00	100.00
62	0	0.00	99.85	0	0.00	100.00	0	0.00	100.00
63	0	0.00	99.85	0	0.00	100.00	0	0.00	100.00
64	1	0.15	100.00	0	0.00	100.00	0	0.00	100.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
66	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
67	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
68	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
69	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
70	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
71	0	0.00	100.00	NA	NA	NA	0	0.00	100.00

Table 4.A.4 Raw Score Frequency Distribution—Grade Six

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
1	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
2	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
3	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
4	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
5	0	0.00	0.00	1	0.21	0.21	0	0.00	0.00
6	1	0.22	0.22	0	0.00	0.21	0	0.00	0.00
7	0	0.00	0.22	0	0.00	0.21	0	0.00	0.00
8	2	0.44	0.66	1	0.21	0.43	0	0.00	0.00
9	2	0.44	1.09	1	0.21	0.64	0	0.00	0.00
10	0	0.00	1.09	0	0.00	0.64	0	0.00	0.00
11	0	0.00	1.09	2	0.43	1.07	0	0.00	0.00
12	0	0.00	1.09	2	0.43	1.50	0	0.00	0.00
13	0	0.00	1.09	1	0.21	1.71	0	0.00	0.00
14	1	0.22	1.31	0	0.00	1.71	0	0.00	0.00
15	1	0.22	1.53	3	0.64	2.36	0	0.00	0.00
16	2	0.44	1.97	0	0.00	2.36	0	0.00	0.00
17	1	0.22	2.18	3	0.64	3.00	0	0.00	0.00
18	6	1.31	3.49	1	0.21	3.21	0	0.00	0.00
19	3	0.66	4.15	2	0.43	3.64	2	5.00	5.00
20	3	0.66	4.80	1	0.21	3.85	1	2.50	7.50
21	3	0.66	5.46	6	1.28	5.14	1	2.50	10.00
22	3	0.66	6.11	6	1.28	6.42	0	0.00	10.00
23	10	2.18	8.30	5	1.07	7.49	2	5.00	15.00
24	17	3.71	12.01	11	2.36	9.85	0	0.00	15.00
25	11	2.40	14.41	12	2.57	12.42	2	5.00	20.00
26	13	2.84	17.25	12	2.57	14.99	0	0.00	20.00
27	13	2.84	20.09	13	2.78	17.77	3	7.50	27.50
28	15	3.28	23.36	12	2.57	20.34	0	0.00	27.50
29	15	3.28	26.64	12	2.57	22.91	1	2.50	30.00
30	15	3.28	29.91	13	2.78	25.70	0	0.00	30.00
31	15	3.28	33.19	13	2.78	28.48	0	0.00	30.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	20	4.37	37.55	16	3.43	31.91	1	2.50	32.50
33	13	2.84	40.39	15	3.21	35.12	3	7.50	40.00
34	16	3.49	43.89	12	2.57	37.69	0	0.00	40.00
35	10	2.18	46.07	13	2.78	40.47	1	2.50	42.50
36	11	2.40	48.47	15	3.21	43.68	1	2.50	45.00
37	13	2.84	51.31	13	2.78	46.47	1	2.50	47.50
38	16	3.49	54.80	11	2.36	48.82	1	2.50	50.00
39	16	3.49	58.30	6	1.28	50.11	3	7.50	57.50
40	8	1.75	60.04	16	3.43	53.53	1	2.50	60.00
41	18	3.93	63.97	16	3.43	56.96	2	5.00	65.00
42	14	3.06	67.03	12	2.57	59.53	1	2.50	67.50
43	14	3.06	70.09	18	3.85	63.38	1	2.50	70.00
44	10	2.18	72.27	18	3.85	67.24	2	5.00	75.00
45	13	2.84	75.11	15	3.21	70.45	2	5.00	80.00
46	12	2.62	77.73	16	3.43	73.88	2	5.00	85.00
47	10	2.18	79.91	13	2.78	76.66	0	0.00	85.00
48	14	3.06	82.97	5	1.07	77.73	0	0.00	85.00
49	8	1.75	84.72	16	3.43	81.16	3	7.50	92.50
50	6	1.31	86.03	12	2.57	83.73	2	5.00	97.50
51	6	1.31	87.34	14	3.00	86.72	0	0.00	97.50
52	7	1.53	88.86	12	2.57	89.29	0	0.00	97.50
53	10	2.18	91.05	13	2.78	92.08	0	0.00	97.50
54	11	2.40	93.45	6	1.28	93.36	0	0.00	97.50
55	10	2.18	95.63	8	1.71	95.07	1	2.50	100.00
56	7	1.53	97.16	4	0.86	95.93	0	0.00	100.00
57	4	0.87	98.03	9	1.93	97.86	0	0.00	100.00
58	2	0.44	98.47	6	1.28	99.14	0	0.00	100.00
59	2	0.44	98.91	2	0.43	99.57	0	0.00	100.00
60	3	0.66	99.56	1	0.21	99.79	0	0.00	100.00
61	1	0.22	99.78	0	0.00	99.79	0	0.00	100.00
62	1	0.22	100.00	0	0.00	99.79	0	0.00	100.00
63	0	0.00	100.00	0	0.00	99.79	0	0.00	100.00
64	0	0.00	100.00	1	0.21	100.00	0	0.00	100.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
66	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
67	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
68	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
69	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
70	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
71	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
72	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
73	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
74	0	0.00	100.00	0	0.00	100.00	NA	NA	NA

Table 4.A.5 Raw Score Frequency Distribution—Grade Seven

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
1	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
2	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
3	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
4	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
5	0	0.00	0.00	1	0.23	0.23	0	0.00	0.00
6	2	0.48	0.48	1	0.23	0.46	0	0.00	0.00
7	0	0.00	0.48	5	1.15	1.61	0	0.00	0.00
8	1	0.24	0.71	1	0.23	1.83	0	0.00	0.00
9	3	0.71	1.43	0	0.00	1.83	0	0.00	0.00
10	2	0.48	1.90	1	0.23	2.06	0	0.00	0.00
11	2	0.48	2.38	3	0.69	2.75	0	0.00	0.00
12	4	0.95	3.33	1	0.23	2.98	0	0.00	0.00
13	2	0.48	3.80	1	0.23	3.21	0	0.00	0.00
14	5	1.19	4.99	1	0.23	3.44	0	0.00	0.00
15	1	0.24	5.23	3	0.69	4.13	0	0.00	0.00
16	3	0.71	5.94	2	0.46	4.59	0	0.00	0.00
17	1	0.24	6.18	4	0.92	5.50	0	0.00	0.00
18	8	1.90	8.08	8	1.83	7.34	0	0.00	0.00
19	12	2.85	10.93	5	1.15	8.49	1	2.86	2.86
20	13	3.09	14.01	16	3.67	12.16	0	0.00	2.86
21	17	4.04	18.05	9	2.06	14.22	0	0.00	2.86
22	22	5.23	23.28	15	3.44	17.66	3	8.57	11.43
23	15	3.56	26.84	11	2.52	20.18	0	0.00	11.43
24	19	4.51	31.35	18	4.13	24.31	1	2.86	14.29
25	23	5.46	36.82	17	3.90	28.21	3	8.57	22.86
26	16	3.80	40.62	16	3.67	31.88	1	2.86	25.71
27	10	2.38	42.99	15	3.44	35.32	4	11.43	37.14
28	12	2.85	45.84	19	4.36	39.68	3	8.57	45.71
29	16	3.80	49.64	18	4.13	43.81	1	2.86	48.57
30	20	4.75	54.39	25	5.73	49.54	2	5.71	54.29
31	12	2.85	57.24	22	5.05	54.59	2	5.71	60.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	14	3.33	60.57	22	5.05	59.63	1	2.86	62.86
33	8	1.90	62.47	10	2.29	61.93	3	8.57	71.43
34	9	2.14	64.61	19	4.36	66.28	3	8.57	80.00
35	10	2.38	66.98	16	3.67	69.95	1	2.86	82.86
36	10	2.38	69.36	10	2.29	72.25	1	2.86	85.71
37	11	2.61	71.97	9	2.06	74.31	0	0.00	85.71
38	17	4.04	76.01	10	2.29	76.61	1	2.86	88.57
39	13	3.09	79.10	13	2.98	79.59	0	0.00	88.57
40	8	1.90	81.00	12	2.75	82.34	1	2.86	91.43
41	11	2.61	83.61	8	1.83	84.17	0	0.00	91.43
42	6	1.43	85.04	11	2.52	86.70	0	0.00	91.43
43	13	3.09	88.12	11	2.52	89.22	1	2.86	94.29
44	6	1.43	89.55	10	2.29	91.51	0	0.00	94.29
45	7	1.66	91.21	5	1.15	92.66	0	0.00	94.29
46	11	2.61	93.82	4	0.92	93.58	2	5.71	100.00
47	4	0.95	94.77	7	1.61	95.18	0	0.00	100.00
48	6	1.43	96.20	6	1.38	96.56	0	0.00	100.00
49	2	0.48	96.67	6	1.38	97.94	0	0.00	100.00
50	2	0.48	97.15	3	0.69	98.62	0	0.00	100.00
51	5	1.19	98.34	2	0.46	99.08	0	0.00	100.00
52	1	0.24	98.57	1	0.23	99.31	0	0.00	100.00
53	0	0.00	98.57	3	0.69	100.00	0	0.00	100.00
54	3	0.71	99.29	0	0.00	100.00	0	0.00	100.00
55	3	0.71	100.00	0	0.00	100.00	0	0.00	100.00
56	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
57	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
58	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
59	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
60	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
61	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
62	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
63	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
64	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
66	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
67	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
68	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
69	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
70	0	0.00	100.00	0	0.00	100.00	0	0.00	100.00
71	0	0.00	100.00	0	0.00	100.00	NA	NA	NA

Table 4.A.6 Raw Score Frequency Distribution—Grade Eight

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
1	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
2	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
3	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
4	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
5	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
6	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
7	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
8	0	0.00	0.00	1	0.71	0.71	NA	NA	NA
9	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
10	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
11	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
12	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
13	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
14	0	0.00	0.00	0	0.00	0.71	NA	NA	NA
15	0	0.00	0.00	1	0.71	1.42	NA	NA	NA
16	1	0.72	0.72	0	0.00	1.42	NA	NA	NA
17	0	0.00	0.72	0	0.00	1.42	NA	NA	NA
18	2	1.45	2.17	1	0.71	2.13	NA	NA	NA
19	0	0.00	2.17	1	0.71	2.84	NA	NA	NA
20	1	0.72	2.90	0	0.00	2.84	NA	NA	NA
21	3	2.17	5.07	0	0.00	2.84	NA	NA	NA
22	4	2.90	7.97	2	1.42	4.26	NA	NA	NA
23	3	2.17	10.14	2	1.42	5.67	NA	NA	NA
24	7	5.07	15.22	5	3.55	9.22	NA	NA	NA
25	7	5.07	20.29	6	4.26	13.48	NA	NA	NA
26	5	3.62	23.91	9	6.38	19.86	NA	NA	NA
27	3	2.17	26.09	6	4.26	24.11	NA	NA	NA
28	7	5.07	31.16	5	3.55	27.66	NA	NA	NA
29	4	2.90	34.06	6	4.26	31.91	NA	NA	NA
30	5	3.62	37.68	5	3.55	35.46	NA	NA	NA
31	4	2.90	40.58	5	3.55	39.01	NA	NA	NA

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	11	7.97	48.55	9	6.38	45.39	NA	NA	NA
33	4	2.90	51.45	11	7.80	53.19	NA	NA	NA
34	9	6.52	57.97	2	1.42	54.61	NA	NA	NA
35	9	6.52	64.49	8	5.67	60.28	NA	NA	NA
36	5	3.62	68.12	6	4.26	64.54	NA	NA	NA
37	3	2.17	70.29	8	5.67	70.21	NA	NA	NA
38	7	5.07	75.36	4	2.84	73.05	NA	NA	NA
39	7	5.07	80.43	2	1.42	74.47	NA	NA	NA
40	5	3.62	84.06	1	0.71	75.18	NA	NA	NA
41	1	0.72	84.78	0	0.00	75.18	NA	NA	NA
42	2	1.45	86.23	4	2.84	78.01	NA	NA	NA
43	2	1.45	87.68	5	3.55	81.56	NA	NA	NA
44	1	0.72	88.41	4	2.84	84.40	NA	NA	NA
45	1	0.72	89.13	2	1.42	85.82	NA	NA	NA
46	4	2.90	92.03	1	0.71	86.52	NA	NA	NA
47	3	2.17	94.20	2	1.42	87.94	NA	NA	NA
48	2	1.45	95.65	3	2.13	90.07	NA	NA	NA
49	2	1.45	97.10	1	0.71	90.78	NA	NA	NA
50	2	1.45	98.55	4	2.84	93.62	NA	NA	NA
51	1	0.72	99.28	4	2.84	96.45	NA	NA	NA
52	0	0.00	99.28	0	0.00	96.45	NA	NA	NA
53	0	0.00	99.28	1	0.71	97.16	NA	NA	NA
54	0	0.00	99.28	1	0.71	97.87	NA	NA	NA
55	0	0.00	99.28	2	1.42	99.29	NA	NA	NA
56	0	0.00	99.28	1	0.71	100.00	NA	NA	NA
57	0	0.00	99.28	0	0.00	100.00	NA	NA	NA
58	0	0.00	99.28	0	0.00	100.00	NA	NA	NA
59	0	0.00	99.28	0	0.00	100.00	NA	NA	NA
60	1	0.72	100.00	0	0.00	100.00	NA	NA	NA
61	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
62	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
63	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
64	0	0.00	100.00	0	0.00	100.00	NA	NA	NA

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
66	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
67	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
68	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
69	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
70	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
71	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
72	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
73	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
74	NA	NA	NA	0	0.00	100.00	NA	NA	NA

Table 4.A.7 Raw Score Frequency Distribution—High School

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
0	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
1	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
2	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
3	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
4	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
5	0	0.00	0.00	0	0.00	0.00	NA	NA	NA
6	0	0.00	0.00	1	0.31	0.31	NA	NA	NA
7	1	0.31	0.31	0	0.00	0.31	NA	NA	NA
8	2	0.61	0.92	0	0.00	0.31	NA	NA	NA
9	1	0.31	1.22	0	0.00	0.31	NA	NA	NA
10	0	0.00	1.22	0	0.00	0.31	NA	NA	NA
11	0	0.00	1.22	1	0.31	0.62	NA	NA	NA
12	2	0.61	1.83	2	0.62	1.24	NA	NA	NA
13	0	0.00	1.83	1	0.31	1.55	NA	NA	NA
14	0	0.00	1.83	1	0.31	1.86	NA	NA	NA
15	0	0.00	1.83	0	0.00	1.86	NA	NA	NA
16	0	0.00	1.83	3	0.93	2.80	NA	NA	NA
17	0	0.00	1.83	3	0.93	3.73	NA	NA	NA
18	1	0.31	2.14	1	0.31	4.04	NA	NA	NA
19	5	1.53	3.67	4	1.24	5.28	NA	NA	NA
20	2	0.61	4.28	7	2.17	7.45	NA	NA	NA
21	6	1.83	6.12	5	1.55	9.01	NA	NA	NA
22	10	3.06	9.17	6	1.86	10.87	NA	NA	NA
23	8	2.45	11.62	5	1.55	12.42	NA	NA	NA
24	3	0.92	12.54	10	3.11	15.53	NA	NA	NA
25	13	3.98	16.51	7	2.17	17.70	NA	NA	NA
26	13	3.98	20.49	11	3.42	21.12	NA	NA	NA
27	11	3.36	23.85	13	4.04	25.16	NA	NA	NA
28	9	2.75	26.61	15	4.66	29.81	NA	NA	NA
29	15	4.59	31.19	13	4.04	33.85	NA	NA	NA
30	17	5.20	36.39	11	3.42	37.27	NA	NA	NA
31	14	4.28	40.67	13	4.04	41.30	NA	NA	NA

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
32	16	4.89	45.57	11	3.42	44.72	NA	NA	NA
33	9	2.75	48.32	12	3.73	48.45	NA	NA	NA
34	23	7.03	55.35	11	3.42	51.86	NA	NA	NA
35	14	4.28	59.63	18	5.59	57.45	NA	NA	NA
36	19	5.81	65.44	13	4.04	61.49	NA	NA	NA
37	17	5.20	70.64	14	4.35	65.84	NA	NA	NA
38	14	4.28	74.92	13	4.04	69.88	NA	NA	NA
39	15	4.59	79.51	6	1.86	71.74	NA	NA	NA
40	13	3.98	83.49	17	5.28	77.02	NA	NA	NA
41	7	2.14	85.63	12	3.73	80.75	NA	NA	NA
42	7	2.14	87.77	7	2.17	82.92	NA	NA	NA
43	3	0.92	88.69	9	2.80	85.71	NA	NA	NA
44	6	1.83	90.52	9	2.80	88.51	NA	NA	NA
45	10	3.06	93.58	7	2.17	90.68	NA	NA	NA
46	3	0.92	94.50	7	2.17	92.86	NA	NA	NA
47	2	0.61	95.11	6	1.86	94.72	NA	NA	NA
48	2	0.61	95.72	3	0.93	95.65	NA	NA	NA
49	3	0.92	96.64	1	0.31	95.96	NA	NA	NA
50	5	1.53	98.17	4	1.24	97.20	NA	NA	NA
51	1	0.31	98.47	3	0.93	98.14	NA	NA	NA
52	1	0.31	98.78	3	0.93	99.07	NA	NA	NA
53	0	0.00	98.78	0	0.00	99.07	NA	NA	NA
54	1	0.31	99.08	2	0.62	99.69	NA	NA	NA
55	1	0.31	99.39	0	0.00	99.69	NA	NA	NA
56	0	0.00	99.39	0	0.00	99.69	NA	NA	NA
57	1	0.31	99.69	1	0.31	100.00	NA	NA	NA
58	0	0.00	99.69	0	0.00	100.00	NA	NA	NA
59	1	0.31	100.00	0	0.00	100.00	NA	NA	NA
60	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
61	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
62	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
63	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
64	0	0.00	100.00	0	0.00	100.00	NA	NA	NA

Raw Score	Form 1 Frequency	Form 1 Percent	Form 1 Cumulative Percent	Form 2 Frequency	Form 2 Percent	Form 2 Cumulative Percent	Accommodated Form Frequency	Accommodated Form Percent	Accommodated Form Cumulative Percent
65	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
66	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
67	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
68	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
69	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
70	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
71	0	0.00	100.00	0	0.00	100.00	NA	NA	NA
72	0	0.00	100.00	NA	NA	NA	NA	NA	NA
73	0	0.00	100.00	NA	NA	NA	NA	NA	NA

Appendix 4.B Classical Item Analyses

Note the following about [Table 4.B.1](#) through [Table 4.B.28](#):

- An “r” indicates polyserial correlation, which is a statistical index of the item-total correlation.
- A hyphen (“-”) indicates a polytomous item.
- NA indicates there were no students taking the test form so no data is available.
- Grade levels reflect students’ enrolled grade levels during the 2017–18 school year.

Table 4.B.1 Dichotomous Item Statistics for Form 1—Grade Three

Item Sequence	<i>p</i> -value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.91	0.25	0.00	0.00
2	0.60	0.45	0.00	0.00
3	0.64	0.28	0.00	0.00
4	-	-	-	-
5	0.60	0.55	0.00	0.00
6	0.35	0.21	0.00	0.00
7	0.74	0.23	0.00	0.00
8	0.45	0.43	0.00	0.00
9	0.46	-0.04	0.00	0.00
10	-	-	-	-
11	0.42	0.29	0.15	0.00
12	0.52	0.19	0.09	0.00
13	0.53	0.22	0.20	0.00
14	0.47	0.31	0.37	0.00
15	0.47	0.30	0.38	0.00
16	0.21	-0.06	0.38	0.00
17	0.37	0.20	0.38	0.00
18	-	-	-	-
19	0.48	0.36	0.54	0.43
20	0.38	0.24	0.75	0.43
21	0.46	0.60	0.86	0.43
22	0.32	0.19	0.97	0.43
23	0.22	0.10	0.97	0.43
24	0.29	0.35	1.29	0.43
25	0.64	0.15	1.40	0.43
26	-	-	-	-
27	0.62	0.57	0.77	1.24
28	-	-	-	-
29	0.30	0.24	0.86	1.24

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
30	0.28	0.09	0.33	2.30
31	0.47	0.55	0.55	2.75
32	0.46	0.47	0.93	2.63
33	-	-	-	-
34	-	-	-	-
35	0.62	0.63	0.99	2.75
36	0.32	0.51	0.97	2.42
37	0.25	0.25	0.99	2.75
38	0.66	0.80	0.97	2.42
39	0.53	0.65	0.97	2.42
40	-	-	-	-
41	0.80	0.59	1.11	3.78
42	0.35	0.51	1.22	3.78
43	0.22	0.48	1.22	3.78
44	0.46	0.05	1.22	3.78
45	0.42	0.14	1.22	3.78
46	-	-	-	-
47	0.45	0.45	1.22	3.78
48	0.15	0.28	1.22	3.78
49	0.51	0.58	0.45	5.06
50	0.29	0.34	0.90	5.06
51	0.42	0.23	0.99	4.55
52	0.61	0.46	0.99	4.55
53	-	-	-	-
54	0.28	0.21	1.12	5.06
55	0.41	0.44	1.12	5.06
56	0.20	0.41	1.09	4.55
57	-	-	-	-
58	0.26	0.50	1.35	5.06
59	0.43	0.52	0.57	6.38

Table 4.B.2 Dichotomous Item Statistics for Form 2—Grade Three

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.52	0.43	0.00	0.00
2	0.60	0.45	0.00	0.00
3	-	-	-	-
4	0.52	0.44	0.00	0.00
5	-	-	-	-
6	0.57	0.39	0.00	0.00
7	0.68	0.48	0.00	0.00
8	0.42	0.06	0.00	0.00
9	0.43	0.45	0.00	0.00
10	0.46	-0.04	0.00	0.00
11	-	-	-	-
12	0.42	0.29	0.15	0.00
13	0.53	0.22	0.20	0.00
14	0.47	0.31	0.37	0.00
15	0.41	0.57	0.47	0.00
16	0.39	0.43	0.47	0.00
17	0.33	0.41	0.47	0.00
18	0.52	0.41	0.21	0.53
19	0.65	0.40	0.21	0.53
20	0.25	0.39	0.21	0.53
21	0.25	0.19	0.32	0.53
22	0.54	0.56	0.43	0.53
23	0.54	0.48	0.53	0.53
24	0.57	0.45	0.53	0.53
25	-	-	-	-
26	0.36	0.29	0.19	0.95
27	-	-	-	-
28	0.76	0.52	0.54	1.07
29	0.46	0.54	0.66	0.95
30	-	-	-	-
31	0.47	0.52	0.48	2.21
32	0.64	0.55	0.67	2.21
33	0.49	0.43	0.76	2.51
34	0.51	0.43	0.86	2.21
35	0.41	0.23	0.98	2.51
36	0.46	0.47	0.93	2.63
37	-	-	-	-
38	0.48	0.53	1.20	2.51

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.66	0.54	1.06	2.21
40	-	-	-	-
41	0.57	0.60	0.55	3.52
42	0.56	0.43	0.66	3.52
43	0.42	0.56	0.77	3.52
44	0.34	0.49	0.88	3.52
45	0.28	-0.09	1.10	3.52
46	0.20	0.22	1.10	3.52
47	0.40	0.08	1.10	3.52
48	-	-	-	-
49	0.53	0.61	0.89	4.67
50	-	-	-	-
51	0.46	0.56	1.08	4.21
52	0.50	0.40	1.33	4.67
53	0.40	0.44	1.56	4.67
54	0.36	0.33	1.78	4.67
55	0.48	0.39	1.66	4.21
56	0.29	0.02	1.76	4.21
57	0.26	0.32	2.11	4.67
58	0.36	0.52	1.76	4.21
59	0.52	0.39	0.23	6.81

Table 4.B.3 Dichotomous Item Statistics for the Accommodated Form—Grade Three

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.52	0.43	0.00	0.00
2	0.60	0.45	0.00	0.00
3	-	-	-	-
4	0.52	0.44	0.00	0.00
5	-	-	-	-
6	0.60	0.55	0.00	0.00
7	0.35	0.21	0.00	0.00
8	0.57	0.39	0.00	0.00
9	0.42	0.06	0.00	0.00
10	0.43	0.45	0.00	0.00
11	0.46	-0.04	0.00	0.00
12	-	-	-	-
13	0.42	0.29	0.15	0.00
14	0.52	0.19	0.09	0.00
15	0.53	0.22	0.20	0.00
16	0.47	0.30	0.38	0.00
17	0.41	0.57	0.47	0.00
18	0.21	-0.06	0.38	0.00
19	0.39	0.43	0.47	0.00
20	0.37	0.20	0.38	0.00
21	0.33	0.41	0.47	0.00
22	0.19	0.39	0.00	0.00
23	0.25	0.31	0.00	0.00
24	0.54	0.49	0.00	0.00
25	0.31	0.44	0.00	0.00
26	0.28	0.43	0.00	0.00
27	0.34	-0.03	0.00	0.00
28	0.20	0.53	0.00	0.00
29	0.36	0.29	0.19	0.95
30	-	-	-	-
31	0.62	0.57	0.77	1.24
32	0.46	0.54	0.66	0.95
33	0.30	0.24	0.86	1.24
34	-	-	-	-
35	0.47	0.52	0.48	2.21
36	-	-	-	-
37	0.64	0.55	0.67	2.21
38	0.51	0.43	0.86	2.21

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.32	0.51	0.97	2.42
40	0.66	0.80	0.97	2.42
41	0.53	0.65	0.97	2.42
42	0.66	0.54	1.06	2.21
43	0.40	0.52	0.00	0.81
44	0.32	0.68	0.00	0.81
45	0.20	0.27	0.00	0.81
46	0.40	-0.19	0.00	0.81
47	0.28	-0.12	0.00	0.81
48	0.28	0.07	0.00	0.81
49	0.37	0.17	0.00	0.81
50	0.42	0.23	0.99	4.55
51	0.46	0.56	1.08	4.21
52	0.61	0.46	0.99	4.55
53	-	-	-	-
54	0.48	0.39	1.66	4.21
55	0.33	0.40	0.00	0.81
56	0.29	0.02	1.76	4.21
57	0.20	0.41	1.09	4.55
58	-	-	-	-
59	0.36	0.52	1.76	4.21

Table 4.B.4 Dichotomous Item Statistics for Form 1—Grade Four

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.69	0.29	0.00	0.00
2	0.61	0.63	0.00	0.00
3	0.18	0.23	0.00	0.00
4	-	-	-	-
5	-	-	-	-
6	0.44	0.36	0.00	0.00
7	-	-	-	-
8	0.63	0.30	0.00	0.00
9	0.60	0.47	0.00	0.00
10	0.42	0.47	0.00	0.00
11	0.74	0.62	0.00	0.00
12	0.38	0.45	0.22	0.00
13	0.29	0.27	0.22	0.00
14	0.31	0.36	0.22	0.00
15	0.41	0.29	0.73	0.17
16	-	-	-	-
17	-	-	-	-
18	-	-	-	-
19	0.47	0.40	1.12	0.22
20	0.43	0.33	1.23	0.22
21	0.29	0.14	1.30	0.17
22	0.34	0.41	1.68	0.22
23	0.45	0.37	1.63	0.17
24	0.15	0.06	2.02	0.22
25	0.48	0.44	0.80	1.94
26	0.52	0.35	0.80	1.94
27	0.36	0.33	1.25	1.94
28	0.75	0.52	1.37	1.94
29	-	-	-	-
30	0.44	0.43	1.71	1.94
31	0.43	0.23	1.94	1.94
32	0.24	0.18	2.05	1.94
33	-	-	-	-
34	-	-	-	-
35	0.24	0.60	0.59	4.04
36	0.46	0.56	0.93	4.32
37	0.40	0.53	1.05	4.32
38	0.64	0.60	1.28	4.32

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.55	0.52	1.28	4.32
40	0.44	0.41	1.28	4.32
41	0.20	0.52	1.00	4.04
42	0.42	0.43	0.12	5.67
43	0.52	0.62	0.21	5.07
44	-	-	-	-
45	0.65	0.50	0.59	5.67
46	0.30	0.05	0.62	5.07
47	0.73	0.71	0.83	5.07
48	-	-	-	-
49	0.46	0.50	0.52	5.72
50	0.29	0.41	1.11	5.62
51	0.24	0.16	1.07	6.43
52	0.37	0.66	1.04	5.72
53	0.56	0.33	1.31	6.43
54	-	-	-	-
55	0.41	0.50	1.43	5.96
56	-	-	-	-
57	0.52	0.49	1.43	6.43
58	-	-	-	-
59	0.41	0.00	0.36	7.97

Table 4.B.5 Dichotomous Item Statistics for Form 2—Grade Four

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.69	0.29	0.00	0.00
2	0.72	0.47	0.00	0.00
3	-	-	-	-
4	0.40	0.41	0.00	0.00
5	0.46	0.43	0.00	0.00
6	0.36	0.36	0.00	0.00
7	-	-	-	-
8	0.55	0.35	0.00	0.00
9	0.54	0.25	0.00	0.00
10	-	-	-	-
11	0.27	0.01	0.11	0.00
12	0.55	0.30	0.11	0.00
13	-	-	-	-
14	-	-	-	-
15	0.41	0.29	0.73	0.17
16	0.29	0.14	1.30	0.17
17	0.31	0.29	0.91	0.11
18	-	-	-	-
19	0.28	0.44	1.36	0.11
20	-	-	-	-
21	0.45	0.37	1.63	0.17
22	0.18	0.18	1.36	0.11
23	0.32	-0.05	1.36	0.11
24	0.29	0.11	1.36	0.11
25	-	-	-	-
26	0.77	0.19	0.92	1.61
27	0.60	0.63	1.04	1.61
28	0.30	0.59	1.15	1.61
29	0.29	-0.07	1.27	1.61
30	0.24	0.35	1.61	1.61
31	0.42	0.36	1.73	1.61
32	-	-	-	-
33	-	-	-	-
34	0.52	0.68	0.24	3.76
35	-	-	-	-
36	0.48	0.45	0.24	3.76
37	0.60	0.68	0.24	3.76
38	0.24	0.60	0.59	4.04

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.53	0.47	0.35	3.76
40	0.65	0.65	0.35	3.76
41	0.20	0.52	1.00	4.04
42	0.53	0.33	0.59	4.13
43	-	-	-	-
44	-	-	-	-
45	0.73	0.69	1.14	3.72
46	0.38	0.33	1.30	4.13
47	0.49	0.57	1.24	3.72
48	-	-	-	-
49	-	-	-	-
50	-	-	-	-
51	0.29	0.41	1.11	5.62
52	0.16	0.47	1.31	5.50
53	0.44	0.73	1.36	4.91
54	0.41	0.50	1.43	5.96
55	0.38	0.51	1.57	4.91
56	-	-	-	-
57	0.29	0.24	2.09	4.91
58	0.37	0.42	2.09	4.91
59	0.51	0.41	0.43	6.91

Table 4.B.6 Dichotomous Item Statistics for the Accommodated Form—Grade Four

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.69	0.29	0.00	0.00
2	-	-	-	-
3	0.61	0.63	0.00	0.00
4	0.72	0.47	0.00	0.00
5	0.40	0.41	0.00	0.00
6	0.18	0.23	0.00	0.00
7	0.46	0.49	0.00	0.00
8	-	-	-	-
9	0.52	0.46	0.00	0.00
10	0.66	0.63	0.00	0.00
11	0.46	0.44	0.00	0.00
12	0.56	0.45	0.00	0.00
13	0.49	0.12	0.00	0.00
14	0.40	0.47	0.00	0.00
15	0.66	0.29	0.00	0.00
16	0.34	0.23	0.00	0.00
17	0.34	-0.03	0.00	0.00
18	-	-	-	-
19	-	-	-	-
20	-	-	-	-
21	0.56	0.21	0.00	0.00
22	0.31	0.46	0.00	0.00
23	0.35	0.06	0.00	0.00
24	0.17	0.36	0.00	0.00
25	0.36	0.17	0.00	0.00
26	0.38	0.13	0.00	0.00
27	0.15	0.01	0.00	0.00
28	0.61	0.64	0.82	0.00
29	0.64	0.68	0.82	0.00
30	0.43	0.58	0.82	0.00
31	0.43	0.53	0.82	0.00
32	0.48	0.46	0.82	0.00
33	0.43	0.34	0.82	0.00
34	-	-	-	-
35	0.72	0.61	0.00	0.83
36	0.45	0.54	0.00	0.83
37	0.55	0.33	0.00	0.83
38	0.20	0.47	0.00	0.83

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.67	0.51	0.00	0.83
40	0.76	0.57	0.00	0.83
41	0.40	0.35	0.00	0.83
42	-	-	-	-
43	0.73	0.69	1.14	3.72
44	0.52	0.62	0.21	5.07
45	0.30	0.05	0.62	5.07
46	0.73	0.71	0.83	5.07
47	0.49	0.57	1.24	3.72
48	-	-	-	-
49	0.51	0.41	0.43	6.91
50	0.46	0.50	0.52	5.72
51	-	-	-	-
52	0.29	0.41	1.11	5.62
53	0.37	0.66	1.04	5.72
54	0.44	0.73	1.36	4.91
55	-	-	-	-
56	-	-	-	-
57	0.38	0.51	1.57	4.91
58	0.29	0.24	2.09	4.91
59	0.37	0.42	2.09	4.91

Table 4.B.7 Dichotomous Item Statistics for Form 1—Grade Five

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.65	0.30	0.00	0.00
2	0.78	0.41	0.00	0.00
3	0.84	0.31	0.00	0.00
4	0.76	0.49	0.00	0.00
5	-	-	-	-
6	-	-	-	-
7	0.56	0.30	0.00	0.00
8	0.42	0.32	0.00	0.00
9	0.29	0.11	0.00	0.00
10	0.33	0.53	0.00	0.00
11	0.70	0.58	0.00	0.00
12	-	-	-	-
13	0.34	0.34	0.00	0.00
14	0.39	0.34	0.00	0.00
15	0.20	0.36	0.00	0.00
16	0.35	0.26	0.00	0.00
17	0.67	0.11	0.31	0.00
18	-	-	-	-
19	0.25	0.06	0.31	0.00
20	-	-	-	-
21	0.54	0.51	0.31	0.00
22	0.18	0.01	0.46	0.00
23	0.39	0.36	0.46	0.00
24	0.37	0.32	0.62	0.00
25	0.30	0.31	0.78	0.47
26	0.54	0.55	0.78	0.78
27	0.49	0.43	1.24	0.47
28	0.51	0.64	1.55	0.47
29	0.28	0.01	1.47	0.78
30	0.33	0.50	1.63	0.78
31	-	-	-	-
32	0.33	0.38	2.17	0.47
33	0.50	0.57	2.25	0.78
34	0.52	0.48	2.33	0.47
35	0.36	0.11	0.63	2.86
36	0.57	0.37	0.79	2.86
37	0.30	0.31	0.79	2.86
38	0.49	0.59	0.79	2.86

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	-	-	-	-
40	0.53	0.42	0.79	2.86
41	0.36	0.39	0.79	2.86
42	0.58	0.40	0.95	2.86
43	-	-	-	-
44	-	-	-	-
45	0.61	0.49	0.75	3.46
46	0.34	0.58	0.80	3.68
47	0.07	0.01	0.75	3.46
48	0.16	0.13	0.32	4.52
49	0.39	0.37	0.65	4.52
50	-	-	-	-
51	0.39	0.41	0.97	4.52
52	0.27	0.55	1.29	4.52
53	0.12	0.13	1.12	4.42
54	-	-	-	-
55	0.39	0.33	1.29	4.52
56	0.24	0.32	1.45	4.52
57	0.58	0.51	1.45	4.52
58	0.71	0.19	1.45	4.52
59	-	-	-	-

Table 4.B.8 Dichotomous Item Statistics for Form 2—Grade Five

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.21	0.13	0.00	0.00
2	0.78	0.41	0.00	0.00
3	0.77	0.49	0.00	0.00
4	-	-	-	-
5	0.60	0.29	0.00	0.00
6	0.29	0.22	0.00	0.00
7	0.53	0.29	0.00	0.00
8	0.29	0.11	0.00	0.00
9	0.53	0.42	0.00	0.00
10	-	-	-	-
11	0.24	0.04	0.00	0.00
12	-	-	-	-
13	0.68	0.57	0.15	0.00
14	0.51	0.62	0.31	0.00
15	0.48	0.39	0.31	0.00
16	0.40	0.34	0.31	0.00
17	0.26	0.30	0.15	0.31
18	0.70	0.08	0.15	0.31
19	-	-	-	-
20	0.18	0.08	0.46	0.31
21	0.38	0.26	0.62	0.31
22	0.42	0.40	0.77	0.31
23	0.45	0.43	0.77	0.31
24	0.53	0.59	0.77	0.31
25	0.54	0.55	0.78	0.78
26	0.46	0.60	1.09	1.09
27	0.28	0.01	1.47	0.78
28	0.51	0.47	1.24	1.09
29	0.33	0.50	1.63	0.78
30	-	-	-	-
31	-	-	-	-
32	0.50	0.57	2.25	0.78
33	-	-	-	-
34	0.26	0.37	2.33	1.09
35	0.56	0.44	0.00	3.49
36	0.57	0.36	0.32	3.49
37	0.33	0.24	0.32	3.49
38	0.41	0.30	0.48	3.49

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.69	0.47	0.48	3.49
40	0.53	0.62	0.48	3.49
41	-	-	-	-
42	0.42	0.20	0.48	3.49
43	0.31	0.13	0.63	3.49
44	-	-	-	-
45	0.71	0.67	0.15	3.75
46	0.73	0.74	0.30	3.75
47	0.41	0.29	0.30	3.75
48	0.45	0.32	0.48	4.15
49	-	-	-	-
50	0.27	0.18	0.80	4.32
51	0.45	0.46	0.80	4.32
52	0.12	0.13	1.12	4.42
53	0.66	0.56	0.96	4.32
54	0.58	0.43	0.96	4.32
55	-	-	-	-
56	0.55	0.54	1.44	4.32
57	0.22	0.17	1.44	4.32
58	0.17	0.48	1.44	4.32
59	0.54	0.42	0.16	5.84

Table 4.B.9 Dichotomous Item Statistics for the Accommodated Form—Grade Five

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.56	0.30	0.00	0.00
2	0.78	0.41	0.00	0.00
3	0.77	0.49	0.00	0.00
4	0.76	0.49	0.00	0.00
5	0.60	0.29	0.00	0.00
6	-	-	-	-
7	-	-	-	-
8	0.56	0.59	0.00	0.00
9	-	-	-	-
10	-	-	-	-
11	0.49	0.64	0.00	0.00
12	0.41	0.38	0.00	0.00
13	0.10	0.34	0.00	0.00
14	0.15	0.56	0.00	0.00
15	0.41	0.20	0.00	0.00
16	0.18	0.30	0.00	0.00
17	0.77	0.02	0.00	0.00
18	0.46	-0.01	0.00	0.00
19	0.36	0.20	0.00	0.00
20	0.54	0.54	0.00	0.00
21	0.41	0.30	0.00	0.00
22	0.56	0.40	0.00	0.00
23	0.46	0.60	0.00	0.00
24	0.51	0.09	0.00	0.00
25	0.36	0.09	0.00	0.00
26	-	-	-	-
27	0.38	0.42	0.00	0.00
28	0.41	0.35	0.00	0.00
29	-	-	-	-
30	0.38	0.39	0.00	0.00
31	-	-	-	-
32	0.36	0.43	0.00	0.00
33	0.67	0.42	0.00	0.00
34	0.49	0.33	0.00	0.00
35	0.38	-0.18	0.00	0.00
36	0.46	0.25	0.00	0.00
37	0.28	0.59	0.00	0.00
38	-	-	-	-

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.28	0.24	0.00	0.00
40	0.65	0.30	0.00	0.00
41	-	-	-	-
42	0.61	0.49	0.75	3.46
43	-	-	-	-
44	0.71	0.67	0.15	3.75
45	0.73	0.74	0.30	3.75
46	0.41	0.29	0.30	3.75
47	0.07	0.01	0.75	3.46
48	0.53	0.29	0.00	0.00
49	-	-	-	-
50	0.21	0.26	0.00	0.00
51	0.41	0.15	0.00	0.00
52	0.51	0.61	0.00	0.00
53	0.10	0.02	0.00	0.00
54	0.59	0.44	0.00	0.00
55	-	-	-	-
56	0.46	0.27	0.00	0.00
57	0.28	0.28	0.00	0.00
58	0.23	0.20	0.00	0.00
59	0.15	0.38	0.00	0.00

Table 4.B.10 Dichotomous Item Statistics for Form 1—Grade Six

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.51	0.31	0.00	0.00
2	0.16	0.37	0.00	0.00
3	-	-	-	-
4	0.70	0.48	0.00	0.00
5	0.81	0.39	0.00	0.00
6	0.81	0.40	0.00	0.00
7	-	-	-	-
8	-	-	-	-
9	0.45	0.44	0.00	0.00
10	0.74	0.63	0.00	0.00
11	-	-	-	-
12	-	-	-	-
13	0.46	0.41	0.40	0.00
14	-	-	-	-
15	-	-	-	-
16	0.60	0.45	0.66	0.00
17	0.45	0.46	0.60	0.00
18	-	-	-	-
19	-	-	-	-
20	0.18	-0.02	0.66	0.66
21	0.80	0.39	0.88	0.66
22	0.36	0.10	0.88	0.66
23	0.38	0.28	0.88	0.66
24	0.56	0.16	0.88	0.66
25	0.34	0.37	0.88	0.66
26	0.44	0.57	1.10	0.66
27	0.41	0.17	1.11	1.55
28	-	-	-	-
29	0.44	0.36	1.33	1.55
30	0.56	0.57	0.95	1.58
31	0.64	0.47	1.05	1.58
32	0.51	0.29	1.22	1.43
33	0.34	0.53	0.00	2.92
34	0.37	0.40	0.00	2.89
35	-	-	-	-
36	0.46	0.33	0.45	3.15
37	0.38	0.17	0.68	3.15
38	0.78	0.68	1.35	3.39

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.36	0.33	1.35	3.39
40	0.36	0.23	1.35	3.39
41	0.51	0.44	2.03	3.39
42	-	-	-	-
43	-	-	-	-
44	0.46	0.12	2.48	3.39
45	0.51	0.48	2.48	3.39
46	0.06	0.02	2.48	3.39
47	0.52	0.26	0.21	5.51
48	-	-	-	-
49	0.14	0.46	1.16	6.02
50	0.51	0.70	1.62	6.02
51	-	-	-	-
52	0.46	0.55	1.85	6.02
53	0.62	0.64	1.70	5.11
54	0.20	0.31	1.85	6.02
55	0.50	0.49	1.82	5.11
56	0.55	0.47	2.08	6.02
57	0.29	0.58	2.08	6.02
58	0.46	0.26	0.47	8.27
59	0.21	0.21	0.65	7.56

Table 4.B.11 Dichotomous Item Statistics for Form 2—Grade Six

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not- Reached Rate
1	0.71	0.50	0.00	0.00
2	0.49	0.26	0.00	0.00
3	0.44	0.32	0.00	0.00
4	0.54	0.43	0.00	0.00
5	0.81	0.39	0.00	0.00
6	0.81	0.40	0.00	0.00
7	-	-	-	-
8	0.22	0.01	0.00	0.00
9	0.54	0.45	0.00	0.00
10	0.47	0.31	0.00	0.00
11	0.39	0.47	0.00	0.00
12	-	-	-	-
13	-	-	-	-
14	0.26	0.33	0.39	0.00
15	0.41	0.16	0.43	0.00
16	-	-	-	-
17	0.48	0.44	0.64	0.00
18	0.68	0.41	0.22	0.43
19	0.43	0.09	0.43	0.43
20	0.52	0.33	0.86	0.43
21	0.62	0.51	0.86	0.43
22	-	-	-	-
23	-	-	-	-
24	0.43	0.59	1.29	0.43
25	0.77	0.41	1.29	0.43
26	0.20	0.16	1.29	0.43
27	0.32	0.34	0.00	1.74
28	0.18	0.33	0.65	1.74
29	0.51	0.39	0.65	1.74
30	0.56	0.57	0.95	1.58
31	-	-	-	-
32	0.64	0.47	1.05	1.58
33	0.33	0.13	0.80	1.60
34	0.33	0.25	0.88	2.41
35	0.45	0.52	0.00	3.26
36	-	-	-	-
37	0.72	0.69	0.20	3.26
38	0.25	0.32	0.20	3.26

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.18	0.06	0.22	3.55
40	0.25	0.17	0.44	3.55
41	0.68	0.68	0.44	3.55
42	0.64	0.44	0.44	3.55
43	-	-	-	-
44	0.20	0.55	0.44	3.55
45	-	-	-	-
46	0.26	0.47	0.67	3.55
47	0.47	0.30	0.67	3.55
48	-	-	-	-
49	0.66	0.63	0.67	4.24
50	-	-	-	-
51	-	-	-	-
52	0.42	0.30	1.34	4.24
53	-	-	-	-
54	0.62	0.64	1.70	5.11
55	-	-	-	-
56	0.50	0.49	1.82	5.11
57	0.46	0.54	1.56	4.24
58	0.56	0.46	1.79	4.24
59	0.25	-0.05	0.62	5.41

Table 4.B.12 Dichotomous Item Statistics for the Accommodated Form—Grade Six

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.71	0.50	0.00	0.00
2	0.51	0.31	0.00	0.00
3	-	-	-	-
4	0.70	0.48	0.00	0.00
5	0.49	0.26	0.00	0.00
6	0.44	0.32	0.00	0.00
7	0.54	0.43	0.00	0.00
8	-	-	-	-
9	0.45	0.44	0.00	0.00
10	-	-	-	-
11	0.46	0.41	0.40	0.00
12	0.54	0.45	0.00	0.00
13	0.47	0.31	0.00	0.00
14	-	-	-	-
15	0.39	0.47	0.00	0.00
16	-	-	-	-
17	0.26	0.33	0.39	0.00
18	0.53	0.21	0.00	0.00
19	0.45	0.46	0.60	0.00
20	0.73	0.39	0.00	0.00
21	0.70	0.04	0.00	0.00
22	0.23	-0.21	0.00	0.00
23	0.40	0.76	0.00	0.00
24	0.35	-0.09	0.00	0.00
25	0.23	0.05	0.00	0.00
26	0.48	0.33	0.00	0.00
27	0.56	0.57	0.95	1.58
28	-	-	-	-
29	0.64	0.47	1.05	1.58
30	0.51	0.29	1.22	1.43
31	0.33	0.13	0.80	1.60
32	0.45	0.52	0.00	3.26
33	0.37	0.40	0.00	2.89
34	-	-	-	-
35	0.72	0.69	0.20	3.26
36	-	-	-	-
37	0.25	0.32	0.20	3.26
38	0.45	0.47	0.00	0.00

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.20	-0.27	0.00	0.00
40	0.63	0.69	0.00	0.00
41	-	-	-	-
42	0.33	0.42	0.00	0.00
43	0.53	0.56	0.00	0.00
44	0.03	0.79	0.00	0.00
45	-	-	-	-
46	0.52	0.26	0.21	5.51
47	0.75	0.88	0.00	0.00
48	-	-	-	-
49	-	-	-	-
50	-	-	-	-
51	0.50	0.30	0.00	0.00
52	0.68	0.44	0.00	0.00
53	0.25	0.37	0.00	0.00
54	0.45	0.48	0.00	0.00
55	0.55	0.53	0.00	0.00
56	0.53	0.48	0.00	0.00
57	0.33	0.48	0.00	0.00
58	0.25	-0.05	0.62	5.41
59	0.21	0.21	0.65	7.56

Table 4.B.13 Dichotomous Item Statistics for Form 1—Grade Seven

Item Sequence	<i>p</i> -value	<i>r</i>	Omit Rate	Not-Reached Rate
1	-	-	-	-
2	-	-	-	-
3	-	-	-	-
4	0.79	0.51	0.00	0.00
5	0.30	0.27	0.00	0.00
6	0.44	0.22	0.00	0.00
7	0.37	0.28	0.00	0.00
8	0.42	0.25	0.00	0.00
9	0.41	0.14	0.00	0.00
10	0.54	0.61	0.23	0.00
11	-	-	-	-
12	0.33	0.43	0.24	0.00
13	0.21	0.10	0.24	0.00
14	0.30	0.26	0.48	0.00
15	0.43	0.34	0.48	0.00
16	0.30	0.37	0.48	0.00
17	-	-	-	-
18	0.05	0.13	0.48	0.48
19	0.41	0.17	0.48	0.48
20	0.19	0.17	0.72	0.48
21	0.33	0.52	0.72	0.48
22	0.46	0.45	1.43	0.48
23	0.40	0.13	1.67	0.48
24	0.57	0.44	2.39	0.48
25	0.62	0.53	1.26	2.18
26	0.17	0.40	2.23	1.79
27	-	-	-	-
28	0.50	0.56	2.42	1.94
29	-	-	-	-
30	0.03	0.16	1.91	2.27
31	0.43	0.21	0.46	4.11
32	0.41	0.16	1.50	4.99
33	0.52	0.69	1.75	4.99
34	0.33	0.20	2.06	4.59
35	0.34	0.45	2.49	4.99
36	0.38	0.35	2.52	4.59
37	0.37	0.48	2.74	4.99
38	-	-	-	-

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.35	0.24	2.52	4.59
40	0.36	0.49	0.51	7.95
41	0.33	0.24	1.03	7.95
42	0.29	0.58	1.28	7.95
43	-	-	-	-
44	0.17	-0.07	1.79	7.95
45	0.67	0.50	1.79	7.95
46	0.46	0.50	1.79	7.95
47	0.45	0.38	2.31	7.95
48	-	-	-	-
49	0.35	0.68	2.09	9.92
50	0.31	0.54	2.35	9.92
51	-	-	-	-
52	0.58	0.29	2.61	9.92
53	0.04	0.36	3.19	9.31
54	0.33	0.22	2.81	9.31
55	0.41	0.55	3.39	9.92
56	0.53	0.57	3.66	9.92
57	0.49	0.52	4.21	9.31
58	-	-	-	-
59	0.09	0.56	2.16	13.78

Table 4.B.14 Dichotomous Item Statistics for Form 2—Grade Seven

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.89	0.49	0.00	0.00
2	0.62	0.26	0.00	0.00
3	-	-	-	-
4	-	-	-	-
5	0.57	0.36	0.00	0.00
6	0.61	0.26	0.00	0.00
7	0.42	0.25	0.00	0.00
8	0.37	0.35	0.00	0.00
9	-	-	-	-
10	0.41	0.14	0.00	0.00
11	0.54	0.61	0.23	0.00
12	0.22	-0.19	0.23	0.00
13	0.46	0.50	0.46	0.00
14	0.42	0.15	0.69	0.00
15	0.35	0.17	0.69	0.00
16	0.48	0.47	0.69	0.00
17	-	-	-	-
18	0.03	0.18	0.92	0.69
19	0.56	0.43	1.39	0.69
20	0.59	0.57	1.39	0.69
21	0.25	0.03	1.39	0.69
22	0.34	0.27	1.85	0.69
23	0.49	0.39	1.85	0.69
24	0.32	0.16	2.54	0.69
25	0.54	0.40	0.87	2.39
26	0.61	0.35	1.09	2.39
27	0.62	0.53	1.26	2.18
28	0.20	0.08	1.09	2.39
29	0.03	0.16	1.91	2.27
30	-	-	-	-
31	0.46	0.33	0.00	3.74
32	0.34	0.46	0.48	4.31
33	0.27	0.28	0.48	4.31
34	-	-	-	-
35	-	-	-	-
36	0.51	0.52	1.67	4.31
37	0.55	0.46	1.67	4.31
38	0.44	0.52	1.91	4.31

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	-	-	-	-
40	0.08	0.42	0.97	6.08
41	-	-	-	-
42	0.33	0.29	1.70	6.08
43	0.18	0.42	1.70	6.08
44	0.62	0.45	1.70	6.08
45	0.41	0.28	2.43	6.08
46	0.44	0.30	2.43	6.08
47	-	-	-	-
48	0.46	0.07	2.68	6.08
49	0.39	0.40	1.25	8.73
50	0.42	0.54	1.50	8.73
51	0.41	0.52	1.75	8.73
52	-	-	-	-
53	0.33	0.22	2.81	9.31
54	0.04	0.36	3.19	9.31
55	0.27	0.49	3.24	8.73
56	0.43	0.62	3.99	8.73
57	0.49	0.52	4.21	9.31
58	-	-	-	-
59	0.25	0.10	0.52	13.25

Table 4.B.15 Dichotomous Item Statistics for the Accommodated Form—Grade Seven

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.89	0.49	0.00	0.00
2	-	-	-	-
3	-	-	-	-
4	0.79	0.51	0.00	0.00
5	0.57	0.36	0.00	0.00
6	0.30	0.27	0.00	0.00
7	-	-	-	-
8	0.44	0.22	0.00	0.00
9	0.61	0.26	0.00	0.00
10	-	-	-	-
11	0.34	0.30	0.00	0.00
12	0.17	-0.35	0.00	0.00
13	0.37	0.20	0.00	0.00
14	0.37	0.23	0.00	0.00
15	0.43	0.18	0.00	0.00
16	0.43	0.32	0.00	0.00
17	0.26	0.19	0.00	0.00
18	-	-	-	-
19	0.09	0.24	0.00	0.00
20	0.40	0.39	0.00	0.00
21	0.29	-0.05	0.00	0.00
22	0.60	0.40	0.00	0.00
23	0.34	0.32	0.00	0.00
24	0.49	0.67	0.00	0.00
25	0.54	0.40	0.87	2.39
26	0.61	0.35	1.09	2.39
27	0.62	0.53	1.26	2.18
28	0.20	0.08	1.09	2.39
29	-	-	-	-
30	-	-	-	-
31	0.17	0.40	2.23	1.79
32	0.43	0.21	0.46	4.11
33	0.46	0.33	0.00	3.74
34	0.33	0.20	2.06	4.59
35	0.38	0.35	2.52	4.59
36	-	-	-	-
37	-	-	-	-
38	0.49	0.88	0.00	0.00

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.35	0.24	2.52	4.59
40	0.00	NA	0.00	0.00
41	0.17	-0.03	0.00	0.00
42	0.20	-0.17	0.00	0.00
43	0.26	-0.11	0.00	0.00
44	0.31	-0.09	0.00	0.00
45	0.34	-0.39	0.00	0.00
46	0.46	0.09	0.00	0.00
47	0.54	0.00	0.00	0.00
48	0.29	0.60	0.00	0.00
49	0.54	0.52	0.00	0.00
50	0.31	0.23	0.00	0.00
51	-	-	-	-
52	0.54	0.20	0.00	0.00
53	0.26	0.67	0.00	0.00
54	-	-	-	-
55	0.26	0.02	0.00	0.00
56	0.06	0.78	0.00	0.00
57	0.29	0.36	0.00	0.00
58	0.49	0.88	0.00	0.00
59	0.51	0.73	0.00	0.00

Table 4.B.16 Dichotomous Item Statistics for Form 1—Grade Eight

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.80	0.43	0.00	0.00
2	-	-	-	-
3	0.80	0.51	0.00	0.00
4	0.56	0.23	0.00	0.00
5	0.50	0.58	0.00	0.00
6	0.88	0.54	0.00	0.00
7	-	-	-	-
8	0.57	0.28	0.00	0.00
9	0.04	0.12	0.00	0.00
10	0.46	0.37	0.00	0.00
11	0.44	0.22	0.00	0.00
12	0.34	0.10	0.00	0.00
13	0.62	0.64	0.00	0.00
14	0.15	-0.28	0.00	0.00
15	0.30	0.24	0.36	0.00
16	-	-	-	-
17	0.44	0.37	0.00	0.00
18	-	-	-	-
19	0.51	0.28	0.00	0.00
20	0.30	-0.08	0.00	0.00
21	0.47	0.34	0.36	0.00
22	-	-	-	-
23	0.22	-0.11	0.00	0.00
24	-	-	-	-
25	-	-	-	-
26	-	-	-	-
27	0.33	0.32	0.00	0.00
28	0.28	0.25	0.00	0.00
29	0.22	-0.08	0.00	0.36
30	0.22	-0.16	0.00	0.00
31	-	-	-	-
32	-	-	-	-
33	0.30	0.45	0.00	0.00
34	0.67	0.41	0.00	0.00
35	0.42	0.78	0.00	0.00
36	0.20	0.16	0.00	0.00
37	0.42	0.42	0.00	0.00
38	-	-	-	-

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
39	0.40	0.39	0.72	0.36
40	0.36	0.28	0.00	0.00
41	0.34	0.42	0.72	0.00
42	-	-	-	-
43	0.11	0.31	0.72	0.00
44	0.34	0.11	0.72	0.00
45	0.60	0.46	0.72	0.00
46	0.37	0.48	1.45	0.00
47	0.30	0.48	1.45	0.00
48	0.35	0.63	2.17	0.00
49	0.29	0.11	1.47	1.47
50	0.15	0.62	1.47	1.47
51	-	-	-	-
52	0.46	0.48	1.47	1.47
53	0.43	0.16	1.47	1.47
54	-	-	-	-
55	0.51	0.60	1.47	1.47
56	0.44	0.52	1.47	1.47
57	0.19	0.29	1.47	1.47
58	0.50	0.58	1.82	1.82
59	0.13	0.51	0.00	2.99

Table 4.B.17 Dichotomous Item Statistics for Form 2—Grade Eight

Item Sequence	p-value	r	Omit Rate	Not-Reached Rate
1	0.84	0.46	0.00	0.00
2	0.50	0.33	0.00	0.00
3	-	-	-	-
4	-	-	-	-
5	0.55	0.45	0.00	0.00
6	0.91	0.53	0.00	0.00
7	0.55	0.28	0.00	0.00
8	-	-	-	-
9	0.62	0.46	0.00	0.00
10	0.52	0.33	0.00	0.00
11	0.50	0.43	0.00	0.00
12	0.13	0.33	0.00	0.00
13	0.40	0.50	0.00	0.00
14	0.36	0.49	0.71	0.00
15	0.64	0.24	0.71	0.00
16	0.30	0.24	0.36	0.00
17	0.26	0.44	0.71	0.00
18	0.56	0.38	0.71	0.00
19	-	-	-	-
20	0.21	0.34	0.71	0.00
21	0.33	0.30	0.71	0.00
22	0.47	0.34	0.36	0.00
23	-	-	-	-
24	-	-	-	-
25	-	-	-	-
26	0.28	-0.02	0.00	0.71
27	0.55	0.29	0.00	0.71
28	0.54	0.21	0.00	0.71
29	0.22	-0.08	0.00	0.36
30	0.44	0.17	0.00	0.71
31	-	-	-	-
32	0.36	0.49	0.71	0.71
33	0.44	0.60	1.43	0.71
34	-	-	-	-
35	-	-	-	-
36	0.24	0.49	1.43	0.71
37	0.29	0.45	1.43	0.71
38	-	-	-	-

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.53	0.52	1.43	0.71
40	0.40	0.39	0.72	0.36
41	0.34	0.15	0.00	2.17
42	0.51	0.45	0.00	2.17
43	0.35	0.19	0.00	2.17
44	0.32	0.40	0.00	2.17
45	-	-	-	-
46	0.66	0.53	0.00	2.17
47	-	-	-	-
48	0.38	0.09	0.72	2.17
49	-	-	-	-
50	0.31	0.49	1.45	2.17
51	0.34	0.64	2.17	2.17
52	0.53	0.19	2.17	2.17
53	0.28	0.07	2.17	2.17
54	0.43	0.51	2.17	2.17
55	-	-	-	-
56	0.57	0.46	2.17	2.17
57	0.64	0.66	2.17	2.17
58	0.50	0.58	1.82	1.82
59	0.24	0.01	0.00	4.44

Table 4.B.18 Dichotomous Item Statistics for the Accommodated Form—Grade Eight

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not- Reached Rate
1	NA	NA	NA	NA
2	NA	NA	NA	NA
3	NA	NA	NA	NA
4	NA	NA	NA	NA
5	NA	NA	NA	NA
6	NA	NA	NA	NA
7	NA	NA	NA	NA
8	NA	NA	NA	NA
9	NA	NA	NA	NA
10	NA	NA	NA	NA
11	NA	NA	NA	NA
12	NA	NA	NA	NA
13	NA	NA	NA	NA
14	NA	NA	NA	NA
15	NA	NA	NA	NA
16	NA	NA	NA	NA
17	NA	NA	NA	NA
18	NA	NA	NA	NA
19	NA	NA	NA	NA
20	NA	NA	NA	NA
21	NA	NA	NA	NA
22	NA	NA	NA	NA
23	NA	NA	NA	NA
24	NA	NA	NA	NA
25	NA	NA	NA	NA
26	NA	NA	NA	NA
27	NA	NA	NA	NA
28	NA	NA	NA	NA
29	NA	NA	NA	NA
30	NA	NA	NA	NA
31	NA	NA	NA	NA
32	NA	NA	NA	NA
33	NA	NA	NA	NA
34	NA	NA	NA	NA
35	NA	NA	NA	NA
36	NA	NA	NA	NA
37	NA	NA	NA	NA
38	NA	NA	NA	NA

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	NA	NA	NA	NA
40	NA	NA	NA	NA
41	NA	NA	NA	NA
42	NA	NA	NA	NA
43	NA	NA	NA	NA
44	NA	NA	NA	NA
45	NA	NA	NA	NA
46	NA	NA	NA	NA
47	NA	NA	NA	NA
48	NA	NA	NA	NA
49	NA	NA	NA	NA
50	NA	NA	NA	NA
51	NA	NA	NA	NA
52	NA	NA	NA	NA
53	NA	NA	NA	NA
54	NA	NA	NA	NA
55	NA	NA	NA	NA
56	NA	NA	NA	NA
57	NA	NA	NA	NA
58	NA	NA	NA	NA
59	NA	NA	NA	NA

Table 4.B.19 Dichotomous Item Statistics for Form 1—High School

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	0.35	0.22	0.00	0.00
2	-	-	-	-
3	0.45	0.22	0.00	0.00
4	0.51	0.32	0.00	0.00
5	0.44	0.39	0.00	0.00
6	0.78	0.29	0.00	0.00
7	0.26	0.24	0.00	0.00
8	-	-	-	-
9	0.58	0.32	0.00	0.00
10	-	-	-	-
11	0.22	0.38	0.00	0.00
12	0.66	0.50	0.31	0.00
13	0.37	0.21	0.31	0.00
14	0.24	-0.01	0.31	0.00
15	0.25	0.13	0.31	0.00
16	0.39	0.19	0.31	0.31
17	-	-	-	-
18	-	-	-	-
19	0.37	0.32	0.61	0.31
20	0.53	0.40	0.61	0.31
21	0.52	0.48	0.92	0.31
22	0.46	0.42	0.62	0.31
23	0.60	0.51	1.23	0.31
24	0.44	0.42	1.23	0.31
25	0.34	0.42	0.31	1.24
26	0.56	0.31	0.00	1.55
27	0.54	0.37	0.00	1.55
28	-	-	-	-
29	0.29	0.32	0.00	1.55
30	-	-	-	-
31	0.25	0.36	0.31	1.55
32	0.32	0.29	0.47	1.41
33	0.32	0.39	0.62	1.55
34	0.60	0.43	0.47	1.41
35	0.64	0.56	0.62	1.55
36	0.25	0.16	0.47	1.41
37	0.20	-0.12	0.47	1.41
38	0.56	0.45	0.62	1.55

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.50	0.42	0.62	1.55
40	-	-	-	-
41	0.19	-0.19	0.00	2.19
42	0.40	0.17	0.00	2.19
43	0.28	0.34	0.00	2.19
44	-	-	-	-
45	0.46	0.14	0.00	2.19
46	0.29	0.22	0.00	2.19
47	-	-	-	-
48	0.34	0.30	0.00	2.19
49	0.40	0.28	0.00	2.19
50	0.27	0.26	0.00	2.04
51	-	-	-	-
52	-	-	-	-
53	0.26	0.28	0.31	2.19
54	0.47	0.15	0.31	2.19
55	0.42	0.17	0.31	2.19
56	-	-	-	-
57	-	-	-	-
58	0.61	0.61	0.31	2.51
59	0.34	0.29	0.63	2.51

Table 4.B.20 Dichotomous Item Statistics for Form 2—High School

Item Sequence	p-value	r	Omit Rate	Not- Reached Rate
1	0.35	0.22	0.00	0.00
2	0.48	0.49	0.00	0.00
3	0.60	0.37	0.00	0.00
4	0.51	0.32	0.00	0.00
5	0.44	0.39	0.00	0.00
6	-	-	-	-
7	-	-	-	-
8	0.62	0.52	0.00	0.00
9	0.54	0.37	0.00	0.00
10	-	-	-	-
11	0.49	0.51	0.00	0.00
12	0.25	0.44	0.31	0.00
13	0.29	0.13	0.31	0.00
14	0.61	0.11	0.31	0.00
15	0.65	0.38	0.31	0.00
16	-	-	-	-
17	0.54	0.39	0.00	0.31
18	0.46	0.12	0.00	0.31
19	-	-	-	-
20	0.46	0.42	0.62	0.31
21	0.54	0.31	0.31	0.31
22	0.53	0.46	0.31	0.31
23	0.46	0.38	0.62	0.31
24	0.50	0.50	0.62	0.31
25	0.78	0.23	0.63	0.63
26	0.50	0.20	0.31	0.94
27	0.47	0.41	0.31	0.94
28	0.39	0.38	0.31	0.94
29	0.48	0.47	0.31	0.94
30	0.50	0.47	0.00	1.26
31	0.32	0.29	0.47	1.41
32	0.60	0.43	0.47	1.41
33	0.25	0.16	0.47	1.41
34	0.49	0.49	0.31	1.26
35	0.20	-0.12	0.47	1.41
36	0.49	0.52	0.31	1.26
37	-	-	-	-
38	0.42	0.46	0.31	1.26

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	0.54	0.51	0.31	1.26
40	0.52	0.43	0.32	1.58
41	-	-	-	-
42	0.37	0.38	0.32	1.58
43	0.50	0.51	0.32	1.58
44	0.18	0.05	0.32	1.58
45	0.19	0.53	0.32	1.58
46	0.21	-0.03	0.32	1.58
47	0.76	0.47	0.00	1.90
48	0.56	0.32	0.00	1.90
49	0.53	0.21	0.00	1.90
50	-	-	-	-
51	0.27	0.26	0.00	2.04
52	-	-	-	-
53	0.28	0.33	0.32	1.90
54	-	-	-	-
55	-	-	-	-
56	0.29	0.21	0.63	1.90
57	-	-	-	-
58	0.17	-0.06	0.00	2.55
59	0.17	0.13	0.00	2.55

Table 4.B.21 Dichotomous Item Statistics for the Accommodated Form—High School

Item Sequence	<i>p</i>-value	<i>r</i>	Omit Rate	Not-Reached Rate
1	NA	NA	NA	NA
2	NA	NA	NA	NA
3	NA	NA	NA	NA
4	NA	NA	NA	NA
5	NA	NA	NA	NA
6	NA	NA	NA	NA
7	NA	NA	NA	NA
8	NA	NA	NA	NA
9	NA	NA	NA	NA
10	NA	NA	NA	NA
11	NA	NA	NA	NA
12	NA	NA	NA	NA
13	NA	NA	NA	NA
14	NA	NA	NA	NA
15	NA	NA	NA	NA
16	NA	NA	NA	NA
17	NA	NA	NA	NA
18	NA	NA	NA	NA
19	NA	NA	NA	NA
20	NA	NA	NA	NA
21	NA	NA	NA	NA
22	NA	NA	NA	NA
23	NA	NA	NA	NA
24	NA	NA	NA	NA
25	NA	NA	NA	NA
26	NA	NA	NA	NA
27	NA	NA	NA	NA
28	NA	NA	NA	NA
29	NA	NA	NA	NA
30	NA	NA	NA	NA
31	NA	NA	NA	NA
32	NA	NA	NA	NA
33	NA	NA	NA	NA
34	NA	NA	NA	NA
35	NA	NA	NA	NA
36	NA	NA	NA	NA
37	NA	NA	NA	NA
38	NA	NA	NA	NA

Item Sequence	<i>p</i>-value	r	Omit Rate	Not-Reached Rate
39	NA	NA	NA	NA
40	NA	NA	NA	NA
41	NA	NA	NA	NA
42	NA	NA	NA	NA
43	NA	NA	NA	NA
44	NA	NA	NA	NA
45	NA	NA	NA	NA
46	NA	NA	NA	NA
47	NA	NA	NA	NA
48	NA	NA	NA	NA
49	NA	NA	NA	NA
50	NA	NA	NA	NA
51	NA	NA	NA	NA
52	NA	NA	NA	NA
53	NA	NA	NA	NA
54	NA	NA	NA	NA
55	NA	NA	NA	NA
56	NA	NA	NA	NA
57	NA	NA	NA	NA
58	NA	NA	NA	NA
59	NA	NA	NA	NA

Item and response type abbreviations used in [Table 4.B.22](#) through [Table 4.B.28](#) are as follows:

- MC = Multiple choice
- MS = Multiple select
- MCMS = Multiple choice, multiple select

Table 4.B.22 Polytomous Items Statistics—Grade Three

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1/2/A	VH703598	Inline Choice	Inline Choice MS	0.68	0.57	0.00	16.91	31.07	52.03
1	VH692715	MC	MCMS	0.67	0.44	0.00	11.13	43.25	45.61
1	VH824422	MC	Inline Text Choice MS	0.43	0.24	0.43	29.25	54.62	16.13
1	VH691746	Inline Choice	Inline Choice MS	0.71	0.58	1.41	11.51	35.18	53.31
1	VH691759	Zone	Zone MS	0.86	0.54	1.41	2.61	23.34	74.05
1/2/A	VH692952	MC	MCMS	0.47	0.48	2.46	22.76	60.23	17.02
1	VH692955	Inline Choice	Inline Choice MS	0.62	0.24	2.75	15.73	45.32	38.94
1	VH824376	Composite	Composite	0.34	0.41	3.78	50.00	32.44	17.56
1	VH824374	MC	MCMS	0.53	0.28	3.78	13.89	66.44	19.67
1/A	VH695013	MC	MCMS	0.41	0.38	4.55	34.68	47.83	17.49
1/A	VH695497	MC	MCMS	0.51	0.31	4.55	18.58	60.77	20.65
2/A	VH703581	MC	MCMS	0.55	0.49	0.00	15.68	58.22	26.10
2/A	VH692724	MC	MCMS	0.49	0.33	0.00	19.53	62.35	18.12
2	VH824451	Composite	Composite	0.71	0.46	0.53	11.22	35.36	53.42
2/A	VH691750	MC	MCMS	0.81	0.71	0.95	4.64	28.25	67.11
2/A	VH700500	MC	MCMS	0.47	0.22	1.33	14.84	76.40	8.75
2	VH824403	Composite	Composite	0.34	0.34	3.52	49.61	32.34	18.04
2	VH824389	MC	MCMS	0.53	0.20	3.52	16.39	60.84	22.77
2	VH694870	Match	Match MS	0.30	0.56	4.67	60.40	18.80	20.80

Table 4.B.23 Polytomous Items Statistics—Grade Four

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1/A	VH704231	MC	MCMS	0.53	0.21	0.00	17.13	58.96	23.92
1/A	VH704427	MC	MCMS	0.77	0.56	0.00	5.02	35.53	59.45
1	VH824533	Inline Choice	Inline Choice MS	0.56	0.32	0.00	14.32	59.73	25.95
1	VH887686	MC	MCMS	0.64	0.29	0.22	10.65	51.23	38.12
1	VH887688	MC	MCMS	0.50	0.32	0.22	12.00	75.11	12.89
1/2	VH887690	MC	MCMS	0.55	0.31	0.17	12.06	66.18	21.76
1	VH824492	Inline Choice	Inline Choice MS	0.55	0.21	1.94	23.15	43.44	33.41
1	VH887682	MC	MCMS	0.63	0.55	4.32	15.17	42.71	42.12
1	VH887672	Composite	Composite	0.48	0.59	4.32	35.12	34.19	30.69
1/2	VH708093	Match	Match MS	0.78	0.51	4.90	5.02	33.83	61.16
1/2/A	VH704909	Composite	Composite	0.44	0.49	5.56	34.74	41.58	23.68
1	VH700976	Composite	Composite	0.32	0.47	6.43	45.83	44.05	10.12
1/A	VH701086	MC	MCMS	0.54	0.21	5.72	21.64	48.70	29.66
1	VH701113	MC	MCMS	0.44	0.33	6.43	21.90	67.50	10.60
2	VH704379	Match	Match MS	0.53	0.30	0.00	9.85	74.75	15.40
2	VH824584	Inline Choice	Inline Choice MS	0.54	0.39	0.00	15.18	60.70	24.12
2	VH824590	MC	Inline Text Choice MS	0.50	0.32	0.00	22.65	54.59	22.76
2	VH824583	Match	MatchMS	0.34	0.17	0.00	46.77	37.49	15.74
2	VH824593	MC	MCMS	0.56	0.30	0.00	10.42	67.38	22.20
2	VH887695	MC	MCMS	0.51	0.31	0.11	13.15	70.75	16.10
2	VH824502	MC	Inline Text Choice MS	0.58	0.19	1.61	9.21	66.40	24.40
2	VH824515	Inline Choice	Inline Choice MS	0.72	0.49	1.61	11.62	32.45	55.93
2	VH887669	Inline Choice	Inline Choice MS	0.47	0.47	3.76	33.73	39.37	26.91

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
2	VH887673	Inline Choice	Inline Choice MS	0.56	0.51	3.76	20.80	46.77	32.43
2	VH708130	Zone	Zone MS	0.76	0.60	4.13	4.95	37.97	57.08
2	VH700758	Composite	Composite	0.33	0.68	5.50	53.17	27.36	19.47
2/A	VH700782	MC	MCMS	0.46	0.46	4.91	23.70	61.27	15.03
2	VH701122	Match	Match MS	0.54	0.59	5.50	15.89	59.62	24.49
A	VH704148	MC	MCMS	0.60	0.37	0.00	9.84	59.84	30.33
A	VH704162	MC	MCMS	0.52	0.42	0.00	10.66	75.41	13.93
A	VH704183	MC	MCMS	0.57	0.21	0.00	10.66	65.57	23.77
A	VH706293	MC	MCMS	0.71	0.61	0.83	11.57	35.54	52.89
A	VH871711	MC	MCMS	0.78	0.55	0.83	4.96	33.88	61.16
A	VH700960	MC	MCMS	0.60	0.51	0.83	19.83	41.32	38.84

Table 4.B.24 Polytomous Items Statistics—Grade Five

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1/2/A	VH704416	MC	MCMS	0.66	0.48	0.00	4.03	59.22	36.74
1/A	VH704410	MC	MCMS	0.36	0.15	0.00	39.16	49.78	11.06
1/2	VH883665	MC	MCMS	0.54	0.33	0.00	14.46	63.77	21.77
1	VH824731	MC	Inline Text Choice MS	0.55	0.32	0.00	15.59	57.87	26.54
1	VH824733	Inline Choice	Inline Choice MS	0.44	0.39	0.00	30.40	51.23	18.36
1	VH883773	Inline Choice	Inline Choice MS	0.34	0.33	0.47	41.09	49.30	9.61
1	VH824664	MC	MCMS	0.47	0.61	2.86	29.21	48.10	22.70
1	VH824658	MC	MCMS	0.42	0.34	2.86	26.35	63.65	10.00
1	VH704605	Inline Choice	Inline Choice MS	0.45	0.45	3.68	29.44	51.36	19.20
1	VH708494	Inline Choice	Inline Choice MS	0.69	0.61	4.52	11.13	39.35	49.52
1	VH708628	MC	MCMS	0.68	0.39	4.52	13.06	38.87	48.06
1/A	VH718143	Inline Choice	Inline Choice MS	0.46	0.36	5.69	26.62	54.00	19.38
2	VH883668	MC	MCMS	0.44	0.30	0.00	21.93	68.71	9.36
2	VH824795	Inline Choice	Inline Choice MS	0.42	0.41	0.31	31.69	52.46	15.85
2	VH883771	MC	MCMS	0.51	0.31	1.09	15.81	65.89	18.29
2	VH883761	MC	MCMS	0.61	0.61	1.09	17.98	41.40	40.62
2	VH883774	MC	MCMS	0.38	-0.06	1.09	30.54	63.10	6.36
2	VH824693	MC	MCMS	0.58	0.36	3.49	13.97	56.83	29.21
2/A	VH704599	MC	MCMS	0.59	0.34	3.75	11.11	59.61	29.28
2	VH883649	MC	MCMS	0.55	0.10	4.32	12.48	65.12	22.40
2	VH883654	MC	MCMS	0.58	0.11	4.32	11.68	60.48	27.84
A	VH700929	MC	MCMS	0.53	0.33	0.00	15.38	64.10	20.51
A	VH700970	MC	MCMS	0.46	0.25	0.00	20.51	66.67	12.82

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
A	VH733196	MC	MCMS	0.46	0.41	0.00	25.64	56.41	17.95
A	VH733126	MC	MCMS	0.78	0.61	0.00	7.69	28.21	64.10
A	VH733226	MC	MCMS	0.35	0.07	0.00	35.90	58.97	5.13
A	VH824845	MC	MCMS	0.63	0.35	0.00	7.69	58.97	33.33
A	VH708503	MC	MCMS	0.53	-0.20	0.00	10.26	74.36	15.38
A	VH708784	MC	MCMS	0.47	0.17	0.00	20.51	64.10	15.38

Table 4.B.25 Polytomous Items Statistics—Grade Six

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1/A	VH718520	Composite	Composite	0.54	0.51	0.00	26.51	38.96	34.54
1/2/A	VH718692	MC	MCMS	0.81	0.29	0.00	1.76	33.58	64.66
1	VH641821	Zone	Zone MS	0.70	0.47	0.00	5.02	49.13	45.85
1/2/A	VH687878	MC	MCMS	0.46	0.15	0.00	24.25	60.21	15.54
1	VH687880	Inline Choice	Inline Choice MS	0.41	0.34	0.00	36.03	45.20	18.78
1/A	VH687895	MC	MCMS	0.56	0.35	0.00	11.65	64.86	23.49
1	VH687904	Composite	Composite	0.28	0.56	0.00	53.49	36.68	9.83
1	VH825346	MC	MCMS	0.71	0.36	0.66	4.84	48.57	46.59
1	VH825351	Inline Choice	Inline Choice MS	0.79	0.56	0.66	6.59	29.45	63.96
1	VH691025	Match	Match MS	0.39	0.59	1.55	42.35	37.69	19.96
1/A	VH703709	Inline Choice	Inline Choice MS	0.72	0.33	2.89	10.12	35.54	54.34
1	VH825432	Inline Choice	Inline Choice MS	0.55	0.49	3.39	18.28	52.60	29.12
1	VH825434	MC	MCMS	0.74	0.68	3.39	6.32	40.18	53.50
1/2	VH887619	MC	MCMS	0.47	0.43	5.11	28.07	50.34	21.59
1/2	VH887621	MC	MCMS	0.67	0.52	5.11	6.14	54.66	39.20
2/A	VH688029	Composite	Composite	0.20	0.24	0.00	61.74	36.49	1.78
2	VH688045	Composite	Composite	0.48	0.45	0.00	25.91	52.68	21.41
2	VH825595	Inline Choice	Inline Choice MS	0.77	0.61	0.43	8.39	29.25	62.37
2	VH825596	MC	MCMS	0.62	0.54	0.43	16.56	42.37	41.08
2/A	VH691097	MC	MCMS	0.64	0.39	1.60	7.21	57.92	34.87
2/A	VH703683	Inline Choice	Inline Choice MS	0.87	0.67	3.26	6.31	14.26	79.43
2	VH825506	Inline Choice	Inline Choice MS	0.61	0.36	3.55	12.64	52.33	35.03
2	VH825508	MC	MCMS	0.75	0.71	3.55	4.43	41.02	54.55

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
2/A	VH633937	MC	MCMS	0.71	0.57	3.89	11.89	34.84	53.28
2	VH887613	MC	MCMS	0.71	0.45	4.24	6.92	43.75	49.33
2	VH887617	MC	MCMS	0.59	0.56	4.24	18.97	43.53	37.50
A	VH825675	MC	MCMS	0.68	0.73	0.00	7.50	50.00	42.50
A	VH703659	MC	MCMS	0.46	0.45	0.00	22.50	62.50	15.00
A	VH703427	MC	MCMS	0.66	0.68	0.00	7.50	52.50	40.00
A	VH703698	MC	MCMS	0.61	0.34	0.00	15.00	47.50	37.50

Table 4.B.26 Polytomous Items Statistics—Grade Seven

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1	VH711296	MC	Inline Text Choice MS	0.84	0.37	0.00	2.85	26.84	70.31
1/A	VH703944	Inline Choice	Inline Choice MS	0.76	0.54	0.00	5.04	37.72	57.24
1/A	VH703977	MC	MCMS	0.82	0.58	0.00	6.14	24.12	69.74
1	VH884051	MC	MCMS	0.42	0.31	0.00	30.40	56.06	13.54
1	VH825294	MC	MCMS	0.71	0.53	0.48	7.64	42.96	49.40
1/A	VH732220	MC	MCMS	0.62	0.50	1.79	7.59	60.94	31.47
1/2/A	VH732299	MC	MCMS	0.68	0.51	2.18	7.33	48.34	44.33
1	VH688816	MC	MCMS	0.48	0.41	4.99	22.94	58.35	18.70
1	VH825449	Inline Choice	Inline Choice MS	0.41	0.33	7.95	33.85	50.51	15.64
1	VH825444	Composite	Composite	0.25	0.17	7.95	54.62	40.26	5.13
1	VH887529	MC	MCMS	0.39	0.26	9.92	33.16	54.83	12.01
1	VH887540	Inline Choice	Inline Choice MS	0.66	0.66	9.92	22.19	24.54	53.26
2/A	VH703980	MC	MCMS	0.63	0.38	0.00	8.70	56.26	35.03
2	VH703957	Inline Choice	Inline Choice MS	0.44	0.39	0.00	36.93	37.39	25.69
2	VH884047	MC	MCMS	0.52	0.25	0.00	18.81	57.80	23.39
2	VH825408	MC	MCMS	0.56	0.33	0.69	12.70	62.59	24.71
2/A	VH689249	MC	MCMS	0.58	0.43	3.97	16.56	50.55	32.89
2	VH689400	MC	MCMS	0.55	0.44	4.31	16.51	56.22	27.27
2/A	VH689409	MC	MCMS	0.49	0.32	3.97	27.15	48.57	24.28
2	VH825559	MC	MCMS	0.51	0.14	6.08	14.11	69.59	16.30
2	VH825571	MC	Inline Text Choice MS	0.64	0.47	6.08	9.98	51.82	38.20
2	VH704821	MC	MCMS	0.50	0.28	8.73	20.95	58.10	20.95
2	VH887539	Inline Choice	Inline Choice MS	0.41	0.32	8.73	39.65	38.40	21.95

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
A	VH700308	MC	MCMS	0.40	0.17	0.00	28.57	62.86	8.57
A	VH825547	MC	MCMS	0.43	0.64	0.00	28.57	57.14	14.29
A	VH704769	MC	MCMS	0.33	0.40	0.00	40.00	54.29	5.71
A	VH878753	MC	MCMS	0.54	0.46	0.00	14.29	62.86	22.86

Table 4.B.27 Polytomous Items Statistics—Grade Eight

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1	VH687601	MC	MCMS	0.67	0.51	0.00	10.14	44.93	44.93
1	VH824929	MC	MCMS	0.61	0.37	0.00	15.94	47.10	36.96
1	VH883894	Inline Choice	Inline Choice IMS	0.71	0.38	0.00	13.77	30.43	55.80
1	VH883884	MC	MCMS	0.52	0.28	0.00	10.87	73.91	15.22
1/2	VH883892	Inline Choice	Inline Choice MS	0.47	0.39	0.00	30.82	44.44	24.73
1	VH729944	Composite	Composite	0.49	0.48	0.00	13.77	73.91	12.32
1/2	VH710292	Inline Choice	Inline Choice MS	0.56	0.27	0.36	19.42	48.20	32.37
1	VH710288	MC	MCMS	0.83	0.49	0.00	5.80	23.19	71.01
1/2	VH711654	MC	Inline Text Choice MS	0.35	0.16	0.36	37.05	56.83	6.12
1/2	VH883952	Composite	Composite	0.51	0.46	0.36	23.02	52.16	24.82
1/2	VH711946	MC	Inline Text Choice MS	0.47	0.40	0.36	23.02	60.07	16.91
1	VH824948	Composite	Composite	0.26	0.13	0.00	53.62	39.86	6.52
1	VH883918	Inline Choice	InlineChoiceMS	0.73	0.57	1.47	11.03	31.62	57.35
1	VH710210	MC	MCMS	0.61	0.25	1.47	7.35	63.24	29.41
2	VH687598	Inline Choice	Inline Choice MS	0.52	0.15	0.00	25.53	44.68	29.79
2	VH687620	MC	MCMS	0.73	0.15	0.00	2.84	48.94	48.23
2	VH824992	Inline Choice	Inline Choice MS	0.75	0.44	0.00	5.67	39.01	55.32
2	VH883887	MC	MCMS	0.62	0.65	0.00	14.89	46.81	38.30
2	VH729923	MC	MCMS	0.54	0.26	0.71	11.43	68.57	20.00
2	VH883955	MC	MCMS	0.55	0.43	0.71	8.57	72.14	19.29
2	VH825048	Composite	Composite	0.28	-0.03	2.17	52.17	40.58	7.25
2	VH825055	MC	MCMS	0.66	0.39	2.17	7.25	52.90	39.86
2	VH883917	Composite	Composite	0.39	0.37	2.17	36.96	47.83	15.22
2	VH883929	Composite	Composite	0.40	0.64	2.17	26.09	68.12	5.80

Table 4.B.28 Polytomous Items Statistics—High School

Form	Item ID	Item Type	Response Type	AIS	r	Not Reached Rate	Proportion Earning 0 Points	Proportion Earning 1 Point	Proportion Earning 2 Points
1	VH685015	MC	MCMS	0.69	0.32	0.00	3.67	54.13	42.20
1	VH824898	Inline Choice	Inline Choice MS	0.58	0.34	0.00	18.96	45.57	35.47
1	VH824902	Inline Choice	Inline Choice MS	0.61	0.37	0.00	15.90	46.48	37.61
1	VH883294	MC	MCMS	0.48	0.53	0.31	24.85	54.60	20.55
1/2	VH883296	Inline Choice	Inline Choice MS	0.49	0.39	0.31	28.28	44.51	27.20
1	VH689833	MC	MCMS	0.55	0.46	1.55	19.88	50.31	29.81
1	VH729847	MC	MCMS	0.51	0.26	1.55	20.50	57.45	22.05
1	VH883337	Zone	Zone MS	0.22	0.18	1.55	62.11	32.61	5.28
1	VH825096	MC	MCMS	0.86	0.50	2.19	3.44	21.88	74.69
1	VH825104	MC	MCMS	0.53	0.23	2.19	14.06	65.63	20.31
1	VH883391	MC	MCMS	0.58	0.43	2.19	6.25	72.19	21.56
1	VH883394	MC	Inline Text Choice MS	0.40	0.33	2.19	28.75	62.50	8.75
1/2	VH883400	MC	MCMS	0.65	0.41	2.04	12.89	44.65	42.45
1	VH883401	MC	MCMS	0.55	0.23	2.19	5.94	78.44	15.63
2	VH685130	MC	MCMS	0.64	0.39	0.00	7.76	56.21	36.02
2	VH825037	MC	MCMS	0.63	0.33	0.00	8.39	56.83	34.78
2	VH825030	Inline Choice	Inline Choice MS	0.54	0.31	0.00	20.50	51.24	28.26
2	VH883306	MC	MCMS	0.43	0.41	0.31	27.73	59.50	12.77
2	VH883335	MC	MCMS	0.68	0.34	1.26	11.32	40.88	47.80
2	VH825100	Inline Choice	Inline Choice MS	0.52	0.39	1.58	30.60	35.02	34.38
2	VH883389	MC	MCMS	0.55	0.25	1.90	18.35	53.80	27.85
2	VH883392	MC	MCMS	0.54	0.28	1.90	12.97	65.19	21.84
2	VH883395	Inline Choice	Inline Choice MS	0.55	0.38	1.90	18.04	53.80	28.16
2	VH883386	Inline Choice	Inline Choice MS	0.29	0.22	1.90	47.47	46.52	6.01

Item and response type abbreviations used in [Table 4.B.29](#) through [Table 4.B.42](#) are as follows:

- MC = Multiple choice
- MS = Multiple select
- MCMS = Multiple choice, multiple select

Table 4.B.29 Summary of the Classical Item Statistics by Item Type—Grade Three

Item Type	No. of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	3	0.46	0.34	0.71	0.40	0.34	0.46
Grid	2	0.47	0.15	0.80	0.44	0.28	0.59
Inline Choice	23	0.47	0.20	0.76	0.30	-0.19	0.58
Match	7	0.50	0.26	0.91	0.36	0.09	0.63
MC	89	0.43	0.19	0.81	0.39	-0.12	0.80
Zone	1	0.86	0.86	0.86	0.54	0.54	0.54

Table 4.B.30 Summary of the Classical Item Statistics by Item Type—Grade Four

Item Type	No. of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	4	0.39	0.32	0.48	0.56	0.47	0.68
Grid	2	0.36	0.31	0.40	0.41	0.36	0.47
Inline Choice	22	0.50	0.15	0.74	0.34	-0.05	0.69
Match	8	0.55	0.34	0.78	0.39	0.17	0.68
MC	105	0.46	0.16	0.78	0.40	-0.07	0.73
Zone	2	0.57	0.38	0.76	0.47	0.33	0.60

Table 4.B.31 Summary of the Classical Item Statistics by Item Type—Grade Five

Item Type	No. of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Inline Choice	28	0.50	0.15	0.78	0.36	0.01	0.61
Match	9	0.45	0.17	0.71	0.31	0.01	0.58
MC	113	0.43	0.07	0.78	0.34	-0.20	0.74
Numeric	1	0.58	0.58	0.58	0.40	0.40	0.40
Zone	1	0.84	0.84	0.84	0.31	0.31	0.31

Table 4.B.32 Summary of the Classical Item Statistics by Item Type—Grade Six

Item Type	No. of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	4	0.37	0.20	0.54	0.44	0.24	0.56
Grid	1	0.16	0.16	0.16	0.37	0.37	0.37
Inline Choice	17	0.61	0.18	0.87	0.39	0.06	0.67
Match	6	0.38	0.20	0.51	0.40	0.10	0.70
MC	102	0.48	0.03	0.81	0.39	-0.27	0.88
Zone	4	0.55	0.32	0.81	0.34	0.17	0.47

Table 4.B.33 Summary of the Classical Item Statistics by Item Type—Grade Seven

Item Type	No. of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	1	0.25	0.25	0.25	0.17	0.17	0.17
Grid	4	0.20	0.04	0.54	0.54	0.36	0.78
Inline Choice	15	0.45	0.25	0.76	0.35	0.10	0.66
Match	7	0.33	0.03	0.67	0.35	0.13	0.57
MC	118	0.42	0.0	0.89	0.33	-0.39	0.88

Table 4.B.34 Summary of the Classical Item Statistics by Item Type—Grade Eight

Item Type	No. of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	6	0.39	0.26	0.51	0.34	-0.03	0.64
Grid	2	0.26	0.22	0.30	0.15	-0.16	0.45
Inline Choice	18	0.51	0.11	0.84	0.31	-0.02	0.57
Match	4	0.44	0.15	0.66	0.51	0.42	0.62
MC	78	0.44	0.04	0.91	0.36	-0.28	0.78

Table 4.B.35 Summary of the Classical Item Statistics by Item Type—High School

Item Type	No. of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Grid	1	0.76	0.76	0.76	0.47	0.47	0.47
Inline Choice	16	0.49	0.18	0.78	0.32	0.05	0.46
Match	7	0.41	0.25	0.78	0.25	0.14	0.49
MC	80	0.45	0.17	0.86	0.33	-0.19	0.61
Zone	3	0.38	0.22	0.51	0.22	0.17	0.32

Table 4.B.36 Summary of the Classical Item Statistics by Item and Response Types—Grade Three

Item Type	Response Type	Number of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	Composite	3	0.46	0.34	0.71	0.40	0.34	0.46
Grid	Grid MS	2	0.47	0.15	0.80	0.44	0.28	0.59
Inline Choice	Inline Choice MS	3	0.67	0.62	0.71	0.47	0.24	0.58
Inline Choice	Inline Choice SS	20	0.44	0.20	0.76	0.27	-0.19	0.55
Match	Match MS	4	0.53	0.26	0.91	0.37	0.15	0.56

Item Type	Response Type	Number of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Match	Match SS	3	0.46	0.28	0.62	0.34	0.09	0.63
MC	Inline Text Choice	1	0.48	0.48	0.48	0.53	0.53	0.53
MC	Inline Text Choice MS	1	0.43	0.43	0.43	0.24	0.24	0.24
MC	MCMS	10	0.55	0.41	0.81	0.38	0.20	0.71
MC	MCSS	77	0.42	0.19	0.68	0.39	-0.12	0.80
Zone	Zone MS	1	0.86	0.86	0.86	0.54	0.54	0.54

Table 4.B.37 Summary of the Classical Item Statistics by Item and Response Types—Grade Four

Item Type	Response Type	Number of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	Composite	4	0.39	0.32	0.48	0.56	0.47	0.68
Grid	Grid MS	2	0.36	0.31	0.40	0.41	0.36	0.47
Inline Choice	Inline Choice MS	7	0.53	0.31	0.72	0.38	0.21	0.51
Inline Choice	Inline Choice SS	15	0.49	0.15	0.74	0.31	-0.05	0.69
Match	Match MS	6	0.57	0.34	0.78	0.35	0.17	0.59
Match	Match SS	2	0.48	0.43	0.52	0.51	0.33	0.68
MC	Inline Text Choice	4	0.36	0.16	0.56	0.24	-0.07	0.47
MC	Inline Text Choice MS	2	0.54	0.50	0.58	0.25	0.19	0.32
MC	MCMS	20	0.53	0.20	0.78	0.39	0.16	0.61
MC	MCSS	79	0.45	0.17	0.76	0.42	-0.03	0.73
Zone	Zone MS	1	0.76	0.76	0.76	0.60	0.60	0.60
Zone	Zone SS	1	0.38	0.38	0.38	0.33	0.33	0.33

Table 4.B.38 Summary of the Classical Item Statistics by Item and Response Types—Grade Five

Item Type	Response Type	Number of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Inline Choice	Inline Choice MS	6	0.47	0.34	0.69	0.43	0.33	0.61
Inline Choice	Inline Choice SS	22	0.51	0.15	0.78	0.34	0.01	0.59
Match	Match MS	3	0.58	0.34	0.71	0.45	0.19	0.58
Match	Match SS	6	0.39	0.18	0.58	0.23	0.01	0.43
MC	InlineText Choice	1	0.30	0.30	0.30	0.31	0.31	0.31
MC	Inline Text Choice MS	1	0.55	0.55	0.55	0.32	0.32	0.32
MC	MCMS	25	0.49	0.10	0.78	0.30	-0.20	0.61
MC	MCSS	86	0.41	0.07	0.77	0.36	-0.18	0.74
Numeric	Numeric	1	0.58	0.58	0.58	0.40	0.40	0.40
Zone	ZoneSS	1	0.84	0.84	0.84	0.31	0.31	0.31

Table 4.B.39 Summary of the Classical Item Statistics by Item and Response Types—Grade Six

Item Type	Response Type	Number of Items	Mean p -value	Minimum p -value	Maximum p -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	Composite	4	0.37	0.20	0.54	0.44	0.24	0.56
Grid	Grid MS	1	0.16	0.16	0.16	0.37	0.37	0.37
Inline Choice	Inline Choice MS	7	0.67	0.41	0.87	0.48	0.33	0.67
Inline Choice	Inline Choice SS	10	0.56	0.18	0.80	0.33	0.06	0.64
Match	Match MS	6	0.38	0.20	0.51	0.40	0.10	0.70
MC	Inline Text Choice	2	0.55	0.46	0.64	0.35	0.26	0.44
MC	MCMS	19	0.58	0.03	0.81	0.47	0.02	0.79
MC	MCSS	81	0.45	0.14	0.81	0.38	-0.27	0.88
Zone	Zone MS	1	0.70	0.70	0.70	0.47	0.47	0.47
Zone	Zone SS	3	0.50	0.32	0.81	0.30	0.17	0.39

Table 4.B.40 Summary of the Classical Item Statistics by Item Type and Response Type—Grade Seven

Item Type	Response Type	Number of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	Composite	1	0.25	0.25	0.25	0.17	0.17	0.17
Grid	Grid MS	4	0.20	0.04	0.54	0.54	0.36	0.78
Inline Choice	Inline Choice MS	6	0.49	0.25	0.76	0.39	0.10	0.66
Inline Choice	Inline Choice SS	9	0.43	0.27	0.58	0.33	0.20	0.47
Match	Match MS	5	0.20	0.03	0.53	0.31	0.13	0.57
Match	Match SS	2	0.64	0.62	0.67	0.47	0.45	0.50
MC	Inline Text Choice	1	0.40	0.40	0.40	0.13	0.13	0.13
MC	Inline Text Choice MS	3	0.52	0.09	0.84	0.47	0.37	0.56
MC	MCMS	22	0.48	0.00	0.82	0.39	0.14	0.64
MC	MCSS	92	0.40	0.09	0.89	0.32	-0.39	0.88

Table 4.B.41 Summary of the Classical Item Statistics by Item and Response Types—Grade Eight

Item Type	Response Type	Number of Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Composite	Composite	6	0.39	0.26	0.51	0.34	-0.03	0.64
Grid	Grid MS	2	0.26	0.22	0.30	0.15	-0.16	0.45
Inline Choice	Inline Choice MS	8	0.52	0.11	0.75	0.37	0.15	0.57
Inline Choice	Inline Choice SS	10	0.50	0.28	0.84	0.26	-0.02	0.46
Match	Match MS	2	0.25	0.15	0.34	0.52	0.42	0.62
Match	Match SS	2	0.63	0.60	0.66	0.49	0.46	0.53
MC	Inline Text Choice MS	2	0.41	0.35	0.47	0.28	0.16	0.40
MC	MCMS	13	0.58	0.34	0.83	0.38	0.15	0.65
MC	MCSS	63	0.42	0.04	0.91	0.35	-0.28	0.78

**Table 4.B.42 Summary of the Classical Item Statistics by Item and Response Types—
High School**

Item Type	Response Type	Number of Items	Mean <i>p</i>-value	Minimum <i>p</i>-value	Maximum <i>p</i>-value	Mean Polyserial	Minimum Polyserial	Maximum Polyserial
Grid	Grid SS	1	0.76	0.76	0.76	0.47	0.47	0.47
Inline Choice	Inline Choice MS	8	0.49	0.29	0.61	0.34	0.22	0.39
Inline Choice	Inline Choice SS	8	0.48	0.18	0.78	0.31	0.05	0.46
Match	Match MS	1	0.48	0.48	0.48	0.49	0.49	0.49
Match	Match SS	6	0.40	0.25	0.78	0.21	0.14	0.32
MC	Inline Text Choice	4	0.33	0.22	0.58	0.37	0.32	0.44
MC	Inline Text Choice MS	2	0.33	0.25	0.40	0.35	0.33	0.36
MC	MCMS	15	0.59	0.43	0.86	0.36	0.23	0.53
MC	MCSS	59	0.43	0.17	0.66	0.32	-0.19	0.61
Zone	Zone MS	1	0.22	0.22	0.22	0.18	0.18	0.18
Zone	Zone SS	2	0.47	0.42	0.51	0.25	0.17	0.32

Appendix 4.C Response Time Analysis

Notes:

- Raw scores for machine-scorable items were used to partition students into quartiles.
- All students who completed the test and have unrounded test time greater than 0 are included.
- Grade levels reflect students’ enrolled grade levels during the 2017–18 school year.

Table 4.C.1 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Three

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	%	%	%	%	%	%	%
							Pt. 1	Pt. 10	Pt. 25	Pt. 50	Pt. 75	Pt. 90	Pt. 99
1	0–26	231	71.49	42.47	6.83	300.32	13.80	22.99	39.72	65.02	93.17	130.32	201.10
1	27–32	223	90.21	43.52	19.43	267.20	29.13	46.89	58.95	80.69	108.50	146.52	242.72
1	33–39	230	95.09	41.51	31.63	337.15	38.50	54.17	66.02	85.93	112.17	145.44	219.94
1	40–70	250	90.63	39.39	33.98	389.21	36.45	52.77	65.08	81.67	108.32	135.79	215.37
2	0–24	227	68.11	42.60	11.07	308.50	12.52	22.46	38.38	61.49	84.94	126.70	191.14
2	25–31	228	88.79	41.69	22.00	309.88	27.10	41.62	59.39	79.52	111.72	141.69	217.88
2	32–40	249	95.97	42.66	23.01	280.94	35.87	53.58	65.01	86.36	114.43	159.79	232.04
2	41–69	237	89.17	35.44	30.72	256.96	37.14	51.80	63.90	79.28	109.07	141.95	180.42
A	0–20	28	67.99	36.06	22.36	195.73	22.36	23.98	45.86	64.17	81.49	104.27	195.73
A	21–25	31	68.99	26.12	27.41	141.09	27.41	39.83	49.34	66.94	81.39	100.33	141.09
A	26–34	33	82.27	39.67	23.62	257.99	23.62	57.26	66.31	73.60	78.91	108.61	257.99
A	35–67	32	69.24	23.80	32.72	145.98	32.72	41.94	53.23	65.78	83.29	95.87	145.98

Table 4.C.2 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Four

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–25	189	55.39	39.86	6.57	215.60	7.37	13.28	20.90	45.07	78.59	109.63	163.62
1	26–32	235	76.66	44.50	7.85	390.47	9.85	29.74	47.03	67.09	98.74	134.25	187.63
1	33–40	218	86.63	30.59	13.11	168.85	30.57	52.02	64.82	82.30	104.60	134.28	163.24
1	41–73	252	91.42	33.72	31.04	242.39	38.81	54.45	68.06	83.99	106.92	136.66	205.28
2	0–26	191	57.11	39.35	6.41	207.63	7.96	16.64	26.59	46.43	75.78	115.55	168.10
2	27–33	230	76.90	38.40	7.24	205.51	12.55	31.24	50.35	69.68	102.65	131.09	181.10
2	34–42	237	88.33	34.79	18.21	261.21	21.36	49.39	64.10	85.98	102.69	137.06	181.06
2	43–76	225	91.56	36.93	41.07	357.01	42.86	55.71	67.56	83.58	103.85	129.48	194.51
A	0–28	28	87.57	48.12	7.89	205.68	7.89	16.12	65.12	88.08	97.48	158.05	205.68
A	29–36	32	95.57	28.12	40.26	160.26	40.26	59.94	76.91	97.53	109.03	125.14	160.26
A	37–42	29	103.09	27.40	54.67	174.99	54.67	75.16	83.64	97.24	117.25	146.59	174.99
A	43–70	33	100.24	40.05	53.42	291.22	53.42	65.67	80.60	96.85	116.65	121.87	291.22

Table 4.C.3 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Five

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–24	141	85.02	36.73	19.87	258.54	27.26	41.08	59.24	81.82	104.00	131.38	192.35
1	25–30	170	94.35	32.75	37.15	194.76	38.52	54.67	69.04	88.79	112.55	141.42	177.14
1	31–37	165	104.90	32.38	19.38	206.50	33.14	66.16	82.53	102.67	123.49	146.87	195.54
1	38–71	172	99.91	30.90	31.84	212.58	41.31	67.67	77.71	97.40	115.62	140.06	200.61
2	0–25	148	85.21	32.75	14.25	180.70	22.51	47.68	63.57	81.23	101.28	134.47	170.16
2	26–32	174	97.09	36.91	25.80	202.05	27.43	52.77	71.25	92.27	118.50	153.86	196.17
2	33–39	165	105.01	36.92	34.45	246.69	36.74	59.77	76.80	104.16	126.87	152.01	199.97
2	40–70	165	97.68	30.60	41.76	218.03	47.18	61.11	74.81	94.54	118.09	136.94	184.69
A	0–27	9	89.08	26.88	60.24	138.07	60.24	60.24	77.69	80.79	90.79	138.07	138.07
A	28–31	10	76.02	23.21	45.75	122.70	45.75	51.32	58.66	72.40	88.18	111.69	122.70
A	32–36	9	91.99	57.75	46.60	237.43	46.60	46.60	66.68	75.94	95.37	237.43	237.43
A	37–71	11	82.20	27.62	45.89	140.08	45.89	51.62	61.10	84.87	102.14	102.73	140.08

Table 4.C.4 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Six

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–28	107	77.91	36.63	10.59	204.03	25.82	40.53	47.07	69.76	101.79	129.64	165.84
1	29–36	115	87.69	39.11	26.98	242.86	29.25	44.60	56.43	85.57	107.47	137.54	201.36
1	37–44	109	91.39	28.46	36.31	170.25	36.64	54.42	73.07	88.11	108.82	130.57	156.77
1	45–74	127	90.87	27.41	44.01	176.31	45.64	58.97	72.06	86.87	103.34	128.18	168.78
2	0–29	107	73.06	36.22	6.54	226.58	11.52	29.26	45.65	67.23	101.72	115.00	154.11
2	30–38	121	90.02	39.35	33.66	309.20	35.69	52.01	68.67	83.15	105.70	131.31	237.11
2	39–46	117	88.60	30.37	34.36	185.51	35.41	51.96	64.24	88.02	107.95	132.84	168.00
2	47–74	122	85.91	27.30	34.65	159.34	36.59	52.11	64.89	82.97	102.21	129.03	158.68
A	0–26	8	52.16	22.57	25.92	93.70	25.92	25.92	36.69	45.93	66.20	93.70	93.70
A	27–38	12	65.11	22.13	41.75	100.71	41.75	42.49	47.26	54.18	86.02	96.73	100.71
A	39–44	10	94.53	23.76	60.52	132.51	60.52	66.83	79.14	90.55	120.68	128.19	132.51
A	45–72	10	78.23	22.00	47.14	117.41	47.14	49.76	62.81	79.96	92.19	107.88	117.41

Table 4.C.5 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Seven

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–22	98	90.06	43.91	6.43	358.82	6.43	53.14	72.56	86.14	102.78	128.31	358.82
1	23–29	111	92.36	39.12	12.74	213.75	17.49	50.55	65.41	87.48	115.02	148.29	206.33
1	30–37	94	103.76	46.95	31.48	341.61	31.48	63.09	77.30	98.37	113.01	152.42	341.61
1	38–71	118	101.32	39.21	32.48	408.22	55.11	65.56	75.90	95.99	120.26	133.80	167.89
2	0–24	106	84.93	34.31	7.32	195.13	13.11	41.17	63.21	85.94	102.73	125.19	166.15
2	25–30	110	89.84	41.61	12.48	243.83	16.90	36.51	65.65	87.24	108.59	150.42	227.07
2	31–37	108	98.08	31.39	34.76	220.87	40.93	65.73	76.90	94.42	112.28	137.90	215.98
2	38–71	112	97.81	29.21	54.46	233.15	55.03	64.82	77.22	93.61	115.58	129.22	198.06
A	0–25	8	129.47	25.17	93.29	162.75	93.29	93.29	111.25	129.47	149.13	162.75	162.75
A	26–29	9	143.01	28.91	110.50	194.68	110.50	110.50	117.90	151.66	159.42	194.68	194.68
A	30–33	8	155.02	22.88	113.05	186.66	113.05	113.05	141.29	159.83	169.10	186.66	186.66
A	34–70	10	160.60	95.08	84.37	418.53	84.37	88.53	129.07	135.82	156.32	302.50	418.53

Table 4.C.6 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Eight

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–26	33	70.82	27.02	27.55	129.43	27.55	38.68	50.67	65.88	88.03	106.25	129.43
1	27–32	34	86.60	25.00	46.16	142.18	46.16	52.20	71.28	84.35	102.89	120.16	142.18
1	33–37	30	95.87	23.51	55.33	141.88	55.33	64.91	78.89	91.13	113.45	125.36	141.88
1	38–73	41	93.97	24.55	45.73	180.13	45.73	62.61	74.08	93.74	107.13	117.31	180.13
2	0–27	34	69.25	23.90	34.25	139.00	34.25	45.56	55.14	61.84	77.64	111.92	139.00
2	28–32	30	76.57	28.35	14.10	135.18	14.10	41.65	58.43	73.01	89.62	118.48	135.18
2	33–39	41	91.49	26.38	47.33	139.73	47.33	62.60	68.92	89.56	106.38	126.29	139.73
2	40–74	36	89.45	22.83	52.73	134.60	52.73	63.86	69.13	84.47	109.30	116.29	134.60
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.C.7 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Nine

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–26	25	76.55	40.19	14.84	157.40	14.84	21.96	55.85	76.06	105.99	138.17	157.40
1	27–32	24	85.64	24.99	46.01	155.70	46.01	60.38	71.60	78.97	96.92	124.72	155.70
1	33–38	28	107.85	39.00	47.97	181.65	47.97	49.72	86.04	99.10	140.99	163.41	181.65
1	39–73	30	95.50	38.81	42.95	192.51	42.95	52.43	62.48	86.65	114.17	148.15	192.51
2	0–27	22	72.55	29.69	9.20	135.89	9.20	44.32	54.17	71.50	89.14	101.59	135.89
2	28–34	25	97.74	43.51	51.04	255.71	51.04	62.34	70.98	87.09	110.33	141.94	255.71
2	35–39	18	96.43	35.80	47.04	173.78	47.04	50.55	73.47	91.32	119.96	163.77	173.78
2	40–71	30	95.60	30.27	45.48	170.74	45.48	65.33	71.30	88.96	111.45	133.05	170.74
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.C.8 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Ten

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–28	34	75.02	32.43	21.71	168.72	21.71	42.77	58.24	71.45	83.80	119.67	168.72
1	29–33	28	75.79	30.24	35.80	164.04	35.80	42.60	53.11	68.66	100.34	112.95	164.04
1	34–37	38	88.39	26.20	43.48	160.78	43.48	59.84	66.92	85.01	109.38	123.44	160.78
1	38–73	41	90.44	24.01	49.51	165.46	49.51	62.92	72.18	87.01	107.66	116.61	165.46
2	0–27	35	69.32	30.67	20.02	143.46	20.02	38.71	47.64	62.56	89.01	112.48	143.46
2	28–34	35	84.71	33.51	43.69	167.38	43.69	47.51	56.21	83.17	104.97	137.69	167.38
2	35–39	30	95.56	39.46	34.87	220.19	34.87	49.73	68.23	89.72	121.27	143.55	220.19
2	40–71	41	96.65	32.25	43.58	170.27	43.58	61.58	75.66	94.43	112.30	146.91	170.27
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.C.9 Total Testing Time (in Minutes) at Each Raw Score Interval—Grade Eleven

Form	Raw Score Interval	N	Mean	Standard Deviation	Minimum	Maximum	% Pt. 1	% Pt. 10	% Pt. 25	% Pt. 50	% Pt. 75	% Pt. 90	% Pt. 99
1	0–26	17	64.06	29.59	15.51	120.70	15.51	19.40	47.65	58.45	81.62	114.78	120.70
1	27–31	18	70.85	33.44	16.88	132.37	16.88	19.50	49.46	64.44	86.78	120.81	132.37
1	32–37	23	73.13	36.03	25.25	160.86	25.25	34.03	48.23	67.14	88.39	124.10	160.86
1	38–73	21	84.29	22.44	50.06	142.66	50.06	60.57	68.02	82.54	93.38	110.36	142.66
2	0–26	20	56.28	24.30	12.85	99.73	12.85	19.95	44.47	52.32	73.11	92.26	99.73
2	27–32	22	75.27	40.33	18.80	209.25	18.80	27.36	57.49	70.71	85.18	105.84	209.25
2	33–38	22	83.53	28.15	35.94	129.11	35.94	48.60	72.78	82.11	93.63	126.46	129.11
2	39–71	22	82.49	29.84	35.95	159.73	35.95	55.59	63.83	76.49	91.78	126.77	159.73
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
A	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Appendix 4.D Differential Item Functioning (DIF) Analyses

Notes:

- Grade levels reflect students’ enrolled grade levels during the 2017–18 school year.
- Meaning of DIF categories are as follows:
 A = Negligible DIF
 B = Moderate DIF
 C = Large DIF
 “+” = Favors the focal group
 “-” = Favors the reference group.
- Small N indicates the DIF analysis was not performed for the item due to insufficient sample size.

Table 4.D.1 Gender DIF Classifications Summary by Grade Level

DIF Category	Grade 3 Number	Grade 3 Percent	Grade 4 Number	Grade 4 Percent	Grade 5 Number	Grade 5 Percent	Grade 6 Number	Grade 6 Percent	Grade 7 Number	Grade 7 Percent	Grade 8 Number	Grade 8 Percent	High School Number	High School Percent
C-	0	0.00	0	0.00	1	0.66	0	0.00	0	0.00	0	0.00	0	0.00
B-	1	0.80	1	0.70	2	1.32	4	2.99	4	2.76	0	0.00	4	2.65
A-	57	45.60	51	35.66	49	32.24	46	34.33	54	37.24	0	0.00	47	31.13
A+	50	40.00	52	36.36	56	36.84	55	41.04	47	32.41	0	0.00	53	35.10
B+	2	1.60	2	1.40	1	0.66	3	2.24	4	2.76	0	0.00	3	1.99
C+	0	0.00	1	0.70	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Small N	15	12.00	36	27.27	43	28.30	26	19.40	36	24.83	152	100.00	46	29.14
Total	125	100.00	143	100.00	152	100.00	134	100.00	145	100.00	152	100.00	153	100.00

Table 4.D.2 Items Exhibiting Significant DIF by Student Group

Grade Level	Item ID	N Focal	N Reference	MH DDIF	Comparison	In Favor Of
4	VH708131	532	473	1.79	Male–Female	Female
5	VH704377	355	336	-2.31	Male–Female	Male

Chapter 5: Quality Control

The California Department of Education (CDE) and Educational Testing Service (ETS) implemented rigorous quality control procedures throughout the test development, administration, scoring, and analyses processes. As part of this effort, ETS staff worked with its Office of Professional Standards Compliance, which publishes and maintains the *ETS Standards for Quality and Fairness* (ETS, 2014). These standards support the goal of delivering technically sound, fair, and useful products and services; and assisting the public and auditors in evaluating those products and services. This chapter highlights the quality control processes used at various stages of administration.

5.1. Quality Control of Item Development

ETS' goal is to provide the best standards-based and innovative items for the California Spanish Assessment (CSA). Items developed for the CSA field test were subject to an extensive item review process. The item writers hired to develop CSA items and tasks, some of whom are current California educators, were trained in California Assessment of Student Performance and Progress (CAASPP) and ETS policies on quality control of item content, sensitivity and bias guidelines, and guidelines for accessibility to ensure that the items allow the widest possible range of students to demonstrate their content knowledge.

Once a written item was accepted for authoring—that is, once it was entered into ETS' item bank and formatted for use in an assessment—ETS employed a series of internal and external reviews. These reviews used established criteria and specifications to judge the quality of item content and to ensure that each item measured what it is intended to measure. These reviews also examined the overall quality of the test items before presentation to the CDE and item reviewers. To finish the process for the field test items, a group of California educators reviewed the items and performance tasks for accessibility, bias and sensitivity, and content, and made recommendations for item enhancement. The details on item development processes for quality control purposes are described in subsection [2.5 Item Review Process](#) of [Chapter 2: Item Development and Test Assembly](#).

When student response data on each item became available, ETS Psychometric Analysis and Research (PAR) staff conducted item analyses to examine whether the items performed as expected. When the CSA field test was completed and the population data was available, psychometric staff conducted a thorough item analysis and evaluated all items carefully using the statistical criteria described in subsection [4.4.4 Summary of Classical Item Analyses Flagging Criteria](#). PAR flagged items that were potentially problematic due to poor item performance, content issues, item bias, and accessibility challenges. After that, a data review process was implemented, where a group of California educators and ETS content staff reviewed the items and performance tasks together with their associated statistical results and made recommendations to the CDE about item disposition.

5.2. Quality Control of Test Form Development

ETS conducted multiple levels of quality assurance checks on each constructed field test form to ensure it met the form-building specifications. Both ETS Assessment Development (AD) and PAR staff reviewed and signed off on the accuracy of forms before the test forms were put into production for administration in the field test. Detailed information related to test assembly can be found in subsection [2.6 Test Assembly and Length](#).

In particular, the assembly of all test forms went through a certification process that included various checks, including verifying that:

- all answers are correct,
- answers are scored correctly in the item bank and incorrect answers are scored as incorrect,
- all items align with the standard,
- all content in the item is correct,
- distractors are plausible,
- multiple-choice item options are parallel in structure,
- language is grade-level appropriate,
- no more than three multiple choice items in a row have the same key,
- all art is correct,
- there are no errors in spelling or grammar, and
- items adhere to the approved style guide.

Reviews were also conducted for functionality and sequencing during the user acceptance testing (UAT) process to ensure all items functioned as expected.

5.3. Quality Control of Test Materials

5.3.1. Developing Test Administration Instructions

ETS staff consult with internal subject matter experts and conduct validation checks to verify that test instructions accurately match the testing processes. Copy editors and content editors review each document for spelling, grammar, accuracy, and adherence to CDE style and usage requirements as well as the CDE accessibility standards. CSA content is incorporated to fit the CAASPP System specifications. All CAASPP documents are approved by the CDE before they can be published to the CAASPP Portal at <http://www.caaspp.org/>. Only nonsecure documents are posted to this website.

5.3.2. Processing Test Materials

Online tests that were submitted by students were transmitted from the American Institutes for Research (AIR) to ETS each day. Each system checked for the completeness of the student record and stopped records that were identified as having an error and were flagged for review.

5.4. Quality Control of Test Administration

The quality of test administration for the CSA, and all assessments administered as part of the CAASPP System, was monitored and controlled through several strategies. A fully staffed support center, the California Technical Assistance Center (CaTAC), supports all local educational agencies (LEAs) in the administration of CAASPP assessments. In addition to providing guidance and answering questions, CaTAC regularly conducts outreach campaigns on particular administration topics to ensure all LEAs understand correct test administration procedures. CaTAC is guided by a core group of LEA Outreach Advocacy staff that manage communications to LEAs; provide regional and web-based

trainings; and host a website, <http://www.caaspp.org/>, that houses a full range of manuals, videos, and other instructional and support materials.

The quality of test administration was further managed through comprehensive rules and guidelines for maintaining the security and standardization of CAASPP assessments, including the CSA field test. LEAs received training on these topics and were provided tools for reporting security incidents and resolving testing discrepancies for specific testing sessions.

The ETS Office of Testing Integrity (OTI) reinforced the quality control procedures for test administration, providing quality assurance services for all testing programs managed by ETS. The detailed procedures OTI developed and applied in quality control are described in subsection [3.2.1 ETS' Office of Testing Integrity \(OTI\)](#).

5.5. Quality Control of Machine Scoring

AIR, the CAASPP subcontractor, provided the test delivery system (TDS) and scored machine-scorable items. AIR psychometric staff members independently reviewed all CSA test forms by taking sample tests. Responses to the test forms were compared with the answer keys for each form to confirm the accuracy of scoring keys. The scores for all applicable items were recorded. A final comparison of the test map to each online form as configured in the UAT environment ensured that no changes to the form were introduced prior to operational deployment.

A real-time, quality-monitoring component was built into the TDS. After a test was administered to a student, the TDS passed the resulting data to the quality assurance (QA) system. QA conducted a series of data integrity checks, ensuring, for example, that the record for each test contained information for each item, keys for multiple-choice items, score points in each item, and the total number of operational items, and that the test record contained no data from items that might have been invalidated.

Data passed directly from the Quality Monitoring System to the Database of Record, which served as the repository for all test information, and from which all test information was pulled and transmitted to ETS in a predetermined results format.

5.6. Quality Control of Psychometric Processes

5.6.1. Development of Psychometric Specifications

ETS scoring specifications for the CSA field test were completed, reviewed, approved, and checked in advance of the receipt of student response data. Before psychometric analysis, PAR developed a psychometric analysis plan and road map, describing each step of psychometric analyses, procedures, and schedules. This plan was submitted to CDE for review and approval. After that, psychometric specifications were developed for ETS data analysts conducting all analyses. Psychometric specifications contained detailed scoring procedures as well as the procedures for determining whether a student attempted a test and whether that student's response data should be included in the statistical analyses and calculations for computing summary data.

5.6.2. Development of Psychometric Procedures

Prior to the test administration, ETS Assessment Development (AD) staff reviewed and verified the keys and scoring rubrics for each item. Then, these keys and rubrics were provided to AIR for implementing machine scoring of the MC items. After AIR finished machine scoring, item scores and responses were delivered to ETS.

ETS's Centralized Repository Distribution System and Enterprise Service Bus departments collected and parsed .xml files that contained student response data from AIR. Following successful validation, the student response statistical extracts were made available to the psychometric team. Classical item analyses and differential item functioning analyses were then conducted using verified data.

All psychometric analyses conducted at ETS underwent comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists and psychometric specifications were developed by members of the team for each of the statistical procedures performed on CSA results data. Items that were flagged for questionable statistical attributes were sent to ETS AD staff for review; their comments were reviewed by the psychometricians before the data review meetings with the CDE.

Reference

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>

Chapter 6: Continuous Improvement

The California Spanish Assessment (CSA) had its field test administration in fall 2018. Since its inception, continuous efforts have been made to improve the CSA in various ways. This chapter summarizes the accomplishments and ongoing improvements to the CSA in terms of test design, item development, test delivery and administration, psychometric analyses, and accessibility.

6.1. Feedback from the Data Review Meeting

Participants at the data review meeting reviewed 172 flagged items, all of which were being considered for potential use on 2018–19 operational forms. Based on item statistics from the field test data analysis, some of the items were removed from the 2018–19 operational forms and replaced with other field-tested items. [Table 6.1](#) presents the number of grade-level items flagged and reviewed by a panel of educators at the data review meeting.

Table 6.1 Summary of Flagged Items

Grade Level	Item Count
3	26
4	22
5	19
6	17
7	17
8	52
High school	19
Total:	172

Due to low n-counts, the entire slate of 52 grade-eight items was reviewed. Reviewers escalated concerns about 10 of the 172 items reviewed to the California Department of Education for further discussions, and three items were removed from the operational pool.

The data review participants were unaccustomed to reviewing the statistics of an item and then focusing on the item’s Spanish language-arts content. Feedback from the data review meeting indicated that participants would like more time devoted to the presentation on statistics, adding more thorough explanation of what the terminology refers to, using relatable examples. For future data review meetings, a simplified format of the item data review sheets will be developed and approved by the CDE to help the data review participants understand the statistical terminology used in the review process.

6.2. Test Delivery and Administration

6.2.1. Survey Results

The California Assessment of Student Performance and Progress (CAASPP) program solicits feedback annually for participants of the suite of CAASPP assessments, through the CAASPP Post-Test Survey. Given the CSA program administered the field test ahead of the 2018–19 CAASPP testing window, local educational agencies (LEAs) were asked to provide their feedback through a specific instrument, the CSA Field Test Post-Test Survey.

Approximately 50 participating LEA staff members provided feedback and actionable insights for the CSA field test. Overall, the state's educators indicated a successful CSA field test administration. Eighty-eight percent of respondents described a smooth administration with respect to planning and managing test sessions and found the resources provided to be helpful or somewhat helpful. Educators also reported positive testing experiences for students with regard to the testing interface and online testing platform. Across all LEA staff roles, 97 percent of respondents deemed the resources provided for administration were useful for a successful test administration.

The one common issue reported on the CSA Field Test Post-Test Survey was that LEAs had trouble with internet connections. In response to this feedback, ETS will further stress the importance of having technology prepared ahead of testing. As the CSA continues its annual administration, the CDE and Educational Testing Service (ETS) will continue their outreach efforts to LEAs to provide test administration support. ETS also will use focus groups, surveys, and evaluations to continually improve overall CSA-related processes, systems, and resources.

A summary of the survey results is included in the CSA Field Test Post-Test Survey results summary document (ETS, 2019).

6.2.2. Training and Communication

As ETS and the CDE continue the incorporation of the CSA into the CAASPP System of assessments, training and communication will be a focal point moving forward. Because the CSA is a new assessment and completely voluntary, ETS will continue to provide statewide training specific to the CSA to LEA staff and test administrators so they are prepared to administer the test. Training will include Usability, Accessibility, and Accommodations Guidelines tutorials in Spanish and English, sections in the test administration manual dedicated to information about the test and scripts for test administrators, and an operational administration video.

Furthermore, like the other CAASPP assessments, to continue familiarizing students with the CSA items, ETS will develop and provide practice tests with a variety of item types that mirror the test length and grade levels on the operational CSA and are developed using the same standards as the operational assessment. The practice tests also include accessibility resources. It also can be used by parents or guardians, students, and LEA staff to determine if the CSA is appropriate for a student. LEAs will be encouraged to use the practice tests to prepare students to become more familiar with utilizing the technology and technology-enhanced items prior to taking the operational assessment.

6.3. Psychometric Analyses

After the 2016–17 CSA pilot test, composite items have been included on the forms for most grade levels except for grade five and high school. Each composite item has two parts. To better understand how composite items perform, not only in terms of overall performance but also with regard to each composite part, ETS conducted item analyses for both composite parts and composite items as a whole after the 2018 field test. It turned out that the composite-part item analyses contained a wealth of data to help assessment developers and teacher reviewers better understand the composite items. ETS will continue both types of analyses for the 2018–19 operational test.

6.4. Accessibility

ETS increased the number of accessibility resources available in the field test to match the upcoming operational assessment. In the interest of increasing the number of items that are “born accessible”—i.e., items that are as universally accessible as possible by all populations—ETS continues to investigate the construct-irrelevant use of item types that have no discernable difference from traditional test questions. As a result, ETS is reducing the development of both Grid single-select items and text-based Zone items, because they function much as multiple-choice items do.

Items also are reviewed by teachers of students with visual impairments separately, to ensure they are accessible to students with sensory disabilities.

Reference

Educational Testing Service. (2019). *Summary of results from the California Assessment of Student Performance and Progress California Spanish Assessment 2018–19 field test post-test survey*. Princeton, NJ: Educational Testing Service.