TECHNICAL DIGEST

2022–2023 School Year

Table of Contents

Chapter 1 Historical Overview of Assessment in Texas

Chapter 2 Building a High-Quality Assessment System

Chapter 3 Standard Technical Processes

Chapter 4 State of Texas Assessments of Academic Readiness

Chapter 5 STAAR Alternate 2

Chapter 6 Texas English Language Proficiency Assessment System

Chapter 7 TELPAS Alternate

Chapter 8 Resources

References

Appendix A Quality Control System

Appendix B STAAR Statistical Tables and Figures

Appendix C STAAR Alternate 2 Statistical Tables and Figures

Appendix D TELPAS Statistical Tables and Figures

Appendix E TELPAS Alternate Statistical Tables and Figures

TECHNICAL DIGEST 2022–2023

Chapter 1

Historical Overview of Assessment in Texas

Assessment Timeline

Changes to the Assessment Program Over Time

Assessment Timeline

Texas has a long history of student assessment dating back to 1979, when its first statewide testing program was required by statute. Over the years, changes in legislation and policy have impacted the size and scope of the Texas Assessment Program. This chapter provides an overview of these changes, including an assessment timeline and a description of changes to the assessment program over time.

—1979–1980

The Texas Assessment of Basic Skills (TABS) was administered for the first time in February 1980. TABS included mathematics, reading, and writing assessments in grades 3, 5, and 9. The final administration of TABS was in fall 1985.

—1986–1987

The Texas Educational Assessment of Minimum Skills (TEAMS) was first administered in fall 1986 and included mathematics, reading, and writing assessments in grades 1, 3, 5, 7, 9, and 11. TEAMS represented the first time that Texas students were required to pass a state assessment to be eligible to receive a high school diploma; students had to pass the TEAMS grade 11 exit-level tests in mathematics and reading to graduate. The final administration of TEAMS was in fall 1989. After that, students who were required to meet TEAMS graduation requirements had to take the Texas Assessment of Academic Skills (TAAS) exit-level assessments with adjusted performance standards.

—1990–1991

First administered in fall 1990, TAAS shifted the focus of assessment from minimum skills to academic skills and included mathematics, reading, and writing assessments in grades 3, 5, 7, 9, and 11. Students had to pass the TAAS grade 11 exit-level assessments in mathematics, reading, and writing to receive their high school diploma.

—1993–1994

Administration of TAAS moved to the spring, and the grades and subjects assessed were reconfigured. From 1994 to 2002, TAAS was administered every spring to students in grades 3–8 and 10 in mathematics and reading; grades 4, 8, and 10 in writing; and grade 8 in science and social studies. Students had to pass the TAAS grade 10 exit-level assessments in mathematics, reading, and writing to be eligible to graduate.

The final administration of TAAS for grades 3–8 was in spring 2002. Because TAAS remained the graduation requirement for students in grade 9 or above on January 1, 2001, exit-level TAAS tests continued to be administered through July 2009. Subsequently, students who were required to meet TAAS graduation requirements were able to take Texas Assessment of Knowledge and Skills (TAKS) exit-level assessments with adjusted performance standards.

—1995–1996

Spanish-language TAAS mathematics and reading assessments were incorporated into the testing program for grades 3 and 4.

Algebra I and Biology end-of-course (EOC) assessments were administered for the first time to students who completed these courses.

—1996–1997

Spanish-language TAAS mathematics and reading assessments were incorporated into the testing program for grades 5 and 6.

—1998–1999

English II and U.S. History EOC assessments were administered for the first time to students who completed these courses. Through spring 2002, the four EOC assessments—Algebra I, English II, Biology, and U.S. History—were administered as state-mandated assessments and as an option for meeting graduation requirements.

—1999–2000

Reading Proficiency Tests in English (RPTE)

The Reading Proficiency Tests in English (RPTE) were first administered in spring 2000 to emergent bilingual (EB) students in grades 3–12.

—2000–2001

State-Developed Alternative Assessment (SDAA)

The State-Developed Alternative Assessment (SDAA) was first administered in spring 2001 to eligible students receiving special education services. SDAA included assessments in mathematics and reading for kindergarten through grade 8 and in writing for kindergarten through grade 7. The final administration of SDAA was in spring 2004.

—2002–2003

TAKS Texas Assessment of Knowledge and Skills

To satisfy legislative requirements, TAKS was designed to be more comprehensive than its predecessors and to measure more of the state-mandated curriculum known as the Texas Essential Knowledge and Skills (TEKS). TAKS was first administered in spring 2003 and included assessments in mathematics in grades 3–11; reading in grades 3–9; writing in grades 4 and 7;

English language arts (ELA) in grades 10 and 11; science in grades 5, 10, and 11; and social studies in grades 8, 10, and 11. Spanish versions of TAKS were administered in grades 3–6. Students had to pass the TAKS grade 11 exit-level tests in mathematics, ELA, science, and social studies to receive a high school diploma.

In compliance with the Student Success Initiative (SSI), satisfactory performance on TAKS grade 3 reading, grade 5 mathematics and reading, and grade 8 mathematics and reading assessments were requirements for promotion to the next grade level. These requirements became effective for grade 3 in the 2002–2003 school year, grade 5 in the 2004–2005 school year, and grade 8 in the 2007–2008 school year. The TAKS grade 3 reading promotion requirements were removed beginning with the 2009–2010 school year.

The final administration of TAKS for grades 3–10 was in spring 2011. Because TAKS remained the graduation requirement for students in grade 9 or above in the 2011–2012 school year, exit-level TAKS tests continued to be administered through June 2017. After that, students who were required to meet TAKS graduation requirements could take the State of Texas Assessments of Academic Readiness (STAAR®) EOC assessments with adjusted performance standards.

—2003–2004

To fulfill requirements of the federal No Child Left Behind Act (NCLB), the Texas Observation Protocol (TOP) was developed to assign holistic English language proficiency ratings for students based on observations during instruction. Holistic ratings were developed in the language domains of listening, speaking, and writing in kindergarten through grade 12 and in reading in kindergarten through grade 2.

TELPAS
Texas English Language Proficiency Assessment System

Together, TOP and RPTE formed the Texas English Language Proficiency Assessment System (TELPAS).

—2004–2005

LAT
Linguistically Accommodated Testing

In response to NCLB regulations, a linguistically accommodated testing (LAT) process was added to TAKS grades 3–8 and 10 mathematics for eligible EB students.

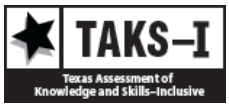
State-Developed Alternative Assessment II (SDAA II)

SDAA was replaced with SDAA II in spring 2005 to better align the alternate assessment to TAKS. SDAA II was available for students who received special education services in mathematics in kindergarten through grade 10, reading in kindergarten through grade 9, writing in kindergarten through grade 9, and ELA in grade 10. The final administration of SDAA II was in spring 2007.

In response to the 2004 Algebra Incentive Program and Executive Order RP53, the Algebra I EOC assessment was revised and made available online in spring 2005.

—2005–2006

Based on legislative requirements, TAKS grade 8 science was added to the testing program.



To meet the requirements of the Individuals with Disabilities Education Act (IDEA) of 2004, TAKS–Inclusive (TAKS–I) was added to the assessment program in spring 2006. TAKS–I was available for eligible students receiving special education services and included science in grades 5, 8, 10 and 11; science in Spanish in grade 5; social studies in grades 8, 10, and 11; and mathematics and ELA in grade 11. The final administration of TAKS–I was in spring 2007.

—2006–2007

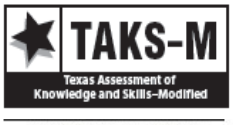
LAT administrations of TAKS grades 3–8 reading and grade 10 ELA were implemented in spring 2007 for eligible EB students.

—2007–2008

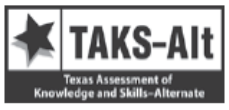
LAT administrations of TAKS grades 5, 8, and 10 science were implemented in spring 2008 for eligible EB students.

TAKS
(Accommodated)

TAKS (Accommodated) replaced TAKS–I for students receiving special education services who met the eligibility requirements for specific accommodations. TAKS (Accommodated) was available for mathematics in grades 3–11; reading in grades 3–9; writing in grades 4 and 7; ELA in grades 10 and 11; science in grades 5, 8, 10, and 11; and social studies in grades 8, 10, and 11. The final administration of TAKS (Accommodated) was in spring 2011.



TAKS–Modified (TAKS–M) was an alternate assessment based on modified academic achievement standards and was first administered in spring 2008. TAKS–M was available for eligible students receiving special education services and included mathematics in grades 3–11; reading in grades 3–9; writing in grades 4 and 7; ELA in grades 10 and 11; science in grades 5, 8, 10, and 11; and social studies in grades 8, 10, and 11. The final administration of TAKS–M was in spring 2011.



To fulfill federal requirements, TAKS–Alternate (TAKS–Alt) was first administered in spring 2008. It was developed for students with significant cognitive disabilities and was based on alternate achievement standards. TAKS–Alt included mathematics in grades 3–11; reading in grades 3–9; writing in grades 4 and 7; ELA in grades 10 and 11; science in grades 5, 8, 10, and 11; and social studies

in grades 8, 10, and 11. The final administration of TAKS-Alt was in spring 2011.

Based on NCLB requirements, TELPAS reading for grades 2–12 was redesigned and administered as an online testing program beginning in spring 2008.

EOC assessments in Geometry and Biology were first administered on a voluntary basis.

—2008–2009

Based on legislation, TAKS grade 6 assessments in Spanish were administered for the final time in spring 2009.

EOC assessments in Chemistry and U.S. History were first administered on a voluntary basis.

—2009–2010

EOC assessments in Physics and World Geography were first administered on a voluntary basis.

—2010–2011

EOC assessments in Algebra II and English I were first administered on a voluntary basis.

—2011–2012

STAAR replaced TAKS as the state academic assessment program beginning in spring 2012. STAAR included mathematics and reading in grades 3–8, writing in grades 4 and 7, science in grades 5 and 8, and social studies in grade 8. For high school, grade-specific assessments were replaced with 15 STAAR EOC assessments: Algebra I, Geometry, Algebra II, English I reading, English I writing, English II reading, English II writing, English III reading, English III writing, Biology, Chemistry, Physics, World Geography, World History, and U.S. History. STAAR Spanish was administered in grades 3–5.

In compliance with SSI, satisfactory performance on STAAR grades 5 and 8 mathematics and reading were requirements for promotion to the next grade level through spring 2021.

Depending on their graduation program, high school students were required to meet the passing standard (or achieve a predetermined minimum score) on at least 11 of the 15 STAAR EOC assessments. Additionally, students needed to meet a cumulative score requirement in each content area.

STAAR L, a linguistically accommodated English version of STAAR, was first administered online in spring 2012. STAAR L was available for EB students for

grades 3–8 and EOC assessments in mathematics, science, and social studies. The final administration of STAAR L was in fall 2016.

STAAR Modified replaced TAKS–M beginning in spring 2012. STAAR Modified originally included mathematics and reading in grades 3–8, writing in grades 4 and 7, science in grades 5 and 8, and social studies in grade 8. The final administration of STAAR Modified was in spring 2014.

STAAR Alternate replaced TAKS–Alt in spring 2012. STAAR Alternate included mathematics and reading in grades 3–8, writing in grades 4 and 7, science in grades 5 and 8, social studies in grade 8, and EOC assessments in Algebra I, Geometry, English I reading, English I writing, English II reading, English II writing, English III reading, English III writing, Biology, World Geography, World History, and U.S. History. The final administration of STAAR Alternate was in spring 2014.

—2012–2013

Based on legislative changes, spring 2013 was the final administration of STAAR Geometry, Chemistry, Physics, World Geography, and World History EOC assessments. STAAR Algebra II and English III post-secondary readiness assessments became optional, and their administration was suspended until spring 2016.

STAAR Modified EOC assessments in Algebra I, Geometry, English I reading, English I writing, English II reading, English II writing, Biology, World Geography, and World History were added to the testing program.

—2013–2014

Based on legislative requirements, STAAR high school English assessments were redesigned to combine reading and writing into a single assessment. The redesigned STAAR English I and English II EOC assessments were first administered in spring 2014.

The STAAR Modified U.S. History EOC assessment was added to the testing program.

—2014–2015

STAAR A was administered online for the first time in spring 2015 with embedded accommodations designed to help students who met eligibility requirements access the content being assessed. STAAR A was available for mathematics and reading in grades 3–8, writing in grades 4 and 7, science in grades 5 and 8, social studies in grade 8, and EOC assessments in Algebra I, English I, English II, Biology, and U.S. History. The final administration of STAAR A was in fall 2016.

STAAR Alternate 2 was administered for the first time in spring 2015 to eligible students with the most significant cognitive disabilities. STAAR Alternate 2 includes assessments for mathematics and reading in grades 3–8, science in grades 5 and 8, social studies in grade 8, and EOC assessments in Algebra I, English I, English II, Biology, and U.S. History.

—2015–2016

STAAR Algebra II and English III post-secondary readiness assessments were administered as optional assessments from spring 2016 through spring 2021.

—2016–2017

STAAR online with embedded supports replaced STAAR A and STAAR L beginning with the spring 2017 administration. This change allowed for a wider range of accessibility features and accommodations, including content supports and language and vocabulary supports, based on each student's needs.

—2017–2018

TELPAS listening and speaking holistic assessments for grades 2–12 were combined and made into standardized item-based assessments to be administered online. In addition, the blueprint for TELPAS reading was shortened.

New optional STAAR Interim Assessments were offered for grades 3–8 mathematics and reading, Spanish grades 3–5 mathematics and reading, and EOC assessments in Algebra I, English I, and English II.

—2018–2019

In the 2018–2019 school year, STAAR was administered online for the first time in American Sign Language (ASL) and refreshable braille. STAAR Spanish grades 3–5 assessments were offered online for the first time.

TELPAS Alternate was first administered in spring 2019 to EB students in grades 2–12 with the most significant cognitive disabilities. TELPAS Alternate is a holistic assessment process that includes the English language domains of listening, speaking, reading, and writing.

—2019–2020

In response to the COVID-19 pandemic, the Texas Education Agency (TEA) launched optional end-of-year (EOY) assessments that school systems and parents could choose to administer free of charge in the absence of STAAR to evaluate the academic progress students made.

—2020–2021

TEA launched optional beginning-of-year (BOY) assessments that school systems could choose to administer free of charge to evaluate the academic progress students made. BOY assessments were available each fall through the 2022–2023 school year.

—2021–2022

A braille version of TELPAS reading was available for the first time for students with visual impairments.

—2022–2023

STAAR transitioned to a primarily online assessment program beginning with the December 2022 administration.

Spring 2023 marked the launch of the STAAR redesign. New non-multiple-choice question types were present across all grades, subjects, and courses. STAAR reading language arts (RLA) assessments included reading and writing components.

TELPAS writing in grades 2–12 moved from a holistic assessment to a standardized item-based assessment administered online and was combined with the reading assessment beginning in spring 2023.

Changes to the Assessment Program Over Time

The Texas Assessment Program must comply with federal regulations and state statutes concerning student assessment. Federal regulations are mandated by NCLB, the Elementary and Secondary Education Act (ESEA), and the Every Student Succeeds Act (ESSA). The majority of state law pertaining to the statewide student assessment program is found in Texas Education Code (TEC) Chapter 39, Subchapter B.

The Texas Assessment Program measures students' understanding of the statewide curriculum. When the statewide curriculum is revised, changes are subsequently made to the assessment program to maintain a strong, direct, and effective link between the statewide curriculum and the state assessments.

The following provides a summary of changes in law and in the statewide curriculum that affected the Texas Assessment Program.

1979

The Texas Assessment Program began in 1979 when the 66th Texas Legislature enacted Senate Bill (SB) 350, which required basic skills competencies in mathematics, reading, and writing for grades 3, 5, and 9. As a result of SB 350, TABS was implemented in 1980.

1981

House Bill (HB) 246, passed by the 67th Texas Legislature, Regular Session, 1981, made changes to the state curriculum. As a result, the State Board of Education (SBOE) adopted the Essential Elements in 1984.

1984

HB 72, passed by the 68th Texas Legislature, Second Called Session, 1984, called for accountability provisions in exit-level testing requirements. HB 72 also led to the implementation of TEAMS, which replaced TABS in 1986.

1991

In 1991, the 72nd Texas Legislature passed SB 7, which required statewide testing of students in grades 3–8 and exit-level tests in high school. As a result, TEAMS was replaced by TAAS, which was administered from 1990 to 2002.

1995

Enacted by the 74th Texas Legislature in 1995, SB 1 overhauled the TEC and required the development of four EOC assessments. Students could use satisfactory performance on the Algebra I, the English II, and either the Biology or the U.S. History EOC assessment in place of TAAS to meet graduation assessment requirements.

1997

In July 1997, the SBOE replaced the Essential Elements with the TEKS. Implemented as the statewide curriculum for Texas in the 1998–1999 school year, the TEKS were developed to be more specific and focused than the Essential Elements, with emphasis placed on the knowledge and skills students were expected to learn rather than on the delivery standards expected of teachers.

1999

In 1999, the 76th Texas Legislature passed SB 103, which required the development of TAKS to replace TAAS. SB 103 also required the development of a system to assess the reading proficiency and language acquisition of EB students, resulting in the development of RPTE.

SSI, enacted by the Texas Legislature in 1999, made satisfactory performance on the grade 3 reading assessment, the grade 5 mathematics and reading assessments, and the grade 8 mathematics and reading assessments a promotion requirement for Texas students. The first cohort of students affected by this law was the grade 3 class of 2002–2003. Passing the grade 5 mathematics and reading assessments was a promotion requirement for the first time in the 2004–2005 school year. Grade 8 promotion requirements became effective in the 2007–2008 school year. In 2009, the Texas Legislature amended SSI to remove the grade 3 promotion requirement.

2005

In response to the governor's 2004 Algebra Incentive Program, the Algebra I EOC assessment was revised and administered online in spring 2005 on a voluntary basis to students who completed Algebra I coursework.

Executive Order RP53, issued by the governor in December 2005, called for increased college readiness programs in Texas schools and authorized the development of a series of EOC assessments in subjects assessed by TAKS in grade 11, including Algebra I, Geometry, Biology, Chemistry, Physics, and U.S. History.

2007

In May 2007, the 80th Texas Legislature enacted SB 1031, which required the implementation of an EOC assessment program. With the expanded role of the EOC assessment program, SB 1031 phased out TAKS grade level–based testing in high school and replaced it with EOC assessments as a component of the new high school graduation requirements that would apply beginning with the incoming freshman class of 2011–2012. The bill required the development of EOC assessments for Algebra I, Geometry, Algebra II, English I, English II, English III, Biology, Chemistry, Physics, World Geography, World History, and U.S. History.

HB 1, also passed in 2007, required TEA and the Texas Higher Education Coordinating Board (THECB) to develop the College and Career Readiness Standards (CCRS). After the CCRS were developed, TEA and THECB linked the CCRS to the TEKS in mathematics, RLA, science,

and social studies. Finally, as part of the TEKS review process, the SBOE incorporated the CCRS into the TEKS, making Texas the first state in the country to adopt a curriculum aligned to college and career readiness.

The SBOE adoption of new English Language Proficiency Standards (ELPS) for EB students in kindergarten through grade 12 was effective in December 2007. Beginning in 2008, TELPAS was aligned to the new ELPS.

2009

In 2009, the 81st Texas Legislature, Regular Session, enacted HB 3, which made further changes to the assessment program. HB 3 required that the performance standards for mathematics and reading assessments in grades 3–8 be linked from grade to grade to the college readiness performance standards for the Algebra II and English III assessments. The required vertical linking, along with the replacement of exit-level TAKS with EOC assessments, necessitated the design of a new series of assessments to indicate college readiness. As a result, STAAR was developed to encompass the EOC assessments mandated by SB 1031 in 2007 and the grades 3–8 assessments mandated by HB 3.

HB 3 also required the commissioner of education, rather than the SBOE, to determine performance levels for assessments and eliminated the exit-level TAAS assessments. As a result, students who had been required to meet TAAS graduation requirements could take TAKS exit-level assessments with adjusted performance standards.

2010

In 2010, the SBOE adopted revised social studies TEKS; alignment with those TEKS was reflected in the 2011–2012 STAAR social studies assessments.

2011

In 2011, the 82nd Texas Legislature, Regular Session, passed HB 2135, which impacted students receiving above-grade-level instruction. The bill allowed students who were enrolled in and taking the assessment for an above-grade-level course to not take the grade-level assessment. Additionally, the bill indicated that a student in an SSI grade could not be denied promotion based on performance on an assessment if the student was taking an above-grade-level assessment instead of the grade-level assessment.

2012

In 2012, the SBOE adopted new mathematics TEKS; alignment with the new TEKS was reflected in the spring 2015 STAAR grades 3–8 mathematics assessments and in the spring 2016 STAAR Algebra I and Algebra II assessments.

2013

In 2013, the 83rd Texas Legislature, Regular Session, enacted HB 5, which reduced the number of STAAR EOC assessments required for graduation from 15 to five: Algebra I, English I, English II, Biology, and U.S. History. The administration of Algebra II and English III was suspended until the 2015–2016 school year and became optional for districts. In addition, the separate reading and writing assessments for the high school English courses were required to be combined into a single assessment for each course with a single reported score. HB 5 removed the requirement to provide a cumulative and minimum score and to include the STAAR EOC assessment results as 15 percent of a course grade.

HB 5 also required changes to the administration of STAAR Alternate, and SB 906 required changes to the performance standards for STAAR Alternate. Based on both bills, STAAR Alternate was redesigned, and STAAR Alternate 2 was administered for the first time in spring 2015.

2015

In 2015, the 84th Texas Legislature passed several bills that affected the assessment program. SB 149 allowed students to qualify for graduation through an individual graduation committee (IGC) beginning in the 2014–2015 school year.

As required by HB 1164 that year, TEA completed a pilot study to examine alternative methods of assessing writing. The pilot study included the collection and scoring of a range of student writing samples produced throughout the school year.

Also passed in 2015, HB 743 required that STAAR be designed so that 85 percent of students taking an assessment in grades 3–5 could complete a test in two hours and 85 percent of students taking an assessment in grades 6–8 could complete the assessment in three hours. In response to HB 743, TEA redesigned the grades 3–8 assessments by reducing the total number of questions and the number of field-test questions on most assessments and redesigned the two-day grades 4 and 7 writing tests as single-day tests that could be completed in a four-hour administration.

The legislature also passed HB 2349, which revised the state's assessment requirements for graduation. Effective beginning with the 2015–2016 school year, a student who earned high school credit for a course for which there was an EOC assessment prior to enrolling in a Texas public school and for which a Texas public school district accepted the credit was not required to take that EOC assessment to receive a Texas diploma. Additionally, HB 2349 required a school district or charter school to report to TEA whether a student assessed with STAAR transferred into a Texas school or district from out of state during the current school year so those students could be excluded in accountability calculations.

2017

In 2017, the SBOE adopted new English and Spanish RLA TEKS, which were implemented in the STAAR RLA assessments beginning in spring 2022. The SBOE also adopted streamlined

TEKS for science, which were first reflected in the STAAR science assessments in December 2018.

2018

In 2018, the SBOE adopted streamlined TEKS for social studies. The streamlined TEKS were first reflected in the 2019–2020 STAAR social studies assessments.

2019

In 2019, the 86th Texas Legislature passed HB 3906, which addressed several components of the assessment program. The bill's key measures included eliminating the STAAR grades 4 and 7 writing assessments, developing a transition plan to administer all STAAR assessments online by the 2022–2023 school year, establishing a cap of no more than 75 percent multiple-choice questions on any STAAR assessment, codifying STAAR Interim Assessments, and developing an integrated formative assessment pilot.

Additionally, HB 1244 required that the STAAR U.S. History EOC assessment include 10 questions randomly selected from the civics test administered by the United States Citizenship and Immigration Services (USCIS). The 10 questions selected were required to align with the TEKS for United States History Studies Since 1877 and were added in the 2019–2020 school year.

2020

In response to the COVID-19 pandemic in spring 2020, the governor used his statutory authority to suspend annual academic assessment requirements for the remainder of the 2019–2020 school year. Therefore, STAAR was not administered in spring or summer 2020, and specific STAAR EOC assessment requirements for graduation were waived for students enrolled in and completing the corresponding course. STAAR Alternate 2 was not administered. Since the administrations of TELPAS and TELPAS Alternate had already begun, completion of these assessments was made optional for districts. TEA received approval from the U.S. Department of Education to waive statewide assessment and accountability requirements for the 2019–2020 school year.

In addition, SSI requirements were waived for the 2020–2021 school year, so retests for STAAR grades 5 and 8 mathematics and reading were not administered.

2021

In 2021, the 87th Texas Legislature, Regular Session, passed HB 4545, which eliminated the grade retention and retesting requirements associated with SSI and established new requirements for accelerated instruction for students who do not pass STAAR.

2023

The 88th Texas Legislature, Regular Session, 2023, passed HB 1225, which permitted districts to provide paper administrations of STAAR to any student whose parent, guardian, or teacher requests it. Requests must be submitted to the district by September 15 for fall administrations and December 1 for spring administrations. The number of students who are administered paper by request is limited to three percent of the total number of students enrolled in the district and is separate and distinct from the students who are eligible to receive a special paper administration of STAAR.

Also passed in 2023, HB 1883 allowed districts to consider the dates of religious holy days likely to be observed by their students when establishing district calendars and days within the testing windows on which students are administered state assessments. For the bill's purposes, holy days were defined as those observed by a religion whose places of worship are exempt from property taxation under Section 11.20 of the Tax Code. HB 1883 required districts to provide alternative dates within the testing window for students who are absent from school on scheduled testing dates to observe a religious holy day. As a result, districts are required to provide make-up testing opportunities for religious holy days observed by students.

TECHNICAL DIGEST 2022–2023

Chapter 2

Building a High-Quality Assessment System

Test Development Activities

Groups Involved

Item Development and Review

Pilot Testing

Field Testing and Data Review

Security

Quality-Control Procedures

Performance Assessments

Test Development Activities

Texas educators, including kindergarten through grade 12 classroom teachers, higher education representatives, curriculum specialists, administrators, and Education Service Center (ESC) staff, play a vital role in every phase of the test development process. Thousands of Texas educators have served on one or more of the educator committees involved in the development of the Texas Assessment Program, including STAAR, STAAR Alternate 2, TELPAS, and TELPAS Alternate. These committees are intended to represent the state geographically, ethnically, by gender, and by type and size of school district. While there are slight differences in the processes for developing different assessments, Table 2.1 outlines the procedures used to develop a test framework and provide for ongoing development of test items for the Texas Assessment Program.

Table 2.1. Test Development Process

Step	Process
1	Committees of Texas educators review the state-mandated curriculum, the TEKS, or the ELPS to develop appropriate assessment categories for a specific grade and subject, course, or domain that is assessed. For each grade and subject, course, or domain, educators provide advice on an assessment model that aligns with best practices in classroom instruction.
2	Educator committees work with TEA both to prepare draft test reporting categories and to determine how these categories would best be assessed. These preliminary recommendations are reviewed by classroom teachers, higher education representatives, curriculum specialists, and assessment specialists.
3	A draft of the reporting categories and TEKS or ELPS student expectations to be assessed is refined based on input from Texas educators. TEA begins to gather statewide opportunity-to-learn information.
4	Prototype test questions are written to measure each reporting category and, when necessary, are pilot-tested with Texas students from volunteer classrooms.
5	Educator committees assist in developing guidelines for assessing each reporting category. These guidelines outline the eligible test content and test question formats and include sample items.
6	With educator input, a preliminary test blueprint is developed that sets the number of questions on the test and the number of test questions measuring each reporting category.
7*	Professional item writers develop test items based on the reporting categories, the TEKS or ELPS student expectations, and the item guidelines.
8*	TEA content specialists review and revise the proposed test items.
9*	Item review committees composed of Texas educators review the revised test items to judge the appropriateness of item content and difficulty and to eliminate potential bias.
10*	Test questions are revised again based on input from Texas educator committees and are then field-tested with large representative samples of Texas students.
11*	Technical processes are used to analyze field-test data for reliability, validity, and possible bias.
12*	Data reviews are held to determine whether items are appropriate for inclusion in the item bank from which test forms are built.
13	A final blueprint for each test is developed to establish the number of questions on the test and the number of test questions measuring each reporting category.

Step	Process
14*	All accepted field-test items and data are entered into an item bank. Tests are built from the item bank so that the tests are comparable in difficulty and content from one administration to the next.
15*	Content validation panels composed of university-level experts in each content area review the EOC assessments for accuracy because of the advanced level of content being assessed.
16*	Tests are administered to Texas students.
17*	Stringent quality control (QC) measures are applied to all stages of developing, administering, scoring, and reporting for both online and paper assessments. Test results are reported at the student, campus, district, regional, and state levels.
18	In accordance with state law, the Texas Assessment Program releases tests to the public.
19	In accordance with state law, the commissioner of education uses impact data, study results, and statewide opportunity-to-learn information, along with recommendations from standard-setting panels, to set a passing standard for state assessments.
20	A technical digest is developed and published annually to provide verified technical information about the tests.

*For STAAR, STAAR Alternate 2, and TELPAS, these steps are repeated annually to ensure that tests of the highest quality are developed.

Groups Involved

The entities described below perform crucial functions in the test development process, and their collaborative efforts significantly contribute to the quality of the Texas Assessment Program.

Assessment Development Division

The TEA Assessment Development Division is composed of content experts and psychometricians. The content experts provide content expertise during the item development and test development processes for all statewide assessments. The psychometricians are responsible for ensuring that assessments meet reliability and validity requirements for a sound assessment system.

Student Assessment Division

The TEA Student Assessment Division is responsible for implementing the provisions of federal and state law for the state assessment program. The Student Assessment Division oversees the planning, scheduling, administration, scoring, and reporting of all major assessment activities. TEA staff members in this division conduct QC activities for the administration, scoring, and reporting of the assessment program.

Cambium Assessment, Inc.

Cambium Assessment, Inc. (CAI) is the test administration, scoring, and reporting contractor for the Texas Assessment Program. CAI also serves as the program integration contractor. This

role includes working with Pearson to make sure that the Texas Assessment Program is managed in accordance with TEA requirements.

Pearson

Pearson is TEA's primary item development contractor. Due to the diverse nature of the services required, Pearson employs highly qualified assessment specialists and independent contractors with diverse experience teaching and assessing students.

Texas Educators

When a new assessment is developed, committees of Texas educators review the state-mandated curriculum, help determine appropriate reporting categories, and provide input on the appropriate alignment of the assessment items to the curriculum standards.

Teachers, curriculum specialists, assessment specialists, and administrators review draft reporting categories with the corresponding TEKS or ELPS student expectations. Texas educator committees assist in the review and revision of the eligible TEKS or ELPS documents that outline the student expectations eligible for assessment. TEA staff members then revise and finalize these draft reporting categories and eligible TEKS or ELPS documents based on input from Texas educators.

Following the development of test items by professional item writers, committees of Texas educators review the items to ensure appropriate content alignment and level of difficulty and to eliminate potential bias. Items are revised based on this input and then field-tested.

Item Development and Review

Pearson assumes the major role for STAAR (including STAAR Spanish), STAAR Alternate 2, and TELPAS item development, and TEA personnel are involved throughout the item development process.

Item Guidelines

Item and performance task specifications provide guidance to item writers on how to translate the TEKS or ELPS into assessment items. Item writers strictly follow these guidelines to ensure the accurate measurement of the TEKS or ELPS student expectations. In addition, guidelines for universal design, bias and sensitivity, accessibility and accommodations, and style help item writers and reviewers establish consistency in the development of test items.

Item Writers

Pearson and its subcontractors employ item writers with extensive experience developing items for standardized achievement tests, large-scale criterion-referenced measurements, and English language proficiency tests. These individuals are selected based on their content-area knowledge, their teaching or curriculum development experience in the relevant grades, or their experience teaching EB students or students with special needs.

For each STAAR (including STAAR Spanish), STAAR Alternate 2, and TELPAS assessment, TEA receives an item inventory indicating the number of test items to be developed for each reporting category and TEKS student expectation (for STAAR and STAAR Alternate 2) or ELPS student expectation (for TELPAS). Item inventories are used throughout the item review process. If necessary, Pearson develops additional items to provide the requisite number of items per student expectation.

For TELPAS Alternate, Texas educators developed the Observable Behaviors during a series of TEA-led meetings. Guided by TEA and Pearson staff, the educators created an inventory of items that align to the ELPS and cover the alternate proficiency level descriptors (PLDs).

Training

Pearson provides extensive training for item writers. Before item development begins, Pearson reviews in detail the content expectations and item specifications for the applicable assessment program and discusses the scope of the testing program, security issues, adherence to the measurement specifications, and avoidance of possible economic, regional, cultural, gender, or ethnic bias.

Contractor Review

Pearson staff members who are content experts in the grades and subject areas for which items are developed participate in the review of each set of newly developed items. The review includes a check for content accuracy and item fairness for various demographic groups. Pearson reviewers also consider the alignment between the items and the reporting categories, range of difficulty, clarity, accuracy of correct answers, and plausibility of incorrect answer choices (or distractors) as well as the more global issues of universal design, passage appropriateness, passage difficulty, readability measures, interactions among items, and appropriateness of artwork, graphics, or charts. Pearson editorial staff members examine the items before submission to TEA for review.

TEA Review

TEA Assessment Development Division staff members who are content experts in the grades and subject areas for which items are developed review each item to verify alignment to a particular student expectation in the TEKS or ELPS, grade appropriateness, clarity of wording, content accuracy, plausibility of the distractors, accessibility, and identification of any potential economic, regional, cultural, gender, or ethnic bias. TEA staff members provide edits and meet with Pearson to discuss the progress of the reviews before each item review committee meeting.

Item Review Committee

Each year the TEA Assessment Development Division convenes committees composed of Texas classroom teachers (including general education teachers, special education teachers, and bilingual and English as a second language [ESL] teachers) and curriculum specialists to work with TEA staff in reviewing newly developed test items.

TEA seeks recommendations for item review committee members from superintendents and other district administrators, district curriculum specialists, ESC executive directors and staff members, and staff from other agency divisions. In addition, TEA has developed an educator committee application database where educators can self-nominate to participate on TEA educator committees. Item review committee members are selected based on their established expertise in a content area or in second-language acquisition. Committee members are selected to represent the 20 ESC regions and various types of districts (e.g., urban, suburban, rural, large, small) in Texas, as well as the major ethnic groups in the state.

TEA staff works with Pearson and its subcontractors to train committee members on the proper procedures and criteria for reviewing newly developed items. Committee members judge each item for alignment, appropriateness, adequacy of student preparation, and any potential bias. Committee members discuss each test item and recommend whether it should be field-tested as written or revised, recoded to a different TEKS or ELPS student expectation, or rejected. In their reviews, committee members consider the effect any item may have on various student populations and work toward eliminating potential bias against any group. Table 2.2 shows the guidelines that item review committee members follow.

Table 2.2. Item and Passage Review Guidelines

Category	Guidelines
Reporting Category/Student Expectation Item Match	• The item measures what it is supposed to assess. • The item poses a clearly defined problem or task.
Appropriateness (Interest Level)	• The item or passage is well written and clear. • The point of view is relevant to students taking the test. • The subject matter is of fairly wide interest to students at the grade being tested. • The artwork is clear, correct, and appropriate.
Appropriateness (Format)	• The format is appropriate for the intended grade. • The format is interesting to the student. • The item is formatted so it is not unnecessarily difficult.
Appropriateness (Answer Choices)	• The answer choices are reasonably parallel in structure. • The answer choices are worded clearly and concisely. • The answer choices do not eliminate each other. • There is only one correct answer.
Appropriateness (Difficulty of Distractors)	• Each distractor is plausible. • There is a rationale for each distractor. • Each distractor is relevant to the knowledge and understanding being measured. • Each distractor is at a difficulty level appropriate for both the objective and the intended grade.
Opportunity to Learn	• The item is a good measure of the curriculum. • The item is suitable for the grade or course.

Category	Guidelines
Sensitivity Concerns and Freedom from Bias	• The item or passage does not assume racial, class, or gender values or suggest such stereotypes. • The item does not provide an advantage or disadvantage to any group of students because of their personal characteristics, such as race, gender, socioeconomic status, or religion. • The item or passage avoids needless reference to topics that are extremely controversial or upsetting. • The item or passage addresses sensitive topics in a careful, fair, and balanced way. • The item fairly represents cultural, ethnic, social, and political diversity.

TEA field-tests the recommended items to collect student responses from representative samples across the state. Items rejected by the item review committee are not considered for field testing.

Annual item review committees are not convened for TELPAS Alternate because the TELPAS Alternate Observable Behaviors that were written and revised by educators during the development of the assessment are used every year.

Pilot Testing

The purpose of pilot testing is to gather information about test item prototypes and administration logistics for a new assessment and to refine item development guidelines as needed. Pilot testing can be conducted to accomplish varying objectives. If the purpose is to gather information about test items of differing types and ranges of difficulty, the pilot test might occur before the extensive item development process described above. If the purpose is to gather information about test administration logistics, the pilot test might occur after major item development but before field testing.

Field Testing and Data Review

Field-Test Procedures

Items are field-tested before they are used on an operational test form. Whenever possible, TEA conducts field tests of new items by embedding them in spring operational tests so that the field-test items are randomly distributed across the state. This procedure ensures that a large representative sample of responses is gathered on each item. Experience has shown that embedded field testing yields sufficient data for precise item evaluation and allows for the collection of statistical data on a large number of items in a realistic testing situation. (Performance on field-test items does not affect students' scores on the operational tests.) TEA also periodically conducts stand-alone field tests of new items (e.g., extended constructed-response items) by administering them to a purposefully selected representative student sample. Refer to Chapter 4, "STAAR," for detailed information about stand-alone field testing.

Typically, for STAAR grades 3-8, six field-test questions are embedded in each form for mathematics, RLA, science, and social studies. For spring STAAR EOC assessments, 13 field-

test questions are embedded in each English I and English II form, eight are embedded in each Algebra I and Biology form, and four are embedded in the U.S. History form.

For STAAR Alternate 2, at least four field-test questions are embedded in each form for all grades and subjects and courses assessed.

For TELPAS, at least seven field-test questions are embedded in each form for listening and speaking and for reading and writing.

The Observable Behaviors for TELPAS Alternate are the same each year; therefore, TELPAS Alternate does not include field-test questions.

To ensure that each item is examined for potential ethnic bias, the sample selection is designed so that the proportions of African American and Hispanic students in the samples are representative of their respective total student populations in Texas. Data obtained from the field test include:

- the number of students by ethnicity and gender in each sample;
- the percentage of students choosing each response for multiple-choice questions or obtaining each score point for non-multiple-choice questions;
- the percentage of students, by gender and by ethnicity, choosing each response for multiple-choice questions or obtaining each score point for non-multiple-choice questions;
- point-biserial correlations to determine the relationship between a student's score on a particular test item and the score obtained on the total assessment;
- Rasch statistical indices to determine the relative difficulty of each test item; and
- Mantel-Haenszel statistics for dichotomous items and standardized mean difference (SMD) for constructed-response items to identify, by gender and ethnicity, greater-than-expected differences in group performance on any single item.

Data Review Procedures

After field testing, TEA content development specialists provide feedback to Pearson on each test item and its associated data regarding reporting category and student expectation match, appropriateness, level of difficulty, and potential gender, ethnic, or other bias. They then recommend acceptance or rejection of each field-test item. Items that pass all stages of development—item review, field testing, and data review—are placed in the item bank and become eligible for use on future test forms. Rejected items are marked as such and eliminated from consideration for use on any summative assessment.

Item Bank

The item bank maintained by CAI for the Texas Assessment Program stores each test item, accompanying artwork, and item data such as the unique item number (UIN), grade and subject or course, reporting category, TEKS or ELPS student expectation measured, dates the item was

administered, and item statistics. The item bank also contains information obtained during data review meetings that specifies whether a test item is acceptable for use. During the test construction process, TEA, CAI, and Pearson use item statistics and other item information to maintain consistent test difficulty and adjust tests for content coverage and balance.

Test Construction

Each grade and subject and course assessment is based on a specific test blueprint that guides how each test is constructed. Test blueprints delineate the number of items and points from each reporting category that will appear on a given test. STAAR, including STAAR Spanish, focuses on the TEKS that are most critical by incorporating readiness and supporting standards into the test blueprints. Readiness standards are emphasized annually; supporting standards, while eligible for assessment as an important part of instruction, may not be tested each year. All decisions about the relative emphasis of each reporting category are based on feedback from Texas educators and are indicated in the assessed curriculum and blueprint documents found on the STAAR Resources webpage. General characteristics of readiness and supporting standards are shown in Table 2.3.

Table 2.3. Comparison of Readiness and Supporting Standards

Readiness Standards	Supporting Standards
• are essential for success in the current grade or course • are important for preparedness for the next grade or course • support college and career readiness • necessitate in-depth instruction • address broad and deep ideas	• may be introduced in the current grade or course and emphasized in a subsequent year • may be reinforced in the current grade or course and emphasized in a previous year • play a role in preparing students for the next grade or course, but not a central role • address more narrowly defined ideas

STAAR Alternate 2 provides access to the grade-level TEKS through vertical alignment and curriculum framework documents. These documents, along with the blueprints for STAAR Alternate 2, can be found on the STAAR Alternate 2 Resources webpage.

TELPAS is based on the ELPS. TELPAS assessed curriculum and blueprints can be found on the TELPAS Resources webpage.

Overall, each STAAR, STAAR Alternate 2, and TELPAS assessment is designed to reflect:

- problem-solving and complex thinking skills,
- the range of content represented in the TEKS or ELPS,
- the level of difficulty of the skills represented in the TEKS or the range of English proficiency represented in the ELPS, and
- the application of content and skills in different contexts, both familiar and unfamiliar.

Tests are constructed from the bank of items determined to be acceptable after data review. Field-test data are used to place the item difficulty values on a common Rasch scale. This scale allows for the comparison of the difficulty of each item with that of all other items in the bank.

Consequently, items are selected not only to meet sound content and test construction practices but also to ensure that tests are approximately comparable in difficulty from one administration to the next. Refer to Chapter 3, “Standard Technical Processes,” for detailed information about Rasch scaling.

Tests are constructed to meet a blueprint for the required number of items and points on the overall test and for each reporting category. For STAAR, including STAAR Spanish, blueprints indicate the number of dichotomous and polytomous items and the number of extended constructed-response items. In addition, blueprints for STAAR, including STAAR Spanish, list a specific number of readiness and supporting standards. Items that test each reporting category are included for every administration, but the array of TEKS student expectations represented might vary from one administration to the next. Although the STAAR, including STAAR Spanish, assessments are constructed to emphasize the readiness standards, they still measure a variety of TEKS student expectations and represent the range of content eligible for each reporting category being assessed.

Before test construction is completed for the STAAR EOC assessments, panels made up of university-level experts review the content to ensure that each assessment is of the highest quality. A content-validation review is critical to the development of the EOC assessments because of the advanced level of content being assessed. Committee members note any issues of concern, and when necessary, replacement items are chosen and reviewed. STAAR Alternate 2, TELPAS, and TELPAS Alternate do not have content validation reviews.

After test construction for STAAR, including STAAR Spanish, is complete, TEA and Pearson work together to develop content and language supports for students who meet eligibility criteria. Content and language supports allow for various types of assistance (e.g., scaffolded directions, assistance with tracking, graphic organizers, simplified language, graphic representations of vocabulary and concepts) to support a student’s understanding of passages, test questions, and answer choices and are mainly in the form of pop-ups, rollovers, prereading text, and supplementary materials. These embedded supports are available for all online STAAR test forms.

For STAAR Alternate 2, accommodations and supports are included as part of the test design. For TELPAS, embedded accommodations are available on writing questions for students who meet eligibility criteria. Embedded accommodations are not provided on TELPAS Alternate.

All test content, including embedded supports, is reviewed and approved by TEA, after which the assessments are ready to be administered.

TELPAS Alternate is a holistic inventory that contains the same Observable Behaviors every year. Thus, there is no annual test construction process. Blueprints for TELPAS Alternate are available on the TELPAS Alternate Resources webpage.

Security

TEA prioritizes test security and confidentiality for all aspects of the Texas Assessment Program, from development and construction to administration and reporting. TEA ensures that every allegation of cheating or breach of confidentiality is properly investigated.

Maintaining the security and confidentiality of the Texas Assessment Program is critical for ensuring valid test scores and providing standardized and comparable testing opportunities for all students. TEA has implemented numerous measures to strengthen test security and confidentiality, including the development of various administrative procedures and manuals to train and support district testing personnel.

Test Administration Materials

The *District and Campus Coordinator Resources* and assessment-specific test administrator manuals provide guidelines on training testing personnel, administering tests, creating secure testing environments, and properly storing test materials. They also instruct testing personnel on how to report any confirmed or alleged testing irregularities that might have occurred. The manuals include information on the test security oaths that all personnel with access to secure test materials are required to sign as well as specific details about the possible penalties for violating test procedures. In addition, Texas Administrative Code (TAC) §101.3031 addresses test administration procedures and includes specific language detailing the requirements of school districts and charter schools to maintain security and confidentiality of assessment instruments, including a list of violations and their consequences.

Training

TEA training materials cover test administration best practices, including test security issues. All district and campus personnel who participate in state-mandated testing or handle secure test materials and content are required to be trained in test security and administration procedures. In addition to this required training, TEA provides optional online training modules. It is strongly recommended that districts and charter schools use these modules to help supplement the mandatory training required of all personnel involved in testing. Trainings are posted on the Learning Management System (LMS).

Security Violations

In accordance with test administration procedures, any person who violates, solicits another to violate, or assists in the violation of test security or confidentiality, and any person who fails to report such a violation, could be penalized. An educator involved with a testing irregularity might face:

- restrictions on the issuance, renewal, or holding of a Texas educator certificate, either indefinitely or for a set term;
- issuance of an inscribed or non-inscribed reprimand;
- suspension of a Texas educator certificate for a set term; or
- revocation or cancellation of a Texas educator certificate without opportunity for reapplication for a set term or permanently.

Students involved in a violation of test security could have their test results invalidated.

Incident Tracking

TEA regularly monitors and tracks testing irregularities and reviews all incidents reported from districts and campuses.

Processes that have been developed to assist in test administration and security include:

- an internal database that allows TEA to track reported testing irregularities and security violations,
- a system to review and respond to each reported testing irregularity, and
- a resolution process that tracks missing secure test materials after each administration and provides suggested best practices that districts can implement for proper handling and return of secure materials.

Quality-Control Procedures

The data provided by the Texas Assessment Program plays an important role in decision-making about student performance and public education accountability. Individual student test scores are used for accelerated instruction and graduation. In addition, the aggregated student performance results from the Texas Assessment Program are a major component of state and federal accountability systems used to rate individual public schools and school districts in Texas. The data are also used in education research and in the establishment of public policy. Therefore, it is essential that assessments are scored correctly and that scores are reported accurately.

TEA uses a comprehensive QC system to review work produced by the testing contractors. This section describes the procedures used to confirm the validity of scoring, reporting, and test development.

Data and Report Processing

TEA undertakes an extensive and comprehensive QC process to verify the quality and accuracy of final Texas Assessment Program results before reporting them. Begun months in advance of an assessment date, the QC process involves internal steps taken by CAI and the implementation of a joint process supported by TEA. This QC process is applied to every operational assessment administered in the school year.

CAI executes an internal QC system for the reporting of test results. QC at the unit level confirms that software modules associated with various business processes—such as online test delivery, scoring, and reporting—are properly developed and that they operate to meet program requirements. Performed by a group that is independent from the software development group, system QC confirms that all the modules work together so that outputs from one module in the system match the proper inputs for the next module. This process allows for independent verification and interpretation of project requirements. Once the independent testing group has completed and approved the test, the system is moved into production mode.

The joint QC process involves a complete scoring and reporting test run. For each test administration, TEA prepares response data for thousands of hypothetical students who serve as test cases. The test-run processing includes scoring the responses and generating student- and district-level reports and data files, and TEA independently verifies information during every step. Reports are not sent to districts until all discrepancies in the QC data set are resolved and the reports generated by TEA and the contractor match. Details of the QC process can be found in Appendix A.

Technical Processing

In addition to the processing of data and generation of reports, psychometric or technical processing of the data also occurs before and after each test administration and includes additional QC measures.

Each technical procedure requires calculations or transformations of the data to be completed and verified by multiple psychometricians and testing experts at CAI and Pearson; TEA also verifies these calculations.

Each year's calculations are also compared to historical values to further validate the reasonableness of the results. Comparisons of technical procedures and assessment results from year to year help verify the quality of the assessments and inform TEA of the program's impact on student achievement.

For more information about the standard technical processes of the Texas Assessment Program, see Chapter 3, "Standard Technical Processes."

Performance Assessments

STAAR, including STAAR Spanish, and TELPAS contain constructed-response items, which require scoring by trained human raters, on the following operational assessments:

- STAAR grades 3–8 RLA, grades 5 and 8 science, grade 8 social studies, English I, English II, Biology, and U.S. History
- STAAR Spanish grades 3–5 RLA and grade 5 science
- TELPAS grades 2–12 speaking and grades 2–12 writing

STAAR, including STAAR Spanish, uses extended constructed responses, which measure the student's ability to synthesize the component skills of writing; that is, the extended constructed-response task requires the student to express ideas effectively in writing for a specified purpose.

The types of writing required vary by grade and subject or course and represent the learning progression evident in the TEKS. RLA assessments include short constructed-response questions as well as an extended constructed-response question at every grade level. Science and social studies assessments include short constructed-response questions.

Extended constructed responses for STAAR, including STAAR Spanish, are evaluated using a holistic scoring process, meaning that the student response is evaluated as a whole according to pre-established criteria. These criteria, which are explained in detail in the scoring rubrics for

each type of writing, are used to determine the effectiveness, and thus the score, of each response.

The 5-point rubric for extended constructed responses includes two main components: 1) development and organization of ideas and 2) conventions. A student response may receive up to 3 points for development and organization of ideas and up to 2 points for the use of writing conventions. The constructed responses are scored independently by two scorers, and the scores are added to create a final score; therefore, a student may receive up to 10 points for his or her essay. Short constructed responses in the reading domain are scored using a 2-point prompt-specific rubric, and short constructed responses in the writing domain are scored using a 1-point rubric. Responses deemed nonscorable are assigned a condition code and receive 0 points. The STAAR writing rubrics for extended constructed responses can be found on the STAAR Resources webpage. Rubrics for the short constructed responses are included in the STAAR constructed-response scoring guides found on the same webpage.

TELPAS grades 2–12 reading and writing assessments include constructed-response and sentence-rewrite items, the scoring rubrics for which are found on the TELPAS Resources webpage. TELPAS writing items are evaluated using a holistic scoring process. Sentence-rewrite items receive a score of 0 or 1 based on the criteria defined in the rubric. Scorers use a 4-point writing rubric to evaluate constructed responses at grades 2 and 3, with two scorers independently evaluating student responses and those scores added together to calculate the students' raw score (from 2 to 8). For grades 4 through 12, a 12-point rubric is used to evaluate constructed responses, which are scored for three traits: vocabulary, usage, and completeness. Each trait is worth 1 to 4 points, and trait scores are added together to calculate a raw score of 3 to 12 points. A second scorer scores 25 percent of constructed-response items and 5 percent of sentence-rewrite items. These QC measures ensure the validity and reliability of scores.

The TELPAS speaking assessment consists of prompts that elicit student speaking responses recorded using a headset with a microphone. Speaking responses are scored according to a 2-point or 4-point rubric, depending on the item type. An automated scoring engine scores all student responses for TELPAS speaking. To ensure continued validity, reliability, and calibration of the assessment scoring process, a second scorer scores 10 percent of engine-scored responses. Data from these two methods are continuously compared to ensure the process is reliable.

Human scorers train the automated scoring engine by assigning points to the responses gathered during field testing. For operational items, human scorers score any responses that are considered uncertain cases or are part of a backread to examine the inter-rater reliability of the automated scoring engine. The TELPAS 2-point and 4-point speaking rubrics can be found on the TELPAS Resources webpage. Human scoring also takes place for responses that the automated engine identifies as nonscorable. These responses often have a unique characteristic—including, for example, background noise, mumbled or unclear speech, or low volume—that makes them appropriate for scoring by a human scorer. All scorers undergo the same extensive training process using the same materials and rubrics. Refer to Chapter 6, "TELPAS," for detailed information about the TELPAS speaking scoring process.

Scoring Staff

All test scorers have at least a four-year college degree and must undergo rigorous TEA-approved training before they are allowed to begin scoring. As part of this training, applicants must review an anchor set, score practice sets, and pass qualification testing. Scorers are monitored daily to produce scores that are accurate and reliable.

Pearson's training and monitoring of scorer performance is conducted by content specialists, supervisors, directors, and program managers, all of whom have demonstrated expertise with scoring constructed responses. Content specialists build the training materials from field-test responses to represent a full range of scores and train scoring leadership on both content and job expectations before scorer training. During operational scoring sessions, supervisors guide, support, and monitor scorers, and directors guide, support, and monitor supervisors; both roles share responsibility for monitoring and managing scoring quality by answering scorers' questions and reviewing scoring reports. Supervisors and directors apply all condition codes and reach out to content specialists when they need guidance. Program managers monitor all aspects of scoring for STAAR, including STAAR Spanish, and TELPAS, specify the configuration of training materials, and oversee the schedule and process for performing the work.

Distributed Scoring

Distributed scoring is used for STAAR, including STAAR Spanish, to allow scorers to participate in the scoring process from any location, provided they qualify and meet strict requirements. Distributed scoring is a secure, web-based model that incorporates several innovations and includes the following benefits:

- The number of scorers available locally can be augmented by other highly credentialed scorers from across the state and country.
- More teachers across the state can participate in the scoring process.
- Paper handling and associated costs and risks are reduced.
- Scorers are trained and qualified using comprehensive, self-paced online training modules that allow them to manage their training more efficiently.
- Distributed scoring uses state-of-the-art approaches to monitor scoring quality and communicate feedback to distributed scorers.

The ePEN Scoring System

STAAR, including STAAR Spanish, and TELPAS constructed responses are scored using the Pearson ePEN system. Scorers have access to TEA-approved rubrics and anchor papers during training, qualification, and operational scoring, and once they have completed training and qualification, they have secure access to students' constructed responses. The ePEN response viewer renders scanned images of students' constructed responses. Scorers can adjust contrast, color, and magnification to improve readability and reduce fatigue.

All constructed responses from a particular student and test are linked throughout Pearson scoring and reporting processes via a unique identifier. To protect student anonymity and prevent potential bias, student identifiers and other demographic information are not visible to scorers in ePEN.

Responses are grouped by grade and subject or course and are stored on the ePEN server. As scorers score the responses, more responses are routed into their scoring queues. Each scorer independently reads a response and selects a score from a menu on the computer screen. Scoring supervisors, scoring directors, and content specialists can identify which scorer reads each response.

Although the automated scoring engine scores most TELPAS speaking responses, sometimes responses require review by a human scorer. Pearson scores these items through ePEN, providing secure access to the students' audio files and scoring reports for content supervisory staff. Each scorer independently listens to a response and selects the appropriate score in the scoring grid. The system provides numerous tools and reports to help supervisory staff monitor scoring, and the rubric and training can be reinforced through qualification sets delivered regularly or when needed to address a scoring issue.

Scorer Training Process

All scorers who work on the STAAR and TELPAS performance task scoring projects receive extensive training through Pearson's online modules. This training covers the materials associated with the performance questions for each assessment and includes orientation in the ePEN system. Scorers receive training on the scoring guide that provides the rubric and anchor sets of each rubric score point for a particular assessment item. Additionally, scorers score training sets and have an opportunity to explain and discuss the scores. Scorers are required to demonstrate a complete understanding of the rubrics and to pass a qualification set before being allowed to score operational student responses.

Extended Constructed Responses

Training materials are selected to clearly differentiate student performance at the different rubric score points and to help scorers learn the difference between score points. To help scorers refine their understanding of differences between adjacent score points, training materials also include responses determined to be on the borderline between two adjacent score points. Supervisors are available during scorer training to assist and answer questions.

Once scorers complete the training sets, they are administered qualification sets of student responses. These student responses have already been scored by TEA and Pearson staff, and scorers must accurately assign scores to the student responses. Scorers are given two opportunities to qualify, with a different set of responses in each set. Any scorer who cannot meet the standards established by TEA and Pearson is dismissed from scoring.

Ongoing Training

After initial training, ongoing training is available to ensure scoring consistency and high scorer agreement. Supervisors and scoring directors monitor scoring and provide mentoring continually

during operational scoring. The ePEN scoring system includes a comprehensive set of scoring and monitoring tools that help identify areas for additional training.

Scoring Process

STAAR, including STAAR Spanish, constructed responses are scored using a holistic approach in which scores can be exact (scorer 1 and scorer 2 agree) or adjacent (scores by scorer 1 and scorer 2 differ by no more than 1 point). During scoring, two scorers independently assign a score from 1 to 4 to each student response. The scores are summed and weighted, if applicable, and the performance is reported to districts on both the STAAR Student Report Card for individuals and on the Constructed Responses Summary Report for campuses and districts.

In instances in which the scores are discrepant (i.e., scores from scorer 1 and scorer 2 differ by more than 1 point), the student response is routed to a resolution queue. A supervisor or scoring director reviews the response and applies a third score, invalidating the two initial scores. This score is then doubled and becomes the reported score.

Throughout scoring, TEA staff members are consulted on responses that are highly unusual or require a policy decision from TEA.

Nonscorable Responses

Only a scoring director may determine if a constructed response is nonscorable. Before a constructed response can be given a nonscorable designation, the supervisor or scoring director thoroughly reviews the response. If the scoring director determines that the response is scorable, it is assigned a score and routed to a second content scoring leader. If the scoring director determines that the response is nonscorable, a nonscorable code is applied, and the response is routed to another scoring director for confirmation. While the response is under review, it is held in a review queue that prevents it from being distributed to other scorers.

Monitoring of Scorer Quality

Scorers can defer student responses to their supervisor, who will provide feedback on how to score the response or pass the question to the scoring director. This allows scorers to receive feedback regularly on their performance. If a scorer is identified as having difficulty applying the criteria, the responses they scored are invalidated and rescored, and that scorer must then complete targeted qualification training. Any scorer who cannot pass the targeted qualification training set is dismissed from scoring.

Validity responses are student responses that have already been assigned a score during anchor approval meetings and are presented to scorers throughout the operational scoring process to monitor their scoring quality. TEA approves all validity responses before they are introduced into the scoring systems. Indistinguishable from operational responses, validity responses are inserted randomly into the scoring queue. Scorers' accuracy is evaluated based on how often their scores on validity responses agree with the scores that have been assigned to them.

For TELPAS, a supervisor using ePEN can back-listen to responses scored and send that scorer feedback through the ePEN messaging system. Scorers can also submit responses for

review so that a scoring supervisor or scoring director can listen and provide feedback. Validity responses with TEA-approved scores are delivered randomly to scorers throughout the project. Scorers failing to meet the standard for validity after remediation are dismissed from the project, and their work is reset and scored again.

Anchor Sets

In addition to the scoring that field-test scorers perform, TEA and Pearson staff members independently score samples of the field-test responses that will be used on the operational assessments. TEA and Pearson content and management staff and Texas educators participate in a series of anchor approval meetings to analyze these responses and assign scores. Assessment specialists select responses from the anchor approval meetings to be included in each scoring guide. Scoring directors then assign the remaining pre-scored responses from the meetings to training sets and qualifying sets for use in scorer training. Educators assist in the review and make recommendations to reach a consensus on the scores. Before scoring, TEA staff members review and approve all scoring guides and training sets.

Score Reliability and Validity Information

TEA regularly reports on the reliability and validity of the performance scoring process. Reliability is expressed in terms of scorer agreement (percentage of exact agreement between scorers' scores) and correlation between first and second scores. Validity is assessed by the inclusion of validity responses throughout the operational scoring process and expressed in terms of exact agreement between the score assigned by the scorer and the score assigned by TEA and Pearson.

Appeals

If a district has questions about the score assigned to a response, a rescore can be requested in TIDE. If the score changes, CAI provides rescore results by posting an updated student report card to the TIDE secure inbox and to the Family Portal. In instances where a rescore improves scores, the fee associated with the rescore request is waived. If the score does not change, the district pays the associated fee. If a district files a formal appeal with TEA related to scores reported on the consolidated accountability file, an analysis of the response in question is provided to explain the final outcome of the appeal and whether the score was changed.

TECHNICAL DIGEST 2022–2023

Chapter 3

Standard Technical Processes

Overview
Technical Details and Procedures
- Performance Standards
- Item Analysis
- Scaling
- Equating
- Reliability
- Validity
- Measures of Student Progress
- Sampling

Overview

The Standards for Educational and Psychological Testing, developed jointly by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME), provides guidelines for evaluating the quality of testing practices. TEA applies these standards to all aspects of the Texas Assessment Program to ensure its assessments are technically defensible and appropriate for the purposes for which they are used.

To promote fairness, accuracy, reliability, and validity in the Texas Assessment Program, TEA uses the following technical concepts, which are discussed in detail in this chapter:

- performance standards
- item analysis
- scaling
- equating
- reliability
- validity
- measures of student progress
- sampling

Program-specific technical processes are covered in subsequent chapters.

Technical Details and Procedures

Performance Standards

A critical aspect of any statewide testing program is the establishment of performance standards that provide a frame of reference for interpreting test scores. Performance standards help relate test performance directly to the student expectations expressed in the state curriculum in terms of what knowledge and skills students are expected to demonstrate upon completion of each grade or course. Performance standards, therefore, describe the level of competence students are expected to demonstrate on an assessment.

STAAR, including STAAR Spanish, has three cut scores that identify the following four performance levels:

- Did Not Meet Grade Level
- Approaches Grade Level

- Meets Grade Level
- Masters Grade Level

STAAR Alternate 2 has two cut scores that identify the following three performance levels:

- Level I: Developing Academic Performance
- Level II: Satisfactory Academic Performance
- Level III: Accomplished Academic Performance

TELPAS has three cut scores that identify the following four English proficiency levels:

- Beginning
- Intermediate
- Advanced
- Advanced High

TELPAS Alternate has four cut scores that identify the following five English proficiency levels:

- Awareness
- Imitation
- Early Independence
- Developing Independence
- Basic Fluency

Standard setting is the process of establishing cut scores that define the performance levels on an assessment. The standard-setting framework and process for the STAAR, STAAR Alternate 2, TELPAS, and TELPAS Alternate programs are described below.

Standard Setting for STAAR

Performance standards for STAAR were originally established in 2012 using an evidence-based standard-setting approach (O'Malley, Keng, & Miles, 2012). Standard setting for STAAR involved a process of combining policy considerations, the PLDs derived from the TEKS content standards, educator knowledge about what students should know and be able to do, and information about how student performance on statewide assessments aligns with performance on other assessments. Standard-setting advisory panels, made up largely of diverse groups of educators, considered the interaction of all these elements for each STAAR assessment.

In 2014, standard-setting committees reset performance standards for the STAAR English I, English II, and English III assessments, which combined the reading and writing components into a single assessment. In 2015, standard-setting committees reset the STAAR grades 3–8 mathematics performance standards due to changes in the TEKS. With the STAAR redesign in the 2022–2023 school year, performance standards for all STAAR assessments were updated using the Modified Angoff (Angoff, 1971) standard-setting method.

Refer to the *STAAR Standard Setting Technical Report* available on the Assessment Reports and Studies webpage for more detailed information.

Standard Setting for STAAR Alternate 2

Performance standards for STAAR Alternate 2 were originally established in spring 2015 using an evidence-based standard-setting approach (O'Malley, Keng, & Miles, 2012). This involved a process of combining considerations regarding policy, the TEKS content standards, educator knowledge about what students should know and be able to do, and information about how student performance on state assessments aligned with student performance on other assessments.

Due to changes in the STAAR Alternate 2 RLA assessments, performance standards for these assessments were reset in spring 2023 using the Modified Angoff (Angoff, 1971) method. This content- and item-based method led panelists through a standardized process in which they considered student expectations, as defined by the PLDs, and the individual items that were administered to students to recommend cut scores for each performance level.

Refer to the *STAAR Alternate 2 Standard Setting Technical Report* available on the Assessment Reports and Studies webpage for more detailed information.

Standard Setting for TELPAS

TELPAS grades 2–12 reading proficiency standards were originally established in 2008. The method consisted of a two-phase process in which an internal work group made initial recommendations and then an external committee of state educators recommended specific cut scores after reviewing the recommendations, the test forms on which the recommendations were based, and impact data.

During the 2013–2014 school year, TEA convened educator committees to review the proficiency standards for TELPAS grades 2–12 reading to align the program with STAAR. TEA used an evidence-based standard-setting approach to determine the cut scores. As with STAAR standard setting, the item mapping with external data method (Ferrara, Lewis, Mercado, D'Brot, Barth, & Egan, 2011; Phillips, 2012) was used for TELPAS, along with validity study information, to recommend the updated proficiency standards.

The TELPAS grades 2–12 reading test redesign in spring 2018 and the first-time administration of an online test for the grades 2–12 listening and speaking domains required establishing new cut scores for TELPAS proficiency levels. A test-centered, criterion-referenced method was used to guide panelists as they determined their proficiency level cut score recommendations.

The applied method was a hybrid of the Angoff method (Angoff, 1971) and the Extended Modified Yes/No Angoff method (Davis & Moyer, 2015; Plake, Ferdous, Impara, & Buckendahl, 2005).

Proficiency standards were established for TELPAS grades 2–12 writing in spring 2012 as the assessment transitioned to a standardized online assessment. The standard-setting methodology used was a modification of the well-known Body of Work method (Kingston, Kahl, Sweeney, & Bay, 2001; Kingston & Tiemann, 2012), which has been used to recommend proficiency level cut scores for various large-scale state assessments.

Refer to the *TELPAS Standard Setting Technical Report* available on the Assessment Reports and Studies webpage for more detailed information.

Standard Setting for TELPAS Alternate

The proficiency standards for TELPAS Alternate were established in 2019. To establish the proficiency levels for each domain, a test-centered, criterion-referenced method was used to guide the panelists. The implemented procedure was a hybrid of the Extended Modified Yes/No Angoff method (Davis & Moyer, 2015; Plake, Ferdous, Impara, & Buckendahl, 2005). The hybrid standard-setting procedure is a systematic method that combines various considerations into the process of recommending cut scores for the different proficiency levels.

Refer to the *TELPAS Alternate Standard Setting Technical Report* available on the Assessment Reports and Studies webpage for more detailed information.

Item Analysis

Statistical analyses are conducted on student performance data to gauge the level of difficulty of the item, examine the degree to which the item appropriately distinguishes between students of different proficiency levels, and assess the item for potential bias. Several statistical analyses, based on both classical test theory and item response theory (IRT), are used to analyze the data collected annually for operational items. Item analyses are also conducted annually for the purpose of reviewing the quality of newly field-tested items to help determine which items may be included as operational items in future test administrations. Statistics generated after each administration of STAAR (including STAAR Spanish), STAAR Alternate 2, and TELPAS include *p*-value, point-biserial correlation, Rasch item difficulty, Rasch fit, and response or score point distribution. In addition, group difference analyses, also known as differential item functioning (DIF), are conducted using the Mantel-Haenszel (MH) alpha and ABC DIF classification.

p-Value

The *p*-value indicates the proportion of the total group of students answering a multiple-choice or dichotomous item correctly. For polytomous items, the *p*-value indicates the average score obtained by students divided by the number of points possible. An item's *p*-value shows how difficult the item was for the students who were administered the item. An item with a high *p*-value, such as 0.90, is a relatively easy item. An item with a low *p*-value, such as 0.30, is a relatively difficult item.

Point-Biserial Correlation

The point-biserial correlation describes the relationship between a student's performance on the item and performance on the assessment as a whole. A high point-biserial correlation indicates that students who answered the item correctly tended to score higher on the entire test than those who answered the item incorrectly. In general, point-biserial correlations less than 0.20 indicate a potentially weaker-than-desired relationship.

Note that the point-biserial correlation may be weak on items with very high or very low *p*-values. For example, if nearly all students perform well (or poorly), that item does not provide useful information for distinguishing between those students with higher performance from those students with lower performance on the entire assessment.

Rasch Item Difficulty

The Rasch item difficulty estimate is another indicator of item difficulty. In contrast to *p*-values, which are influenced by the ability level of the students who were administered the item, Rasch item difficulties can be compared across test forms and administrations. Items with low Rasch item difficulty values (e.g., -1.5) are relatively easy, and items with higher values (e.g., +1.5) are relatively difficult.

Rasch Fit

The Rasch fit statistic indicates the extent to which student performance on an item is similar to what would be expected under the Rasch measurement model. Specifically, items with good Rasch fit have relatively few unexpected responses (e.g., low-scoring students answering difficult items correctly, high-scoring students answering easy items incorrectly). In general, a Rasch fit value lower than 0.7 or greater than 1.3 may indicate that the item fits the Rasch model poorly.

Response or Score Point Distribution

The response or score point distribution represents the percentage of students responding to each of the answer choices (i.e., A, B, C, or D) for a multiple-choice item, the percentage of students who responded correctly or incorrectly for a dichotomous item, or the percentage of students who received each of the score points for a polytomous item. Response or score point distributions are provided for the entire group of students and for various demographic groups (e.g., gender, ethnicity for STAAR) or for proficiency level groups (e.g., Beginning, Intermediate, Advanced, Advanced High for TELPAS).

Group Difference Analysis

Statistics from a group-difference analysis provide information about how different student groups (e.g., male, female, African American, Hispanic, White) performed on an item. Such analyses help identify items on which a group of students performed unexpectedly well or poorly. Both the MH alpha and the ABC DIF classification, also known as the Educational

Testing Service (ETS) DIF classification (Petersen, 1987; Zieky, 1993), are used for the Texas Assessment Program.

It should be noted that DIF analyses serve to merely identify test items that have unusual statistical characteristics related to student group performance; the DIF analyses alone do not prove that specific items are biased. Such judgments are made by item reviewers who are knowledgeable about the state's content standards, instructional methodology, and student testing behavior.

Mantel-Haenszel Alpha

To calculate the MH alpha, students are first divided into categories of similar proficiency. An odds ratio is calculated for each of those proficiency categories, where the odds ratio equals the odds of answering correctly for the designated reference group (e.g., males) divided by the odds of answering correctly for the focal group (e.g., females). These odds ratios are combined across proficiency categories to obtain a common odds ratio, known as the MH alpha. If the value of the MH alpha is 1, students of similar proficiency, regardless of group membership (e.g., males, females), are equally likely to answer the item correctly. If the MH alpha value is statistically significantly greater than 1, the chance of success on the item is better for the reference group (e.g., males) than for the focal group (e.g., females) when comparing students of similar proficiency. Statistically, an MH alpha value significantly less than 1 indicates the item is easier for the focal group compared to similarly proficient students in the reference group.

ABC DIF Classification

The ABC DIF classification is based on the MH alpha, but it considers both statistical and practical significance when examining an item for DIF. Each item is classified into one of three categories based on each group comparison: "A" means negligible or no DIF, "B" means moderate DIF, and "C" means large DIF (refer to Zieky, 1993, for more information). Plus and minus signs (+/-) indicate the direction of DIF. A plus sign indicates that the item is unexpectedly easy for the focal group (e.g., females), and a minus sign indicates that the item is unexpectedly easy for the reference group (e.g., males).

Scaling

Scaling associates numbers with characteristics of interest to provide information about measurable quantities for those characteristics. For example, temperature can be described using the Fahrenheit scale or the Celsius scale. Different numbers refer to the same temperature, but they describe it using different scales. Similarly, test scores can also be reported using more than one scale.

The number of items that a student answers correctly on a given test is known as the raw score, and this raw score is interpreted in terms of the specific set of answered test questions. In general, raw scores from different test forms are not comparable. For example, suppose there are two forms of an assessment that are not equally difficult: Form A is harder than Form B. One student takes Form A and earns a raw score of 34 out of 50, while another takes Form B

and also earns a raw score of 34 out of 50. Here, the first student's performance on the harder test reflects greater achievement than the second student's performance on the easier one, even though both students receive the same raw score.

When a new form of an assessment is administered, the questions on the new form are generally different from those on older forms. Despite the fact that different test forms target the same knowledge and skills, some forms will be slightly easier or slightly more difficult than others. As a result, in most cases student performance cannot be compared directly across test administrations using raw scores. To facilitate comparisons, raw scores from different test forms and administrations are placed onto a common scale resulting in scale scores. Unlike raw scores, scale scores allow for direct comparisons of student performance across separate test forms and different test administrations. A scale score considers the difficulty level of the specific set of questions on a test form, and it describes students' performance relative to each other and relative to the performance standards across separate test forms.

Three scales underlie the STAAR (including STAAR Spanish), STAAR Alternate 2, TELPAS, and TELPAS Alternate assessments: the raw score scale, the Rasch scale, and the reporting scale. The scales are defined as follows:

- The raw score scale is defined as the number of items answered correctly, regardless of difficulty.
- The Rasch scale is a transformation of the raw scores onto a scale that considers the difficulty of the items and is comparable across different test forms and administrations.
- The reporting scale is a linear transformation of the Rasch scale, through scaling constants, onto a user-friendly scale. Because the transformation is linear, the reporting scale also considers item difficulty. The reported scale scores are comparable and maintain performance standards across test forms and administrations.

The following sections detail the scaling process in terms of establishing the Rasch scale and transforming the scores on the Rasch scale into the reported scale scores.

The Scaling Process

The scaling process places test score data from different tests onto a common scale. There are three primary approaches to scaling: subject-centered, stimulus-centered, and response-centered (Crocker & Algina, 2006; Torgerson, 1958). Subject-centered approaches locate students on a scale according to the amount of knowledge each student demonstrates, while stimulus-centered approaches place the test items or stimuli on a scale according to the amount of knowledge required to answer each item correctly. Response-centered approaches simultaneously locate students and items on a scale based on how students respond to the items and how difficult the items are and can be thought of as a combination of subject-centered and stimulus-centered approaches; therefore, they are the most complex approaches.

TEA scales assessments using a response-centered approach that involves specialized statistical methods that can estimate both student proficiency and the difficulty of a particular set of test items. Specifically, the Texas Assessment Program uses a statistical model known as the Rasch Partial-Credit Model (RPCM) to place test items and measures of student proficiency on the same Rasch scale across test forms and test administrations. Scores on the Rasch scale are then transformed to more user-friendly scale scores to facilitate interpretation.

Rasch Partial-Credit Model

Test items (whether dichotomous or polytomous) for the Texas Assessment Program are scaled and equated using the RPCM. The RPCM is an extension of the Rasch one-parameter IRT model attributed to Georg Rasch (1966), and extended by Wright and Stone (1979), Masters (1982), Wright and Masters (1982), and Linacre (2018). The RPCM was selected because of its flexibility in accommodating dichotomous or polytomous items. The RPCM maintains a one-to-one relationship between scale scores and raw scores, meaning each raw score is associated with a unique scale score. An advantage to the underlying Rasch scale over the raw score scale is that it allows for comparisons of student performance across years. Additionally, the underlying Rasch scale enables the maintenance of equivalent performance standards across test forms.

The RPCM is defined by the following equation:

$$p_{im}(\theta) = \frac{\exp\left[\sum_{k=0}^{m}(\theta - \delta_{ik})\right]}{\sum_{v=0}^{M_i-1}\exp\left[\sum_{k=0}^{v}(\theta - \delta_{ik})\right]}, \quad (1)$$

where M_i is the number of score categories of item i, θ is a student's proficiency (ability) score, m=(0, 1,… M_i–1) is a raw score of item i, $p_{im}(\theta)$ is the probability of getting score m on item i conditional on θ, δ_{ik} is the step difficulty parameter of score k on item i, and denote θ–δ_{i0} =0.

The RPCM provides the probability of scoring each value of m on item i as a function of a student's proficiency score θ and the step difficulties δ_{ik}, which indicate the proficiency score at which the probability of scoring k equals the probability of scoring k–1 (refer to Masters, 1982, for an example). Note that for multiple-choice and dichotomous technology-enhanced items, there are only two score categories: 0 for an incorrect response and 1 for a correct response. In this case, the RPCM reduces to the standard Rasch one-parameter IRT model, and the resulting single-step difficulty is more properly referred to as an item difficulty.

Some of the advantages of RPCM scaling are as follows:

- All items, regardless of type, are placed on the same common Rasch scale.
- Students' achievement results are placed onto the same scale as the items, so it is possible to make inferences about which items a student is likely to respond to

correctly or incorrectly based on the student's proficiency. This facet is helpful in describing test results to students, parents, and teachers.

- Field-test items can be placed on the same Rasch scale as items on the operational assessment. This enables student performance on the field-test items to be linked to all items in the item bank, which is useful in the construction of future test forms.
- The RPCM allows for the pre-equating of future test forms, which can help test builders evaluate test forms during the test construction process.
- The RPCM also supports post-equating of the test, which establishes a link between the current and previous test forms. Linking the current test form to previous test forms enables comparisons of test difficulties and passing rates in test forms given across different administrations. Because both pre-equated and post-equated item difficulty estimates are available, any drift in scale or difficulty can be quantified.

The Texas Assessment Program uses two types of scale scores—horizontal and vertical. Horizontal scale scores are used for STAAR grades 5 and 8 science (including STAAR Spanish grade 5 science), STAAR grade 8 social studies, STAAR EOC assessments, STAAR Alternate 2, TELPAS, and TELPAS Alternate. Vertical scale scores are used for STAAR grades 3–8 mathematics, STAAR grades 3–8 RLA, STAAR Spanish grades 3–5 mathematics, and STAAR Spanish grades 3–5 RLA.

Horizontal Scaling

Scale scores (SS_θ) for the Texas Assessment Program represent linear transformations of Rasch-based proficiency estimates (θ). For horizontal scale scores, this transformation is made by first multiplying any given θ by a slope (A) and then adding an intercept (B). This operation is represented by the following equation:

$$SS_\theta = A \times \theta + B \quad (2)$$

The slope and intercept in equation 2 are scaling constants, and they are derived using a method described by Kolen and Brennan (2004). For STAAR grades 5 and 8 science, STAAR grade 8 social studies, STAAR EOC assessments, TELPAS, and TELPAS Alternate, two scale score values at two specific standards were established in advance. These standards are Meets Grade Level and Approaches Grade Level for STAAR, Advanced and Advanced High for TELPAS, and Early Independence and Developing Independence for TELPAS Alternate. The A scaling constant is calculated as follows:

$$A = \frac{SS_2 - SS_1}{\theta_2 - \theta_1} \quad (3)$$

In equation 3, SS_2 represents the desired scale score at the higher of the two standards desired to be fixed, and SS_1 represents the desired scale score at the lower standard, where θ_2 and θ_1

are the corresponding Rasch-based proficiency estimates at the selected standards. The *B* scaling constant is calculated as follows:

$$B = SS_2 - A \times \theta_2 \quad (4)$$

For STAAR Alternate 2, the scale score value at the passing standard (Satisfactory) and the standard deviation of the reportable scale score were established in advance. The *A* scaling constant is calculated as follows:

$$A = \frac{\sigma_{SS}}{\sigma_\theta} \quad (5)$$

In equation 5, σ_{SS} represents the desired standard deviation of the scale score, and σ_θ represents the standard deviation of the Rasch-based θ values among a sample group. For all STAAR Alternate 2 assessments except grades 3–8 RLA, English I, and English II, the horizontal scales sample group comprised all students who took that assessment in spring 2015. For grades 3–8 RLA, English I, and English II, the sample group comprised all students in the spring 2023 administration. The *B* scaling constant is calculated as follows:

$$B = SS_{Satisfactory} - A \times \theta_{Satisfactory} \quad (6)$$

In equation 6, $SS_{Satisfactory}$ and $\theta_{Satisfactory}$ represent the selected scale score to be fixed at the passing standard and its corresponding Rasch-based proficiency estimate, respectively.

Because each assessment's horizontal scale is derived using its own sample group, σ_θ varies across assessments. Likewise, each assessment has a unique Meets Grade Level performance standard on STAAR in Rasch units, so θ_{Meets} varies across assessments. SS_{Meets} and σ_{SS} are set to be consistent within content areas but not across all assessments. Similarly, the STAAR Alternate 2 Level II: Satisfactory performance standards are also unique for each assessment; $\theta_{Level\ II}$ varies across assessments, and $SS_{Level\ II}$ and σ_{SS} are set to be consistent within content areas. Once these constants are established, the same transformations are applied each year to the Rasch proficiency estimates derived from performance on that year's test questions.

Vertical Scaling

A vertical scale score system allows for direct comparison of student test scores across grade levels within a content area. Vertical scaling refers to the process of placing scores of tests in the same content area at different grade levels onto a common scale. In order to implement a vertical scale, research studies were needed to determine differences in difficulty across grade levels. Such studies were conducted for STAAR grades 3–8 mathematics and RLA and STAAR Spanish grades 3–5 RLA in spring 2023. For these studies, embedded field-test positions (refer to the Field-Test Equating section) were also used to administer vertical linking items. The studies assumed a common-item nonequivalent groups design (refer to the Equating section) in which items from different grade levels appear together on adjacent grade-level tests, allowing for direct comparison of item difficulties across grade levels. By embedding vertical linking items

across grade levels, it is possible to calculate linking constants equal to the average differences in item difficulties of vertical linking items between adjacent grade pairs. These linking constants are used to create a vertical scale.

Similar to the horizontally scaled assessments, vertically scaled scores also reflect linear transformations of Rasch-based proficiency scores (θ). Vertically scaled scores, however, include an extra scaling constant (V_g) that varies across each grade (g). This is given by the equation below:

$$SS_\theta = A \times \left(\theta + V_g\right) + B, \quad (7)$$

where SS_θ is the scale score for a Rasch proficiency score (θ). The scaling constants *A* and *B* in equation 7 are derived in the same way as for horizontal scale score systems, except that the scale score for one of the performance standards (e.g., Meets Grade Level) is fixed only for one of the assessments in the vertical scale (e.g., STAAR grade 3 mathematics for the STAAR mathematics vertical scale), and the standard deviation is calculated using the calibration sample of the base grade. The *A* scaling constant is calculated as follows:

$$A = \frac{\sigma_{SS}}{\sigma_\theta} \quad (8)$$

In equation 8, σ_{SS} represents the desired standard deviation of the scale across all assessments, while σ_θ represents the standard deviation of Rasch-based θ values for the calibration sample in the base grade. The STAAR grades 3–8 mathematics, grades 3–8 RLA, and Spanish grades 3–5 RLA vertical scale sample group comprised all students who took a test form with embedded vertical scale items in spring 2023. Like field-test items, these vertical scale items are not used to calculate student scores.

The *B* scaling constant is calculated as follows:

$$B = SS_{Approaches} - \frac{\sigma_{SS}}{\sigma_\theta} \times \theta_{Approaches} \quad (9)$$

In equation 9, $SS_{Approaches}$ represents the desired scale score at the STAAR Approaches Grade Level cut score for the final assessment in the vertical scale, and $\theta_{Approaches}$ represents the approved STAAR Approaches Grade Level performance standard in Rasch units for the final assessment in the vertical scale.

Equating

Used in conjunction with the scaling process, equating is the process that considers the differences in difficulty across test forms and administrations and allows scores to be placed onto a common scale. The Texas Assessment Program uses the common-item nonequivalent groups design to equate most assessments because of its relative ease of implementation and, more importantly, because it is less burdensome on students and campuses. Under the common-item nonequivalent groups design, each student sample takes a different form of the

test with a set of items that is common across tests. The common items, sometimes referred to as equating items, can be embedded within the test or can stand alone as a separate test. The specific data-collection designs and equating methods used for the Texas Assessment Program are described in this section. Refer to Kolen and Brennan (2004) or Petersen, Kolen, and Hoover (1989) for a more detailed explanation of equating designs and methods.

With the spring 2023 administrations, new scales were established for STAAR; STAAR Alternate 2 grades 3–8 RLA, English I, and English II; and TELPAS writing. When new scales are created, there is no equating methodology employed since there is no link to a previous scale necessary during calibrations. All items are freely calibrated to establish the new scale.

During test construction, pre-equating based on the previous scales for STAAR and STAAR Alternate 2 and the field-test results for TELPAS were used as a guide. However, these scales were not used for reporting. Thus, for these assessments, equating was mainly employed for field-test analyses.

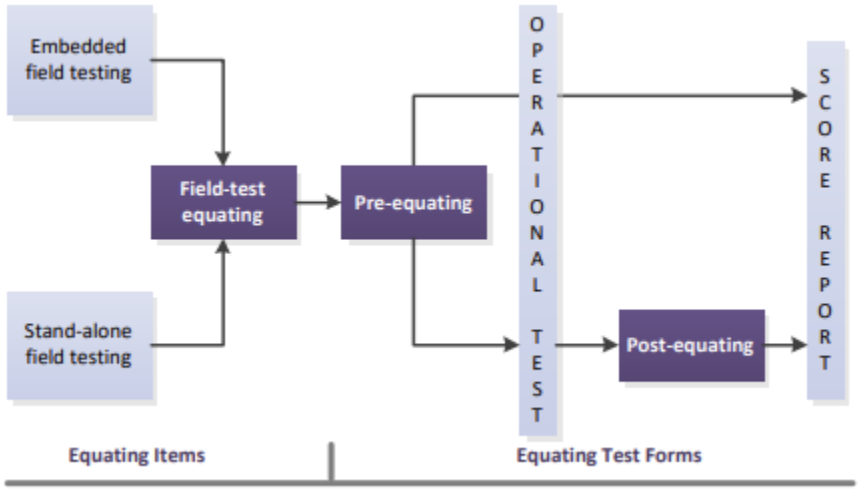
Types of Equating

The following are the three types of equating used in the item and test development process:

1. pre-equating test forms that are under construction
2. post-equating operational test forms after administration
3. equating field-test items after administration

One or more of these three types of equating is used on each component of the Texas Assessment Program, allowing the established performance standards for the assessments to be maintained on all subsequent test forms. Figure 3.1 illustrates the three types of equating used for the Texas Assessment Program. While field-test equating focuses on equating individual items to the Rasch scale of the item bank, pre-equating and post-equating both focus on equating test forms to maintain score comparability and consistent performance standards. Pre-equating and post-equating methods take into account differences in the difficulty of test forms.

Figure 3.1. Three Types of Equating Used



Pre-Equating

The pre-equating process occurs when a newly developed test form is placed onto the Rasch scale prior to administration. The goal of pre-equating is to produce a table that establishes the link between raw scores and scale scores before the test is administered. Because the difficulty of the items is established in advance (the items appeared previously on one or more test forms as field-test or operational items), the difficulty level of newly developed test forms can be estimated, and the anticipated connection among the raw scores, scale scores, and performance level standards can be identified. Once the anticipated connection among raw scores, scale scores, and performance levels has been established, a raw score to scale score (RSSS) conversion table can be produced that maps each raw score to a scale score and indicates the performance level cut scores.

The pre-equating process involves the following four steps:

1. Items are selected that have been equated to the Rasch scale from the item bank.
2. A new test form is constructed that meets the content specifications and statistical guidelines.
3. The test form under construction is evaluated against Rasch-based difficulty targets.
4. An RSSS conversion table for the operational test form is developed using the Rasch-based item difficulties.

Pre-equating is conducted as part of the test construction process for all assessments for which scale scores are reported (i.e., STAAR, STAAR Spanish, STAAR Alternate 2, TELPAS grades 2–12). In many cases, post-equating is also conducted. For some assessments, however, post-equating is not conducted, and the pre-equated RSSS conversion tables are used to assign scale scores. A pre-equating-only model might be preferred when a small or non-representative

sample of students is taking the operational test form or when faster reporting of scores is a priority.

Post-Equating

Post-equating might be preferred when changes in item presentation (e.g., position, formatting) or instructional practice have occurred since an item was field-tested because those changes might impact the estimated difficulty of the item. STAAR, STAAR Spanish, STAAR Alternate 2, and TELPAS grades 2–12 are post-equated. Post-equating in the Texas Assessment Program employs a conventional common-item nonequivalent groups equating design whereby an equating constant is calculated and used to transform the Rasch difficulty obtained from the current calibration to the Rasch difficulty established by the original test form. This equating constant is defined as:

$$t_{a,b} = \frac{\sum_{i=1}^{k}(d_{i,a} - d_{i,b})}{k}, \quad (10)$$

where $t_{a,b}$ is the equating constant, $d_{i,a}$ is the Rasch difficulty of item i on the current form a, $d_{i,b}$ is the Rasch difficulty of item i on the item bank scale, and k is the number of common items (Wright, 1977). Once the equating constant is calculated, it is applied to all item difficulties, transforming them to the item bank scale. After this transformation, the item difficulties from the current administration of the test are directly comparable to the item difficulties from all past administrations because equating was also performed on those items. These updated item difficulty estimates are then used to create the RSSS conversion table that is used to report scale scores. Both item difficulty and student proficiency are on the same scale under the Rasch model. Therefore, the resulting scale scores are comparable from year to year.

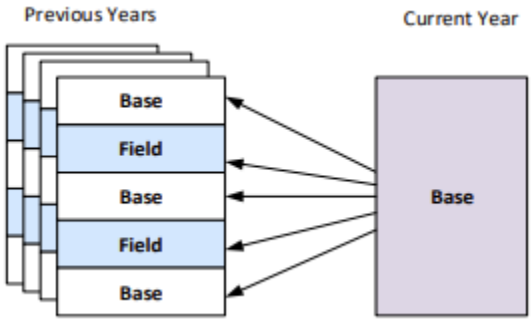
The post-equating procedure for STAAR involves the following steps:

1. Tests are assembled and evaluated using Rasch-based difficulty targets.
2. Data from the test administrations are sampled (where applicable).
3. Rasch item difficulty calibrations are conducted using the sampled data.
4. A post-equating constant is calculated as the difference in mean Rasch item difficulty of items in the equating item set on the scale of the item bank versus the operational scale.
5. The post-equating constant is applied to the Rasch difficulty estimates for the operational test items, and RSSS conversion tables are produced.

The redesigned STAAR assessments were first administered in spring 2023 with updated performance standards. Subsequent STAAR test forms will be equated to the spring 2023 administration. However, the June 2022 and December 2022 STAAR EOC assessments followed the previous scale and employed the post-equating steps listed above. For these English I and English II assessments, all multiple-choice items on each assessment were used as the equating item set, and post-equating was conducted on the entire population to ensure

representativeness. Figure 3.2 illustrates the source of the common item sets for these tests. The base-test items in the current year form were field-test items in previous years.

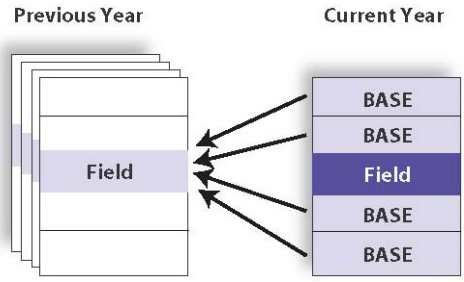
Figure 3.2. STAAR Common-Item Post-Equating Design



The initial equating item set comprised all multiple-choice items. However, the stability of the Rasch item difficulty estimates for the equating items is monitored from year to year. If a Rasch item difficulty is less stable than expected, the item will be excluded from the equating item set during the stability check. Prior to applying the final equating constant, the number of items in the equating set is compared to the base test, and the content representation of the common item set is compared to that of the base test to verify that the reporting categories are appropriately represented.

For STAAR Alternate 2 and TELPAS, the equating item set comprises all the base-test items, and the base-test items' Rasch difficulty values from field testing are compared to their values from operational testing to calculate the equating constant. Figure 3.3 illustrates the source of the equating items for STAAR Alternate 2 and TELPAS. The arrows in Figure 3.3 indicate the transformation of the base-test Rasch item difficulties for the current year onto the Rasch scale for an assessment through the same items' field-test Rasch item difficulties from their appearance in previous assessments.

Figure 3.3. STAAR Alternate 2 and TELPAS Common-Item Post-Equating Design



STAAR Alternate 2 and TELPAS post-equating is conducted using all or nearly all of the student data, so no sampling is needed. However, the stability of the Rasch item difficulty estimates is monitored from field test to base test, and if an item's Rasch item difficulty appears less stable than expected, the item will be excluded from the equating item set during the stability check. Prior to applying the final equating constant, the number of items in the equating set and the

content representation of the equating item set are compared to the base test to verify that the test content is appropriately represented in the equating item set.

The full equating process is independently replicated by multiple psychometricians from TEA and external vendors for verification.

Field-Test Equating

To replenish the item bank as new tests are created and released, newly developed items must be field-tested and equated to the Rasch scale of the assessment. STAAR (including STAAR Spanish), STAAR Alternate 2, and TELPAS use embedded field-test designs to collect data on field-test items. A stand-alone field test is occasionally conducted for STAAR.

After a newly constructed field-test item has cleared the review process, it is embedded in a test form along with operational items. There are two ways in which field-test items may be embedded.

STAAR (including STAAR Spanish) field-test items are randomly administered to students using a linear-on-the-fly test (LOFT) design in which all students are presented the same set of operational items that count toward their score. The LOFT design also achieves a representative sample of test takers for each item while eliminating the need for spiraling of forms.

STAAR Alternate 2 and TELPAS field-test items are placed on fixed forms along with operational items. Each field-test item appears on only a small number of test forms (typically one form) and does not count toward students' scores. Test forms containing field-test items are distributed so that a representative sample of test takers responds to the field-test items.

Regardless of which method is used to field-test items, all items are combined into a single data matrix, and a calibration of the Rasch item difficulties for both the operational items and the field-test items is conducted.

STAAR and TELPAS use a fixed common-items parameter approach to place the field-test items on the same Rasch scale as the operational items. In this procedure, all operational or base-test items are anchored to their bank values, and field-test items are calibrated and equated to the bank scale in a single step. STAAR Alternate 2 uses Wright's (1977) common-items equating procedure to transform the Rasch difficulty of the field-test items to the same Rasch scale as the common items. Because the Rasch scale of the common items had previously been equated to the base scale, the equated field-test items are also on the base scale.

Reliability

Reliability indicates the precision of test scores, which also reflects the consistency of test results across testing conditions. The degree to which results are consistent is assessed using a reliability coefficient. The concept of reliability is based on the idea that repeated administrations of the same assessment should generate consistent results. Reliability is a critical technical

characteristic of any measurement instrument because unreliable scores cannot be interpreted in a valid way. There are many methods for estimating test score reliability, including some that require multiple assessments to be administered to the same sample of students. Because obtaining these types of reliability estimates is burdensome on schools and students, reliability estimation methods that require only one test administration have been developed and are commonly used for large-scale assessments, including STAAR, STAAR Alternate 2, TELPAS, and TELPAS Alternate.

Internal Consistency Estimates

Reliability coefficients based on one test administration are known as internal consistency measures because they measure the consistency with which students respond to the items within the test. As a general rule, reliability coefficients from 0.70 to 0.79 are considered adequate, those from 0.80 to 0.89 are considered good, and those at 0.90 or above are considered excellent. However, what is considered appropriate might vary in accordance with how assessment results are used (e.g., for low-stakes or high-stakes purposes). The following types of internal consistency measures are used to estimate the reliability of the components of the Texas Assessment Program:

- Kuder-Richardson 20 (KR20) is used for tests with only dichotomously scored items.
- Stratified coefficient alpha is used for tests containing a mixture of dichotomously scored and polytomously scored items.

KR_{20} is a mathematical expression of the classical test theory definition of test score reliability as the ratio of true score variance (i.e., no measurement error) to observed score variance (i.e., measurement error included). The classical test theory concept of reliability, in general, can be expressed as:

$$P'_{XX} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2+\sigma_E^2}\,, \tag{11}$$

where the reliability P'_{XX} of test X is a function of the ratio between true score variance σ_T^2 and observed score variance σ_X^2, which is further defined as the sum of the true score variance and error variance $\sigma_T^2 + \sigma_E^2$. As error variance is reduced, reliability increases (i.e., students' observed scores are more precise estimates of their true scores). KR_{20} can be mathematically represented as:

$$KR_{20} = \left[\frac{k}{k-1}\right]\left[\frac{\sigma_X^2-\sum_{i=1}^{k} p_i(1-p_i)}{\sigma_X^2}\right], \tag{12}$$

where KR_{20} is a lower-bound estimate of the true reliability, k is the number of items in test X, σ_X^2 is the observed score variance of test X, and p_i is the proportion of students who answered item i correctly. This formula is used when test items are dichotomously scored.

Coefficient alpha (also known as Cronbach's alpha) is an extension of KR_{20} to cases where items are polytomously scored (in more than two possible score categories) and is computed as follows:

$$\alpha = \left[\frac{k}{k-1}\right]\left[1 - \frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma_X^2}\right], \quad (13)$$

where α is a lower-bound estimate of the true reliability, k is the number of items in test X, σ_X^2 is the observed score variance of test X, and σ_i^2 is the observed score variance of item i.

The stratified coefficient alpha is an extension of coefficient alpha used when a mixture of item types appears on the same test. In computing the stratified coefficient alpha as an estimate of reliability, each item-type component is treated as a subtest. Given the small N counts for non-multiple-choice items, items are subset by multiple choice versus non-multiple choice. A separate measure of reliability is computed for each component and combined as follows:

$$Stratified\ \alpha = 1 - \frac{\sum_{j=1}^{c}\sigma_{X_j}^2(1-\alpha_j)}{\sigma_X^2}, \quad (14)$$

where c is the number of item-type components, α_j is the estimate of reliability for each item-type component, $\sigma_{X_j}^2$ is the observed score variance for each item-type component j, and σ_X^2 is the observed score variance for the total score. For components comprising multiple-choice and non-multiple-choice items, coefficient alpha is used as the estimate of component reliability. The correlation between ratings of the first two raters (i.e., inter-rater reliability) is used as the estimate of component reliability for written responses.

Inter-Rater Reliability

Some assessments require different types of reliability evidence than those described above. For example, STAAR RLA assessments include an extended constructed-response question at all grade levels. As part of the process for evaluating the reliability of such assessments, TEA provides evidence that the evaluation of student performance is appropriately conducted.

To gather such evidence of inter-rater reliability, two evaluators independently score the same student response. If the scores from the two scorers differ by more than one point, then a third evaluation is conducted by a supervisor or scoring director to resolve the discrepancy. These scores can then be analyzed, and the extent of agreement (or correlation) between the two sets of scores can be calculated. The correlation between the two sets of ratings is considered to be a measure of the reliability of the test scores.

Measurement Error

Test scores for the Texas Assessment Program are typically highly reliable; however, each test score contains an associated measurement error, which is the part of the test score that is not associated with the characteristic of interest. The measurement error associated with test scores can be broadly categorized as systematic or random. Systematic errors are caused by a

particular characteristic of the student or test that has nothing to do with the construct being measured, and they affect scores in a consistent manner (i.e., making scores lower or higher). An example of a systematic error would be a language barrier that caused a student to incorrectly answer questions to which the student knew the answer. By contrast, random errors are chance occurrences that may increase or decrease test scores. An example of a random error would be a student guessing the correct answer to a test question. TEA computes the classical standard error of measurement (SEM), the conditional standard error of measurement (CSEM), and classification consistency and classification accuracy for the purpose of estimating the amount of random error in test scores.

Standard Error of Measurement

The SEM reflects the amount of random variance in a score resulting from factors other than what the assessment is designed to measure. Because underlying traits such as academic achievement cannot be measured with perfect precision, the SEM is used to quantify the margin of uncertainty in test scores. For example, factors such as chance error and differential testing conditions can cause a student's observed score (the score achieved on a test) to fluctuate above or below his or her true score (the student's expected score). The SEM is calculated using both the standard deviation and the reliability of test scores, as follows:

$$SEM = \sigma_X \sqrt{(1 - P'_{XX})}\,, \tag{15}$$

where P'_{XX} is the reliability estimate (e.g., KR_{20}, coefficient alpha, stratified alpha) and σ_X is the standard deviation of raw scores on test X. A standard error provides some sense of the uncertainty or error in the estimate of the true score using the observed score. For example, suppose a student achieves a raw score of 50 on a test with a SEM of 3. Placing a one-SEM band around this student's score would result in a raw score range of 47 to 53. If the student takes the test 100 times, about 68 of those test raw scores will fall into the range of 47 to 53. In other words, the student's true score has a 68 percent probability of being in this range.

It is important to note that the SEM provides an estimate of the average test score error for all students regardless of their individual proficiency scores. It is generally accepted (e.g., refer to Peterson, Kolen, & Hoover, 1989) that the SEM varies across the range of student proficiencies. For this reason, it is useful to report not only a test-level SEM estimate but also individual score-level estimates. Individual score-level SEMs are commonly referred to as CSEMs.

Conditional Standard Error of Measurement

Like the SEM, the CSEM reflects the amount of variance in a score resulting from random factors other than what the assessment is designed to measure, but it provides an estimate conditional on proficiency. In other words, the CSEM provides a measurement error estimate at each score point on an assessment. The CSEM is usually smallest (and thus scores are most reliable) near the middle of the score distribution because achievement tests typically include a relatively large number of moderately difficult items (compared to easy or difficult items), and such items provide more precise information about student proficiency near the middle of the score distribution.

IRT methods for estimating score-level CSEM are used because test- and item-level difficulties for STAAR, STAAR Alternate 2, and TELPAS are calibrated using the Rasch measurement model. By using CSEMs that are specific to each scale score, a more precise error band can be placed around each student's observed score.

Classification Consistency and Accuracy

Test scores are used to classify students into performance levels. Because all test scores contain errors, the classifications also have errors. Usually there are two indicators to evaluate the quality of classifications: consistency and accuracy. Consistency refers to the percentage of students who are classified into the same performance levels if they took two parallel forms of a test, while accuracy refers to the percentage of students who are correctly classified into their true performance levels based on their observed scores on a test. Classification consistency and accuracy are two related but different concepts; high consistency does not necessarily lead to high accuracy, and vice versa. To better understand the classification quality, TEA conducts an analysis of the consistency and accuracy of student classifications into performance levels based on results of tests for which performance standards have been previously established.

The classification consistency index developed for IRT models (Lee, 2010) is used in this section. The basic idea is to estimate the probability of classifying into each performance level conditional on each test raw score based on an IRT model. For a performance level and a raw score, the probability that the raw score is classified into the same performance level on two parallel forms is just the square of the above probability for one test. Across all performance levels, the probability that a raw score is consistently classified on two parallel forms is the sum of the above probabilities for two tests and one performance level. The consistency index for a test is then the sum of the above probabilities over all raw scores weighted by the observed percentages of students on each raw score. The mathematical formula of consistency index can be expressed as:

$$\hat{\varphi} = \sum_{r=0}^{r_5} \left\{ \sum_{l=1}^{3} \left[\sum_{x=r_l}^{r_{l+1}-1} \hat{p}(x \mid \hat{\theta}_r) \right]^2 + \left[\sum_{x=r_4}^{r_5} \hat{p}(x \mid \hat{\theta}_r) \right]^2 \right\} f_r \,, \tag{16}$$

where l is the performance level (for STAAR, 1 = Did Not Meet, 2 = Approaches, 3 = Meets, 4 = Masters); r_l and r_{l+1} are the raw score cut scores for level l and l+1, respectively, with $r_1 = 0$ and r_5 = maximum possible test raw score; $\hat{\theta}_r$ is the estimated proficiency score associated with raw score r ; $\hat{p}(x \mid \hat{\theta}_r)$ is the estimated probability of getting raw score x conditional on $\hat{\theta}_r$; and f_r is the percentage of students with raw score r. The probability, $\hat{p}(x \mid \hat{\theta}_r)$, can be estimated based on the following recursive algorithm:

$$\hat{p}(x \mid \hat{\theta}_r) = \sum_{m=0}^{M_i - 1} \hat{p}_{i-1}(x - m \mid \hat{\theta}_r) \hat{p}_{im}(\hat{\theta}_r) \,, \tag{17}$$

where i refers to the ith item in a test; x is a raw score in a performance level, which is between the minimum ($\min_i$) and maximum ($\max_i$) scores after adding the ith item; M_i is the number of score categories for item i; $\hat{p}_{im}(\hat{\theta}_r)$ is the estimated probability of getting score m on item i conditional on $\hat{\theta}_r$ which is calculated based on the RPCM (equation 1); and $\hat{p}_i(x \mid \hat{\theta}_r)$ is the estimated probability of getting score x conditional on $\hat{\theta}_r$ after adding the ith item.

Note that $\hat{p}(x \mid \hat{\theta}_r) = 1$, and when $x - m < \min_{i-1}$ or $x - m < \max_{i-1}$ for $i > 1$, then define $\hat{p}_{i-1}(x-m \mid \hat{\theta}_r) = 0$.

The method recommended by Rudner (2000, 2005) is adapted here for computing classification accuracy. Under an IRT model, for an estimated proficiency score $\hat{\theta}_r$ associated with raw test score r, the true proficiency score θ_r is expected to be normally distributed with a mean of $\hat{\theta}_r$ and an estimated standard deviation of $\hat{\sigma}_{\theta_r}$ (i.e., the CSEM). The estimated proficiency score cut score $\hat{\theta}_l$ for each performance level l is also available. For each raw score point in a performance level, the probability of correctly classifying into this level can then be estimated. The accuracy index is the sum of these probabilities across all raw scores weighted by the observed percentages of students on each raw score point, f_r. In particular, the estimation formula is written as:

$$\hat{\psi} = \sum_{l=1}^{3} \sum_{r=r_l}^{r_{l+1}-1} \left[\phi\left(\frac{\hat{\theta}_{l+1} - \hat{\theta}_r}{\hat{\sigma}_{\theta_r}} \right) - \phi\left(\frac{\hat{\theta}_l - \hat{\theta}_r}{\hat{\sigma}_{\theta_r}} \right) \right] f_r + \sum_{r=r_4}^{r_5} \left[\phi\left(\frac{\hat{\theta}_{l+1} - \hat{\theta}_r}{\hat{\sigma}_{\theta_r}} \right) - \phi\left(\frac{\hat{\theta}_l - \hat{\theta}_r}{\hat{\sigma}_{\theta_r}} \right) \right] f_r , \quad (18)$$

where ϕ is the cumulative standard normal distribution function and θ_l is the proficiency score cut score for level l with $\theta_{l=1} = -10$ and $\theta_{l=5} = 10$.

Note that each STAAR EOC assessment has three different Approaches level cut scores: one for students who first took an EOC assessment before the December 2015 administration, one for students who first took an EOC assessment on or after the December 2015 administration and before spring 2023, and one for students who first took an EOC assessment in spring 2023. Therefore, for each EOC assessment, first the classification consistency and accuracy for each group of students who have the same Approaches cut score (i.e., "Approaches 2012–2015," "Approaches 2016–2022," or "Approaches") are estimated, and then the classification consistency and accuracy indexes weighted by proportion of students in each group as the overall classification consistency and accuracy estimate for a test are summed.

Validity

Validity refers to the extent to which test scores accurately measure what the test is intended to measure. The results of STAAR, including STAAR Spanish, and STAAR Alternate 2 are used to make inferences about how well students know and understand the TEKS curriculum. Similarly, TELPAS and TELPAS Alternate test results are used to make inferences regarding English language acquisition aligned with the ELPS. When test scores are used to make inferences about student achievement, it is important that the assessment supports those inferences. In

other words, the assessment should measure what it was intended to measure in order for inferences about test results to be valid.

Validity evidence can be organized into five categories: test content, response processes, internal structure, relations to other variables, and consequences of testing (AERA, APA, NCME, 2014; Schafer, Wang, & Wang, 2009). Such evidence supports the valid interpretation and use of test scores; however, validation is a matter of degree and is an ongoing process.

Evidence Based on Test Content

Validity evidence based on test content supports the assumption that the content of the test adequately reflects the intended construct. For example, STAAR and STAAR Spanish test scores are designed to help make inferences about students' knowledge and understanding of the statewide curriculum standards, the TEKS. Therefore, evidence supporting the content validity of STAAR maps the test content to the TEKS. Validity evidence supporting test content comes from the established test development process and the judgment of content experts about the relationship between the items and the test construct.

The test development process starts with a review of the TEKS by Texas educators. The educators then work with TEA to identify the readiness and supporting standards in the TEKS and help determine how each standard can best be assessed. A test blueprint developed with educator input maps the items to the reporting categories they are intended to represent. Items are then developed based on the test blueprints.

The steps in the test development process followed each year to support the validity of test content for the Texas Assessment Program are:

- Items are developed based on the TEKS curriculum standards and item guidelines.
- Items are reviewed for appropriateness of item content and difficulty, for alignment to the TEKS, and to eliminate potential bias.
- Data on field-test items is collected and reviewed to determine appropriateness for inclusion on a test.
- Tests are built to pre-defined criteria.
- University-level experts review high-school assessments for accuracy of the advanced content.

A more comprehensive description of the test development process is available in Chapter 2, "Building a High-Quality Assessment System."

Evidence Based on Response Processes

Response processes refer to the cognitive behaviors required to respond to a test item. Texas collects evidence showing that the manner in which students are required to respond to test items supports an accurate measurement of the construct of interest.

For example, STAAR RLA assessments include extended constructed-response items because requiring students to respond to open-ended writing questions reflects an appropriate manner for students to demonstrate their writing abilities. Student response processes for the components of the Texas Assessment Program differ by item type.

STAAR requires students to respond to various item types, including multiple choice, technology enhanced, short constructed response, and extended constructed response. STAAR Alternate 2 involves test administrators observing students as they respond to standardized items and scoring the items based on item-specific rubrics. TELPAS grades 2–12 requires students to respond to multiple-choice items, technology-enhanced items, and performance-based speaking tasks. Holistic assessment for TELPAS kindergarten and grade 1 and TELPAS Alternate do not contain traditional items; instead, students are evaluated and assigned holistic ratings based on ongoing classroom observations.

TEA gathers evidence to support validity based on response processes from several sources. When new item types or changes to the format of existing item types are considered for any assessments, cognitive labs are used to study the way students engage with the various item presentations. In this setting, students “think aloud” while responding to assessment items. This can provide evidence that students’ cognitive processes are consistent with those expected for a given item type and that they reflect the knowledge and skills described in the TEKS. After evaluation in the cognitive lab setting, test items are pilot-tested with a larger sample of students to gather information about performance on new item types and formats. Once new item types and formats are determined to be appropriate, evidence including statistical information (e.g., item difficulty, point-biserial correlations, DIF) is gathered about student responses through field testing. The evidence is then submitted for content expert review.

The process used to score items can provide validity evidence related to response processes. For assessments with constructed-response items, human scorers use rubrics to score student responses. For TELPAS speaking, the responses are scored by an automated scoring engine. The validity of student scores is supported if such rubrics accurately describe the characteristics of student responses on a continuum from low to high quality. All rubrics for STAAR, including STAAR Spanish, have been validated by educator committees and content experts. In addition, TEA has implemented a rigorous scoring process for constructed-response items that includes training and qualification requirements for human scorers, ongoing monitoring during scoring, adjudication and resolution processes for student responses that do not meet the perfect or adjacent scoring requirements, and rescoring of responses as needed. A more comprehensive description of the scoring process for constructed-response items is available in Chapter 2, “Building a High-Quality Assessment System.”

Evidence Based on Internal Structure

When a test is designed to measure a single construct, the internal components of the test should exhibit a high level of homogeneity that can be quantified in terms of the internal consistency reliability coefficients. Internal consistency estimates are evaluated for reported groups, including all students, female students, male students, Black or African American

students, Hispanic or Latino students, and White students. Estimates are made for the full assessment, as well as for each reporting category within a content area.

Validity studies have also been conducted to evaluate the structural composition of assessments, such as the comparability between two language versions of the same test. For example, a study conducted on the structural equivalence of transadapted tests (Davies, O'Malley, & Wu, 2007) provided evidence that the English and Spanish versions of the components of the Texas Assessment Program were measuring the same construct, which supports the internal structure validity of the tests.

Evidence Based on Relationships to Other Variables

Another source of validity evidence is the relationship between test performance and performance on another measure, sometimes referred to as criterion-related validity. The relationship can be concurrent, predictive, convergent, or discriminant:

- Concurrent—The performances on two measures taken at the same time are correlated.
- Predictive—The current performance on one measure predicts performance on a future measure.
- Convergent—The performances on two measures that are meant to assess the same or similar construct should be strongly correlated.
- Discriminant—The performances on two measures that are meant to assess unrelated constructs should have a weak correlation or no correlation.

Several past and current research studies have been designed to evaluate the relationship between performance on STAAR and performance on other related tests or criteria, including the following:

- STAAR to TAKS comparison studies, which link performance on STAAR to performance on TAKS (e.g., STAAR grade 7 mathematics to TAKS grade 7 mathematics)
- STAAR linking studies, which link performance on STAAR across grade levels or courses in the same content areas (e.g., grade 4 RLA to grade 5 RLA, English I to English II)
- STAAR intercorrelation estimates, which evaluate the strength of the relationship (or lack thereof) among scores on STAAR across different content areas (e.g., grade 4 mathematics to grade 4 RLA, English I to Biology)
- grade correlation studies, which link performance on STAAR EOC assessments to course grades
- validity studies, which link performance on STAAR to other measures (e.g., Scholastic

Aptitude Test [SAT], American College Testing [ACT], Lexiles, Quantiles, STAAR Interim Assessments)

- college students taking STAAR studies, which link performance on STAAR EOC assessments to grades in college courses

For detailed descriptions and results of such studies, refer to the Assessment Reports and Studies webpage.

STAAR Alternate 2 intercorrelation estimates are calculated to evaluate the strength of the relationship between scores on STAAR Alternate 2 across different content areas. Results from all these analyses are provided in Appendix C.

To examine validity evidence based on external measures for TELPAS, an annual analysis is conducted on the relationship between TELPAS reading and writing performance and STAAR RLA performance. For each grade level and TELPAS proficiency level breakout group, the following two types of performance data are examined:

- average STAAR scale scores
- STAAR passing rates (Approaches Grade Level performance)

Refer to Chapter 6, “TELPAS,” for more details. The same analysis is also conducted on the relationship between TELPAS Alternate and STAAR Alternate 2. Refer to Chapter 7, “TELPAS Alternate,” for more details.

Evidence Based on Consequences of Testing

Consequential validity refers to the idea that the validity of an assessment program should account for both intended and unintended consequences resulting from inferences based on test scores. For example, STAAR is intended to have an effect on instructional content and delivery strategies; however, an unintended consequence could be the narrowing of instruction, a phenomenon sometimes referred to as "teaching to the test." Consequential validity studies in Texas use surveys to collect input from various assessment program stakeholders to measure the intended and unintended consequences of the assessments.

Given the important stakes associated with the Texas Assessment Program, the validity of interpretations and uses of test scores are critical. The intended interpretations of test results are stated in the policy definitions of the performance levels, which are provided on the STAAR Performance Standards webpage.

Measures of Student Progress

Measures of student progress describe changes in student performance across time. The overall description of student achievement can be enhanced by providing student progress measures that convey information about how performance in the current year compares to performance in the prior year. For example, consider a student who achieves Approaches

Grade Level on a STAAR assessment. The interpretation of Approaches Grade Level performance would depend on the performance that the student achieved in the previous year. If the student achieved Did Not Meet Grade Level in the previous year, then the student made notable progress this year by advancing a performance level. However, if the student achieved Meets Grade Level in the previous year, then the interpretation of Approaches Grade Level performance this year would be quite different because the student regressed.

Development of Progress Measures

Several types of progress measures were considered for use with STAAR and STAAR Alternate 2, including student growth models based on regression, student growth percentile, growth to proficiency, value/transition tables, and gain scores. These student growth models differ in the types of information used, the complexity of the calculations, the feedback provided, and the ease with which they can be explained. These factors are all important to consider when selecting a model for measuring student progress.

As part of the development of STAAR and STAAR Alternate 2 progress measures, several factors were considered, including:

- the suitability of different models for measuring student progress given the characteristics of STAAR and STAAR Alternate 2,
- the appropriateness of progress measures given the content relationships among STAAR and STAAR Alternate 2,
- the usability of progress measures for accountability given federal and state requirements, and
- the effectiveness of communicating progress-measure results given various reporting options.

Additionally, input was sought from a number of advisory groups regarding the development of progress measures for STAAR and STAAR Alternate 2. Several options for progress measures were presented to the Texas Technical Advisory Committee (TTAC), a national group of educational measurement experts who provided recommendations and guidance. Progress measures were also discussed with the Accountability Technical Advisory Committee (ATAC) and the Accountability Policy Advisory Committee (APAC), which are groups composed of educators from various Texas campuses, districts, and ESCs, as well as parents, higher education representatives, business leaders, and legislative representatives. Input from these groups was requested at several points during the development of progress measures for STAAR and STAAR Alternate 2.

Implementation

Based on the input and considerations described earlier, gain score was selected as the progress measure for STAAR. The STAAR Progress Measure was implemented for the first time in the 2012–2013 school year beginning with STAAR and STAAR Spanish mathematics

and reading. Since then, Algebra I, English I, and English II have been added to the STAAR Progress Measure, which has been reported every year except for the years when performance standards have been reestablished.

In addition to the STAAR Progress Measure, TEA also produces an on-track measure, which provides information about whether a student is on track to be at or above the Meets Grade Level performance standard in a future target year. Using gain scores, individual students are categorized as Not On Track or On Track toward the target year. On-track measures are available for STAAR and STAAR Spanish mathematics and RLA.

The STAAR Alternate 2 Progress Measure employs a transition table approach and was reported for the first time in in 2016 with the mathematics and reading assessments. STAAR Alternate 2 progress measures were calculated and reported for mathematics and RLA assessments.

Details about these progress measures can be found in Chapter 4, "STAAR," and Chapter 5, “STAAR Alternate 2,” and on the Progress Measures webpage.

Sampling

Sampling is a procedure that is used to select and examine a small set that is representative of the population from which it is drawn. The results from well-drawn samples allow TEA to estimate characteristics of the Texas student population as a whole. Through the careful selection of student samples, TEA is able to make reliable and valid inferences about student performance on its assessments while minimizing the burden on campuses and districts.

Key Concepts of Sampling

A target population is the set of students to which the results should generalize, also known as the complete collection of objects of interest (Lohr, 1999). For example, consider a study with the goal of understanding how grade 3 EB students perform on a set of test questions. In that case, the target population would be all grade 3 EB students in Texas. Careful consideration is given to defining the target population before sampling takes place.

A sampling unit is the unit to be sampled from the target population. A sampling unit could be a student, a campus, a district, or even a region. For example, if 20 campuses are randomly chosen from a list of all campuses in the state, then the campus is the sampling unit.

An observation unit is the unit on which data are actually collected. An observation unit might or might not be the same as the sampling unit. For example, a study designed to estimate the number of computers per campus in the entire state might involve requesting that each of 20 randomly selected campuses report the number of computers it has. In this case, the campus is both the sampling unit and the observation unit. By comparison, consider a study designed to estimate student computer access in the entire state, in which each of the same 20 sampled campuses is requested to report student data on how many students have computer access at home. In that case, even though the sampling unit is still the campus

(because 20 campuses were selected), the observation unit is the student (because the data being collected reflect student characteristics).

Reasons for Sampling

The Texas Assessment Program employs sampling instead of studying entire target populations for several reasons, including the following:

- Accessibility—There are situations where collecting data on every member of the target population is not feasible.
- Burden—Sampling minimizes the participation requirements for the campus and district, thereby reducing the testing burden.
- Cost—It is more cost efficient to obtain data for a carefully selected subset of a population than it is to collect the same data for the entire population.
- Size—It is more efficient to examine a representative sample when there is a large target population.
- Time—Using sampling to study the target population is less time consuming. Sampling might be needed when the timeline of the analysis is important.

Sampling Designs

The Texas Assessment Program uses sampling to collect data for the purpose of field testing, audits, and research studies (e.g., linking studies, cognitive labs, comparability studies). Results from field testing are used to evaluate statistical properties of newly developed test items that have not yet been used on an operational test form. Audits allow for the collection of information from school districts that can be used to evaluate training, administration, and scoring of the assessments. Research studies generally involve assessing a sample of students under various testing conditions to collect evidence to support the technical quality of and make improvements to the Texas Assessment Program. TEA uses the following sample designs.

Probability Sampling

In a probability sample, all sampling units have a known probability of being selected. Probability sampling requires that the number of sampling units in the target population is known. For example, if the student is the sampling unit, probability sampling would require an accurate list of all the students in the target population. The following are the major types of probability sampling designs:

- Simple Random Sampling—All sampling units in the target population have the same probability of being selected.
- Stratified Sampling— First the sampling units are grouped (i.e., stratified) according to variables of interest such as gender and ethnicity; then a random sample is selected from each group.

- Cluster Sampling— First the sampling units are grouped into clusters according to variables of interest; then, unlike stratified sampling, a predetermined number of clusters is randomly selected. All sampling units within the selected clusters are observed.

Regardless of the type of probability sampling used, a decision about whether to sample with or without replacement must be made. To help clarify this distinction, consider simple random sampling with replacement and simple random sampling without replacement. First, suppose that a simple random sample of size *n* with replacement is drawn from a population of size *N*. In this case, when a sampling unit is randomly selected, that unit remains eligible to be selected again. In other words, after the sampling unit is picked, it is put back and can be selected again. When sampling with replacement, a sampling unit might be selected multiple times and its data would be duplicated in the resulting sample of size *n*.

By comparison, suppose that a simple random sample of size *n* without replacement is drawn from a population of size *N*. In this case, once a sampling unit is chosen, it is ineligible to be selected again. In other words, after the sampling unit is picked, it is not put back. Thus, when sampling without replacement, each sample comprises *n* distinct, non-duplicate units from the population of size *N*.

Typically, sampling without replacement is preferred over sampling with replacement because duplicate data add no new information to the sample (Lohr, 1999). The method of sampling with replacement, however, is important in re-sampling and replication methods, such as bootstrapping.

Re-Sampling and Replication Methods: Bootstrapping

Bootstrapping is one of the re-sampling and replication methods that treats the sample like a population. These methods repeatedly draw pseudo-samples from samples to estimate the parameters of distributions. Thus, sampling with replacement is assumed with these methods. The bootstrap method was developed by Efron (1979) and described in Efron and Tibshirani (1993). The Texas Assessment Program uses bootstrapping methods when conducting comparability studies that compare online and paper versions of a test form.

Convenience (Nonprobability) Sampling

A sample that is created without the use of random selection is a convenience (or nonprobability) sample. Convenience samples are selected when it is impractical or impossible to collect a complete list of sampling units. When using convenience sampling, the list of sampling units is incomplete, and sampling units have no known probability of being selected. Convenience sampling introduces sources of potential bias into the resulting data, which makes it difficult to generalize results to the target populations.

TECHNICAL DIGEST 2022–2023

Chapter 4

State of Texas Assessments of Academic Readiness

Overview

Testing Requirements

Test Development

Accommodations

Training

Test Administration

Performance Standards

Scores and Reports

Measures of Student Progress

Scaling

Equating

Reliability

Validity

Sampling

Test Results

Overview

TEA, in collaboration with THECB and Texas educators, developed the STAAR program in accordance with educational requirements set forth by the Texas legislature in 2007 and 2009.

STAAR was implemented in the 2011–2012 school year and included Spanish versions of the assessments for grades 3–5.

STAAR is designed to measure the extent to which a student has learned and is able to apply the knowledge and skills defined in the TEKS. Every item is directly aligned to the TEKS currently in effect for the tested grade and subject or course. STAAR includes the following assessments:

- grades 3–8 mathematics,
- grades 3–8 RLA,
- grades 5 and 8 science,
- grade 8 social studies, and
- EOC assessments for:
    - Algebra I,
    - English I,
    - English II,
    - Biology, and
    - U.S. History.

Based on legislation passed in 2019, STAAR was redesigned to align more closely with effective classroom instruction. The redesign was implemented in the 2022–2023 school year and included:

- the addition of new non-multiple-choice questions that give students more ways to show their understanding and better reflect questions teachers ask in the classroom,
- the addition of a writing component to reading assessments for grades 3–8 to better support the interconnected way these subjects are taught, and
- the incorporation of more cross-curricular passages into the new RLA assessments so that test questions can reference topics students have learned about in other classes.

STAAR Spanish

STAAR Spanish is administered to eligible students for whom the language proficiency assessment committee (LPAC) determines that STAAR Spanish is the most appropriate way to measure those students' mastery of skills and is also available for students who receive academic instruction in Spanish while they learn English. The STAAR Spanish assessments are offered for grades 3–5 mathematics and RLA and for grade 5 science. The English and Spanish

versions of STAAR have the same test blueprint and assess the same TEKS student expectations for mathematics and science and similar student expectations in RLA.

STAAR Interim Assessments

STAAR Interim Assessments are a set of optional online assessments aligned to the TEKS; the purpose of the interim assessments is to monitor student progress and predict student performance on STAAR summative assessments. The interim assessments are available at no cost to districts and are not tied to accountability. More information is available on the STAAR Interim Assessments webpage.

Testing Requirements

All students enrolled in Texas public schools and open-enrollment charter schools are required to take STAAR unless the student meets the participation requirements for STAAR Alternate 2.

Students enrolled in grade 9 or below for the first time in the 2011–2012 school year or later are required to meet STAAR EOC assessment graduation requirements.

In 2015, legislation revised the state's assessment graduation requirements to allow an eligible student to receive a Texas high school diploma by means of an IGC if the student fails to pass no more than two STAAR EOC assessments. Eligibility criteria for an IGC can be found in TEC §28.0258.

The admission, review, and dismissal (ARD) committee makes educational decisions, including decisions related to state assessments and graduation requirements as described in TAC §89.1070, for students receiving special education services.

Due to the impact of the COVID-19 pandemic, STAAR testing was suspended for spring and summer 2020, and a STAAR EOC assessment waiver reduced the number of EOC assessments that a student was required to pass to meet assessment graduation requirements. To qualify for the waiver, a student must have:

- been enrolled in the course during spring or summer 2020,
- completed the full course by the end of spring or summer 2020, and
- earned full course credit by the end of the spring or summer 2020.

Test Development

Maintaining a high-quality student assessment program involves a complex and detailed test-development process, and TEA relies on input from educators to ensure that all measures of learning for Texas public school students are equitable and accurate. Test items for STAAR, including STAAR Spanish, are developed annually, reviewed by educator committees, field-tested, reviewed with their data, and, if approved, added to the STAAR item bank. In most cases, newly developed items are embedded in STAAR operational assessments each spring. However, stand-alone field tests are periodically required and have been administered in 2011, 2015, 2019, and 2022. For more information regarding each step of the STAAR test-development process, refer to Chapter 2, "Building a High-Quality Assessment System,"

which outlines the processes used to develop each STAAR assessment's framework and explains ongoing test development.

STAAR English-Spanish Alignment

TEA staff, Texas educators, and Spanish-language experts collaborate to develop STAAR Spanish test materials. STAAR Spanish RLA assessments are composed entirely of passages and items in Spanish. This development approach allows the Spanish RLA curriculum to be assessed in a more authentic and meaningful manner. Items for STAAR Spanish mathematics and science are transadapted. Transadaptation involves translating items from English and adapting them as necessary to ensure cultural and linguistic accessibility. Spanish bilingual educators then review all original and transadapted test items in accordance with the educator review process described in Chapter 2, "Building a High-Quality Assessment System."

The following practices reinforce alignment of the STAAR English and Spanish assessments:

- When the performance standards for STAAR were established, standard-setting panels reviewed both the English and Spanish grades 3–5 RLA assessments to establish comparable performance standards.
- The development and review processes for the RLA assessments in English and Spanish are parallel (i.e., item reviews for English and Spanish include judgments related to each item's alignment to the TEKS). Field-test data reviews for English and Spanish items also include item statistics reviews based on actual student performance. These safeguards ensure that only psychometrically sound items are selected for inclusion in the STAAR item banks.
- Each year, STAAR development staff review the newly developed test items, focusing on the best ways to assess the TEKS and further enhancing the alignment between the English and Spanish assessments.
- The RLA assessments in English and Spanish are constructed concurrently and in coordination, and they adhere to the same test construction guidelines regarding the range of item content and cognitive complexity.
- The Spanish mathematics and science assessments are transadapted from the corresponding English assessments. The item-writing and review processes for transadapted items ensure that the Spanish items are linguistically and culturally appropriate and that the interpretations of grade-level performance expectations are the same for English and Spanish.
- The test blueprints for the English and Spanish assessments are the same, including the number of items that assess each reporting category and the number of items on the test.

Accommodations

The goal of STAAR accommodations is to ensure that each student can interact appropriately with the content, presentation, and response modes of the state assessments. To meet this goal, STAAR accommodations are designed to allow all students to demonstrate their

knowledge of the content being assessed without the format of the assessment, the non-tested language, or the type of response needed to answer the questions being barriers. The various accommodations made available on STAAR are also designed to be the same or similar to those accommodations commonly used during classroom instruction.

Accommodation policies for STAAR, including STAAR Spanish, are divided into three main categories: accessibility features, locally-approved designated supports, and designated supports requiring TEA approval. More information is available on the Accommodations page of the *Coordinator Resources*.

Accessibility Features

Accessibility features may be provided to students based on their needs. In general, these procedures and materials are available to any student who regularly benefits from their use during instruction; however, a student cannot be required to use them during STAAR. Coordinators are responsible for ensuring that test administrators understand the proper implementation of these procedures and use of these materials. In some cases, a student may need to complete the test in a separate setting to eliminate distractions to other students and to ensure that the security and confidentiality of the test are maintained.

Locally-Approved Designated Supports

Locally-approved designated supports include accommodations that may be made available to students who meet eligibility criteria. The appropriate team of people at the campus level (e.g., Response to Intervention [RtI] team; LPAC; Section 504 committee; ARD committee) determines eligibility as indicated in each policy document. The decision to allow the use of a designated support during STAAR should be made on an individual student basis, taking into consideration the needs of the student and whether the student routinely receives the support during classroom instruction and classroom testing. In addition, the support has to have been proven effective in meeting the student's specific needs, as evidenced by student scores or teacher observations.

Designated Supports Requiring TEA Approval

These designated supports require the submission of an Accommodation Request Form to TEA. The appropriate team of people at the campus level, as indicated in each policy document, determines whether the student meets all the specific eligibility criteria and, if so, submits an Accommodation Request Form to TEA. TEA must receive Accommodation Request Forms according to the posted deadlines. Late requests will not be processed unless circumstances involving the student change after the deadline (e.g., newly enrolled student, medical emergency, updated ARD committee decision). The request must be approved by TEA before a student can use the designated support on STAAR.

Training

TEA develops instructional materials, including manuals, guides, presentations, online modules, and videos, to support the training of all testing personnel on test security and administration procedures. Preparation for test administration begins every year with a TEA-provided training-of-trainers session for testing coordinators from each of the 20 Texas regional ESCs as well as district testing coordinators from the state's 25 largest districts. Using materials and information

provided in the TEA training session, ESC regional testing coordinators train the district coordinators in their respective regions. District coordinators then train their campus testing coordinators, who are responsible for training test administrators.

Test security and administration procedures provided in the *Coordinator Resources* and the *STAAR Test Administrator Manual* must be followed so that all students have an equal opportunity to demonstrate their academic knowledge and skills. The *Coordinator Resources* guide district and campus coordinators through their responsibilities as they oversee the administration of the Texas Assessment Program. This online resource contains preparation and administration procedures for each state-required assessment and is available prior to the annual ESC training.

Test Administration

All STAAR assessments—grades 3–8 mathematics, grades 3–8 RLA, grades 5 and 8 science, grade 8 social studies, Spanish grades 3–5 mathematics, Spanish grades 3–5 RLA, Spanish grade 5 science, Algebra I, English I, English II, Biology, and U.S. History—are administered online in the spring. STAAR EOC assessments are also administered online in June and December. A paper version of STAAR is available for students with a special circumstance. The number of students tested for each STAAR assessment is shown in Table 4.1.

Table 4.1. STAAR Assessments Administered in 2022–2023

STAAR Assessment	Assessments Administered
Grade 3 mathematics	370,006
Grade 3 RLA	356,558
Grade 3 Spanish mathematics	16,454
Grade 3 Spanish RLA	30,213
Grade 4 mathematics	373,988
Grade 4 RLA	365,035
Grade 4 Spanish mathematics	11,497
Grade 4 Spanish RLA	21,694
Grade 5 mathematics	378,663
Grade 5 RLA	372,677
Grade 5 science	378,742
Grade 5 Spanish mathematics	8,483
Grade 5 Spanish RLA	15,991
Grade 5 Spanish science	9,775
Grade 6 mathematics	384,766
Grade 6 RLA	391,376
Grade 7 mathematics	331,698
Grade 7 RLA	400,416

STAAR Assessment	Assessments Administered
Grade 8 mathematics	364,110
Grade 8 RLA	410,472
Grade 8 science	407,847
Grade 8 social studies	414,692
Algebra I	608,559
English I	712,965
English II	612,128
Biology	544,720
U.S. History	427,007

NOTE: For the STAAR EOC assessments, the table includes the sum total of the December, spring, and June administrations.

The Test Delivery System

STAAR online assessments are administered using the Test Delivery System (TDS). TDS includes the Test Administrator Interface, which is used for scheduling test sessions; the Student Interface, which allows students to participate in testing; and the Secure Browser application, which provides a secure online environment for testing. TDS allows for the secure transfer and storage of test data while remaining scalable to support the student testing population. The TDS architecture has demonstrated stability and efficiency by seamlessly handling over 1.2 million concurrent users.

Make-up Testing

Make-up testing opportunities for students who are absent on the day of testing are available during the STAAR testing window for all grades, subjects, and courses.

Out-of-District Testing

For STAAR EOC assessments, students who are unable to test in their home district are allowed to test out-of-district (OOD). For example, a student from Houston who spends the summer in Dallas could register to test in Dallas. OOD students are required to complete registration within a set window so that receiving districts are aware of the student's intent and have the resources to administer the assessment. Students must present photo identification at the test administration site on the day of the test.

Out-of-School Testing

Examinees who have not passed a STAAR EOC assessment and are no longer enrolled in school but have otherwise completed requirements for graduation may take an assessment during a test administration at a participating district.

Performance Standards

Performance standards directly relate levels of test performance to what students are expected to learn, as defined in the statewide curriculum. Standard setting is the process of establishing cut scores that define the performance levels on an assessment.

Performance Levels and Policy Definitions

For STAAR, including STAAR Spanish, the performance levels and policy definitions are as follows:

Did Not Meet Grade Level

Performance in this category indicates that students are unlikely to succeed in the next grade or course without significant ongoing academic intervention. Students in this category do not demonstrate a sufficient understanding of the assessed knowledge and skills.

Approaches Grade Level

Performance in this category indicates that students are likely to succeed in the next grade or course with targeted academic intervention. Students in this category generally demonstrate the ability to apply the assessed knowledge and skills in familiar contexts.

Meets Grade Level

Performance in this category indicates that students have a high likelihood of success in the next grade or course but may still need some short-term, targeted academic intervention. Students in this category generally demonstrate the ability to think critically and apply the assessed knowledge and skills in familiar contexts.

Masters Grade Level

Performance in this category indicates that students are expected to succeed in the next grade or course with little or no academic intervention. Students in this category demonstrate the ability to think critically and apply the assessed knowledge and skills in varied contexts, both familiar and unfamiliar.

Standard Setting

The STAAR program's goal was to have a comprehensive assessment system with curriculum standards and performance standards that were vertically aligned within a content area (i.e., the curriculum and performance standards linked from the high school courses back to the middle school and elementary school grades and subjects). Standard setting for STAAR took into consideration a variety of factors, such as policy, TEKS content standards, educator knowledge about what students should know and be able to do, and information about how student performance on state assessments aligns with performance on other assessments. Standard-setting committees comprised diverse groups of stakeholders who carefully considered the interaction of these elements for each STAAR assessment. The task of each standard-setting committee was to recommend cut scores that would define the performance levels for each STAAR assessment.

Initial performance standards for all STAAR assessments were established in 2012, and performance standards were reset in 2023 with the redesign of STAAR. The current performance standards for STAAR are provided in Tables 4.2 and 4.3.

Table 4.2. STAAR Grades 3–8 Performance Standards

Assessment	Approaches Grade Level (Scale Score)	Meets Grade Level (Scale Score)	Masters Grade Level (Scale Score)
Grade 3 Mathematics	1360	1471	1600
Grade 4 Mathematics	1462	1557	1690
Grade 5 Mathematics	1515	1634	1776
Grade 6 Mathematics	1616	1745	1889
Grade 7 Mathematics	1703	1793	1965
Grade 8 Mathematics	1754	1859	2009
Grade 3 RLA	1345	1467	1596
Grade 4 RLA	1414	1552	1663
Grade 5 RLA	1475	1592	1700
Grade 6 RLA	1535	1634	1749
Grade 7 RLA	1564	1669	1771
Grade 8 RLA	1592	1698	1803
Grade 5 Science	3550	4000	4380
Grade 8 Science	3550	4000	4619
Grade 8 Social Studies	3550	4000	4352
Grade 3 Spanish Mathematics	1360	1471	1600
Grade 4 Spanish Mathematics	1462	1557	1690
Grade 5 Spanish Mathematics	1515	1634	1776
Grade 3 Spanish RLA	1318	1447	1515
Grade 4 Spanish RLA	1408	1488	1581
Grade 5 Spanish RLA	1431	1556	1662
Grade 5 Spanish Science	3550	4000	4380

Table 4.3. STAAR EOC Assessments Performance Standards

Assessment	Approaches Grade Level 2012–2015 (Scale Score)	Approaches Grade Level 2016–2022 (Scale Score)	Approaches Grade Level (Scale Score)	Meets Grade Level (Scale Score)	Masters Grade Level (Scale Score)
Algebra I	3489	3541	3550	4000	4345
English I	3775	3775	3775	4000	4606
English II	3766	3775	3775	4000	4734
Biology	3516	3550	3550	4000	4531
U.S. History	3486	3536	3550	4000	4424

Refer to the STAAR standard-setting technical reports, which are available on the Assessment Reports and Studies webpage, for more information.

Scores and Reports

TEA publishes resources on both the TEA and Texas Assessment websites to assist school personnel in understanding and interpreting student performance data and to help parents understand their child's STAAR results. School personnel can access STAAR test results through the Centralized Reporting System (CRS), parents can access their child's STAAR results in the Family Portal, and the public can access STAAR statewide, region, district, and campus data using the Research Portal.

TEC §39.030 and TAC §101.3014 specify the requirements for maintaining the confidentiality of individual student results and for reporting district-level and campus-level results. The results of individual student performance on state assessments are confidential and may be released only in accordance with the Family Educational Rights and Privacy Act (FERPA). Districts must provide each student's state assessment results to the student, to his or her parent or guardian, and to his or her teacher for the applicable subject area. In addition, all state assessment results must be included in each student's academic achievement record.

Description of Scores

Scores for STAAR and STAAR Spanish include raw scores, scale scores, and the resulting performance level associated with the student's score. Additionally, percentiles, Lexiles, Quantiles, and English learner (EL) performance measures are provided.

The number of points that a student earns on a STAAR assessment is the student's raw score. A scale score is a conversion of the raw score onto a scale that is common to all test forms for that assessment. The scale score is used to determine whether a student achieved the Did Not Meet Grade Level, Approaches Grade Level, Meets Grade Level, or Masters Grade Level performance standard. Refer to Chapter 3, "Standard Technical Processes," for more information about raw scores and scale scores.

Percentiles represent the percentage of students across the state who took the assessment and received a scale score at or below the scale score of interest. Percentiles are calculated based on all students (except out-of-school [OOS] testers) who received valid scale scores on the assessment in the previous year's spring administration.

Students receive a Lexile Measure on the STAAR RLA assessments, including grades 3–8 RLA, Spanish grades 3–5 RLA, English I, and English II. Lexile measures indicate the level of difficulty of materials a student can read and range from below 0L for beginning readers to above 1600L. Similarly, students receive a Quantile Measure on STAAR mathematics assessments, including grades 3–8 mathematics, Spanish grades 3–5 mathematics, and Algebra I. Quantile measures indicate the mathematics concepts a student has learned and the concepts they are ready to learn next. These measures range from below 0Q to above 1400Q. More information about Lexiles and Quantiles is available on the Texas Assessment website.

Beginning in the 2018–2019 school year, qualifying EB students who tested in English also received an EL performance measure, which showed whether an eligible EB student was making sufficient progress on each STAAR content-area assessment based on predetermined

performance measure progress expectations. The EL performance measure was calculated and reported for all STAAR assessments except STAAR Spanish.

Assessment Reports

TEA provides reports of student performance on STAAR, including STAAR Spanish, to all Texas public school districts and open-enrollment charter schools. For each STAAR administration, student report cards, student labels, campus rosters, summary reports, and reporting data files are provided.

The spring administration of each assessment for STAAR, including STAAR Spanish, is released to the public through the Practice Test Site. To correspond with the released tests, TEA provides student item analysis reports and item analysis summary reports. These summary reports are available at the campus, district, region, and state level.

For more information about scoring and reporting for STAAR, refer to the Interpreting Results page of the *Coordinator Resources*.

Use of Test Results

Test results can be used to evaluate the performance of a group over time. Average scale scores and the percentage of students meeting the Approaches Grade Level, Meets Grade Level, and Masters Grade Level performance standards can be analyzed by grade and content area across administrations to provide insight into whether student performance is improving across years. For example, the average scale score for students who took the STAAR grade 4 RLA test can be compared over time.

Test results can also be used to compare the performance of different demographic or program groups. STAAR scores can be analyzed within the same content area of any single administration to determine, for example, which demographic or program group has the highest average scale score, which group has the lowest percentage meeting the Approaches Grade Level performance standard, and which group has the highest percentage achieving the Meets Grade Level performance standard. Other scores can be used to help evaluate the academic performances of demographic or program groups in core academic areas. For example, reporting category data can help districts and campuses identify areas of potential academic weakness for a group of students. The same methodology can be applied to an entire district or campus. Test results for groups of students can be used when evaluating instruction or programs that require average-score or year-to-year comparisons. The tests are designed to measure content areas within the required state curriculum, and so the consideration of test results by content area and reporting category might be helpful when evaluating curriculum and instructional programs. All test scores can be compared with statewide and regional performance within the same content area for any administration.

Test scores can also be used to identify where an individual student needs additional instruction or support in each subject. Other scores can provide information about a student's relative strengths or weaknesses in core academic areas. For example, reporting category–level data can provide information about a student's relative strengths or weaknesses and can be used to identify areas where a student might be having difficulty. This identification can help educators plan the most effective instructional intervention. Finally, individual student test scores are also used in conjunction with other performance indicators to assist in making placement decisions.

While scores can contribute to decisions regarding placement, educational planning for a student should take into account as much student information as possible.

Generalizations from test results can be made from the specific content area being measured on the test. However, because each test measures a finite set of skills with a limited set of items, any generalizations about student achievement derived solely from a particular test should be made with great caution and with full reference to the fact that the conclusions are based only on that test. Instruction and program evaluations should take into account as much information as possible, rather than relying on test scores alone, to provide a more complete picture of student performance.

Measures of Student Progress

Student progress measures provide information beyond performance levels by providing a comparison of performance over time. Whereas performance-level information describes students' current achievement, progress measures describe students' achievement across multiple years.

STAAR Progress Measure

The STAAR Progress Measure is legislatively mandated and was reported for the first time in the 2012–2013 school year. For STAAR, progress is measured as a student's gain score, which represents the difference between the scale score a student achieved in the prior school year and the scale score a student achieved in the current school year. These gain scores are then classified as Limited, Expected, and Accelerated in relation to progress targets. The progress targets define the expectation of annual progress for each grade and content area. These progress targets are grounded in the STAAR performance standards, the goal of having all students achieve Meets Grade Level or above, and having high-performing students maintain Masters Grade Level performance.

Specifically, for students who achieve Did Not Meet Grade Level, Approaches Grade Level, or Meets Grade Level performance standards in the prior year, the Expected progress target is defined as the distance between the Meets Grade Level performance standard on the prior year test and the Meets Grade Level performance standard on the current year test in the same content area. For students who achieve the Masters Grade Level performance standard in the prior year, the progress target is based on the distance between Masters Grade Level on the prior year test and Masters Grade Level on the current year test in the same content area.

The Accelerated progress classification is a designation reserved for those students who have demonstrated significant growth over the course of the year beyond that of the Expected progress target. The Accelerated progress target is defined as the distance between Meets Grade Level on the prior year test and Masters Grade Level on the current year test.

Students with gain scores less than the Expected progress target are classified as having achieved Limited progress. Students with gain scores greater than or equal to the Expected progress target and less than or equal to the Accelerated progress target are classified as having achieved Expected progress. Students with gain scores greater than the Accelerated progress target are classified as having achieved Accelerated progress.

At the extreme high and low ends of the scale, the application of the Limited, Expected, and Accelerated definitions would not be appropriate. At the extreme ends of the scale, unlike the rest of the scale, answering one more question correctly results in significant differences in scale scores. For this reason, several places on the scale have been identified as exceptions to the Limited, Expected, and Accelerated definitions.

Because the performance standards do not have the same numerical value across grades and content areas, the Expected and Accelerated progress targets differ from grade to grade and across content areas. Steps for calculating progress measures and progress targets for each STAAR grade and content area, including when students skip grade levels, can be found on the Progress Measures webpage.

Due to the STAAR redesign and updated performance standards, progress measures were not reported for STAAR assessments for the 2022–2023 school year.

STAAR On-Track Measure

While the STAAR Progress Measure accounted for performance from the prior year and the current year, it did not include any information about how the student was likely to perform in the future. Because this additional information may be helpful to students, teachers, and other stakeholders, TEA developed the STAAR on-track measure, which was reported for the first time in 2013–2014. The on-track measure used the STAAR Progress Measure and extrapolated performance into future years to determine if a student was on-track to achieve Meets Grade Level in a later grade or course. To calculate the STAAR on-track measure, three assessments covering the same content area must be available (i.e., previous year, current year, and target year). For example, the on-track measure can be calculated for STAAR grade 7 RLA (current year assessment) because the previous year assessment was STAAR grade 6 RLA and the target year assessment will be STAAR grade 8 RLA. Additional information about on-track measures can be found on the Progress Measures webpage.

Due to the STAAR redesign and updated performance standards, on-track measures were not reported for STAAR assessments for the 2022–2023 school year.

Scaling

Scaling is a statistical procedure that places raw scores on a common scoring metric to make test scores comparable across test administrations. Scaling associates numbers with characteristics of interest to provide information about measurable quantities for those characteristics. STAAR, including STAAR Spanish, uses the RPCM to place test items on the same Rasch scale across administrations for a given STAAR assessment. Once performance standards have been set for an assessment, the Rasch scale is then transformed to a more user-friendly metric to ease interpretation of the test scores. Details of the RPCM scaling method are provided in Chapter 3, “Standard Technical Processes.”

Reporting Scales

STAAR scale scores are reported on either a horizontal scale or a vertical scale. Horizontal scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations. Vertical scale scores allow for direct comparisons of student scores across grades within a content area. Student increases in vertical scale scores

provide information on the year-to-year growth of students. Refer to Chapter 3, "Standard Technical Processes," for detailed information about the scaling process.

Horizontal Reporting Scales

The following STAAR assessments are reported on horizontal scales:

- grade 5 science
- grade 8 science
- grade 8 social studies
- Spanish grade 5 science
- Algebra I, English I, English II, Biology, and U.S. History

For all STAAR assessments reported on a horizontal scale, a scale score of 4000 represents the Meets Grade Level performance standard. The Approaches Grade Level cut score was set to 3550 for all STAAR assessments except for English I and English II, for which the cut score was set to 3775. The Masters Grade Level cut scores vary across STAAR assessments, but for any given assessment, performance standards remain constant over time.

The STAAR scale scores represent linear transformations of the Rasch proficiency-level estimate (θ). Specifically, the transformation is made by first multiplying θ by a slope constant (*A*) and then adding an intercept constant (*B*). This operation is described by the following equation:

$$SS_\theta = A \times \theta + B,$$

where SS_θ is the scale score for a Rasch proficiency score estimate (θ) and *A* and *B* are referred to as the horizontal scaling constants. These same transformations are applied each year to the Rasch proficiency score estimates (θ) for that year's set of test items. Values for the horizontal scaling constants are provided in Tables 4.4 and 4.5 for the horizontally scaled STAAR grades 3–8 and EOC assessments, respectively.

Table 4.4. Horizontal Scaling Constants for STAAR Grades 3–8

Assessment			*A*	*B*
Grade	**Language**	**Content Area**		
5	English	Science	555.8300	3661.6663
8	English	Science	630.2521	3873.5084
8	English	Social Studies	571.3560	3726.2633
5	Spanish	Science	555.8300	3661.6663

Table 4.5. Horizontal Scaling Constants for STAAR EOC Assessments

Assessment	*A*	*B*
Algebra I	460.7351	3919.0028
English I	429.3074	3845.4064
English II	444.4006	3852.8590
Biology	435.9620	4042.0267
U.S. History	487.6991	4073.2524

Vertical Reporting Scales

As required by TEC §39.036, TEA developed vertical scales for assessing student performance in grades 3–8 for mathematics and RLA. Vertical scales were developed for the following grades and subjects:

- grades 3–8 mathematics
- grades 3–8 RLA
- Spanish grades 3–5 mathematics
- Spanish grades 3–5 RLA

The vertical scale established for the English versions of grades 3–5 mathematics was also applied to the Spanish versions of grades 3–5 mathematics because the Spanish versions were transadapted from the English. For the STAAR grades 3–8 mathematics scale, a scale score of 1360 represented the Approaches Grade Level performance standard for the grade 3 assessment. The scale's standard deviation was set to 150.

For the STAAR grades 3–8 RLA scale, a scale score of 1345 represented the Approaches Grade Level performance standard for the grade 3 assessment. For the STAAR Spanish grades 3–5 RLA scale, a scale score of 1318 represented the Approaches Grade Level performance standard for the grade 3 assessment. The RLA vertical scales' standard deviations were also set to 150 for the assessment in both languages.

It is important to note that although Approaches Grade Level scale score values are fixed for the lowest grade in the vertical scale, the Approaches Grade Level scale score for the other assessments in the vertical scale will vary across grades. However, these Approaches Grade Level scale score values, as well as the Meets Grade Level and Masters Grade Level scale score values, remain constant over time.

The linear transformation of the underlying Rasch proficiency score estimate (θ) for vertical scale scores is described by the following equation for a vertically scaled test at grade *g*:

$$SS_\theta = A \times (\theta + V_g) + B,$$

where SS_θ is the scale score for a Rasch proficiency score estimate (θ), *A* and *B* are the vertical scale score transformation constants, and V_g is the vertical scaling constant for the grade *g* test. The values of *A*, *B*, and V_g for the vertically scaled STAAR assessments are provided in

Table 4.6. Once established, these same transformations are applied each year to the proficiency level estimates for that year's set of test questions.

Table 4.6. Vertical Scale Score Transformation and Scaling Constants for STAAR Grades 3–8 Mathematics and RLA

Assessment			*A*	*B*	V_g
Grade	**Language**	**Content Area**			
3	English/Spanish	Mathematics	130.0052	1454.3188	0
4	English/Spanish	Mathematics	130.0052	1454.3188	0.5911
5	English/Spanish	Mathematics	130.0052	1454.3188	1.0884
6	English	Mathematics	130.0052	1454.3188	1.9965
7	English	Mathematics	130.0052	1454.3188	2.4185
8	English	Mathematics	130.0052	1454.3188	3.0511
3	English	RLA	143.7195	1398.5930	0
4	English	RLA	143.7195	1398.5930	0.6921
5	English	RLA	143.7195	1398.5930	0.6641
6	English	RLA	143.7195	1398.5930	1.4135
7	English	RLA	143.7195	1398.5930	1.2939
8	English	RLA	143.7195	1398.5930	1.9002
3	Spanish	RLA	153.0768	1318.1531	0
4	Spanish	RLA	153.0768	1318.1531	0.4323
5	Spanish	RLA	153.0768	1318.1531	0.6918

Equating

Used in conjunction with the scaling process, equating is the process that considers the differences in difficulty across test forms and administrations and allows scores to be placed onto a common scale. By using statistical methods, TEA equates the results of different test forms so that scale scores across test forms and testing administrations can be compared. TEA uses pre-equating for all STAAR assessments and post-equating for STAAR assessments that include constructed-response items.

To replenish the item bank as new tests are created each year, newly developed items must be field-tested and equated to the item bank scale. During each spring administration, field-test equating is conducted for STAAR, including STAAR Spanish, through an embedded-field-test design for all tests. In some years, stand-alone field tests are conducted for STAAR. Each stand-alone field test also includes some items from the item bank as anchor items, and the field-tested items are equated to the item bank scale through these items.

Refer to Chapter 3, "Standard Technical Processes," for detailed information about equating.

Reliability

Reliability indicates the precision of test scores, referring to the expectation that repeated administrations of the same test should generate consistent results. Reliability for STAAR test

scores is estimated using statistical measures including internal consistency, classical SEM, CSEM, and classification consistency and accuracy. Data for each of these statistical measures from the spring STAAR administration are provided in Appendix B. Refer to Chapter 3, "Standard Technical Processes," for detailed information about reliability.

Validity

Validity refers to the extent to which test scores accurately measure what the test is intended to measure. TEA follows national standards of best practice and annually collects validity evidence to support the interpretations and uses of STAAR, including STAAR Spanish, test scores. TTAC, a panel of national testing experts created specifically for the Texas Assessment Program, provides ongoing input to TEA about STAAR validity evidence. The following sections describe how validity evidence has been collected for STAAR. Refer to Chapter 3, "Standard Technical Processes," for additional information about validity.

Evidence Based on Test Content

Validity evidence based on test content refers to evidence of the relationship between tested content and the construct that the assessment is intended to measure. STAAR, including STAAR Spanish, has been developed to align with content as defined by the TEKS. Content validity evidence is collected at all stages of the test-development process. Nationally established test-development processes for the Texas Assessment Program are followed while developing STAAR. This supports the use of STAAR scores in making inferences about students' knowledge and understanding of the TEKS.

Relationship to the Statewide Curriculum

The TEKS are designed to ensure that Texas students receive a solid education that will enable them to be successful in life, whether they choose to pursue higher education or enter the workforce directly after graduation. The TEKS are specifically aligned to the CCRS. The CCRS specify the knowledge and skills necessary to succeed in entry-level community college and university courses. The CCRS have been incorporated into the secondary TEKS to form a vertically articulated set of curriculum standards. STAAR focuses on fewer skills and addresses those skills in a deeper manner through the identification of readiness and supporting standards in the TEKS and the inclusion of a larger number of items that assess readiness standards in the test blueprint. STAAR, therefore, focuses on the TEKS that are most critical to success in the next grade or course and ultimately on postsecondary readiness.

Educator Input

As part of the development of STAAR, teachers, curriculum specialists, test development specialists, college educators, and TEA staff worked together in advisory committees to identify appropriate assessment reporting categories for STAAR. The input of the advisory committees was reflected in the assessed curricula and test blueprints. In addition, prototype items were developed for the assessments early in the development process. The educator advisory committees reviewed these prototypes to identify how well the items would measure the student expectations to which the items were aligned. These early reviews provided valuable suggestions for item development guidelines and item types. Item development guidelines continued to be refined through the test development process as various STAAR item-review

educator committees shared their feedback about how the student expectations could be effectively assessed.

As part of the annual process of item development, committees of Texas educators meet to review STAAR items and confirm that each item appropriately measures the TEKS to which it is aligned. These committees also review items for content and bias. Two distinct types of educator committee meetings are regularly held to support the validity of test content: item review committees and content validation committees. Item review committees are composed of Texas educators, and these committees revise and edit items, as appropriate, prior to field testing. Item review committees are convened for all STAAR assessments. Content validation committees, by comparison, comprise university faculty who are experts in the relevant subject matter. Though these committees do not edit or revise items prior to field testing, they can recommend that certain items not be placed on STAAR operational assessments. Content validation is conducted for all STAAR EOC assessments before assessments are administered to students.

Test Developer Input

Item writers and reviewers follow test development guidelines that explain how content aligned to given TEKS should be measured. At each stage of development, writers and reviewers verify the alignment of the items with the assessed student expectations.

Evidence Based on Response Processes

Response processes refer to the cognitive behaviors that are required to respond to a test item. TEA collects evidence to show that the way students respond to items on STAAR, including STAAR Spanish, reflects accurate measurement of the construct.

Items

Student response processes on STAAR vary per item type. Across STAAR, 15 types of response interactions are available to measure student learning. For more information about the question types, refer to the STAAR Resources webpage.

TEA gathers theoretical and empirical evidence to confirm that the type of response required for each item does not add construct-irrelevant variance. TEA also gathers evidence from several sources to confirm that response processes do not result in an advantage or disadvantage for any student group. When new item types or changes to the format of existing item types are considered for STAAR, cognitive labs are used to study the way students engage with the various item presentations. After item types are determined to be appropriate for STAAR, evidence about student responses is gathered annually through educator and expert reviews and analyses of individual student responses to these items. During item reviews, educators evaluate whether the content for a given item type is being appropriately assessed and whether students will be able to accurately demonstrate their knowledge of the construct given the items' planned format. When items are field-tested, additional data are gathered about students' responses. Data such as item difficulty, item point-biserial correlations, and DIF are all evaluated regarding the item type. For additional information, refer to the Item Analysis section of Chapter 3, "Standard Technical Processes."

Scoring Process

The process used to score items can provide additional validity evidence based on response processes. This type of validity evidence is predicated on accurate scoring.

For all multiple-choice, multipart, and multiselect items on STAAR, statistical key checks are conducted during the equating process. A statistical key check is a procedure in which the statistical properties of all items on every test form are computed. Items whose statistics do not meet predetermined criteria are flagged for further review by content experts to verify that the items are correctly keyed and scored.

An adjudication process is used to ensure scoring reliability and validity for technology-enhanced items. During adjudication, data files that include all unique responses for each test question are analyzed to identify responses or questions that require more detailed analysis to ensure accurate, consistent scoring. Evaluators who specialize in STAAR content then review student responses to resolve scoring discrepancies or uncertainties.

For short and extended constructed-response questions, rubrics are used by human scorers to evaluate student responses. All rubrics for STAAR are validated by educator committees and content experts. In addition, TEA has implemented a rigorous scoring process for constructed responses that includes training and qualification requirements for scorers, ongoing monitoring during scoring, adjudication and resolution processes for student responses that do not meet the exact or adjacent scoring requirements, and rescoring of responses for which concerns have been raised by districts, campuses, or teachers regarding the assigned score. A more comprehensive description of the scoring process for constructed-response items is available in Chapter 2, “Building a High-Quality Assessment System.”

Score reliability for every STAAR assessment is generated and evaluated in terms of scorer agreement rates and the commonly used kappa with quadratic weights (Fleiss & Cohen, 1973). Constructed responses are scored with the adjacent agreement model. The exact agreement rate, adjacent agreement rate, and total agreement rate (exact and adjacent) between both scores are generated (refer to Table 4.7 and Table 4.8). When both scores are not in exact or adjacent agreement, the student response is adjudicated by a scoring leader.

Table 4.7. Summary of Scorer Agreement (Reliability) for Spring 2023 STAAR Extended Constructed Responses

Item Type	Number of Responses	Agreement Rate (%) after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Grade 3 RLA					
Ideas	355,155	77%	21%	98%	0.77
Conventions	355,155	76%	23%	99%	0.71
Grade 4 RLA					
Ideas	363,493	69%	27%	96%	0.74
Conventions	363,493	73%	23%	96%	0.69

Item Type	Number of Responses	Agreement Rate (%) after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Grade 5 RLA					
Ideas	371,792	63%	34%	97%	0.75
Conventions	371,792	67%	31%	98%	0.67
Grade 6 RLA					
Ideas	389,737	65%	32%	97%	0.80
Conventions	389,737	68%	29%	97%	0.72
Grade 7 RLA					
Ideas	398,777	65%	33%	98%	0.79
Conventions	398,777	67%	31%	98%	0.71
Grade 8 RLA					
Ideas	407,717	66%	32%	98%	0.83
Conventions	407,717	72%	26%	98%	0.77
English I					
Ideas	509,861	67%	31%	98%	0.83
Conventions	509,861	70%	28%	98%	0.75
English II					
Ideas	463,238	65%	32%	97%	0.81
Conventions	463,238	69%	28%	97%	0.72
Grade 3 Spanish RLA					
Ideas	30,059	82%	17%	99%	0.81
Conventions	30,059	82%	17%	99%	0.71
Grade 4 Spanish RLA					
Ideas	21,259	73%	25%	98%	0.79
Conventions	21,259	75%	24%	99%	0.75
Grade 5 Spanish RLA					
Ideas	15,961	68%	30%	98%	0.70
Conventions	15,961	73%	25%	98%	0.69

Table 4.8. Summary of Scorer Agreement (Reliability) for Spring 2023 STAAR Short Constructed Responses

Item Type	Number of Responses	Agreement Rate (%) after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Grade 3 RLA					
Writing	354,892	92%	7%	99%	0.86
Grade 4 RLA					
Writing	363,116	92%	7%	99%	0.85
Grade 5 RLA					
Writing	371,782	89%	7%	100%	0.77
Grade 6 RLA					
Writing	389,880	89%	11%	100%	0.78
Grade 7 RLA					
Writing	399,103	87%	13%	100%	0.75
Grade 8 RLA					
Reading	408,977	77%	22%	99%	0.76
Writing	408,815	90%	10%	100%	0.79
English I					
Reading	512,642	85%	16%	100%	0.88
Writing	512,403	94%	6%	100%	0.89
English II					
Writing	465,604	93%	7%	100%	0.87
Grade 3 Spanish RLA					
Writing	30,298	92%	7%	99%	0.86
Grade 4 Spanish RLA					
Reading	21,663	82%	27%	99%	0.83
Writing	22,004	93%	7%	100%	0.86
Grade 5 Spanish RLA					
Reading	15,975	82%	18%	100%	0.83
Writing	16,181	92%	8%	100%	0.83
Grade 5 Science					
Question 1	378,366	87%	13%	100%	0.91
Grade 8 Science					
Question 1	405,505	84%	15%	99%	0.97
Question 2	406,151	97%	2%	99%	0.84

Item Type	Number of Responses	Agreement Rate (%) after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Biology					
Question 1	455,603	92%	8%	100%	0.90
Question 2	459,859	89%	11%	99%	0.88
Grade 5 Spanish Science					
Question 1	9,763	91%	9%	100%	0.83
Social Studies Grade 8					
Question 1	411,961	81%	18%	99%	0.83
Question 2	411,625	83%	16%	99%	0.85
U.S. History					
Question 1	375,688	63%	37%	97%	0.62
Question 2	375,890	64%	36%	98%	0.62

Validity is evaluated through validity papers, which are student responses from the field test and current administrations that are representative of different levels of writing performance based on the scoring rubrics. Validity papers are identified by scoring leaders and are then systematically given to scorers throughout the scoring project. An important feature of validity papers is that they are not identifiable as such; in fact, they are indistinguishable from unscored student responses. Each person's daily scores on validity papers are compared with the approved scores. Validity papers are used throughout the scoring project as a primary quality-control measure, the purpose of which is to ensure that scorers are accurately and reliably scoring on a daily basis and across time. Validity agreement rate in Table 4.9 and Table 4.10 is expressed in terms of exact agreement between the score assigned by a given person and the true score approved by scoring leaders.

Table 4.9. Summary of Validity Results for Spring 2023 STAAR Extended Constructed Responses

STAAR Assessment and Trait	Exact Agreement Rate (%)
Grade 3 RLA	
Ideas	92%
Conventions	91%
Grade 4 RLA	
Ideas	85%
Conventions	88%
Grade 5 RLA	
Ideas	89%
Conventions	89%

STAAR Assessment and Trait	Exact Agreement Rate (%)
Grade 6 RLA	
Ideas	82%
Conventions	88%
Grade 7 RLA	
Ideas	81%
Conventions	78%
Grade 8 RLA	
Ideas	80%
Conventions	84%
English I	
Ideas	84%
Conventions	83%
English II	
Ideas	83%
Conventions	85%
Grade 3 Spanish RLA	
Ideas	94%
Conventions	92%
Grade 4 Spanish RLA	
Ideas	86%
Conventions	82%
Grade 5 Spanish RLA	
Ideas	82%
Conventions	83%

Table 4.10. Summary of Validity Results for Spring 2023 STAAR Short Constructed Responses

STAAR Assessment	Exact Agreement Rate (%)
English I Reading	95%
English I Writing	96%
English II Writing	97%
Biology Question 1	97%
Biology Question 2	98%
U.S. History Question 1	89%
U.S. History Question 2	94%

Evidence Based on Internal Structure

TEA collects evidence that shows the relationship of students' responses between items, within reporting categories of items, and within the full tests to verify that the elements of an assessment conform to the intended test construct and conducts internal consistency studies to gather evidence based on internal structure. The internal consistency of STAAR is evaluated using KR20 for assessments that have only dichotomously scored items. For the STAAR assessments that have a combination of dichotomous and polytomous items, internal consistency is evaluated using stratified coefficient alpha. These internal consistency evaluations are made for all students and for student groups such as female, male, Black or African American, Hispanic or Latino, and White students. Estimates of internal consistency are made for the full test, as well as for each reporting category within a content area, and can be found in Appendix B.

Evidence Based on Relationships to Other Variables

Another method TEA uses to provide validity evidence for STAAR, including STAAR Spanish, is analyzing the relationship between performance on STAAR and performance on other assessments, a process that supports criterion-related validity. Evidence can be collected to show that the empirical relationships are consistent with the expected relationships. Several past and current research studies have been designed to evaluate the relationship between performance on STAAR and performance on other related assessments or criteria, including the following:

- STAAR to TAKS comparison studies, which link performance on STAAR to performance on TAKS (e.g., STAAR grade 7 mathematics to TAKS grade 7 mathematics)
- STAAR linking studies, which link performance on STAAR across grade levels or courses in the same content areas (e.g., grade 4 RLA to grade 5 RLA, English I to English II)
- STAAR intercorrelation estimates, which evaluate the strength of the relationship (or lack thereof) among scores on STAAR across different content areas (e.g., grade 4 mathematics to grade 4 RLA, English I to Biology)
- grade correlation studies, which link performance on the STAAR EOC assessments to course grades
- validity studies, which link performance on STAAR to other measures (e.g., SAT, ACT, Lexiles, Quantiles, STAAR Interim Assessments)
- college students taking STAAR studies, which link performance on STAAR EOC assessments to college course grades

More information on comparisons between STAAR operational assessments and STAAR Interim Assessments is available in the STAAR Interim Assessments Summary Report. In addition, STAAR correlation estimates based on student performance on the spring administration are provided in Appendix B.

Evidence Based on the Consequences of Testing

Another method for providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. The collection of consequential validity evidence typically occurs after a program has been in place for some time and on a regular basis.

Given the important stakes associated with STAAR (including STAAR Spanish), valid test scores are critical in supporting their intended interpretations and uses. The intended interpretations of STAAR results are stated in the policy definitions of the four performance levels. Refer to the Performance Standards section in this chapter for the policy definitions of the STAAR performance levels. Each performance level describes a student's knowledge and skills in a content area and a student's level of preparation for the next grade or course.

Student-Level Performance

The following are the intended uses of STAAR test scores based on the policy definitions for student-level performance:

- Performance on STAAR is one indicator of a student's level of proficiency in a content area or specific course.
- Performance on STAAR is one indicator of a student's readiness for the next grade level or course in the same content area.
- Performance on STAAR is one indicator of a student's possible need for academic intervention.
- Performance on STAAR across years provides one indicator of a student's academic progress within a content area.
- Performance on STAAR may provide information about expected student performance on external assessments, such as the ACT or SAT, that measure similar knowledge and skills.

District- or Campus-Level Performance

The following are the intended uses of STAAR test scores based on the policy definitions for district- or campus-level performance:

- STAAR performance results can be aggregated to provide one indicator of overall student proficiency at a district or campus.
- STAAR performance results can be aggregated to provide one indicator of overall student readiness (for the next grade level or course in the same content area) at a district or campus.
- STAAR performance results can be aggregated across years to provide one indicator of overall student academic progress at a district or campus.

Sampling

Sampling is a procedure that is used to select and examine a small set that is representative of the population from which it was drawn. STAAR uses two types of sampling: stratified random sampling and simple random sampling. Stratified random sampling used in stand-alone field testing ensures that subgroups of a given population are adequately represented within the whole sample. Simple random sampling is used to sample responses for field-test items that are scored by human scorers.

Test Results

Appendix B provides consistency and accuracy data, scale score correlations, CSEMs, mean *p*-values, scale score descriptive statistics, and frequency distributions for the spring STAAR administration. Pass rates for STAAR are available on the Statewide Summary Reports webpage.

TECHNICAL DIGEST 2022–2023

Chapter 5

STAAR Alternate 2

Overview

Participation Requirements

Test Development

Accommodations

Training

Test Administration

Performance Standards

Scores and Reports

Measures of Student Progress

Scaling

Equating

Reliability

Validity

Sampling

Test Results

Overview

STAAR Alternate 2 is a standardized alternate academic achievement assessment based on alternate academic achievement standards and designed to measure the extent to which a student has learned and is able to apply the defined knowledge and skills in the TEKS. STAAR Alternate 2 is administered individually to students with the most significant cognitive disabilities who meet the participation requirements. STAAR Alternate 2 fulfills ESSA and IDEA. ESSA requires that all students be assessed in specific grades and subjects throughout their academic careers, whereas IDEA requires that students with disabilities have access to the same standards as their nondisabled peers and that they be included in statewide assessments.

STAAR Alternate 2 is not a traditional paper-pencil or multiple-choice test. Instead, it involves test administrators observing students as they respond to standardized, state-developed assessment questions that align to the grade-level TEKS through prerequisite skills. Teachers evaluate student performance based on standard scoring instructions embedded into each item on STAAR Alternate 2.

STAAR Alternate 2 was implemented in the 2014–2015 school year and includes the following assessments:

- grades 3–8 mathematics,
- grades 3–8 RLA,
- grades 5 and 8 science,
- grade 8 social studies, and
- EOC assessments for:
    - Algebra I,
    - English I,
    - English II,
    - Biology, and
    - U.S. History.

Due to the redesign of STAAR, STAAR Alternate 2 reading assessments were also redesigned to combine reading and writing into an RLA assessment for each grade. The redesigned STAAR Alternate 2 RLA assessments were implemented in spring 2023.

Participation Requirements

Students who receive special education services and have the most significant cognitive disabilities are eligible to participate in STAAR Alternate 2. These students exhibit significant intellectual and adaptive behavior deficits that limit their ability to plan, comprehend, and reason as well as adaptive behavior deficits that limit their ability to apply social and practical skills

(e.g., personal care, social problem-solving skills, dressing, eating, using money) across all life domains. Students with the most significant cognitive disabilities require extensive, direct, individualized instruction and have a need for substantial supports that are neither temporary nor content specific.

STAAR Alternate 2 has specific participation requirements that an ARD committee must carefully review and consider annually. The STAAR Alternate 2 Participation Requirements, available in English and Spanish on the STAAR Alternate 2 Resources webpage, detail the ARD committee's responsibility for ensuring that a student is eligible for STAAR Alternate 2. Prior to reviewing the eligibility criteria for STAAR Alternate 2, the ARD committee must understand all assessment options, including the characteristics of each assessment and the potential implications of each assessment choice. If STAAR Alternate 2 is being considered, the ARD committee must review the participation requirements against the supporting documentation within the student's individualized education program (IEP), such as in the present levels of academic achievement and functional performance, to determine eligibility.

Students in grades 3–8 who meet the participation requirements will take all applicable STAAR Alternate 2 subject assessments at their enrolled grade level.

Students in grades 9–12 who meet the participation requirements will take STAAR Alternate 2 EOC assessments—Algebra I, English I, English II, Biology, and U.S. History—as they are completing the corresponding course. The ARD committee makes educational decisions for a student with a disability, including decisions related to graduation requirements as described in TAC §89.1070.

In rare circumstances a student's ARD committee may determine prior to the administration of the assessment that the student will not participate in STAAR Alternate 2 because the student meets the eligibility criteria for a medical exception or no authentic academic response (NAAR). For both exceptions, the ARD committee reviews educational records and eligibility requirements. For more information, refer to the eligibility criteria on the STAAR Alternate 2 Resources webpage.

Test Development

Maintaining a high-quality student assessment program involves a complex and detailed test-development process, and TEA relies on input from educators to ensure that all measures of learning for Texas public school students are equitable and accurate. Test items for STAAR Alternate 2 are developed annually, reviewed by educator committees, field-tested, reviewed with their data, and, if approved, added to the STAAR Alternate 2 item bank. Newly developed items are embedded in STAAR Alternate 2 operational assessments each spring. For more information regarding each step of the STAAR Alternate 2 test-development process, refer to Chapter 2, which outlines the processes used to develop each STAAR Alternate 2 assessment's framework and explains ongoing test development.

For the initial development of STAAR Alternate 2, TEA sought input from educator committees and a statewide steering committee that included state assessment experts, parents, advocacy group representatives, related service providers, administrators, and ESC professionals. Consistent with the idea of universal design, particular attention was given to:

- students' response modes to allow students to show what they know and can do,
- differentiated supports and materials to allow students to access the content of the assessment, and
- multiple means of engagement to allow students more time to complete each task.

To ensure STAAR Alternate 2 is linked to grade-level TEKS, TEA worked with experts in test development, special education, and specific subject areas to develop vertical alignments for each content area and curriculum framework tools. The vertical alignments link content standards across grades, and the curriculum frameworks list the grade-level TEKS and the associated prerequisite skills for each grade and subject area. Essence statements, also known as strand statements in RLA, act as a bridge between grade-level content standards and STAAR Alternate 2 prerequisite skills. Specific essence statements are selected each year and provided to educators in the fall, giving them time to plan instruction and develop standards-based IEPs for that school year.

Accommodations

The goal of accommodations for STAAR Alternate 2 is to ensure that each student can interact appropriately with the content, presentation, and response modes of the state assessments. STAAR Alternate 2 is a standardized assessment intended to be appropriate for eligible students in its original, intact form. However, it is critical that students with disabilities be provided access to the assessment through careful use of accommodations wherever appropriate. Therefore, allowable accommodations may be provided to enable students with disabilities to participate meaningfully in the assessment.

Test administrators may use accommodations only if they are routinely provided in classroom instruction and listed in the student's IEP. Some accommodations provided during classroom instruction may not be allowed during testing, as certain accommodations used in the classroom would invalidate the content being assessed or compromise the security and integrity of the test. A list of allowable accommodations can be found in the *STAAR Alternate 2 Test Administrator Manual*, which is available on the STAAR Alternate 2 Resources webpage.

Training

TEA develops instructional materials, including manuals, presentations, online modules, and videos, to support the training of all testing personnel on test security and administration procedures. Preparation for test administration begins every year with a training-of-trainers session for testing coordinators from each of the 20 Texas regional ESCs as well as district testing coordinators from the state's 25 largest districts. Using materials and information provided in the TEA training session, ESC regional testing coordinators train the district coordinators in their respective regions. District coordinators then train their campus testing coordinators, who are responsible for training test administrators.

Test security and administration procedures provided in the *Coordinator Resources* and the *STAAR Alternate 2 Test Administrator Manual* must be followed so that all students have an equal opportunity to demonstrate their academic knowledge and skills. The *Coordinator*

Resources guide district and campus coordinators through their responsibilities as they oversee the administration of the Texas Assessment Program. This online resource contains preparation and administration procedures for each state-required assessment and is available prior to the annual ESC training.

In addition, TEA produces the *STAAR Alternate 2 Educator Guide*, available on the STAAR Alternate 2 Resources webpage, to familiarize educators with the assessment. The guide includes information on test design, alignment with state curriculum, training, and test results.

Test Administration

All STAAR Alternative 2 assessments—grades 3–8 mathematics, grades 3–8 RLA, grades 5 and 8 science, grade 8 social studies, Algebra I, English I, English II, Biology, and U.S. History—are administered on paper. The STAAR Alternate 2 testing window occurs over a five-week period during the spring, and retest opportunities are not offered. The number of students tested for each STAAR Alternate 2 assessment is shown in Table 5.1.

Table 5.1. STAAR Alternate 2 Assessments Administered in 2022–2023

STAAR Assessment	Assessments Administered
Grade 3 mathematics	7,615
Grade 3 RLA	7,617
Grade 4 mathematics	7,544
Grade 4 RLA	7,546
Grade 5 mathematics	7,084
Grade 5 RLA	7,083
Grade 5 science	7,081
Grade 6 mathematics	6,722
Grade 6 RLA	6,723
Grade 7 mathematics	6,557
Grade 7 RLA	6,561
Grade 8 mathematics	6,445
Grade 8 RLA	6,448
Grade 8 science	6,445
Grade 8 social studies	6,446
Algebra I	6,310
English I	6,325
English II	6,077
Biology	6,325
U.S. History	5,502

Each STAAR Alternate 2 test question measures a targeted prerequisite skill. A cluster of four test questions tests a common skill or concept at varying levels of difficulty. Five clusters make up a test form of 20 base test questions. Test forms also include one field-test cluster.

The assessment is designed with scripted presentation instructions that mirror instructional techniques for a student with the most significant cognitive disabilities. Student responses during a STAAR Alternate 2 test administration may be verbal, physical, or visual as appropriate for the student at the time of testing. Each question has a unique set of scoring instructions that describe what the student must do for his or her response to be marked correct. The test administrator must refer to the scoring instructions for each question to determine how to score the student's response.

STAAR Alternate 2 is scored polytomously using a standard scoring rubric with item score ranges from 0 to 2. Each item is scored according to the level of independence with which a student responds. The scoring rubric is as follows:

- If a student responds correctly to the first presentation of an item, he or she receives a score point of 2. If the student does not respond or responds incorrectly, the item is presented again with allowable teacher assists.
- If the student responds correctly to the second presentation of the item, he or she receives a score point of 1.
- If the student does not respond or responds incorrectly to the second presentation, he or she receives a score point of 0.

Performance Standards

Performance standards directly relate levels of test performance to what students are expected to learn, as defined in the statewide curriculum. Standard setting is the process of establishing cut scores that define the performance levels on an assessment.

Performance Levels and Policy Definitions

For STAAR Alternate 2, the performance levels and policy definitions are as follows:

Level I: Developing Academic Performance

Performance in this category indicates that students require additional instructional supports for accessing the curriculum through prerequisite skills. Students are able to acknowledge some concepts, but they demonstrate a minimal or inconsistent understanding of the knowledge and skills that are linked to content measured in this grade or course. Even with continued support, students in this category need significant intervention to show progress in the next grade or course.

Level II: Satisfactory Academic Performance

Performance in this category indicates that students are sufficiently prepared for the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. Students demonstrate sufficient understanding of the knowledge and skills that are linked to

content measured at this grade or course. Students exhibit the ability to determine relationships, integrate multiple pieces of information, extend details, identify concepts, and match concepts that are similar. With continued support, students in this category have a reasonable likelihood of showing progress in the next grade or course.

Level III: Accomplished Academic Performance

Performance in this category indicates that students are well prepared for the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. Students demonstrate a strong understanding of the knowledge and skills that are linked to content measured at this grade or course. Students exhibit the ability to use higher-level thinking and more complex skills, which includes making inferences and comparisons and solving multi-step problems. With support, students in this category have a high likelihood of showing progress in the next grade or course.

Standard Setting

Standards for all assessments were originally set for STAAR Alternate 2 in 2015. Standard setting for STAAR Alternate 2 involved combining considerations regarding policy, the TEKS content standards, educator knowledge about what students should know and be able to do, and information about how student performance on state assessments aligns with student performance on other assessments. Standard-setting committees comprised diverse groups of stakeholders who carefully considered the interaction of these elements for each STAAR Alternate 2 assessment. The task of each standard-setting committee was to recommend cut scores that would define the performance levels for each STAAR Alternate 2 assessment. In 2023, performance standards were reset for STAAR Alternate 2 RLA based on the redesign of STAAR.

The current performance standards for STAAR Alternate 2 are provided in Table 5.2.

Table 5.2. Performance Levels for STAAR Alternate 2

Subject Area	Grade/Course	Level II: Satisfactory	Level III: Accomplished
Reading Language Arts	Grade 3	300	388
	Grade 4	300	380
	Grade 5	300	374
	Grade 6	300	370
	Grade 7	300	378
	Grade 8	300	371
	English I	300	365
	English II	300	370

Subject Area	Grade/Course	Level II: Satisfactory	Level III: Accomplished
Mathematics	Grade 3	300	375
	Grade 4	300	387
	Grade 5	300	379
	Grade 6	300	373
	Grade 7	300	375
	Grade 8	300	365
	Algebra I	300	361
Science	Grade 5	300	387
	Grade 8	300	382
	Biology	300	383
Social Studies	Grade 8	300	372
	U.S. History	300	368

Refer to the standard setting technical reports, which are available on the Assessment Reports and Studies webpage, for more information.

Scores and Reports

TEA publishes resources on both the TEA and Texas Assessment websites to assist school personnel in understanding and interpreting student performance data and to help parents understand their child's STAAR Alternate 2 results. School personnel can access STAAR Alternate 2 results through CRS, parents can access their child's STAAR Alternate 2 results in the Family Portal, and the public can access STAAR Alternate 2 statewide, region, district, and campus data using the Research Portal.

TEC §39.030 and TAC §101.3014 specify the requirements for maintaining the confidentiality of individual student results and for reporting district-level and campus-level results. The results of individual student performance on state assessments are confidential and may be released only in accordance with FERPA. Districts must provide each student's state assessment results to the student, to his or her parent or guardian, and to his or her teacher for the applicable subject area. In addition, all state assessment results must be included in each student's academic achievement record.

Description of Scores

Scores for STAAR Alternate 2 include raw scores, scale scores, and the resulting performance level associated with the student's score. The number of points that a student earns on a STAAR Alternate 2 assessment is the student's raw score. A scale score is a conversion of the raw score onto a scale that is common to all test forms for that assessment. The scale score is used to determine whether a student achieved the Level I: Developing Academic Performance, Level II: Satisfactory Academic Performance, or Level III: Accomplished Academic Performance standard. Refer to Chapter 3, "Standard Technical Processes," for more information about raw scores and scale scores.

Assessment Reports

TEA provides reports of student performance on STAAR Alternate 2 to all Texas public school districts and open-enrollment charter schools. For STAAR Alternate 2, TEA provides student report cards, student labels, campus rosters, summary reports, and reporting data files. In addition, TEA periodically releases STAAR Alternate 2 assessments, which can be found on the STAAR Alternate 2 Released Test Questions webpage.

For more information about scoring and reporting for STAAR Alternate 2, refer to the Interpreting Results page of the *Coordinator Resources*.

Use of Test Results

Test results can be used to evaluate the performance of a group over time. Average scale scores and the percentage of students meeting the Level I, Level II, and Level III performance standards can be analyzed by grade and content area across administrations to provide insight into whether student performance is improving across years. For example, the average scale score for students who took the STAAR Alternate 2 grade 4 mathematics test can be compared over time.

Test results can be used when evaluating instruction or programs that require average-score or year-to-year comparisons. The tests are designed to measure content areas within the required state curriculum, so the consideration of test results by content area and reporting category might be helpful when evaluating curriculum and instructional programs. All test scores can be compared with statewide and regional performance within the same content area for any administration.

Test scores can also be used to identify where an individual student needs additional instruction or support in each subject. This identification can help educators plan the most effective instructional intervention. Finally, individual student test scores are also used in conjunction with other performance indicators to assist in making placement decisions. While scores can contribute to decisions regarding placement, educational planning for a student should take into account as much student information as possible.

Generalizations from test results can be made from the specific content area being measured on the test. However, because each test measures a finite set of skills with a limited set of items, any generalizations about student achievement derived solely from a particular test should be made with great caution and with full reference to the fact that the conclusions are based only on that test. Instruction and program evaluations should take into account as much information as possible, rather than relying on test scores alone, to provide a more complete picture of student performance.

Measures of Student Progress

Student progress measures provide information beyond performance level by considering performance over time. Whereas performance-level information describes students' current achievement, progress measures describe students' achievement in adjacent years.

STAAR Alternate 2 Progress Measure

The STAAR Alternate 2 Progress Measure was reported for the first time in the 2015–2016 school year. For STAAR Alternate 2, progress is measured based on a student's stage change from the prior year to the current year. Stage change is determined by: 1) classifying the student's scores from the previous school year and the current school year in terms of the stage of performance achieved and then 2) comparing the stages from year to year. Student progress is then categorized as Did Not Meet, Met, or Exceeded. These progress targets define the expectation of annual progress for each grade and content area. The progress targets are grounded in the STAAR Alternate 2 performance standards.

Steps for calculating a student's stage change and progress indicator for the STAAR Alternate 2 progress measure can be found in *STAAR Alternate 2 Progress Measure Questions and Answers* on the Progress Measures webpage.

Progress measures were not calculated or reported for the 2022–2023 school year.

STAAR Alternate 2 On-Track Measure

The STAAR Alternate 2 on-track measure examines a student's progress and projects where that student will be in a future target year if that student continues to make progress at the same rate over future years. The student is then classified as On Track or Not On Track to achieve Level II: Satisfactory in the target year.

If a student has scores for STAAR Alternate 2 in two consecutive grades and subject or courses in two consecutive years, the on-track measure can be calculated for the student. If any of the required information for STAAR Alternate 2 on-track measure calculation is lacking, the on-track measure is not available. This includes students who have received exceptions through the medical exception or NAAR policies in the previous or current grade.

Steps for calculating a student's STAAR Alternate 2 on-track measure can be found in *STAAR Alternate 2 On-Track Measure Questions and Answers* on the Progress Measures webpage.

Scaling

Scaling is a statistical procedure that places raw scores on a common scoring metric to make test scores comparable across test administrations. Scaling associates numbers with characteristics of interest to provide information about measurable quantities for those characteristics. STAAR Alternate 2 uses the RPCM to place test items on the same Rasch scale across administrations for a given assessment. Once performance standards have been set for an assessment, the Rasch scale is then transformed to a more user-friendly metric to facilitate interpretation of the test scores. Details of the RPCM scaling method are provided in Chapter 3, "Standard Technical Processes."

Reporting Scales

STAAR Alternate 2 scale scores are reported on a horizontal scale. Horizontal scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations for a specific grade and subject or course.

For all STAAR Alternate 2 assessments, a scale score of 300 represents the Level II: Satisfactory performance standard. The desired standard deviation for each grade and subject and course is 50. The Level III scale score values vary across STAAR Alternate 2 assessments, but for any given assessment, performance standards remain constant over time.

STAAR Alternate 2 scale scores represent linear transformations of Rasch proficiency-level estimates (θ). Specifically, the transformation is made by first multiplying θ by a slope constant (*A*) and then adding an intercept constant (*B*). This operation is described by the following equation:

$$SS_\theta = A \times \theta + B,$$

where SS_θ is the scale score for a Rasch proficiency-level estimate (θ) and *A* and *B* are the horizontal scaling constants. These same transformations will be applied each year to the Rasch proficiency-level estimates (θ) for that year's set of test items. Values for the horizontal scaling constants for STAAR Alternate 2 are provided in Table 5.3.

Table 5.3. Horizontal Scaling Constants for STAAR Alternate 2

Subject Area	Grade/Course	*A*	*B*
Mathematics	Grade 3	43.9599	297.2305
	Grade 4	42.3406	297.9677
	Grade 5	42.9221	293.4758
	Grade 6	47.3082	293.8972
	Grade 7	45.0653	292.6994
	Grade 8	45.9897	283.5357
	Algebra I	46.1042	287.8285
Reading Language Arts	Grade 3	51.7409	300.6002
	Grade 4	52.5281	289.7045
	Grade 5	52.1646	285.3261
	Grade 6	52.7711	284.9813
	Grade 7	53.4243	290.3676
	Grade 8	50.2019	283.8651
	English I	49.3225	294.3526
	English II	49.5385	294.8480
Science	Grade 5	43.8943	291.6601
	Grade 8	38.5892	298.4950
	Biology	38.2614	293.1129
Social Studies	Grade 8	41.4662	282.7501
	U.S. History	41.3565	283.7055

Equating

Used in conjunction with the scaling process, equating is the process that considers the differences in difficulty across test forms and administrations and allows scores to be placed onto a common scale. By using statistical methods, TEA equates the results of different test forms so that scale scores across test forms and test administrations can be compared. TEA uses pre-equating and post-equating for all STAAR Alternate 2 assessments.

To replenish the item bank as new tests are created each year, newly developed items must be field-tested and equated to the item bank scale. During each spring administration, field-test equating is conducted for STAAR Alternate 2 through an embedded-field-test design for all tests.

Refer to Chapter 3, "Standard Technical Processes," for detailed information about equating.

Reliability

Reliability indicates the precision of test scores, referring to the expectation that repeated administrations of the same test should generate consistent results. Reliability for STAAR Alternate 2 test scores is estimated using statistical measures including internal consistency, classical SEM, CSEM, and classification consistency and accuracy. Data for these statistical measures from the spring STAAR Alternate 2 administration are provided in Appendix C. Refer to Chapter 3, “Standard Technical Processes,” for detailed information about reliability.

Validity

Validity refers to the extent to which test scores accurately measure what the test is intended to measure. TEA follows national standards of best practice and annually collects validity evidence to support the interpretations and uses of STAAR Alternate 2 test scores. TTAC, a panel of national testing experts created specifically for the Texas Assessment Program, provides ongoing input to TEA about STAAR Alternate 2 validity evidence. The following sections describe how validity evidence has been collected for STAAR Alternate 2. Refer to Chapter 3, “Standard Technical Processes,” for additional information about validity.

Evidence Based on Test Content

Validity evidence based on test content refers to evidence of the relationship between tested content and the construct that the assessment is intended to measure. STAAR Alternate 2 has been developed to align with content as defined by the TEKS through prerequisite skills. Content validity evidence is collected at all stages of the test-development process. Nationally established test-development processes for the Texas Assessment Program are followed while developing STAAR Alternate 2. This supports the use of STAAR Alternate 2 scores in making inferences about students’ knowledge and understanding of the TEKS.

Relationship to the Statewide Curriculum

The TEKS are designed to ensure that Texas students receive a solid education that will enable them to be successful in life, whether they choose to pursue higher education or enter the workforce directly after graduation. STAAR Alternate 2 assesses the TEKS through prerequisite skills.

In 2015–2016, an independent third-party analysis of the alignment between items on the 2016 STAAR Alternate 2 tests and the TEKS was conducted to inform TEA about the degree of alignment between the test items and curriculum standards. The study concluded that the 2016 STAAR Alternate 2 items demonstrated strong linkages across all grades and subjects and

courses. All items were found to have an academic foundation and to have content connections to the grade-level student expectations.

Educator Input

As part of the initial development of STAAR Alternate 2, teachers, curriculum specialists, special education experts, test development specialists, and TEA staff worked together in advisory committees to identify the best way to assess students with the most significant cognitive disabilities. The input of the advisory committees was reflected in the vertical alignment documents, prerequisite skills, essence statements, and test blueprints. In addition, prototype items were developed for the assessments early in the development process. The educator advisory committees reviewed these prototypes to identify how well the items would measure the TEKS through the prerequisite skills to which the items were aligned. These early reviews provided valuable suggestions for item development guidelines and item types. Item development guidelines continued to be refined through the test development process as various STAAR Alternate 2 item review committees of educators shared their feedback about how the TEKS could be effectively assessed.

As part of the annual process of item development, committees of Texas educators meet to review STAAR Alternate 2 items and confirm that each item appropriately measures the TEKS through prerequisite skills. These committees also review items for content and bias. Item review committees are composed of Texas educators, including special education teachers, and these committees revise and edit items, as appropriate, prior to field testing. Item review committees are convened for all STAAR Alternate 2 assessments.

Test Developer Input

Item writers and reviewers, including content experts and special education experts, follow test development guidelines and item specifications that explain how the content of the assessed TEKS should be measured. At each stage of development, writers and reviewers verify the alignment of the test items with the assessed essence statements.

Evidence Based on Response Processes

Response processes refers to the cognitive behaviors that are required to respond to a test item. TEA collects evidence to show that the way students respond to items on STAAR Alternate 2 reflects accurate measurement of the construct.

Items

TEA gathers theoretical and empirical evidence supporting the expectation that the way students respond to test items does not add construct-irrelevant variance. During yearly item reviews, educators evaluate whether the content for a given item is being appropriately assessed and whether students will be able to accurately demonstrate their knowledge of the construct given the items’ planned format. When items are field-tested, additional data are gathered about students' responses. Data such as item difficulty, item-total correlations, and item fit are all evaluated. For additional information, see the Item Analysis section of Chapter 3, “Standard Technical Processes.”

Scoring Process

The process used to score items can provide additional validity evidence based on response processes. This type of validity evidence is predicated on accurate scoring. The test administrator booklet provides test administrators with exact scoring rules and scripted instructions on how to present every item to a student. Test administrators are provided with resources to prepare for a STAAR Alternate 2 test administration, including a scheduled period directly before the testing window in which they can preview the test booklet to prepare for a valid test administration.

Evidence Based on Internal Structure

TEA collects evidence that shows the relationship between items and reporting categories to verify that the elements of an assessment conform to the intended test construct and conducts internal consistency studies to gather evidence based on internal structure. The internal consistency of STAAR Alternate 2 is evaluated using coefficient alpha. These internal consistency evaluations are made for all students and for student groups such as female, male, Black or African American, Hispanic or Latino, and White.

Evidence Based on Relationships to Other Variables

Another method TEA uses to provide validity evidence for STAAR Alternate 2 is analyzing the relationship between performance on STAAR Alternate 2 and performance on other assessments, a process that supports criterion-related validity. Evidence can be collected to show that the empirical relationships are consistent with the expected relationships. STAAR Alternate 2 correlation estimates are calculated to evaluate the strength of the relationship (or lack thereof) among scores on STAAR Alternate 2 assessments across different content areas (e.g., grade 4 mathematics to grade 4 RLA, English I to Biology).

Evidence Based on Consequences of Testing

Another method for providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. The collection of consequential validity evidence typically occurs on a regular basis after a program has been in place for some time. Some of the intended consequences of STAAR Alternate 2 are as follows:

- Students with the most significant cognitive disabilities can receive challenging instruction that is linked to state content standards.
- Students with the most significant cognitive disabilities can be included in state assessment programs.
- STAAR Alternate 2 can assess the achievement of students with the most significant cognitive disabilities.

Given the important stakes associated with STAAR Alternate 2, valid test scores are critical in supporting their intended interpretations and uses. The intended interpretations of STAAR Alternate 2 results are stated in the policy definitions of the three performance levels. Refer to the Performance Standards section in this chapter for the policy definitions of the STAAR

Alternate 2 performance levels. Each performance level describes a student's knowledge and skills in a content area and a student's level of preparation for the next grade or course.

Student-Level Performance

The following are the intended uses of STAAR Alternate 2 test scores based on the policy definitions for student-level performance:

- Performance on STAAR Alternate 2 is one indicator of a student's level of proficiency in a content area or specific course.
- Performance on STAAR Alternate 2 is one indicator of a student's readiness for the next grade or course in the same content area.
- Performance on STAAR Alternate 2 is one indicator of a student's possible need for academic intervention.
- Performance on STAAR Alternate 2 across years provides one indicator of a student's academic progress within a content area.

District- or Campus-Level Performance

The following are the intended uses of STAAR Alternate 2 test scores based on the policy definitions for district- or campus-level performance:

- STAAR Alternate 2 performance results can be aggregated to provide one indicator of overall student proficiency at a district or campus.
- STAAR Alternate 2 performance results can be aggregated to provide one indicator of overall student readiness (for the next grade or course in the same content area) at a district or campus.
- STAAR Alternate 2 performance results can be aggregated across years to provide one indicator of overall student academic progress at a district or campus.

Sampling

Sampling is a procedure that is used to select and examine a small set that is representative of the population from which it was drawn. For the STAAR Alternate 2 administration, campus assignment of forms uses an annual sampling process wherein a single form is assigned to each campus. A sample of students who represent the state demographic makeup respond to each form. This approach ensures that each campus administers the same form to all students and that teachers need only administer a single form.

Test Results

Appendix C provides consistency and accuracy data, scale score correlations, CSEMs, mean *p*-values, scale score descriptive statistics, and frequency distributions for the spring STAAR Alternate 2 administration. Pass rates for STAAR Alternate 2 are available on the Statewide Summary Reports webpage.

TECHNICAL DIGEST 2022–2023

Chapter 6

Texas English Language Proficiency Assessment System

Overview

Participation Requirements

Test Development

Training

Test Administration

Proficiency Standards

Scores and Reports

Audits

Scaling

Equating

Reliability

Validity

Sampling

Test Results

Overview

TELPAS is an English language proficiency assessment that measures the progress that EB students make in acquiring the English language. It fulfills the requirements of ESSA, which requires that all EB students be assessed annually until they are determined to be proficient in the English language.

TELPAS assesses EB students in kindergarten through grade 12 in the language domains of listening, speaking, reading, and writing. For grades 2–12, TELPAS consists of online assessments for listening and speaking and for reading and writing. For kindergarten and grade 1, holistically rated assessments based on ongoing classroom observations are used for all four language domains.

Participation Requirements

All EB students in kindergarten through grade 12 are required to participate in TELPAS unless the student meets the participation requirements for TELPAS Alternate. EB students are assessed annually in English language proficiency until they are determined to be proficient by meeting the EB reclassification criteria. This includes students classified as emergent bilingual (EB)/English learner (EL) in the Public Education Information Management System (PEIMS) whose parents have declined bilingual or ESL program services.

In rare cases, it might be necessary for the ARD committee, in conjunction with the LPAC, to determine that an EB student receiving special education services should not be assessed in listening, speaking, reading, or writing for reasons associated with the student's disability. Participation must be considered on a domain-by-domain basis. The reason for not assessing the student must be related to the student's disability and be well supported and documented in the student's IEP by the ARD committee and in the student's permanent record file by the LPAC.

Test Development

Maintaining a high-quality student assessment program involves a complex and detailed test-development process, and TEA relies on input from educators to ensure that all measures of learning for Texas public school students are equitable and accurate. For more information regarding each step of the TELPAS test-development process, refer to Chapter 2, "Building a High-Quality Assessment System," which outlines the processes used to develop each TELPAS assessment's framework and explains ongoing test development.

Test items for TELPAS online assessments are developed annually, reviewed by educator committees, embedded in operational assessments each spring for field testing, reviewed with their data, and, if approved, added to the TELPAS item bank. TELPAS grades 2–12 online assessments were developed as combined listening and speaking assessments for multiple grade bands and combined reading and writing assessments for specific grades and grade bands, as shown in Table 6.1.

Table 6.1. TELPAS Grades 2–12 Online Assessments

Listening and Speaking	Reading and Writing
Grades 2–3	Grade 2
Grades 4–5	Grade 3
Grades 6–8	Grades 4–5
Grades 9–12	Grades 6–7
	Grades 8–9
	Grades 10–12

TEA developed the TELPAS holistically rated components in collaboration with test development experts, bilingual and ESL consultants, and focus group members including teachers, bilingual and ESL directors, assessment directors, campus administrators, and university professors. Like the TELPAS grades 2–12 assessments, these assessments align with the ELPS, assessing the English communication skills that EB students need to engage meaningfully and effectively in learning the academic knowledge and skills required by the TEKS. The holistically rated assessments draw on second language acquisition research, research-based standards, the experience of Texas educators, and observational assessment practices.

More information about the development of TELPAS is available in the *TELPAS Educator Guide* available on the TELPAS Resources webpage. Provided to familiarize educators with TELPAS, the guide shows the integral relationship between TELPAS and the ELPS. It explains the TELPAS language domains of listening, speaking, reading, and writing and provides examples of classroom instruction and annotated test item descriptions.

Training

TEA develops instructional materials, including manuals, guides, presentations, online modules, and videos, to support the training of all testing personnel on test security and administration procedures. Preparation for test administration begins every year with a TEA-provided training-of-trainers session for testing coordinators from each of the 20 Texas regional ESCs as well as district testing coordinators from the state's 25 largest districts. Using materials and information provided in the TEA training session, ESC regional testing coordinators train the district coordinators in their respective regions. District coordinators then train their campus testing coordinators, who are responsible for training test administrators.

Test security and administration procedures provided in the *Coordinator Resources*, the *TELPAS Test Administrator Manual*, and the *TELPAS Rater Manual* must be followed so that all students have an equal opportunity to demonstrate their knowledge and skills. The *Coordinator Resources* guide district and campus coordinators through their responsibilities as they oversee the administration of the Texas Assessment Program. This online resource contains preparation and administration procedures for each state-required assessment and is available prior to the annual ESC training.

In addition, TEA produces the *TELPAS Educator Guide* to familiarize educators with the assessment. The guide includes information on test design, alignment with the ELPS, training, and test results.

TELPAS raters must have trained and calibrated successfully before rating students. The training that TELPAS raters receive supports the administration of TELPAS and provides teachers with ongoing professional development to support effective implementation of the ELPS.

The Online Basic Training course teaches new raters the essentials of second language acquisition theory and how to use the ELPS PLDs to accurately identify the English language proficiency levels of EB students based on how well the students understand and use English during daily academic instruction and classroom interaction. The trainings are specific to grade clusters, and raters should complete the holistic rating training in the grade cluster that corresponds to the grade levels of the students they will rate. Online courses for kindergarten through grade 1 contain numerous practice rating activities composed of student writing samples and video segments in which EB students demonstrate their listening, speaking, reading, and writing skills in authentic Texas classroom settings. The courses give raters practice applying the scoring rubrics (i.e., PLDs) and provide detailed feedback about their rating accuracy.

New raters are required to successfully complete the online holistic rating trainings and separate practice activities for the grade cluster they are assigned before they access rater calibration activities and must then complete the online calibration activities to demonstrate their ability to apply the PLD rubrics consistently and accurately before they rate students for the operational assessment. There are two sets of calibration activities, and all applicable language domains are represented. Raters finish the calibration activities when they demonstrate sufficient accuracy. If sufficient accuracy is not obtained on the first set, the rater attempts a second and final online calibration set. Individuals not successful on the final set are either not used as raters or are provided rater support in accordance with test administration procedures. More information about TELPAS rater training can be found on the TELPAS Resources webpage.

Test Administration

TELPAS is administered once a year, in the spring, during a six-week testing window. The number of TELPAS assessments that were administered to eligible students in 2022–2023 is indicated in Table 6.2.

Table 6.2. TELPAS Assessments Administered in 2022–2023

Grade	Listening	Speaking	Reading	Writing
Kindergarten	96,717	96,519	96,425	96,449
Grade 1	103,723	103,510	103,341	103,338
Grade 2	100,204	100,196	100,259	100,251
Grade 3	100,597	100,591	100,651	100,650

Grade	Listening	Speaking	Reading	Writing
Grade 4	101,659	101,649	101,693	101,692
Grade 5	103,145	103,142	103,176	103,176
Grade 6	100,388	100,383	100,455	100,455
Grade 7	98,437	98,432	98,513	98,514
Grade 8	95,401	95,395	95,483	95,480
Grade 9	96,472	96,467	96,465	96,466
Grade 10	77,503	77,502	77,622	77,616
Grade 11	55,395	55,394	55,459	55,460
Grade 12	45,346	45,346	45,476	45,476
Total	1,174,987	1,174,526	1,175,018	1,175,023

Holistic Assessments

A holistically rated assessment process is used for kindergarten and grade 1 for all four language domains. To conduct these assessments, raters are specially trained to use the ELPS PLDs as holistic rating rubrics and determine the English language proficiency of EB students based on classroom observations and daily interactions with students. EB students in grades 2–12 may qualify for special holistic administrations of TELPAS listening, speaking, or writing, which follow this same process.

Online Assessments

EB students in grades 2–12 take two online TELPAS assessments—one combined assessment for listening and speaking and one combined assessment for reading and writing. In addition to a special holistic administration of listening, speaking, and writing, an EB student may qualify for a special paper administration of TELPAS reading.

The Test Delivery System

TELPAS online assessments are administered using TDS. TDS includes the Test Administrator Interface, which is used for scheduling test sessions; the Student Interface, which enables students to participate in testing; and the Secure Browser application, which provides a secure online environment for testing. TDS allows for the secure transfer and storage of test data while remaining scalable to support the student testing population. The TDS architecture has demonstrated stability and efficiency by seamlessly handling over 1.2 million concurrent users.

Make-up Testing

Make-up testing opportunities for students who are absent on the scheduled day of testing are available during the TELPAS testing window for all grades and domains.

Proficiency Standards

For TELPAS holistically rated assessments, proficiency standards are established through descriptions of student performance in the scoring rubrics and student exemplars used in scorer training. The scoring rubrics are the ELPS PLDs, and the student exemplars are the student writing collections and student videos used in rater training.

For TELPAS online assessments, proficiency standards are established by determining the score students need to obtain to be classified into specified performance categories. The proficiency categories are the proficiency levels described in the ELPS.

Proficiency Levels and Policy Definitions

As an English language proficiency assessment, TELPAS provides an indicator of where EB students are on a continuum of English language development. This continuum is divided into four proficiency levels: Beginning, Intermediate, Advanced, and Advanced High.

Beginning

Beginning students have little or no ability to understand and use English. They may know a little English but not enough to function meaningfully in social or academic settings.

Intermediate

Intermediate students have some ability to understand and use English. They can function in social and academic settings as long as the tasks require them to understand and use simple language structures and high-frequency vocabulary in routine contexts.

Advanced

Advanced students are able to engage in grade-appropriate academic instruction in English, although ongoing second language acquisition support is needed to help them understand and use grade-appropriate language. These students function beyond the level of simple, routinely used English.

Advanced High

Advanced high students have attained the command of English that enables them, with minimal second language acquisition support, to engage in regular all-English academic instruction at their grade level.

Standard Setting

Initial proficiency standards for TELPAS reading were established in 2008. Proficiency standards for TELPAS listening and speaking were established in 2018 and were reset for TELPAS reading with the shift to online assessments for listening and speaking. In 2023, proficiency standards for TELPAS writing were established with the shift to online writing assessments. The current proficiency standard ranges for TELPAS are provided in Table 6.3.

Table 6.3. TELPAS Proficiency Standards

Domain	Grade or Grade Band	Beginning	Intermediate	Advanced	Advanced High
Listening	Grades 2–3	1000 to 1441	1442 to 1524	1525 to 1599	1600 to 2000
	Grades 4–5	1000 to 1455	1456 to 1524	1525 to 1599	1600 to 2000
	Grades 6–8	1000 to 1430	1431 to 1524	1525 to 1599	1600 to 2000
	Grades 9–12	1000 to 1447	1448 to 1524	1525 to 1599	1600 to 2000
Speaking	Grades 2–3	1000 to 1410	1411 to 1524	1525 to 1599	1600 to 2000
	Grades 4–5	1000 to 1466	1467 to 1524	1525 to 1599	1600 to 2000
	Grades 6–8	1000 to 1459	1460 to 1524	1525 to 1599	1600 to 2000
	Grades 9–12	1000 to 1484	1485 to 1524	1525 to 1599	1600 to 2000
Reading	Grade 2	1000 to 1439	1440 to 1524	1525 to 1599	1600 to 2000
	Grade 3	1000 to 1434	1435 to 1524	1525 to 1599	1600 to 2000
	Grades 4–5	1000 to 1430	1431 to 1524	1525 to 1599	1600 to 2000
	Grades 6–7	1000 to 1446	1447 to 1524	1525 to 1599	1600 to 2000
	Grades 8–9	1000 to 1437	1438 to 1524	1525 to 1599	1600 to 2000
	Grades 10–12	1000 to 1426	1427 to 1524	1525 to 1599	1600 to 2000
Writing	Grade 2	1000 to 1431	1432 to 1524	1525 to 1599	1600 to 2000
	Grade 3	1000 to 1400	1401 to 1524	1525 to 1599	1600 to 2000
	Grades 4–5	1000 to 1408	1409 to 1524	1525 to 1599	1600 to 2000
	Grades 6–7	1000 to 1428	1429 to 1524	1525 to 1599	1600 to 2000
	Grades 8–9	1000 to 1412	1413 to 1524	1525 to 1599	1600 to 2000
	Grades 10–12	1000 to 1445	1446 to 1524	1525 to 1599	1600 to 2000

Refer to the TELPAS standard-setting technical reports, which are available on the Assessment Reports and Studies webpage, for more information.

Scores and Reports

TEA publishes resources on both the TEA and Texas Assessment websites to assist school personnel in understanding and interpreting student performance data and to help parents understand their child's TELPAS results. School personnel can access TELPAS test results through CRS, parents can access their child's TELPAS results in the Family Portal, and the public can access TELPAS statewide, region, district, and campus data using the Research Portal.

TEC §39.030 and TAC §101.3014 specify the requirements for maintaining the confidentiality of individual student results and for reporting district-level and campus-level results. The results of individual student performance on state assessments are confidential and may be released only in accordance with FERPA. Districts must provide each student's state assessment results to the student, to his or her parent or guardian, and to his or her teacher for the applicable subject

area. In addition, all state assessment results must be included in each student's academic achievement record.

Description of Scores

Results for TELPAS include proficiency-level ratings for each domain, composite scores, composite ratings, and yearly progress indicators.

For TELPAS online assessments, scores include raw scores and scale scores. The number of points that a student earns on a TELPAS assessment is the student's raw score. A scale score is a conversion of the raw score onto a scale that is common to all test forms for that assessment. The scale score is used to determine whether a student achieved the Beginning, Intermediate, Advanced, or Advanced High proficiency level. Refer to Chapter 3, "Standard Technical Processes," for more information about raw scores and scale scores.

Yearly Progress Indicator

The student's yearly progress indicator provides information about the yearly proficiency-level progress that an EB student makes in acquiring the English language. This measure is based on a comparison of a student's composite rating in the previous year with his or her composite rating in the current year. The yearly statewide summary reports provide the number and percentage of students who progressed one, two, or three proficiency levels. The yearly statewide summary reports also provide the number and percentage of students who progressed at least one proficiency level. The yearly progress indicator is set as follows:

- If a student received a composite rating one level higher than the previous year, the student's yearly progress indicator is 1. Additionally, if a student received an Advanced High composite rating in the current year and in the previous year, the student's yearly progress indicator is also 1.
- If a student received a composite rating two levels higher than the previous year, the student's yearly progress indicator is 2.
- If a student received a composite rating three levels higher than the previous year, the student's yearly progress indicator is 3.
- If a student received a current year composite rating that is the same as the previous year's composite rating (excluding an Advanced High composite rating) or lower than the previous year's rating, the yearly progress indicator is 0.

The yearly progress indicator was not calculated in 2023 due to the shift to online writing assessments.

Composite Score and Rating

In addition to receiving a proficiency-level rating for each domain, students also receive a composite score and composite rating. TELPAS composite scores and ratings indicate a student's overall level of English language proficiency and are determined from the student's

listening, speaking, reading, and writing proficiency ratings. To calculate the composite score, the proficiency rating for each of the domains is converted to a domain score from 1 (Beginning) to 4 (Advanced High). The domain scores are equally weighted, as shown in Table 6.4, and added for one composite score.

Table 6.4. Language Domain Weights for TELPAS Composite Scores

Listening	Speaking	Reading	Writing
25%	25%	25%	25%

After a composite score is calculated, a composite rating is determined according to the rules below. All criteria listed for a particular rating must be met for a student to receive that rating.

Beginning:

- a composite score that fails to meet the Intermediate requirements

Intermediate:

- a composite score of 1.5 or higher
- a minimum proficiency level of Intermediate in at least half of the domains in which the student was assessed

Advanced:

- a composite score of 2.5 or higher
- a minimum proficiency level of Intermediate in all domains
- a minimum proficiency level of Advanced in at least half of the domains in which the student was assessed

Advanced High:

- a composite score of 3.5 or higher
- a minimum proficiency level of Advanced in all domains

Figure 6.1 provides a student example to show how composite results are generated. Each domain rating is converted to a domain score from 1 (Beginning) to 4 (Advanced High).

Figure 6.1. Sample Calculation of Composite Results
All Domains Assessed

Domain	Proficiency Level	Domain Score
Listening	Advanced	3
Speaking	Intermediate	2
Reading	Advanced	3
Writing	Intermediate	2

The domain scores are multiplied by the appropriate weight in Table 6.3 and then added together to obtain the composite score, as shown:

Composite Score = (Listening × 0.25) + (Speaking × 0.25) + (Reading × 0.25) + (Writing × 0.25)

Using the sample scores from the chart above, the composite score is calculated as follows:

Composite Score = (3 × 0.25) + (2 × 0.25) + (3 × 0.25) + (2 × 0.25) = 2.5

TELPAS composite scores are converted to TELPAS composite ratings. In this example, the composite score of 2.5 results in a composite rating of Advanced due to the ratings profile having:

- a TELPAS composite score of 2.5 or higher,
- a minimum proficiency level of Intermediate in all domains, and
- a minimum proficiency level of Advanced in at least half of the domains in which the student was assessed.

A small subset of EB students with disabilities who cannot be assessed in all four domains will receive a composite score if they have results for at least two domains. This is applicable only to students who have a decision from the ARD committee, in conjunction with the LPAC, to not be evaluated in one or two domains. In such instances when not all four domains are assessed, the composite score will be calculated based on the number of domains assessed.

Figure 6.2 provides a student example to show how composite results are generated when one domain is not assessed.

Figure 6.2. Sample Calculation of Composite Results
One Domain Not Assessed

Domain	Proficiency Level	Domain Score
Listening	Intermediate	2
Speaking	Intermediate	2
Reading	Beginning	1
Writing	Not Assessed	Not Assessed

The domain scores are multiplied by the appropriate weight and then added together to obtain the composite score, as shown:

Composite Score = (Listening × $\frac{1}{3}$) + (Speaking × $\frac{1}{3}$) + (Reading × $\frac{1}{3}$)

Using the sample scores from the chart above, the composite score is calculated as follows:

Composite Score = (2 × $\frac{1}{3}$) + (2 × $\frac{1}{3}$) + (1 × $\frac{1}{3}$) = 1.7

TELPAS composite scores are converted to TELPAS composite ratings. In this example, the composite score of 1.7 results in a composite rating of Intermediate due to the ratings profile having:

- a TELPAS composite score of 1.5 or higher, and
- a minimum proficiency level of Intermediate in at least half of the domains in which the student was assessed.

Assessment Reports

TEA provides reports of student performance on TELPAS to all Texas public school districts and open-enrollment charter schools. For TELPAS, TEA provides student report cards, student labels, campus rosters, summary reports, and reporting data files. In addition, TEA periodically releases TELPAS online assessments to the public through the Practice Test Site.

For more information about scoring and reporting for TELPAS, refer to the Interpreting Results page of the *Coordinator Resources*.

Use of Test Results

Test results can be used to evaluate the performance of a group over time. Average scale scores and the percentage of students achieving each proficiency level can be analyzed by grade and domain across administrations to provide insight into whether student performance is improving across years. Test results can be used when evaluating instruction or programs that require average-score or year-to-year comparisons. The tests are designed to measure English language proficiency based on the ELPS, and so the consideration of test results by domain and reporting category might be helpful when evaluating curriculum and instructional programs.

All test scores can be compared with statewide and regional performance within the same domain for any administration.

TELPAS student performance reports are used to:

- help families monitor the progress their child is making in acquiring English;
- inform instructional planning for individual students;
- report results to local school boards, school professionals, and the community;
- evaluate programs, resources, and staffing patterns; and
- evaluate district effectiveness in accountability measures.

Audits

Since the 2004–2005 school year, TEA has conducted periodic audits of the TELPAS assessment processes as a means of collecting reliability and validity evidence for the assessment program. Audits allow for the collection of information from school districts that can be used to evaluate the training, administration, and scoring of the holistically rated assessments. Information collected during TELPAS audits has been useful in the refinement of TELPAS holistic rating training and administration procedures.

Last conducted in spring 2011, an audit process for the listening and speaking domains was used in which documentation was collected from teachers at selected sites to evaluate the accuracy of holistic ratings. Due to the replacement of holistically scored assessments with an online assessment, no further audits are needed for TELPAS listening and speaking.

In the TELPAS writing audit conducted in 2019, expert raters provided second ratings of student writing samples and testing personnel at the sampled sites completed questionnaires that allowed for a conformity evaluation of training and administration procedures. Due to the replacement of the holistically rated writing assessment with an online assessment in 2023, audits are no longer necessary for TELPAS writing.

Scaling

Scaling is a statistical procedure that places raw scores on a common scoring metric to make test scores comparable across test administrations. Scaling associates numbers with characteristics of interest to provide information about measurable quantities for those characteristics. TELPAS uses the RPCM to place test items on the same Rasch scale across administrations for a given assessment. Once performance standards have been set for an assessment, the Rasch scale is then transformed to a more user-friendly metric to facilitate interpretation of the test scores. Details of the RPCM scaling method are provided in Chapter 3, “Standard Technical Processes.”

Reporting Scales

TELPAS scale scores are reported on a horizontal scale. Horizontal scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations. Refer to Chapter 3, "Standard Technical Processes," for detailed information about the scaling process.

Scale for Online Assessments

The reporting scales for each domain (listening, speaking, reading, and writing) are independent horizontal scales with lowest obtainable scale scores of 1000 and highest obtainable scale scores of 2000. The cut scores on the reporting scale for the Advanced and Advanced High proficiency levels are 1525 and 1600, respectively, to create common points of reference across the assessments for each grade and domain. It is important to note that although the Advanced and Advanced High scale score values are fixed across horizontally scaled assessments, the Intermediate scale score values vary across assessments. For any given assessment, the proficiency standards remain constant over time.

TELPAS scale scores represent linear transformations of Rasch proficiency-level estimates (θ). Specifically, the transformation is made by first multiplying θ by a slope constant (*A*) and then adding an intercept constant (*B*). This operation is described by the following equation:

$$SS_\theta = A \times \theta + B,$$

where SS_θ is the scale score for a Rasch proficiency score estimate (θ) and *A* and *B* are referred to as the horizontal scaling constants. These same transformations are applied each year to the Rasch proficiency level estimates (θ) for that year's set of test items. Values for the horizontal scaling constants are provided in Table 6.5.

Table 6.5. Horizontal Scaling Constants for TELPAS

Domain	Grade or Grade Band	*A*	*B*
Listening	Grades 2–3	67.4946	1497.4015
	Grades 4–5	64.5661	1482.9804
	Grades 6–8	67.6285	1486.0798
	Grades 9–12	53.7172	1497.3517
Speaking	Grades 2–3	35.0533	1511.4519
	Grades 4–5	24.6208	1522.0652
	Grades 6–8	19.5008	1530.4446
	Grades 9–12	21.0574	1545.1456

Domain	Grade or Grade Band	*A*	*B*
Reading	Grade 2	66.7438	1423.0422
	Grade 3	88.0488	1396.6160
	Grades 4–5	86.5951	1391.3838
	Grades 6–7	79.5756	1380.2599
	Grades 8–9	68.8452	1408.3486
	Grades 10–12	64.4607	1389.4972
Writing	Grade 2	37.5921	1452.6615
	Grade 3	41.4342	1450.0496
	Grades 4–5	57.2738	1484.0778
	Grades 6–7	58.9855	1504.2252
	Grades 8–9	58.3794	1480.8360
	Grades 10–12	68.8389	1508.4649

Scale for Holistically Rated Assessments

The scale for TELPAS holistically rated assessments (all domains for kindergarten and grade 1) ranges from 1 to 4 and is defined by the four proficiency levels: Beginning, Intermediate, Advanced, and Advanced High.

Scale for Composite Ratings

TELPAS composite ratings use a scale from 1 to 4.

Equating

Used in conjunction with the scaling process, equating is the statistical process that considers the differences in difficulty across test forms and administrations and allows scores to be placed onto a common scale. By using statistical methods, TEA equates the results of different test forms so that scale scores across test forms and testing administrations can be compared. TEA uses pre-equating, post-equating, and field-test equating for all online TELPAS assessments. During each administration, field-test equating is conducted for online TELPAS assessments through an embedded-field-test design.

Equating is not necessary for TELPAS holistically rated assessments. The difficulty level of these assessments is maintained using consistent rating rubrics developed to define the proficiency levels. Prior to test administration, raters complete training activities that provide consistency in the way the rubrics are applied. By calibrating the raters to the assessment rubric, the training maintains the difficulty of the assessment across administrations.

Refer to Chapter 3, “Standard Technical Processes,” for detailed information about equating.

Reliability

Reliability indicates the precision of test scores, referring to the expectation that repeated administrations of the same test should generate consistent results. Reliability for TELPAS test scores is estimated using statistical measures including internal consistency, classical SEM, CSEM, and classification consistency and accuracy. Data for each of these statistical measures from the spring TELPAS administration is provided in Appendix D.

In addition to the statistical measures mentioned above, TEA also collects inter-rater reliability evidence, median response times for speaking, and composite score reliability estimates.

Inter-Rater Reliability

Evidence that the holistically rated components of TELPAS result in reliable observation and rating of student performance is collected through periodic inter-rater reliability studies. Evidence of inter-rater reliability is collected through the audit process by having a second rater provide independent ratings for a sample of students.

For the TELPAS grades 2–12 writing audit conducted in 2019, districts were required to submit writing collections for approximately 2,200 EB students selected for the pure random sample, which was spread across grade levels and stratified across proficiency levels. The writing collections included writing from classroom instruction in a variety of core content areas and were rescored by Pearson scorers after the original scores were collected from the TELPAS raters. Audit results were documented in the TELPAS Writing Audit Report, available on the Assessment Reports and Studies webpage, and add to the body of validity and reliability evidence collected to support the assessment system. This process enables the evaluation of classroom activities on which the assessments are based and the way raters statewide interpret the PLD rubrics. The same information collected during TELPAS audits has been useful in the refinement of TELPAS rater training and administration procedures.

For TELPAS speaking, field-test items are examined for human-to-human and human-to-machine agreement. Evidence of inter-rater reliability is gathered by examining the perfect agreement rates and the Pearson correlations. An additional validity check is performed on the automated scoring of the responses to check inter-rater reliability between automated and human scoring. A random sample of 15 percent of students per grade band is selected for human scoring. The grade band correlations between the total raw scores on the human-scored and automated scored samples are presented in Table 6.6.

Table 6.6. TELPAS Speaking Inter-Rater Correlations

Grade	*N*	Inter-Rater Correlation
Grades 2–3	29,925	0.85
Grades 4–5	30,973	0.82
Grades 6–8	44,052	0.86
Grades 9–12	41,491	0.90

Median Response Time

When students are ready to respond to a speaking prompt, they use a speech capture tool in the online testing interface to record their responses. They have 45 seconds to respond to simple prompts and 90 seconds to respond to more complex prompts. Students are allowed two recording attempts per item; they may listen to their first recorded response and, if desired, delete it and record a second response.

Analysis was conducted on student response time to speaking items and the relationship to the overall student proficiency level on the speaking domain. This information is useful to educators and students to help demonstrate how the time spent responding impacts student performance.

Table 6.7 shows the median response time per item (for both 45-second and 90-second responses), by proficiency level, for a random sample of 5,000 students per grade band from this year's administration.

Table 6.7. TELPAS Speaking
Median Response Time per Item

Proficiency Level	Time per Item in Seconds			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Beginning	0.0	0.0	0.0	0.0
Intermediate	13.4	11.4	8.2	8.5
Advanced	29.6	23.2	23.9	18.2
Advanced High	44.2	38.6	42.2	33.3

Composite Score Reliability Estimates

TELPAS composite score reliability estimates are analyzed annually to evaluate the impact of the reliability of the listening, speaking, reading, and writing domains on the TELPAS composite reliability estimates. The composite score reliability estimates are calculated using a stratified alpha approach. This approach is described by the equation below:

$$\alpha_{Strat} = 1 - \frac{\sum_{i=1}^{k} w_i^2 \sigma_{X_i}^2 \left(1 - \rho_{X_i X_i'}\right)}{\sigma_Z^2} ,$$

where k is the number of the components or domains, w_i is the weight of each domain, X_i represents the domain score of each domain, $\rho_{X_i X_i'}$ is the internal consistency of each domain, and z is the composite score. The internal consistency values of listening, speaking, reading, and writing on the categorical scale were estimated based on their internal consistency values on the continuous scale. The results of these analyses, presented in Table 6.8, show that the weighted TELPAS composite scores have reliability estimates of at least 0.913.

Table 6.8. TELPAS Composite Score Reliability Estimates

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 2 (*n*=97926)	Listening	2.834	1.001	0.786	0.913
	Speaking	1.962	0.746	0.814	
	Writing	1.882	0.822	0.783	
	Reading	1.894	0.860	0.756	
Grade 3 (*n*=99466)	Listening	3.363	0.897	0.789	0.929
	Speaking	2.248	0.812	0.821	
	Writing	2.080	0.879	0.814	
	Reading	2.450	1.120	0.837	
Grade 4 (*n*=100728)	Listening	2.685	1.012	0.811	0.929
	Speaking	2.442	0.891	0.825	
	Writing	2.180	0.834	0.775	
	Reading	2.728	1.054	0.822	
Grade 5 (*n*=102472)	Listening	2.987	0.995	0.809	0.929
	Speaking	2.514	0.914	0.831	
	Writing	2.467	0.853	0.767	
	Reading	3.072	1.015	0.823	
Grade 6 (*n*=99576)	Listening	2.928	0.891	0.735	0.913
	Speaking	2.389	0.788	0.820	
	Writing	2.389	0.838	0.759	
	Reading	2.662	1.017	0.793	
Grade 7 (*n*=97447)	Listening	3.013	0.908	0.743	0.917
	Speaking	2.319	0.817	0.831	
	Writing	2.519	0.858	0.762	
	Reading	2.799	1.026	0.803	
Grade 8 (*n*=94274)	Listening	3.128	0.908	0.762	0.922
	Speaking	2.322	0.842	0.851	
	Writing	2.347	0.833	0.757	
	Reading	2.880	0.937	0.815	
Grade 9 (n=92661)	Listening	2.746	0.901	0.778	0.931
	Speaking	2.178	0.974	0.879	
	Writing	2.265	0.860	0.767	
	Reading	2.801	0.962	0.820	

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 10 (n=74724)	Listening	2.842	0.884	0.756	0.927
	Speaking	2.268	0.984	0.880	
	Writing	2.408	0.914	0.763	
	Reading	2.642	0.935	0.814	
Grade 11 (n=53384)	Listening	2.917	0.865	0.762	0.924
	Speaking	2.329	1.000	0.880	
	Writing	2.500	0.899	0.750	
	Reading	2.729	0.929	0.808	
Grade 12 (n=44100)	Listening	2.879	0.846	0.749	0.918
	Speaking	2.260	0.995	0.883	
	Writing	2.466	0.886	0.741	
	Reading	2.700	0.901	0.796	

Refer to Chapter 3, “Standard Technical Processes,” for detailed information about reliability.

Validity

Validity refers to the extent to which test scores accurately measure what the test is intended to measure. TEA follows national standards of best practice and annually collects validity evidence to support the interpretations and uses of TELPAS results. TTAC, a panel of national testing experts created specifically for the Texas Assessment Program, provides ongoing input to TEA about TELPAS validity evidence. The following sections describe how validity evidence has been collected for TELPAS. Refer to Chapter 3, “Standard Technical Processes,” for additional information about validity.

Evidence Based on Test Content

Validity evidence based on test content refers to evidence of the relationship between tested content and the construct that the assessment is intended to measure. Content validity evidence is collected at all stages of the test-development process. Nationally established test-development processes for the Texas Assessment Program are followed while developing TELPAS. This supports the use of TELPAS results in making inferences about students’ English language proficiency. TELPAS measures student performance in direct alignment with the English language acquisition skills and PLDs defined by the Texas ELPS that are part of the TEKS curriculum. The ELPS outline the instruction that EB students must receive to support their ability to develop academic English language proficiency.

Online Assessments

TELPAS online assessments are designed to assess English language proficiency in a manner that provides information about how well EB students understand and produce the English they need for academic success in Texas schools, as well as the types of language supports they require to comprehend written or spoken English independently.

As part of the development of TELPAS online assessments, teachers, curriculum specialists, test development specialists, and TEA staff worked together in advisory committees to identify appropriate assessment reporting categories. The input of the advisory committees was reflected in the assessed curricula and test blueprints. In addition, prototype items were developed for the assessments early in the development process. The educator advisory committees reviewed these prototypes to identify how well the items would measure the student expectations to which the items were aligned. These early reviews provided valuable suggestions for item development guidelines and item types. Item development guidelines continued to be refined through the test development process as various TELPAS item-review educator committees shared their feedback about how the student expectations in the ELPS could be effectively assessed.

As part of the annual process of item development, committees of Texas educators meet to review TELPAS items and confirm that each item appropriately measures the ELPS to which it is aligned. These committees also review items for bias. Item review committees are composed of Texas educators, and these committees revise and edit items, as appropriate, prior to field testing. Item review committees are convened for all TELPAS online assessments.

Item writers and reviewers follow test development guidelines that explain how content aligned to given ELPS should be measured. At each stage of development, writers and reviewers verify the alignment of the items with the assessed student expectations.

TELPAS online assessments are built using four levels of built-in linguistic support addressing the gradually reduced degree of linguistic accommodation that EB students need as they progress from knowing little or no English to becoming fluent in English. The levels of linguistic support are integrally related to the four proficiency levels assessed, as each proficiency level described in the ELPS is characterized by the degree of linguistic accommodation that students at that level need to understand and speak English. The staged linguistic accommodation test design is shown in Table 6.9.

Table 6.9. Staged Linguistic Accommodation Test Design

Proficiency Level	Degree of Linguistic Accommodation Applied to Stimulus and Item Development	
Beginning	Extensive	• maximum picture support • short stimuli that require comprehension of words, phrases, and short sentences that use the type of high-frequency, concrete vocabulary first acquired by learners of a second language
Intermediate	Substantial	• frequent picture support • short stimuli written primarily on familiar topics • commonly used everyday English and routine academic English
Advanced	Moderate	• occasional picture support • contextual aids and organizational features support comprehension of longer stimuli on both familiar and unfamiliar social and content-area topics
Advanced High	Minimal	• minimal linguistic accommodation • stimuli highly comparable to those intended for native English speakers

This test design supports the validity of TELPAS online assessments in that it provides built-in, staged linguistic accommodations validated by second language acquisition theory and empirical data as it measures skills in the ELPS that students need for academic success in all content areas.

Holistic Assessments

TELPAS holistically rated assessments are aligned with the ELPS and are designed to assess the English communication skills that EB students need to engage meaningfully and successfully in learning the TEKS. They draw on second language acquisition research, research-based standards, the experience of Texas educators, and observational assessment practices.

The TELPAS holistically rated components are based on ongoing observations of the ability of EB students to understand and use English during the grade-level content-area instruction required by the state-mandated curriculum and assessed by STAAR. TELPAS holistically rated assessments measure the ELPS student expectations from the cross-curricular second language acquisition knowledge and skills and use the ELPS PLDs as assessment rubrics. Rater training and administration procedures require these ratings to be based on the ability of students to use English in a variety of content areas.

Evidence Based on Response Processes

Examining students' response processes provides an additional source of validity evidence.

Online Assessments

Student response processes on TELPAS online assessments vary per item type. Across TELPAS, a variety of question types (e.g., multiple-choice, fill in the blank, drag and drop, hot spots) and response interactions are available to measure second language acquisition.

TEA gathers theoretical and empirical evidence to confirm that the type of response required for each item does not add construct-irrelevant variance. TEA also gathers evidence from several sources to confirm that response processes do not result in an advantage or disadvantage for any student group. When new item types or changes to the format of existing item types are considered for TELPAS, cognitive labs are used to study the way students engage with the various item presentations. After item types are determined to be appropriate for TELPAS, evidence about student responses is gathered annually through educator and expert reviews and analyses of individual student responses to these items. During item reviews, educators evaluate whether the content for a given item type is being appropriately assessed and whether students will be able to accurately demonstrate their knowledge of the construct given the items' planned format. When items are field-tested, additional data are gathered about students' responses. Data such as item difficulty, item point-biserial correlations, and DIF are all evaluated regarding the item type. For additional information, refer to the Item Analysis section of Chapter 3, "Standard Technical Processes."

The process used to score items can provide additional validity evidence based on response processes. This type of validity evidence is predicated on accurate scoring. For all multiple-choice and multiselect items on TELPAS, statistical key checks are conducted during the equating process. A statistical key check is a procedure in which the statistical properties of all items on every test form are computed. Items whose statistics do not meet predetermined criteria are flagged for further review by content experts to verify that the items are correctly keyed and scored. An adjudication process is used to ensure scoring reliability and validity for technology enhanced items. During adjudication, data files that include all unique responses for each test question are analyzed to identify responses or questions that require more detailed analysis to ensure accurate, consistent scoring. Evaluators who specialize in English language proficiency then review student responses to resolve scoring discrepancies or uncertainties.

For constructed-response questions, rubrics are used to evaluate student responses. All rubrics for TELPAS are validated by educator committees and content experts. In addition, TEA has implemented a rigorous scoring process for constructed responses that includes training and qualification requirements for scorers, ongoing monitoring during scoring, adjudication and resolution processes for student responses that do not meet the exact or adjacent scoring requirements, and rescoring of responses for which concerns have been raised by districts, campuses, or teachers regarding the assigned score. A more comprehensive description of the scoring process for constructed-response items is available in Chapter 2, "Building a High-Quality Assessment System."

Validity is evaluated through validity papers, which are student responses from the field test and current administrations that are representative of different levels of writing performance based on the scoring rubrics. Validity papers are identified by scoring leaders and are then

systematically given to scorers throughout the scoring project. An important feature of validity papers is that they are not identifiable as such; in fact, they are indistinguishable from unscored student responses. Each person's daily scores on validity papers are compared with the approved scores. Validity papers are used throughout the scoring project as a primary quality-control measure, the purpose of which is to ensure that scorers are accurately and reliably scoring on a daily basis and across time.

Holistic Assessments

TELPAS holistically rated assessments are based on ongoing classroom observations and daily interaction with students. As is typical of holistically scored assessments, students are evaluated on their overall performance in a global and direct way. TELPAS holistically rated assessments meet the goal of English language proficiency assessments to effectively assess the extent to which EB students are making progress in attaining academic language proficiency by serving as direct measures of the ability of students to understand and use English while engaging in state-required academic instruction. As such, the assessments provide strong validity evidence related to the response process.

Evidence Based on Internal Structure

TEA collects evidence that reflects the relationship between item performance and proficiency levels to verify that patterns of item performance are consistent with the constructs the assessment is intended to measure.

Online Assessments

Internal consistency reliability estimates provide a measure of the consistency with which students respond to the items in an assessment and show the relationship of students' responses between items, within reporting categories of items, and within domains to verify that the elements of an assessment conform to the intended test construct. The internal consistency of TELPAS online assessments is evaluated using KR20 for assessments that have only dichotomously scored items. For TELPAS online assessments that have a combination of dichotomous and polytomous items, internal consistency is evaluated using coefficient alpha and stratified alpha. These internal consistency evaluations are made for all students and for female and male student groups. Estimates of internal consistency can be found in Appendix D.

Holistic Assessments

Evidence of the validity of TELPAS holistic assessments is supported by comprehensive training and administration procedures that prepare raters to perform their duties and district administrators to follow procedures to maintain the integrity of the test administration. In addition to holistic rating training, raters must perform calibration activities to demonstrate high accuracy in rating student activities across all TELPAS holistically rated domains they will assess. Additional support is provided to raters who cannot calibrate on their first two attempts in order to help them assess assigned students in a manner consistent with the PLDs.

TELPAS holistic rating audits provide both validity and reliability evidence based on the internal structure by examining the extent to which raters follow the defined protocol for rating these components. As part of the audit, reports of rater adherence to the assessment protocol are made and used to provide evidence that the internal structure of the assessment is intact and that educators are administering the assessment and applying the scoring rubrics appropriately.

In addition to directly supporting the state's goal of having a valid and authentic assessment, TELPAS holistically rated assessments also serve an ongoing critical role as a professional development tool that supports effective instruction, enabling teachers to better understand and meet the educational needs of EB students.

Evidence Based on Consequences of Testing

Another method for providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. The collection of consequential validity evidence typically occurs after a program has been in place for some time and on a regular basis.

Given the important stakes associated with TELPAS, valid test scores are critical in supporting their intended interpretations and uses. The intended interpretations of TELPAS results are stated in the policy definitions of the four proficiency levels. Refer to the Proficiency Standards section for the policy definitions of the TELPAS proficiency levels. The ELPS PLDs describe a student's English language acquisition skills in each domain based on the student's proficiency level.

Student-Level Performance

The following are the intended uses of TELPAS results based on the policy definitions for student-level performance:

- Proficiency on TELPAS is one indicator of a student's level of proficiency in learning English.
- Proficiency on TELPAS is one indicator of a student's possible need for academic intervention.
- Proficiency on TELPAS across years provides one indicator of a student's English language acquisition within a domain.

District- or Campus-Level Performance

The following are the intended uses of TELPAS test results based on the policy definitions for district- or campus-level performance:

- TELPAS results provide an indicator of overall student English language proficiency at a district or campus.
- TELPAS results can be aggregated across years to provide an indicator of overall student progress in English language acquisition at a district or campus.

Evidence based on the consequences of testing can also be found by comparing performance from past administrations, which is represented in Appendix D. The proficiency-level classifications of students for the listening, speaking, reading, and writing domains of TELPAS have been continually collected since the first administration. In general, long-term trends show a gradual increase in student performance after the introduction of TELPAS, and such improvement may have resulted, in part, from the use of test data to inform instruction.

While TELPAS has continued to assess the same ELPS, changes to the assessment design over time make comparisons to earlier results difficult to interpret. Comparisons in performance are only appropriate across certain years. For example, TELPAS writing results for all grades can be compared from 2005 until 2022. For grades 2–12 listening and speaking, results can be compared within the periods of 2005–2017 and 2018–present. For reading, results are comparable within the periods of 2005–2013, 2014–2017, and 2018–present. However, direct comparisons across these distinct periods are not appropriate. If historical trends hold, however, over time the percentages of students across proficiency levels are expected to remain relatively stable, with the possibility of a gradual increase in performance.

In addition to district and campus consequences, based on what educators learn during rater training and from the observation process, the administration of TELPAS holistically rated assessments leads to improvements in students' language acquisition for both formative and summative purposes. For example, educators learn how developing academic language proficiency in English relates to and supports academic achievement in English.

Sampling

Sampling is a procedure that is used to select and examine a small set that is representative of the population from which it was drawn. For TELPAS, sampling occurs when observed *n*-counts for handscored field-test items exceed 3,000.

Test Results

Appendix D provides consistency and accuracy data, scale score correlations, CSEMs, mean *p*-values, scale score descriptive statistics, frequency distributions, and proficiency level distributions for the spring TELPAS administration. The percentages of students in each proficiency level for all four domains as well as for the composite rating are available on the Statewide Summary Reports webpage.

TECHNICAL DIGEST 2022–2023

Chapter 7

TELPAS Alternate

Overview

Participation Requirements

Test Development

Training

Test Administration

Proficiency Standards

Scores and Reports

Scaling

Reliability

Validity

Sampling

Test Results

Overview

TELPAS Alternate is an alternate English language proficiency assessment that measures the progress that EB students in grades 2–12 with the most significant cognitive disabilities make in acquiring the English language. It fulfills the ESSA requirements that EB students who cannot participate in the general English language proficiency assessment even with allowable accommodations be assessed annually until they are determined to be proficient in the English language.

TELPAS Alternate is a holistic inventory based on ongoing classroom observations used for the four language domains of listening, speaking, reading, and writing.

Participation Requirements

All EB students in kindergarten through grade 12 are required to participate in TELPAS or TELPAS Alternate. EB students are assessed annually in English language proficiency until they are determined to be proficient by meeting the EB reclassification criteria. This includes students classified as emergent bilingual (EB)/English learner (EL) in PEIMS whose parents have declined bilingual or ESL program services.

EB students in kindergarten and grade 1 are assessed with the general TELPAS in which they are holistically rated in all four language domains. EB students in grades 2–12 who receive special education services and have the most significant cognitive disabilities are eligible to participate in TELPAS Alternate. These students exhibit significant intellectual and adaptive behavior deficits that limit their ability to plan, comprehend, and reason as well as adaptive behavior deficits that limit their ability to apply social and practical skills (e.g., personal care, social problem-solving skills, dressing, eating, using money) across all life domains. Students with the most significant cognitive disabilities require extensive, direct, individualized instruction and have a need for substantial supports that are neither temporary nor content specific.

TELPAS Alternate has specific participation requirements that an ARD committee, in conjunction with the LPAC, must carefully review and consider annually. The TELPAS Alternate Participation Requirements, available in English and Spanish on the TELPAS Alternate Resources webpage, detail the ARD committee's responsibility for ensuring that a student is eligible for TELPAS Alternate. Prior to reviewing the eligibility criteria for TELPAS Alternate, the ARD committee must understand all assessment options, including the characteristics of each assessment and the potential implications of each assessment choice. If TELPAS Alternate is being considered, the ARD committee, in conjunction with the LPAC, must review the participation requirements against the supporting documentation within the student's IEP, such as in the present levels of academic achievement and functional performance, to determine eligibility.

In rare circumstances a student's ARD committee might determine prior to the administration of the assessment that the student will not participate in TELPAS Alternate because the student meets the eligibility criteria for a medical exception or NAAR. For both exceptions, the ARD committee reviews educational records and eligibility requirements. For more information, refer to the eligibility criteria on the TELPAS Alternate Resources webpage.

Test Development

Maintaining a high-quality student assessment program involves a complex and detailed test-development process, and TEA relies on input from educators to ensure that all measures of learning for Texas public school students are equitable and accurate. For more information regarding each step of the TELPAS Alternate test-development process, refer to Chapter 2, “Building a High-Quality Assessment System," which outlines the processes used to develop the TELPAS Alternate assessment’s framework.

For the initial development of TELPAS Alternate, TEA sought input from educator committees, state assessment experts, administrators, second language acquisition experts and researchers, and bilingual, ESL, and special education teachers and coordinators. Particular attention was given to:

- eligibility requirements,
- alternate PLDs,
- Observable Behaviors for the holistic inventory, and
- classroom examples for each Observable Behavior.

TELPAS Alternate is aligned to the ELPS and measures English language proficiency in listening, speaking, reading, and writing. TELPAS Alternate is based on alternate PLDs that were created to address the specific access needs of this student population.

More information about the development of TELPAS Alternate is available in the *TELPAS Alternate Educator Guide* available on the TELPAS Alternate Resources webpage. Provided to familiarize educators with TELPAS Alternate, the guide shows the integral relationship between TELPAS Alternate and the ELPS and includes explanatory information on the TELPAS Alternate language domains of listening, speaking, reading, and writing, as well as a sample of Observable Behaviors for reading.

Training

TEA develops instructional materials, including manuals, guides, presentations, online modules, and videos, to support the training of all testing personnel on test security and administration procedures. Preparation for test administration begins every year with a TEA-provided training-of-trainers session for testing coordinators from each of the 20 Texas regional ESCs as well as district testing coordinators from the state’s 25 largest districts. Using materials and information provided in the TEA training session, ESC regional testing coordinators train the district coordinators in their respective regions. District coordinators then train their campus testing coordinators, who are responsible for training test administrators.

Test security and administration procedures provided in the *Coordinator Resources* and the *TELPAS Alternate Test Administrator Manual* must be followed so that all students have an equal opportunity to demonstrate their knowledge and skills. The *Coordinator Resources* guide district and campus coordinators through their responsibilities as they oversee the administration of the Texas Assessment Program. This online resource contains preparation

and administration procedures for each state-required assessment and is available prior to the annual ESC training.

In addition, TEA produces the *TELPAS Alternate Educator Guide* to familiarize educators with the assessment. The guide includes information on test design, alignment with state curriculum, training, and test results.

Training presentations for each of the four domains are also available to help test administrators understand the language in the holistic inventory, provide rating examples, and suggest ways that the Observable Behaviors can be demonstrated in the classroom.

Test Administration

TELPAS Alternate is administered once a year, in the spring, during a six-week testing window. TELPAS Alternate includes 40 descriptions of behaviors, called Observable Behaviors, for test administrators to consider regarding each student's use of the English language in the four language domains. Each Observable Behavior includes descriptions of characteristics that EB students with the most significant cognitive disabilities learning English demonstrate as they gain proficiency. Test administrators read the continuum of descriptions for each Observable Behavior and indicate the description that most accurately describes a student's skills for that Observable Behavior at the time of the test administration.

The intent of this observational design is to increase student access to TELPAS Alternate and account for a student's communication mode, thereby resulting in a more accurate measurement of a student's English proficiency level. For purposes of TELPAS Alternate, "English" is not limited to the typical spoken or written English of other state assessments. Test administrators should consider whether another method of communication as a substitute for traditional English is a more appropriate way for some students to demonstrate proficiency in a specific language domain.

The number of TELPAS Alternate assessments that were administered to eligible students in 2022–2023 is indicated in Table 7.1.

Table 7.1. TELPAS Alternate Assessments Administered in 2022–2023

Grade	Assessments Administered
Grade 2	1,641
Grade 3	1,669
Grade 4	1,513
Grade 5	1,277
Grade 6	1,222
Grade 7	1,017
Grade 8	860
Grade 9	756
Grade 10	573

Grade	Assessments Administered
Grade 11	410
Grade 12	402

Proficiency Standards

TELPAS Alternate is a holistic inventory based on alternate PLDs that describe the English language acquisition progress of a student with the most significant cognitive disabilities. The inventory includes 40 descriptions of behaviors, called Observable Behaviors, for test administrators to consider regarding each student's use of the English language in the four language domains.

Proficiency Levels and Policy Definitions

As an English language proficiency assessment, TELPAS Alternate provides an indicator of where EB students with the most significant cognitive disabilities are on a continuum of English language development. This continuum is divided into five proficiency levels: Awareness, Imitation, Early Independence, Developing Independence, and Basic Fluency.

Awareness

Students who receive this rating may be aware of English sounds or print; however, they have little or no functional ability to participate in communication activities in English.

Imitation

Students who receive this rating match, imitate, or approximate some English in their environment; however, they are not able to independently understand or produce English. They participate in routine communication activities in a familiar environment when the activities are significantly linguistically accommodated.

Early Independence

Students who receive this rating understand short, simple messages and produce messages of one or two high-need, high-frequency words (e.g., book, cafeteria, teacher). They are starting to participate in linguistically accommodated communication activities in English in familiar environments.

Developing Independence

Students who receive this rating understand longer messages of multiple sentences in English and produce simple, descriptive, original messages by combining two or more words (e.g., new red bike, big fast truck). They participate meaningfully in linguistically accommodated communication activities in English in familiar environments.

Basic Fluency

Students who receive this rating understand and produce more detailed, complex, and elaborate messages with multiple sentences in English. They participate independently in communication activities in English in familiar environments.

Standard Setting

TELPAS Alternate proficiency standards were established in 2019. Table 7.2 shows the scale score ranges as established by the standard-setting committee.

Table 7.2. TELPAS Alternate Proficiency Standards

Domain	Awareness Level	Imitation Level	Early Independence Level	Developing Independence Level	Basic Fluency Level
Listening	600–698	699–749	750–799	800–857	858–1000
Speaking	600–681	682–749	750–799	800–863	864–1000
Reading	600–703	704–749	750–799	800–848	849–1000
Writing	600–700	701–749	750–799	800–852	853–1000

Refer to the TELPAS Alternate Standard Setting Technical Report, which is available on the Assessment Reports and Studies webpage, for more information.

Scores and Reports

TEA publishes resources on both the TEA and Texas Assessment websites to assist school personnel in understanding and interpreting student performance data and to help parents understand their child's TELPAS Alternate results. School personnel can access TELPAS Alternate test results through CRS, parents can access their child's TELPAS Alternate results in the Family Portal, and the public can access TELPAS Alternate statewide, region, district, and campus data using the Research Portal.

TEC §39.030 and TAC §101.3014 specify the requirements for maintaining the confidentiality of individual student results and for reporting district-level and campus-level results. The results of individual student performance on state assessments are confidential and may be released only in accordance with FERPA. Districts must provide each student's state assessment results to the student, to his or her parent or guardian, and to his or her teacher for the applicable subject area. In addition, all state assessment results must be included in each student's academic achievement record.

Description of Scores

Results for TELPAS Alternate include raw scores, scale scores, proficiency ratings, composite scores, composite ratings, and yearly progress indicators.

The number of points that a student earns in the holistic inventory for each domain is the student's raw score. A scale score is a conversion of the raw score onto a scale that is common to all test forms for that assessment. The scale score is used to determine whether a student

achieved the Awareness, Imitation, Early Independence, Developing Independence, or Basic Fluency proficiency level. Refer to Chapter 3, "Standard Technical Processes," for more information about raw scores and scale scores.

Composite Score and Rating

In addition to receiving a proficiency-level rating for each domain, students also receive a composite score and composite rating. If students are not rated in all four domains, they will not receive any TELPAS Alternate scores or ratings, including composite scores and ratings. TELPAS Alternate composite scores and ratings indicate a student's overall level of English language proficiency and are determined from the student's listening, speaking, reading, and writing proficiency ratings. To calculate the composite score, the proficiency rating for each of the domains is converted to a domain score from 1 (Awareness) to 5 (Basic Fluency). The domain scores are equally weighted, as shown in Table 7.3, and added for one composite score.

Table 7.3. Language Domain Weights for TELPAS Alternate Composite Scores

Listening	Speaking	Reading	Writing
25%	25%	25%	25%

After a composite score is calculated, a composite rating is determined according to the rules below. All criteria listed for a particular rating must be met for a student to receive that rating.

- Awareness
  - a composite score that fails to meet the Imitation requirements
- Imitation
  - a composite score of 1.5 or higher
  - a minimum proficiency level of Imitation in at least two domains
- Early Independence
  - a composite score of 2.25 or higher
  - a minimum proficiency level of Early Independence in at least two domains
  - a minimum proficiency level of Imitation in at least three domains
- Developing Independence
  - a composite score of 3.25 or higher
  - a minimum proficiency level of Developing Independence in at least two domains
  - a minimum proficiency level of Imitation in all domains

- Basic Fluency
  - a composite score of 4 or higher
  - a minimum proficiency level of Basic Fluency in at least two domains
  - a minimum proficiency level of Early Independence in all domains

Figure 7.1 provides a student example to show how composite results are generated. Each domain rating is converted to a domain score from 1 (Awareness) to 5 (Basic Fluency).

Figure 7.1. Sample Calculation of Composite Results

Domain	Proficiency Level	Domain Score
Listening	Developing Independence	4
Speaking	Early Independence	3
Reading	Developing Independence	4
Writing	Imitation	2

The domain scores are multiplied by the appropriate weight in Table 7.3 and then added together to obtain the composite score, as shown:

Composite Score = (Listening × 0.25) + (Speaking × 0.25) + (Reading × 0.25) + (Writing × 0.25).

Using the sample domain scores from the chart above, the composite score is calculated as follows:

Composite Score = (4 × 0.25) + (3 × 0.25) + (4 × 0.25) + (2 × 0.25) = 3.25

TELPAS Alternate composite scores are converted to TELPAS Alternate composite ratings. In this example, the composite score of 3.25 results in a composite rating of Developing Independence due to the ratings profile having:

- a TELPAS Alternate composite score of 3.25 or higher,
- a minimum proficiency level of Developing Independence in at least two domains, and
- a minimum proficiency level of Imitation in all domains.

Yearly Progress Indicator

The student's yearly progress indicator provides information about the yearly proficiency level progress that an EB student with the most significant cognitive disabilities makes in acquiring the English language. This measure is based on a comparison of a student's composite rating in the previous year with his or her composite rating in the current year. The yearly statewide summary reports provide the number and percentage of students who progressed one, two, three, or four proficiency levels. The yearly statewide summary reports also provide the number and percentage of students who progressed at least one proficiency level. The yearly progress indicator is set as follows:

- If a student received a composite rating one level higher than the previous year, the student's yearly progress indicator is 1. Additionally, if a student received a Basic Fluency composite rating in the current year and in the previous year, the student's yearly progress indicator is also 1.
- If a student received a composite rating two levels higher than the previous year, the student's yearly progress indicator is 2.
- If a student received a composite rating three levels higher than the previous year, the student's yearly progress indicator is 3.
- If a student received a composite rating four levels higher than the previous year, the student's yearly progress indicator is 4.
- If a student received a current year composite rating that is the same as the previous year's composite rating (excluding a Basic Fluency composite rating) or lower than the previous year's composite rating, the yearly progress indicator is 0.

The yearly progress indicator is provided on the summary reports for each assessed grade that contain information about every student for whom a TELPAS Alternate record was submitted.

Assessment Reports

TEA provides reports of student performance on TELPAS Alternate to all Texas public school districts and open-enrollment charter schools. For TELPAS Alternate, TEA provides student report cards, student labels, campus rosters, summary reports, and reporting data files.

For more information about scoring and reporting for TELPAS Alternate, refer to the Interpreting Results page of the *Coordinator Resources*.

Use of Test Results

Test results can be used to evaluate the performance of a group over time. Average scale scores and the percentage of students achieving each proficiency level can be analyzed by domain across administrations to provide insight into whether student performance is improving across years. Test results can be used when evaluating instruction or programs that require average-score or year-to-year comparisons. The tests are designed to measure English language proficiency based on the ELPS, and so the consideration of test results by domain might be helpful when evaluating curriculum and instructional programs. All test scores can be compared with statewide and regional performance within the same domain for any administration.

TELPAS Alternate student performance reports are used to:

- help families monitor the progress their child is making in acquiring English;
- inform instructional planning for individual students;
- report results to local school boards, school professionals, and the community;

- evaluate programs, resources, and staffing patterns; and
- evaluate district effectiveness in accountability measures.

Scaling

Scaling is a statistical procedure that places raw scores on a common scoring metric to make test scores comparable across test administrations. Scaling associates numbers with characteristics of interest to provide information about measurable quantities for those characteristics. TELPAS Alternate uses the RPCM to place test questions (Observable Behaviors) on the same Rasch scale across administrations for a given assessment. Once performance standards have been set for an assessment, the Rasch scale is then transformed to a more user-friendly metric to facilitate interpretation of the test scores. Details of the RPCM scaling method are provided in Chapter 3, "Standard Technical Processes."

Reporting Scales

TELPAS Alternate scale scores are reported on a horizontal scale. Horizontal scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations. Refer to Chapter 3, "Standard Technical Processes," for detailed information about the scaling process.

The reporting scales for the four domains are independent scales with the lowest obtainable scale scores of 600 and the highest obtainable scale scores of 1000. The cut scores on the reporting scale for the Early Independence and Developing Independence proficiency levels are 750 and 800, respectively, to create common points of reference across the assessments for each domain. It is important to note that although the Early Independence and Developing Independence scale score values are fixed across horizontally scaled assessments, the Imitation and Basic Fluency scale score values vary across TELPAS Alternate domains. For each domain, the scale score values of the proficiency-level cut score remain constant over time.

TELPAS Alternate scale scores represent linear transformations of Rasch proficiency-level estimates (θ). Specifically, the transformation is made by first multiplying θ by a slope constant (A) and then adding an intercept constant (B). This operation is described by the following equation:

$$SS_\theta = A \times \theta + B,$$

where SS_θ is the scale score for a Rasch proficiency-level estimate (θ) and A and B are referred to as the horizontal scaling constants. Once established, these same transformations are applied each year to the Rasch proficiency-level estimates for that year's set of test questions. Values for the horizontal scaling constants are provided in Table 7.4.

Table 7.4. Horizontal Scaling Constants for TELPAS Alternate

Domain	*A*	*B*
Listening	22.6974	770.8089
Speaking	20.9486	772.9659
Reading	22.0080	778.9801
Writing	20.3990	785.4575

TELPAS Alternate composite ratings use a scale from 1 to 5.

Reliability

Reliability indicates the precision of test scores, referring to the expectation that repeated administrations of the same test should generate consistent results. Reliability for TELPAS Alternate test scores is estimated using statistical measures including internal consistency, classical SEM, CSEM, and classification consistency and accuracy. Data for each of these statistical measures from the spring TELPAS Alternate administration is provided in Appendix E.

In addition to the statistical measures mentioned above, TEA also collects composite score reliability estimates.

Composite Score Reliability Estimates

TELPAS Alternate composite score reliability estimates are analyzed annually to evaluate the impact of the reliability of the listening, speaking, reading, and writing domains on the TELPAS Alternate composite reliability estimates. The composite score reliability estimates are calculated using a stratified alpha approach. The internal consistency values of listening, speaking, reading, and writing on the categorical scale are estimated based on their internal consistency values on the continuous scale. The results of these analyses are shown in Table 7.5.

Table 7.5. TELPAS Alternate Composite Score Reliability Estimates

Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Listening	3.173	1.279	0.927	0.984
Speaking	3.101	1.394	0.947	
Reading	2.558	1.337	0.950	
Writing	2.859	1.411	0.945	

Refer to Chapter 3, “Standard Technical Processes,” for detailed information about reliability.

Validity

Validity refers to the extent to which test scores accurately measure what the test is intended to measure. TEA follows national standards of best practice and annually collects validity evidence to support the interpretations and uses of TELPAS Alternate results. TTAC, a panel of national

testing experts created specifically for the Texas Assessment Program, provides ongoing input to TEA about TELPAS Alternate validity evidence. The following sections describe how validity evidence has been collected for TELPAS Alternate. Refer to Chapter 3, “Standard Technical Processes,” for additional information about validity.

Evidence Based on Test Content

Validity evidence based on test content refers to evidence of the relationship between tested content and the construct that the assessment is intended to measure. TELPAS Alternate measures student performance in direct alignment with the English language acquisition skills defined by the Texas ELPS that are part of the TEKS curriculum. The ELPS outline the instruction that EB students must receive to support their ability to develop academic English language proficiency.

TELPAS Alternate is designed to assess English language proficiency in a manner that provides information about how well EB students with the most significant cognitive disabilities understand and produce the English they need for academic success in Texas schools, as well as the types of language supports they require to independently comprehend written or spoken English.

TELPAS Alternate's holistic inventory of 40 Observable Behaviors, 10 per domain, is based on ongoing observations of the ability of EB students with the most significant disabilities to understand and use English during the content-area instruction required by the state-mandated curriculum and assessed by STAAR Alternate 2. TELPAS Alternate measures the ELPS student expectations from the cross-curricular second language acquisition knowledge and skills and uses the alternate PLDs as assessment rubrics. Rater training and administration procedures require that holistic inventories be based on the ability of students to use English in a variety of content areas.

Evidence Based on Response Processes

Examining students' response processes provides an additional source of validity evidence. For TELPAS Alternate, this validity evidence is gathered to confirm that the way administrators engage with the Observable Behaviors does not add construct-irrelevant variance.

TELPAS Alternate's holistic inventory is based on ongoing classroom observations and daily interaction with students through the Observable Behaviors. As is typical of holistically scored assessments, students are evaluated on their overall performance in a global and direct way.

During the development of TELPAS Alternate, cognitive labs were used to study the way test administrators engaged with the Observable Behaviors and classroom examples. Student response evidence was also gathered about Observable Behaviors through educator and expert reviews and analyses of responses.

TELPAS Alternate's holistic inventory meets the goal of an English language proficiency assessment to effectively assess the extent to which EB students are making progress in attaining academic language proficiency by serving as a direct measure of the ability of EB students with the most significant cognitive disabilities to understand and use English while

engaging in state-required academic instruction. As such, the holistic inventory of Observable Behaviors provides strong validity evidence related to the response process.

Evidence Based on Internal Structure

TEA collects evidence that reflects the relationship between test performance and proficiency levels to verify that patterns of test performance are consistent with the constructs that the assessment is intended to measure.

Evidence of the validity of TELPAS Alternate is supported by training and administration procedures that prepare administrators to perform their duties and district administrators to follow procedures to maintain the integrity of the test administration.

Internal structure is evaluated annually by estimating the internal consistency reliability for the TELPAS Alternate ratings of student performance. Internal consistency reliability estimates provide a measure of the consistency with which test administrators evaluate students. The internal consistency of TELPAS Alternate is evaluated each year using coefficient alpha statistics that can be found in Appendix E.

In addition to directly supporting the state's goal of having a valid and authentic assessment for EB students with the most significant cognitive disabilities, TELPAS Alternate also serves an ongoing critical role as a professional development tool that supports effective instruction, enabling administrators to better understand and meet the educational needs of EB students with the most significant cognitive disabilities.

Evidence Based on Relationships to Other Variables

TEA also provides validity evidence for TELPAS Alternate by analyzing the relationship between test performance and performance on external measures. By examining this relationship, evidence can be collected to show that the relationships are consistent with those expected at the level of the construct underlying the proposed score interpretations.

Evidence based on external measures comes from analyses of the relationship between performance on TELPAS Alternate reading and writing and STAAR Alternate 2 RLA. TELPAS Alternate measures English language proficiency while STAAR Alternate 2 assumes that students already understand the English language and focuses instead on assessing the degree to which students can apply the skills in the RLA TEKS. Because of the differences in the designs and purposes of these two assessments, one would not expect EB students to perform at the same level of proficiency on the two assessments. One would, however, expect EB students who have comparatively little difficulty understanding English to score higher on the STAAR Alternate 2 RLA tests when compared with EB students who are in earlier stages of English fluency.

To examine the relationship between performance on the assessments, the average scale score and pass rate for each STAAR Alternate 2 RLA assessment was calculated based on the TELPAS Alternate proficiency level achieved.

Table 7.6 shows the TELPAS Alternate reading data for students who participated in both STAAR Alternate 2 and TELPAS Alternate. Table 7.7 shows the TELPAS Alternate writing data for students who participated in both assessments.

Table 7.6. Reading Proficiency Compared to RLA Performance

Grade or Course	TELPAS Alternate Reading Proficiency Level	Number of Students	STAAR Alternate 2 RLA Average Scale Score*	STAAR Alternate 2 Passing Rate
3	Awareness	500	322	78
	Imitation	362	342	91
	Early Independence	424	355	96
	Developing Independence	248	359	98
	Basic Fluency	107	368	98
4	Awareness	315	316	71
	Imitation	283	343	91
	Early Independence	382	351	95
	Developing Independence	323	364	98
	Basic Fluency	199	376	98
5	Awareness	227	309	68
	Imitation	208	337	91
	Early Independence	288	350	94
	Developing Independence	301	363	98
	Basic Fluency	238	374	99
6	Awareness	222	321	75
	Imitation	201	341	89
	Early Independence	275	355	95
	Developing Independence	251	366	98
	Basic Fluency	255	380	99
7	Awareness	194	328	78
	Imitation	141	347	91
	Early Independence	218	364	98
	Developing Independence	224	369	98
	Basic Fluency	227	375	99
8	Awareness	139	331	79
	Imitation	95	347	89
	Early Independence	174	351	95
	Developing Independence	173	364	99
	Basic Fluency	259	379	99
English I	Awareness	153	336	83
	Imitation	102	354	97
	Early Independence	154	363	99
	Developing Independence	166	378	99
	Basic Fluency	171	386	100
English II	Awareness	121	337	85
	Imitation	80	354	95
	Early Independence	95	361	97
	Developing Independence	136	375	99
	Basic Fluency	141	384	100

* A scale score of 300 is necessary to meet the STAAR Alternate 2 Level II: Satisfactory performance level.

Table 7.7. Writing Proficiency Compared to RLA Performance

Grade or Course	TELPAS Alternate Writing Proficiency Level	Number of Students	STAAR Alternate 2 RLA Average Scale Score*	STAAR Alternate 2 Passing Rate
3	Awareness	619	325	80
	Imitation	462	348	95
	Early Independence	361	357	96
	Developing Independence	162	363	98
	Basic Fluency	37	383	100
4	Awareness	408	320	73
	Imitation	350	347	95
	Early Independence	417	357	97
	Developing Independence	233	367	98
	Basic Fluency	94	389	100
5	Awareness	275	313	70
	Imitation	288	344	92
	Early Independence	329	357	98
	Developing Independence	240	365	98
	Basic Fluency	130	377	99
6	Awareness	264	322	77
	Imitation	250	346	91
	Early Independence	306	363	96
	Developing Independence	220	366	99
	Basic Fluency	164	384	99
7	Awareness	220	327	78
	Imitation	200	355	96
	Early Independence	237	370	98
	Developing Independence	195	369	98
	Basic Fluency	152	378	99
8	Awareness	150	329	98
	Imitation	148	350	93
	Early Independence	167	359	98
	Developing Independence	173	364	99
	Basic Fluency	202	381	99
English I	Awareness	175	337	84
	Imitation	136	355	99
	Early Independence	160	372	99
	Developing Independence	148	380	99
	Basic Fluency	127	388	100
English II	Awareness	132	337	86
	Imitation	102	352	95
	Early Independence	114	371	97
	Developing Independence	110	380	100
	Basic Fluency	115	385	99

* A scale score of 300 is necessary to meet the STAAR Alternate 2 Level II: Satisfactory performance level.

As shown in Tables 7.6 and 7.7, the average STAAR Alternate 2 scale score increases as students' TELPAS Alternate proficiency levels increase within a grade. In addition, passing rates on STAAR Alternate 2 tend to increase as TELPAS Alternate proficiency levels increase within each grade.

Evidence Based on Consequences of Testing

Another method for providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. The collection of consequential validity evidence typically occurs after a program has been in place for some time and on a regular basis.

Given the important stakes associated with TELPAS Alternate, valid test scores are critical in supporting their intended interpretations and uses. The intended interpretations of TELPAS Alternate results are stated in the policy definitions of the five proficiency levels. Refer to the Proficiency Standards section for the policy definitions of the TELPAS Alternate proficiency levels. The alternate PLDs describe a student's English language acquisition skills in each domain based on the student's proficiency level.

Student-Level Performance

The following are the intended uses of TELPAS Alternate results based on the policy definitions for student-level performance:

- Proficiency on TELPAS Alternate is one indicator of a student's level of proficiency in learning English.
- Proficiency on TELPAS Alternate is one indicator of a student's possible need for academic intervention.
- Proficiency on TELPAS Alternate across years provides one indicator of a student's English language acquisition within a domain.

District- or Campus-Level Performance

The following are the intended uses of TELPAS Alternate test results for district- or campus-level performance based on the policy definitions:

- TELPAS Alternate results provide an indicator of overall English language proficiency for students with the most significant cognitive disabilities at a district or campus.
- TELPAS Alternate results can be aggregated across years to provide an indicator of overall progress in English language acquisition for students with the most significant cognitive disabilities at a district or campus.

In addition to district and campus consequences, based on what educators learn during rater training and from the observation process, the administration of TELPAS Alternate leads to improvements in students' language acquisition for both formative and summative purposes. For example, educators learn how developing academic language proficiency in English relates to and supports academic achievement in English. A 2019 survey of TELPAS Alternate test administrators suggested that the administration of TELPAS Alternate led to an increase in students receiving support from bilingual and ESL staff and an increase in special education teachers obtaining ESL certification to better serve this student population.

Sampling

Sampling is a procedure that is used to select and examine a small set that is representative of the population from which it was drawn. Because TELPAS Alternate has been administered for only a few years, audits and sampling have not been conducted.

Test Results

Appendix E provides consistency and accuracy data, scale score correlations, CSEMs, mean *p*-values, scale score descriptive statistics, and frequency distributions for the spring TELPAS Alternate administration. The percentages of students in each proficiency level for all four domains as well as for the composite rating are available on the Statewide Summary Reports webpage.

TECHNICAL DIGEST 2022–2023

Chapter 8

Resources

Resources

Information about the Texas Assessment Program can be found on the TEA Student Assessment webpages and the Texas Assessment website. A summary of some available resources is provided in Table 8.1.

Table 8.1 Texas Assessment Program Online Resources

Topic	URL
Texas Assessment Systems	http://texasassessment.gov/testing-personnel.html
District and Campus Coordinator Resources	https://txassessmentdocs.atlassian.net/wiki/spaces/ODCCM/overview
STAAR Resources	https://tea.texas.gov/student-assessment/testing/staar/staar-resources
STAAR Alternate 2 Resources	https://tea.texas.gov/student-assessment/testing/staar-alternate/staar-alternate-2-resources
TELPAS Resources	https://tea.texas.gov/student-assessment/testing/telpas/telpas-resources
TELPAS Alternate Resources	https://tea.texas.gov/student-assessment/testing/telpas/telpas-alternate-resources
Assessments for Special Populations	https://tea.texas.gov/student-assessment/testing/student-assessment-overview/assessments-for-special-populations
Accommodation Resources	https://tea.texas.gov/student-assessment/testing/student-assessment-overview/accommodation-resources
Test Administration Resources	https://tea.texas.gov/student-assessment/testing/student-assessment-overview/test-administration-resources
Student Assessment Results	https://tea.texas.gov/student-assessment/testing/student-assessment-results
Assessment Reports and Studies	https://tea.texas.gov/student-assessment/testing/student-assessment-overview/assessment-reports-and-studies
Assessment Resources for Educators	https://www.texasassessment.gov/educators.html
Assessment Resources for Students and Families	https://www.texasassessment.gov/families.html

TECHNICAL DIGEST 2022–2023

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508–600). American Council on Education.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory.* Cengage Learning.

Davies, S., O'Malley, K., & Wu, B. (2007, April 9-13). *Establishing measurement equivalence of transadapted reading and mathematics tests* [Paper presentation]. American Educational Research Association Annual Meeting, Chicago, IL, United States. https://convention2.allacademic.com/one/aera/aera07/

Davis, L. L., & Moyer, E. L. (2015). *Performance level setting technical report.* Partnership for Assessment of Readiness for College and Careers.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*(1), 1–26. DOI: 10.1214/aos/1176344552

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall.

Ferrara, S., Lewis, D., Mercado, R., D'Brot, J., Barth, J., & Egan, K. (2011, April 9-11). *A method for setting benchmarked performance standards: Workshop procedures, panelist judgments, and empirical results* [Paper presentation]. National Council on Measurement in Education Annual Meeting, New Orleans, LA, United States.

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, *33*(3), 613–619. https://doi.org/10.1177/001316447303300309

Kingston, N. M., Kahl, S. R., Sweeney, K. P., & Bay, L. (2001). Setting performance standards using the body of work method. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (1st ed., pp. 219–248). Lawrence Erlbaum.

Kingston N. M., & Tiemann G. C. (2012). Setting performance standards on complex assessments: The body of work method. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 201–224). Routledge.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). Springer.

Lee, W.-C. (2010). Classification consistency and accuracy for complex assessments using item response theory. *Journal of Educational Measurement, 47*(1), 1–17. https://doi.org/10.1111/j.1745-3984.2009.00096.x

Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.

Lohr, S. L. (1999). *Sampling: Design and analysis* (1st ed.). Cengage Learning.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174. https://doi.org/10.1007/BF02296272

O'Malley, K., Keng, L., & Miles, J. (2012). Using validity evidence to set performance standards. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 301–322). Routledge.

Petersen, N. S. (1987, September 25). *DIF procedures for use in statistical analysis.* ETS internal memorandum.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221–262). Macmillan; American Council on Education.

Phillips, G. W. (2012). The benchmark method of standard setting. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 323–346). Routledge.

Plake, B. S., Ferdous, A. A., Impara, J. C., & Buckendahl, C. W. (2005, November). *Setting multiple performance standards using the Yes/No Method: An alternative item mapping procedure* [Paper presentation]. National Council on Measurement in Education Annual Meeting, Montreal, Canada.

Rasch, G. (1966). An individualistic approach to item analysis. In P. Lazarsfeld & N. W. Henry (Eds.), *Readings in mathematical social science* (pp. 89–107). Science Research Associates.

Rudner, L. M. (2000). Computing the expected proportions of misclassified examinees. *Practical Assessment, Research & Evaluation*, *7*. https://doi.org/10.7275/an9m-2035

Rudner, L. M. (2005). Expected classification accuracy. *Practical Assessment, Research & Evaluation*, *10*. https://doi.org/10.7275/56a5-6b14

Schafer, W. D., Wang, J., & Wang, V. (2009). Validity in action: State assessment validity evidence for compliance with NCLB. In R. W. Lissitz (Ed.) *The concept of validity: Revisions, new directions, and applications* (pp. 173–193). Information Age.

Torgerson, W. S. (1958). *Theory and methods of scaling*. Wiley.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, *14*(2), 97–116. https://doi.org/10.1111/j.1745-3984.1977.tb00031.x

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. MESA Press.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Lawrence Erlbaum Associates, Inc.

TECHNICAL DIGEST 2022–2023

Appendix A

Quality Control System

Quality Control

Prior to reporting test results, TEA performs an extensive and comprehensive QC process to verify the quality and accuracy of final reports for Texas assessments.

QC Process

1. Prepare a test design for each test administration. A test design is a set of specific instructions for preparing fictitious student responses that are used as test cases.
    - Check the proposed test design document for the upcoming administration for any design changes that might affect the QC process (e.g., new or revised fields in the online system).
    - Determine whether any new policies have been established since the last test administration that would affect how student responses are edited or how test scores are reported. Establish how these policies affect the QC process and how these changes should be tested with modified or additional test cases.
    - Create a new database of test cases. A new test administration will have most of the same test cases as the previous administration of the same assessment but will include additions or changes necessary to reflect new reporting policies or new conditions that should be tested.
    - Prepare a written test design consisting of coding and gridding instructions for the test contractors, CAI and Pearson.
2. Prepare, enter, and submit online test cases for fictitious students.
    - Create a database of test cases that simulate the exact steps a student and a test administrator would take on a testing day.
    - Verify that the score codes, accommodations, and other test indicators are appropriate for the given administration.
    - Verify that the information on the test ticket matches the online test that is generated upon logging in.
    - Take tests by selecting the answers as outlined in the spreadsheet of test cases.
    - Submit all tests for scoring.
    - Ensure that responses, score codes, accommodations, test information attributes, and demographic information are appropriately recorded and reported.
3. Receive RSSS conversion tables from CAI that show the scale score corresponding to each raw score point for an assessment.
    - Verify, approve, and incorporate the tables into computer programs that produce the student and district/campus files and reports. Refer to the Equating section in Chapter 3, “Standard Technical Processes,” for more information about RSSS

conversion tables.

4. Create a student-level data file with data from the simulated processing of test-case student responses.
    - Verify that the resolved fields are correct in the database of test cases. The resolved fields simulate the changes that would be made in the contractors' editing process if coding errors are made in the online system or if any defaulting conditions apply.
    - Export the student-level data file from the database.
5. Receive from CAI an independently generated student-level data file with student names, demographic data, and scores for the tests administered.
6. Compare TEA student-level data files to CAI's student-level data files.
    - For each record in the data file:
        - compare every variable in CAI's student-level data files with the corresponding variable in TEA student-level data files,
        - investigate and determine the source of and reason for any discrepancies, and
        - make any necessary corrections in accordance with established policies and business rules.
    - Repeat the process by regenerating student-level data files, comparing variables again, and resolving discrepancies until CAI's data files match TEA data files.
7. Produce reports.
    - Receive electronically from CAI all standard reports for three fictitious districts for the tests administered. In addition, receive printed copies of student report cards for spring administrations and confidential student labels for all administrations.
    - TEA and CAI generate reports using independently produced computer programs.
8. Verify reports.
    - Compare TEA reports to CAI's reports to verify that reports contain identical information.
    - Investigate and correct any discrepancies.
9. Approve reports.
    - Verify that all the reports for the fictitious districts are error-free.
    - Notify CAI that production reports can be posted to the Secure File Center and shipped to school districts when their internal quality assurance (QA) process is complete.
10. Verify the Family and Analytic portals by selecting a random sample of student records

for each test case and executing a QA checklist confirming the following functionalities:

- Reports are accessible by portal access code and date of birth.
- Reports populate as expected for each test case spanning the administration type, grade, subjects, and courses.
- Data for the selected student record appear correctly, including but not limited to student identifying information, demographic information, and both current and historical test result information.
- All functionality within the portal performs as expected, including the toggle to the Spanish version.
- PDF reports download and render as expected.
- PDF report data and portal data match in accuracy and consistency.
- Where applicable, item reports populate correctly, verifying the number of correct and incorrect responses, number of items, and percentage of students responding correctly.
- Test cases run correctly in Chrome and Edge.
- Test cases run correctly on mobile devices such as iPads.

11. Verify the Centralized Reporting System by executing a QA checklist confirming the following functionalities:

    - For each program, reports populate as expected at each level (e.g., district/campus, roster, student).
    - All functionality performs as expected, including aggregate summaries and the display of *n*-counts and/or percentages where expected.
    - CSV files download and populate as expected.
    - Users can drill down from summary to roster and from roster to individual student.
    - ID searches return results for the appropriate students and display the correct scores.
    - The menu works as expected.
    - Users can generate and print requested reports.
    - Users can access the correct help guide.
    - All user roles return information as expected.
    - Test cases run correctly in Chrome and Edge.
    - Test cases run correctly on mobile devices such as iPads.

TECHNICAL DIGEST 2022–2023

Appendix B

STAAR Statistical Tables and Figures

STAAR Statistical Tables and Figures

Spring 2023 STAAR Classification Consistency and Accuracy

- Table B.1.1. STAAR Grades 3–8 Assessments
- Table B.1.2. STAAR Spanish Grades 3–5 Assessments
- Table B.1.3. STAAR EOC Assessments

Spring 2023 STAAR Scale Score Correlations

- Table B.2.1. STAAR Grades 3–8 Assessments
- Table B.2.2. STAAR Spanish Grades 3–5 Assessments
- Table B.2.3. STAAR EOC Assessments

Spring 2023 STAAR Conditional Standard Error of Measurement for Scale Scores

- Table B.3.1. STAAR Grades 3–5 Mathematics
- Table B.3.2. STAAR Grades 6–8 Mathematics
- Table B.3.3. STAAR Grades 3–5 RLA
- Table B.3.4. STAAR Grades 6–8 RLA
- Table B.3.5. STAAR Grade 5 and Grade 8 Science
- Table B.3.6. STAAR Grade 8 Social Studies
- Table B.3.7. STAAR Spanish Grades 3–5 Mathematics
- Table B.3.8. STAAR Spanish Grades 3–5 RLA
- Table B.3.9. STAAR Spanish Grade 5 Science
- Table B.3.10. STAAR Algebra I
- Table B.3.11. STAAR English I and English II
- Table B.3.12. STAAR Biology
- Table B.3.13. STAAR U.S. History

Spring 2023 STAAR Mean P-Values

- Table B.4.1. STAAR Grade 3 Total Group
- Table B.4.2. STAAR Grade 3 Female
- Table B.4.3. STAAR Grade 3 Male
- Table B.4.4. STAAR Grade 3 Black or African-American
- Table B.4.5. STAAR Grade 3 Hispanic or Latino
- Table B.4.6. STAAR Grade 3 White
- Table B.4.7. STAAR Grade 3 Students Tested with Accommodations
- Table B.4.8. STAAR Grade 3 Students Tested without Accommodations
- Table B.4.9. STAAR Grade 4 Total Group
- Table B.4.10. STAAR Grade 4 Female

- Table B.4.11. STAAR Grade 4 Male
- Table B.4.12. STAAR Grade 4 Black or African-American
- Table B.4.13. STAAR Grade 4 Hispanic or Latino
- Table B.4.14. STAAR Grade 4 White
- Table B.4.15. STAAR Grade 4 Students Tested with Accommodations
- Table B.4.16. STAAR Grade 4 Students Tested without Accommodations
- Table B.4.17. STAAR Grade 5 Total Group
- Table B.4.18. STAAR Grade 5 Female
- Table B.4.19. STAAR Grade 5 Male
- Table B.4.20. STAAR Grade 5 Black or African-American
- Table B.4.21. STAAR Grade 5 Hispanic or Latino
- Table B.4.22. STAAR Grade 5 White
- Table B.4.23. STAAR Grade 5 Students Tested with Accommodations
- Table B.4.24. STAAR Grade 5 Students Tested without Accommodations
- Table B.4.25. STAAR Grade 6 Total Group
- Table B.4.26. STAAR Grade 6 Female
- Table B.4.27. STAAR Grade 6 Male
- Table B.4.28. STAAR Grade 6 Black or African-American
- Table B.4.29. STAAR Grade 6 Hispanic or Latino
- Table B.4.30. STAAR Grade 6 White
- Table B.4.31. STAAR Grade 6 Students Tested with Accommodations
- Table B.4.32. STAAR Grade 6 Students Tested without Accommodations
- Table B.4.33. STAAR Grade 7 Total Group
- Table B.4.34. STAAR Grade 7 Female
- Table B.4.35. STAAR Grade 7 Male
- Table B.4.36. STAAR Grade 7 Black or African-American
- Table B.4.37. STAAR Grade 7 Hispanic or Latino
- Table B.4.38. STAAR Grade 7 White
- Table B.4.39. STAAR Grade 7 Students Tested with Accommodations
- Table B.4.40. STAAR Grade 7 Students Tested without Accommodations
- Table B.4.41. STAAR Grade 8 Total Group
- Table B.4.42. STAAR Grade 8 Female
- Table B.4.43. STAAR Grade 8 Male
- Table B.4.44. STAAR Grade 8 Black or African-American
- Table B.4.45. STAAR Grade 8 Hispanic or Latino
- Table B.4.46. STAAR Grade 8 White

- Table B.4.47. STAAR Grade 8 Students Tested with Accommodations
- Table B.4.48. STAAR Grade 8 Students Tested without Accommodations
- Table B.4.49. STAAR Spanish Grade 3 Total Group
- Table B.4.50. STAAR Spanish Grade 3 Female
- Table B.4.51. STAAR Spanish Grade 3 Male
- Table B.4.52. STAAR Spanish Grade 3 Black or African-American
- Table B.4.53. STAAR Spanish Grade 3 Hispanic or Latino
- Table B.4.54. STAAR Spanish Grade 3 White
- Table B.4.55. STAAR Spanish Grade 3 Students Tested with Accommodations
- Table B.4.56. STAAR Spanish Grade 3 Students Tested without Accommodations
- Table B.4.57. STAAR Spanish Grade 4 Total Group
- Table B.4.58. STAAR Spanish Grade 4 Female
- Table B.4.59. STAAR Spanish Grade 4 Male
- Table B.4.60. STAAR Spanish Grade 4 Black or African-American
- Table B.4.61. STAAR Spanish Grade 4 Hispanic or Latino
- Table B.4.62. STAAR Spanish Grade 4 White
- Table B.4.63. STAAR Spanish Grade 4 Students Tested with Accommodations
- Table B.4.64. STAAR Spanish Grade 4 Students Tested without Accommodations
- Table B.4.65. STAAR Spanish Grade 5 Total Group
- Table B.4.66. STAAR Spanish Grade 5 Female
- Table B.4.67. STAAR Spanish Grade 5 Male
- Table B.4.68. STAAR Spanish Grade 5 Black or African-American
- Table B.4.69. STAAR Spanish Grade 5 Hispanic or Latino
- Table B.4.70. STAAR Spanish Grade 5 White
- Table B.4.71. STAAR Spanish Grade 5 Students Tested with Accommodations
- Table B.4.72. STAAR Spanish Grade 5 Students Tested without Accommodations
- Table B.4.73. STAAR Algebra I Total Group
- Table B.4.74. STAAR Algebra I Female
- Table B.4.75. STAAR Algebra I Male
- Table B.4.76. STAAR Algebra I Black or African-American
- Table B.4.77. STAAR Algebra I Hispanic or Latino
- Table B.4.78. STAAR Algebra I White
- Table B.4.79. STAAR Algebra I Students Tested with Accommodations
- Table B.4.80. STAAR Algebra I Students Tested without Accommodations
- Table B.4.81. STAAR English I Total Group
- Table B.4.82. STAAR English I Female

- Table B.4.83. STAAR English I Male
- Table B.4.84. STAAR English I Black or African-American
- Table B.4.85. STAAR English I Hispanic or Latino
- Table B.4.86. STAAR English I White
- Table B.4.87. STAAR English I Students Tested with Accommodations
- Table B.4.88. STAAR English I Students Tested without Accommodations
- Table B.4.89. STAAR English II Total Group
- Table B.4.90. STAAR English II Female
- Table B.4.91. STAAR English II Male
- Table B.4.92. STAAR English II Black or African-American
- Table B.4.93. STAAR English II Hispanic or Latino
- Table B.4.94. STAAR English II White
- Table B.4.95. STAAR English II Students Tested with Accommodations
- Table B.4.96. STAAR English II Students Tested without Accommodations
- Table B.4.97. STAAR Biology Total Group
- Table B.4.98. STAAR Biology Female
- Table B.4.99. STAAR Biology Male
- Table B.4.100. STAAR Biology Black or African-American
- Table B.4.101. STAAR Biology Hispanic or Latino
- Table B.4.102. STAAR Biology White
- Table B.4.103. STAAR Biology Students Tested with Accommodations
- Table B.4.104. STAAR Biology Students Tested without Accommodations
- Table B.4.105. STAAR U.S. History Total Group
- Table B.4.106. STAAR U.S. History Female
- Table B.4.107. STAAR U.S. History Male
- Table B.4.108. STAAR U.S. History Black or African-American
- Table B.4.109. STAAR U.S. History Hispanic or Latino
- Table B.4.110. STAAR U.S. History White
- Table B.4.111. STAAR U.S. History Students Tested with Accommodations
- Table B.4.112. STAAR U.S. History Students Tested without Accommodations

Spring 2023 STAAR Progress Measure

Progress measures were not calculated or reported for the 2022–2023 school year.

Spring 2023 STAAR On-Track Results

On-Track results were not calculated or reported for the 2022–2023 school year.

Spring 2023 STAAR Scale Score Descriptive Statistics

- Table B.7.1. STAAR Grades 3–8 Assessments

- Table B.7.2. STAAR Spanish Grades 3–5 Assessments
- Table B.7.3. STAAR EOC Assessments

Spring 2023 STAAR Frequency Distribution of Scale Scores

- Figure B.7.1. STAAR Grade 3 Mathematics
- Figure B.7.2. STAAR Grade 4 Mathematics
- Figure B.7.3. STAAR Grade 5 Mathematics
- Figure B.7.4. STAAR Grade 6 Mathematics
- Figure B.7.5. STAAR Grade 7 Mathematics
- Figure B.7.6. STAAR Grade 8 Mathematics
- Figure B.7.7. STAAR Grade 3 RLA
- Figure B.7.8. STAAR Grade 4 RLA
- Figure B.7.9. STAAR Grade 5 RLA
- Figure B.7.10. STAAR Grade 6 RLA
- Figure B.7.11. STAAR Grade 7 RLA
- Figure B.7.12. STAAR Grade 8 RLA
- Figure B.7.13. STAAR Grade 5 Science
- Figure B.7.14. STAAR Grade 8 Science
- Figure B.7.15. STAAR Grade 8 Social Studies
- Figure B.7.16. STAAR Spanish Grade 3 Mathematics
- Figure B.7.17. STAAR Spanish Grade 4 Mathematics
- Figure B.7.18. STAAR Spanish Grade 5 Mathematics
- Figure B.7.19. STAAR Spanish Grade 3 RLA
- Figure B.7.20. STAAR Spanish Grade 4 RLA
- Figure B.7.21. STAAR Spanish Grade 5 RLA
- Figure B.7.22. STAAR Spanish Grade 5 Science
- Figure B.7.23. STAAR Algebra I
- Figure B.7.24. STAAR English I
- Figure B.7.25. STAAR English II
- Figure B.7.26. STAAR Biology
- Figure B.7.27. STAAR U.S. History

Spring 2023 STAAR English Learner Performance Measure

- Table B.8.1. STAAR Grades 3–8 Mathematics
- Table B.8.2. STAAR Grades 3–8 RLA
- Table B.8.3. STAAR Grade 5 and Grade 8 Science
- Table B.8.4. STAAR Grade 8 Social Studies
- Table B.8.5. STAAR EOC Assessments

Spring 2023 STAAR Classification Consistency and Accuracy

Table B.1.1. Spring 2023 STAAR Grades 3–8 Assessments
Classification Consistency and Accuracy

	Decision Consistency				Decision Accuracy			
Grade	**Mathematics**	**RLA**	**Science**	**Social Studies**	**Mathematics**	**RLA**	**Science**	**Social Studies**
3	66.3	70.0			75.3	78.4		
4	69.6	69.8			77.9	78.2		
5	69.1	68.6	65.7		77.7	77.1	74.4	
6	68.8	69.7			77.4	78.1		
7	70.6	69.7			78.6	78.0		
8	70.1	69.3	68.2	70.7	78.4	77.8	76.8	78.4

Table B.1.2. Spring 2023 STAAR Spanish Grades 3–5 Assessments
Classification Consistency and Accuracy

	Decision Consistency			Decision Accuracy		
Grade	**Mathematics**	**RLA**	**Science**	**Mathematics**	**RLA**	**Science**
3	66.6	72.9		75.5	79.9	
4	72.5	71.8		80.0	79.0	
5	70.1	71.2	75.3	78.5	79.0	81.9

Notes for both tables:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

Table B.1.3. Spring 2023 STAAR EOC Assessments
Classification Consistency and Accuracy

Test	Decision Consistency	Decision Accuracy
Algebra I	71.6	79.5
English I	76.0	82.6
English II	76.4	82.8
Biology	72.1	79.9
U.S. History	74.8	81.9

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

Spring 2023 STAAR
Scale Score
Correlations

Table B.2.1. Spring 2023 STAAR Grades 3–8 Assessments
Scale Score Correlations within Grade

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	354,663	0.74
4	Mathematics	RLA	361,988	0.75
5	Mathematics	RLA	368,369	0.72
	Mathematics	Science	373,554	0.75
	RLA	Science	371,039	0.73
6	Mathematics	RLA	380,287	0.74
7	Mathematics	RLA	323,997	0.73
8	Mathematics	RLA	290,288	0.70
	Mathematics	Science	296,002	0.77
	Mathematics	Social Studies	292,013	0.71
	RLA	Science	393,859	0.76
	RLA	Social Studies	406,777	0.78
	Science	Social Studies	396,906	0.80

Table B.2.2. Spring 2023 STAAR Spanish Grades 3–5 Assessments Scale Score Correlations within Grade

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	15,916	0.67
4	Mathematics	RLA	10,976	0.65
5	Mathematics	RLA	8,109	0.66
	Mathematics	Science	7,018	0.62
	RLA	Science	9,202	0.65

Table B.2.3. Spring 2023 STAAR EOC Assessments
Scale Score Correlations

STAAR	STAAR	N	Correlation
Algebra I	English I	317,118	0.69
Algebra I	English II	50,795	0.46
Algebra I	Biology	275,141	0.73
Algebra I	U.S. History	20,570	0.44
English I	English II	87,016	0.67
English I	Biology	389,471	0.81
English I	U.S. History	35,893	0.58
English II	Biology	85,167	0.76
English II	U.S. History	83,119	0.76
Biology	U.S. History	26,044	0.74

Spring 2023 STAAR Conditional Standard Error of Measurement for Scale Scores

Table B.3.1. Spring 2023 STAAR Grades 3–5 Mathematics Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	860		910		1000	
1	955	131	1009	134	1094	132
2	1047	94	1108	97	1189	95
3	1103	78	1168	81	1247	79
4	1144	68	1213	72	1289	70
5	1176	62	1248	65	1323	63
6	1204	58	1279	60	1351	59
7	1229	55	1305	56	1376	55
8	1251	53	1328	53	1398	52
9	1272	51	1349	51	1418	50
10	1291	49	1368	49	1437	48
11	1309	48	1385	47	1454	47
12	1327	47	1402	46	1470	45
13	1344	46	1418	45	1486	44
14	*1360	46	1433	44	1500	43
15	1376	45	1448	43	*1515	43
16	1392	45	*1462	43	1528	42
17	1407	45	1476	42	1542	42
18	1423	45	1489	42	1555	41
19	1439	45	1503	42	1568	41
20	1454	45	1516	42	1581	41
21	**1471	46	1530	42	1594	41
22	1487	46	1544	42	1607	41
23	1504	47	**1557	43	1620	41
24	1521	48	1572	43	**1634	42
25	1539	49	1586	44	1647	42
26	1558	50	1601	45	1661	43
27	1579	52	1617	46	1675	44
28	***1600	54	1634	47	1690	44
29	1624	56	1651	48	1706	45
30	1649	59	1670	50	1722	46
31	1678	63	***1690	52	1739	48
32	1710	68	1711	54	1757	49

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1748	74	1735	57	***1776	51
34	1796	84	1762	61	1797	53
35	1860	100	1792	65	1819	56
36	1963	137	1828	72	1845	59
37	2070		1872	81	1873	64
38			1932	97	1907	70
39			2029	134	1950	79
40			2130		2008	95
41					2103	132
42					2200	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.2. Spring 2023 STAAR Grades 6–8 Mathematics
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1070		1150		1240	
1	1165	135	1246	132	1332	131
2	1264	98	1339	94	1425	94
3	1325	82	1395	78	1480	77
4	1370	72	1436	68	1520	68
5	1407	66	1468	62	1552	61
6	1438	61	1495	57	1579	56
7	1465	57	1519	53	1602	53
8	1489	55	1539	51	1622	50
9	1511	52	1558	49	1640	48
10	1531	50	1576	47	1657	46
11	1550	49	1592	45	1673	45
12	1568	47	1608	44	1688	43
13	1584	46	1622	43	1702	42
14	1601	45	1637	43	1716	42
15	*1616	45	1650	42	1729	41
16	1631	44	1664	42	1741	40
17	1646	43	1677	41	*1754	40
18	1660	43	1690	41	1766	39
19	1675	43	*1703	41	1778	39
20	1689	43	1716	41	1789	39
21	1703	43	1728	41	1801	39
22	1717	43	1741	41	1813	39
23	1731	43	1754	41	1824	39
24	**1745	43	1767	41	1836	39
25	1759	43	1780	41	1848	39
26	1774	44	**1793	42	**1859	39
27	1788	44	1806	42	1871	39
28	1804	45	1820	42	1883	40
29	1819	46	1834	43	1896	40
30	1836	47	1848	43	1908	41
31	1853	48	1863	44	1921	41
32	1870	49	1878	45	1934	42

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	***1889	50	1894	46	1948	43
34	1910	52	1910	47	1962	43
35	1932	54	1927	48	1977	44
36	1955	57	1946	49	1993	46
37	1982	61	***1965	51	***2009	47
38	2013	65	1986	53	2027	49
39	2049	72	2009	56	2046	50
40	2093	81	2035	60	2066	53
41	2153	97	2064	64	2089	56
42	2251	134	2099	70	2114	59
43	2350		2142	80	2143	64
44			2200	96	2177	70
45			2296	133	2220	80
46			2400		2278	96
47					2374	133
48					2470	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.3. Spring 2023 STAAR Grades 3–5 RLA
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	720		820		830	
1	829	147	934	146	925	146
2	935	106	1038	105	1030	105
3	1000	88	1102	87	1093	87
4	1047	78	1148	77	1140	77
5	1085	71	1186	70	1177	70
6	1117	65	1217	65	1209	65
7	1145	61	1245	61	1236	61
8	1170	58	1269	58	1261	58
9	1193	56	1292	56	1284	56
10	1213	54	1312	54	1304	54
11	1233	52	1332	52	1324	52
12	1251	51	1350	50	1342	50
13	1269	49	1367	49	1359	49
14	1285	48	1383	48	1376	48
15	1301	47	1399	47	1391	47
16	1316	46	*1414	46	1406	46
17	1331	45	1429	45	1421	45
18	*1345	45	1443	45	1435	45
19	1359	44	1457	44	1449	44
20	1372	43	1470	43	1462	43
21	1385	43	1483	43	*1475	43
22	1397	42	1495	42	1487	42
23	1409	41	1507	41	1499	42
24	1421	41	1519	41	1511	41
25	1433	41	1530	40	1523	41
26	1444	40	1541	39	1534	41
27	1456	40	**1552	39	1546	41
28	**1467	40	1562	39	1557	41
29	1478	41	1573	39	1569	41
30	1490	41	1583	39	1580	41
31	1502	41	1593	39	**1592	41
32	1514	42	1604	39	1604	42

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1526	43	1615	40	1617	43
34	1539	44	1626	41	1629	43
35	1553	44	1638	41	1643	44
36	1567	45	1650	42	1656	45
37	1581	46	***1663	44	1670	45
38	***1596	47	1676	45	1685	46
39	1612	48	1691	46	***1700	47
40	1628	48	1706	48	1716	48
41	1644	49	1723	49	1732	50
42	1662	51	1740	51	1750	51
43	1680	52	1759	54	1769	53
44	1700	55	1780	56	1790	56
45	1722	58	1804	60	1812	59
46	1748	62	1830	64	1838	63
47	1777	68	1861	69	1867	68
48	1813	76	1897	76	1903	75
49	1859	87	1942	86	1948	86
50	1921	104	2004	104	2010	104
51	2022	144	2107	145	2113	146
52	2120		2210		2220	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.4. Spring 2023 STAAR Grades 6–8 RLA
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	880		890		980	
1	1016	146	983	147	1084	146
2	1121	105	1089	106	1189	105
3	1184	87	1154	89	1252	87
4	1231	77	1203	78	1298	77
5	1268	70	1241	71	1335	70
6	1299	64	1274	66	1366	64
7	1326	60	1303	62	1393	61
8	1350	57	1329	59	1417	57
9	1372	55	1352	57	1439	55
10	1392	52	1373	54	1459	53
11	1410	51	1393	52	1478	51
12	1428	49	1412	51	1495	49
13	1444	47	1429	49	1512	48
14	1459	46	1446	48	1527	46
15	1473	45	1461	47	1541	45
16	1487	44	1476	46	1555	44
17	1500	43	1490	45	1568	43
18	1512	42	1503	44	1580	41
19	1524	41	1516	43	*1592	40
20	*1535	40	1529	42	1603	39
21	1546	39	1541	41	1614	39
22	1557	39	1552	40	1624	38
23	1567	38	*1564	40	1634	37
24	1577	38	1575	39	1643	37
25	1587	37	1585	39	1652	36
26	1596	37	1596	39	1662	36
27	1606	37	1606	39	1671	36
28	1615	37	1617	39	1680	36
29	1625	37	1627	39	1689	36
30	**1634	37	1637	39	**1698	36
31	1644	37	1648	39	1707	37
32	1653	37	1658	39	1717	37

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1663	37	**1669	39	1727	37
34	1673	38	1679	39	1736	38
35	1683	38	1690	39	1747	38
36	1693	39	1701	40	1757	39
37	1703	39	1712	40	1768	40
38	1714	40	1723	40	1779	41
39	1726	41	1734	41	1791	41
40	1737	41	1746	41	***1803	42
41	***1749	42	1758	42	1816	43
42	1762	43	***1771	43	1829	45
43	1776	45	1784	44	1843	46
44	1790	46	1798	46	1858	47
45	1805	48	1814	47	1875	49
46	1822	50	1830	49	1892	51
47	1840	52	1847	52	1911	53
48	1860	55	1867	54	1931	56
49	1882	58	1889	58	1954	59
50	1907	62	1913	62	1980	63
51	1936	67	1942	67	2010	68
52	1971	75	1977	74	2046	75
53	2015	86	2021	85	2090	86
54	2077	104	2081	103	2152	104
55	2179	145	2184	145	2256	145
56	2280		2290		2360	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.5. Spring 2023 STAAR Grade 5 and Grade 8 Science Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1140		1000	
1	1532	559	1458	639
2	1924	398	1913	458
3	2158	330	2186	379
4	2330	291	2386	333
5	2470	267	2544	301
6	2590	251	2677	279
7	2699	240	2793	261
8	2798	231	2895	247
9	2890	223	2988	236
10	2977	217	3073	227
11	3060	211	3152	220
12	3138	206	3226	213
13	3212	202	3297	208
14	3284	198	3364	203
15	3353	194	3428	199
16	3420	192	3490	196
17	3486	190	*3550	193
18	*3550	188	3609	191
19	3613	187	3666	190
20	3677	187	3723	188
21	3740	187	3779	187
22	3803	188	3834	187
23	3867	190	3890	186
24	3933	192	3945	187
25	**4000	195	**4000	187
26	4069	198	4056	188
27	4141	202	4112	189
28	4216	207	4169	190
29	4296	213	4227	192
30	***4380	220	4287	195
31	4471	229	4348	198
32	4570	240	4411	202

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	4680	254	4477	206
34	4804	273	4546	211
35	4951	299	***4619	218
36	5131	338	4697	225
37	5376	406	4780	234
38	5778	564	4872	245
39	6200		4972	259
40			5086	277
41			5217	300
42			5375	331
43			5573	378
44			5845	458
45			6300	639
46			6800	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.6. Spring 2023 STAAR Grade 8 Social Studies
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Grade 8
Raw	**SS**	**CSEM**
0	1050	
1	1489	579
2	1900	414
3	2146	342
4	2325	300
5	2466	271
6	2584	250
7	2686	234
8	2777	221
9	2858	211
10	2933	202
11	3002	195
12	3067	190
13	3128	185
14	3186	181
15	3242	177
16	3297	175
17	3349	172
18	3401	170
19	3451	169
20	3501	168
21	*3550	167
22	3599	167
23	3648	167
24	3697	167
25	3746	168
26	3795	168
27	3845	170
28	3896	171
29	3947	173
30	**4000	174
31	4054	177
32	4109	179

(Continued)

		Grade 8
Raw	**SS**	**CSEM**
33	4167	182
34	4226	186
35	4288	190
36	***4352	194
37	4420	200
38	4492	206
39	4569	213
40	4651	221
41	4740	231
42	4839	244
43	4950	259
44	5077	280
45	5227	308
46	5415	350
47	5670	420
48	6090	583
49	6550	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.7. Spring 2023 STAAR Spanish Grades 3–5 Mathematics Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	860		910		1000	
1	955	131	1009	134	1094	132
2	1047	94	1108	97	1189	95
3	1103	78	1168	81	1247	79
4	1144	68	1213	72	1289	70
5	1176	62	1248	65	1323	63
6	1204	58	1279	60	1351	59
7	1229	55	1305	56	1376	55
8	1251	53	1328	53	1398	52
9	1272	51	1349	51	1418	50
10	1291	49	1368	49	1437	48
11	1309	48	1385	47	1454	47
12	1327	47	1402	46	1470	45
13	1344	46	1418	45	1486	44
14	*1360	46	1433	44	1500	43
15	1376	45	1448	43	*1515	43
16	1392	45	*1462	43	1528	42
17	1407	45	1476	42	1542	42
18	1423	45	1489	42	1555	41
19	1439	45	1503	42	1568	41
20	1454	45	1516	42	1581	41
21	**1471	46	1530	42	1594	41
22	1487	46	1544	42	1607	41
23	1504	47	**1557	43	1620	41
24	1521	48	1572	43	**1634	42
25	1539	49	1586	44	1647	42
26	1558	50	1601	45	1661	43
27	1579	52	1617	46	1675	44
28	***1600	54	1634	47	1690	44
29	1624	56	1651	48	1706	45
30	1649	59	1670	50	1722	46
31	1678	63	***1690	52	1739	48
32	1710	68	1711	54	1757	49

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1748	74	1735	57	***1776	51
34	1796	84	1762	61	1797	53
35	1860	100	1792	65	1819	56
36	1963	137	1828	72	1845	59
37	2070		1872	81	1873	64
38			1932	97	1907	70
39			2029	134	1950	79
40			2130		2008	95
41					2103	132
42					2200	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.8. Spring 2023 STAAR Spanish Grades 3–5 RLA
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	600		680		720	
1	712	156	795	156	828	156
2	823	112	906	112	939	112
3	891	93	974	93	1006	93
4	941	82	1023	82	1056	82
5	981	75	1062	74	1095	74
6	1014	69	1095	69	1128	69
7	1044	65	1124	64	1157	65
8	1070	62	1150	61	1183	61
9	1094	59	1173	58	1207	59
10	1117	57	1194	56	1229	57
11	1137	56	1214	54	1249	55
12	1157	54	1232	52	1268	53
13	1176	53	1249	51	1286	52
14	1194	52	1266	49	1303	51
15	1211	51	1281	48	1319	49
16	1227	50	1296	47	1335	48
17	1244	49	1310	46	1350	48
18	1259	49	1323	45	1364	47
19	1274	48	1336	44	1379	46
20	1289	47	1349	44	1392	45
21	1304	47	1361	43	1405	45
22	*1318	46	1373	42	1418	44
23	1332	46	1385	42	*1431	44
24	1345	45	1396	42	1444	44
25	1359	45	*1408	42	1456	43
26	1372	44	1419	42	1468	43
27	1385	44	1430	42	1481	43
28	1397	44	1442	42	1493	43
29	1410	44	1453	42	1505	44
30	1422	44	1465	42	1518	44
31	1435	44	1476	43	1530	44
32	**1447	44	**1488	43	1543	45

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1460	44	1501	44	**1556	45
34	1473	45	1513	44	1570	46
35	1487	46	1526	45	1584	47
36	1500	46	1539	45	1599	48
37	***1515	47	1553	46	1614	48
38	1529	48	1567	47	1629	49
39	1545	49	***1581	48	1645	50
40	1561	50	1596	49	***1662	51
41	1578	51	1613	51	1680	53
42	1595	53	1630	53	1698	54
43	1615	55	1649	55	1718	56
44	1635	58	1669	58	1740	59
45	1659	61	1693	61	1763	62
46	1685	66	1719	66	1790	66
47	1715	71	1750	71	1821	71
48	1752	79	1787	79	1857	79
49	1798	90	1833	91	1904	90
50	1862	109	1898	110	1968	110
51	1968	153	2006	154	2076	154
52	2070		2110		2180	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.9. Spring 2023 STAAR Spanish Grade 5 Science
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Grade 5
Raw	**SS**	**CSEM**
0	1140	
1	1532	559
2	1924	398
3	2158	330
4	2330	291
5	2470	267
6	2590	251
7	2699	240
8	2798	231
9	2890	223
10	2977	217
11	3060	211
12	3138	206
13	3212	202
14	3284	198
15	3353	194
16	3420	192
17	3486	190
18	*3550	188
19	3613	187
20	3677	187
21	3740	187
22	3803	188
23	3867	190
24	3933	192
25	**4000	195
26	4069	198
27	4141	202
28	4216	207
29	4296	213
30	***4380	220
31	4471	229
32	4570	240

(Continued)

		Grade 5
Raw	**SS**	**CSEM**
33	4680	254
34	4804	273
35	4951	299
36	5131	338
37	5376	406
38	5778	564
39	6200	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.10. Spring 2023 STAAR EOC Assessments Mathematics
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Algebra I
Raw	**SS**	**CSEM**
0	1500	
1	1844	472
2	2185	340
3	2392	282
4	2543	248
5	2663	225
6	2764	208
7	2852	195
8	2930	184
9	3000	176
10	3065	169
11	3124	163
12	3181	158
13	3234	154
14	3284	151
15	3332	148
16	3379	145
17	3423	142
18	3467	140
19	3489	138
20	*3550	137
21	3590	135
22	3630	134
23	3668	133
24	3706	132
25	3744	131
26	3781	131
27	3818	130
28	3855	130
29	3891	129
30	3927	129
31	3964	129
32	**4000	129

(Continued)

	Algebra I	
Raw	**SS**	**CSEM**
33	4036	130
34	4073	130
35	4110	131
36	4147	132
37	4185	133
38	4224	134
39	4263	135
40	4303	137
41	***4345	139
42	4388	141
43	4432	144
44	4478	147
45	4526	151
46	4577	155
47	4631	160
48	4688	165
49	4750	172
50	4817	180
51	4891	189
52	4973	201
53	5066	215
54	5175	233
55	5305	258
56	5469	294
57	5694	354
58	6061	487
59	6430	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.11. Spring 2023 STAAR EOC Assessments RLA
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1750		1650	
1	2073	435	1994	451
2	2382	311	2315	323
3	2567	258	2509	268
4	2702	226	2650	235
5	2810	204	2762	213
6	2899	189	2856	197
7	2977	177	2938	184
8	3046	167	3010	175
9	3108	160	3076	167
10	3165	153	3136	160
11	3218	148	3191	154
12	3267	143	3242	149
13	3313	139	3291	144
14	3357	135	3336	140
15	3398	131	3380	137
16	3437	128	3421	133
17	3474	125	3460	130
18	3510	122	3497	128
19	3544	120	3533	125
20	3577	117	3568	122
21	3608	115	3601	120
22	3638	113	3632	118
23	3667	111	3663	115
24	3695	109	3692	113
25	3723	107	3721	111
26	3749	106	3748	110
27	*3775	105	*3775	108
28	3800	104	3801	107
29	3825	103	3827	106
30	3850	103	3852	105
31	3875	103	3876	105
32	3899	103	3901	104

(Continued)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	3924	103	3925	104
34	3949	104	3950	105
35	3974	104	3975	105
36	**4000	105	**4000	106
37	4026	106	4026	107
38	4052	107	4052	108
39	4079	108	4078	110
40	4107	109	4106	111
41	4135	111	4134	113
42	4164	112	4163	114
43	4193	113	4193	116
44	4224	115	4224	118
45	4255	117	4255	120
46	4287	119	4288	122
47	4321	121	4323	125
48	4355	123	4358	127
49	4392	126	4396	130
50	4430	129	4435	134
51	4469	133	4476	138
52	4512	137	4520	142
53	4557	142	4567	147
54	***4606	147	4618	153
55	4659	154	4673	160
56	4717	162	***4734	169
57	4781	172	4802	179
58	4854	184	4879	192
59	4940	200	4969	208
60	5042	221	5076	231
61	5172	253	5213	264
62	5353	308	5402	320
63	5657	432	5719	449
64	6000		6050	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.12. Spring 2023 STAAR EOC Assessments Science
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Biology
Raw	**SS**	**CSEM**
0	1900	
1	2208	443
2	2524	318
3	2715	264
4	2855	232
5	2966	210
6	3060	195
7	3142	183
8	3215	174
9	3281	166
10	3342	160
11	3398	155
12	3451	150
13	3502	146
14	*3550	143
15	3596	140
16	3641	138
17	3684	136
18	3726	134
19	3767	133
20	3807	132
21	3846	131
22	3885	130
23	3924	129
24	3962	129
25	**4000	129
26	4038	129
27	4076	129
28	4114	129
29	4153	130
30	4192	131
31	4231	131
32	4271	133

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	4312	134
34	4353	135
35	4396	137
36	4440	139
37	4484	141
38	***4531	143
39	4579	146
40	4629	149
41	4681	153
42	4736	157
43	4794	162
44	4856	167
45	4923	175
46	4997	184
47	5079	196
48	5174	211
49	5286	233
50	5427	265
51	5620	320
52	5939	445
53	6260	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.13. Spring 2023 STAAR EOC Assessments Social Studies Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	U.S. History	
Raw	**SS**	**CSEM**
0	1420	
1	1769	494
2	2120	354
3	2330	292
4	2484	256
5	2605	232
6	2707	214
7	2795	200
8	2873	190
9	2943	181
10	3007	173
11	3066	167
12	3122	162
13	3174	157
14	3223	153
15	3269	149
16	3314	146
17	3357	143
18	3398	140
19	3438	138
20	3476	136
21	3486	134
22	*3550	132
23	3586	131
24	3620	130
25	3654	128
26	3688	127
27	3721	126
28	3753	125
29	3785	124
30	3816	124
31	3848	123
32	3879	123

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
33	3909	122
34	3940	122
35	3970	121
36	**4000	121
37	4030	121
38	4060	121
39	4090	121
40	4120	121
41	4149	121
42	4179	121
43	4209	121
44	4239	121
45	4270	122
46	4300	122
47	4331	122
48	4361	123
49	4393	124
50	***4424	124
51	4456	125
52	4488	126
53	4521	127
54	4555	128
55	4589	130
56	4624	131
57	4660	133
58	4696	135
59	4734	137
60	4773	139
61	4813	142
62	4855	145
63	4899	148
64	4945	152

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
65	4993	156
66	5045	161
67	5099	166
68	5158	173
69	5222	180
70	5292	189
71	5370	200
72	5458	214
73	5560	232
74	5682	257
75	5836	294
76	6048	355
77	6401	495
78	6750	

Notes: * Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Spring 2023 STAAR
Mean P-Values

Table B.4.1. Spring 2023 STAAR Grade 3 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	369,987	19.40	7.75	0.88	2.68	50.16
	Numerical Representations and Relationships	10	369,987	5.28	2.69	0.73	1.40	53.90
	Computations and Algebraic Relationships	15	369,987	7.28	3.54	0.77	1.71	46.05
	Geometry and Measurement	7	369,987	3.45	1.68	0.52	1.17	46.02
	Data Analysis and Personal Financial Literacy	5	369,987	3.38	1.18	0.31	0.97	63.68
RLA	OVERALL TEST	52	356,509	27.14	10.75	0.92	3.03	57.92
	Reading	26	356,509	15.86	5.39	0.82	2.26	61.37
	Writing	26	356,509	11.29	6.04	0.89	2.00	53.05

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.2. Spring 2023 STAAR Grade 3 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	181,001	18.77	7.48	0.87	2.69	48.34
	Numerical Representations and Relationships	10	181,001	5.03	2.62	0.71	1.41	51.44
	Computations and Algebraic Relationships	15	181,001	7.03	3.44	0.75	1.71	44.35
	Geometry and Measurement	7	181,001	3.36	1.64	0.49	1.17	44.54
	Data Analysis and Personal Financial Literacy	5	181,001	3.35	1.15	0.29	0.97	62.60
RLA	OVERALL TEST	52	174,079	28.19	10.72	0.92	3.01	59.79
	Reading	26	174,079	16.28	5.31	0.82	2.24	62.99
	Writing	26	174,079	11.91	6.10	0.89	2.00	55.28

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.3. Spring 2023 STAAR Grade 3 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	187,930	20.02	7.96	0.89	2.67	51.94
	Numerical Representations and Relationships	10	187,930	5.53	2.73	0.74	1.39	56.29
	Computations and Algebraic Relationships	15	187,930	7.53	3.61	0.78	1.70	47.72
	Geometry and Measurement	7	187,930	3.55	1.71	0.54	1.16	47.48
	Data Analysis and Personal Financial Literacy	5	187,930	3.41	1.20	0.33	0.98	64.73
RLA	OVERALL TEST	52	181,413	26.16	10.69	0.92	3.03	56.15
	Reading	26	181,413	15.45	5.44	0.83	2.28	59.83
	Writing	26	181,413	10.70	5.92	0.89	1.99	50.94

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.4. Spring 2023 STAAR Grade 3 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	47,854	16.35	7.13	0.85	2.73	41.95
	Numerical Representations and Relationships	10	47,854	4.31	2.51	0.68	1.41	44.27
	Computations and Algebraic Relationships	15	47,854	6.09	3.23	0.72	1.72	38.07
	Geometry and Measurement	7	47,854	2.91	1.59	0.45	1.18	38.75
	Data Analysis and Personal Financial Literacy	5	47,854	3.04	1.26	0.31	1.04	56.92
RLA	OVERALL TEST	52	47,899	23.96	10.47	0.92	3.04	51.78
	Reading	26	47,899	14.43	5.47	0.82	2.31	55.80
	Writing	26	47,899	9.52	5.70	0.88	1.97	46.10

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.5. Spring 2023 STAAR Grade 3 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	184,900	18.18	7.25	0.86	2.71	46.92
	Numerical Representations and Relationships	10	184,900	4.89	2.56	0.70	1.41	50.18
	Computations and Algebraic Relationships	15	184,900	6.76	3.37	0.74	1.73	42.79
	Geometry and Measurement	7	184,900	3.26	1.62	0.47	1.17	43.12
	Data Analysis and Personal Financial Literacy	5	184,900	3.27	1.15	0.27	0.98	61.31
RLA	OVERALL TEST	52	171,364	25.42	10.43	0.91	3.07	54.36
	Reading	26	171,364	15.06	5.35	0.82	2.30	58.21
	Writing	26	171,364	10.36	5.79	0.88	2.01	48.93

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.6. Spring 2023 STAAR Grade 3 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	100,098	21.74	7.62	0.88	2.63	56.43
	Numerical Representations and Relationships	10	100,098	6.07	2.66	0.74	1.37	61.48
	Computations and Algebraic Relationships	15	100,098	8.23	3.50	0.77	1.68	52.18
	Geometry and Measurement	7	100,098	3.81	1.66	0.52	1.16	51.17
	Data Analysis and Personal Financial Literacy	5	100,098	3.64	1.12	0.30	0.94	68.77
RLA	OVERALL TEST	52	100,139	30.04	10.20	0.92	2.97	63.96
	Reading	26	100,139	17.22	5.01	0.81	2.19	66.82
	Writing	26	100,139	12.82	5.89	0.89	1.98	59.93

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.7. Spring 2023 STAAR Grade 3 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	119,924	14.89	6.23	0.81	2.73	38.33
	Numerical Representations and Relationships	10	119,924	3.88	2.26	0.61	1.41	40.10
	Computations and Algebraic Relationships	15	119,924	5.34	2.90	0.64	1.74	34.11
	Geometry and Measurement	7	119,924	2.71	1.49	0.38	1.17	35.77
	Data Analysis and Personal Financial Literacy	5	119,924	2.97	1.20	0.27	1.02	55.02
RLA	OVERALL TEST	52	110,972	19.52	8.65	0.87	3.08	42.66
	Reading	26	110,972	12.23	4.90	0.77	2.36	47.09
	Writing	26	110,972	7.29	4.54	0.81	1.96	36.40

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.8. Spring 2023 STAAR Grade 3 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	250,063	21.56	7.47	0.87	2.65	55.84
	Numerical Representations and Relationships	10	250,063	5.96	2.62	0.72	1.38	60.51
	Computations and Algebraic Relationships	15	250,063	8.21	3.43	0.76	1.68	51.78
	Geometry and Measurement	7	250,063	3.81	1.65	0.51	1.16	50.94
	Data Analysis and Personal Financial Literacy	5	250,063	3.58	1.11	0.27	0.95	67.83
RLA	OVERALL TEST	52	245,537	30.59	9.79	0.91	2.98	64.81
	Reading	26	245,537	17.49	4.77	0.79	2.19	67.82
	Writing	26	245,537	13.10	5.75	0.88	2.00	60.57

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.9. Spring 2023 STAAR Grade 4 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	373,970	21.82	9.15	0.91	2.80	52.60
	Numerical Representations and Relationships	12	373,970	6.90	3.18	0.78	1.50	54.79
	Computations and Algebraic Relationships	13	373,970	6.60	3.43	0.77	1.65	48.52
	Geometry and Measurement	11	373,970	5.69	2.66	0.70	1.47	49.92
	Data Analysis and Personal Financial Literacy	4	373,970	2.62	1.11	0.48	0.80	65.52
RLA	OVERALL TEST	52	365,003	26.01	11.12	0.92	3.10	54.93
	Reading	26	365,003	14.27	5.58	0.83	2.28	55.04
	Writing	26	365,003	11.74	6.21	0.89	2.09	54.78

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.10. Spring 2023 STAAR Grade 4 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	183,694	21.10	8.90	0.90	2.82	50.85
	Numerical Representations and Relationships	12	183,694	6.58	3.09	0.76	1.53	51.92
	Computations and Algebraic Relationships	13	183,694	6.44	3.40	0.76	1.66	47.16
	Geometry and Measurement	11	183,694	5.46	2.59	0.68	1.47	47.90
	Data Analysis and Personal Financial Literacy	4	183,694	2.62	1.10	0.46	0.81	65.49
RLA	OVERALL TEST	52	178,902	26.88	11.06	0.92	3.11	56.46
	Reading	26	178,902	14.57	5.53	0.83	2.27	56.25
	Writing	26	178,902	12.30	6.21	0.88	2.11	56.77

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.11. Spring 2023 STAAR Grade 4 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	189,258	22.52	9.34	0.91	2.77	54.34
	Numerical Representations and Relationships	12	189,258	7.22	3.24	0.79	1.48	57.61
	Computations and Algebraic Relationships	13	189,258	6.77	3.45	0.78	1.63	49.87
	Geometry and Measurement	11	189,258	5.91	2.71	0.71	1.46	51.91
	Data Analysis and Personal Financial Literacy	4	189,258	2.62	1.12	0.50	0.79	65.56
RLA	OVERALL TEST	52	185,098	25.19	11.11	0.92	3.10	53.48
	Reading	26	185,098	13.98	5.61	0.83	2.29	53.90
	Writing	26	185,098	11.21	6.15	0.89	2.07	52.89

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.12. Spring 2023 STAAR Grade 4 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	47,866	18.11	8.40	0.88	2.86	43.35
	Numerical Representations and Relationships	12	47,866	5.79	2.97	0.73	1.54	45.26
	Computations and Algebraic Relationships	13	47,866	5.33	3.30	0.75	1.66	38.84
	Geometry and Measurement	11	47,866	4.71	2.40	0.62	1.49	41.10
	Data Analysis and Personal Financial Literacy	4	47,866	2.28	1.10	0.42	0.84	56.99
RLA	OVERALL TEST	52	47,886	22.47	10.34	0.91	3.11	48.43
	Reading	26	47,886	12.74	5.32	0.81	2.33	49.19
	Writing	26	47,886	9.73	5.70	0.87	2.05	47.35

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.13. Spring 2023 STAAR Grade 4 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	189,989	20.55	8.71	0.89	2.84	49.38
	Numerical Representations and Relationships	12	189,989	6.54	3.08	0.75	1.53	51.72
	Computations and Algebraic Relationships	13	189,989	6.17	3.31	0.75	1.66	45.03
	Geometry and Measurement	11	189,989	5.33	2.53	0.66	1.48	46.70
	Data Analysis and Personal Financial Literacy	4	189,989	2.51	1.10	0.45	0.82	62.71
RLA	OVERALL TEST	52	180,166	24.06	10.52	0.91	3.14	51.03
	Reading	26	180,166	13.26	5.32	0.81	2.32	51.16
	Writing	26	180,166	10.79	5.90	0.87	2.10	50.85

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.14. Spring 2023 STAAR Grade 4 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	99,424	24.44	8.95	0.91	2.74	59.20
	Numerical Representations and Relationships	12	99,424	7.68	3.14	0.78	1.46	61.22
	Computations and Algebraic Relationships	13	99,424	7.50	3.30	0.75	1.64	55.57
	Geometry and Measurement	11	99,424	6.39	2.65	0.70	1.45	56.28
	Data Analysis and Personal Financial Literacy	4	99,424	2.87	1.06	0.48	0.76	71.73
RLA	OVERALL TEST	52	99,890	29.30	10.85	0.92	3.06	61.56
	Reading	26	99,890	15.92	5.43	0.83	2.22	61.41
	Writing	26	99,890	13.38	6.12	0.88	2.09	61.78

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.15. Spring 2023 STAAR Grade 4 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	118,965	16.20	7.38	0.85	2.88	38.84
	Numerical Representations and Relationships	12	118,965	5.16	2.68	0.66	1.56	40.38
	Computations and Algebraic Relationships	13	118,965	4.57	2.97	0.69	1.67	33.31
	Geometry and Measurement	11	118,965	4.38	2.18	0.52	1.50	38.17
	Data Analysis and Personal Financial Literacy	4	118,965	2.10	1.08	0.38	0.85	52.46
RLA	OVERALL TEST	52	112,805	17.83	7.94	0.85	3.10	39.27
	Reading	26	112,805	10.34	4.28	0.70	2.36	39.94
	Writing	26	112,805	7.49	4.48	0.80	2.00	38.32

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.16. Spring 2023 STAAR Grade 4 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	255,005	24.43	8.71	0.90	2.75	59.03
	Numerical Representations and Relationships	12	255,005	7.72	3.07	0.77	1.47	61.51
	Computations and Algebraic Relationships	13	255,005	7.55	3.21	0.74	1.63	55.61
	Geometry and Measurement	11	255,005	6.30	2.65	0.70	1.45	55.40
	Data Analysis and Personal Financial Literacy	4	255,005	2.86	1.04	0.45	0.77	71.60
RLA	OVERALL TEST	52	252,198	29.67	10.36	0.91	3.09	61.94
	Reading	26	252,198	16.03	5.18	0.81	2.23	61.80
	Writing	26	252,198	13.65	5.92	0.87	2.12	62.14

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.17. Spring 2023 STAAR Grade 5 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	378,653	23.61	9.29	0.90	2.88	56.18
	Numerical Representations and Relationships	9	378,653	5.62	2.11	0.64	1.26	58.74
	Computations and Algebraic Relationships	20	378,653	11.18	4.93	0.84	2.00	58.04
	Geometry and Measurement	9	378,653	4.40	2.35	0.63	1.43	48.48
	Data Analysis and Personal Financial Literacy	4	378,653	2.40	1.15	0.45	0.85	60.11
RLA	OVERALL TEST	52	372,636	30.96	10.49	0.92	2.98	64.39
	Reading	26	372,636	16.77	5.01	0.82	2.11	65.57
	Writing	26	372,636	14.19	6.14	0.89	2.08	62.71
Science	OVERALL TEST	39	378,696	20.93	7.48	0.87	2.74	52.98
	Matter and Energy	6	378,696	3.16	1.46	0.48	1.05	51.28
	Force, Motion, and Energy	9	378,696	4.64	2.32	0.68	1.32	53.34
	Earth and Space	11	378,696	4.88	2.39	0.56	1.59	45.89
	Organisms and Environments	13	378,696	8.25	2.80	0.73	1.47	59.34

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.18. Spring 2023 STAAR Grade 5 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	185,535	23.27	9.03	0.90	2.89	55.45
	Numerical Representations and Relationships	9	185,535	5.58	2.08	0.63	1.26	58.44
	Computations and Algebraic Relationships	20	185,535	11.01	4.79	0.82	2.01	57.50
	Geometry and Measurement	9	185,535	4.32	2.31	0.62	1.43	47.19
	Data Analysis and Personal Financial Literacy	4	185,535	2.36	1.13	0.43	0.85	59.07
RLA	OVERALL TEST	52	182,255	32.27	10.22	0.92	2.96	66.49
	Reading	26	182,255	17.19	4.87	0.82	2.09	67.32
	Writing	26	182,255	15.08	6.01	0.88	2.08	65.31
Science	OVERALL TEST	39	185,318	20.23	7.28	0.86	2.76	51.00
	Matter and Energy	6	185,318	3.04	1.46	0.48	1.05	49.10
	Force, Motion, and Energy	9	185,318	4.38	2.26	0.65	1.34	50.06
	Earth and Space	11	185,318	4.70	2.34	0.54	1.58	43.95
	Organisms and Environments	13	185,318	8.11	2.76	0.71	1.48	58.23

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.19. Spring 2023 STAAR Grade 5 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	192,208	23.95	9.52	0.91	2.87	56.92
	Numerical Representations and Relationships	9	192,208	5.66	2.13	0.66	1.25	59.07
	Computations and Algebraic Relationships	20	192,208	11.35	5.05	0.85	1.99	58.60
	Geometry and Measurement	9	192,208	4.48	2.38	0.64	1.42	49.75
	Data Analysis and Personal Financial Literacy	4	192,208	2.45	1.17	0.47	0.85	61.14
RLA	OVERALL TEST	52	189,477	29.71	10.59	0.92	2.99	62.39
	Reading	26	189,477	16.37	5.11	0.83	2.13	63.92
	Writing	26	189,477	13.34	6.15	0.89	2.08	60.23
Science	OVERALL TEST	39	192,466	21.61	7.61	0.87	2.72	54.92
	Matter and Energy	6	192,466	3.28	1.46	0.49	1.04	53.40
	Force, Motion, and Energy	9	192,466	4.90	2.34	0.70	1.29	56.52
	Earth and Space	11	192,466	5.05	2.44	0.57	1.59	47.77
	Organisms and Environments	13	192,466	8.38	2.83	0.74	1.46	60.44

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.20. Spring 2023 STAAR Grade 5 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	48,326	19.58	8.54	0.88	2.94	46.83
	Numerical Representations and Relationships	9	48,326	4.86	2.12	0.63	1.30	50.09
	Computations and Algebraic Relationships	20	48,326	9.13	4.55	0.80	2.04	48.24
	Geometry and Measurement	9	48,326	3.60	2.12	0.54	1.44	39.73
	Data Analysis and Personal Financial Literacy	4	48,326	2.00	1.15	0.39	0.90	50.00
RLA	OVERALL TEST	52	48,491	27.63	10.51	0.92	3.04	58.25
	Reading	26	48,491	15.26	5.18	0.82	2.20	59.42
	Writing	26	48,491	12.37	6.02	0.88	2.09	56.60
Science	OVERALL TEST	39	48,468	17.72	6.75	0.83	2.78	44.41
	Matter and Energy	6	48,468	2.74	1.37	0.41	1.05	43.63
	Force, Motion, and Energy	9	48,468	3.69	2.15	0.61	1.34	42.68
	Earth and Space	11	48,468	4.02	2.16	0.47	1.58	37.63
	Organisms and Environments	13	48,468	7.27	2.72	0.68	1.53	51.42

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.21. Spring 2023 STAAR Grade 5 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	193,929	22.29	8.76	0.89	2.92	53.15
	Numerical Representations and Relationships	9	193,929	5.42	2.06	0.62	1.28	56.44
	Computations and Algebraic Relationships	20	193,929	10.53	4.68	0.81	2.03	55.05
	Geometry and Measurement	9	193,929	4.07	2.23	0.58	1.44	44.91
	Data Analysis and Personal Financial Literacy	4	193,929	2.27	1.13	0.40	0.87	56.76
RLA	OVERALL TEST	52	186,834	29.38	10.23	0.91	3.04	61.39
	Reading	26	186,834	16.06	4.98	0.81	2.16	62.74
	Writing	26	186,834	13.32	5.95	0.87	2.12	59.49
Science	OVERALL TEST	39	192,912	19.64	7.10	0.85	2.77	49.58
	Matter and Energy	6	192,912	2.98	1.44	0.46	1.06	47.94
	Force, Motion, and Energy	9	192,912	4.34	2.22	0.64	1.33	50.25
	Earth and Space	11	192,912	4.53	2.28	0.52	1.58	42.67
	Organisms and Environments	13	192,912	7.78	2.74	0.70	1.50	55.56

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.22. Spring 2023 STAAR Grade 5 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	100,245	26.33	9.03	0.90	2.84	62.47
	Numerical Representations and Relationships	9	100,245	6.08	2.01	0.62	1.23	63.92
	Computations and Algebraic Relationships	20	100,245	12.53	4.84	0.83	1.97	64.36
	Geometry and Measurement	9	100,245	5.01	2.33	0.63	1.42	54.96
	Data Analysis and Personal Financial Literacy	4	100,245	2.71	1.08	0.44	0.80	67.86
RLA	OVERALL TEST	52	100,688	33.83	9.78	0.91	2.89	69.96
	Reading	26	100,688	18.17	4.57	0.80	2.02	71.22
	Writing	26	100,688	15.67	5.90	0.88	2.04	68.18
Science	OVERALL TEST	39	100,663	23.78	7.16	0.86	2.69	60.48
	Matter and Energy	6	100,663	3.53	1.42	0.47	1.03	58.16
	Force, Motion, and Energy	9	100,663	5.39	2.25	0.67	1.29	61.24
	Earth and Space	11	100,663	5.65	2.39	0.55	1.60	53.01
	Organisms and Environments	13	100,663	9.21	2.60	0.71	1.39	67.15

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.23. Spring 2023 STAAR Grade 5 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	117,526	17.94	7.50	0.85	2.95	43.25
	Numerical Representations and Relationships	9	117,526	4.47	1.99	0.56	1.32	45.56
	Computations and Algebraic Relationships	20	117,526	8.46	4.08	0.75	2.05	45.51
	Geometry and Measurement	9	117,526	3.17	1.90	0.45	1.42	35.55
	Data Analysis and Personal Financial Literacy	4	117,526	1.85	1.11	0.32	0.92	46.16
RLA	OVERALL TEST	52	110,713	22.53	9.08	0.88	3.12	48.50
	Reading	26	110,713	13.04	4.77	0.77	2.28	50.54
	Writing	26	110,713	9.50	5.12	0.83	2.10	45.61
Science	OVERALL TEST	39	114,750	16.43	6.43	0.81	2.80	41.16
	Matter and Energy	6	114,750	2.55	1.37	0.39	1.07	39.79
	Force, Motion, and Energy	9	114,750	3.50	2.05	0.58	1.32	40.98
	Earth and Space	11	114,750	3.76	2.08	0.43	1.57	35.41
	Organisms and Environments	13	114,750	6.62	2.66	0.65	1.58	46.61

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.24. Spring 2023 STAAR Grade 5 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	261,127	26.16	8.88	0.90	2.84	62.00
	Numerical Representations and Relationships	9	261,127	6.14	1.95	0.60	1.23	64.68
	Computations and Algebraic Relationships	20	261,127	12.41	4.79	0.83	1.97	63.68
	Geometry and Measurement	9	261,127	4.95	2.32	0.63	1.41	54.31
	Data Analysis and Personal Financial Literacy	4	261,127	2.66	1.08	0.42	0.82	66.38
RLA	OVERALL TEST	52	261,923	34.52	8.89	0.89	2.91	71.10
	Reading	26	261,923	18.34	4.21	0.77	2.04	71.93
	Writing	26	261,923	16.17	5.42	0.86	2.06	69.94
Science	OVERALL TEST	39	263,946	22.88	7.05	0.85	2.71	58.12
	Matter and Energy	6	263,946	3.43	1.42	0.47	1.03	56.28
	Force, Motion, and Energy	9	263,946	5.14	2.25	0.66	1.31	58.71
	Earth and Space	11	263,946	5.36	2.36	0.54	1.59	50.44
	Organisms and Environments	13	263,946	8.95	2.56	0.70	1.41	64.87

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.25. Spring 2023 STAAR Grade 6 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	384,765	21.45	8.87	0.90	2.86	49.39
	Numerical Representations and Relationships	10	384,765	5.21	2.51	0.75	1.27	48.98
	Computations and Algebraic Relationships	17	384,765	8.07	3.78	0.77	1.81	47.45
	Geometry and Measurement	7	384,765	3.06	1.62	0.50	1.14	44.34
	Data Analysis and Personal Financial Literacy	9	384,765	5.11	2.30	0.63	1.40	58.38
RLA	OVERALL TEST	56	391,376	29.57	11.80	0.92	3.25	54.11
	Reading	28	391,376	15.47	5.64	0.83	2.34	54.50
	Writing	28	391,376	14.10	6.83	0.89	2.24	53.57

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.26. Spring 2023 STAAR Grade 6 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	188,857	21.02	8.64	0.89	2.86	48.47
	Numerical Representations and Relationships	10	188,857	5.08	2.47	0.74	1.26	47.64
	Computations and Algebraic Relationships	17	188,857	7.80	3.66	0.76	1.81	46.14
	Geometry and Measurement	7	188,857	3.04	1.60	0.49	1.14	43.84
	Data Analysis and Personal Financial Literacy	9	188,857	5.09	2.28	0.62	1.40	58.39
RLA	OVERALL TEST	56	191,639	30.91	11.60	0.92	3.25	55.91
	Reading	28	191,639	15.96	5.52	0.82	2.33	56.24
	Writing	28	191,639	14.95	6.74	0.89	2.25	55.46

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.27. Spring 2023 STAAR Grade 6 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	194,836	21.89	9.06	0.90	2.85	50.32
	Numerical Representations and Relationships	10	194,836	5.34	2.54	0.75	1.27	50.32
	Computations and Algebraic Relationships	17	194,836	8.33	3.87	0.78	1.81	48.74
	Geometry and Measurement	7	194,836	3.09	1.64	0.52	1.14	44.84
	Data Analysis and Personal Financial Literacy	9	194,836	5.13	2.32	0.64	1.40	58.41
RLA	OVERALL TEST	56	198,680	28.30	11.86	0.93	3.24	52.40
	Reading	28	198,680	15.01	5.72	0.83	2.34	52.86
	Writing	28	198,680	13.29	6.81	0.89	2.23	51.77

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.28. Spring 2023 STAAR Grade 6 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	49,173	18.33	7.69	0.86	2.88	42.07
	Numerical Representations and Relationships	10	49,173	4.46	2.28	0.70	1.26	41.09
	Computations and Algebraic Relationships	17	49,173	6.88	3.26	0.69	1.82	40.41
	Geometry and Measurement	7	49,173	2.59	1.44	0.39	1.13	37.38
	Data Analysis and Personal Financial Literacy	9	49,173	4.40	2.23	0.58	1.45	50.75
RLA	OVERALL TEST	56	49,382	26.39	11.13	0.91	3.28	48.99
	Reading	28	49,382	14.13	5.39	0.81	2.38	49.59
	Writing	28	49,382	12.27	6.45	0.88	2.25	48.17

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.29. Spring 2023 STAAR Grade 6 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	203,019	19.66	8.00	0.87	2.90	45.14
	Numerical Representations and Relationships	10	203,019	4.78	2.35	0.71	1.28	44.39
	Computations and Algebraic Relationships	17	203,019	7.36	3.40	0.71	1.83	43.30
	Geometry and Measurement	7	203,019	2.80	1.48	0.41	1.14	40.61
	Data Analysis and Personal Financial Literacy	9	203,019	4.73	2.24	0.59	1.44	53.81
RLA	OVERALL TEST	56	205,499	27.23	11.21	0.91	3.29	50.11
	Reading	28	205,499	14.38	5.42	0.81	2.37	50.53
	Writing	28	205,499	12.84	6.50	0.88	2.26	49.53

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.30. Spring 2023 STAAR Grade 6 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	98,213	24.76	8.85	0.90	2.81	57.25
	Numerical Representations and Relationships	10	98,213	6.00	2.49	0.74	1.27	57.27
	Computations and Algebraic Relationships	17	98,213	9.32	3.87	0.79	1.78	54.81
	Geometry and Measurement	7	98,213	3.55	1.67	0.53	1.15	51.30
	Data Analysis and Personal Financial Literacy	9	98,213	5.89	2.13	0.61	1.33	67.53
RLA	OVERALL TEST	56	100,160	33.59	11.23	0.92	3.19	60.98
	Reading	28	100,160	17.36	5.38	0.82	2.28	61.52
	Writing	28	100,160	16.22	6.55	0.89	2.22	60.24

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.31. Spring 2023 STAAR Grade 6 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	106,713	16.03	6.41	0.80	2.89	36.54
	Numerical Representations and Relationships	10	106,713	3.89	1.96	0.59	1.25	35.22
	Computations and Algebraic Relationships	17	106,713	6.05	2.78	0.57	1.83	35.27
	Geometry and Measurement	7	106,713	2.39	1.31	0.27	1.12	34.50
	Data Analysis and Personal Financial Literacy	9	106,713	3.71	2.07	0.49	1.48	42.50
RLA	OVERALL TEST	56	104,485	20.14	8.78	0.86	3.27	38.34
	Reading	28	104,485	11.36	4.58	0.73	2.39	39.67
	Writing	28	104,485	8.78	5.06	0.81	2.21	36.52

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.32. Spring 2023 STAAR Grade 6 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	278,052	23.53	8.80	0.90	2.84	54.33
	Numerical Representations and Relationships	10	278,052	5.72	2.51	0.74	1.27	54.25
	Computations and Algebraic Relationships	17	278,052	8.84	3.82	0.78	1.80	52.12
	Geometry and Measurement	7	278,052	3.32	1.65	0.52	1.14	48.11
	Data Analysis and Personal Financial Literacy	9	278,052	5.64	2.15	0.60	1.37	64.47
RLA	OVERALL TEST	56	286,891	33.00	10.86	0.91	3.22	59.85
	Reading	28	286,891	16.96	5.24	0.81	2.31	59.91
	Writing	28	286,891	16.04	6.34	0.88	2.24	59.77

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.33. Spring 2023 STAAR Grade 7 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	331,698	22.75	9.21	0.89	3.03	47.96
	Probablility and Numerical Representations	8	331,698	4.75	2.13	0.63	1.29	55.93
	Computations and Algebraic Relationships	17	331,698	8.09	3.72	0.76	1.83	47.83
	Geometry and Measurement	12	331,698	4.97	2.60	0.67	1.48	41.63
	Data Analysis and Personal Financial Literacy	9	331,698	4.94	2.20	0.60	1.39	51.93
RLA	OVERALL TEST	56	400,416	32.40	11.74	0.93	3.17	60.57
	Reading	28	400,416	16.90	5.49	0.82	2.34	60.77
	Writing	28	400,416	15.49	6.93	0.91	2.13	60.31

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.34. Spring 2023 STAAR Grade 7 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	162,392	22.46	8.96	0.89	3.03	47.18
	Probablility and Numerical Representations	8	162,392	4.74	2.07	0.61	1.29	55.35
	Computations and Algebraic Relationships	17	162,392	7.87	3.59	0.74	1.83	46.39
	Geometry and Measurement	12	162,392	4.96	2.59	0.67	1.48	41.52
	Data Analysis and Personal Financial Literacy	9	162,392	4.88	2.18	0.59	1.39	51.34
RLA	OVERALL TEST	56	195,365	33.77	11.41	0.92	3.16	62.49
	Reading	28	195,365	17.19	5.35	0.81	2.33	61.91
	Writing	28	195,365	16.58	6.73	0.90	2.12	63.28

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.35. Spring 2023 STAAR Grade 7 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	168,426	23.05	9.43	0.90	3.03	48.75
	Probablility and Numerical Representations	8	168,426	4.76	2.19	0.65	1.29	56.52
	Computations and Algebraic Relationships	17	168,426	8.30	3.83	0.77	1.82	49.24
	Geometry and Measurement	12	168,426	4.99	2.60	0.68	1.48	41.78
	Data Analysis and Personal Financial Literacy	9	168,426	5.00	2.22	0.62	1.38	52.55
RLA	OVERALL TEST	56	204,024	31.09	11.89	0.93	3.18	58.77
	Reading	28	204,024	16.64	5.61	0.83	2.35	59.70
	Writing	28	204,024	14.46	6.95	0.91	2.13	57.50

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.36. Spring 2023 STAAR Grade 7 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	43,307	19.68	7.85	0.85	3.07	40.99
	Probablility and Numerical Representations	8	43,307	4.22	2.03	0.56	1.34	48.81
	Computations and Algebraic Relationships	17	43,307	6.92	3.25	0.68	1.84	40.73
	Geometry and Measurement	12	43,307	4.20	2.25	0.57	1.48	34.88
	Data Analysis and Personal Financial Literacy	9	43,307	4.34	1.99	0.49	1.42	45.01
RLA	OVERALL TEST	56	50,218	29.27	11.20	0.92	3.22	55.61
	Reading	28	50,218	15.71	5.34	0.80	2.38	56.53
	Writing	28	50,218	13.55	6.59	0.89	2.16	54.35

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.37. Spring 2023 STAAR Grade 7 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	181,030	21.10	8.42	0.87	3.08	44.35
	Probablility and Numerical Representations	8	181,030	4.38	2.09	0.59	1.34	51.22
	Computations and Algebraic Relationships	17	181,030	7.52	3.46	0.72	1.84	44.42
	Geometry and Measurement	12	181,030	4.61	2.37	0.60	1.50	38.49
	Data Analysis and Personal Financial Literacy	9	181,030	4.60	2.10	0.54	1.42	48.07
RLA	OVERALL TEST	56	211,278	30.07	11.41	0.92	3.23	56.57
	Reading	28	211,278	15.89	5.41	0.81	2.38	57.12
	Writing	28	211,278	14.17	6.71	0.90	2.17	55.81

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.38. Spring 2023 STAAR Grade 7 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	80,968	26.18	9.16	0.90	2.95	55.48
	Probablility and Numerical Representations	8	80,968	5.56	1.95	0.62	1.20	66.03
	Computations and Algebraic Relationships	17	80,968	9.27	3.74	0.77	1.80	55.03
	Geometry and Measurement	12	80,968	5.70	2.68	0.70	1.48	47.88
	Data Analysis and Personal Financial Literacy	9	80,968	5.65	2.17	0.63	1.33	59.96
RLA	OVERALL TEST	56	102,824	36.51	10.76	0.92	3.08	67.62
	Reading	28	102,824	18.67	5.04	0.80	2.27	67.10
	Writing	28	102,824	17.85	6.45	0.90	2.06	68.33

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.39. Spring 2023 STAAR Grade 7 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	96,147	17.40	6.65	0.78	3.10	36.24
	Probablility and Numerical Representations	8	96,147	3.51	1.94	0.49	1.38	40.51
	Computations and Algebraic Relationships	17	96,147	6.16	2.83	0.58	1.84	36.32
	Geometry and Measurement	12	96,147	3.88	1.97	0.43	1.49	32.07
	Data Analysis and Personal Financial Literacy	9	96,147	3.84	1.82	0.37	1.44	39.41
RLA	OVERALL TEST	56	99,444	22.46	9.60	0.88	3.27	43.64
	Reading	28	99,444	12.70	4.94	0.76	2.44	45.80
	Writing	28	99,444	9.75	5.50	0.84	2.17	40.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.40. Spring 2023 STAAR Grade 7 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	235,551	24.94	9.21	0.89	2.99	52.74
	Probablility and Numerical Representations	8	235,551	5.26	2.00	0.61	1.25	62.22
	Computations and Algebraic Relationships	17	235,551	8.87	3.76	0.77	1.81	52.52
	Geometry and Measurement	12	235,551	5.42	2.69	0.70	1.48	45.53
	Data Analysis and Personal Financial Literacy	9	235,551	5.39	2.18	0.61	1.36	57.05
RLA	OVERALL TEST	56	300,972	35.68	10.46	0.91	3.12	66.17
	Reading	28	300,972	18.29	4.93	0.78	2.30	65.71
	Writing	28	300,972	17.39	6.28	0.89	2.10	66.79

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.41. Spring 2023 STAAR Grade 8 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	364,110	24.71	10.30	0.91	3.10	49.96
	Numerical Representations and Relationships	5	364,110	3.23	1.36	0.46	1.00	62.05
	Computations and Algebraic Relationships	18	364,110	8.45	4.17	0.79	1.90	45.44
	Geometry and Measurement	18	364,110	9.30	4.02	0.79	1.86	49.41
	Data Analysis and Personal Financial Literacy	7	364,110	3.73	2.11	0.67	1.20	55.64
RLA	OVERALL TEST	56	410,468	30.86	11.62	0.93	3.16	55.64
	Reading	28	410,468	16.41	5.72	0.84	2.25	57.74
	Writing	28	410,468	14.45	6.61	0.89	2.21	52.77
Science	OVERALL TEST	46	407,845	23.65	9.64	0.89	3.14	51.92
	Matter and Energy	14	407,845	7.73	3.58	0.80	1.59	54.64
	Force, Motion, and Energy	10	407,845	4.25	2.17	0.65	1.28	43.46
	Earth and Space	11	407,845	6.21	2.81	0.70	1.55	55.55
	Organisms and Environments	11	407,845	5.46	2.69	0.65	1.60	53.32
Social Studies	OVERALL TEST	49	414,692	24.15	9.90	0.89	3.22	49.91
	History	18	414,692	8.64	4.03	0.80	1.82	49.63
	Geography and Culture	11	414,692	6.35	2.82	0.69	1.58	57.44
	Government and Citizenship	13	414,692	5.78	2.73	0.62	1.68	45.61
	Economics, Science, Technology, and Society	7	414,692	3.38	1.83	0.69	1.02	45.84

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.42. Spring 2023 STAAR Grade 8 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	176,349	24.91	9.98	0.90	3.11	50.34
	Numerical Representations and Relationships	5	176,349	3.24	1.33	0.43	1.01	62.29
	Computations and Algebraic Relationships	18	176,349	8.51	4.09	0.78	1.90	45.64
	Geometry and Measurement	18	176,349	9.46	3.90	0.77	1.87	50.23
	Data Analysis and Personal Financial Literacy	7	176,349	3.71	2.06	0.65	1.21	55.23
RLA	OVERALL TEST	56	199,809	32.62	11.28	0.92	3.15	58.03
	Reading	28	199,809	16.96	5.58	0.84	2.23	59.56
	Writing	28	199,809	15.67	6.38	0.88	2.21	55.95
Science	OVERALL TEST	46	198,883	23.21	9.34	0.89	3.16	50.83
	Matter and Energy	14	198,883	7.75	3.50	0.79	1.60	54.67
	Force, Motion, and Energy	10	198,883	4.13	2.09	0.63	1.28	42.14
	Earth and Space	11	198,883	6.03	2.78	0.68	1.57	54.10
	Organisms and Environments	11	198,883	5.30	2.63	0.63	1.61	51.61
Social Studies	OVERALL TEST	49	202,501	23.88	9.59	0.89	3.22	49.17
	History	18	202,501	8.39	3.91	0.78	1.82	48.02
	Geography and Culture	11	202,501	6.51	2.77	0.68	1.57	58.49
	Government and Citizenship	13	202,501	5.67	2.67	0.60	1.68	44.90
	Economics, Science, Technology, and Society	7	202,501	3.31	1.78	0.67	1.02	44.67

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.43. Spring 2023 STAAR Grade 8 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	186,774	24.52	10.58	0.91	3.09	49.62
	Numerical Representations and Relationships	5	186,774	3.22	1.39	0.49	1.00	61.83
	Computations and Algebraic Relationships	18	186,774	8.40	4.24	0.80	1.89	45.26
	Geometry and Measurement	18	186,774	9.15	4.13	0.80	1.86	48.64
	Data Analysis and Personal Financial Literacy	7	186,774	3.75	2.15	0.70	1.19	56.05
RLA	OVERALL TEST	56	209,652	29.19	11.69	0.93	3.17	53.38
	Reading	28	209,652	15.90	5.80	0.85	2.27	56.04
	Writing	28	209,652	13.29	6.62	0.89	2.20	49.76
Science	OVERALL TEST	46	207,970	24.09	9.90	0.90	3.13	52.99
	Matter and Energy	14	207,970	7.72	3.66	0.81	1.59	54.64
	Force, Motion, and Energy	10	207,970	4.36	2.24	0.67	1.28	44.73
	Earth and Space	11	207,970	6.39	2.84	0.71	1.53	56.97
	Organisms and Environments	11	207,970	5.63	2.73	0.66	1.58	54.97
Social Studies	OVERALL TEST	49	211,185	24.41	10.18	0.90	3.22	50.65
	History	18	211,185	8.87	4.12	0.81	1.81	51.18
	Geography and Culture	11	211,185	6.21	2.86	0.69	1.59	56.47
	Government and Citizenship	13	211,185	5.89	2.78	0.63	1.68	46.32
	Economics, Science, Technology, and Society	7	211,185	3.44	1.87	0.71	1.02	46.99

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.44. Spring 2023 STAAR Grade 8 Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	49,109	21.14	8.95	0.88	3.16	42.52
	Numerical Representations and Relationships	5	49,109	2.89	1.38	0.41	1.06	55.33
	Computations and Algebraic Relationships	18	49,109	7.14	3.65	0.73	1.90	37.86
	Geometry and Measurement	18	49,109	8.11	3.55	0.71	1.92	42.69
	Data Analysis and Personal Financial Literacy	7	49,109	3.01	1.94	0.60	1.23	45.68
RLA	OVERALL TEST	56	52,131	27.53	11.08	0.92	3.20	50.45
	Reading	28	52,131	14.98	5.59	0.83	2.30	52.68
	Writing	28	52,131	12.56	6.26	0.87	2.22	47.41
Science	OVERALL TEST	46	52,181	20.27	8.62	0.86	3.18	44.73
	Matter and Energy	14	52,181	6.67	3.37	0.77	1.62	47.11
	Force, Motion, and Energy	10	52,181	3.66	1.93	0.55	1.29	37.92
	Earth and Space	11	52,181	5.28	2.66	0.64	1.60	47.20
	Organisms and Environments	11	52,181	4.66	2.44	0.57	1.59	46.12
Social Studies	OVERALL TEST	49	52,568	21.13	9.12	0.87	3.25	43.63
	History	18	52,568	7.48	3.69	0.75	1.84	43.15
	Geography and Culture	11	52,568	5.61	2.78	0.66	1.62	50.34
	Government and Citizenship	13	52,568	5.22	2.51	0.55	1.69	41.17
	Economics, Science, Technology, and Society	7	52,568	2.82	1.75	0.66	1.02	38.53

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.45. Spring 2023 STAAR Grade 8 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	194,119	22.91	9.48	0.89	3.15	46.29
	Numerical Representations and Relationships	5	194,119	3.08	1.36	0.42	1.03	59.00
	Computations and Algebraic Relationships	18	194,119	7.83	3.88	0.76	1.91	42.01
	Geometry and Measurement	18	194,119	8.60	3.71	0.73	1.91	45.51
	Data Analysis and Personal Financial Literacy	7	194,119	3.41	2.03	0.64	1.22	51.45
RLA	OVERALL TEST	56	218,677	28.69	11.16	0.92	3.20	51.96
	Reading	28	218,677	15.34	5.58	0.83	2.29	53.88
	Writing	28	218,677	13.35	6.33	0.88	2.23	49.34
Science	OVERALL TEST	46	216,652	21.71	8.97	0.87	3.18	47.72
	Matter and Energy	14	216,652	7.12	3.45	0.78	1.62	50.24
	Force, Motion, and Energy	10	216,652	3.86	1.98	0.58	1.28	39.68
	Earth and Space	11	216,652	5.72	2.74	0.66	1.59	51.15
	Organisms and Environments	11	216,652	5.00	2.55	0.60	1.60	49.13
Social Studies	OVERALL TEST	49	221,167	22.21	9.20	0.87	3.25	45.93
	History	18	221,167	7.96	3.76	0.76	1.83	45.85
	Geography and Culture	11	221,167	5.89	2.79	0.67	1.61	53.32
	Government and Citizenship	13	221,167	5.28	2.51	0.55	1.69	41.51
	Economics, Science, Technology, and Society	7	221,167	3.07	1.77	0.66	1.03	41.70

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.46. Spring 2023 STAAR Grade 8 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	90,619	28.30	10.21	0.91	3.02	57.26
	Numerical Representations and Relationships	5	90,619	3.56	1.28	0.46	0.94	68.56
	Computations and Algebraic Relationships	18	90,619	9.67	4.17	0.79	1.89	52.17
	Geometry and Measurement	18	90,619	10.61	4.06	0.81	1.79	56.62
	Data Analysis and Personal Financial Literacy	7	90,619	4.46	2.03	0.67	1.18	65.51
RLA	OVERALL TEST	56	104,025	34.93	10.89	0.92	3.10	62.47
	Reading	28	104,025	18.45	5.23	0.82	2.19	65.09
	Writing	28	104,025	16.48	6.38	0.88	2.18	58.88
Science	OVERALL TEST	46	103,721	27.50	9.27	0.89	3.09	60.18
	Matter and Energy	14	103,721	8.91	3.41	0.79	1.56	63.06
	Force, Motion, and Energy	10	103,721	4.94	2.23	0.68	1.27	50.19
	Earth and Space	11	103,721	7.27	2.61	0.68	1.47	64.91
	Organisms and Environments	11	103,721	6.39	2.65	0.63	1.60	61.71
Social Studies	OVERALL TEST	49	104,821	27.72	9.58	0.89	3.17	57.34
	History	18	104,821	9.85	4.02	0.80	1.79	56.52
	Geography and Culture	11	104,821	7.27	2.58	0.65	1.53	65.81
	Government and Citizenship	13	104,821	6.60	2.78	0.64	1.67	52.40
	Economics, Science, Technology, and Society	7	104,821	4.01	1.73	0.66	1.00	54.22

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.47. Spring 2023 STAAR Grade 8 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	94,514	18.44	7.94	0.84	3.16	37.24
	Numerical Representations and Relationships	5	94,514	2.64	1.36	0.36	1.09	50.70
	Computations and Algebraic Relationships	18	94,514	6.16	3.24	0.67	1.87	32.78
	Geometry and Measurement	18	94,514	7.09	3.20	0.63	1.95	37.41
	Data Analysis and Personal Financial Literacy	7	94,514	2.55	1.83	0.58	1.19	39.30
RLA	OVERALL TEST	56	95,835	20.76	9.07	0.87	3.21	39.12
	Reading	28	95,835	11.67	4.90	0.77	2.34	40.98
	Writing	28	95,835	9.09	5.07	0.81	2.18	36.59
Science	OVERALL TEST	46	93,347	17.10	7.37	0.82	3.16	38.18
	Matter and Energy	14	93,347	5.44	2.94	0.69	1.63	38.59
	Force, Motion, and Energy	10	93,347	3.15	1.66	0.41	1.28	32.95
	Earth and Space	11	93,347	4.59	2.53	0.60	1.60	41.48
	Organisms and Environments	11	93,347	3.92	2.18	0.49	1.55	39.27
Social Studies	OVERALL TEST	49	94,139	17.37	7.65	0.82	3.24	36.01
	History	18	94,139	6.27	3.09	0.65	1.84	36.32
	Geography and Culture	11	94,139	4.48	2.64	0.62	1.63	40.76
	Government and Citizenship	13	94,139	4.34	2.14	0.40	1.67	33.67
	Economics, Science, Technology, and Society	7	94,139	2.28	1.60	0.60	1.02	31.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.48. Spring 2023 STAAR Grade 8 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	269,596	26.91	10.12	0.91	3.07	54.43
	Numerical Representations and Relationships	5	269,596	3.44	1.31	0.46	0.96	66.03
	Computations and Algebraic Relationships	18	269,596	9.26	4.16	0.79	1.89	49.88
	Geometry and Measurement	18	269,596	10.08	4.00	0.79	1.83	53.62
	Data Analysis and Personal Financial Literacy	7	269,596	4.14	2.04	0.65	1.20	61.37
RLA	OVERALL TEST	56	314,633	33.93	10.52	0.91	3.14	60.67
	Reading	28	314,633	17.86	5.14	0.81	2.22	62.85
	Writing	28	314,633	16.08	6.15	0.87	2.20	57.69
Science	OVERALL TEST	46	314,498	25.60	9.38	0.89	3.13	55.99
	Matter and Energy	14	314,498	8.41	3.47	0.79	1.58	59.40
	Force, Motion, and Energy	10	314,498	4.57	2.19	0.66	1.27	46.58
	Earth and Space	11	314,498	6.69	2.71	0.69	1.52	59.72
	Organisms and Environments	11	314,498	5.92	2.66	0.63	1.61	57.49
Social Studies	OVERALL TEST	49	320,553	26.14	9.59	0.89	3.20	53.99
	History	18	320,553	9.33	4.01	0.80	1.80	53.53
	Geography and Culture	11	320,553	6.90	2.63	0.65	1.56	62.34
	Government and Citizenship	13	320,553	6.20	2.73	0.62	1.68	49.12
	Economics, Science, Technology, and Society	7	320,553	3.70	1.77	0.67	1.02	50.02

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.49. Spring 2023 STAAR Spanish Grade 3 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	16,435	15.89	6.45	0.82	2.75	40.89
	Numerical Representations and Relationships	10	16,435	4.13	2.33	0.63	1.42	42.54
	Computations and Algebraic Relationships	15	16,435	5.71	2.97	0.66	1.74	36.45
	Geometry and Measurement	7	16,435	2.97	1.57	0.44	1.18	39.26
	Data Analysis and Personal Financial Literacy	5	16,435	3.09	1.21	0.23	1.07	56.94
RLA	OVERALL TEST	52	30,152	23.94	10.11	0.91	3.00	52.00
	Reading	26	30,152	13.65	5.16	0.81	2.26	52.63
	Writing	26	30,152	10.29	5.63	0.88	1.94	51.12

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.50. Spring 2023 STAAR Spanish Grade 3 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	8,436	15.42	6.14	0.80	2.74	39.47
	Numerical Representations and Relationships	10	8,436	3.94	2.25	0.60	1.41	40.64
	Computations and Algebraic Relationships	15	8,436	5.51	2.85	0.63	1.74	35.19
	Geometry and Measurement	7	8,436	2.87	1.53	0.41	1.18	37.72
	Data Analysis and Personal Financial Literacy	5	8,436	3.10	1.18	0.21	1.05	56.52
RLA	OVERALL TEST	52	15,444	24.74	10.07	0.91	3.00	53.51
	Reading	26	15,444	13.98	5.07	0.80	2.26	53.96
	Writing	26	15,444	10.76	5.70	0.88	1.96	52.88

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.51. Spring 2023 STAAR Spanish Grade 3 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	7,971	16.41	6.73	0.83	2.76	42.42
	Numerical Representations and Relationships	10	7,971	4.33	2.40	0.65	1.42	44.60
	Computations and Algebraic Relationships	15	7,971	5.92	3.07	0.68	1.74	37.81
	Geometry and Measurement	7	7,971	3.08	1.61	0.46	1.18	40.90
	Data Analysis and Personal Financial Literacy	5	7,971	3.08	1.25	0.25	1.08	57.39
RLA	OVERALL TEST	52	14,659	23.11	10.09	0.91	2.99	50.43
	Reading	26	14,659	13.32	5.23	0.81	2.27	51.25
	Writing	26	14,659	9.79	5.53	0.88	1.93	49.28

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.52. Spring 2023 STAAR Spanish Grade 3 Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	37	22	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	10	22	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	15	22	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	7	22	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	5	22	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	21	N/A	N/A	N/A	N/A	N/A
	Reading	26	21	N/A	N/A	N/A	N/A	N/A
	Writing	26	21	N/A	N/A	N/A	N/A	N/A

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Statistics not calculated due to small n-counts.

Table B.4.53. Spring 2023 STAAR Spanish Grade 3 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	16,094	15.89	6.44	0.82	2.75	40.89
	Numerical Representations and Relationships	10	16,094	4.13	2.32	0.63	1.42	42.55
	Computations and Algebraic Relationships	15	16,094	5.71	2.97	0.66	1.74	36.45
	Geometry and Measurement	7	16,094	2.97	1.57	0.43	1.18	39.26
	Data Analysis and Personal Financial Literacy	5	16,094	3.09	1.21	0.23	1.07	56.93
RLA	OVERALL TEST	52	29,672	23.96	10.11	0.91	3.00	52.05
	Reading	26	29,672	13.66	5.16	0.81	2.26	52.68
	Writing	26	29,672	10.30	5.64	0.88	1.94	51.16

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.54. Spring 2023 STAAR Spanish Grade 3 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	190	16.42	6.85	0.84	2.73	42.17
	Numerical Representations and Relationships	10	190	4.32	2.53	0.70	1.38	43.89
	Computations and Algebraic Relationships	15	190	5.93	3.00	0.66	1.74	37.98
	Geometry and Measurement	7	190	3.05	1.61	0.49	1.14	40.18
	Data Analysis and Personal Financial Literacy	5	190	3.12	1.28	0.35	1.03	57.72
RLA	OVERALL TEST	52	264	23.36	10.47	0.92	2.98	50.51
	Reading	26	264	13.30	5.34	0.82	2.25	51.22
	Writing	26	264	10.06	5.67	0.88	1.95	49.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.55. Spring 2023 STAAR Spanish Grade 3 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	4,088	12.89	4.82	0.68	2.74	32.91
	Numerical Representations and Relationships	10	4,088	3.22	1.85	0.42	1.41	33.29
	Computations and Algebraic Relationships	15	4,088	4.56	2.32	0.45	1.71	29.21
	Geometry and Measurement	7	4,088	2.36	1.35	0.26	1.16	31.27
	Data Analysis and Personal Financial Literacy	5	4,088	2.75	1.26	0.22	1.11	49.87
RLA	OVERALL TEST	52	7,905	18.19	8.13	0.86	2.99	40.37
	Reading	26	7,905	10.99	4.52	0.74	2.30	41.93
	Writing	26	7,905	7.20	4.39	0.81	1.90	38.17

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.56. Spring 2023 STAAR Spanish Grade 3 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	12,347	16.89	6.61	0.83	2.75	43.53
	Numerical Representations and Relationships	10	12,347	4.43	2.39	0.65	1.42	45.60
	Computations and Algebraic Relationships	15	12,347	6.09	3.06	0.67	1.75	38.85
	Geometry and Measurement	7	12,347	3.17	1.59	0.45	1.18	41.90
	Data Analysis and Personal Financial Literacy	5	12,347	3.20	1.18	0.21	1.05	59.28
RLA	OVERALL TEST	52	22,247	25.98	9.96	0.91	2.99	56.14
	Reading	26	22,247	14.60	5.04	0.80	2.25	56.44
	Writing	26	22,247	11.38	5.62	0.88	1.96	55.72

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.57. Spring 2023 STAAR Spanish Grade 4 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	11,479	16.61	7.71	0.86	2.86	39.90
	Numerical Representations and Relationships	12	11,479	5.18	2.79	0.70	1.54	40.97
	Computations and Algebraic Relationships	13	11,479	4.86	3.10	0.72	1.65	35.28
	Geometry and Measurement	11	11,479	4.38	2.21	0.55	1.48	38.04
	Data Analysis and Personal Financial Literacy	4	11,479	2.19	1.09	0.38	0.85	54.68
RLA	OVERALL TEST	52	21,659	25.26	10.67	0.92	3.09	52.89
	Reading	26	21,659	14.44	5.46	0.83	2.28	56.26
	Writing	26	21,659	10.82	5.93	0.88	2.07	48.12

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.58. Spring 2023 STAAR Spanish Grade 4 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	5,780	16.06	7.32	0.85	2.86	38.56
	Numerical Representations and Relationships	12	5,780	4.94	2.62	0.65	1.54	38.90
	Computations and Algebraic Relationships	13	5,780	4.69	3.05	0.71	1.64	33.93
	Geometry and Measurement	11	5,780	4.21	2.12	0.51	1.48	36.43
	Data Analysis and Personal Financial Literacy	4	5,780	2.22	1.07	0.36	0.86	55.38
RLA	OVERALL TEST	52	11,098	26.40	10.58	0.91	3.10	54.73
	Reading	26	11,098	14.91	5.32	0.82	2.27	57.96
	Writing	26	11,098	11.49	5.98	0.88	2.09	50.16

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.59. Spring 2023 STAAR Spanish Grade 4 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	5,669	17.17	8.05	0.87	2.85	41.30
	Numerical Representations and Relationships	12	5,669	5.43	2.93	0.73	1.53	43.14
	Computations and Algebraic Relationships	13	5,669	5.03	3.14	0.72	1.65	36.67
	Geometry and Measurement	11	5,669	4.56	2.29	0.58	1.49	39.69
	Data Analysis and Personal Financial Literacy	4	5,669	2.16	1.10	0.40	0.85	53.98
RLA	OVERALL TEST	52	10,525	24.07	10.65	0.92	3.07	50.96
	Reading	26	10,525	13.95	5.56	0.83	2.28	54.48
	Writing	26	10,525	10.12	5.79	0.88	2.05	45.98

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.60. Spring 2023 STAAR Spanish Grade 4 Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	40	14	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	12	14	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	13	14	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	11	14	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	4	14	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	11	N/A	N/A	N/A	N/A	N/A
	Reading	26	11	N/A	N/A	N/A	N/A	N/A
	Writing	26	11	N/A	N/A	N/A	N/A	N/A

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Statistics not calculated due to small n-counts.

Table B.4.61. Spring 2023 STAAR Spanish Grade 4 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	11,214	16.58	7.67	0.86	2.86	39.83
	Numerical Representations and Relationships	12	11,214	5.17	2.78	0.69	1.54	40.90
	Computations and Algebraic Relationships	13	11,214	4.85	3.10	0.72	1.65	35.19
	Geometry and Measurement	11	11,214	4.37	2.20	0.54	1.49	37.96
	Data Analysis and Personal Financial Literacy	4	11,214	2.19	1.08	0.38	0.85	54.63
RLA	OVERALL TEST	52	21,272	25.25	10.67	0.92	3.09	52.86
	Reading	26	21,272	14.43	5.45	0.83	2.28	56.24
	Writing	26	21,272	10.82	5.93	0.88	2.07	48.10

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.62. Spring 2023 STAAR Spanish Grade 4 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	133	19.47	9.75	0.92	2.79	47.43
	Numerical Representations and Relationships	12	133	6.23	3.38	0.81	1.48	50.33
	Computations and Algebraic Relationships	13	133	5.69	3.41	0.74	1.76	42.21
	Geometry and Measurement	11	133	5.14	2.93	0.76	1.44	45.41
	Data Analysis and Personal Financial Literacy	4	133	2.42	1.18	0.49	0.84	60.53
RLA	OVERALL TEST	52	224	26.93	11.05	0.92	3.05	56.43
	Reading	26	224	15.34	5.69	0.84	2.24	59.86
	Writing	26	224	11.58	5.97	0.88	2.05	51.58

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.63. Spring 2023 STAAR Spanish Grade 4 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	2,666	13.22	5.65	0.75	2.82	31.72
	Numerical Representations and Relationships	12	2,666	4.18	2.14	0.49	1.53	32.74
	Computations and Algebraic Relationships	13	2,666	3.55	2.52	0.60	1.60	25.95
	Geometry and Measurement	11	2,666	3.67	1.79	0.33	1.47	31.67
	Data Analysis and Personal Financial Literacy	4	2,666	1.83	1.02	0.25	0.88	45.64
RLA	OVERALL TEST	52	5,668	18.79	8.55	0.87	3.07	40.74
	Reading	26	5,668	11.19	4.79	0.76	2.32	44.12
	Writing	26	5,668	7.60	4.56	0.81	1.99	35.96

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.64. Spring 2023 STAAR Spanish Grade 4 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	8,813	17.63	7.95	0.87	2.86	42.38
	Numerical Representations and Relationships	12	8,813	5.48	2.89	0.72	1.54	43.46
	Computations and Algebraic Relationships	13	8,813	5.25	3.15	0.72	1.66	38.10
	Geometry and Measurement	11	8,813	4.60	2.28	0.57	1.49	39.96
	Data Analysis and Personal Financial Literacy	4	8,813	2.30	1.08	0.39	0.84	57.41
RLA	OVERALL TEST	52	15,991	27.56	10.41	0.91	3.08	57.19
	Reading	26	15,991	15.59	5.21	0.81	2.25	60.57
	Writing	26	15,991	11.97	5.93	0.88	2.09	52.43

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.65. Spring 2023 STAAR Spanish Grade 5 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	8,473	17.64	7.81	0.86	2.95	42.16
	Numerical Representations and Relationships	9	8,473	4.31	2.00	0.58	1.29	43.52
	Computations and Algebraic Relationships	20	8,473	8.37	4.40	0.79	2.02	44.54
	Geometry and Measurement	9	8,473	3.30	1.87	0.40	1.45	36.84
	Data Analysis and Personal Financial Literacy	4	8,473	1.66	1.08	0.26	0.93	41.50
RLA	OVERALL TEST	52	15,961	26.70	10.80	0.92	3.04	56.72
	Reading	26	15,961	15.65	5.89	0.86	2.18	61.01
	Writing	26	15,961	11.04	5.59	0.86	2.10	50.66
Science	OVERALL TEST	39	9,729	15.66	5.68	0.76	2.80	39.31
	Matter and Energy	6	9,729	2.37	1.29	0.34	1.05	36.54
	Force, Motion, and Energy	9	9,729	3.25	1.83	0.49	1.31	39.09
	Earth and Space	11	9,729	3.59	1.95	0.36	1.56	33.99
	Organisms and Environments	13	9,729	6.45	2.44	0.57	1.60	45.06

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.66. Spring 2023 STAAR Spanish Grade 5 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	4,162	17.52	7.57	0.85	2.94	41.96
	Numerical Representations and Relationships	9	4,162	4.26	1.99	0.57	1.30	43.22
	Computations and Algebraic Relationships	20	4,162	8.35	4.28	0.78	2.02	44.72
	Geometry and Measurement	9	4,162	3.27	1.83	0.36	1.46	36.16
	Data Analysis and Personal Financial Literacy	4	4,162	1.64	1.06	0.24	0.93	41.01
RLA	OVERALL TEST	52	8,031	27.62	10.77	0.92	3.04	58.14
	Reading	26	8,031	15.97	5.80	0.86	2.18	62.12
	Writing	26	8,031	11.66	5.65	0.86	2.11	52.53
Science	OVERALL TEST	39	4,843	15.21	5.39	0.73	2.79	38.05
	Matter and Energy	6	4,843	2.28	1.26	0.32	1.04	34.86
	Force, Motion, and Energy	9	4,843	3.08	1.75	0.43	1.32	36.96
	Earth and Space	11	4,843	3.49	1.90	0.34	1.55	32.97
	Organisms and Environments	13	4,843	6.36	2.38	0.55	1.60	44.35

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.67. Spring 2023 STAAR Spanish Grade 5 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	4,287	17.78	8.03	0.87	2.94	42.40
	Numerical Representations and Relationships	9	4,287	4.36	2.02	0.60	1.28	43.84
	Computations and Algebraic Relationships	20	4,287	8.40	4.52	0.80	2.03	44.43
	Geometry and Measurement	9	4,287	3.34	1.91	0.43	1.44	37.52
	Data Analysis and Personal Financial Literacy	4	4,287	1.68	1.09	0.28	0.93	42.01
RLA	OVERALL TEST	52	7,905	25.77	10.74	0.92	3.03	55.30
	Reading	26	7,905	15.34	5.96	0.87	2.18	59.91
	Writing	26	7,905	10.43	5.47	0.86	2.08	48.80
Science	OVERALL TEST	39	4,865	16.12	5.92	0.78	2.81	40.58
	Matter and Energy	6	4,865	2.45	1.32	0.36	1.05	38.23
	Force, Motion, and Energy	9	4,865	3.43	1.90	0.53	1.31	41.23
	Earth and Space	11	4,865	3.70	1.99	0.37	1.58	35.01
	Organisms and Environments	13	4,865	6.54	2.51	0.59	1.60	45.78

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.68. Spring 2023 STAAR Spanish Grade 5 Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	42	9	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	9	9	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	20	9	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	9	9	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	4	9	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	3	N/A	N/A	N/A	N/A	N/A
	Reading	26	3	N/A	N/A	N/A	N/A	N/A
	Writing	26	3	N/A	N/A	N/A	N/A	N/A
Science	OVERALL TEST	39	3	N/A	N/A	N/A	N/A	N/A
	Matter and Energy	6	3	N/A	N/A	N/A	N/A	N/A
	Force, Motion, and Energy	9	3	N/A	N/A	N/A	N/A	N/A
	Earth and Space	11	3	N/A	N/A	N/A	N/A	N/A
	Organisms and Environments	13	3	N/A	N/A	N/A	N/A	N/A

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Statistics not calculated due to small n-counts.

Table B.4.69. Spring 2023 STAAR Spanish Grade 5 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	8,264	17.65	7.78	0.86	2.95	42.18
	Numerical Representations and Relationships	9	8,264	4.31	2.00	0.58	1.29	43.55
	Computations and Algebraic Relationships	20	8,264	8.37	4.39	0.79	2.03	44.55
	Geometry and Measurement	9	8,264	3.30	1.87	0.39	1.45	36.82
	Data Analysis and Personal Financial Literacy	4	8,264	1.66	1.08	0.26	0.93	41.59
RLA	OVERALL TEST	52	15,700	26.71	10.79	0.92	3.04	56.74
	Reading	26	15,700	15.66	5.89	0.86	2.18	61.03
	Writing	26	15,700	11.05	5.59	0.86	2.10	50.69
Science	OVERALL TEST	39	9,525	15.68	5.68	0.76	2.80	39.37
	Matter and Energy	6	9,525	2.37	1.30	0.35	1.05	36.59
	Force, Motion, and Energy	9	9,525	3.26	1.83	0.49	1.31	39.18
	Earth and Space	11	9,525	3.59	1.95	0.36	1.56	34.03
	Organisms and Environments	13	9,525	6.46	2.45	0.57	1.60	45.11

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.70. Spring 2023 STAAR Spanish Grade 5 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	97	18.87	9.51	0.91	2.86	45.10
	Numerical Representations and Relationships	9	97	4.56	2.30	0.70	1.26	46.91
	Computations and Algebraic Relationships	20	97	9.12	5.28	0.87	1.89	48.35
	Geometry and Measurement	9	97	3.61	2.12	0.52	1.47	40.27
	Data Analysis and Personal Financial Literacy	4	97	1.58	1.11	0.33	0.91	39.43
RLA	OVERALL TEST	52	131	26.51	11.09	0.93	2.99	56.73
	Reading	26	131	15.47	6.11	0.88	2.14	60.58
	Writing	26	131	11.04	5.71	0.87	2.06	51.31
Science	OVERALL TEST	39	93	14.57	5.83	0.77	2.78	36.81
	Matter and Energy	6	93	2.29	1.19	0.18	1.08	35.05
	Force, Motion, and Energy	9	93	2.87	1.81	0.52	1.25	35.41
	Earth and Space	11	93	3.30	2.14	0.50	1.52	31.30
	Organisms and Environments	13	93	6.11	2.41	0.55	1.62	43.01

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.71. Spring 2023 STAAR Spanish Grade 5 Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	1,900	14.92	6.31	0.79	2.90	36.18
	Numerical Representations and Relationships	9	1,900	3.70	1.83	0.50	1.29	36.99
	Computations and Algebraic Relationships	20	1,900	6.96	3.66	0.71	1.98	38.09
	Geometry and Measurement	9	1,900	2.82	1.60	0.24	1.40	32.03
	Data Analysis and Personal Financial Literacy	4	1,900	1.44	1.01	0.15	0.93	35.88
RLA	OVERALL TEST	52	3,991	20.56	9.44	0.90	3.06	44.97
	Reading	26	3,991	12.41	5.58	0.83	2.27	48.68
	Writing	26	3,991	8.15	4.58	0.80	2.03	39.72
Science	OVERALL TEST	39	2,379	13.92	5.12	0.70	2.79	34.90
	Matter and Energy	6	2,379	2.15	1.20	0.23	1.05	32.48
	Force, Motion, and Energy	9	2,379	2.91	1.68	0.40	1.30	35.16
	Earth and Space	11	2,379	3.12	1.80	0.29	1.52	29.60
	Organisms and Environments	13	2,379	5.75	2.35	0.51	1.64	40.18

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.72. Spring 2023 STAAR Spanish Grade 5 Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	6,573	18.43	8.02	0.86	2.95	43.89
	Numerical Representations and Relationships	9	6,573	4.48	2.02	0.59	1.29	45.40
	Computations and Algebraic Relationships	20	6,573	8.78	4.51	0.80	2.03	46.41
	Geometry and Measurement	9	6,573	3.45	1.92	0.42	1.46	38.23
	Data Analysis and Personal Financial Literacy	4	6,573	1.72	1.09	0.28	0.93	43.12
RLA	OVERALL TEST	52	11,970	28.74	10.44	0.92	3.02	60.64
	Reading	26	11,970	16.73	5.58	0.85	2.14	65.12
	Writing	26	11,970	12.01	5.57	0.86	2.11	54.31
Science	OVERALL TEST	39	7,350	16.23	5.74	0.76	2.80	40.73
	Matter and Energy	6	7,350	2.44	1.32	0.37	1.05	37.85
	Force, Motion, and Energy	9	7,350	3.37	1.86	0.50	1.32	40.36
	Earth and Space	11	7,350	3.74	1.97	0.36	1.57	35.41
	Organisms and Environments	13	7,350	6.68	2.43	0.57	1.59	46.64

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.73. Spring 2023 STAAR EOC Assessments Algebra I Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	476,739	30.91	12.31	0.93	3.31	52.10
	Number and Algebraic Methods	13	476,739	6.70	2.78	0.67	1.59	52.60
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	476,739	7.39	3.02	0.73	1.55	56.33
	Writing and Solving Linear Functions, Equations, and Inequalities	14	476,739	7.29	3.42	0.78	1.61	51.97
	Quadratic Functions and Equations	12	476,739	6.35	2.95	0.74	1.51	50.85
	Exponential Functions and Equations	7	476,739	3.18	1.85	0.65	1.10	45.76

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.74. Spring 2023 STAAR EOC Assessments Algebra I Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	226,861	31.59	11.90	0.92	3.31	53.23
	Number and Algebraic Methods	13	226,861	6.89	2.71	0.66	1.58	54.15
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	226,861	7.57	2.95	0.72	1.56	57.90
	Writing and Solving Linear Functions, Equations, and Inequalities	14	226,861	7.51	3.31	0.77	1.60	53.52
	Quadratic Functions and Equations	12	226,861	6.48	2.88	0.73	1.51	51.65
	Exponential Functions and Equations	7	226,861	3.14	1.80	0.62	1.11	45.06

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.75. Spring 2023 STAAR EOC Assessments Algebra I Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	248,486	30.31	12.65	0.93	3.31	51.10
	Number and Algebraic Methods	13	248,486	6.52	2.82	0.68	1.59	51.22
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	248,486	7.24	3.06	0.74	1.55	54.94
	Writing and Solving Linear Functions, Equations, and Inequalities	14	248,486	7.10	3.50	0.79	1.61	50.59
	Quadratic Functions and Equations	12	248,486	6.24	3.02	0.75	1.51	50.16
	Exponential Functions and Equations	7	248,486	3.21	1.90	0.67	1.10	46.44

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.76. Spring 2023 STAAR EOC Assessments Algebra I Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	65,433	26.74	10.96	0.91	3.37	44.91
	Number and Algebraic Methods	13	65,433	5.96	2.57	0.60	1.62	46.80
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	65,433	6.40	2.82	0.69	1.56	48.91
	Writing and Solving Linear Functions, Equations, and Inequalities	14	65,433	6.27	3.11	0.72	1.64	44.50
	Quadratic Functions and Equations	12	65,433	5.46	2.74	0.69	1.53	43.16
	Exponential Functions and Equations	7	65,433	2.66	1.63	0.53	1.11	37.82

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.77. Spring 2023 STAAR EOC Assessments Algebra I Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	259,099	29.03	11.45	0.91	3.36	48.93
	Number and Algebraic Methods	13	259,099	6.37	2.63	0.63	1.61	49.97
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	259,099	7.02	2.92	0.71	1.57	53.59
	Writing and Solving Linear Functions, Equations, and Inequalities	14	259,099	6.78	3.23	0.75	1.63	48.50
	Quadratic Functions and Equations	12	259,099	5.98	2.82	0.71	1.53	47.74
	Exponential Functions and Equations	7	259,099	2.88	1.72	0.58	1.12	41.33

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.78. Spring 2023 STAAR EOC Assessments Algebra I White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	113,170	34.86	12.11	0.93	3.26	58.80
	Number and Algebraic Methods	13	113,170	7.31	2.75	0.68	1.55	57.57
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	113,170	8.28	2.89	0.72	1.53	62.80
	Writing and Solving Linear Functions, Equations, and Inequalities	14	113,170	8.34	3.35	0.77	1.59	59.23
	Quadratic Functions and Equations	12	113,170	7.16	2.90	0.74	1.48	57.64
	Exponential Functions and Equations	7	113,170	3.77	1.90	0.67	1.09	54.81

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.79. Spring 2023 STAAR EOC Assessments Algebra I Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	110,233	23.39	9.24	0.87	3.39	39.29
	Number and Algebraic Methods	13	110,233	5.40	2.39	0.53	1.64	42.31
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	110,233	5.72	2.58	0.63	1.58	43.83
	Writing and Solving Linear Functions, Equations, and Inequalities	14	110,233	5.25	2.69	0.63	1.64	37.61
	Quadratic Functions and Equations	12	110,233	4.76	2.43	0.60	1.54	37.53
	Exponential Functions and Equations	7	110,233	2.26	1.40	0.37	1.11	31.77

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.80. Spring 2023 STAAR EOC Assessments Algebra I Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	366,506	33.17	12.22	0.93	3.28	55.95
	Number and Algebraic Methods	13	366,506	7.08	2.77	0.68	1.57	55.70
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	366,506	7.90	2.95	0.73	1.54	60.09
	Writing and Solving Linear Functions, Equations, and Inequalities	14	366,506	7.90	3.38	0.78	1.60	56.29
	Quadratic Functions and Equations	12	366,506	6.83	2.93	0.74	1.50	54.86
	Exponential Functions and Equations	7	366,506	3.45	1.89	0.66	1.10	49.96

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.81. Spring 2023 STAAR EOC Assessments English I Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	517,369	36.47	14.60	0.95	3.35	58.98
	Reading	32	517,369	18.89	7.24	0.89	2.41	59.43
	Writing	32	517,369	17.59	7.99	0.92	2.31	58.42

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.82. Spring 2023 STAAR EOC Assessments English I Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	241,593	38.99	14.11	0.94	3.32	62.34
	Reading	32	241,593	19.85	6.97	0.88	2.39	62.22
	Writing	32	241,593	19.14	7.75	0.91	2.29	62.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.83. Spring 2023 STAAR EOC Assessments English I Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	274,141	34.27	14.66	0.95	3.37	56.05
	Reading	32	274,141	18.04	7.37	0.89	2.43	57.00
	Writing	32	274,141	16.23	7.94	0.91	2.33	54.85

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.84. Spring 2023 STAAR EOC Assessments English I Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	70,324	32.44	13.57	0.94	3.44	53.21
	Reading	32	70,324	17.10	6.87	0.87	2.48	53.82
	Writing	32	70,324	15.34	7.43	0.90	2.36	52.45

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.85. Spring 2023 STAAR EOC Assessments English I Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	289,427	33.68	14.08	0.94	3.42	54.88
	Reading	32	289,427	17.55	7.05	0.88	2.47	55.25
	Writing	32	289,427	16.14	7.70	0.91	2.35	54.41

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.86. Spring 2023 STAAR EOC Assessments English I White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	117,120	43.04	13.25	0.94	3.20	68.61
	Reading	32	117,120	22.01	6.53	0.88	2.29	69.23
	Writing	32	117,120	21.04	7.37	0.91	2.22	67.84

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.87. Spring 2023 STAAR EOC Assessments English I Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	138,193	25.05	11.24	0.90	3.49	42.11
	Reading	32	138,193	13.68	6.05	0.82	2.54	43.29
	Writing	32	138,193	11.37	6.04	0.84	2.38	40.61

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.88. Spring 2023 STAAR EOC Assessments English I Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	379,176	40.63	13.41	0.94	3.28	65.13
	Reading	32	379,176	20.78	6.69	0.88	2.36	65.31
	Writing	32	379,176	19.85	7.38	0.91	2.27	64.91

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.89. Spring 2023 STAAR EOC Assessments English II Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	469,418	36.96	13.94	0.94	3.47	59.02
	Reading	32	469,418	19.30	6.63	0.85	2.56	59.56
	Writing	32	469,418	17.65	7.98	0.91	2.33	58.33

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.90. Spring 2023 STAAR EOC Assessments English II Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	222,966	39.21	13.47	0.93	3.44	61.92
	Reading	32	222,966	20.10	6.37	0.84	2.53	62.01
	Writing	32	222,966	19.11	7.74	0.91	2.31	61.81

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.91. Spring 2023 STAAR EOC Assessments English II Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	245,012	34.91	14.05	0.94	3.49	56.38
	Reading	32	245,012	18.58	6.78	0.85	2.58	57.35
	Writing	32	245,012	16.33	7.95	0.91	2.34	55.17

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.92. Spring 2023 STAAR EOC Assessments English II Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	62,251	33.19	13.03	0.93	3.56	53.64
	Reading	32	62,251	17.70	6.43	0.83	2.64	54.50
	Writing	32	62,251	15.50	7.34	0.90	2.38	52.55

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.93. Spring 2023 STAAR EOC Assessments English II Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	258,937	34.10	13.51	0.93	3.54	54.76
	Reading	32	258,937	18.03	6.52	0.84	2.61	55.60
	Writing	32	258,937	16.07	7.69	0.90	2.37	53.71

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.94. Spring 2023 STAAR EOC Assessments English II White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	111,252	43.30	12.45	0.93	3.32	68.45
	Reading	32	111,252	22.14	5.83	0.83	2.43	68.44
	Writing	32	111,252	21.16	7.33	0.91	2.24	68.46

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.95. Spring 2023 STAAR EOC Assessments English II Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	112,437	25.12	10.82	0.89	3.60	41.77
	Reading	32	112,437	14.18	5.78	0.78	2.69	43.76
	Writing	32	112,437	10.94	5.88	0.84	2.37	39.26

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.96. Spring 2023 STAAR EOC Assessments English II Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	356,981	40.68	12.68	0.93	3.42	64.45
	Reading	32	356,981	20.92	6.03	0.83	2.51	64.54
	Writing	32	356,981	19.77	7.36	0.90	2.30	64.34

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.97. Spring 2023 STAAR EOC Assessments Biology Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	461,492	27.43	11.16	0.92	3.16	52.70
	Cell Structure and Function	11	461,492	5.57	2.55	0.71	1.37	52.21
	Mechanisms of Genetics	10	461,492	5.47	2.53	0.77	1.20	57.47
	Biological Evolution and Classification	10	461,492	4.85	2.35	0.65	1.38	46.54
	Biological Processes and Systems	11	461,492	5.13	2.62	0.66	1.53	49.02
	Interdependence within Environmental Systems	11	461,492	6.41	2.81	0.74	1.44	58.27

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.98. Spring 2023 STAAR EOC Assessments Biology Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	222,713	27.62	10.84	0.91	3.17	53.08
	Cell Structure and Function	11	222,713	5.58	2.47	0.70	1.36	52.36
	Mechanisms of Genetics	10	222,713	5.64	2.47	0.77	1.20	59.12
	Biological Evolution and Classification	10	222,713	4.85	2.33	0.65	1.38	46.50
	Biological Processes and Systems	11	222,713	5.13	2.56	0.64	1.54	49.07
	Interdependence within Environmental Systems	11	222,713	6.43	2.74	0.72	1.45	58.36

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.99. Spring 2023 STAAR EOC Assessments Biology Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	237,283	27.26	11.44	0.92	3.15	52.38
	Cell Structure and Function	11	237,283	5.57	2.61	0.72	1.37	52.10
	Mechanisms of Genetics	10	237,283	5.32	2.56	0.78	1.20	55.95
	Biological Evolution and Classification	10	237,283	4.85	2.36	0.66	1.38	46.61
	Biological Processes and Systems	11	237,283	5.12	2.67	0.67	1.53	49.00
	Interdependence within Environmental Systems	11	237,283	6.40	2.88	0.76	1.43	58.21

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.100. Spring 2023 STAAR EOC Assessments Biology Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	61,293	23.66	10.08	0.90	3.21	45.75
	Cell Structure and Function	11	61,293	4.86	2.36	0.65	1.40	45.54
	Mechanisms of Genetics	10	61,293	4.72	2.38	0.73	1.24	50.06
	Biological Evolution and Classification	10	61,293	4.15	2.14	0.57	1.41	39.99
	Biological Processes and Systems	11	61,293	4.32	2.41	0.59	1.54	41.78
	Interdependence within Environmental Systems	11	61,293	5.60	2.70	0.71	1.45	51.38

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.101. Spring 2023 STAAR EOC Assessments Biology Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	250,802	25.05	10.36	0.90	3.21	48.22
	Cell Structure and Function	11	250,802	5.13	2.42	0.67	1.40	48.05
	Mechanisms of Genetics	10	250,802	5.03	2.41	0.74	1.23	53.08
	Biological Evolution and Classification	10	250,802	4.40	2.18	0.59	1.40	41.99
	Biological Processes and Systems	11	250,802	4.62	2.44	0.60	1.55	44.51
	Interdependence within Environmental Systems	11	250,802	5.87	2.77	0.72	1.46	53.46

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.102. Spring 2023 STAAR EOC Assessments Biology White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	110,793	32.43	10.38	0.91	3.07	62.05
	Cell Structure and Function	11	110,793	6.48	2.43	0.70	1.32	60.79
	Mechanisms of Genetics	10	110,793	6.40	2.40	0.77	1.15	66.66
	Biological Evolution and Classification	10	110,793	5.77	2.29	0.66	1.34	55.64
	Biological Processes and Systems	11	110,793	6.19	2.52	0.64	1.50	58.59
	Interdependence within Environmental Systems	11	110,793	7.59	2.46	0.68	1.38	68.58

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.103. Spring 2023 STAAR EOC Assessments Biology Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	108,625	19.51	8.41	0.85	3.23	37.78
	Cell Structure and Function	11	108,625	4.13	2.13	0.55	1.43	38.66
	Mechanisms of Genetics	10	108,625	3.89	2.06	0.62	1.27	41.71
	Biological Evolution and Classification	10	108,625	3.45	1.83	0.41	1.41	32.88
	Biological Processes and Systems	11	108,625	3.54	2.07	0.47	1.51	34.62
	Interdependence within Environmental Systems	11	108,625	4.50	2.56	0.68	1.45	41.03

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.104. Spring 2023 STAAR EOC Assessments Biology Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	352,867	29.86	10.76	0.92	3.13	57.29
	Cell Structure and Function	11	352,867	6.02	2.50	0.71	1.35	56.38
	Mechanisms of Genetics	10	352,867	5.96	2.46	0.77	1.18	62.32
	Biological Evolution and Classification	10	352,867	5.28	2.32	0.66	1.36	50.74
	Biological Processes and Systems	11	352,867	5.61	2.58	0.65	1.53	53.45
	Interdependence within Environmental Systems	11	352,867	7.00	2.62	0.70	1.43	63.58

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.105. Spring 2023 STAAR EOC Assessments U.S. History Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	380,315	44.33	13.81	0.92	3.83	57.44
	History	37	380,315	19.77	6.63	0.83	2.71	53.86
	Geography and Culture	13	380,315	8.09	3.05	0.77	1.47	63.46
	Government and Citizenship	10	380,315	5.71	2.24	0.66	1.31	57.18
	Economics, Science, Technology, and Society	18	380,315	10.76	3.37	0.72	1.79	60.10

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.106. Spring 2023 STAAR EOC Assessments U.S. History Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	187,734	43.65	13.31	0.92	3.86	56.39
	History	37	187,734	19.22	6.38	0.82	2.72	52.18
	Geography and Culture	13	187,734	8.15	3.04	0.76	1.49	63.64
	Government and Citizenship	10	187,734	5.75	2.20	0.64	1.32	57.30
	Economics, Science, Technology, and Society	18	187,734	10.53	3.23	0.69	1.81	58.66

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.107. Spring 2023 STAAR EOC Assessments U.S. History Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	190,902	44.99	14.26	0.93	3.79	58.48
	History	37	190,902	20.30	6.82	0.85	2.68	55.53
	Geography and Culture	13	190,902	8.03	3.07	0.77	1.46	63.27
	Government and Citizenship	10	190,902	5.67	2.28	0.67	1.30	57.07
	Economics, Science, Technology, and Society	18	190,902	10.99	3.48	0.74	1.77	61.51

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.108. Spring 2023 STAAR EOC Assessments U.S. History Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	48,209	40.19	13.30	0.91	3.90	51.98
	History	37	48,209	17.95	6.26	0.81	2.74	48.70
	Geography and Culture	13	48,209	7.30	3.09	0.76	1.51	57.41
	Government and Citizenship	10	48,209	5.16	2.18	0.63	1.32	51.92
	Economics, Science, Technology, and Society	18	48,209	9.77	3.32	0.69	1.86	54.37

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.109. Spring 2023 STAAR EOC Assessments U.S. History Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	202,014	41.83	13.37	0.92	3.88	54.14
	History	37	202,014	18.61	6.34	0.81	2.73	50.56
	Geography and Culture	13	202,014	7.61	3.05	0.76	1.50	59.75
	Government and Citizenship	10	202,014	5.33	2.21	0.64	1.33	53.40
	Economics, Science, Technology, and Society	18	202,014	10.28	3.32	0.70	1.83	57.38

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.110. Spring 2023 STAAR EOC Assessments U.S. History White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	98,526	49.38	12.86	0.92	3.72	64.17
	History	37	98,526	22.03	6.42	0.83	2.65	60.37
	Geography and Culture	13	98,526	9.05	2.74	0.73	1.42	71.04
	Government and Citizenship	10	98,526	6.49	2.08	0.62	1.27	64.89
	Economics, Science, Technology, and Society	18	98,526	11.81	3.13	0.70	1.71	66.06

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.111. Spring 2023 STAAR EOC Assessments U.S. History Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	69,164	33.40	11.92	0.89	3.95	43.10
	History	37	69,164	15.17	5.58	0.76	2.74	40.87
	Geography and Culture	13	69,164	5.79	2.86	0.71	1.53	45.93
	Government and Citizenship	10	69,164	4.15	2.05	0.56	1.35	42.11
	Economics, Science, Technology, and Society	18	69,164	8.29	3.20	0.64	1.92	45.92

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.112. Spring 2023 STAAR EOC Assessments U.S. History Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	311,151	46.75	13.01	0.91	3.80	60.63
	History	37	311,151	20.79	6.40	0.82	2.69	56.75
	Geography and Culture	13	311,151	8.60	2.85	0.74	1.46	67.35
	Government and Citizenship	10	311,151	6.06	2.13	0.63	1.30	60.53
	Economics, Science, Technology, and Society	18	311,151	11.31	3.15	0.69	1.76	63.25

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR-20 was computed instead.
3. Mean of percent correct (0–100%) across all items.

Spring 2023 STAAR Progress Measure

During the 2022-2023 school year, Progress Measures were not calculated or reported. This section is left as a placeholder to keep the numberings consistent.

Spring 2023 STAAR
On-Track Results

During the 2022-2023 school year, On-Track Results were not calculated or reported. This section is left as a placeholder to keep the numberings consistent.

Spring 2023 STAAR
Scale Score
Descriptive Statistics

Glossary

This glossary provides definitions for the statistical terms that appear in the tables and graphs in this section. Definitions of statistical terms and concepts in the other sections are given in Chapter 3 and Chapter 4.

Descriptive Statistics

Mean The mean is a measure of central tendency. It is the average score for the assessment. It is computed by summing the scores of all students and dividing the sum by the total number of students (*N*).

Median The median is another measure of central tendency. It is the score at the middle of the frequency distribution for the assessment. It is computed by finding the score at which there are the same number of scores above as there are below.

Mode The mode is another measure of central tendency. It is the most frequently obtained score for the assessment. It is determined by computing the frequency distribution and finding the score point with the highest frequency (*n*-count).

Range The range is a measure of statistical dispersion (variability or spread). It is the difference between the lowest and highest scores obtained by students on the assessment. It is computed by subtracting the lowest score from the highest score.

Interquartile Range The interquartile range is another measure of statistical dispersion (variability or spread). It is the difference between the first and third quartiles (or 25th and 75th percentiles) of the score distribution for the assessment. It is computed by subtracting the score at the first quartile (the point that splits the lowest 25 percent of the scores) from the score at the third quartile (the point that splits the highest 25 percent of the scores).

Standard Deviation The standard deviation is another measure of statistical dispersion (variability or spread). It is an indicator of the degree of score variation around the mean. It is computed using the following formula:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, and N is the total number of students who took the assessment.

Variance The variance is another measure of statistical dispersion (variability or spread) around the mean. It is computed as the square of the standard deviation.

Skewness The skewness is an indicator of the shape of the score distribution. It measures the extent to which the score distribution "leans" to one side of the mean. A positive skewness indicates that the score

distribution leans below the mean. A negative skewness indicates that the score distribution leans above the mean. A skewness of zero indicates that the score distribution is symmetric around the mean. It is computed using the following formula:

$$Skewness = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^3,$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Kurtosis The kurtosis is another indicator of the shape of the score distribution. It measures the "peakedness" of the score distribution. A positive kurtosis is referred to as *leptokurtic*, meaning that the distribution has a more acute peak around the mean and fatter tails. A negative kurtosis is called *platykurtic*, meaning the distribution has a lower, wider peak around the mean and thinner tails. It is computed using the following formula:

$$Kurtosis = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Frequency Distributions

Frequency (Freq) The frequency is the number of students who obtained the particular score point on the assessment.

Cumulative Frequency (CumFreq) The cumulative frequency is the number of students who obtained a score that is less than or equal to the particular score point on the assessment.

Percentage (Pct) The percentage is the percentage of students who obtained the particular score point on the assess-ment. It is computed as: Pct = Freq ÷ N × 100.

Cumulative Percentage (CumPct) The cumulative percentage is the percentage of students who obtained a score that is less than or equal to the particular score point on the assessment. It is computed as: CumPct = CumFreq ÷ N × 100.

Table B.7.1. Spring 2023 STAAR Grades 3–8 Assessments
Scale Score Descriptive Statistics

	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	369,987	1456.14	1439	1327	1210	214	151.41	22926.05	0.65	0.67
Grade 4	Mathematics	373,970	1557.57	1530	1402	1220	218	167.34	28002.27	0.67	0.61
Grade 5	Mathematics	378,653	1645.41	1620	1500	1200	211	161.73	26156.73	0.67	0.75
Grade 6	Mathematics	384,765	1716.43	1689	1601	1280	203	154.55	23885.87	0.86	1.27
Grade 7	Mathematics	331,698	1760.72	1728	1650	1250	184	148.39	22018.74	1.13	2.20
Grade 8	Mathematics	364,110	1856.20	1836	1741	1230	193	155.28	24112.65	0.89	1.53
Grade 3	RLA	356,509	1458.01	1467	1539	1400	222	152.02	23111.24	-0.02	0.18
Grade 4	RLA	365,003	1536.26	1530	1383	1390	224	151.84	23054.84	0.19	0.25
Grade 5	RLA	372,636	1601.71	1604	1670	1390	201	150.49	22646.39	-0.17	0.32
Grade 6	RLA	391,376	1633.32	1634	1673	1400	191	142.33	20257.82	0.14	0.79
Grade 7	RLA	400,416	1668.66	1669	1734	1400	207	151.37	22914.22	-0.06	0.49
Grade 8	RLA	410,468	1714.37	1707	1757	1380	179	141.30	19966.00	0.07	0.74
Grade 5	Science	378,696	3753.38	3740	3677	5060	788	557.83	311169.40	0.22	0.39
Grade 8	Science	407,845	3952.39	3890	3428	5800	858	662.96	439513.49	0.58	1.12
Grade 8	Social Studies	414,692	3727.54	3648	3242	5500	812	588.42	346242.27	0.54	0.99

Table B.7.2. Spring 2023 STAAR Spanish Grades 3–5 Assessments Scale Score Descriptive Statistics

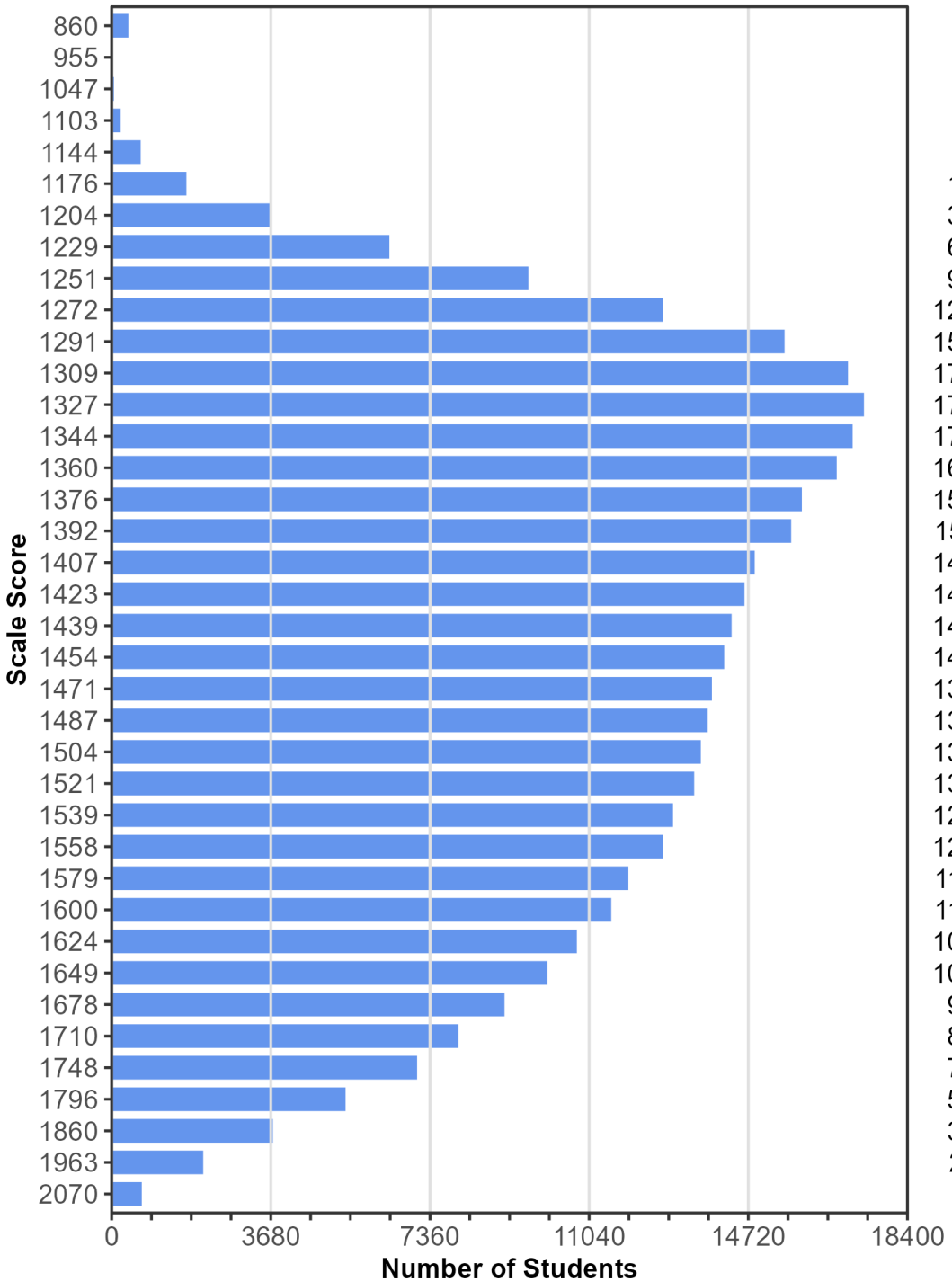
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	16,435	1390.50	1360	1309	1210	145	116.57	13587.80	0.93	1.59
Grade 4	Mathematics	11,479	1467.89	1448	1385	1220	159	127.27	16196.69	1.00	1.63
Grade 5	Mathematics	8,473	1547.95	1528	1454	1200	166	119.99	14398.07	0.84	1.32
Grade 3	RLA	30,152	1335.56	1332	1194	1470	236	150.73	22719.52	0.19	-0.25
Grade 4	RLA	21,659	1408.33	1408	1266	1430	217	145.89	21282.43	0.22	-0.15
Grade 5	RLA	15,961	1478.49	1481	1505	1460	220	160.32	25701.23	0.19	-0.10
Grade 5	Science	9,729	3376.09	3353	3138	3811	553	404.07	163275.10	0.43	0.38

Table B.7.3. Spring 2023 STAAR EOC Assessments
Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	476,739	3995.04	3927	3489	4930	713	572.74	328036.0	0.89	1.68
English I	517,369	4053.64	4052	4430	4250	697	485.23	235446.3	0.15	0.20
English II	469,418	4064.34	4052	4323	4400	610	469.57	220500.2	0.09	0.38
Biology	461,492	4118.43	4076	3641	4360	714	518.70	269048.8	0.57	0.78
U.S. History	380,315	4269.25	4270	4488	5330	649	469.23	220181.5	0.03	0.40

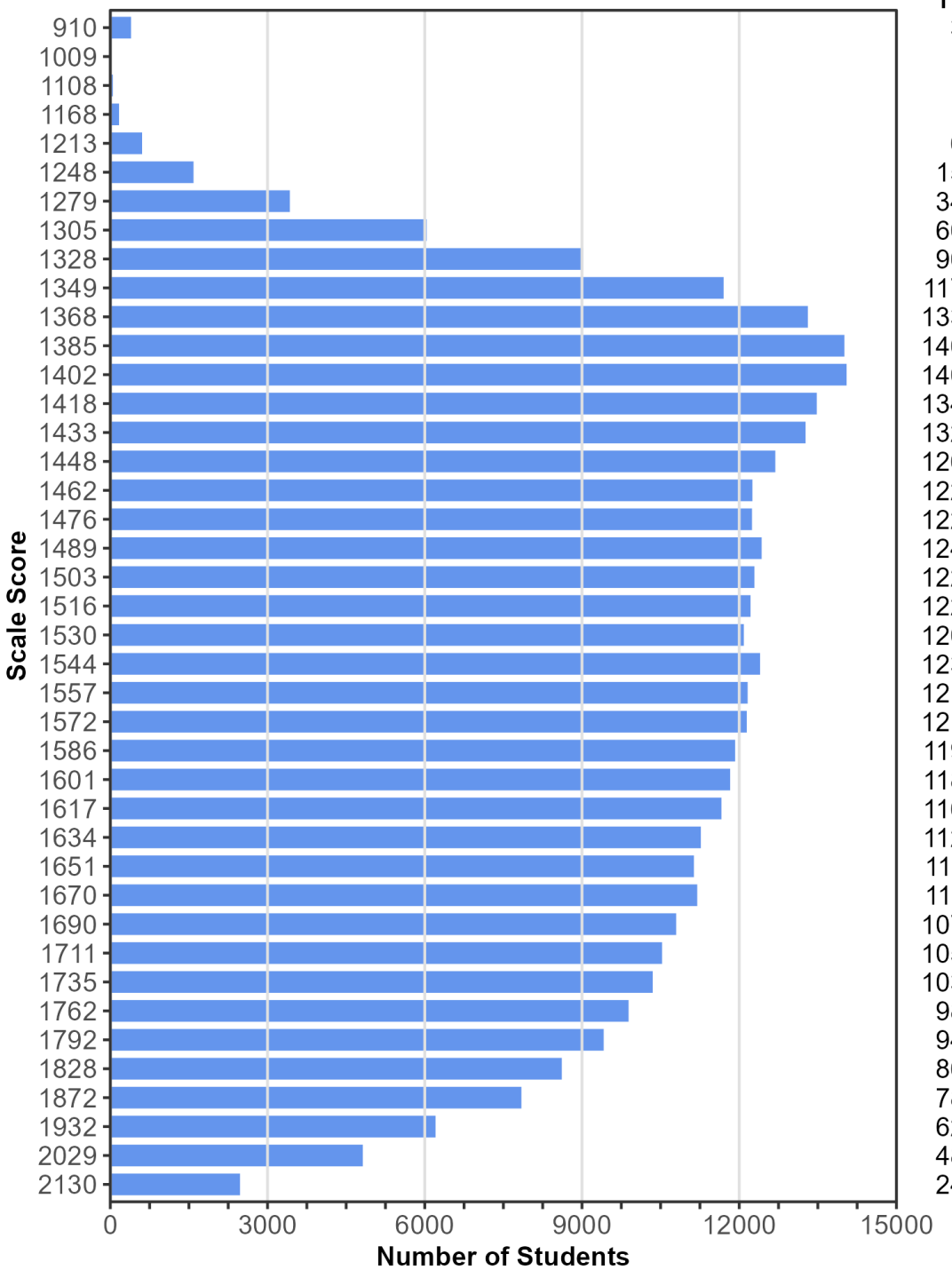
Spring 2023 STAAR Frequency Distribution of Scale Scores

Figure B.7.1. Spring 2023 STAAR Grade 3 Mathematics
Frequency Distribution of Scale Scores
All Students



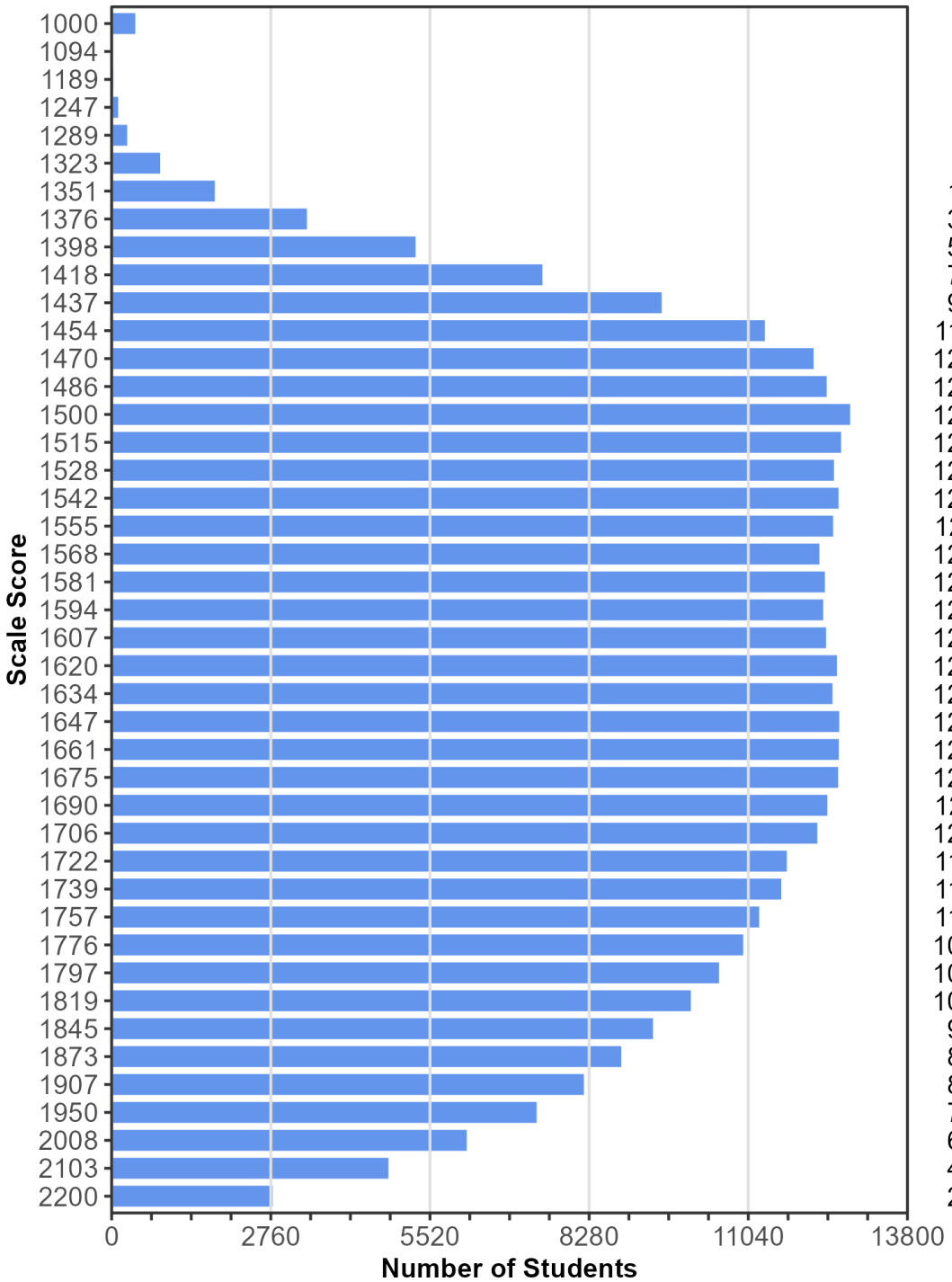
Scale Score	Freq	Cum Freq	Pct	Cum Pct
860	387	387	0.10	0.10
955	13	400	0.00	0.11
1047	52	452	0.01	0.12
1103	209	661	0.06	0.18
1144	671	1332	0.18	0.36
1176	1728	3060	0.47	0.83
1204	3674	6734	0.99	1.82
1229	6422	13156	1.74	3.56
1251	9635	22791	2.60	6.16
1272	12740	35531	3.44	9.60
1291	15558	51089	4.21	13.81
1309	17025	68114	4.60	18.41
1327	17394	85508	4.70	23.11
1344	17132	102640	4.63	27.74
1360	16764	119404	4.53	32.27
1376	15958	135362	4.31	36.59
1392	15711	151073	4.25	40.83
1407	14868	165941	4.02	44.85
1423	14632	180573	3.95	48.81
1439	14334	194907	3.87	52.68
1454	14163	209070	3.83	56.51
1471	13878	222948	3.75	60.26
1487	13780	236728	3.72	63.98
1504	13622	250350	3.68	67.66
1521	13469	263819	3.64	71.30
1539	12979	276798	3.51	74.81
1558	12750	289548	3.45	78.26
1579	11947	301495	3.23	81.49
1600	11551	313046	3.12	84.61
1624	10756	323802	2.91	87.52
1649	10076	333878	2.72	90.24
1678	9083	342961	2.45	92.70
1710	8015	350976	2.17	94.86
1748	7062	358038	1.91	96.77
1796	5408	363446	1.46	98.23
1860	3728	367174	1.01	99.24
1963	2117	369291	0.57	99.81
2070	696	369987	0.19	100.00

Figure B.7.2. Spring 2023 STAAR Grade 4 Mathematics
Frequency Distribution of Scale Scores
All Students



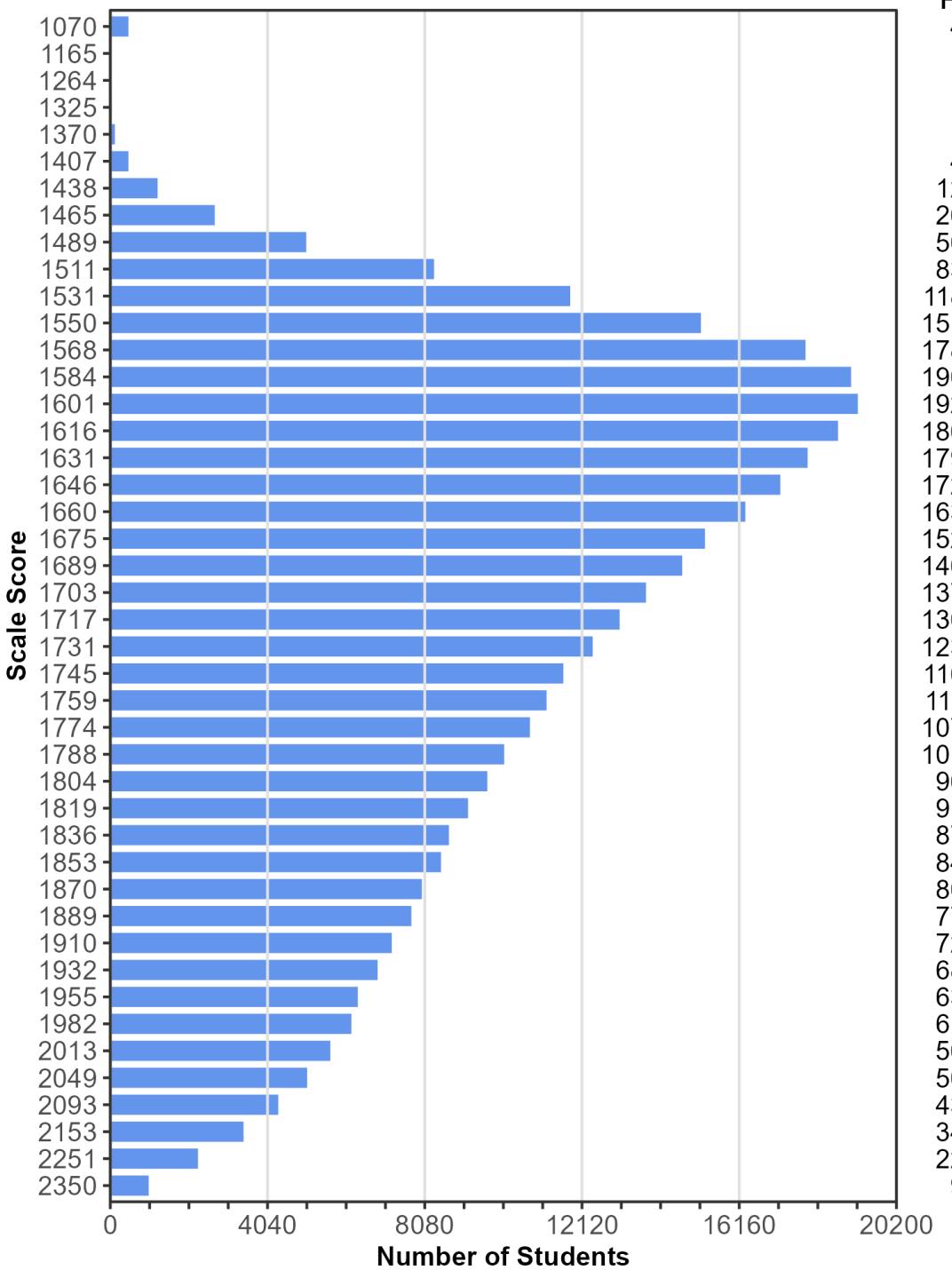
Scale Score	Freq	Cum Freq	Pct	Cum Pct
910	396	396	0.11	0.11
1009	9	405	0.00	0.11
1108	48	453	0.01	0.12
1168	167	620	0.04	0.17
1213	607	1227	0.16	0.33
1248	1588	2815	0.42	0.75
1279	3426	6241	0.92	1.67
1305	6036	12277	1.61	3.28
1328	9021	21298	2.41	5.70
1349	11702	33000	3.13	8.82
1368	13310	46310	3.56	12.38
1385	14007	60317	3.75	16.13
1402	14049	74366	3.76	19.89
1418	13480	87846	3.60	23.49
1433	13266	101112	3.55	27.04
1448	12689	113801	3.39	30.43
1462	12252	126053	3.28	33.71
1476	12244	138297	3.27	36.98
1489	12428	150725	3.32	40.30
1503	12290	163015	3.29	43.59
1516	12216	175231	3.27	46.86
1530	12088	187319	3.23	50.09
1544	12398	199717	3.32	53.40
1557	12161	211878	3.25	56.66
1572	12144	224022	3.25	59.90
1586	11922	235944	3.19	63.09
1601	11827	247771	3.16	66.25
1617	11661	259432	3.12	69.37
1634	11268	270700	3.01	72.39
1651	11136	281836	2.98	75.36
1670	11199	293035	2.99	78.36
1690	10797	303832	2.89	81.25
1711	10528	314360	2.82	84.06
1735	10349	324709	2.77	86.83
1762	9889	334598	2.64	89.47
1792	9414	344012	2.52	91.99
1828	8615	352627	2.30	94.29
1872	7844	360471	2.10	96.39
1932	6206	366677	1.66	98.05
2029	4818	371495	1.29	99.34
2130	2475	373970	0.66	100.00

Figure B.7.3. Spring 2023 STAAR Grade 5 Mathematics
Frequency Distribution of Scale Scores
All Students



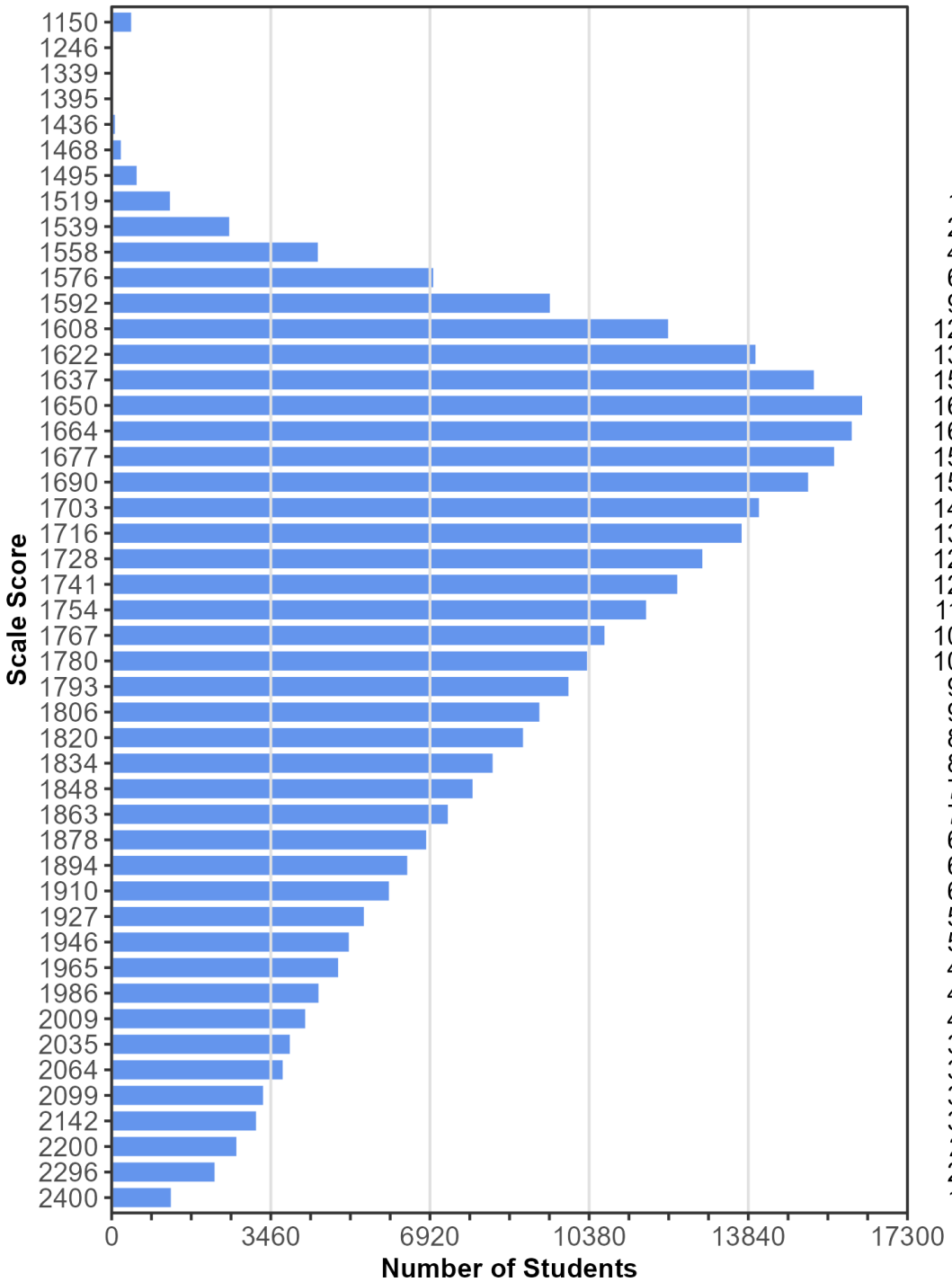
Freq	Cum Freq	Pct	Cum Pct
410	410	0.11	0.11
11	421	0.00	0.11
28	449	0.01	0.12
113	562	0.03	0.15
269	831	0.07	0.22
841	1672	0.22	0.44
1789	3461	0.47	0.91
3386	6847	0.89	1.81
5271	12118	1.39	3.20
7470	19588	1.97	5.17
9539	29127	2.52	7.69
11327	40454	2.99	10.68
12172	52626	3.21	13.90
12400	65026	3.27	17.17
12807	77833	3.38	20.56
12649	90482	3.34	23.90
12526	103008	3.31	27.20
12605	115613	3.33	30.53
12511	128124	3.30	33.84
12273	140397	3.24	37.08
12368	152765	3.27	40.34
12339	165104	3.26	43.60
12389	177493	3.27	46.87
12576	190069	3.32	50.20
12501	202570	3.30	53.50
12617	215187	3.33	56.83
12609	227796	3.33	60.16
12598	240394	3.33	63.49
12411	252805	3.28	66.76
12237	265042	3.23	70.00
11708	276750	3.09	73.09
11610	288360	3.07	76.15
11229	299589	2.97	79.12
10952	310541	2.89	82.01
10533	321074	2.78	84.79
10044	331118	2.65	87.45
9388	340506	2.48	89.93
8837	349343	2.33	92.26
8189	357532	2.16	94.42
7370	364902	1.95	96.37
6158	371060	1.63	97.99
4800	375860	1.27	99.26
2793	378653	0.74	100.00

Figure B.7.4. Spring 2023 STAAR Grade 6 Mathematics
Frequency Distribution of Scale Scores
All Students



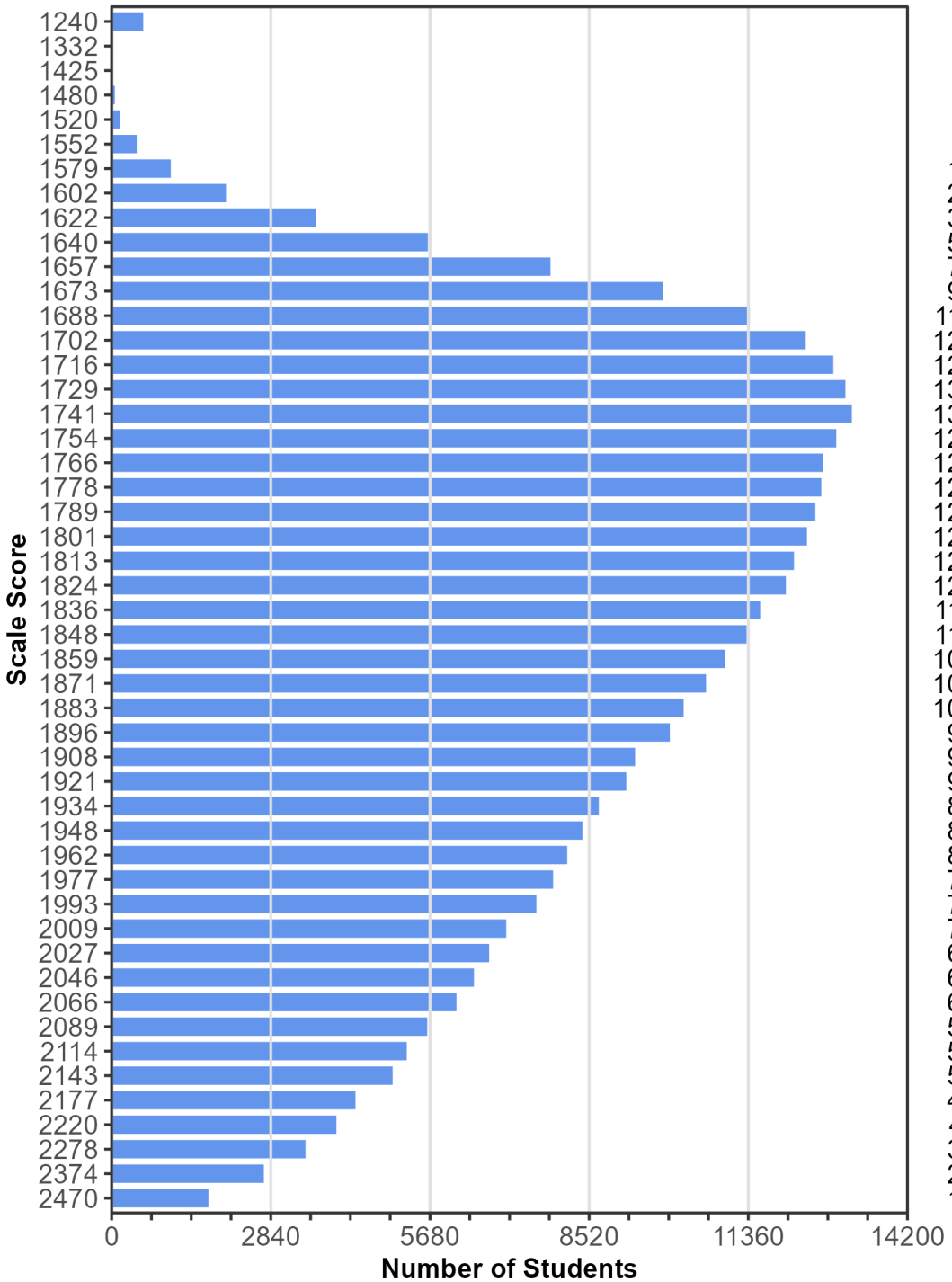
Freq	Cum Freq	Pct	Cum Pct
466	466	0.12	0.12
9	475	0.00	0.12
17	492	0.00	0.13
34	526	0.01	0.14
118	644	0.03	0.17
466	1110	0.12	0.29
1216	2326	0.32	0.60
2685	5011	0.70	1.30
5035	10046	1.31	2.61
8319	18365	2.16	4.77
11817	30182	3.07	7.84
15176	45358	3.94	11.79
17864	63222	4.64	16.43
19035	82257	4.95	21.38
19209	101466	4.99	26.37
18699	120165	4.86	31.23
17918	138083	4.66	35.89
17216	155299	4.47	40.36
16318	171617	4.24	44.60
15280	186897	3.97	48.57
14696	201593	3.82	52.39
13765	215358	3.58	55.97
13088	228446	3.40	59.37
12396	240842	3.22	62.59
11641	252483	3.03	65.62
11211	263694	2.91	68.53
10783	274477	2.80	71.34
10120	284597	2.63	73.97
9688	294285	2.52	76.48
9192	303477	2.39	78.87
8701	312178	2.26	81.13
8498	320676	2.21	83.34
8006	328682	2.08	85.42
7736	336418	2.01	87.43
7232	343650	1.88	89.31
6869	350519	1.79	91.10
6360	356879	1.65	92.75
6195	363074	1.61	94.36
5654	368728	1.47	95.83
5057	373785	1.31	97.15
4316	378101	1.12	98.27
3424	381525	0.89	99.16
2253	383778	0.59	99.74
987	384765	0.26	100.00

Figure B.7.5. Spring 2023 STAAR Grade 7 Mathematics
Frequency Distribution of Scale Scores
All Students



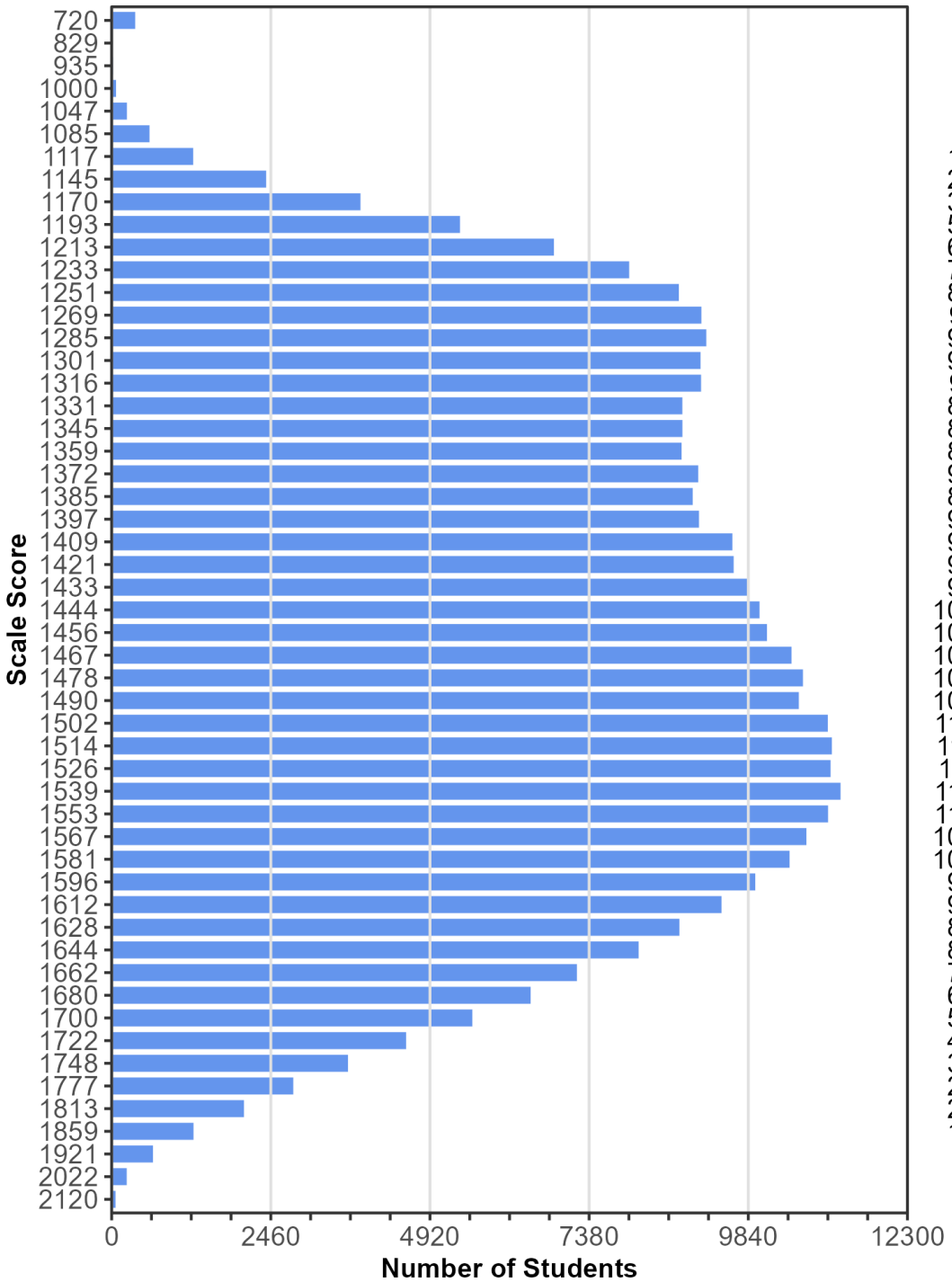
Freq	Cum Freq	Pct	Cum Pct
423	423	0.13	0.13
10	433	0.00	0.13
15	448	0.00	0.14
25	473	0.01	0.14
70	543	0.02	0.16
200	743	0.06	0.22
543	1286	0.16	0.39
1267	2553	0.38	0.77
2555	5108	0.77	1.54
4482	9590	1.35	2.89
6990	16580	2.11	5.00
9526	26106	2.87	7.87
12097	38203	3.65	11.52
13996	52199	4.22	15.74
15266	67465	4.60	20.34
16315	83780	4.92	25.26
16089	99869	4.85	30.11
15706	115575	4.74	34.84
15140	130715	4.56	39.41
14072	144787	4.24	43.65
13695	158482	4.13	47.78
12841	171323	3.87	51.65
12296	183619	3.71	55.36
11618	195237	3.50	58.86
10713	205950	3.23	62.09
10334	216284	3.12	65.21
9930	226214	2.99	68.20
9298	235512	2.80	71.00
8942	244454	2.70	73.70
8282	252736	2.50	76.19
7846	260582	2.37	78.56
7308	267890	2.20	80.76
6837	274727	2.06	82.82
6425	281152	1.94	84.76
6027	287179	1.82	86.58
5482	292661	1.65	88.23
5156	297817	1.55	89.79
4923	302740	1.48	91.27
4496	307236	1.36	92.63
4207	311443	1.27	93.89
3873	315316	1.17	95.06
3718	319034	1.12	96.18
3291	322325	0.99	97.17
3138	325463	0.95	98.12
2711	328174	0.82	98.94
2237	330411	0.67	99.61
1287	331698	0.39	100.00

Figure B.7.6. Spring 2023 STAAR Grade 8 Mathematics
Frequency Distribution of Scale Scores
All Students



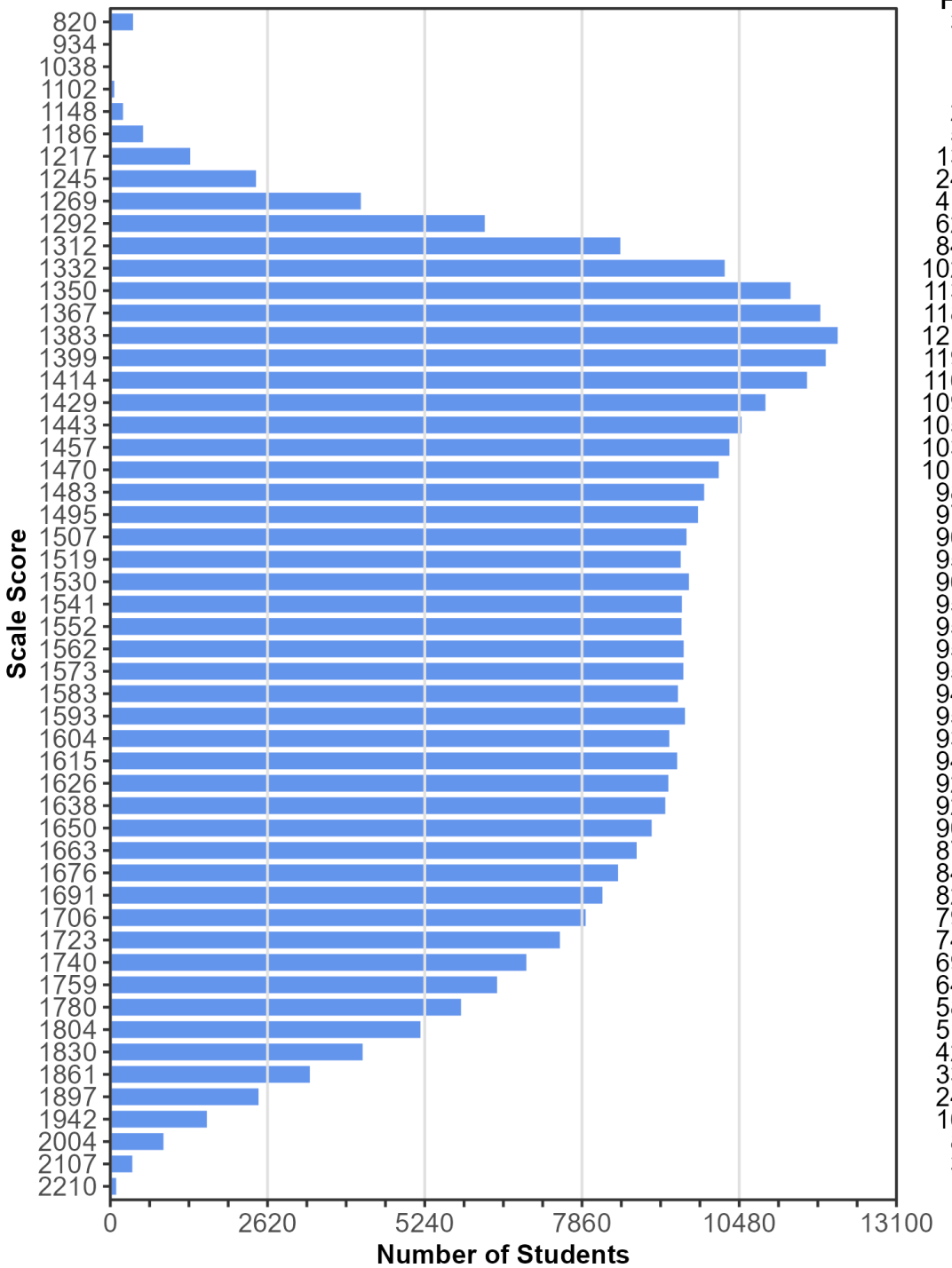
Freq	Cum Freq	Pct	Cum Pct
564	564	0.15	0.15
21	585	0.01	0.16
13	598	0.00	0.16
57	655	0.02	0.18
151	806	0.04	0.22
447	1253	0.12	0.34
1055	2308	0.29	0.63
2040	4348	0.56	1.19
3648	7996	1.00	2.20
5643	13639	1.55	3.75
7830	21469	2.15	5.90
9837	31306	2.70	8.60
11385	42691	3.13	11.72
12384	55075	3.40	15.13
12877	67952	3.54	18.66
13093	81045	3.60	22.26
13208	94253	3.63	25.89
12930	107183	3.55	29.44
12697	119880	3.49	32.92
12663	132543	3.48	36.40
12556	145099	3.45	39.85
12407	157506	3.41	43.26
12175	169681	3.34	46.60
12031	181712	3.30	49.91
11572	193284	3.18	53.08
11329	204613	3.11	56.20
10953	215566	3.01	59.20
10606	226172	2.91	62.12
10206	236378	2.80	64.92
9960	246338	2.74	67.65
9340	255678	2.57	70.22
9183	264861	2.52	72.74
8693	273554	2.39	75.13
8401	281955	2.31	77.44
8129	290084	2.23	79.67
7877	297961	2.16	81.83
7579	305540	2.08	83.91
7041	312581	1.93	85.85
6736	319317	1.85	87.70
6465	325782	1.78	89.47
6153	331935	1.69	91.16
5631	337566	1.55	92.71
5266	342832	1.45	94.16
5014	347846	1.38	95.53
4351	352197	1.19	96.73
4011	356208	1.10	97.83
3459	359667	0.95	98.78
2716	362383	0.75	99.53
1727	364110	0.47	100.00

Figure B.7.7. Spring 2023 STAAR Grade 3 RLA
Frequency Distribution of Scale Scores
All Students



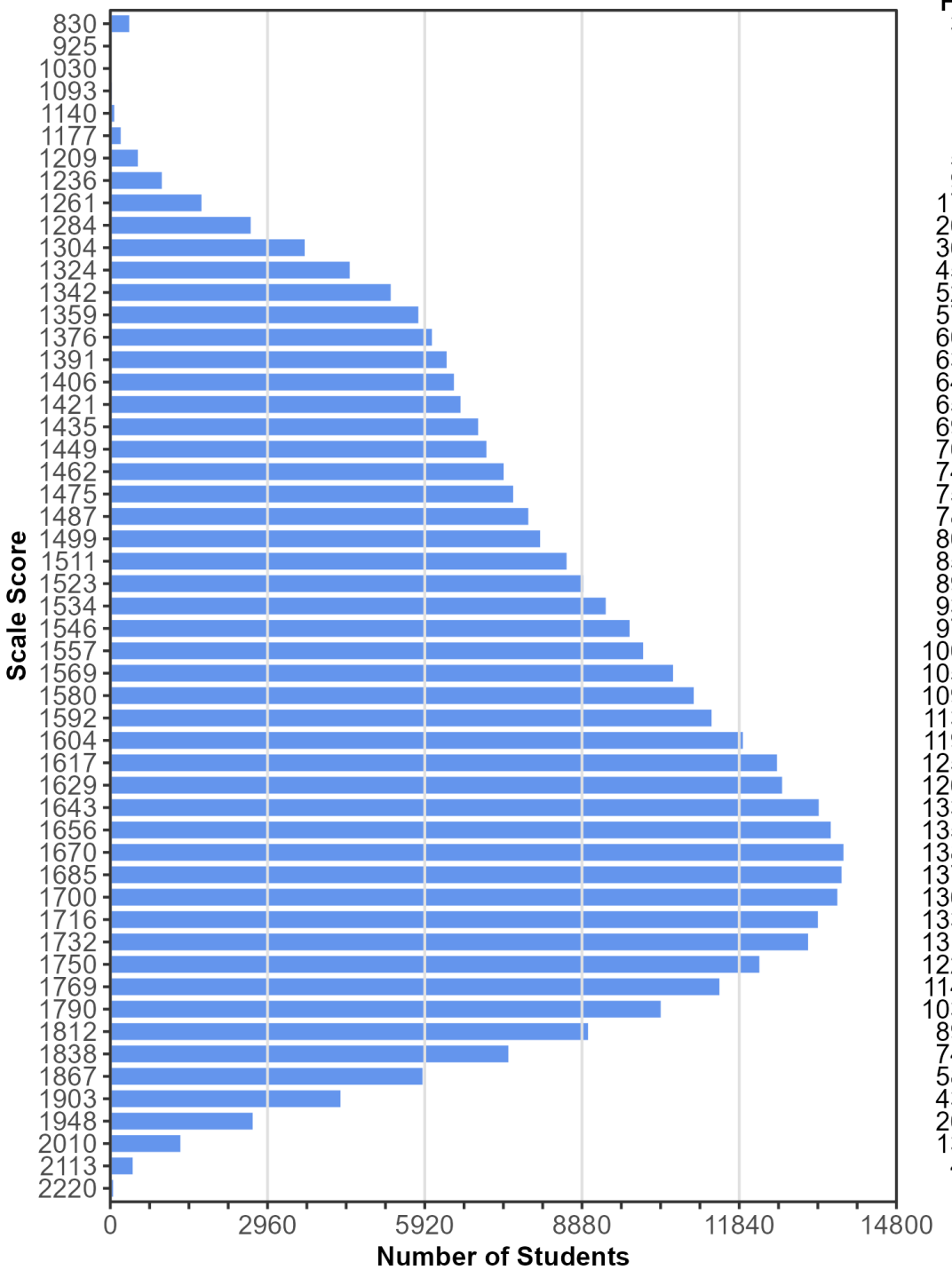
Scale Score	Freq	Cum Freq	Pct	Cum Pct
720	364	364	0.10	0.10
829	7	371	0.00	0.10
935	21	392	0.01	0.11
1000	68	460	0.02	0.13
1047	236	696	0.07	0.20
1085	585	1281	0.16	0.36
1117	1260	2541	0.35	0.71
1145	2388	4929	0.67	1.38
1170	3846	8775	1.08	2.46
1193	5384	14159	1.51	3.97
1213	6836	20995	1.92	5.89
1233	7999	28994	2.24	8.13
1251	8767	37761	2.46	10.59
1269	9115	46876	2.56	13.15
1285	9191	56067	2.58	15.73
1301	9102	65169	2.55	18.28
1316	9110	74279	2.56	20.84
1331	8821	83100	2.47	23.31
1345	8822	91922	2.47	25.78
1359	8809	100731	2.47	28.25
1372	9066	109797	2.54	30.80
1385	8981	118778	2.52	33.32
1397	9078	127856	2.55	35.86
1409	9594	137450	2.69	38.55
1421	9614	147064	2.70	41.25
1433	9831	156895	2.76	44.01
1444	10014	166909	2.81	46.82
1456	10129	177038	2.84	49.66
1467	10508	187546	2.95	52.61
1478	10685	198231	3.00	55.60
1490	10621	208852	2.98	58.58
1502	11069	219921	3.10	61.69
1514	11132	231053	3.12	64.81
1526	11112	242165	3.12	67.93
1539	11266	253431	3.16	71.09
1553	11073	264504	3.11	74.19
1567	10738	275242	3.01	77.20
1581	10478	285720	2.94	80.14
1596	9948	295668	2.79	82.93
1612	9426	305094	2.64	85.58
1628	8775	313869	2.46	88.04
1644	8145	322014	2.28	90.32
1662	7191	329205	2.02	92.34
1680	6476	335681	1.82	94.16
1700	5575	341256	1.56	95.72
1722	4551	345807	1.28	97.00
1748	3653	349460	1.02	98.02
1777	2807	352267	0.79	98.81
1813	2045	354312	0.57	99.38
1859	1265	355577	0.35	99.74
1921	639	356216	0.18	99.92
2022	233	356449	0.07	99.98
2120	60	356509	0.02	100.00

Figure B.7.8. Spring 2023 STAAR Grade 4 RLA
Frequency Distribution of Scale Scores
All Students



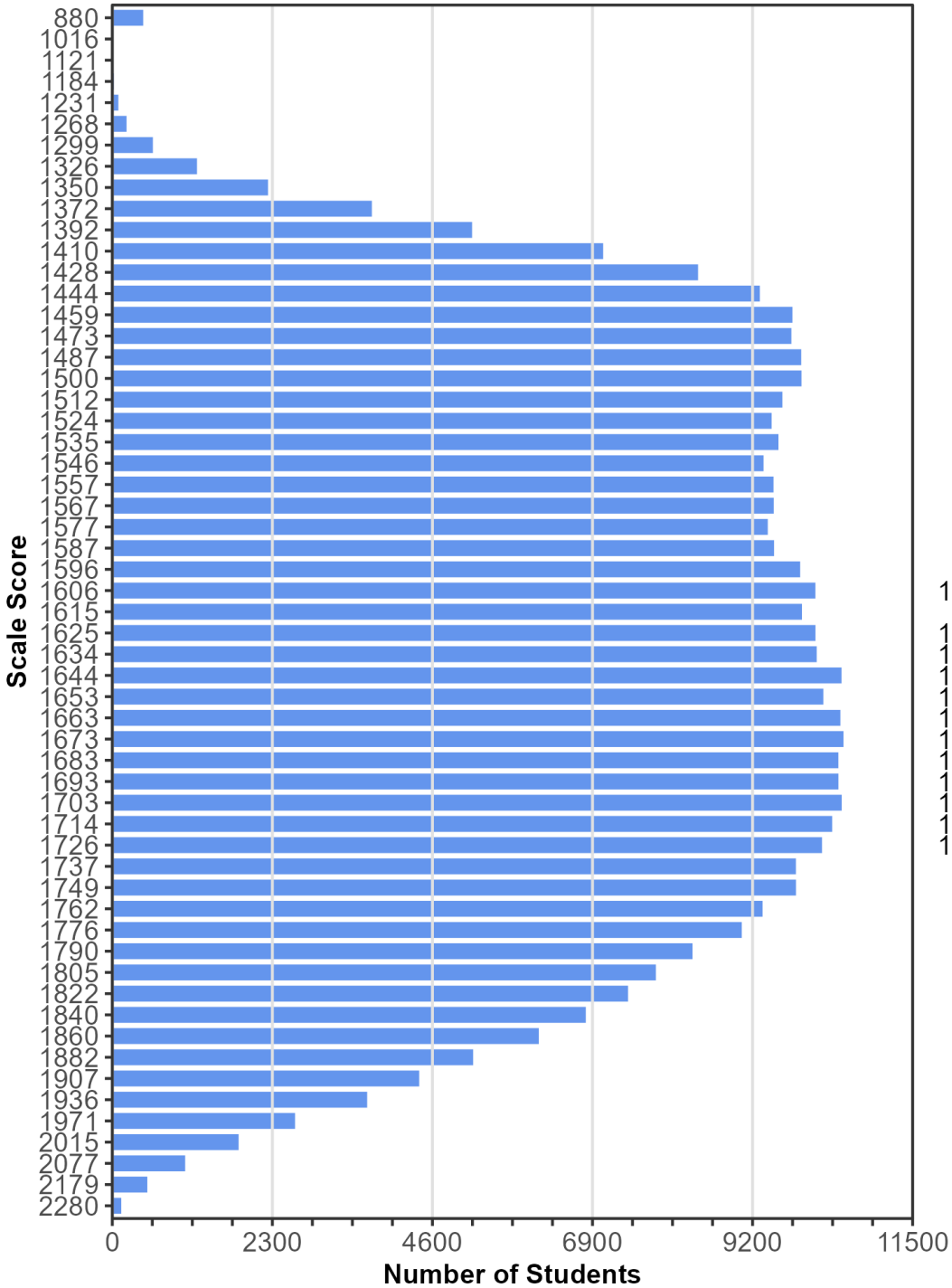
Freq	Cum Freq	Pct	Cum Pct
379	379	0.10	0.10
9	388	0.00	0.11
19	407	0.01	0.11
66	473	0.02	0.13
212	685	0.06	0.19
546	1231	0.15	0.34
1332	2563	0.36	0.70
2428	4991	0.67	1.37
4175	9166	1.14	2.51
6240	15406	1.71	4.22
8499	23905	2.33	6.55
10238	34143	2.80	9.35
11336	45479	3.11	12.46
11832	57311	3.24	15.70
12121	69432	3.32	19.02
11922	81354	3.27	22.29
11610	92964	3.18	25.47
10917	103881	2.99	28.46
10520	114401	2.88	31.34
10317	124718	2.83	34.17
10139	134857	2.78	36.95
9896	144753	2.71	39.66
9794	154547	2.68	42.34
9602	164149	2.63	44.97
9505	173654	2.60	47.58
9642	183296	2.64	50.22
9525	192821	2.61	52.83
9519	202340	2.61	55.44
9554	211894	2.62	58.05
9549	221443	2.62	60.67
9459	230902	2.59	63.26
9576	240478	2.62	65.88
9315	249793	2.55	68.44
9446	259239	2.59	71.02
9299	268538	2.55	73.57
9248	277786	2.53	76.11
9021	286807	2.47	78.58
8771	295578	2.40	80.98
8461	304039	2.32	83.30
8201	312240	2.25	85.54
7921	320161	2.17	87.71
7492	327653	2.05	89.77
6933	334586	1.90	91.67
6445	341031	1.77	93.43
5844	346875	1.60	95.03
5168	352043	1.42	96.45
4205	356248	1.15	97.60
3325	359573	0.91	98.51
2470	362043	0.68	99.19
1609	363652	0.44	99.63
886	364538	0.24	99.87
367	364905	0.10	99.97
98	365003	0.03	100.00

Figure B.7.9. Spring 2023 STAAR Grade 5 RLA
Frequency Distribution of Scale Scores
All Students



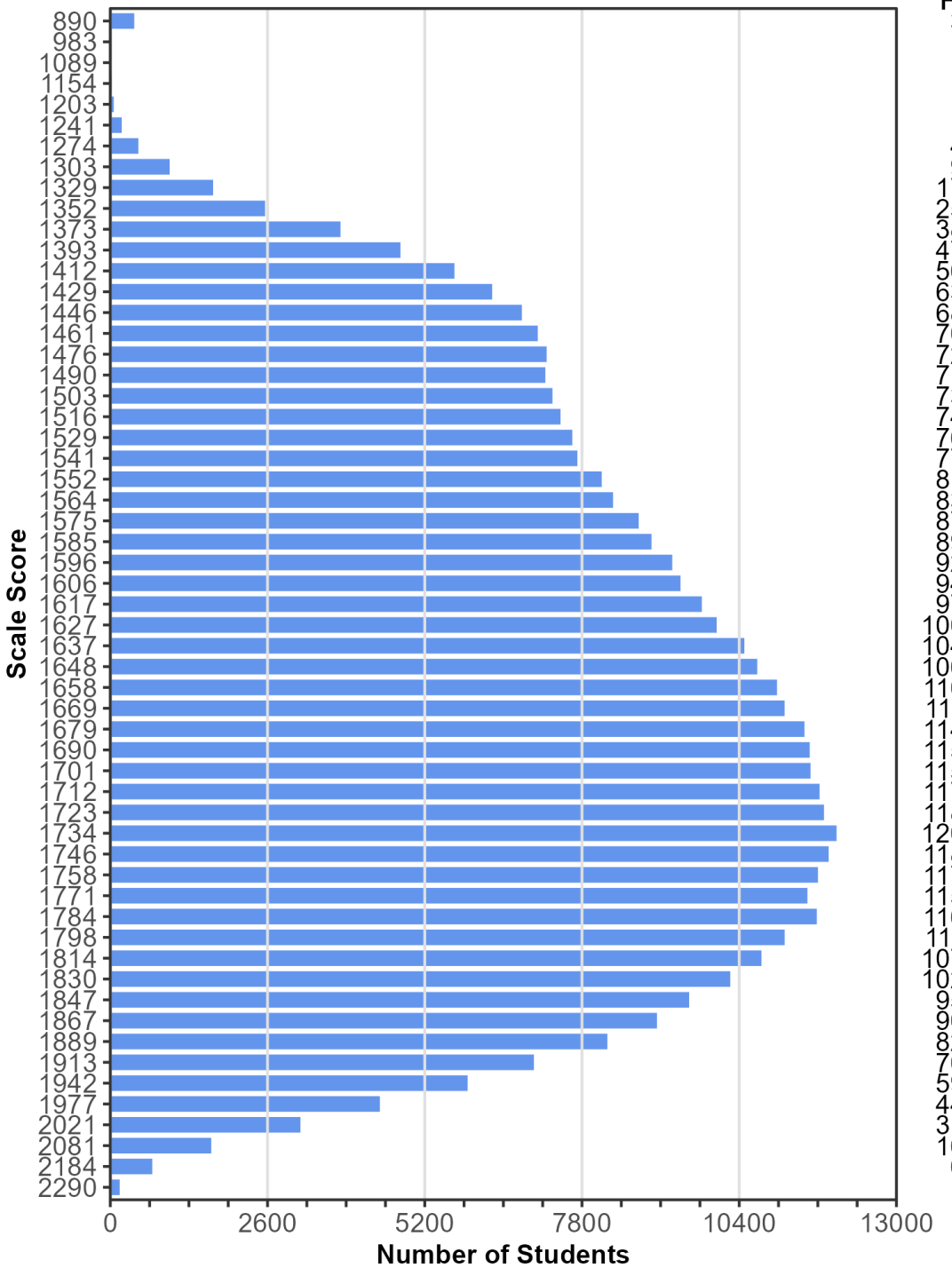
Freq	Cum Freq	Pct	Cum Pct
356	356	0.10	0.10
6	362	0.00	0.10
9	371	0.00	0.10
27	398	0.01	0.11
74	472	0.02	0.13
196	668	0.05	0.18
519	1187	0.14	0.32
970	2157	0.26	0.58
1716	3873	0.46	1.04
2643	6516	0.71	1.75
3660	10176	0.98	2.73
4507	14683	1.21	3.94
5279	19962	1.42	5.36
5799	25761	1.56	6.91
6055	31816	1.62	8.54
6331	38147	1.70	10.24
6469	44616	1.74	11.97
6593	51209	1.77	13.74
6925	58134	1.86	15.60
7081	65215	1.90	17.50
7405	72620	1.99	19.49
7584	80204	2.04	21.52
7869	88073	2.11	23.64
8092	96165	2.17	25.81
8590	104755	2.31	28.11
8910	113665	2.39	30.50
9327	122992	2.50	33.01
9776	132768	2.62	35.63
10031	142799	2.69	38.32
10594	153393	2.84	41.16
10984	164377	2.95	44.11
11319	175696	3.04	47.15
11909	187605	3.20	50.35
12551	200156	3.37	53.71
12647	212803	3.39	57.11
13337	226140	3.58	60.69
13563	239703	3.64	64.33
13803	253506	3.70	68.03
13767	267273	3.69	71.72
13687	280960	3.67	75.40
13321	294281	3.57	78.97
13136	307417	3.53	82.50
12219	319636	3.28	85.78
11467	331103	3.08	88.85
10362	341465	2.78	91.64
8994	350459	2.41	94.05
7494	357953	2.01	96.06
5878	363831	1.58	97.64
4333	368164	1.16	98.80
2680	370844	0.72	99.52
1319	372163	0.35	99.87
421	372584	0.11	99.99
52	372636	0.01	100.00

Figure B.7.10. Spring 2023 STAAR Grade 6 RLA
Frequency Distribution of Scale Scores
All Students



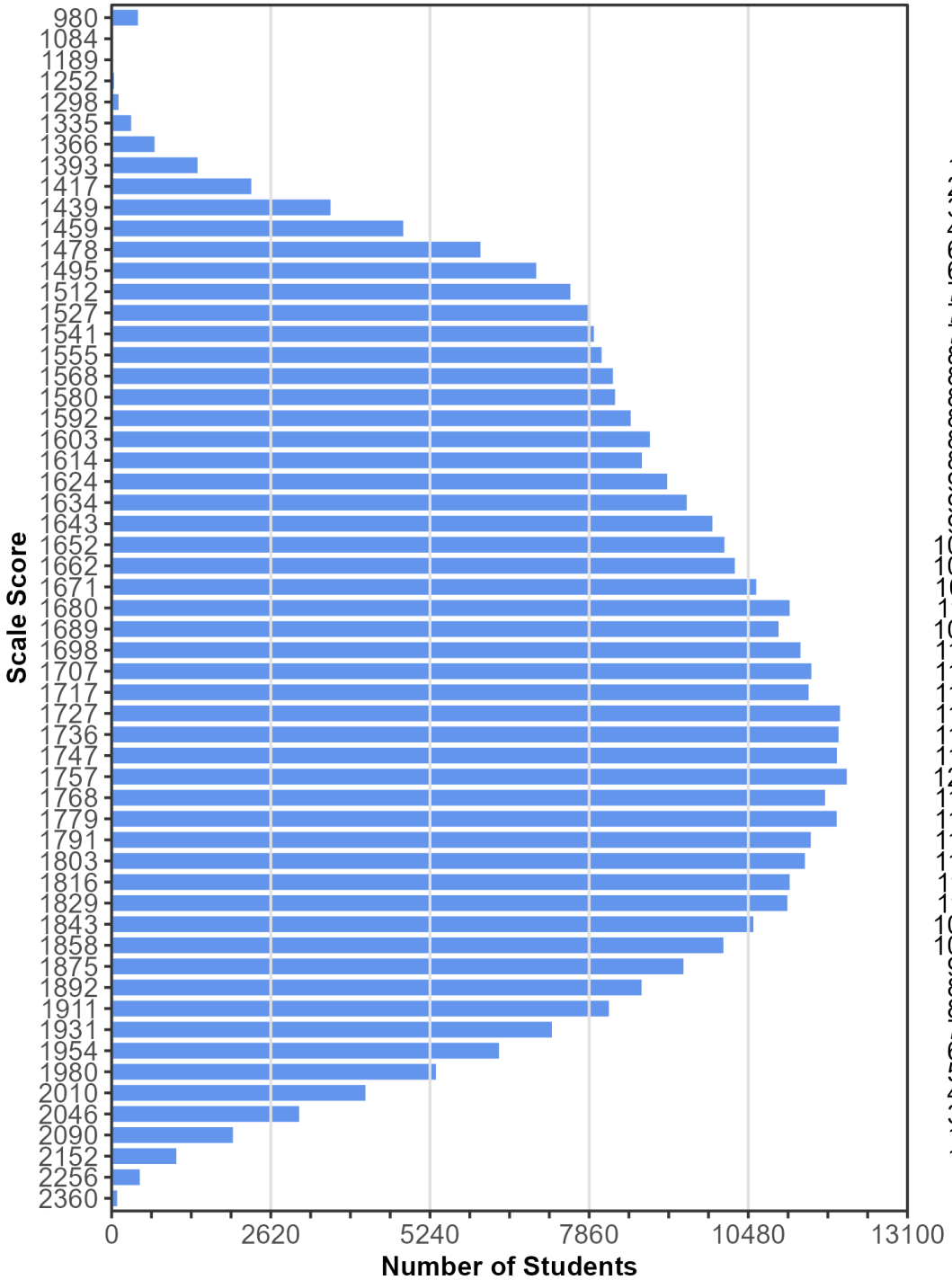
Freq	Cum Freq	Pct	Cum Pct
444	444	0.11	0.11
11	455	0.00	0.12
11	466	0.00	0.12
25	491	0.01	0.13
87	578	0.02	0.15
205	783	0.05	0.20
583	1366	0.15	0.35
1217	2583	0.31	0.66
2238	4821	0.57	1.23
3730	8551	0.95	2.18
5170	13721	1.32	3.51
7053	20774	1.80	5.31
8417	29191	2.15	7.46
9305	38496	2.38	9.84
9773	48269	2.50	12.33
9758	58027	2.49	14.83
9899	67926	2.53	17.36
9903	77829	2.53	19.89
9629	87458	2.46	22.35
9474	96932	2.42	24.77
9572	106504	2.45	27.21
9358	115862	2.39	29.60
9501	125363	2.43	32.03
9504	134867	2.43	34.46
9419	144286	2.41	36.87
9509	153795	2.43	39.30
9884	163679	2.53	41.82
10103	173782	2.58	44.40
9910	183692	2.53	46.93
10104	193796	2.58	49.52
10122	203918	2.59	52.10
10478	214396	2.68	54.78
10217	224613	2.61	57.39
10462	235075	2.67	60.06
10507	245582	2.68	62.75
10433	256015	2.67	65.41
10434	266449	2.67	68.08
10481	276930	2.68	70.76
10344	287274	2.64	73.40
10198	297472	2.61	76.01
9823	307295	2.51	78.52
9824	317119	2.51	81.03
9344	326463	2.39	83.41
9045	335508	2.31	85.73
8337	343845	2.13	87.86
7811	351656	2.00	89.85
7411	359067	1.89	91.74
6804	365871	1.74	93.48
6128	371999	1.57	95.05
5185	377184	1.32	96.37
4410	381594	1.13	97.50
3662	385256	0.94	98.44
2626	387882	0.67	99.11
1815	389697	0.46	99.57
1047	390744	0.27	99.84
503	391247	0.13	99.97
129	391376	0.03	100.00

Figure B.7.11. Spring 2023 STAAR Grade 7 RLA
Frequency Distribution of Scale Scores
All Students



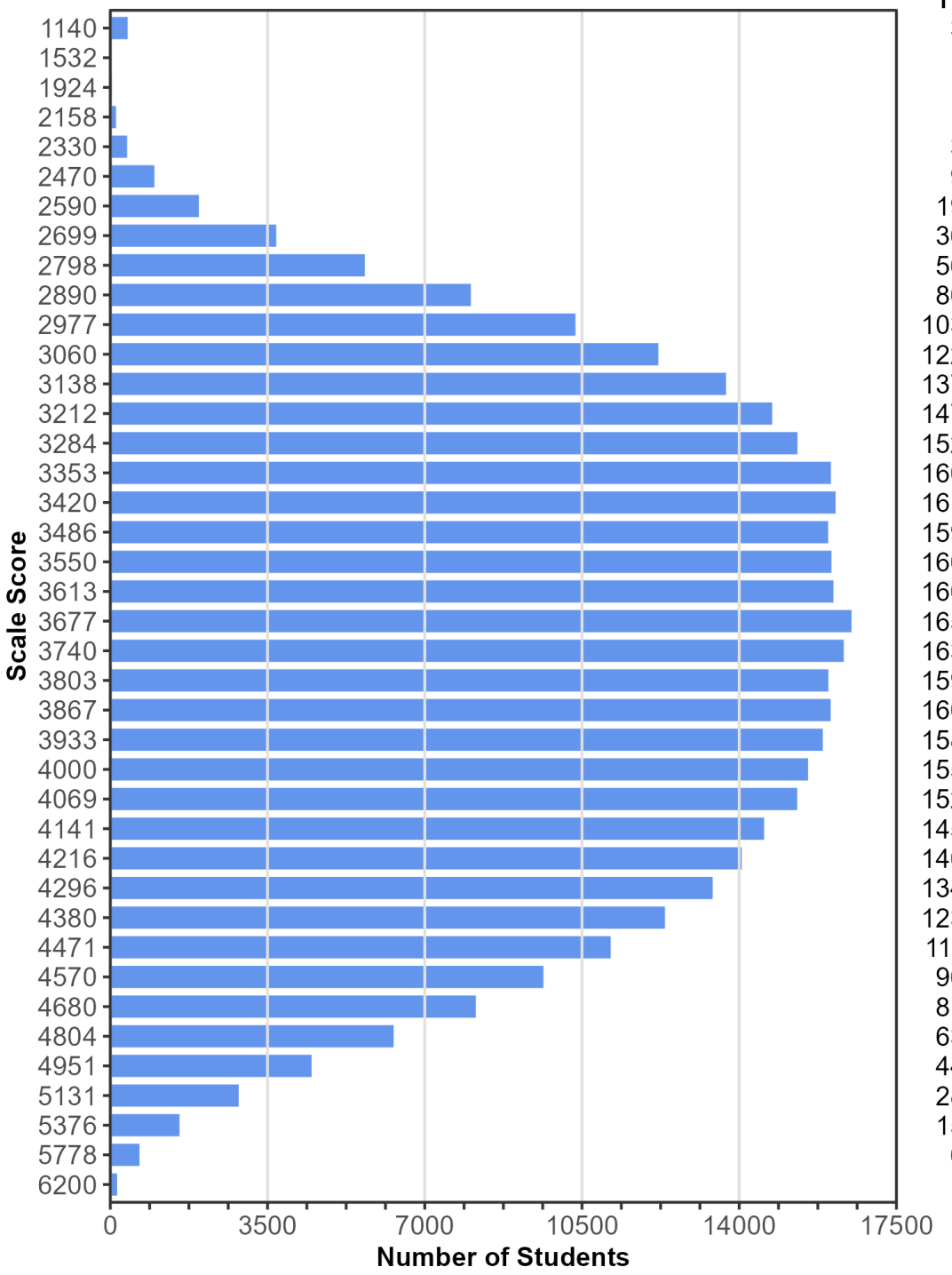
Freq	Cum Freq	Pct	Cum Pct
396	396	0.10	0.10
5	401	0.00	0.10
10	411	0.00	0.10
24	435	0.01	0.11
57	492	0.01	0.12
189	681	0.05	0.17
464	1145	0.12	0.29
981	2126	0.24	0.53
1700	3826	0.42	0.96
2560	6386	0.64	1.59
3807	10193	0.95	2.55
4798	14991	1.20	3.74
5691	20682	1.42	5.17
6315	26997	1.58	6.74
6806	33803	1.70	8.44
7069	40872	1.77	10.21
7216	48088	1.80	12.01
7194	55282	1.80	13.81
7312	62594	1.83	15.63
7445	70039	1.86	17.49
7640	77679	1.91	19.40
7724	85403	1.93	21.33
8127	93530	2.03	23.36
8314	101844	2.08	25.43
8738	110582	2.18	27.62
8950	119532	2.24	29.85
9293	128825	2.32	32.17
9429	138254	2.35	34.53
9782	148036	2.44	36.97
10027	158063	2.50	39.47
10485	168548	2.62	42.09
10698	179246	2.67	44.76
11025	190271	2.75	47.52
11151	201422	2.78	50.30
11479	212901	2.87	53.17
11566	224467	2.89	56.06
11579	236046	2.89	58.95
11728	247774	2.93	61.88
11802	259576	2.95	64.83
12010	271586	3.00	67.83
11879	283465	2.97	70.79
11702	295167	2.92	73.72
11528	306695	2.88	76.59
11683	318378	2.92	79.51
11152	329530	2.79	82.30
10768	340298	2.69	84.99
10252	350550	2.56	87.55
9572	360122	2.39	89.94
9040	369162	2.26	92.19
8220	377382	2.05	94.25
7005	384387	1.75	96.00
5909	390296	1.48	97.47
4458	394754	1.11	98.59
3144	397898	0.79	99.37
1669	399567	0.42	99.79
694	400261	0.17	99.96
155	400416	0.04	100.00

Figure B.7.12. Spring 2023 STAAR Grade 8 RLA
Frequency Distribution of Scale Scores
All Students



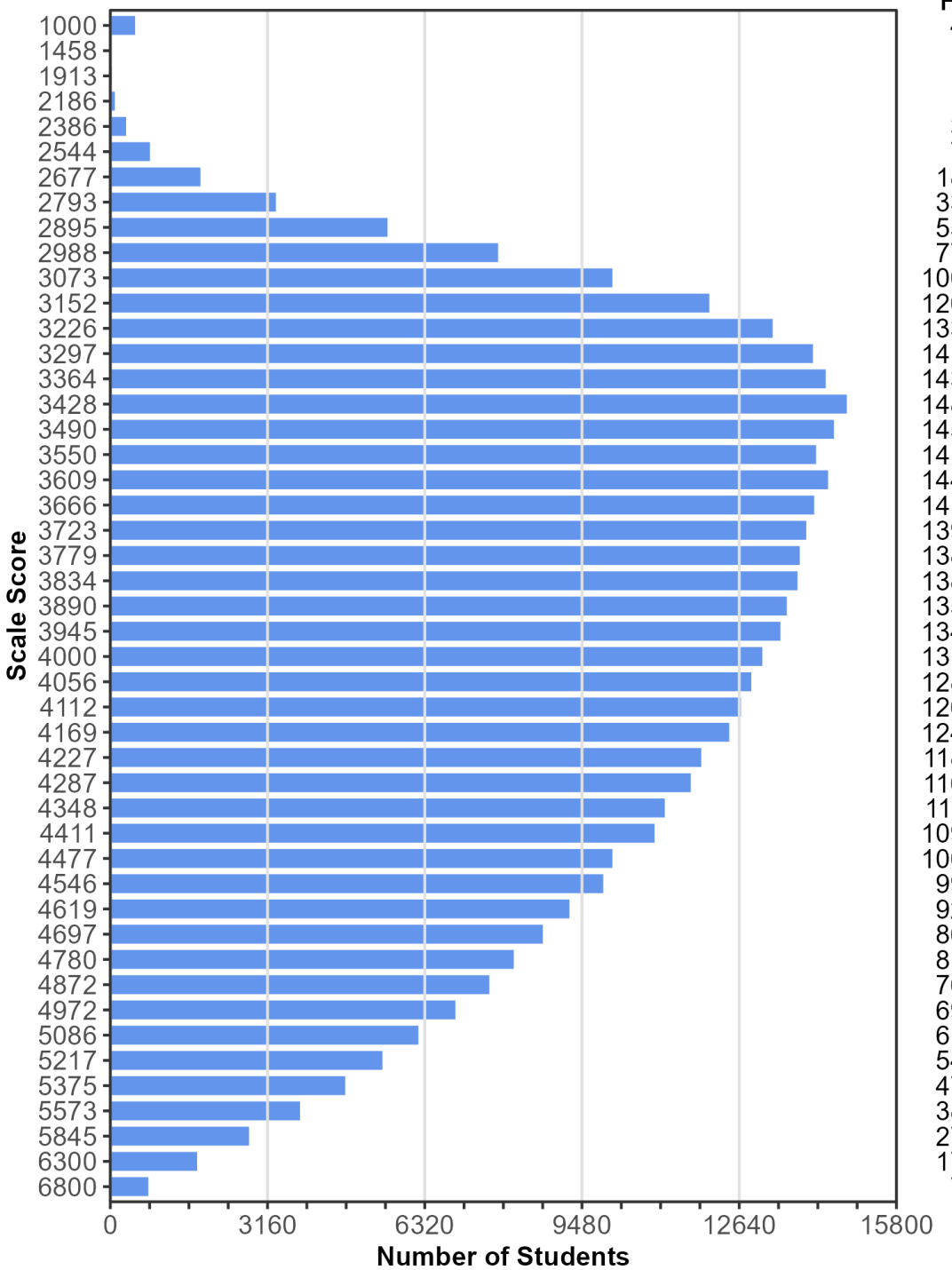
Freq	Cum Freq	Pct	Cum Pct
433	433	0.11	0.11
7	440	0.00	0.11
10	450	0.00	0.11
39	489	0.01	0.12
113	602	0.03	0.15
321	923	0.08	0.22
706	1629	0.17	0.40
1413	3042	0.34	0.74
2298	5340	0.56	1.30
3603	8943	0.88	2.18
4800	13743	1.17	3.35
6071	19814	1.48	4.83
6989	26803	1.70	6.53
7551	34354	1.84	8.37
7841	42195	1.91	10.28
7938	50133	1.93	12.21
8065	58198	1.96	14.18
8252	66450	2.01	16.19
8287	74737	2.02	18.21
8544	83281	2.08	20.29
8860	92141	2.16	22.45
8729	100870	2.13	24.57
9145	110015	2.23	26.80
9467	119482	2.31	29.11
9889	129371	2.41	31.52
10087	139458	2.46	33.98
10258	149716	2.50	36.47
10611	160327	2.59	39.06
11161	171488	2.72	41.78
10979	182467	2.67	44.45
11339	193806	2.76	47.22
11518	205324	2.81	50.02
11473	216797	2.80	52.82
11989	228786	2.92	55.74
11966	240752	2.92	58.65
11938	252690	2.91	61.56
12101	264791	2.95	64.51
11744	276535	2.86	67.37
11935	288470	2.91	70.28
11508	299978	2.80	73.08
11412	311390	2.78	75.86
11161	322551	2.72	78.58
11123	333674	2.71	81.29
10564	344238	2.57	83.86
10070	354308	2.45	86.32
9413	363721	2.29	88.61
8723	372444	2.13	90.74
8185	380629	1.99	92.73
7247	387876	1.77	94.50
6376	394252	1.55	96.05
5338	399590	1.30	97.35
4178	403768	1.02	98.37
3086	406854	0.75	99.12
1996	408850	0.49	99.61
1065	409915	0.26	99.87
463	410378	0.11	99.98
90	410468	0.02	100.00

Figure B.7.13. Spring 2023 STAAR Grade 5 Science
Frequency Distribution of Scale Scores
All Students



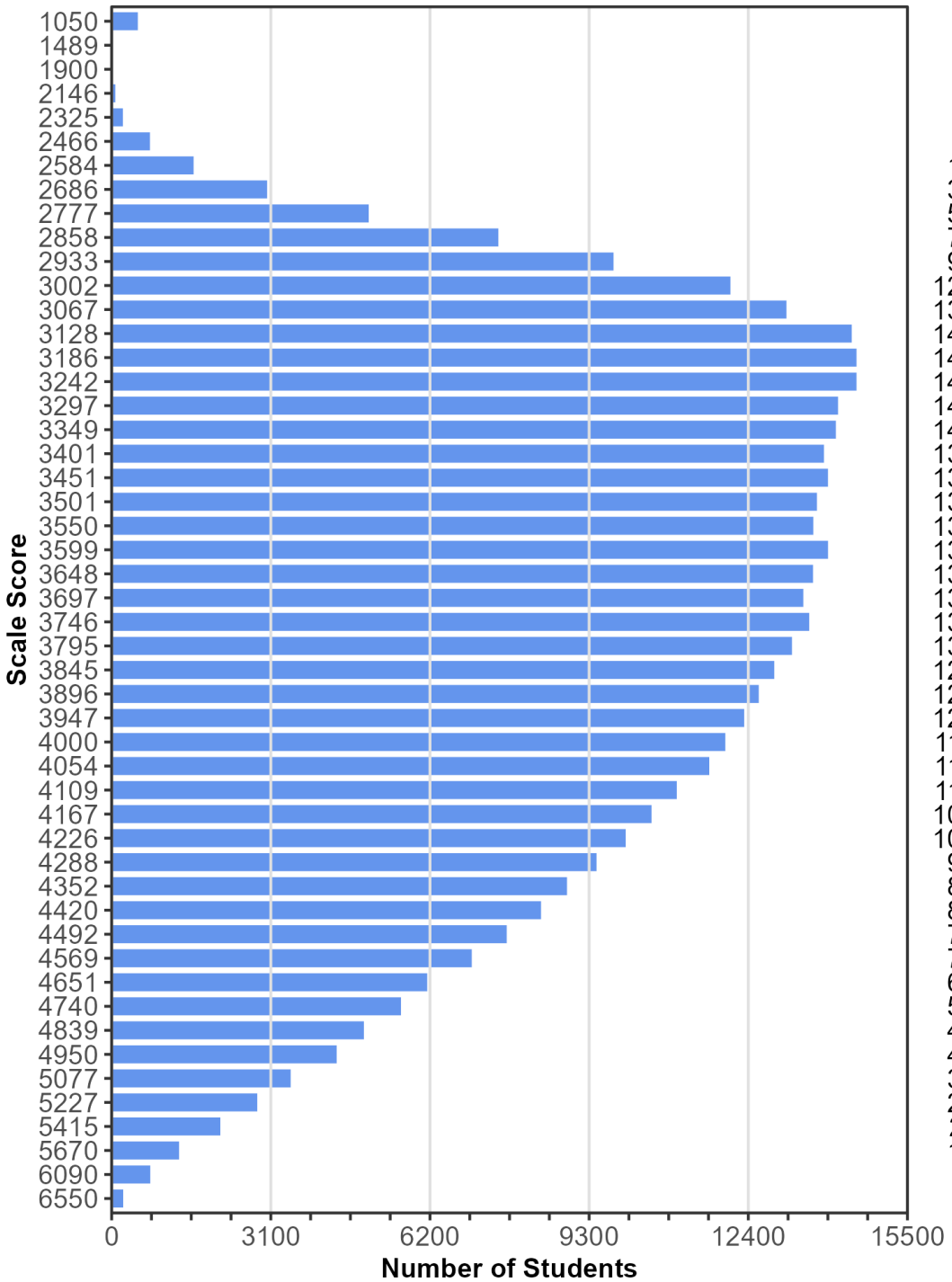
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1140	387	387	0.10	0.10
1532	12	399	0.00	0.11
1924	34	433	0.01	0.11
2158	128	561	0.03	0.15
2330	374	935	0.10	0.25
2470	982	1917	0.26	0.51
2590	1972	3889	0.52	1.03
2699	3694	7583	0.98	2.00
2798	5667	13250	1.50	3.50
2890	8026	21276	2.12	5.62
2977	10355	31631	2.73	8.35
3060	12201	43832	3.22	11.57
3138	13707	57539	3.62	15.19
3212	14734	72273	3.89	19.08
3284	15297	87570	4.04	23.12
3353	16041	103611	4.24	27.36
3420	16146	119757	4.26	31.62
3486	15980	135737	4.22	35.84
3550	16054	151791	4.24	40.08
3613	16098	167889	4.25	44.33
3677	16500	184389	4.36	48.69
3740	16328	200717	4.31	53.00
3803	15984	216701	4.22	57.22
3867	16036	232737	4.23	61.46
3933	15861	248598	4.19	65.65
4000	15533	264131	4.10	69.75
4069	15291	279422	4.04	73.79
4141	14554	293976	3.84	77.63
4216	14056	308032	3.71	81.34
4296	13410	321442	3.54	84.88
4380	12347	333789	3.26	88.14
4471	11139	344928	2.94	91.08
4570	9642	354570	2.55	93.63
4680	8137	362707	2.15	95.78
4804	6308	369015	1.67	97.44
4951	4480	373495	1.18	98.63
5131	2858	376353	0.75	99.38
5376	1541	377894	0.41	99.79
5778	650	378544	0.17	99.96
6200	152	378696	0.04	100.00

Figure B.7.14. Spring 2023 STAAR Grade 8 Science
Frequency Distribution of Scale Scores
All Students



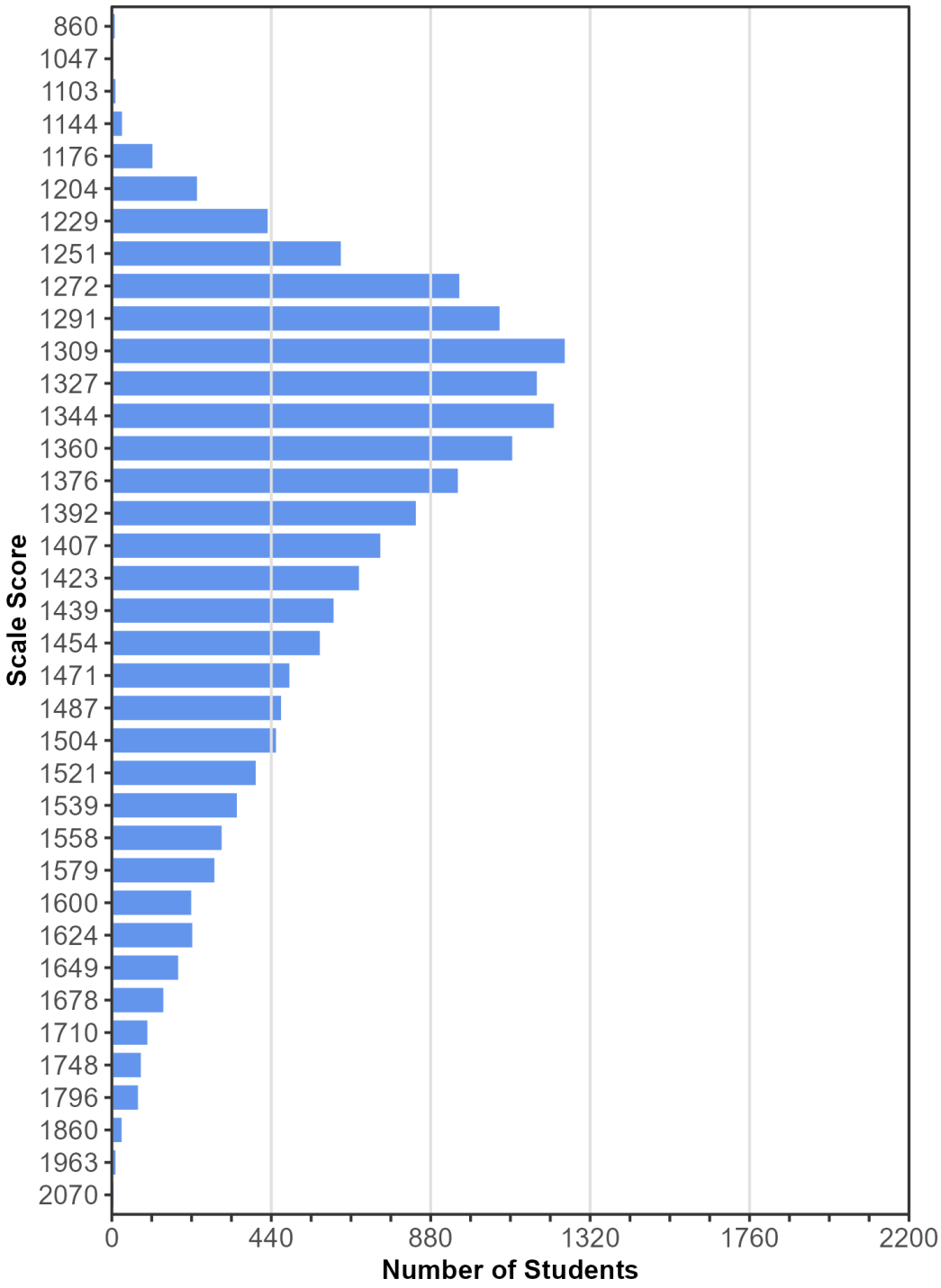
Freq	Cum Freq	Pct	Cum Pct
499	499	0.12	0.12
17	516	0.00	0.13
33	549	0.01	0.13
91	640	0.02	0.16
318	958	0.08	0.23
797	1755	0.20	0.43
1813	3568	0.44	0.87
3329	6897	0.82	1.69
5572	12469	1.37	3.06
7794	20263	1.91	4.97
10092	30355	2.47	7.44
12039	42394	2.95	10.39
13313	55707	3.26	13.66
14122	69829	3.46	17.12
14380	84209	3.53	20.65
14803	99012	3.63	24.28
14544	113556	3.57	27.84
14185	127741	3.48	31.32
14426	142167	3.54	34.86
14147	156314	3.47	38.33
13987	170301	3.43	41.76
13857	184158	3.40	45.15
13813	197971	3.39	48.54
13597	211568	3.33	51.87
13468	225036	3.30	55.18
13105	238141	3.21	58.39
12883	251024	3.16	61.55
12678	263702	3.11	64.66
12440	276142	3.05	67.71
11877	288019	2.91	70.62
11667	299686	2.86	73.48
11144	310830	2.73	76.21
10939	321769	2.68	78.89
10092	331861	2.47	81.37
9907	341768	2.43	83.80
9228	350996	2.26	86.06
8693	359689	2.13	88.19
8109	367798	1.99	90.18
7618	375416	1.87	92.05
6936	382352	1.70	93.75
6192	388544	1.52	95.27
5469	394013	1.34	96.61
4722	398735	1.16	97.77
3813	402548	0.93	98.70
2789	405337	0.68	99.39
1744	407081	0.43	99.81
764	407845	0.19	100.00

Figure B.7.15. Spring 2023 STAAR Grade 8 Social Studies
Frequency Distribution of Scale Scores
All Students



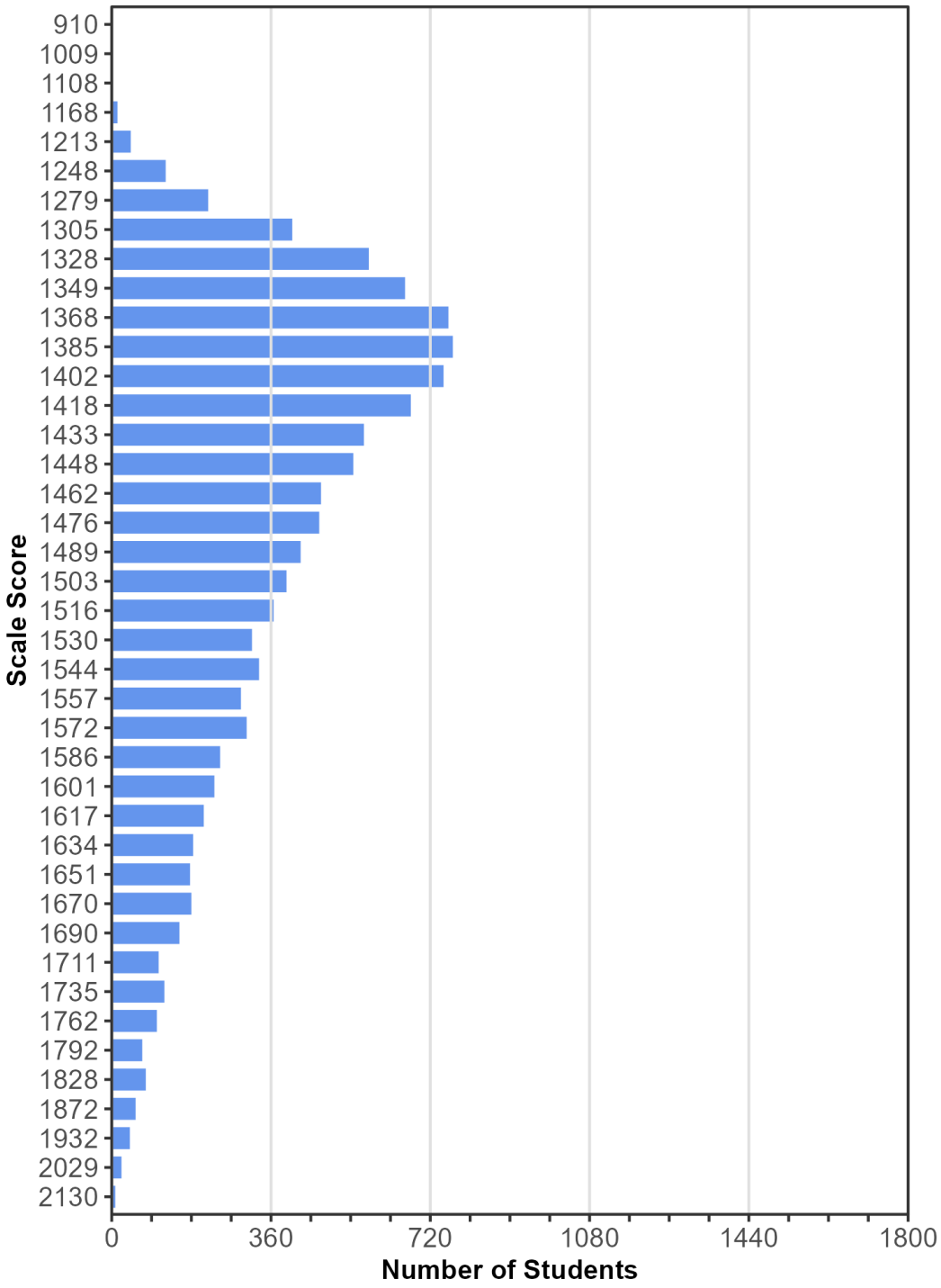
Freq	Cum Freq	Pct	Cum Pct
510	510	0.12	0.12
9	519	0.00	0.13
17	536	0.00	0.13
72	608	0.02	0.15
219	827	0.05	0.20
746	1573	0.18	0.38
1596	3169	0.38	0.76
3028	6197	0.73	1.49
5007	11204	1.21	2.70
7532	18736	1.82	4.52
9775	28511	2.36	6.88
12052	40563	2.91	9.78
13143	53706	3.17	12.95
14415	68121	3.48	16.43
14508	82629	3.50	19.93
14508	97137	3.50	23.42
14148	111285	3.41	26.84
14106	125391	3.40	30.24
13874	139265	3.35	33.58
13952	153217	3.36	36.95
13737	166954	3.31	40.26
13666	180620	3.30	43.56
13953	194573	3.36	46.92
13661	208234	3.29	50.21
13472	221706	3.25	53.46
13587	235293	3.28	56.74
13251	248544	3.20	59.93
12906	261450	3.11	63.05
12604	274054	3.04	66.09
12320	286374	2.97	69.06
11952	298326	2.88	71.94
11640	309966	2.81	74.75
11007	320973	2.65	77.40
10515	331488	2.54	79.94
10012	341500	2.41	82.35
9444	350944	2.28	84.63
8869	359813	2.14	86.77
8363	368176	2.02	88.78
7695	375871	1.86	90.64
7014	382885	1.69	92.33
6146	389031	1.48	93.81
5635	394666	1.36	95.17
4913	399579	1.18	96.36
4383	403962	1.06	97.41
3487	407449	0.84	98.25
2836	410285	0.68	98.94
2117	412402	0.51	99.45
1313	413715	0.32	99.76
752	414467	0.18	99.95
225	414692	0.05	100.00

Figure B.7.16. Spring 2023 STAAR Spanish Grade 3 Mathematics
Frequency Distribution of Scale Scores
All Students



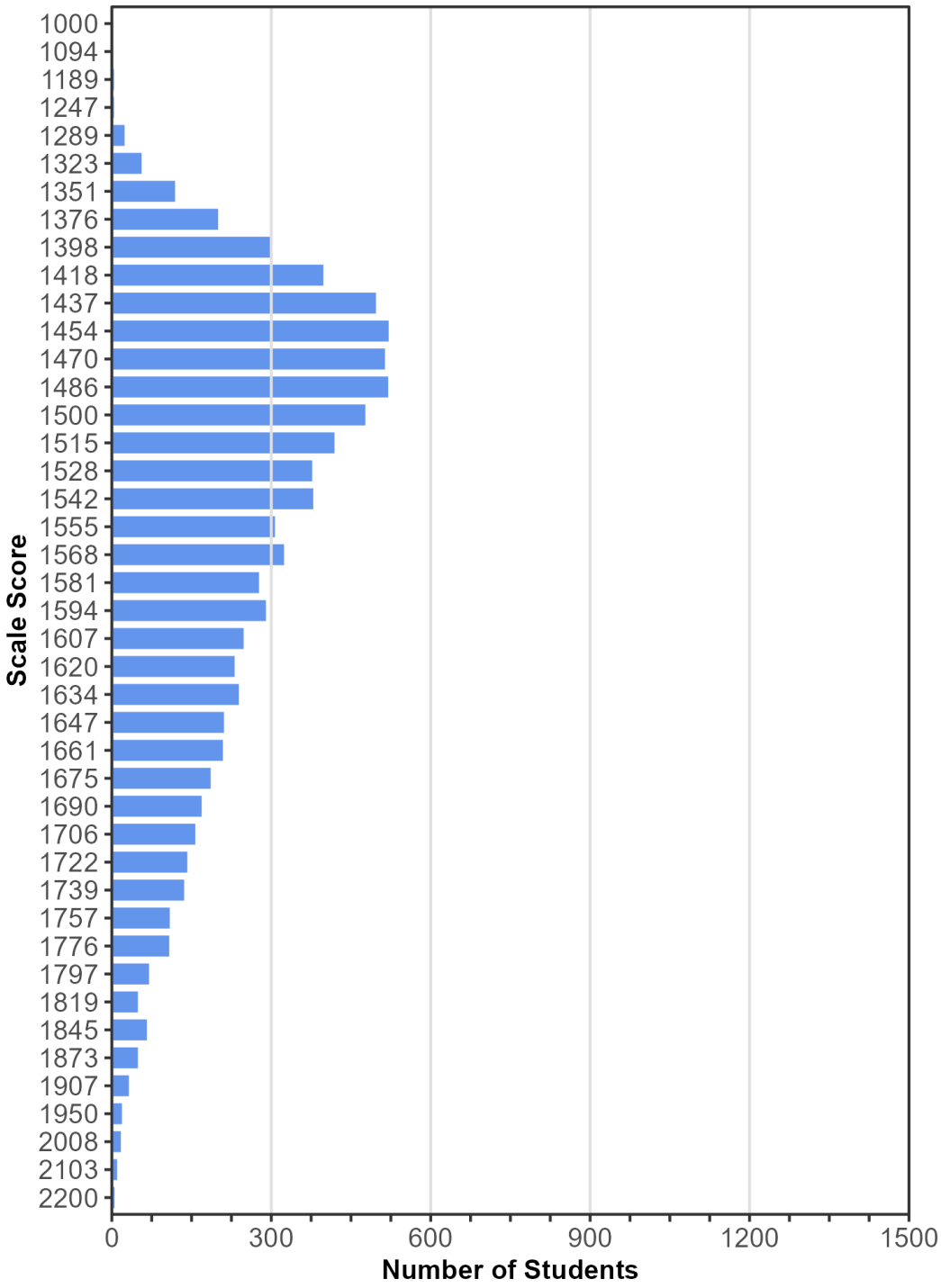
Scale Score	Freq	Cum Freq	Pct	Cum Pct
860	8	8	0.05	0.05
1047	4	12	0.02	0.07
1103	10	22	0.06	0.13
1144	28	50	0.17	0.30
1176	112	162	0.68	0.99
1204	235	397	1.43	2.42
1229	430	827	2.62	5.03
1251	632	1459	3.85	8.88
1272	959	2418	5.84	14.71
1291	1070	3488	6.51	21.22
1309	1250	4738	7.61	28.83
1327	1173	5911	7.14	35.97
1344	1220	7131	7.42	43.39
1360	1105	8236	6.72	50.11
1376	955	9191	5.81	55.92
1392	839	10030	5.10	61.03
1407	741	10771	4.51	65.54
1423	682	11453	4.15	69.69
1439	612	12065	3.72	73.41
1454	574	12639	3.49	76.90
1471	490	13129	2.98	79.88
1487	467	13596	2.84	82.73
1504	453	14049	2.76	85.48
1521	397	14446	2.42	87.90
1539	345	14791	2.10	90.00
1558	303	15094	1.84	91.84
1579	283	15377	1.72	93.56
1600	219	15596	1.33	94.90
1624	222	15818	1.35	96.25
1649	183	16001	1.11	97.36
1678	142	16143	0.86	98.22
1710	98	16241	0.60	98.82
1748	80	16321	0.49	99.31
1796	72	16393	0.44	99.74
1860	27	16420	0.16	99.91
1963	10	16430	0.06	99.97
2070	5	16435	0.03	100.00

Figure B.7.17. Spring 2023 STAAR Spanish Grade 4 Mathematics
Frequency Distribution of Scale Scores
All Students



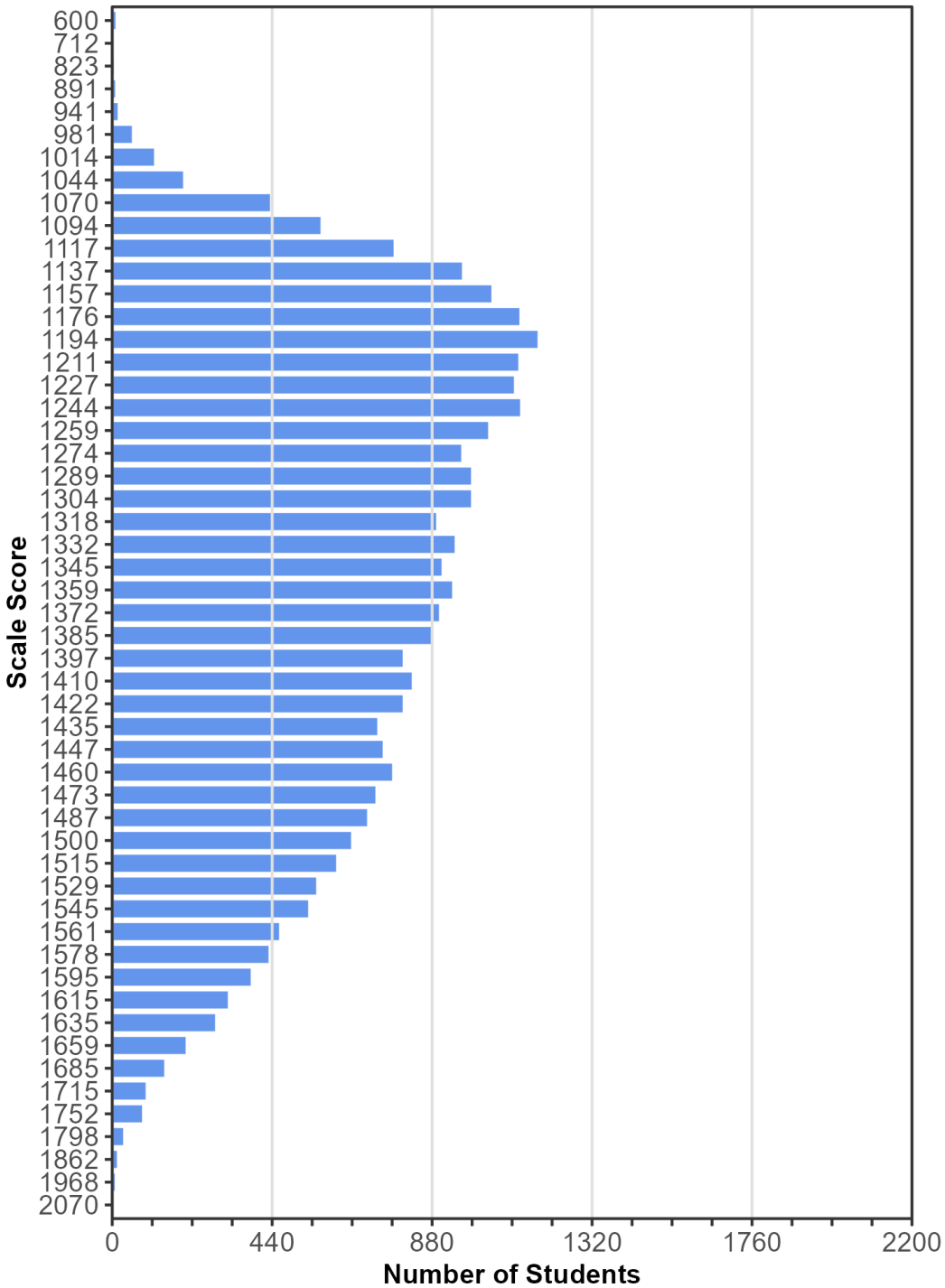
Scale Score	Freq	Cum Freq	Pct	Cum Pct
910	1	1	0.01	0.01
1009	1	2	0.01	0.02
1108	1	3	0.01	0.03
1168	13	16	0.11	0.14
1213	43	59	0.37	0.51
1248	122	181	1.06	1.58
1279	218	399	1.90	3.48
1305	408	807	3.55	7.03
1328	581	1388	5.06	12.09
1349	663	2051	5.78	17.87
1368	761	2812	6.63	24.50
1385	771	3583	6.72	31.21
1402	750	4333	6.53	37.75
1418	676	5009	5.89	43.64
1433	570	5579	4.97	48.60
1448	546	6125	4.76	53.36
1462	473	6598	4.12	57.48
1476	469	7067	4.09	61.56
1489	427	7494	3.72	65.28
1503	395	7889	3.44	68.73
1516	366	8255	3.19	71.91
1530	317	8572	2.76	74.68
1544	333	8905	2.90	77.58
1557	292	9197	2.54	80.12
1572	305	9502	2.66	82.78
1586	245	9747	2.13	84.91
1601	232	9979	2.02	86.93
1617	208	10187	1.81	88.74
1634	184	10371	1.60	90.35
1651	177	10548	1.54	91.89
1670	180	10728	1.57	93.46
1690	153	10881	1.33	94.79
1711	106	10987	0.92	95.71
1735	119	11106	1.04	96.75
1762	102	11208	0.89	97.64
1792	69	11277	0.60	98.24
1828	77	11354	0.67	98.91
1872	54	11408	0.47	99.38
1932	41	11449	0.36	99.74
2029	22	11471	0.19	99.93
2130	8	11479	0.07	100.00

Figure B.7.18. Spring 2023 STAAR Spanish Grade 5 Mathematics
Frequency Distribution of Scale Scores
All Students



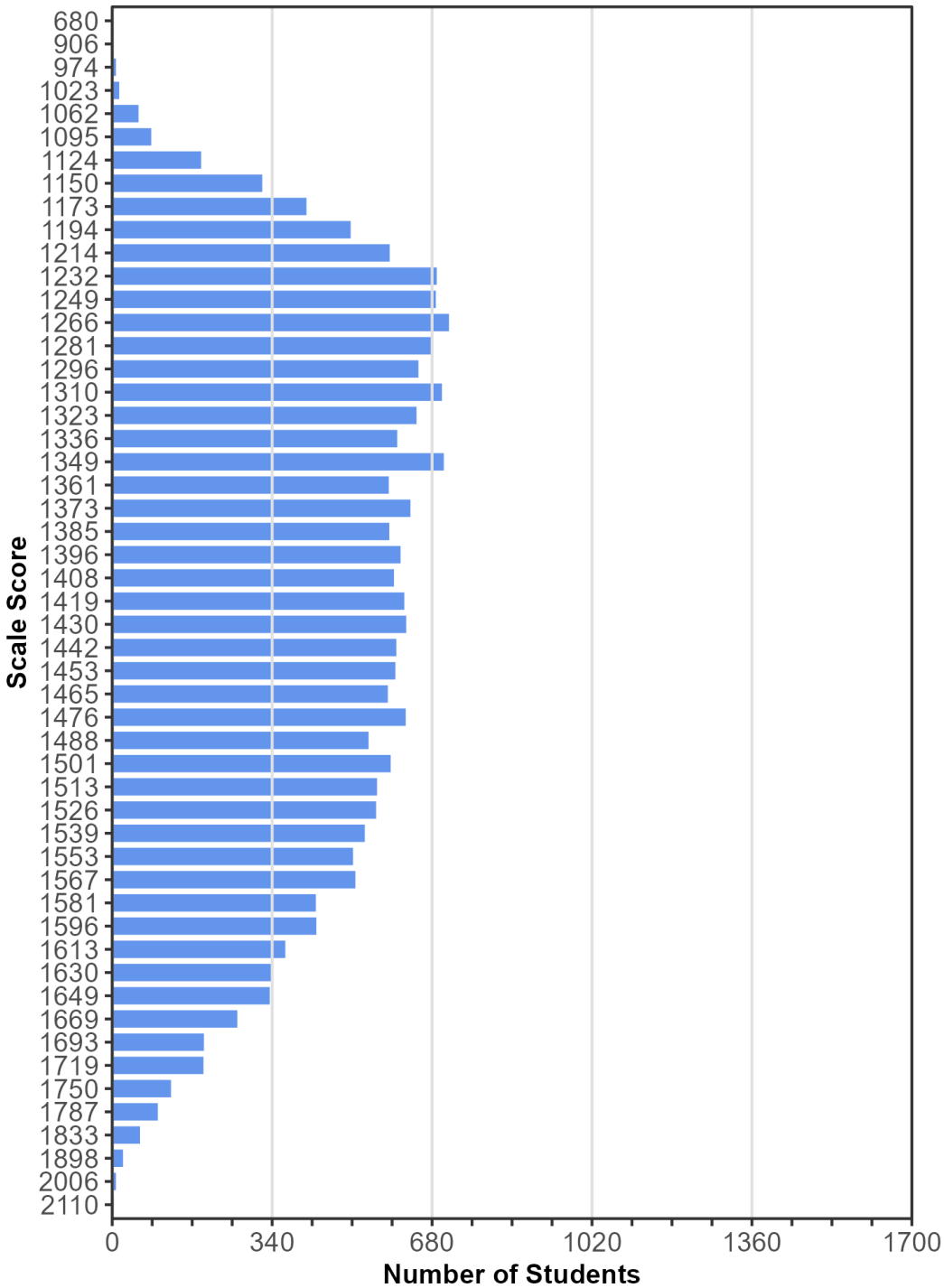
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1000	1	1	0.01	0.01
1094	1	2	0.01	0.02
1189	4	6	0.05	0.07
1247	4	10	0.05	0.12
1289	24	34	0.28	0.40
1323	56	90	0.66	1.06
1351	119	209	1.40	2.47
1376	200	409	2.36	4.83
1398	302	711	3.56	8.39
1418	398	1109	4.70	13.09
1437	497	1606	5.87	18.95
1454	521	2127	6.15	25.10
1470	514	2641	6.07	31.17
1486	520	3161	6.14	37.31
1500	477	3638	5.63	42.94
1515	419	4057	4.95	47.88
1528	377	4434	4.45	52.33
1542	379	4813	4.47	56.80
1555	307	5120	3.62	60.43
1568	324	5444	3.82	64.25
1581	277	5721	3.27	67.52
1594	290	6011	3.42	70.94
1607	248	6259	2.93	73.87
1620	231	6490	2.73	76.60
1634	239	6729	2.82	79.42
1647	211	6940	2.49	81.91
1661	209	7149	2.47	84.37
1675	186	7335	2.20	86.57
1690	169	7504	1.99	88.56
1706	157	7661	1.85	90.42
1722	142	7803	1.68	92.09
1739	136	7939	1.61	93.70
1757	109	8048	1.29	94.98
1776	108	8156	1.27	96.26
1797	70	8226	0.83	97.08
1819	49	8275	0.58	97.66
1845	66	8341	0.78	98.44
1873	49	8390	0.58	99.02
1907	32	8422	0.38	99.40
1950	19	8441	0.22	99.62
2008	17	8458	0.20	99.82
2103	10	8468	0.12	99.94
2200	5	8473	0.06	100.00

Figure B.7.19. Spring 2023 STAAR Spanish Grade 3 RLA
Frequency Distribution of Scale Scores
All Students



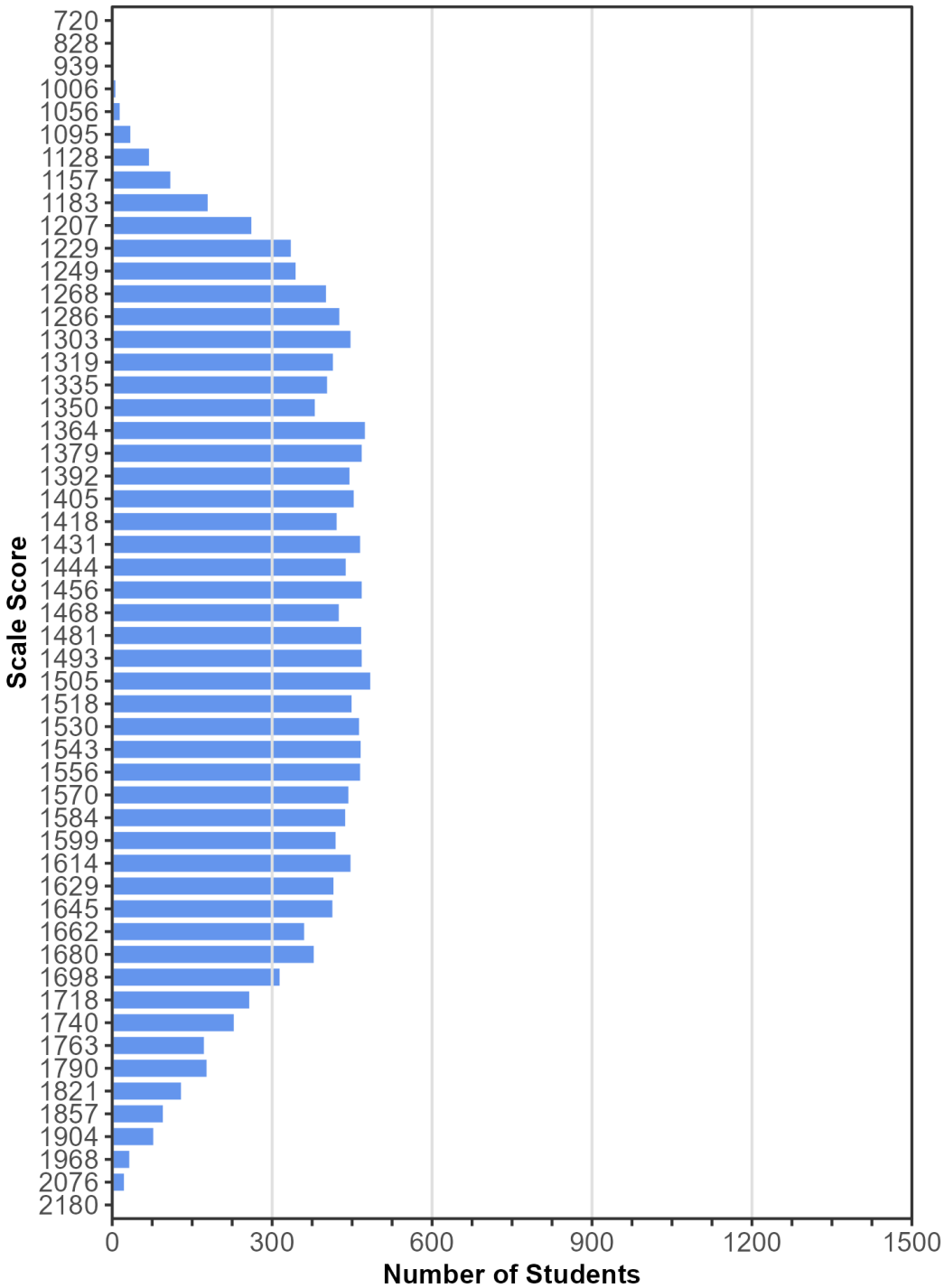
Scale Score	Freq	Cum Freq	Pct	Cum Pct
600	9	9	0.03	0.03
712	3	12	0.01	0.04
823	4	16	0.01	0.05
891	8	24	0.03	0.08
941	15	39	0.05	0.13
981	54	93	0.18	0.31
1014	115	208	0.38	0.69
1044	195	403	0.65	1.34
1070	434	837	1.44	2.78
1094	573	1410	1.90	4.68
1117	774	2184	2.57	7.24
1137	962	3146	3.19	10.43
1157	1043	4189	3.46	13.89
1176	1120	5309	3.71	17.61
1194	1170	6479	3.88	21.49
1211	1117	7596	3.70	25.19
1227	1105	8701	3.66	28.86
1244	1122	9823	3.72	32.58
1259	1034	10857	3.43	36.01
1274	960	11817	3.18	39.19
1289	987	12804	3.27	42.46
1304	987	13791	3.27	45.74
1318	891	14682	2.96	48.69
1332	942	15624	3.12	51.82
1345	906	16530	3.00	54.82
1359	935	17465	3.10	57.92
1372	899	18364	2.98	60.90
1385	884	19248	2.93	63.84
1397	799	20047	2.65	66.49
1410	824	20871	2.73	69.22
1422	799	21670	2.65	71.87
1435	729	22399	2.42	74.29
1447	744	23143	2.47	76.75
1460	770	23913	2.55	79.31
1473	724	24637	2.40	81.71
1487	701	25338	2.32	84.03
1500	657	25995	2.18	86.21
1515	616	26611	2.04	88.26
1529	561	27172	1.86	90.12
1545	539	27711	1.79	91.90
1561	459	28170	1.52	93.43
1578	430	28600	1.43	94.85
1595	381	28981	1.26	96.12
1615	318	29299	1.05	97.17
1635	283	29582	0.94	98.11
1659	202	29784	0.67	98.78
1685	143	29927	0.47	99.25
1715	92	30019	0.31	99.56
1752	82	30101	0.27	99.83
1798	30	30131	0.10	99.93
1862	13	30144	0.04	99.97
1968	7	30151	0.02	100.00
2070	1	30152	0.00	100.00

Figure B.7.20. Spring 2023 STAAR Spanish Grade 4 RLA
Frequency Distribution of Scale Scores
All Students



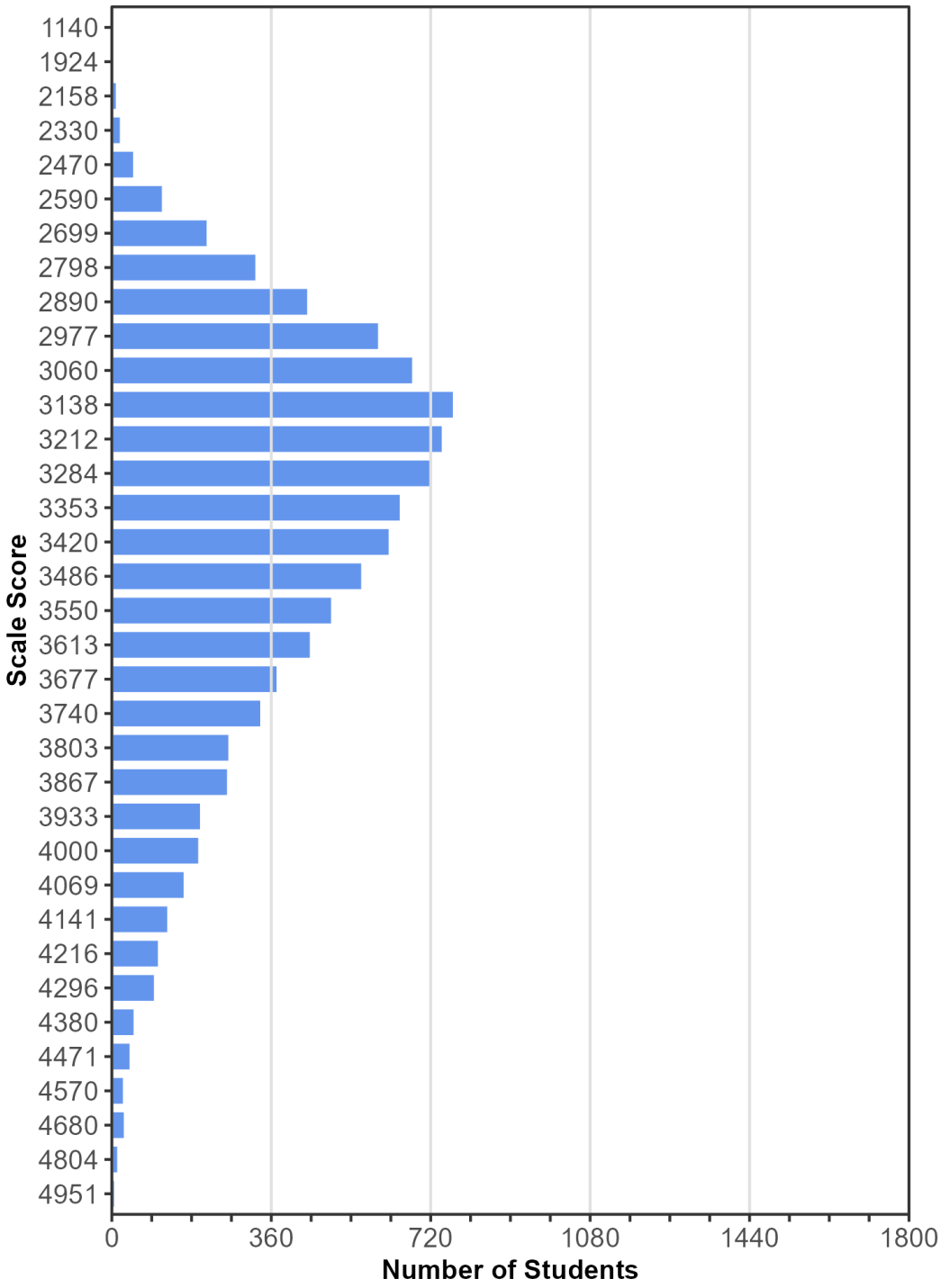
Freq	Cum Freq	Pct	Cum Pct
2	2	0.01	0.01
1	3	0.00	0.01
8	11	0.04	0.05
15	26	0.07	0.12
56	82	0.26	0.38
83	165	0.38	0.76
189	354	0.87	1.63
319	673	1.47	3.11
413	1086	1.91	5.01
507	1593	2.34	7.35
590	2183	2.72	10.08
690	2873	3.19	13.26
688	3561	3.18	16.44
716	4277	3.31	19.75
680	4957	3.14	22.89
651	5608	3.01	25.89
701	6309	3.24	29.13
647	6956	2.99	32.12
606	7562	2.80	34.91
705	8267	3.25	38.17
588	8855	2.71	40.88
634	9489	2.93	43.81
589	10078	2.72	46.53
613	10691	2.83	49.36
599	11290	2.77	52.13
621	11911	2.87	54.99
625	12536	2.89	57.88
604	13140	2.79	60.67
602	13742	2.78	63.45
586	14328	2.71	66.15
624	14952	2.88	69.03
545	15497	2.52	71.55
592	16089	2.73	74.28
563	16652	2.60	76.88
561	17213	2.59	79.47
537	17750	2.48	81.95
512	18262	2.36	84.32
517	18779	2.39	86.70
433	19212	2.00	88.70
434	19646	2.00	90.71
368	20014	1.70	92.41
340	20354	1.57	93.97
335	20689	1.55	95.52
266	20955	1.23	96.75
195	21150	0.90	97.65
194	21344	0.90	98.55
125	21469	0.58	99.12
97	21566	0.45	99.57
59	21625	0.27	99.84
23	21648	0.11	99.95
8	21656	0.04	99.99
3	21659	0.01	100.00

Figure B.7.21. Spring 2023 STAAR Spanish Grade 5 RLA
Frequency Distribution of Scale Scores
All Students



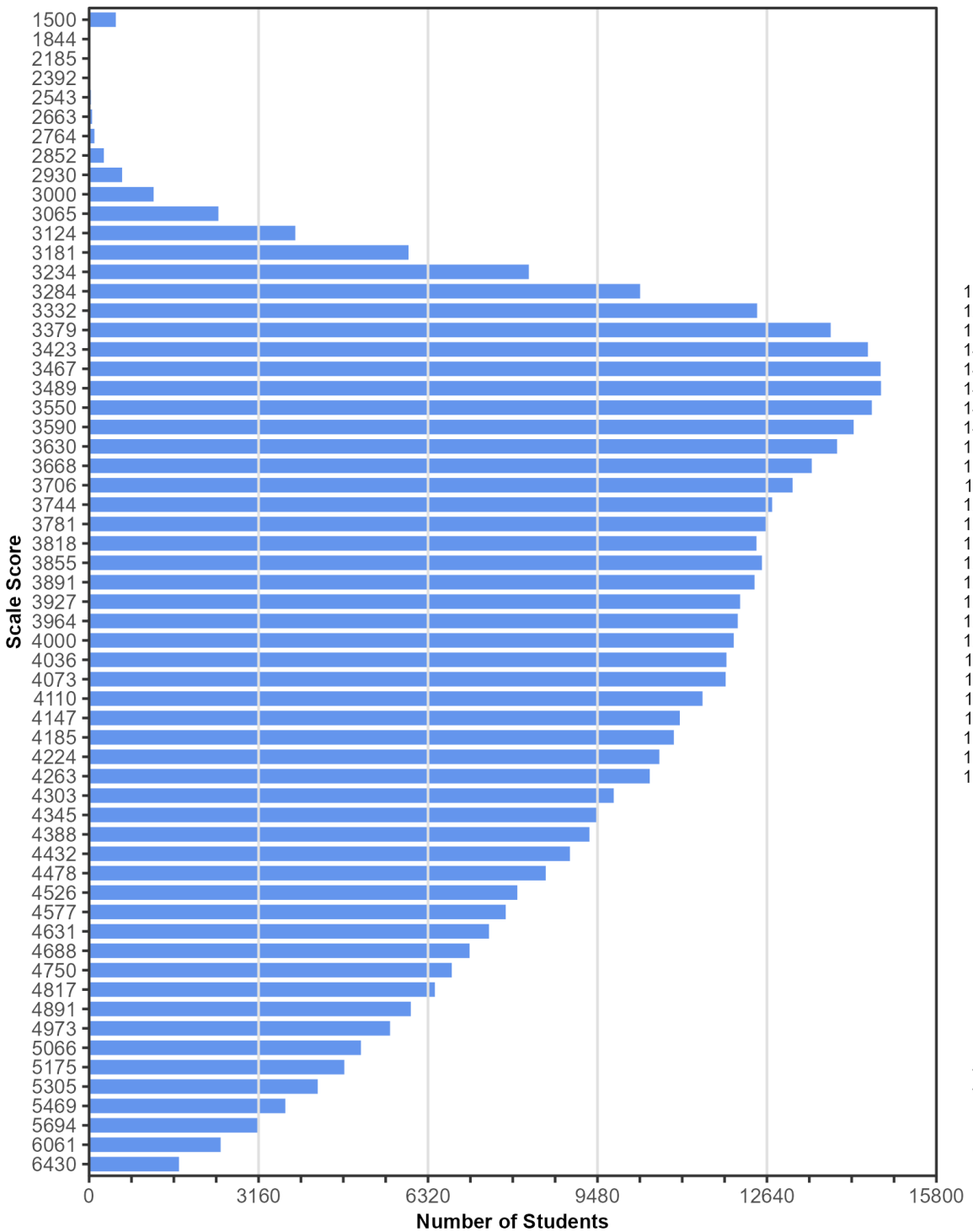
Freq	Cum Freq	Pct	Cum Pct
1	1	0.01	0.01
1	2	0.01	0.01
1	3	0.01	0.02
6	9	0.04	0.06
14	23	0.09	0.14
34	57	0.21	0.36
69	126	0.43	0.79
109	235	0.68	1.47
179	414	1.12	2.59
261	675	1.64	4.23
335	1010	2.10	6.33
344	1354	2.16	8.48
401	1755	2.51	11.00
426	2181	2.67	13.66
447	2628	2.80	16.47
414	3042	2.59	19.06
403	3445	2.52	21.58
380	3825	2.38	23.96
474	4299	2.97	26.93
468	4767	2.93	29.87
445	5212	2.79	32.65
453	5665	2.84	35.49
421	6086	2.64	38.13
465	6551	2.91	41.04
438	6989	2.74	43.79
468	7457	2.93	46.72
425	7882	2.66	49.38
467	8349	2.93	52.31
468	8817	2.93	55.24
484	9301	3.03	58.27
449	9750	2.81	61.09
463	10213	2.90	63.99
466	10679	2.92	66.91
465	11144	2.91	69.82
443	11587	2.78	72.60
437	12024	2.74	75.33
419	12443	2.63	77.96
447	12890	2.80	80.76
415	13305	2.60	83.36
413	13718	2.59	85.95
360	14078	2.26	88.20
378	14456	2.37	90.57
314	14770	1.97	92.54
257	15027	1.61	94.15
228	15255	1.43	95.58
172	15427	1.08	96.65
177	15604	1.11	97.76
129	15733	0.81	98.57
95	15828	0.60	99.17
77	15905	0.48	99.65
32	15937	0.20	99.85
22	15959	0.14	99.99
2	15961	0.01	100.00

Figure B.7.22. Spring 2023 STAAR Spanish Grade 5 Science
Frequency Distribution of Scale Scores
All Students



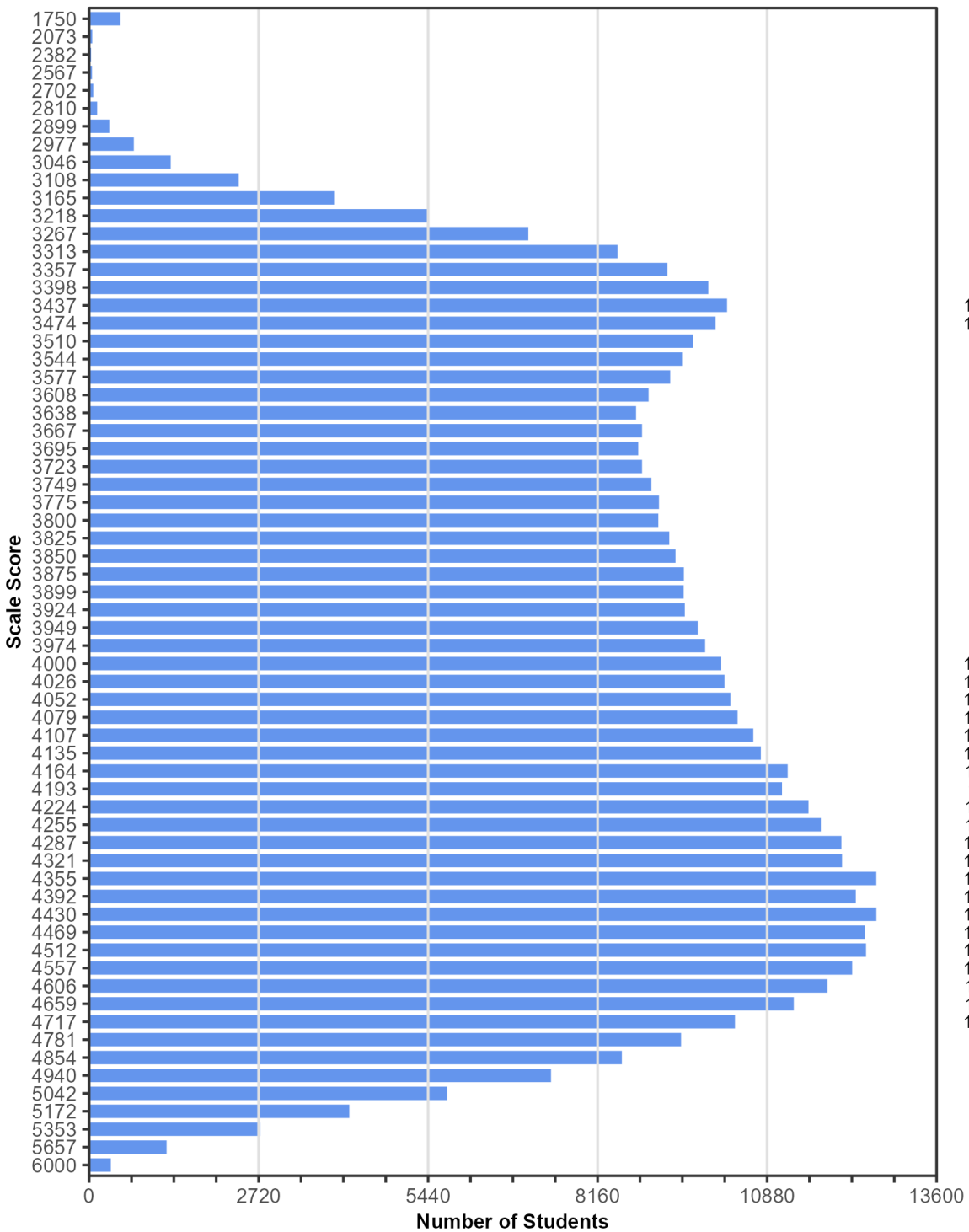
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1140	1	1	0.01	0.01
1924	1	2	0.01	0.02
2158	9	11	0.09	0.11
2330	18	29	0.19	0.30
2470	48	77	0.49	0.79
2590	113	190	1.16	1.95
2699	214	404	2.20	4.15
2798	324	728	3.33	7.48
2890	441	1169	4.53	12.02
2977	601	1770	6.18	18.19
3060	678	2448	6.97	25.16
3138	770	3218	7.91	33.08
3212	745	3963	7.66	40.73
3284	718	4681	7.38	48.11
3353	650	5331	6.68	54.79
3420	625	5956	6.42	61.22
3486	563	6519	5.79	67.01
3550	495	7014	5.09	72.09
3613	447	7461	4.59	76.69
3677	372	7833	3.82	80.51
3740	335	8168	3.44	83.96
3803	263	8431	2.70	86.66
3867	260	8691	2.67	89.33
3933	199	8890	2.05	91.38
4000	195	9085	2.00	93.38
4069	162	9247	1.67	95.05
4141	125	9372	1.28	96.33
4216	104	9476	1.07	97.40
4296	95	9571	0.98	98.38
4380	49	9620	0.50	98.88
4471	40	9660	0.41	99.29
4570	25	9685	0.26	99.55
4680	27	9712	0.28	99.83
4804	12	9724	0.12	99.95
4951	5	9729	0.05	100.00

Figure B.7.23. Spring 2023 STAAR EOC Assessments Algebra I
Frequency Distribution of Scale Scores
All Students



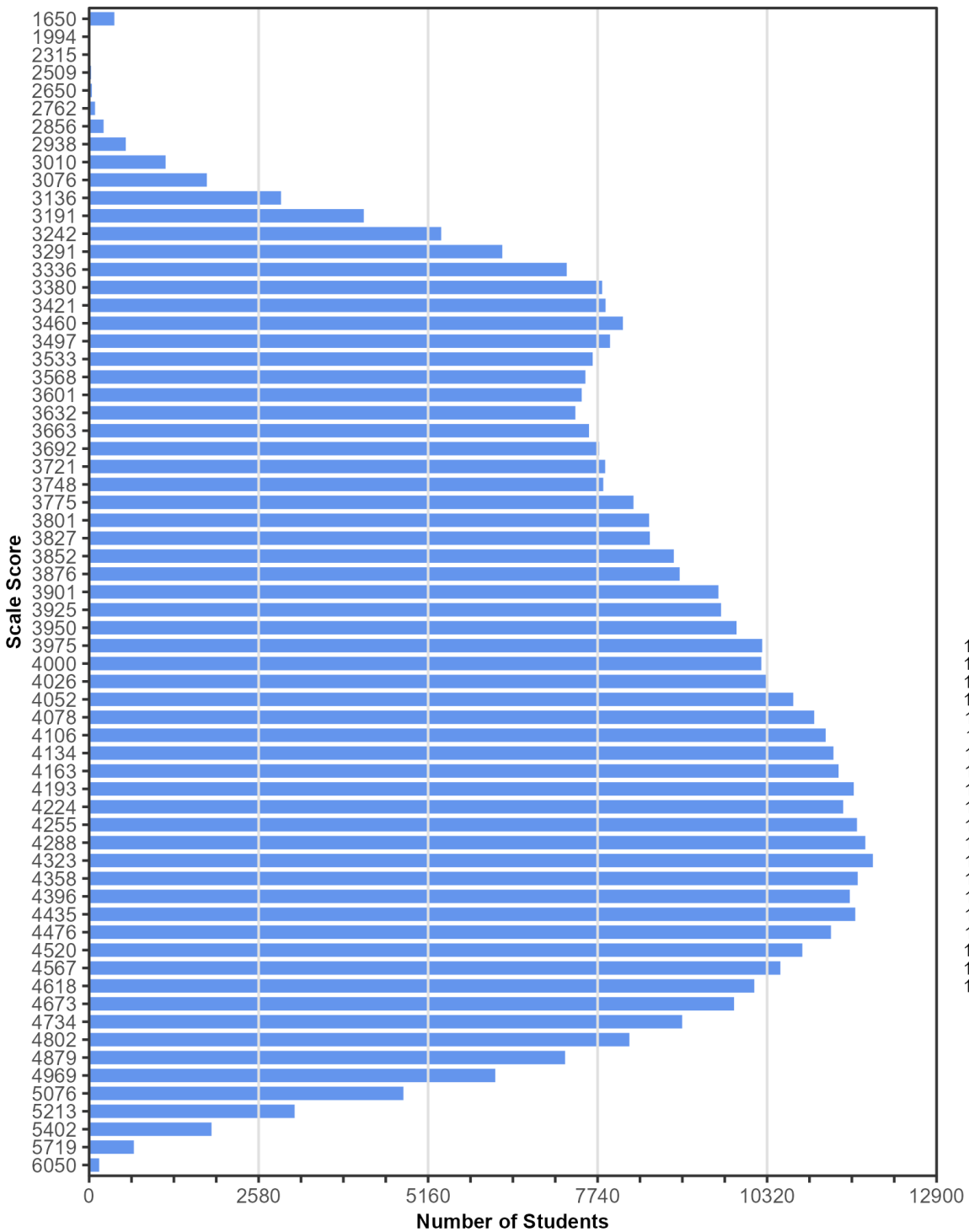
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1500	498	498	0.10	0.10
1844	20	518	0.00	0.11
2185	24	542	0.01	0.11
2392	23	565	0.00	0.12
2543	33	598	0.01	0.13
2663	55	653	0.01	0.14
2764	99	752	0.02	0.16
2852	275	1027	0.06	0.22
2930	615	1642	0.13	0.34
3000	1203	2845	0.25	0.60
3065	2411	5256	0.51	1.10
3124	3845	9101	0.81	1.91
3181	5956	15057	1.25	3.16
3234	8200	23257	1.72	4.88
3284	10275	33532	2.16	7.03
3332	12457	45989	2.61	9.65
3379	13829	59818	2.90	12.55
3423	14523	74341	3.05	15.59
3467	14760	89101	3.10	18.69
3489	14767	103868	3.10	21.79
3550	14596	118464	3.06	24.85
3590	14257	132721	2.99	27.84
3630	13948	146669	2.93	30.77
3668	13476	160145	2.83	33.59
3706	13118	173263	2.75	36.34
3744	12737	186000	2.67	39.02
3781	12640	198640	2.65	41.67
3818	12445	211085	2.61	44.28
3855	12547	223632	2.63	46.91
3891	12409	236041	2.60	49.51
3927	12140	248181	2.55	52.06
3964	12096	260277	2.54	54.60
4000	12021	272298	2.52	57.12
4036	11885	284183	2.49	59.61
4073	11870	296053	2.49	62.10
4110	11439	307492	2.40	64.50
4147	11015	318507	2.31	66.81
4185	10904	329411	2.29	69.10
4224	10634	340045	2.23	71.33
4263	10455	350500	2.19	73.52
4303	9783	360283	2.05	75.57
4345	9471	369754	1.99	77.56
4388	9330	379084	1.96	79.52
4432	8966	388050	1.88	81.40
4478	8515	396565	1.79	83.18
4526	7985	404550	1.67	84.86
4577	7769	412319	1.63	86.49
4631	7457	419776	1.56	88.05
4688	7095	426871	1.49	89.54
4750	6763	433634	1.42	90.96
4817	6448	440082	1.35	92.31
4891	5999	446081	1.26	93.57
4973	5612	451693	1.18	94.75
5066	5069	456762	1.06	95.81
5175	4759	461521	1.00	96.81
5305	4263	465784	0.89	97.70
5469	3659	469443	0.77	98.47
5694	3167	472610	0.66	99.13
6061	2453	475063	0.51	99.65
6430	1676	476739	0.35	100.00

Figure B.7.24. Spring 2023 STAAR EOC Assessments English I
Frequency Distribution of Scale Scores
All Students



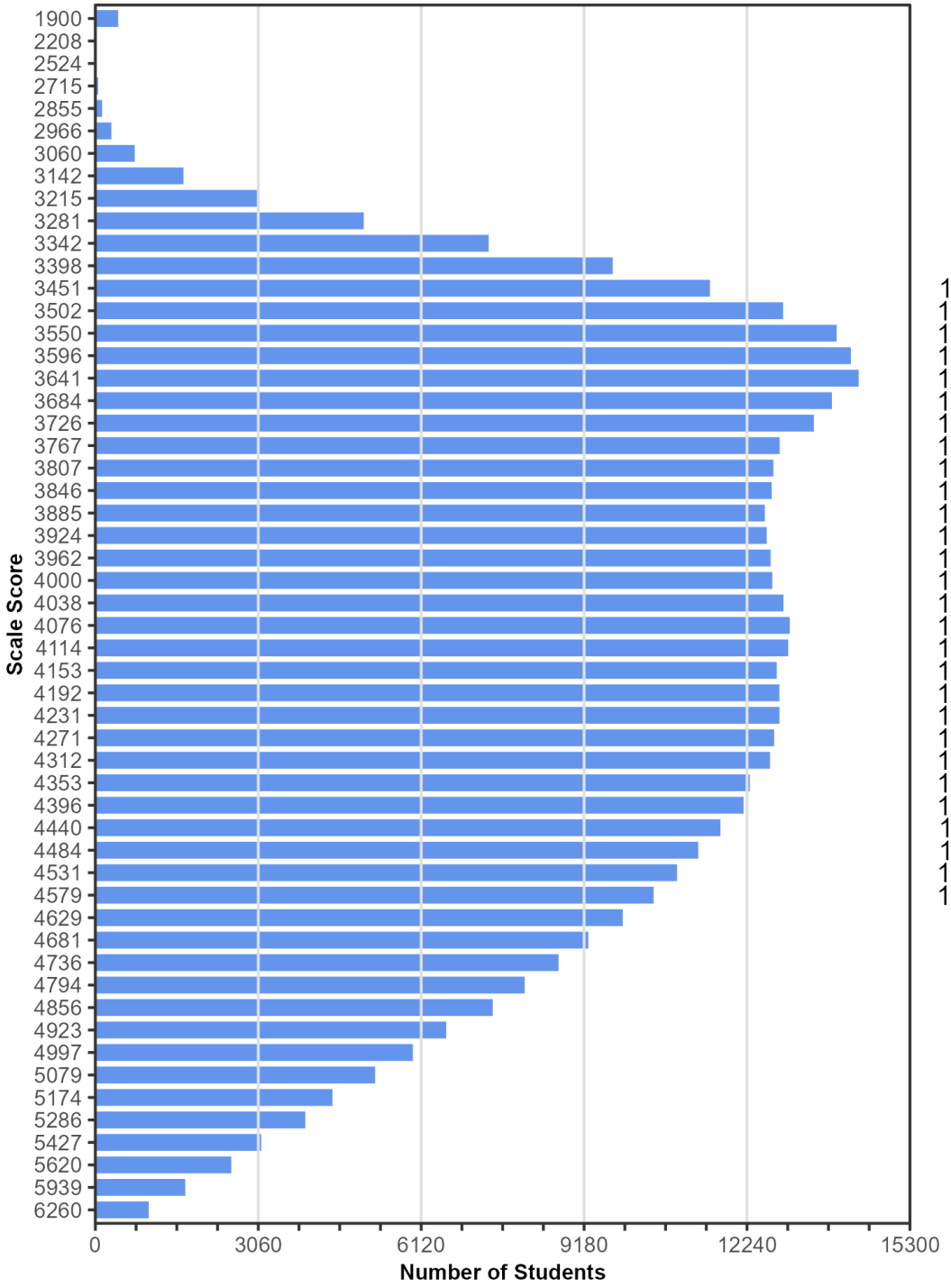
Freq	Cum Freq	Pct	Cum Pct
502	502	0.10	0.10
50	552	0.01	0.11
32	584	0.01	0.11
46	630	0.01	0.12
66	696	0.01	0.13
129	825	0.02	0.16
323	1148	0.06	0.22
717	1865	0.14	0.36
1309	3174	0.25	0.61
2401	5575	0.46	1.08
3931	9506	0.76	1.84
5453	14959	1.05	2.89
7048	22007	1.36	4.25
8479	30486	1.64	5.89
9280	39766	1.79	7.69
9936	49702	1.92	9.61
10237	59939	1.98	11.59
10052	69991	1.94	13.53
9696	79687	1.87	15.40
9515	89202	1.84	17.24
9325	98527	1.80	19.04
8978	107505	1.74	20.78
8777	116282	1.70	22.48
8873	125155	1.72	24.19
8812	133967	1.70	25.89
8873	142840	1.72	27.61
9023	151863	1.74	29.35
9145	161008	1.77	31.12
9134	170142	1.77	32.89
9310	179452	1.80	34.69
9410	188862	1.82	36.50
9544	198406	1.84	38.35
9541	207947	1.84	40.19
9559	217506	1.85	42.04
9765	227271	1.89	43.93
9885	237156	1.91	45.84
10144	247300	1.96	47.80
10197	257497	1.97	49.77
10290	267787	1.99	51.76
10406	278193	2.01	53.77
10658	288851	2.06	55.83
10779	299630	2.08	57.91
11210	310840	2.17	60.08
11119	321959	2.15	62.23
11544	333503	2.23	64.46
11742	345245	2.27	66.73
12072	357317	2.33	69.06
12082	369399	2.34	71.40
12632	382031	2.44	73.84
12303	394334	2.38	76.22
12633	406967	2.44	78.66
12451	419418	2.41	81.07
12467	431885	2.41	83.48
12245	444130	2.37	85.84
11850	455980	2.29	88.13
11308	467288	2.19	90.32
10364	477652	2.00	92.32
9499	487151	1.84	94.16
8549	495700	1.65	95.81
7412	503112	1.43	97.24
5744	508856	1.11	98.35
4175	513031	0.81	99.16
2747	515778	0.53	99.69
1243	517021	0.24	99.93
348	517369	0.07	100.00

Figure B.7.25. Spring 2023 STAAR EOC Assessments English II
Frequency Distribution of Scale Scores
All Students



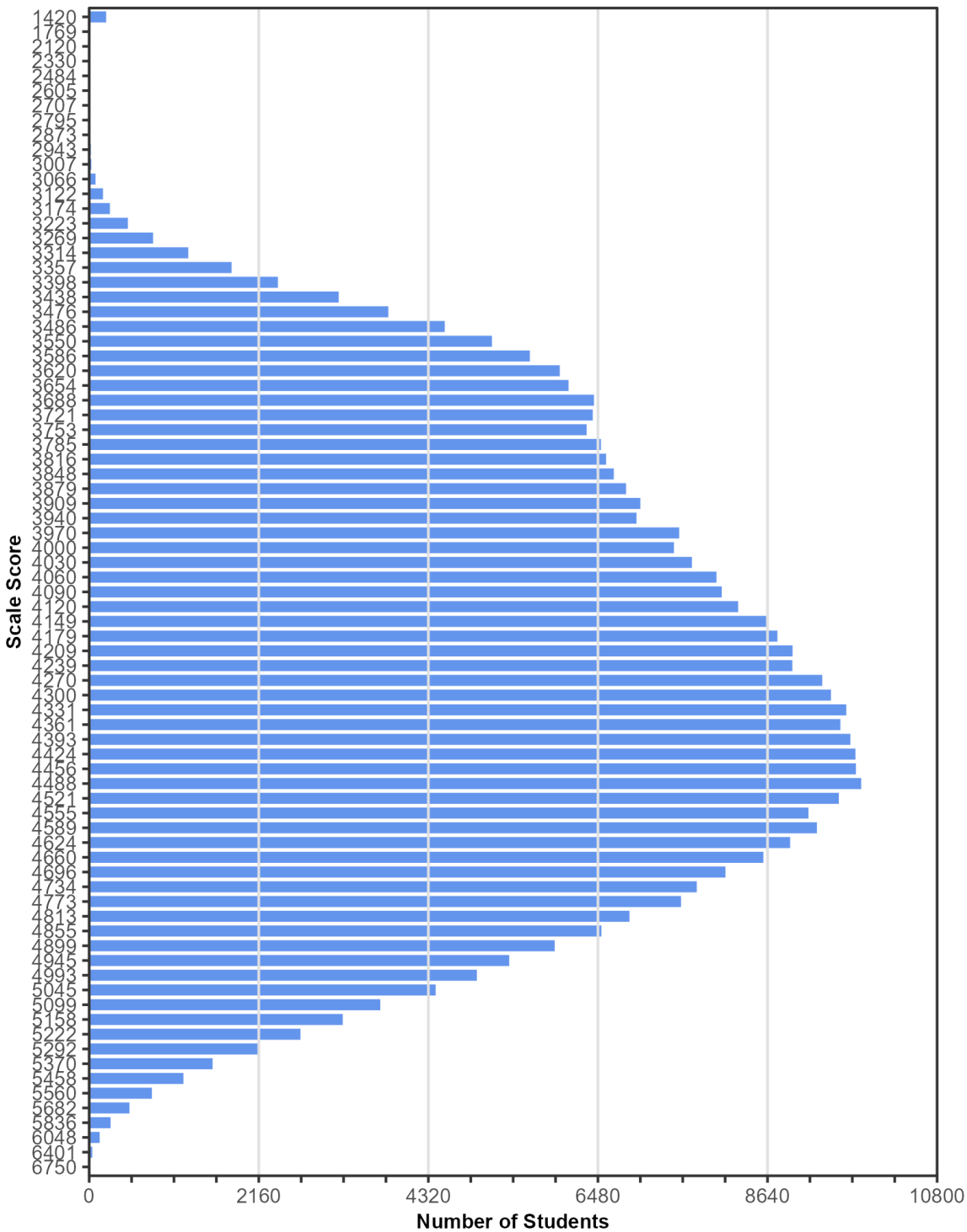
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1650	384	384	0.08	0.08
1994	19	403	0.00	0.09
2315	16	419	0.00	0.09
2509	29	448	0.01	0.10
2650	41	489	0.01	0.10
2762	90	579	0.02	0.12
2856	220	799	0.05	0.17
2938	558	1357	0.12	0.29
3010	1164	2521	0.25	0.54
3076	1792	4313	0.38	0.92
3136	2921	7234	0.62	1.54
3191	4181	11415	0.89	2.43
3242	5359	16774	1.14	3.57
3291	6289	23063	1.34	4.91
3336	7270	30333	1.55	6.46
3380	7809	38142	1.66	8.13
3421	7861	46003	1.67	9.80
3460	8125	54128	1.73	11.53
3497	7929	62057	1.69	13.22
3533	7665	69722	1.63	14.85
3568	7554	77276	1.61	16.46
3601	7498	84774	1.60	18.06
3632	7401	92175	1.58	19.64
3663	7609	99784	1.62	21.26
3692	7766	107550	1.65	22.91
3721	7855	115405	1.67	24.58
3748	7826	123231	1.67	26.25
3775	8286	131517	1.77	28.02
3801	8524	140041	1.82	29.83
3827	8535	148576	1.82	31.65
3852	8900	157476	1.90	33.55
3876	8989	166465	1.91	35.46
3901	9579	176044	2.04	37.50
3925	9619	185663	2.05	39.55
3950	9855	195518	2.10	41.65
3975	10245	205763	2.18	43.83
4000	10232	215995	2.18	46.01
4026	10330	226325	2.20	48.21
4052	10718	237043	2.28	50.50
4078	11037	248080	2.35	52.85
4106	11211	259291	2.39	55.24
4134	11330	270621	2.41	57.65
4163	11408	282029	2.43	60.08
4193	11639	293668	2.48	62.56
4224	11478	305146	2.45	65.01
4255	11688	316834	2.49	67.50
4288	11815	328649	2.52	70.01
4323	11930	340579	2.54	72.55
4358	11699	352278	2.49	75.05
4396	11579	363857	2.47	77.51
4435	11662	375519	2.48	80.00
4476	11292	386811	2.41	82.40
4520	10856	397667	2.31	84.71
4567	10522	408189	2.24	86.96
4618	10125	418314	2.16	89.11
4673	9817	428131	2.09	91.20
4734	9027	437158	1.92	93.13
4802	8224	445382	1.75	94.88
4879	7245	452627	1.54	96.42
4969	6182	458809	1.32	97.74
5076	4784	463593	1.02	98.76
5213	3128	466721	0.67	99.43
5402	1863	468584	0.40	99.82
5719	681	469265	0.15	99.97
6050	153	469418	0.03	100.00

Figure B.7.26. Spring 2023 STAAR EOC Assessments Biology
Frequency Distribution of Scale Scores
All Students



Scale Score	Freq	Cum Freq	Pct	Cum Pct
1900	429	429	0.09	0.09
2208	26	455	0.01	0.10
2524	15	470	0.00	0.10
2715	51	521	0.01	0.11
2855	127	648	0.03	0.14
2966	301	949	0.07	0.21
3060	740	1689	0.16	0.37
3142	1655	3344	0.36	0.72
3215	3081	6425	0.67	1.39
3281	5041	11466	1.09	2.48
3342	7388	18854	1.60	4.09
3398	9718	28572	2.11	6.19
3451	11544	40116	2.50	8.69
3502	12918	53034	2.80	11.49
3550	13924	66958	3.02	14.51
3596	14192	81150	3.08	17.58
3641	14338	95488	3.11	20.69
3684	13835	109323	3.00	23.69
3726	13496	122819	2.92	26.61
3767	12853	135672	2.79	29.40
3807	12735	148407	2.76	32.16
3846	12704	161111	2.75	34.91
3885	12575	173686	2.72	37.64
3924	12610	186296	2.73	40.37
3962	12683	198979	2.75	43.12
4000	12715	211694	2.76	45.87
4038	12925	224619	2.80	48.67
4076	13043	237662	2.83	51.50
4114	13014	250676	2.82	54.32
4153	12798	263474	2.77	57.09
4192	12848	276322	2.78	59.88
4231	12849	289171	2.78	62.66
4271	12749	301920	2.76	65.42
4312	12671	314591	2.75	68.17
4353	12291	326882	2.66	70.83
4396	12175	339057	2.64	73.47
4440	11741	350798	2.54	76.01
4484	11324	362122	2.45	78.47
4531	10926	373048	2.37	80.84
4579	10486	383534	2.27	83.11
4629	9906	393440	2.15	85.25
4681	9259	402699	2.01	87.26
4736	8700	411399	1.89	89.15
4794	8064	419463	1.75	90.89
4856	7464	426927	1.62	92.51
4923	6590	433517	1.43	93.94
4997	5964	439481	1.29	95.23
5079	5255	444736	1.14	96.37
5174	4453	449189	0.96	97.33
5286	3943	453132	0.85	98.19
5427	3116	456248	0.68	98.86
5620	2553	458801	0.55	99.42
5939	1689	460490	0.37	99.78
6260	1002	461492	0.22	100.00

Figure B.7.27. Spring 2023 STAAR EOC Assessments U.S. History
Frequency Distribution of Scale Scores
All Students



Freq	Cum Freq	Pct	Cum Pct
216	216	0.06	0.06
7	223	0.00	0.06
4	227	0.00	0.06
9	236	0.00	0.06
10	246	0.00	0.06
6	252	0.00	0.07
8	260	0.00	0.07
11	271	0.00	0.07
9	280	0.00	0.07
22	302	0.01	0.08
27	329	0.01	0.09
81	410	0.02	0.11
176	586	0.05	0.15
264	850	0.07	0.22
493	1343	0.13	0.35
814	2157	0.21	0.57
1262	3419	0.33	0.90
1814	5233	0.48	1.38
2404	7637	0.63	2.01
3178	10815	0.84	2.84
3810	14625	1.00	3.85
4529	19154	1.19	5.04
5130	24284	1.35	6.39
5613	29897	1.48	7.86
5994	35891	1.58	9.44
6106	41997	1.61	11.04
6429	48426	1.69	12.73
6414	54840	1.69	14.42
6336	61176	1.67	16.09
6518	67694	1.71	17.80
6584	74278	1.73	19.53
6682	80960	1.76	21.29
6838	87798	1.80	23.09
7020	94818	1.85	24.93
6971	101789	1.83	26.76
7515	109304	1.98	28.74
7448	116752	1.96	30.70
7677	124429	2.02	32.72
7990	132419	2.10	34.82
8057	140476	2.12	36.94
8266	148742	2.17	39.11
8624	157366	2.27	41.38
8766	166132	2.30	43.68
8959	175091	2.36	46.04
8956	184047	2.35	48.39
9337	193384	2.46	50.85
9448	202832	2.48	53.33
9643	212475	2.54	55.87
9568	222043	2.52	58.38
9695	231738	2.55	60.93
9760	241498	2.57	63.50
9766	251264	2.57	66.07
9833	261097	2.59	68.65
9548	270645	2.51	71.16
9161	279806	2.41	73.57
9269	289075	2.44	76.01
8928	298003	2.35	78.36
8587	306590	2.26	80.61
8104	314694	2.13	82.75
7739	322433	2.03	84.78
7538	329971	1.98	86.76
6882	336853	1.81	88.57
6524	343377	1.72	90.29
5929	349306	1.56	91.85
5350	354656	1.41	93.25
4938	359594	1.30	94.55
4413	364007	1.16	95.71
3708	367715	0.97	96.69
3230	370945	0.85	97.54
2691	373636	0.71	98.24
2146	375782	0.56	98.81
1572	377354	0.41	99.22
1200	378554	0.32	99.54
799	379353	0.21	99.75
513	379866	0.13	99.88
273	380139	0.07	99.95
134	380273	0.04	99.99
39	380312	0.01	100.00
3	380315	0.00	100.00

Spring 2023 STAAR English Learner Performance Measure

Table B.8.1. Spring 2023 STAAR Grades 3–8 Mathematics
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approach	Meets	Masters
Grade 3	6,164	820 13%	2523 41%	1601 26%	1220 20%
Grade 4	2,365	472 20%	840 36%	546 23%	507 21%
Grade 5	2,209	315 14%	853 39%	602 27%	439 20%
Grade 6	3,588	627 17%	2006 56%	635 18%	320 9%
Grade 7	3,130	1117 36%	1452 46%	392 13%	169 5%
Grade 8	3,900	826 21%	1914 49%	773 20%	387 10%

Table B.8.2. Spring 2023 STAAR Grades 3–8 RLA
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approach	Meets	Masters
Grade 3	5,221	1127 22%	1509 29%	1585 30%	1000 19%
Grade 4	1,900	545 29%	624 33%	463 24%	268 14%
Grade 5	1,867	508 27%	639 34%	425 23%	295 16%
Grade 6	3,609	1592 44%	1280 35%	520 14%	217 6%
Grade 7	3,440	1876 55%	937 27%	428 12%	199 6%
Grade 8	3,982	2033 51%	1168 29%	568 14%	213 5%

Notes for both tables:
1. The reported data are STAAR results only and do not include STAAR Alternate 2 students.
2. Due to rounding, values may not sum to 100.

Table B.8.3. Spring 2023 STAAR Grade 5 and Grade 8 Science
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approach	Meets	Masters
Grade 5	2,187	725 33%	979 45%	315 14%	168 8%
Grade 8	3,909	1304 33%	1806 46%	646 17%	153 4%

Table B.8.4. Spring 2023 STAAR Grade 8 Social Studies
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approach	Meets	Masters
Grade 8	3,967	2059 52%	1483 37%	276 7%	149 4%

Table B.8.5. Spring 2023 STAAR EOC Assessments
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approach	Meets	Masters
Algebra I	5,982	948 16%	2972 50%	1304 22%	758 13%
English I	7,055	3282 47%	2644 37%	938 13%	191 3%
English II	8,129	3443 42%	3400 42%	1208 15%	78 1%
Biology	7,012	1618 23%	3433 49%	1545 22%	416 6%
U.S. History	3,876	133 3%	1792 46%	1343 35%	608 16%

Notes for all tables:
1. The reported data are STAAR results only and do not include STAAR Alternate 2 students.
2. Due to rounding, values may not sum to 100.

TECHNICAL DIGEST 2022–2023

Appendix C

STAAR Alternate 2 Statistical Tables and Figures

STAAR Alternate 2 Statistical Tables and Figures

2023 STAAR Alternate 2 Classification Consistency and Accuracy

- Table C.1.1. STAAR Alternate 2 Grades 3–8 Assessments
- Table C.1.2. STAAR Alternate 2 EOC Assessments

2023 STAAR Alternate 2 Scale Score Correlations

- Table C.2.1. STAAR Alternate 2 Grades 3–8 Assessments
- Table C.2.2. STAAR Alternate 2 EOC Assessments

2023 STAAR Alternate 2 Conditional Standard Error of Measurement for Scale Scores

- Table C.3.1. STAAR Alternate 2 Grades 3–5 Mathematics
- Table C.3.2. STAAR Alternate 2 Grades 6–8 Mathematics
- Table C.3.3. STAAR Alternate 2 Grades 3–5 RLA
- Table C.3.4. STAAR Alternate 2 Grades 6–8 RLA
- Table C.3.5. STAAR Alternate 2 Grade 5 and Grade 8 Science
- Table C.3.6. STAAR Alternate 2 Grade 8 Social Studies
- Table C.3.7. STAAR Alternate 2 Algebra I
- Table C.3.8. STAAR Alternate 2 English I and English II
- Table C.3.9. STAAR Alternate 2 Biology
- Table C.3.10. STAAR Alternate 2 U.S. History

2023 STAAR Alternate 2 Mean P-Values

- Table C.4.1. STAAR Alternate 2 Grade 3 Total Group
- Table C.4.2. STAAR Alternate 2 Grade 3 Female
- Table C.4.3. STAAR Alternate 2 Grade 3 Male
- Table C.4.4. STAAR Alternate 2 Grade 3 Black or African-American
- Table C.4.5. STAAR Alternate 2 Grade 3 Hispanic or Latino
- Table C.4.6. STAAR Alternate 2 Grade 3 White
- Table C.4.7. STAAR Alternate 2 Grade 4 Total Group
- Table C.4.8. STAAR Alternate 2 Grade 4 Female
- Table C.4.9. STAAR Alternate 2 Grade 4 Male
- Table C.4.10. STAAR Alternate 2 Grade 4 Black or African-American
- Table C.4.11. STAAR Alternate 2 Grade 4 Hispanic or Latino
- Table C.4.12. STAAR Alternate 2 Grade 4 White
- Table C.4.13. STAAR Alternate 2 Grade 5 Total Group
- Table C.4.14. STAAR Alternate 2 Grade 5 Female

- Table C.4.15. STAAR Alternate 2 Grade 5 Male
- Table C.4.16. STAAR Alternate 2 Grade 5 Black or African-American
- Table C.4.17. STAAR Alternate 2 Grade 5 Hispanic or Latino
- Table C.4.18. STAAR Alternate 2 Grade 5 White
- Table C.4.19. STAAR Alternate 2 Grade 6 Total Group
- Table C.4.20. STAAR Alternate 2 Grade 6 Female
- Table C.4.21. STAAR Alternate 2 Grade 6 Male
- Table C.4.22. STAAR Alternate 2 Grade 6 Black or African-American
- Table C.4.23. STAAR Alternate 2 Grade 6 Hispanic or Latino
- Table C.4.24. STAAR Alternate 2 Grade 6 White
- Table C.4.25. STAAR Alternate 2 Grade 7 Total Group
- Table C.4.26. STAAR Alternate 2 Grade 7 Female
- Table C.4.27. STAAR Alternate 2 Grade 7 Male
- Table C.4.28. STAAR Alternate 2 Grade 7 Black or African-American
- Table C.4.29. STAAR Alternate 2 Grade 7 Hispanic or Latino
- Table C.4.30. STAAR Alternate 2 Grade 7 White
- Table C.4.31. STAAR Alternate 2 Grade 8 Total Group
- Table C.4.32. STAAR Alternate 2 Grade 8 Female
- Table C.4.33. STAAR Alternate 2 Grade 8 Male
- Table C.4.34. STAAR Alternate 2 Grade 8 Black or African-American
- Table C.4.35. STAAR Alternate 2 Grade 8 Hispanic or Latino
- Table C.4.36. STAAR Alternate 2 Grade 8 White
- Table C.4.37. STAAR Alternate 2 Algebra I Total Group
- Table C.4.38. STAAR Alternate 2 Algebra I Female
- Table C.4.39. STAAR Alternate 2 Algebra I Male
- Table C.4.40. STAAR Alternate 2 Algebra I Black or African-American
- Table C.4.41. STAAR Alternate 2 Algebra I Hispanic or Latino
- Table C.4.42. STAAR Alternate 2 Algebra I White
- Table C.4.43. STAAR Alternate 2 English Total Group
- Table C.4.44. STAAR Alternate 2 English Female
- Table C.4.45. STAAR Alternate 2 English Male
- Table C.4.46. STAAR Alternate 2 English Black or African-American
- Table C.4.47. STAAR Alternate 2 English Hispanic or Latino
- Table C.4.48. STAAR Alternate 2 English White
- Table C.4.49. STAAR Alternate 2 Biology Total Group
- Table C.4.50. STAAR Alternate 2 Biology Female

- Table C.4.51. STAAR Alternate 2 Biology Male
- Table C.4.52. STAAR Alternate 2 Biology Black or African-American
- Table C.4.53. STAAR Alternate 2 Biology Hispanic or Latino
- Table C.4.54. STAAR Alternate 2 Biology White
- Table C.4.55. STAAR Alternate 2 U.S. History Total Group
- Table C.4.56. STAAR Alternate 2 U.S. History Female
- Table C.4.57. STAAR Alternate 2 U.S. History Male
- Table C.4.58. STAAR Alternate 2 U.S. History Black or African-American
- Table C.4.59. STAAR Alternate 2 U.S. History Hispanic or Latino
- Table C.4.60. STAAR Alternate 2 U.S. History White

2023 STAAR Alternate 2 Progress Measure

Progress measures were not calculated or reported for the 2022–2023 school year.

2023 STAAR Alternate 2 Scale Score Descriptive Statistics

- Table C.6.1. STAAR Alternate 2 Grades 3–8 Assessments
- Table C.6.2. STAAR Alternate 2 EOC Assessments

2023 STAAR Alternate 2 Frequency Distribution of Scale Scores

- Figure C.6.1. STAAR Alternate 2 Grade 3 Mathematics
- Figure C.6.2. STAAR Alternate 2 Grade 4 Mathematics
- Figure C.6.3. STAAR Alternate 2 Grade 5 Mathematics
- Figure C.6.4. STAAR Alternate 2 Grade 6 Mathematics
- Figure C.6.5. STAAR Alternate 2 Grade 7 Mathematics
- Figure C.6.6. STAAR Alternate 2 Grade 8 Mathematics
- Figure C.6.7. STAAR Alternate 2 Grade 3 RLA
- Figure C.6.8. STAAR Alternate 2 Grade 4 RLA
- Figure C.6.9. STAAR Alternate 2 Grade 5 RLA
- Figure C.6.10. STAAR Alternate 2 Grade 6 RLA
- Figure C.6.11. STAAR Alternate 2 Grade 7 RLA
- Figure C.6.12. STAAR Alternate 2 Grade 8 RLA
- Figure C.6.13. STAAR Alternate 2 Grade 5 Science
- Figure C.6.14. STAAR Alternate 2 Grade 8 Science
- Figure C.6.15. STAAR Alternate 2 Grade 8 Social Studies
- Figure C.6.16. STAAR Alternate 2 Algebra I
- Figure C.6.17. STAAR Alternate 2 English I
- Figure C.6.18. STAAR Alternate 2 English II
- Figure C.6.19. STAAR Alternate 2 Biology

- Figure C.6.20. STAAR Alternate 2 U.S. History

2023 STAAR Alternate 2 Classification Consistency and Accuracy

Table C.1.1. Spring 2023 STAAR Alternate 2 Grades 3–8 Assessments
Classification Consistency and Accuracy

	Decision Consistency				Decision Accuracy			
Grade	**Mathematics**	**RLA**	**Science**	**Social Studies**	**Mathematics**	**RLA**	**Science**	**Social Studies**
3	74.8	76.8			84.8	85.5		
4	78.3	73.3			87.0	83.2		
5	73.4	69.9	76.0		83.9	81.0	85.8	
6	71.5	69.1			82.3	80.3		
7	72.4	71.1			83.1	81.7		
8	71.4	69.6	75.4	70.9	82.1	80.7	85.2	81.9

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

Table C.1.2. Spring 2023 STAAR Alternate 2 EOC Assessments
Classification Consistency and Accuracy

Test	Decision Consistency	Decision Accuracy
Algebra I	69.7	81.1
English I	69.9	81.3
English II	70.4	81.5
Biology	74.0	84.5
U.S. History	71.3	82.2

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

2023 STAAR Alternate 2 Scale Score Correlations

Table C.2.1. Spring 2023 STAAR Alternate 2 Grades 3–8 Assessments Scale Score Correlations within Grade

Grade	STAAR Alternate 2	STAAR Alternate 2	N	Correlation
3	Mathematics	RLA	7,383	0.79
4	Mathematics	RLA	7,291	0.77
5	Mathematics	RLA	6,814	0.73
	Mathematics	Science	6,816	0.75
	RLA	Science	6,813	0.75
6	Mathematics	RLA	6,473	0.73
7	Mathematics	RLA	6,299	0.70
8	Mathematics	RLA	6,154	0.70
	Mathematics	Science	6,152	0.71
	Mathematics	Social Studies	6,148	0.70
	RLA	Science	6,151	0.70
	RLA	Social Studies	6,147	0.70
	Science	Social Studies	6,154	0.71

Table C.2.2. Spring 2023 STAAR Alternate 2 EOC Assessments Scale Score Correlations

STAAR Alternate 2	STAAR Alternate 2	N	Correlation
Algebra I	English I	5,883	0.71
Algebra I	English II	151	0.64
Algebra I	Biology	4,426	0.66
Algebra I	US History	131	0.71
English I	English II	103	0.82
English I	Biology	4,444	0.68
English I	US History	125	0.65
English II	Biology	1,497	0.65
English II	US History	621	0.71
Biology	US History	340	0.75

2023 STAAR Alternate 2 Conditional Standard Error of Measurement for Scale Scores

Table C.3.1. Spring 2023 STAAR Alternate 2 Grades 3–5 Mathematics Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	116		114		111	
1	162	40	158	39	155	39
2	187	28	183	28	180	27
3	202	23	198	23	194	23
4	213	21	209	21	205	20
5	221	19	218	19	214	19
6	229	18	226	18	221	18
7	236	17	233	17	228	17
8	243	16	240	16	234	16
9	249	16	246	16	240	16
10	254	15	252	15	246	15
11	259	15	257	15	251	15
12	265	15	263	15	256	15
13	269	15	268	15	261	15
14	274	14	273	15	266	14
15	279	14	278	14	271	14
16	283	14	282	14	276	14
17	288	14	287	14	280	14
18	292	14	292	14	284	14
19	296	14	297	14	289	14
20	**300	14	**300	14	293	14
21	305	14	306	14	297	14
22	309	14	310	14	**300	14
23	313	14	315	14	306	14
24	317	14	319	14	310	14
25	322	14	324	14	315	14
26	326	14	329	14	319	14
27	331	14	333	14	324	14
28	335	15	338	15	329	14
29	340	15	343	15	333	15
30	346	15	349	15	339	15
31	351	16	354	16	344	15
32	357	16	360	16	350	16

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	363	17	367	17	356	17
34	371	18	374	18	363	18
35	***375	20	382	19	371	19
36	388	22	***387	21	***379	21
37	400	25	403	24	392	24
38	417	30	419	29	409	29
39	445	43	446	41	437	42
40	496		495		486	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.2. Spring 2023 STAAR Alternate 2 Grades 6–8 Mathematics Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		103		100	
1	138	43	152	42	134	41
2	166	30	178	29	160	29
3	182	25	193	24	175	25
4	193	22	205	21	187	22
5	203	20	214	19	197	20
6	211	19	221	18	205	19
7	218	18	228	17	212	18
8	225	18	235	17	219	18
9	232	17	241	16	226	17
10	238	17	246	16	232	17
11	244	17	252	16	238	17
12	249	16	257	15	244	16
13	255	16	262	15	250	16
14	260	16	267	15	255	16
15	266	16	272	15	261	16
16	271	16	277	15	266	16
17	277	16	282	15	272	16
18	282	16	287	15	277	16
19	287	16	291	15	282	16
20	293	16	296	14	288	15
21	**300	16	**300	14	293	15
22	303	16	305	14	**300	15
23	309	16	310	14	303	15
24	314	16	315	15	308	15
25	320	16	319	15	313	15
26	325	16	324	15	319	16
27	331	16	329	15	324	16
28	337	17	334	15	329	16
29	343	17	340	16	335	16
30	349	17	345	16	341	17
31	355	18	351	17	347	17
32	362	18	357	17	353	18

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	***373	19	364	18	360	18
34	378	20	***375	19	***365	20
35	387	22	381	21	377	21
36	399	24	391	23	388	23
37	412	27	404	26	401	27
38	431	33	422	32	420	32
39	464	47	453	45	451	45
40	520		506		505	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.3. Spring 2023 STAAR Alternate 2 Grades 3–5 RLA Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100		100	
1	141	49	122	49	122	49
2	173	34	154	34	154	34
3	191	28	172	29	172	28
4	204	24	186	25	185	25
5	214	22	197	23	196	22
6	223	20	206	22	205	21
7	230	19	215	21	213	20
8	237	19	223	20	220	19
9	244	18	230	19	227	18
10	250	17	236	18	233	18
11	255	17	243	18	239	17
12	261	17	248	17	245	17
13	266	16	254	17	250	17
14	271	16	260	17	255	17
15	276	16	265	17	261	16
16	281	16	270	16	266	16
17	286	16	275	16	271	16
18	291	16	280	16	276	16
19	295	16	285	16	281	16
20	**300	16	290	16	285	16
21	305	16	295	16	290	16
22	309	16	**300	16	295	16
23	314	16	305	16	**300	16
24	319	16	310	16	305	16
25	324	16	315	16	310	16
26	329	16	320	16	315	16
27	334	16	325	17	320	17
28	339	17	331	17	325	17
29	345	17	336	17	331	17
30	350	17	342	18	337	18
31	357	18	348	18	343	18
32	363	19	355	19	349	19

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	370	20	363	20	357	20
34	378	21	371	22	365	21
35	***388	23	***380	23	***374	23
36	399	25	392	26	385	25
37	412	29	406	30	399	29
38	432	35	426	36	418	35
39	465	50	461	51	452	51
40	525		522		512	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.4. Spring 2023 STAAR Alternate 2 Grades 6–8 RLA Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100		100	
1	116	50	118	50	114	47
2	148	35	150	35	145	33
3	166	29	169	29	163	27
4	180	25	182	25	176	24
5	191	23	193	23	186	22
6	200	21	202	21	196	21
7	208	20	211	20	204	20
8	215	19	218	20	211	19
9	222	19	225	19	218	19
10	229	18	231	18	225	18
11	235	18	238	18	231	18
12	241	17	244	18	238	18
13	246	17	250	18	244	17
14	252	17	255	18	250	17
15	258	17	261	17	256	17
16	263	17	267	17	261	17
17	268	17	272	17	267	17
18	274	17	278	17	273	17
19	279	17	283	17	278	17
20	284	17	289	17	284	17
21	289	17	294	17	289	17
22	295	17	**300	17	295	17
23	**300	17	306	17	**300	17
24	305	17	311	17	305	17
25	311	17	317	17	311	17
26	316	17	322	18	317	17
27	322	17	328	18	322	17
28	328	18	334	18	328	17
29	334	18	341	18	334	18
30	340	18	347	19	341	18
31	346	19	354	19	347	19
32	354	20	361	20	354	19

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	361	21	369	21	362	20
34	***370	22	***378	22	***371	21
35	380	24	388	24	381	23
36	391	26	401	27	392	25
37	406	30	416	31	407	29
38	426	36	437	37	427	35
39	461	52	473	53	462	50
40	523		536		521	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.5. Spring 2023 STAAR Alternate 2 Grade 5 and Grade 8 Science Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	106		139	
1	149	39	175	34
2	172	27	196	24
3	187	23	208	20
4	197	21	217	18
5	206	19	225	16
6	214	18	231	15
7	222	18	237	15
8	229	17	243	14
9	235	17	248	14
10	242	17	253	14
11	248	17	257	13
12	254	16	262	13
13	260	16	267	13
14	266	16	271	13
15	272	16	275	13
16	277	15	280	13
17	283	15	284	13
18	288	15	289	13
19	293	15	293	13
20	**300	15	297	13
21	303	15	**300	13
22	308	14	306	13
23	312	14	311	13
24	317	14	315	13
25	322	14	320	13
26	327	15	325	13
27	332	15	329	14
28	337	15	334	14
29	342	15	339	14
30	347	16	344	14
31	353	16	350	15
32	359	17	356	15

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	365	17	362	16
34	372	18	369	17
35	381	20	376	18
36	***387	22	***382	20
37	403	25	396	22
38	420	30	412	27
39	450	43	437	38
40	501		483	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.6. Spring 2023 STAAR Alternate 2 Grade 8 Social Studies Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Grade 8
Raw	**SS**	**CSEM**
0	103	
1	145	37
2	168	26
3	182	22
4	192	19
5	200	18
6	207	17
7	213	16
8	219	15
9	225	15
10	230	15
11	235	15
12	241	15
13	246	15
14	252	15
15	257	15
16	262	15
17	268	15
18	273	15
19	278	15
20	284	15
21	289	15
22	294	15
23	**300	15
24	305	15
25	310	15
26	315	15
27	320	15
28	325	15
29	331	15
30	336	15
31	342	16
32	348	16

(Continued)

	Grade 8	
Raw	**SS**	**CSEM**
33	355	17
34	362	18
35	***372	19
36	380	21
37	392	24
38	409	29
39	437	41
40	486	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.7. Spring 2023 STAAR Alternate 2 EOC Assessments Mathematics Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Algebra I	
Raw	**SS**	**CSEM**
0	100	
1	137	44
2	165	31
3	182	25
4	194	22
5	204	20
6	213	19
7	220	18
8	227	17
9	233	17
10	240	17
11	245	16
12	251	16
13	256	16
14	261	15
15	267	15
16	272	15
17	277	15
18	281	15
19	286	15
20	291	15
21	296	15
22	**300	15
23	305	15
24	310	15
25	315	15
26	320	15
27	325	15
28	331	16
29	336	16
30	342	17
31	348	17
32	355	18

(Continued)

	Algebra I	
Raw	**SS**	**CSEM**
33	***361	19
34	370	20
35	379	21
36	389	23
37	403	27
38	422	32
39	453	46
40	508	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.8. Spring 2023 STAAR Alternate 2 EOC Assessments RLA Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100	
1	124	47	128	47
2	156	33	159	33
3	174	27	177	27
4	187	24	190	24
5	198	22	201	22
6	207	21	210	20
7	215	20	218	19
8	222	19	225	19
9	229	18	232	18
10	236	18	239	18
11	242	18	245	17
12	249	17	251	17
13	255	17	257	17
14	261	17	263	17
15	267	17	268	17
16	272	17	274	16
17	278	17	279	16
18	284	17	284	16
19	289	16	290	16
20	295	16	295	16
21	**300	16	**300	16
22	305	16	305	16
23	311	16	310	16
24	316	16	316	16
25	322	16	321	16
26	327	17	326	16
27	333	17	332	17
28	339	17	337	17
29	345	17	343	17
30	351	18	349	18
31	358	18	355	18
32	***365	19	362	19

(Continued)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	372	20	***370	20
34	381	21	378	21
35	390	23	388	23
36	402	25	399	25
37	416	29	414	29
38	436	35	434	35
39	470	49	468	49
40	528		527	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.9. Spring 2023 STAAR Alternate 2 EOC Assessments Science Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

		Biology
Raw	**SS**	**CSEM**
0	135	
1	177	36
2	200	25
3	213	21
4	223	18
5	231	16
6	237	15
7	243	14
8	248	14
9	253	13
10	258	13
11	262	13
12	266	13
13	270	12
14	274	12
15	278	12
16	282	12
17	286	12
18	290	12
19	294	12
20	297	12
21	**300	12
22	305	12
23	309	12
24	313	12
25	317	12
26	321	13
27	325	13
28	329	13
29	334	13
30	338	14
31	343	14
32	349	15

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	354	15
34	361	16
35	368	17
36	377	19
37	***383	22
38	403	27
39	430	38
40	475	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table C.3.10. Spring 2023 STAAR Alternate 2 EOC Assessments Social Studies Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	U.S. History	
Raw	**SS**	**CSEM**
0	105	
1	148	37
2	171	26
3	184	21
4	194	19
5	202	17
6	209	17
7	215	16
8	221	15
9	226	15
10	232	15
11	237	15
12	243	15
13	248	15
14	253	15
15	258	15
16	264	15
17	269	15
18	274	15
19	279	15
20	284	14
21	289	14
22	294	14
23	**300	14
24	304	14
25	309	15
26	314	15
27	320	15
28	325	15
29	330	15
30	336	15
31	342	16
32	348	16

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
33	355	17
34	363	18
35	***368	19
36	381	21
37	394	24
38	411	29
39	440	41
40	489	

Notes:
** Level II: Satisfactory Academic Performance
*** Level III: Accomplished Academic Performance
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

2023 STAAR Alternate 2
Mean P-Values

Table C.4.1. Spring 2023 STAAR Alternate 2 Grade 3 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	7,385	29.39	7.28	0.85	2.80	1.47
	Numerical Representations and Relationships	8	7,385	6.61	1.82	0.66	1.05	1.65
	Computations and Algebraic Relationships	16	7,385	11.27	3.27	0.68	1.85	1.41
	Geometry and Measurement	8	7,385	5.53	1.79	0.47	1.30	1.38
	Data Analysis and Personal Financial Literacy	8	7,385	5.99	1.82	0.52	1.27	1.50
RLA	OVERALL TEST	40	7,389	27.97	7.28	0.83	2.97	1.40
	Reading	24	7,389	17.44	4.57	0.77	2.20	1.45
	Writing	16	7,389	10.53	3.36	0.65	1.99	1.32

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.2. Spring 2023 STAAR Alternate 2 Grade 3 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,274	29.20	7.06	0.84	2.82	1.46
	Numerical Representations and Relationships	8	2,274	6.58	1.80	0.65	1.06	1.65
	Computations and Algebraic Relationships	16	2,274	11.19	3.17	0.66	1.86	1.40
	Geometry and Measurement	8	2,274	5.49	1.75	0.44	1.31	1.37
	Data Analysis and Personal Financial Literacy	8	2,274	5.94	1.82	0.51	1.28	1.48
RLA	OVERALL TEST	40	2,274	28.20	7.06	0.82	2.96	1.41
	Reading	24	2,274	17.64	4.47	0.76	2.18	1.47
	Writing	16	2,274	10.57	3.28	0.63	1.98	1.32

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.3. Spring 2023 STAAR Alternate 2 Grade 3 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	5,102	29.47	7.37	0.86	2.79	1.47
	Numerical Representations and Relationships	8	5,102	6.62	1.82	0.67	1.05	1.65
	Computations and Algebraic Relationships	16	5,102	11.30	3.31	0.69	1.85	1.41
	Geometry and Measurement	8	5,102	5.54	1.81	0.49	1.30	1.39
	Data Analysis and Personal Financial Literacy	8	5,102	6.01	1.83	0.52	1.26	1.50
RLA	OVERALL TEST	40	5,106	27.86	7.37	0.84	2.98	1.39
	Reading	24	5,106	17.35	4.62	0.77	2.20	1.45
	Writing	16	5,106	10.50	3.40	0.66	1.99	1.31

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.4. Spring 2023 STAAR Alternate 2 Grade 3 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,251	29.14	7.39	0.85	2.83	1.46
	Numerical Representations and Relationships	8	1,251	6.55	1.87	0.68	1.05	1.64
	Computations and Algebraic Relationships	16	1,251	11.13	3.37	0.69	1.87	1.39
	Geometry and Measurement	8	1,251	5.48	1.81	0.47	1.32	1.37
	Data Analysis and Personal Financial Literacy	8	1,251	5.98	1.83	0.53	1.26	1.50
RLA	OVERALL TEST	40	1,251	28.25	7.23	0.83	2.96	1.41
	Reading	24	1,251	17.59	4.55	0.77	2.20	1.47
	Writing	16	1,251	10.67	3.31	0.65	1.96	1.33

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.5. Spring 2023 STAAR Alternate 2 Grade 3 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,120	29.92	7.00	0.85	2.75	1.50
	Numerical Representations and Relationships	8	4,120	6.73	1.74	0.66	1.02	1.68
	Computations and Algebraic Relationships	16	4,120	11.47	3.16	0.67	1.83	1.43
	Geometry and Measurement	8	4,120	5.63	1.75	0.47	1.28	1.41
	Data Analysis and Personal Financial Literacy	8	4,120	6.09	1.77	0.50	1.25	1.52
RLA	OVERALL TEST	40	4,125	28.37	6.99	0.82	2.94	1.42
	Reading	24	4,125	17.69	4.39	0.76	2.17	1.47
	Writing	16	4,125	10.68	3.27	0.63	1.98	1.34

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.6. Spring 2023 STAAR Alternate 2 Grade 3 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,362	28.96	7.45	0.86	2.82	1.45
	Numerical Representations and Relationships	8	1,362	6.50	1.84	0.65	1.09	1.63
	Computations and Algebraic Relationships	16	1,362	11.12	3.26	0.67	1.88	1.39
	Geometry and Measurement	8	1,362	5.45	1.82	0.50	1.29	1.36
	Data Analysis and Personal Financial Literacy	8	1,362	5.88	1.87	0.53	1.28	1.47
RLA	OVERALL TEST	40	1,361	27.54	7.57	0.85	2.98	1.38
	Reading	24	1,361	17.29	4.76	0.79	2.19	1.44
	Writing	16	1,361	10.26	3.48	0.67	2.00	1.28

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.7. Spring 2023 STAAR Alternate 2 Grade 4 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	7,292	30.41	6.37	0.82	2.70	1.52
	Numerical Representations and Relationships	8	7,292	5.86	1.68	0.40	1.30	1.46
	Computations and Algebraic Relationships	8	7,292	5.64	1.64	0.37	1.31	1.41
	Geometry and Measurement	16	7,292	12.86	2.78	0.68	1.57	1.61
	Data Analysis and Personal Financial Literacy	8	7,292	6.06	1.84	0.58	1.20	1.51
RLA	OVERALL TEST	40	7,295	29.37	6.56	0.81	2.88	1.47
	Reading	24	7,295	18.48	4.10	0.74	2.10	1.54
	Writing	16	7,295	10.90	3.16	0.62	1.95	1.36

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.8. Spring 2023 STAAR Alternate 2 Grade 4 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,317	30.24	6.12	0.80	2.72	1.51
	Numerical Representations and Relationships	8	2,317	5.80	1.65	0.37	1.31	1.45
	Computations and Algebraic Relationships	8	2,317	5.56	1.62	0.34	1.32	1.39
	Geometry and Measurement	16	2,317	12.85	2.66	0.65	1.58	1.61
	Data Analysis and Personal Financial Literacy	8	2,317	6.03	1.82	0.56	1.21	1.51
RLA	OVERALL TEST	40	2,318	29.53	6.26	0.79	2.87	1.48
	Reading	24	2,318	18.59	3.92	0.71	2.10	1.55
	Writing	16	2,318	10.94	3.05	0.59	1.95	1.37

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.9. Spring 2023 STAAR Alternate 2 Grade 4 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,969	30.48	6.49	0.83	2.69	1.52
	Numerical Representations and Relationships	8	4,969	5.88	1.68	0.41	1.29	1.47
	Computations and Algebraic Relationships	8	4,969	5.67	1.66	0.38	1.30	1.42
	Geometry and Measurement	16	4,969	12.86	2.84	0.70	1.56	1.61
	Data Analysis and Personal Financial Literacy	8	4,969	6.07	1.85	0.58	1.19	1.52
RLA	OVERALL TEST	40	4,971	29.30	6.69	0.81	2.88	1.47
	Reading	24	4,971	18.42	4.18	0.75	2.10	1.54
	Writing	16	4,971	10.88	3.20	0.63	1.95	1.36

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.10. Spring 2023 STAAR Alternate 2 Grade 4 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,249	30.16	6.58	0.83	2.72	1.51
	Numerical Representations and Relationships	8	1,249	5.82	1.73	0.42	1.32	1.46
	Computations and Algebraic Relationships	8	1,249	5.57	1.71	0.42	1.30	1.39
	Geometry and Measurement	16	1,249	12.81	2.84	0.69	1.57	1.60
	Data Analysis and Personal Financial Literacy	8	1,249	5.95	1.88	0.59	1.20	1.49
RLA	OVERALL TEST	40	1,249	29.25	6.74	0.82	2.88	1.46
	Reading	24	1,249	18.45	4.16	0.74	2.11	1.54
	Writing	16	1,249	10.79	3.22	0.63	1.95	1.35

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.11. Spring 2023 STAAR Alternate 2 Grade 4 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,060	30.91	6.26	0.82	2.64	1.55
	Numerical Representations and Relationships	8	4,060	5.94	1.65	0.40	1.28	1.49
	Computations and Algebraic Relationships	8	4,060	5.73	1.61	0.36	1.29	1.43
	Geometry and Measurement	16	4,060	13.04	2.72	0.69	1.52	1.63
	Data Analysis and Personal Financial Literacy	8	4,060	6.21	1.79	0.57	1.17	1.55
RLA	OVERALL TEST	40	4,061	29.84	6.40	0.81	2.83	1.49
	Reading	24	4,061	18.80	3.99	0.74	2.05	1.57
	Writing	16	4,061	11.05	3.10	0.61	1.93	1.38

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.12. Spring 2023 STAAR Alternate 2 Grade 4 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,355	29.78	6.40	0.81	2.78	1.49
	Numerical Representations and Relationships	8	1,355	5.74	1.70	0.40	1.32	1.43
	Computations and Algebraic Relationships	8	1,355	5.52	1.67	0.37	1.32	1.38
	Geometry and Measurement	16	1,355	12.63	2.86	0.68	1.63	1.58
	Data Analysis and Personal Financial Literacy	8	1,355	5.88	1.87	0.57	1.22	1.47
RLA	OVERALL TEST	40	1,356	28.95	6.55	0.80	2.92	1.45
	Reading	24	1,356	18.12	4.14	0.73	2.14	1.51
	Writing	16	1,356	10.83	3.13	0.60	1.97	1.35

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.13. Spring 2023 STAAR Alternate 2 Grade 5 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,825	31.49	6.66	0.85	2.62	1.57
	Numerical Representations and Relationships	8	6,825	5.89	1.66	0.40	1.29	1.47
	Computations and Algebraic Relationships	8	6,825	6.44	1.66	0.56	1.10	1.61
	Geometry and Measurement	16	6,825	13.19	2.98	0.73	1.55	1.65
	Data Analysis and Personal Financial Literacy	8	6,825	5.97	1.72	0.44	1.29	1.49
RLA	OVERALL TEST	40	6,822	29.96	6.66	0.82	2.84	1.50
	Reading	24	6,822	18.48	4.12	0.74	2.10	1.54
	Writing	16	6,822	11.47	3.13	0.63	1.91	1.43
Science	OVERALL TEST	40	6,820	31.51	6.08	0.81	2.62	1.58
	Matter and Energy	8	6,820	6.26	1.67	0.48	1.20	1.57
	Force, Motion, and Energy	8	6,820	6.43	1.52	0.42	1.16	1.61
	Earth and Space	8	6,820	6.12	1.54	0.40	1.19	1.53
	Organisms and Environments	16	6,820	12.70	2.77	0.64	1.66	1.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.14. Spring 2023 STAAR Alternate 2 Grade 5 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,181	31.36	6.49	0.84	2.61	1.57
	Numerical Representations and Relationships	8	2,181	5.81	1.60	0.34	1.30	1.45
	Computations and Algebraic Relationships	8	2,181	6.45	1.63	0.56	1.08	1.61
	Geometry and Measurement	16	2,181	13.19	2.95	0.73	1.54	1.65
	Data Analysis and Personal Financial Literacy	8	2,181	5.91	1.70	0.43	1.29	1.48
RLA	OVERALL TEST	40	2,179	30.27	6.52	0.81	2.80	1.51
	Reading	24	2,179	18.61	4.00	0.73	2.08	1.55
	Writing	16	2,179	11.66	3.10	0.63	1.88	1.46
Science	OVERALL TEST	40	2,178	31.47	5.85	0.80	2.61	1.57
	Matter and Energy	8	2,178	6.33	1.65	0.49	1.18	1.58
	Force, Motion, and Energy	8	2,178	6.40	1.51	0.41	1.15	1.60
	Earth and Space	8	2,178	6.05	1.50	0.36	1.20	1.51
	Organisms and Environments	16	2,178	12.69	2.67	0.62	1.64	1.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.15. Spring 2023 STAAR Alternate 2 Grade 5 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,633	31.56	6.74	0.85	2.62	1.58
	Numerical Representations and Relationships	8	4,633	5.94	1.68	0.42	1.28	1.48
	Computations and Algebraic Relationships	8	4,633	6.44	1.68	0.57	1.10	1.61
	Geometry and Measurement	16	4,633	13.19	3.00	0.73	1.56	1.65
	Data Analysis and Personal Financial Literacy	8	4,633	6.00	1.73	0.44	1.29	1.50
RLA	OVERALL TEST	40	4,632	29.81	6.72	0.82	2.85	1.49
	Reading	24	4,632	18.43	4.17	0.74	2.11	1.54
	Writing	16	4,632	11.38	3.14	0.63	1.92	1.42
Science	OVERALL TEST	40	4,631	31.53	6.18	0.82	2.62	1.58
	Matter and Energy	8	4,631	6.23	1.67	0.48	1.20	1.56
	Force, Motion, and Energy	8	4,631	6.44	1.53	0.43	1.16	1.61
	Earth and Space	8	4,631	6.15	1.55	0.41	1.19	1.54
	Organisms and Environments	16	4,631	12.71	2.82	0.65	1.66	1.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.16. Spring 2023 STAAR Alternate 2 Grade 5 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,181	31.43	6.75	0.85	2.63	1.57
	Numerical Representations and Relationships	8	1,181	5.85	1.69	0.42	1.29	1.46
	Computations and Algebraic Relationships	8	1,181	6.46	1.67	0.57	1.10	1.61
	Geometry and Measurement	16	1,181	13.21	3.05	0.75	1.53	1.65
	Data Analysis and Personal Financial Literacy	8	1,181	5.92	1.75	0.44	1.31	1.48
RLA	OVERALL TEST	40	1,178	30.02	6.78	0.82	2.84	1.50
	Reading	24	1,178	18.54	4.19	0.75	2.10	1.55
	Writing	16	1,178	11.48	3.17	0.64	1.91	1.44
Science	OVERALL TEST	40	1,179	31.52	6.15	0.82	2.62	1.58
	Matter and Energy	8	1,179	6.34	1.68	0.50	1.19	1.59
	Force, Motion, and Energy	8	1,179	6.38	1.55	0.44	1.16	1.59
	Earth and Space	8	1,179	6.09	1.49	0.34	1.21	1.52
	Organisms and Environments	16	1,179	12.72	2.77	0.64	1.67	1.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.17. Spring 2023 STAAR Alternate 2 Grade 5 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,781	31.93	6.48	0.84	2.56	1.60
	Numerical Representations and Relationships	8	3,781	5.99	1.63	0.40	1.26	1.50
	Computations and Algebraic Relationships	8	3,781	6.54	1.60	0.55	1.07	1.63
	Geometry and Measurement	16	3,781	13.36	2.89	0.73	1.51	1.67
	Data Analysis and Personal Financial Literacy	8	3,781	6.04	1.69	0.44	1.27	1.51
RLA	OVERALL TEST	40	3,782	30.31	6.46	0.81	2.79	1.52
	Reading	24	3,782	18.71	3.99	0.73	2.06	1.56
	Writing	16	3,782	11.61	3.07	0.62	1.88	1.45
Science	OVERALL TEST	40	3,778	31.87	5.83	0.80	2.58	1.59
	Matter and Energy	8	3,778	6.31	1.61	0.46	1.18	1.58
	Force, Motion, and Energy	8	3,778	6.49	1.48	0.40	1.14	1.62
	Earth and Space	8	3,778	6.20	1.51	0.40	1.17	1.55
	Organisms and Environments	16	3,778	12.87	2.67	0.63	1.63	1.61

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.18. Spring 2023 STAAR Alternate 2 Grade 5 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,335	30.95	6.81	0.85	2.68	1.55
	Numerical Representations and Relationships	8	1,335	5.76	1.67	0.37	1.32	1.44
	Computations and Algebraic Relationships	8	1,335	6.30	1.73	0.57	1.14	1.58
	Geometry and Measurement	16	1,335	13.00	3.05	0.73	1.59	1.63
	Data Analysis and Personal Financial Literacy	8	1,335	5.89	1.72	0.42	1.31	1.47
RLA	OVERALL TEST	40	1,333	29.63	6.82	0.82	2.86	1.48
	Reading	24	1,333	18.34	4.22	0.75	2.12	1.53
	Writing	16	1,333	11.29	3.17	0.63	1.93	1.41
Science	OVERALL TEST	40	1,333	31.17	6.41	0.83	2.65	1.56
	Matter and Energy	8	1,333	6.18	1.74	0.51	1.21	1.55
	Force, Motion, and Energy	8	1,333	6.40	1.57	0.45	1.17	1.60
	Earth and Space	8	1,333	6.05	1.60	0.42	1.21	1.51
	Organisms and Environments	16	1,333	12.54	2.91	0.67	1.66	1.57

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.19. Spring 2023 STAAR Alternate 2 Grade 6 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,481	31.10	5.97	0.81	2.60	1.55
	Numerical Representations and Relationships	8	6,481	6.55	1.59	0.53	1.09	1.64
	Computations and Algebraic Relationships	16	6,481	12.28	2.75	0.62	1.69	1.53
	Geometry and Measurement	8	6,481	6.41	1.44	0.38	1.13	1.60
	Data Analysis and Personal Financial Literacy	8	6,481	5.85	1.63	0.44	1.22	1.46
RLA	OVERALL TEST	40	6,479	30.18	6.01	0.79	2.78	1.51
	Reading	24	6,479	18.39	3.66	0.67	2.09	1.53
	Writing	16	6,479	11.79	2.93	0.61	1.83	1.47

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.20. Spring 2023 STAAR Alternate 2 Grade 6 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,126	30.92	5.89	0.80	2.60	1.55
	Numerical Representations and Relationships	8	2,126	6.46	1.56	0.49	1.11	1.62
	Computations and Algebraic Relationships	16	2,126	12.20	2.75	0.62	1.69	1.53
	Geometry and Measurement	8	2,126	6.38	1.43	0.36	1.14	1.59
	Data Analysis and Personal Financial Literacy	8	2,126	5.87	1.60	0.43	1.20	1.47
RLA	OVERALL TEST	40	2,126	30.37	5.95	0.79	2.75	1.52
	Reading	24	2,126	18.47	3.63	0.67	2.09	1.54
	Writing	16	2,126	11.90	2.91	0.62	1.80	1.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.21. Spring 2023 STAAR Alternate 2 Grade 6 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,344	31.19	6.01	0.81	2.60	1.56
	Numerical Representations and Relationships	8	4,344	6.60	1.61	0.55	1.08	1.65
	Computations and Algebraic Relationships	16	4,344	12.32	2.74	0.62	1.69	1.54
	Geometry and Measurement	8	4,344	6.43	1.44	0.39	1.12	1.61
	Data Analysis and Personal Financial Literacy	8	4,344	5.85	1.64	0.44	1.23	1.46
RLA	OVERALL TEST	40	4,342	30.10	6.03	0.79	2.79	1.51
	Reading	24	4,342	18.36	3.67	0.67	2.10	1.53
	Writing	16	4,342	11.74	2.94	0.60	1.85	1.47

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.22. Spring 2023 STAAR Alternate 2 Grade 6 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,172	31.09	5.74	0.79	2.63	1.55
	Numerical Representations and Relationships	8	1,172	6.63	1.56	0.53	1.07	1.66
	Computations and Algebraic Relationships	16	1,172	12.22	2.72	0.60	1.71	1.53
	Geometry and Measurement	8	1,172	6.44	1.40	0.33	1.15	1.61
	Data Analysis and Personal Financial Literacy	8	1,172	5.79	1.64	0.45	1.22	1.45
RLA	OVERALL TEST	40	1,170	30.26	5.84	0.77	2.77	1.51
	Reading	24	1,170	18.49	3.54	0.65	2.09	1.54
	Writing	16	1,170	11.77	2.89	0.60	1.83	1.47

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.23. Spring 2023 STAAR Alternate 2 Grade 6 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,583	31.51	5.91	0.82	2.54	1.58
	Numerical Representations and Relationships	8	3,583	6.62	1.55	0.53	1.07	1.65
	Computations and Algebraic Relationships	16	3,583	12.45	2.70	0.63	1.65	1.56
	Geometry and Measurement	8	3,583	6.49	1.40	0.38	1.10	1.62
	Data Analysis and Personal Financial Literacy	8	3,583	5.96	1.61	0.45	1.19	1.49
RLA	OVERALL TEST	40	3,584	30.53	5.91	0.79	2.74	1.53
	Reading	24	3,584	18.59	3.58	0.67	2.07	1.55
	Writing	16	3,584	11.94	2.90	0.62	1.80	1.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.24. Spring 2023 STAAR Alternate 2 Grade 6 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,214	30.47	6.06	0.81	2.67	1.52
	Numerical Representations and Relationships	8	1,214	6.43	1.63	0.52	1.13	1.61
	Computations and Algebraic Relationships	16	1,214	12.02	2.80	0.62	1.74	1.50
	Geometry and Measurement	8	1,214	6.28	1.46	0.38	1.15	1.57
	Data Analysis and Personal Financial Literacy	8	1,214	5.75	1.62	0.39	1.27	1.44
RLA	OVERALL TEST	40	1,214	29.74	6.17	0.79	2.83	1.49
	Reading	24	1,214	18.10	3.81	0.69	2.13	1.51
	Writing	16	1,214	11.64	2.95	0.60	1.87	1.45

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.25. Spring 2023 STAAR Alternate 2 Grade 7 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,300	31.74	6.02	0.82	2.57	1.59
	Numerical Representations and Relationships	8	6,300	6.70	1.59	0.55	1.06	1.68
	Computations and Algebraic Relationships	16	6,300	13.13	2.76	0.68	1.57	1.64
	Geometry and Measurement	8	6,300	5.97	1.60	0.44	1.19	1.49
	Data Analysis and Personal Financial Literacy	8	6,300	5.93	1.59	0.38	1.25	1.48
RLA	OVERALL TEST	40	6,308	30.15	5.89	0.78	2.75	1.51
	Reading	24	6,308	18.93	3.71	0.70	2.04	1.58
	Writing	16	6,308	11.22	2.86	0.59	1.84	1.40

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.26. Spring 2023 STAAR Alternate 2 Grade 7 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,117	31.67	5.86	0.81	2.58	1.58
	Numerical Representations and Relationships	8	2,117	6.69	1.56	0.53	1.08	1.67
	Computations and Algebraic Relationships	16	2,117	13.08	2.69	0.65	1.58	1.63
	Geometry and Measurement	8	2,117	5.99	1.58	0.42	1.20	1.50
	Data Analysis and Personal Financial Literacy	8	2,117	5.91	1.58	0.38	1.25	1.48
RLA	OVERALL TEST	40	2,119	30.35	5.82	0.78	2.73	1.52
	Reading	24	2,119	19.00	3.70	0.71	2.01	1.58
	Writing	16	2,119	11.34	2.82	0.58	1.84	1.42

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.27. Spring 2023 STAAR Alternate 2 Grade 7 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,176	31.77	6.10	0.82	2.57	1.59
	Numerical Representations and Relationships	8	4,176	6.71	1.61	0.57	1.05	1.68
	Computations and Algebraic Relationships	16	4,176	13.16	2.80	0.69	1.57	1.64
	Geometry and Measurement	8	4,176	5.96	1.61	0.45	1.19	1.49
	Data Analysis and Personal Financial Literacy	8	4,176	5.94	1.59	0.38	1.25	1.49
RLA	OVERALL TEST	40	4,182	30.05	5.93	0.78	2.77	1.50
	Reading	24	4,182	18.89	3.72	0.70	2.05	1.57
	Writing	16	4,182	11.16	2.88	0.59	1.85	1.39

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.28. Spring 2023 STAAR Alternate 2 Grade 7 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,154	31.76	6.04	0.82	2.56	1.59
	Numerical Representations and Relationships	8	1,154	6.76	1.57	0.57	1.04	1.69
	Computations and Algebraic Relationships	16	1,154	13.13	2.74	0.67	1.56	1.64
	Geometry and Measurement	8	1,154	5.98	1.63	0.47	1.18	1.49
	Data Analysis and Personal Financial Literacy	8	1,154	5.89	1.64	0.42	1.25	1.47
RLA	OVERALL TEST	40	1,157	30.26	5.92	0.78	2.75	1.51
	Reading	24	1,157	19.07	3.74	0.71	2.02	1.59
	Writing	16	1,157	11.19	2.87	0.58	1.85	1.40

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.29. Spring 2023 STAAR Alternate 2 Grade 7 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,436	32.22	5.73	0.81	2.52	1.61
	Numerical Representations and Relationships	8	3,436	6.79	1.52	0.54	1.03	1.70
	Computations and Algebraic Relationships	16	3,436	13.33	2.64	0.66	1.53	1.67
	Geometry and Measurement	8	3,436	6.07	1.55	0.42	1.18	1.52
	Data Analysis and Personal Financial Literacy	8	3,436	6.03	1.54	0.36	1.23	1.51
RLA	OVERALL TEST	40	3,439	30.62	5.62	0.77	2.70	1.53
	Reading	24	3,439	19.19	3.54	0.68	2.00	1.60
	Writing	16	3,439	11.43	2.77	0.57	1.81	1.43

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.30. Spring 2023 STAAR Alternate 2 Grade 7 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,255	31.04	6.36	0.83	2.66	1.55
	Numerical Representations and Relationships	8	1,255	6.55	1.70	0.57	1.12	1.64
	Computations and Algebraic Relationships	16	1,255	12.86	2.92	0.69	1.63	1.61
	Geometry and Measurement	8	1,255	5.81	1.63	0.44	1.23	1.45
	Data Analysis and Personal Financial Literacy	8	1,255	5.82	1.63	0.39	1.28	1.45
RLA	OVERALL TEST	40	1,257	29.53	6.19	0.79	2.81	1.48
	Reading	24	1,257	18.58	3.91	0.72	2.08	1.55
	Writing	16	1,257	10.95	2.94	0.59	1.88	1.37

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.31. Spring 2023 STAAR Alternate 2 Grade 8 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,162	32.11	5.67	0.80	2.54	1.61
	Numerical Representations and Relationships	8	6,162	6.37	1.47	0.37	1.16	1.59
	Computations and Algebraic Relationships	16	6,162	12.68	2.67	0.62	1.65	1.58
	Geometry and Measurement	8	6,162	6.56	1.55	0.51	1.09	1.64
	Data Analysis and Personal Financial Literacy	8	6,162	6.50	1.40	0.36	1.12	1.63
RLA	OVERALL TEST	40	6,167	30.84	5.71	0.78	2.66	1.54
	Reading	24	6,167	19.17	3.53	0.69	1.96	1.60
	Writing	16	6,167	11.67	2.78	0.58	1.79	1.46
Science	OVERALL TEST	40	6,163	32.47	5.17	0.77	2.49	1.62
	Matter and Energy	8	6,163	6.45	1.49	0.41	1.14	1.61
	Force, Motion, and Energy	16	6,163	13.11	2.39	0.57	1.58	1.64
	Earth and Space	8	6,163	6.79	1.24	0.28	1.06	1.70
	Organisms and Environments	8	6,163	6.12	1.43	0.29	1.21	1.53
Social Studies	OVERALL TEST	40	6,157	32.85	5.60	0.81	2.42	1.64
	History	8	6,157	6.55	1.46	0.44	1.09	1.64
	Geography and Culture	8	6,157	6.97	1.36	0.49	0.97	1.74
	Government and Citizenship	8	6,157	6.42	1.46	0.38	1.15	1.61
	Economics, Science, Technology, and Society	16	6,157	12.90	2.56	0.62	1.58	1.61

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.32. Spring 2023 STAAR Alternate 2 Grade 8 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,117	31.87	5.74	0.80	2.54	1.59
	Numerical Representations and Relationships	8	2,117	6.31	1.49	0.39	1.16	1.58
	Computations and Algebraic Relationships	16	2,117	12.63	2.67	0.62	1.65	1.58
	Geometry and Measurement	8	2,117	6.48	1.57	0.52	1.09	1.62
	Data Analysis and Personal Financial Literacy	8	2,117	6.45	1.38	0.34	1.12	1.61
RLA	OVERALL TEST	40	2,119	31.04	5.64	0.78	2.64	1.55
	Reading	24	2,119	19.27	3.45	0.68	1.95	1.61
	Writing	16	2,119	11.77	2.77	0.59	1.77	1.47
Science	OVERALL TEST	40	2,117	32.45	5.18	0.77	2.47	1.62
	Matter and Energy	8	2,117	6.39	1.51	0.43	1.14	1.60
	Force, Motion, and Energy	16	2,117	13.14	2.36	0.56	1.56	1.64
	Earth and Space	8	2,117	6.83	1.24	0.30	1.04	1.71
	Organisms and Environments	8	2,117	6.09	1.45	0.33	1.19	1.52
Social Studies	OVERALL TEST	40	2,116	32.73	5.61	0.81	2.43	1.64
	History	8	2,116	6.53	1.46	0.44	1.10	1.63
	Geography and Culture	8	2,116	6.91	1.38	0.48	1.00	1.73
	Government and Citizenship	8	2,116	6.46	1.46	0.39	1.14	1.61
	Economics, Science, Technology, and Society	16	2,116	12.83	2.56	0.62	1.59	1.60

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.33. Spring 2023 STAAR Alternate 2 Grade 8 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,037	32.23	5.63	0.80	2.54	1.61
	Numerical Representations and Relationships	8	4,037	6.40	1.46	0.36	1.16	1.60
	Computations and Algebraic Relationships	16	4,037	12.70	2.66	0.62	1.65	1.59
	Geometry and Measurement	8	4,037	6.60	1.54	0.50	1.08	1.65
	Data Analysis and Personal Financial Literacy	8	4,037	6.53	1.41	0.37	1.12	1.63
RLA	OVERALL TEST	40	4,040	30.73	5.74	0.78	2.67	1.54
	Reading	24	4,040	19.11	3.57	0.70	1.97	1.59
	Writing	16	4,040	11.62	2.78	0.58	1.80	1.45
Science	OVERALL TEST	40	4,038	32.48	5.17	0.77	2.50	1.62
	Matter and Energy	8	4,038	6.48	1.47	0.40	1.14	1.62
	Force, Motion, and Energy	16	4,038	13.09	2.42	0.57	1.58	1.64
	Earth and Space	8	4,038	6.77	1.24	0.27	1.07	1.69
	Organisms and Environments	8	4,038	6.14	1.42	0.27	1.22	1.53
Social Studies	OVERALL TEST	40	4,033	32.91	5.60	0.82	2.41	1.65
	History	8	4,033	6.57	1.46	0.44	1.09	1.64
	Geography and Culture	8	4,033	7.01	1.36	0.50	0.96	1.75
	Government and Citizenship	8	4,033	6.40	1.46	0.38	1.15	1.60
	Economics, Science, Technology, and Society	16	4,033	12.94	2.57	0.63	1.57	1.62

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.34. Spring 2023 STAAR Alternate 2 Grade 8 Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,130	32.29	5.30	0.77	2.54	1.61
	Numerical Representations and Relationships	8	1,130	6.44	1.40	0.34	1.14	1.61
	Computations and Algebraic Relationships	16	1,130	12.66	2.56	0.58	1.67	1.58
	Geometry and Measurement	8	1,130	6.63	1.50	0.50	1.06	1.66
	Data Analysis and Personal Financial Literacy	8	1,130	6.55	1.35	0.32	1.12	1.64
RLA	OVERALL TEST	40	1,131	30.92	5.54	0.77	2.66	1.55
	Reading	24	1,131	19.27	3.45	0.68	1.95	1.61
	Writing	16	1,131	11.65	2.70	0.55	1.80	1.46
Science	OVERALL TEST	40	1,132	32.68	4.96	0.75	2.46	1.63
	Matter and Energy	8	1,132	6.51	1.43	0.38	1.12	1.63
	Force, Motion, and Energy	16	1,132	13.24	2.31	0.55	1.55	1.66
	Earth and Space	8	1,132	6.79	1.24	0.30	1.04	1.70
	Organisms and Environments	8	1,132	6.13	1.41	0.28	1.20	1.53
Social Studies	OVERALL TEST	40	1,129	33.14	5.30	0.80	2.40	1.66
	History	8	1,129	6.62	1.41	0.43	1.07	1.65
	Geography and Culture	8	1,129	6.99	1.33	0.46	0.98	1.75
	Government and Citizenship	8	1,129	6.50	1.43	0.37	1.13	1.62
	Economics, Science, Technology, and Society	16	1,129	13.04	2.39	0.57	1.57	1.63

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.35. Spring 2023 STAAR Alternate 2 Grade 8 Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,319	32.40	5.69	0.81	2.48	1.62
	Numerical Representations and Relationships	8	3,319	6.43	1.45	0.38	1.14	1.61
	Computations and Algebraic Relationships	16	3,319	12.78	2.68	0.64	1.61	1.60
	Geometry and Measurement	8	3,319	6.62	1.52	0.51	1.06	1.66
	Data Analysis and Personal Financial Literacy	8	3,319	6.57	1.38	0.38	1.09	1.64
RLA	OVERALL TEST	40	3,321	31.21	5.68	0.79	2.60	1.56
	Reading	24	3,321	19.36	3.50	0.70	1.92	1.61
	Writing	16	3,321	11.85	2.78	0.60	1.75	1.48
Science	OVERALL TEST	40	3,317	32.68	5.17	0.78	2.45	1.63
	Matter and Energy	8	3,317	6.49	1.47	0.43	1.11	1.62
	Force, Motion, and Energy	16	3,317	13.17	2.39	0.58	1.55	1.65
	Earth and Space	8	3,317	6.85	1.22	0.29	1.03	1.71
	Organisms and Environments	8	3,317	6.17	1.42	0.29	1.19	1.54
Social Studies	OVERALL TEST	40	3,315	33.11	5.49	0.81	2.37	1.66
	History	8	3,315	6.62	1.43	0.44	1.07	1.65
	Geography and Culture	8	3,315	7.03	1.33	0.50	0.94	1.76
	Government and Citizenship	8	3,315	6.47	1.43	0.38	1.13	1.62
	Economics, Science, Technology, and Society	16	3,315	13.00	2.52	0.62	1.55	1.63

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.36. Spring 2023 STAAR Alternate 2 Grade 8 White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,319	31.72	5.59	0.78	2.62	1.59
	Numerical Representations and Relationships	8	1,319	6.28	1.47	0.33	1.21	1.57
	Computations and Algebraic Relationships	16	1,319	12.59	2.62	0.58	1.70	1.57
	Geometry and Measurement	8	1,319	6.41	1.62	0.51	1.13	1.60
	Data Analysis and Personal Financial Literacy	8	1,319	6.43	1.39	0.31	1.15	1.61
RLA	OVERALL TEST	40	1,322	30.38	5.68	0.77	2.72	1.52
	Reading	24	1,322	18.91	3.56	0.68	2.01	1.58
	Writing	16	1,322	11.47	2.75	0.55	1.83	1.43
Science	OVERALL TEST	40	1,320	32.29	5.14	0.75	2.56	1.61
	Matter and Energy	8	1,320	6.39	1.52	0.40	1.17	1.60
	Force, Motion, and Energy	16	1,320	13.10	2.38	0.54	1.62	1.64
	Earth and Space	8	1,320	6.73	1.24	0.22	1.09	1.68
	Organisms and Environments	8	1,320	6.07	1.44	0.27	1.23	1.52
Social Studies	OVERALL TEST	40	1,319	32.47	5.71	0.81	2.48	1.62
	History	8	1,319	6.46	1.49	0.43	1.13	1.61
	Geography and Culture	8	1,319	6.92	1.39	0.48	1.00	1.73
	Government and Citizenship	8	1,319	6.35	1.50	0.39	1.17	1.59
	Economics, Science, Technology, and Society	16	1,319	12.74	2.63	0.63	1.61	1.59

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.37. Spring 2023 STAAR Alternate 2 Algebra I Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	6,015	30.17	6.20	0.81	2.70	1.51
	Number and Algebraic Methods	8	6,015	5.89	1.70	0.47	1.24	1.47
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	6,015	5.81	1.59	0.40	1.23	1.45
	Writing and Solving Linear Functions, Equations, and Inequalities	8	6,015	6.04	1.73	0.47	1.26	1.51
	Quadratic Functions and Equations	8	6,015	6.19	1.60	0.47	1.17	1.55
	Exponential Functions and Equations	8	6,015	6.24	1.54	0.41	1.18	1.56

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.38. Spring 2023 STAAR Alternate 2 Algebra I Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	2,009	30.20	6.12	0.81	2.69	1.51
	Number and Algebraic Methods	8	2,009	5.85	1.68	0.46	1.24	1.46
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	2,009	5.85	1.56	0.38	1.22	1.46
	Writing and Solving Linear Functions, Equations, and Inequalities	8	2,009	6.00	1.72	0.46	1.26	1.50
	Quadratic Functions and Equations	8	2,009	6.25	1.58	0.48	1.14	1.56
	Exponential Functions and Equations	8	2,009	6.24	1.52	0.38	1.20	1.56

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.39. Spring 2023 STAAR Alternate 2 Algebra I Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	3,989	30.14	6.25	0.81	2.71	1.51
	Number and Algebraic Methods	8	3,989	5.91	1.70	0.47	1.24	1.48
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	3,989	5.79	1.60	0.41	1.23	1.45
	Writing and Solving Linear Functions, Equations, and Inequalities	8	3,989	6.05	1.73	0.47	1.26	1.51
	Quadratic Functions and Equations	8	3,989	6.16	1.62	0.46	1.18	1.54
	Exponential Functions and Equations	8	3,989	6.23	1.54	0.43	1.17	1.56

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.40. Spring 2023 STAAR Alternate 2 Algebra I Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	1,070	30.06	6.17	0.81	2.72	1.50
	Number and Algebraic Methods	8	1,070	5.87	1.70	0.46	1.25	1.47
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	1,070	5.77	1.58	0.38	1.25	1.44
	Writing and Solving Linear Functions, Equations, and Inequalities	8	1,070	5.99	1.70	0.44	1.27	1.50
	Quadratic Functions and Equations	8	1,070	6.24	1.58	0.46	1.16	1.56
	Exponential Functions and Equations	8	1,070	6.19	1.57	0.43	1.18	1.55

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.41. Spring 2023 STAAR Alternate 2 Algebra I Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	3,188	30.49	6.14	0.81	2.66	1.52
	Number and Algebraic Methods	8	3,188	5.96	1.66	0.45	1.22	1.49
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	3,188	5.91	1.57	0.41	1.21	1.48
	Writing and Solving Linear Functions, Equations, and Inequalities	8	3,188	6.09	1.72	0.47	1.24	1.52
	Quadratic Functions and Equations	8	3,188	6.22	1.57	0.46	1.15	1.56
	Exponential Functions and Equations	8	3,188	6.31	1.50	0.40	1.16	1.58

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.42. Spring 2023 STAAR Alternate 2 Algebra I White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	1,318	29.91	6.17	0.80	2.74	1.50
	Number and Algebraic Methods	8	1,318	5.79	1.74	0.48	1.25	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	1,318	5.74	1.58	0.38	1.24	1.44
	Writing and Solving Linear Functions, Equations, and Inequalities	8	1,318	6.05	1.74	0.48	1.25	1.51
	Quadratic Functions and Equations	8	1,318	6.17	1.64	0.47	1.19	1.54
	Exponential Functions and Equations	8	1,318	6.16	1.54	0.40	1.20	1.54

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.43. Spring 2023 STAAR Alternate 2 English Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	6,030	30.33	6.13	0.81	2.67	1.52
	Reading	24	6,030	18.64	3.80	0.72	1.99	1.55
	Writing	16	6,030	11.70	2.91	0.63	1.78	1.46
English II	OVERALL TEST	40	5,770	30.71	6.04	0.80	2.67	1.54
	Reading	24	5,770	19.08	3.71	0.72	1.96	1.59
	Writing	16	5,770	11.64	2.92	0.61	1.81	1.45

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.44. Spring 2023 STAAR Alternate 2 English Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	2,015	30.48	6.18	0.82	2.65	1.52
	Reading	24	2,015	18.76	3.80	0.73	1.98	1.56
	Writing	16	2,015	11.72	2.94	0.64	1.76	1.47
English II	OVERALL TEST	40	1,998	30.79	5.95	0.80	2.67	1.54
	Reading	24	1,998	19.09	3.65	0.71	1.96	1.59
	Writing	16	1,998	11.69	2.89	0.61	1.81	1.46

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.45. Spring 2023 STAAR Alternate 2 English Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	3,997	30.25	6.11	0.81	2.69	1.51
	Reading	24	3,997	18.57	3.79	0.72	2.00	1.55
	Writing	16	3,997	11.67	2.89	0.62	1.79	1.46
English II	OVERALL TEST	40	3,769	30.67	6.08	0.81	2.68	1.53
	Reading	24	3,769	19.07	3.74	0.72	1.97	1.59
	Writing	16	3,769	11.60	2.93	0.62	1.81	1.45

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.46. Spring 2023 STAAR Alternate 2 English Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	1,076	30.19	6.11	0.80	2.70	1.51
	Reading	24	1,076	18.61	3.79	0.72	2.00	1.55
	Writing	16	1,076	11.58	2.93	0.62	1.80	1.45
English II	OVERALL TEST	40	1,065	30.71	5.97	0.80	2.68	1.54
	Reading	24	1,065	19.04	3.72	0.72	1.97	1.59
	Writing	16	1,065	11.66	2.84	0.59	1.81	1.46

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.47. Spring 2023 STAAR Alternate 2 English Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	3,201	30.61	6.13	0.82	2.63	1.53
	Reading	24	3,201	18.75	3.79	0.73	1.97	1.56
	Writing	16	3,201	11.85	2.88	0.63	1.75	1.48
English II	OVERALL TEST	40	3,025	31.03	6.07	0.81	2.62	1.55
	Reading	24	3,025	19.27	3.71	0.73	1.92	1.61
	Writing	16	3,025	11.77	2.93	0.63	1.77	1.47

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.48. Spring 2023 STAAR Alternate 2 English White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	1,321	30.19	6.02	0.80	2.68	1.51
	Reading	24	1,321	18.59	3.72	0.71	2.00	1.55
	Writing	16	1,321	11.60	2.88	0.61	1.79	1.45
English II	OVERALL TEST	40	1,283	30.33	6.04	0.79	2.74	1.52
	Reading	24	1,283	18.93	3.68	0.70	2.00	1.58
	Writing	16	1,283	11.40	2.96	0.60	1.87	1.43

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.49. Spring 2023 STAAR Alternate 2 Biology Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	6,041	33.40	5.99	0.84	2.36	1.67
	Cell Structure and Function	8	6,041	6.61	1.66	0.60	1.04	1.65
	Mechanisms of Genetics	8	6,041	6.64	1.56	0.50	1.10	1.66
	Biological Evolution and Classification	8	6,041	7.21	1.32	0.58	0.86	1.80
	Biological Processes and Systems	8	6,041	6.27	1.54	0.41	1.18	1.57
	Interdependence within Environmental Systems	8	6,041	6.67	1.56	0.53	1.07	1.67

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.50. Spring 2023 STAAR Alternate 2 Biology Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	2,083	33.37	6.02	0.85	2.35	1.67
	Cell Structure and Function	8	2,083	6.59	1.66	0.61	1.04	1.65
	Mechanisms of Genetics	8	2,083	6.68	1.55	0.51	1.09	1.67
	Biological Evolution and Classification	8	2,083	7.24	1.33	0.60	0.84	1.81
	Biological Processes and Systems	8	2,083	6.24	1.51	0.39	1.18	1.56
	Interdependence within Environmental Systems	8	2,083	6.62	1.59	0.54	1.07	1.66

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.51. Spring 2023 STAAR Alternate 2 Biology Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	3,942	33.41	5.98	0.84	2.37	1.67
	Cell Structure and Function	8	3,942	6.62	1.65	0.60	1.05	1.66
	Mechanisms of Genetics	8	3,942	6.62	1.57	0.50	1.11	1.66
	Biological Evolution and Classification	8	3,942	7.19	1.32	0.57	0.87	1.80
	Biological Processes and Systems	8	3,942	6.28	1.55	0.43	1.18	1.57
	Interdependence within Environmental Systems	8	3,942	6.69	1.54	0.52	1.06	1.67

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.52. Spring 2023 STAAR Alternate 2 Biology Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	1,059	33.37	6.06	0.85	2.37	1.67
	Cell Structure and Function	8	1,059	6.66	1.63	0.59	1.05	1.66
	Mechanisms of Genetics	8	1,059	6.66	1.59	0.53	1.09	1.67
	Biological Evolution and Classification	8	1,059	7.20	1.31	0.56	0.87	1.80
	Biological Processes and Systems	8	1,059	6.22	1.55	0.40	1.20	1.55
	Interdependence within Environmental Systems	8	1,059	6.63	1.59	0.54	1.07	1.66

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.53. Spring 2023 STAAR Alternate 2 Biology Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	3,170	33.56	5.88	0.84	2.33	1.68
	Cell Structure and Function	8	3,170	6.65	1.62	0.60	1.03	1.66
	Mechanisms of Genetics	8	3,170	6.65	1.55	0.50	1.09	1.66
	Biological Evolution and Classification	8	3,170	7.25	1.28	0.58	0.83	1.81
	Biological Processes and Systems	8	3,170	6.31	1.52	0.42	1.16	1.58
	Interdependence within Environmental Systems	8	3,170	6.71	1.53	0.52	1.05	1.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.54. Spring 2023 STAAR Alternate 2 Biology White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	1,353	33.61	5.94	0.84	2.34	1.68
	Cell Structure and Function	8	1,353	6.64	1.68	0.63	1.02	1.66
	Mechanisms of Genetics	8	1,353	6.73	1.51	0.48	1.09	1.68
	Biological Evolution and Classification	8	1,353	7.23	1.35	0.62	0.83	1.81
	Biological Processes and Systems	8	1,353	6.31	1.52	0.42	1.16	1.58
	Interdependence within Environmental Systems	8	1,353	6.70	1.55	0.53	1.06	1.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.55. Spring 2023 STAAR Alternate 2 U.S. History Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	5,260	32.73	5.49	0.81	2.40	1.64
	History	8	5,260	6.52	1.42	0.40	1.10	1.63
	Geography and Culture	16	5,260	13.10	2.62	0.66	1.53	1.64
	Government and Citizenship	8	5,260	6.36	1.43	0.37	1.13	1.59
	Economics, Science, Technology, and Society	8	5,260	6.74	1.35	0.43	1.02	1.69

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.56. Spring 2023 STAAR Alternate 2 U.S. History Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	1,819	32.58	5.41	0.80	2.41	1.63
	History	8	1,819	6.51	1.41	0.40	1.09	1.63
	Geography and Culture	16	1,819	12.99	2.56	0.64	1.54	1.62
	Government and Citizenship	8	1,819	6.32	1.42	0.35	1.14	1.58
	Economics, Science, Technology, and Society	8	1,819	6.75	1.34	0.42	1.02	1.69

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.57. Spring 2023 STAAR Alternate 2 U.S. History Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	3,435	32.81	5.53	0.81	2.40	1.64
	History	8	3,435	6.52	1.42	0.40	1.10	1.63
	Geography and Culture	16	3,435	13.16	2.65	0.67	1.52	1.65
	Government and Citizenship	8	3,435	6.38	1.43	0.38	1.13	1.60
	Economics, Science, Technology, and Society	8	3,435	6.74	1.35	0.44	1.01	1.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.58. Spring 2023 STAAR Alternate 2 U.S. History Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	958	32.68	5.42	0.80	2.42	1.63
	History	8	958	6.52	1.40	0.38	1.10	1.63
	Geography and Culture	16	958	13.12	2.61	0.65	1.53	1.64
	Government and Citizenship	8	958	6.32	1.42	0.37	1.13	1.58
	Economics, Science, Technology, and Society	8	958	6.73	1.37	0.43	1.03	1.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.59. Spring 2023 STAAR Alternate 2 U.S. History Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	2,794	32.93	5.46	0.81	2.36	1.65
	History	8	2,794	6.57	1.42	0.42	1.08	1.64
	Geography and Culture	16	2,794	13.18	2.58	0.66	1.49	1.65
	Government and Citizenship	8	2,794	6.40	1.42	0.39	1.11	1.60
	Economics, Science, Technology, and Society	8	2,794	6.77	1.34	0.44	1.00	1.69

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.60. Spring 2023 STAAR Alternate 2 U.S. History White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	1,164	32.70	5.44	0.80	2.43	1.64
	History	8	1,164	6.47	1.41	0.37	1.12	1.62
	Geography and Culture	16	1,164	13.09	2.64	0.66	1.55	1.64
	Government and Citizenship	8	1,164	6.38	1.39	0.31	1.15	1.60
	Economics, Science, Technology, and Society	8	1,164	6.76	1.32	0.41	1.01	1.69

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient Alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

2023 STAAR Alternate 2 Progress Measure

During the 2022-2023 school year, Progress Measures were not calculated or reported. This section is left as a placeholder to keep the numberings consistent.

2023 STAAR Alternate 2
Scale Score
Descriptive Statistics

Glossary

This glossary provides definitions for the statistical terms that appear in the tables and graphs in this section. Definitions of statistical terms and concepts in the other sections are given in Chapter 3 and Chapter 5.

Descriptive Statistics

Mean The mean is a measure of central tendency. It is the average score for the assessment. It is computed by summing the scores of all students and dividing the sum by the total number of students (*N*).

Median The median is another measure of central tendency. It is the score at the middle of the frequency distribution for the assessment. It is computed by finding the score at which there are the same number of scores above as there are below.

Mode The mode is another measure of central tendency. It is the most frequently obtained score for the assessment. It is determined by computing the frequency distribution and finding the score point with the highest frequency (*n*-count).

Range The range is a measure of statistical dispersion (variability or spread). It is the difference between the lowest and highest scores obtained by students on the assessment. It is computed by subtracting the lowest score from the highest score.

Interquartile Range The interquartile range is another measure of statistical dispersion (variability or spread). It is the difference between the first and third quartiles (or 25th and 75th percentiles) of the score distribution for the assessment. It is computed by subtracting the score at the first quartile (the point that splits the lowest 25 percent of the scores) from the score at the third quartile (the point that splits the highest 25 percent of the scores).

Standard Deviation The standard deviation is another measure of statistical dispersion (variability or spread). It is an indicator of the degree of score variation around the mean. It is computed using the following formula:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, and N is the total number of students who took the assessment.

Variance The variance is another measure of statistical dispersion (variability or spread) around the mean. It is computed as the square of the standard deviation.

Skewness The skewness is an indicator of the shape of the score distribution. It measures the extent to which the score distribution "leans" to one side of the mean. A positive skewness indicates that the score

distribution leans below the mean. A negative skewness indicates that the score distribution leans above the mean. A skewness of zero indicates that the score distribution is symmetric around the mean. It is computed using the following formula:

$$Skewness = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^3,$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Kurtosis The kurtosis is another indicator of the shape of the score distribution. It measures the "peakedness" of the score distribution. A positive kurtosis is referred to as *leptokurtic*, meaning that the distribution has a more acute peak around the mean and fatter tails. A negative kurtosis is called *platykurtic*, meaning the distribution has a lower, wider peak around the mean and thinner tails. It is computed using the following formula:

$$Kurtosis = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Frequency Distributions

Frequency (Freq) The frequency is the number of students who obtained the particular score point on the assessment.

Cumulative Frequency (CumFreq) The cumulative frequency is the number of students who obtained a score that is less than or equal to the particular score point on the assessment.

Percentage (Pct) The percentage is the percentage of students who obtained the particular score point on the assess-ment. It is computed as: Pct = Freq ÷ N × 100.

Cumulative Percentage (CumPct) The cumulative percentage is the percentage of students who obtained a score that is less than or equal to the particular score point on the assessment. It is computed as: CumPct = CumFreq ÷ N × 100.

Table C.6.1. Spring 2023 STAAR Alternate 2 Grades 3–8 Assessments Scale Score Descriptive Statistics

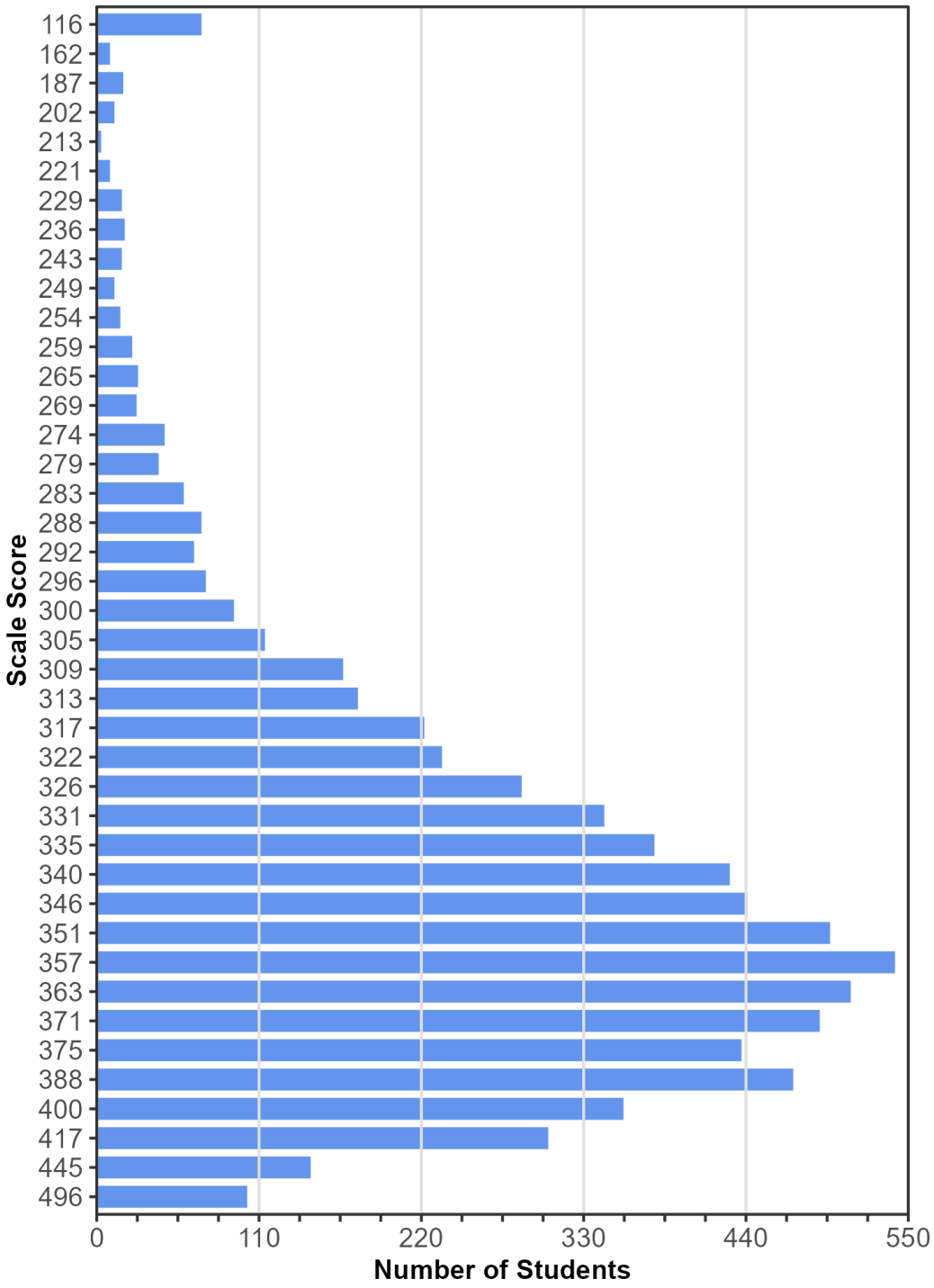
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	7,385	348.93	351	357	380	45	47.99	2303.06	-0.93	5.56
Grade 4	Mathematics	7,292	357.49	354	374	381	49	44.03	1938.68	-0.68	5.79
Grade 5	Mathematics	6,825	358.77	356	379	375	50	50.83	2584.16	-0.07	2.76
Grade 6	Mathematics	6,481	366.44	362	362	420	50	51.52	2653.86	0.16	3.08
Grade 7	Mathematics	6,300	368.18	364	375	403	51	51.79	2682.15	0.16	3.04
Grade 8	Mathematics	6,162	365.30	360	388	405	53	49.99	2498.61	0.16	3.01
Grade 3	RLA	7,389	343.99	345	350	425	46	50.11	2510.71	-1.05	6.09
Grade 4	RLA	7,295	345.89	342	355	422	51	49.44	2444.04	-0.27	4.85
Grade 5	RLA	6,822	345.01	343	374	412	54	49.62	2462.29	-0.27	3.90
Grade 6	RLA	6,479	350.05	346	346	423	48	49.78	2478.07	0.34	3.46
Grade 7	RLA	6,308	356.11	354	361	436	50	49.57	2457.35	0.06	4.24
Grade 8	RLA	6,167	355.44	354	354	421	53	49.79	2478.70	0.35	3.53
Grade 5	Science	6,820	364.60	365	365	395	45	46.63	2174.75	-0.42	4.95
Grade 8	Science	6,163	366.59	362	369	344	38	39.70	1575.95	0.18	3.84
Grade 8	Social Studies	6,157	365.70	362	380	383	56	48.37	2339.97	-0.04	2.79

Table C.6.2. Spring 2023 STAAR Alternate 2 EOC Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	6,015	350.15	348	342	408	45	47.04	2212.57	-0.55	5.92
English I	6,030	361.25	358	358	428	57	49.98	2498.27	-0.28	4.55
English II	5,770	362.59	362	362	427	51	49.74	2474.46	-0.13	4.51
Biology	6,041	371.36	368	377	340	60	47.78	2282.79	0.01	2.08
US History	5,260	363.86	363	368	384	58	47.17	2225.42	-0.08	3.50

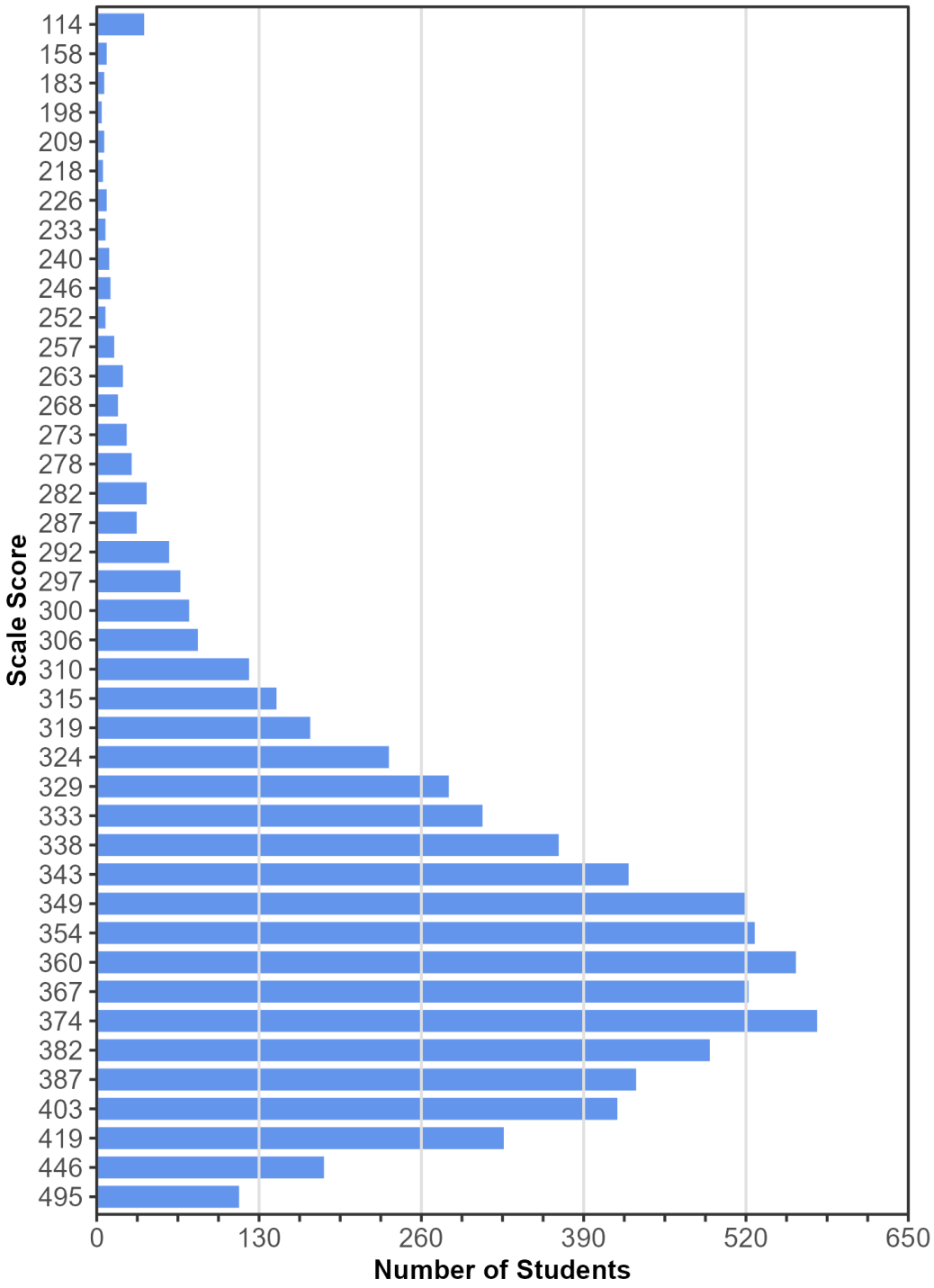
2023 STAAR Alternate 2 Frequency Distribution of Scale Scores

Figure C.6.1. Spring 2023 STAAR Alternate 2 Grade 3 Mathematics
Frequency Distribution of Scale Scores
All Students



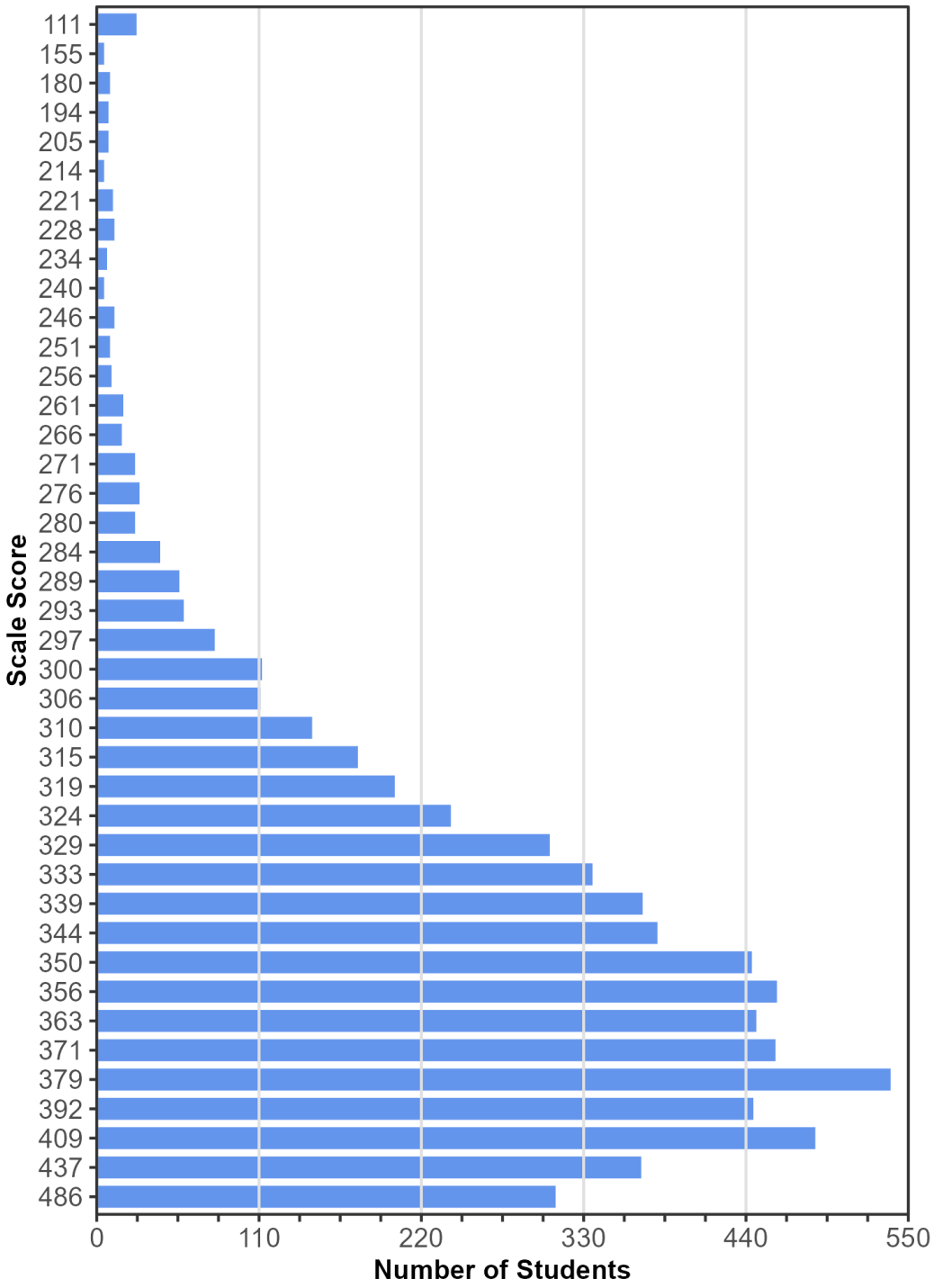
Scale Score	Freq	Cum Freq	Pct	Cum Pct
116	71	71	0.96	0.96
162	9	80	0.12	1.08
187	18	98	0.24	1.33
202	12	110	0.16	1.49
213	3	113	0.04	1.53
221	9	122	0.12	1.65
229	17	139	0.23	1.88
236	19	158	0.26	2.14
243	17	175	0.23	2.37
249	12	187	0.16	2.53
254	16	203	0.22	2.75
259	24	227	0.32	3.07
265	28	255	0.38	3.45
269	27	282	0.37	3.82
274	46	328	0.62	4.44
279	42	370	0.57	5.01
283	59	429	0.80	5.81
288	71	500	0.96	6.77
292	66	566	0.89	7.66
296	74	640	1.00	8.67
300	93	733	1.26	9.93
305	114	847	1.54	11.47
309	167	1014	2.26	13.73
313	177	1191	2.40	16.13
317	222	1413	3.01	19.13
322	234	1647	3.17	22.30
326	288	1935	3.90	26.20
331	344	2279	4.66	30.86
335	378	2657	5.12	35.98
340	429	3086	5.81	41.79
346	441	3527	5.97	47.76
351	497	4024	6.73	54.49
357	541	4565	7.33	61.81
363	511	5076	6.92	68.73
371	490	5566	6.64	75.37
375	437	6003	5.92	81.29
388	472	6475	6.39	87.68
400	357	6832	4.83	92.51
417	306	7138	4.14	96.66
445	145	7283	1.96	98.62
496	102	7385	1.38	100.00

Figure C.6.2. Spring 2023 STAAR Alternate 2 Grade 4 Mathematics
Frequency Distribution of Scale Scores
All Students



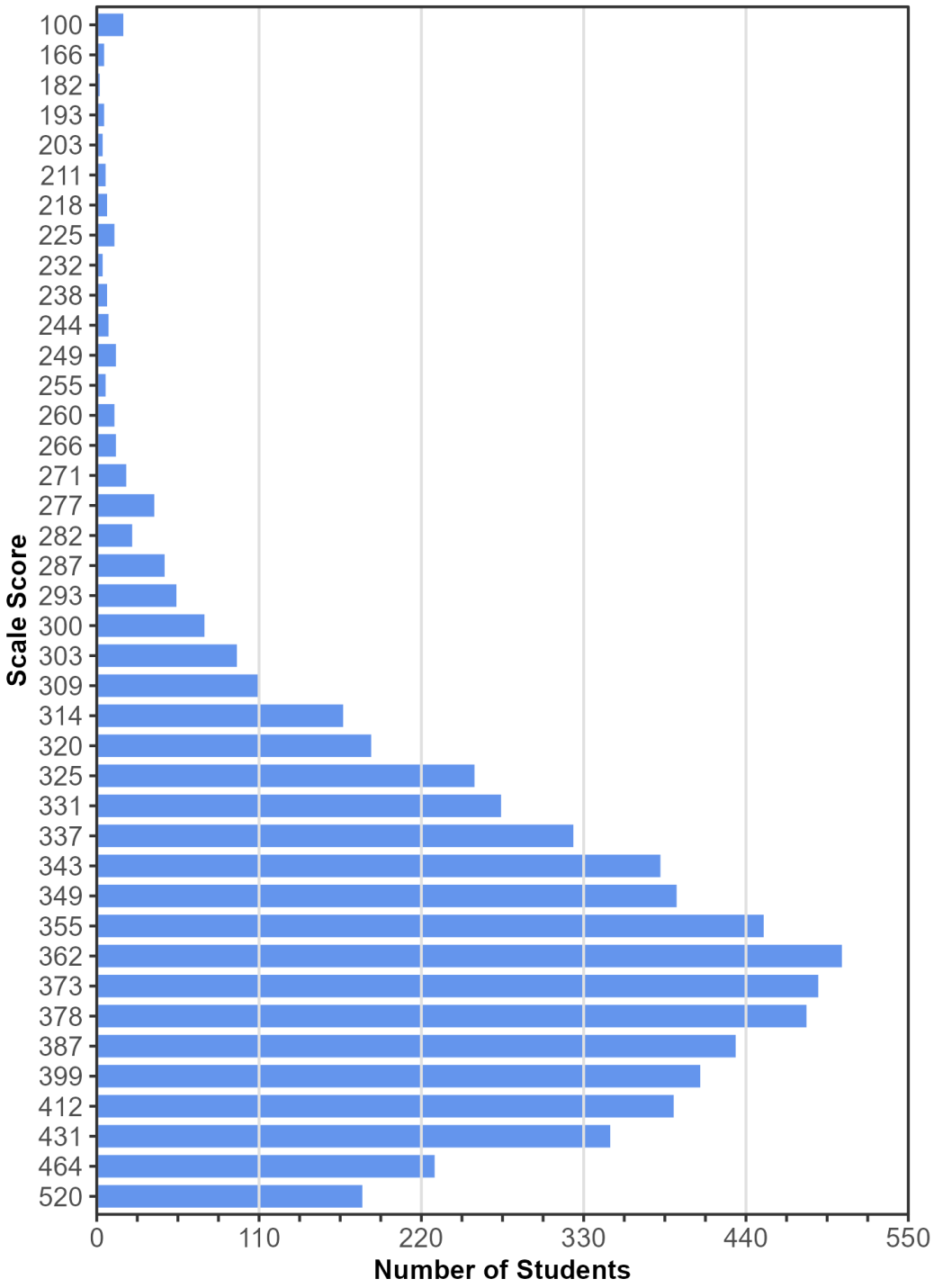
	Freq	Cum Freq	Pct	Cum Pct
114	38	38	0.52	0.52
158	8	46	0.11	0.63
183	6	52	0.08	0.71
198	4	56	0.05	0.77
209	6	62	0.08	0.85
218	5	67	0.07	0.92
226	8	75	0.11	1.03
233	7	82	0.10	1.12
240	10	92	0.14	1.26
246	11	103	0.15	1.41
252	7	110	0.10	1.51
257	14	124	0.19	1.70
263	21	145	0.29	1.99
268	17	162	0.23	2.22
273	24	186	0.33	2.55
278	28	214	0.38	2.93
282	40	254	0.55	3.48
287	32	286	0.44	3.92
292	58	344	0.80	4.72
297	67	411	0.92	5.64
300	74	485	1.01	6.65
306	81	566	1.11	7.76
310	122	688	1.67	9.43
315	144	832	1.97	11.41
319	171	1003	2.35	13.75
324	234	1237	3.21	16.96
329	282	1519	3.87	20.83
333	309	1828	4.24	25.07
338	370	2198	5.07	30.14
343	426	2624	5.84	35.98
349	520	3144	7.13	43.12
354	527	3671	7.23	50.34
360	560	4231	7.68	58.02
367	522	4753	7.16	65.18
374	577	5330	7.91	73.09
382	491	5821	6.73	79.83
387	432	6253	5.92	85.75
403	417	6670	5.72	91.47
419	326	6996	4.47	95.94
446	182	7178	2.50	98.44
495	114	7292	1.56	100.00

Figure C.6.3. Spring 2023 STAAR Alternate 2 Grade 5 Mathematics
Frequency Distribution of Scale Scores
All Students



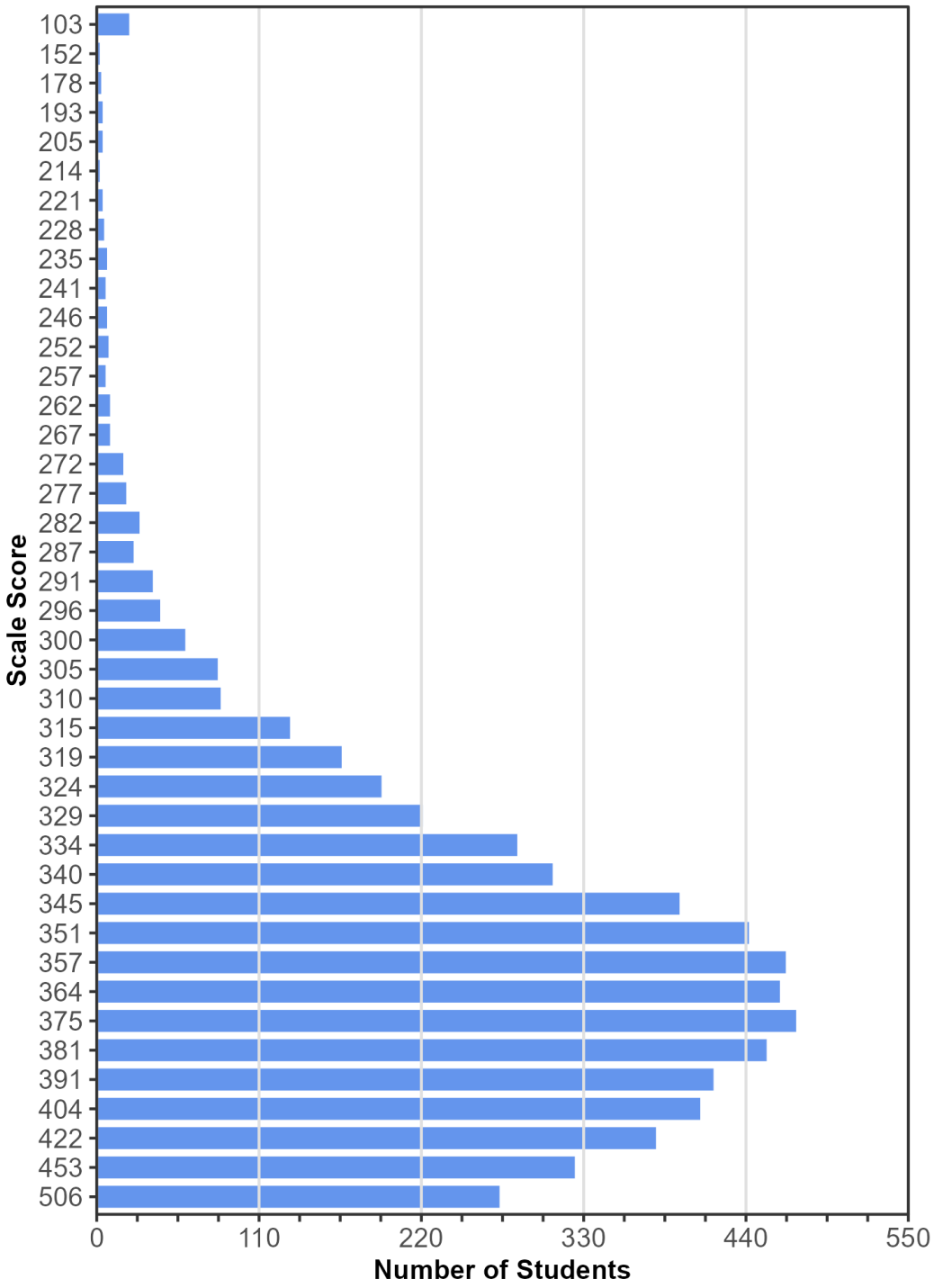
Scale Score	Freq	Cum Freq	Pct	Cum Pct
111	27	27	0.40	0.40
155	5	32	0.07	0.47
180	9	41	0.13	0.60
194	8	49	0.12	0.72
205	8	57	0.12	0.84
214	5	62	0.07	0.91
221	11	73	0.16	1.07
228	12	85	0.18	1.25
234	7	92	0.10	1.35
240	5	97	0.07	1.42
246	12	109	0.18	1.60
251	9	118	0.13	1.73
256	10	128	0.15	1.88
261	18	146	0.26	2.14
266	17	163	0.25	2.39
271	26	189	0.38	2.77
276	29	218	0.42	3.19
280	26	244	0.38	3.58
284	43	287	0.63	4.21
289	56	343	0.82	5.03
293	59	402	0.86	5.89
297	80	482	1.17	7.06
300	112	594	1.64	8.70
306	111	705	1.63	10.33
310	146	851	2.14	12.47
315	177	1028	2.59	15.06
319	202	1230	2.96	18.02
324	240	1470	3.52	21.54
329	307	1777	4.50	26.04
333	336	2113	4.92	30.96
339	370	2483	5.42	36.38
344	380	2863	5.57	41.95
350	444	3307	6.51	48.45
356	461	3768	6.75	55.21
363	447	4215	6.55	61.76
371	460	4675	6.74	68.50
379	538	5213	7.88	76.38
392	445	5658	6.52	82.90
409	487	6145	7.14	90.04
437	369	6514	5.41	95.44
486	311	6825	4.56	100.00

Figure C.6.4. Spring 2023 STAAR Alternate 2 Grade 6 Mathematics
Frequency Distribution of Scale Scores
All Students



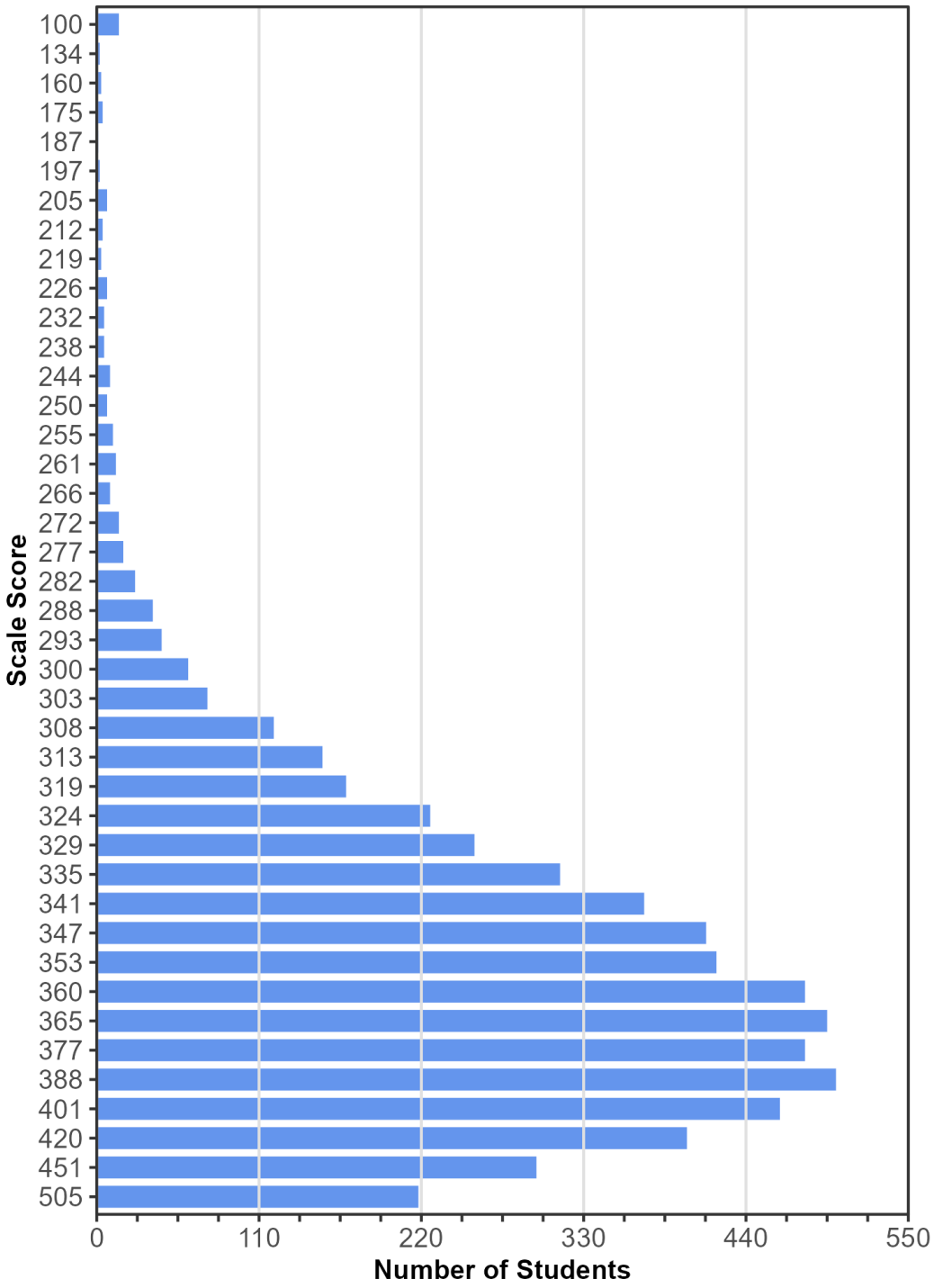
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	18	18	0.28	0.28
166	5	23	0.08	0.35
182	2	25	0.03	0.39
193	5	30	0.08	0.46
203	4	34	0.06	0.52
211	6	40	0.09	0.62
218	7	47	0.11	0.73
225	12	59	0.19	0.91
232	4	63	0.06	0.97
238	7	70	0.11	1.08
244	8	78	0.12	1.20
249	13	91	0.20	1.40
255	6	97	0.09	1.50
260	12	109	0.19	1.68
266	13	122	0.20	1.88
271	20	142	0.31	2.19
277	39	181	0.60	2.79
282	24	205	0.37	3.16
287	46	251	0.71	3.87
293	54	305	0.83	4.71
300	73	378	1.13	5.83
303	95	473	1.47	7.30
309	110	583	1.70	9.00
314	167	750	2.58	11.57
320	186	936	2.87	14.44
325	256	1192	3.95	18.39
331	274	1466	4.23	22.62
337	323	1789	4.98	27.60
343	382	2171	5.89	33.50
349	393	2564	6.06	39.56
355	452	3016	6.97	46.54
362	505	3521	7.79	54.33
373	489	4010	7.55	61.87
378	481	4491	7.42	69.29
387	433	4924	6.68	75.98
399	409	5333	6.31	82.29
412	391	5724	6.03	88.32
431	348	6072	5.37	93.69
464	229	6301	3.53	97.22
520	180	6481	2.78	100.00

Figure C.6.5. Spring 2023 STAAR Alternate 2 Grade 7 Mathematics
Frequency Distribution of Scale Scores
All Students



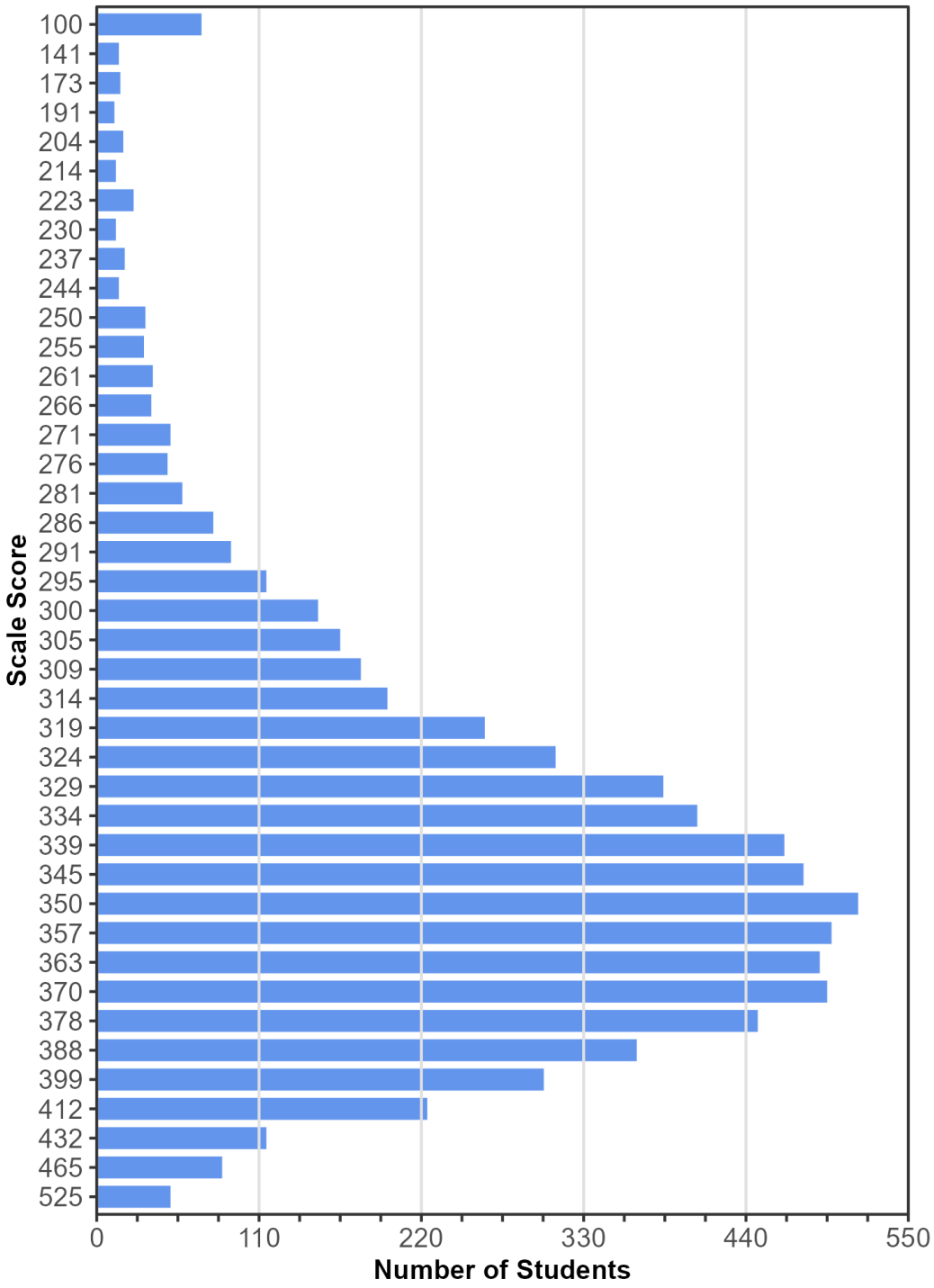
Scale Score	Freq	Cum Freq	Pct	Cum Pct
103	22	22	0.35	0.35
152	2	24	0.03	0.38
178	3	27	0.05	0.43
193	4	31	0.06	0.49
205	4	35	0.06	0.56
214	2	37	0.03	0.59
221	4	41	0.06	0.65
228	5	46	0.08	0.73
235	7	53	0.11	0.84
241	6	59	0.10	0.94
246	7	66	0.11	1.05
252	8	74	0.13	1.17
257	6	80	0.10	1.27
262	9	89	0.14	1.41
267	9	98	0.14	1.56
272	18	116	0.29	1.84
277	20	136	0.32	2.16
282	29	165	0.46	2.62
287	25	190	0.40	3.02
291	38	228	0.60	3.62
296	43	271	0.68	4.30
300	60	331	0.95	5.25
305	82	413	1.30	6.56
310	84	497	1.33	7.89
315	131	628	2.08	9.97
319	166	794	2.63	12.60
324	193	987	3.06	15.67
329	221	1208	3.51	19.17
334	285	1493	4.52	23.70
340	309	1802	4.90	28.60
345	395	2197	6.27	34.87
351	442	2639	7.02	41.89
357	467	3106	7.41	49.30
364	463	3569	7.35	56.65
375	474	4043	7.52	64.17
381	454	4497	7.21	71.38
391	418	4915	6.63	78.02
404	409	5324	6.49	84.51
422	379	5703	6.02	90.52
453	324	6027	5.14	95.67
506	273	6300	4.33	100.00

Figure C.6.6. Spring 2023 STAAR Alternate 2 Grade 8 Mathematics
Frequency Distribution of Scale Scores
All Students



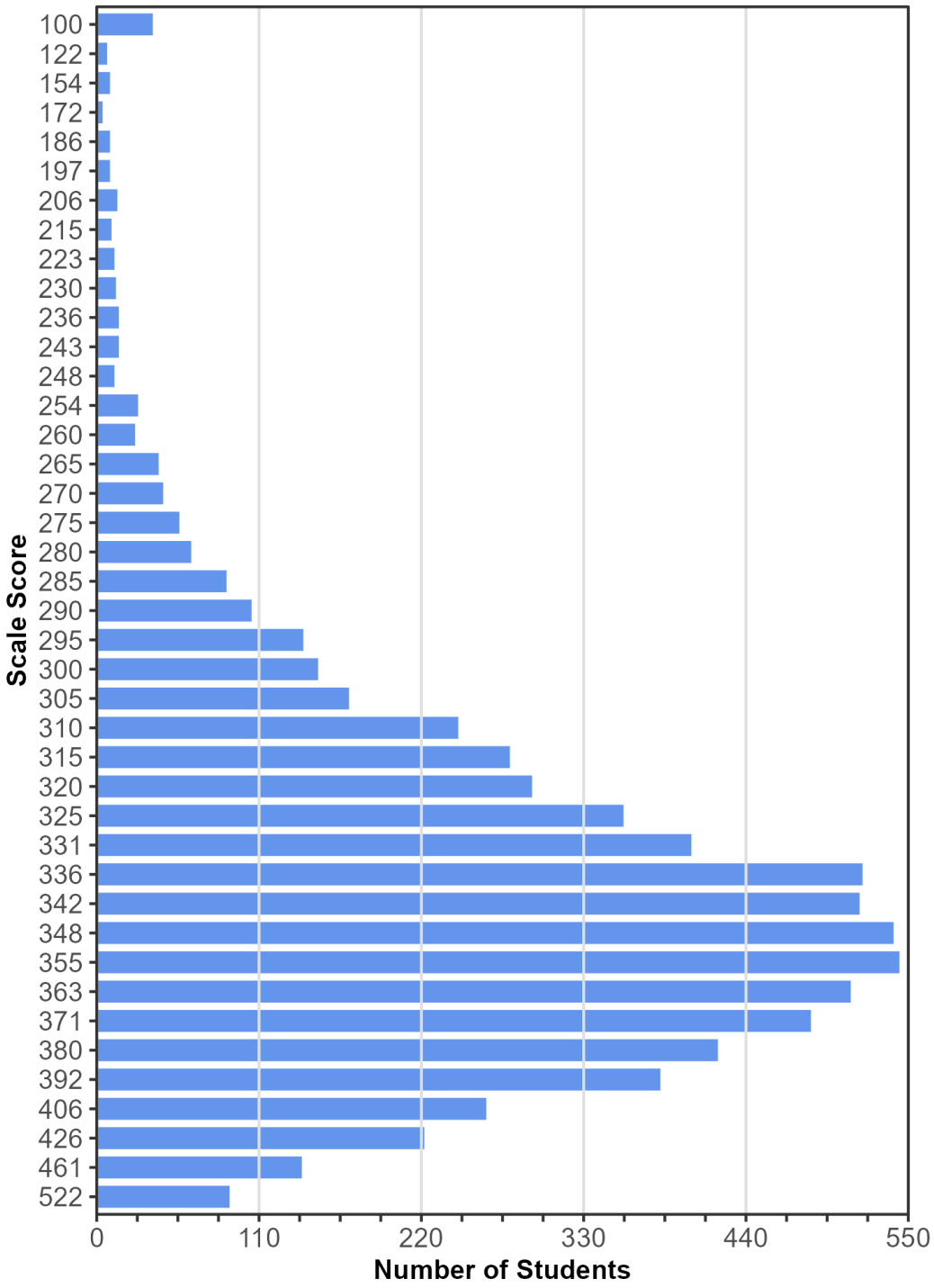
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	15	15	0.24	0.24
134	2	17	0.03	0.28
160	3	20	0.05	0.32
175	4	24	0.06	0.39
187	1	25	0.02	0.41
197	2	27	0.03	0.44
205	7	34	0.11	0.55
212	4	38	0.06	0.62
219	3	41	0.05	0.67
226	7	48	0.11	0.78
232	5	53	0.08	0.86
238	5	58	0.08	0.94
244	9	67	0.15	1.09
250	7	74	0.11	1.20
255	11	85	0.18	1.38
261	13	98	0.21	1.59
266	9	107	0.15	1.74
272	15	122	0.24	1.98
277	18	140	0.29	2.27
282	26	166	0.42	2.69
288	38	204	0.62	3.31
293	44	248	0.71	4.02
300	62	310	1.01	5.03
303	75	385	1.22	6.25
308	120	505	1.95	8.20
313	153	658	2.48	10.68
319	169	827	2.74	13.42
324	226	1053	3.67	17.09
329	256	1309	4.15	21.24
335	314	1623	5.10	26.34
341	371	1994	6.02	32.36
347	413	2407	6.70	39.06
353	420	2827	6.82	45.88
360	480	3307	7.79	53.67
365	495	3802	8.03	61.70
377	480	4282	7.79	69.49
388	501	4783	8.13	77.62
401	463	5246	7.51	85.13
420	400	5646	6.49	91.63
451	298	5944	4.84	96.46
505	218	6162	3.54	100.00

Figure C.6.7. Spring 2023 STAAR Alternate 2 Grade 3 RLA
Frequency Distribution of Scale Scores
All Students



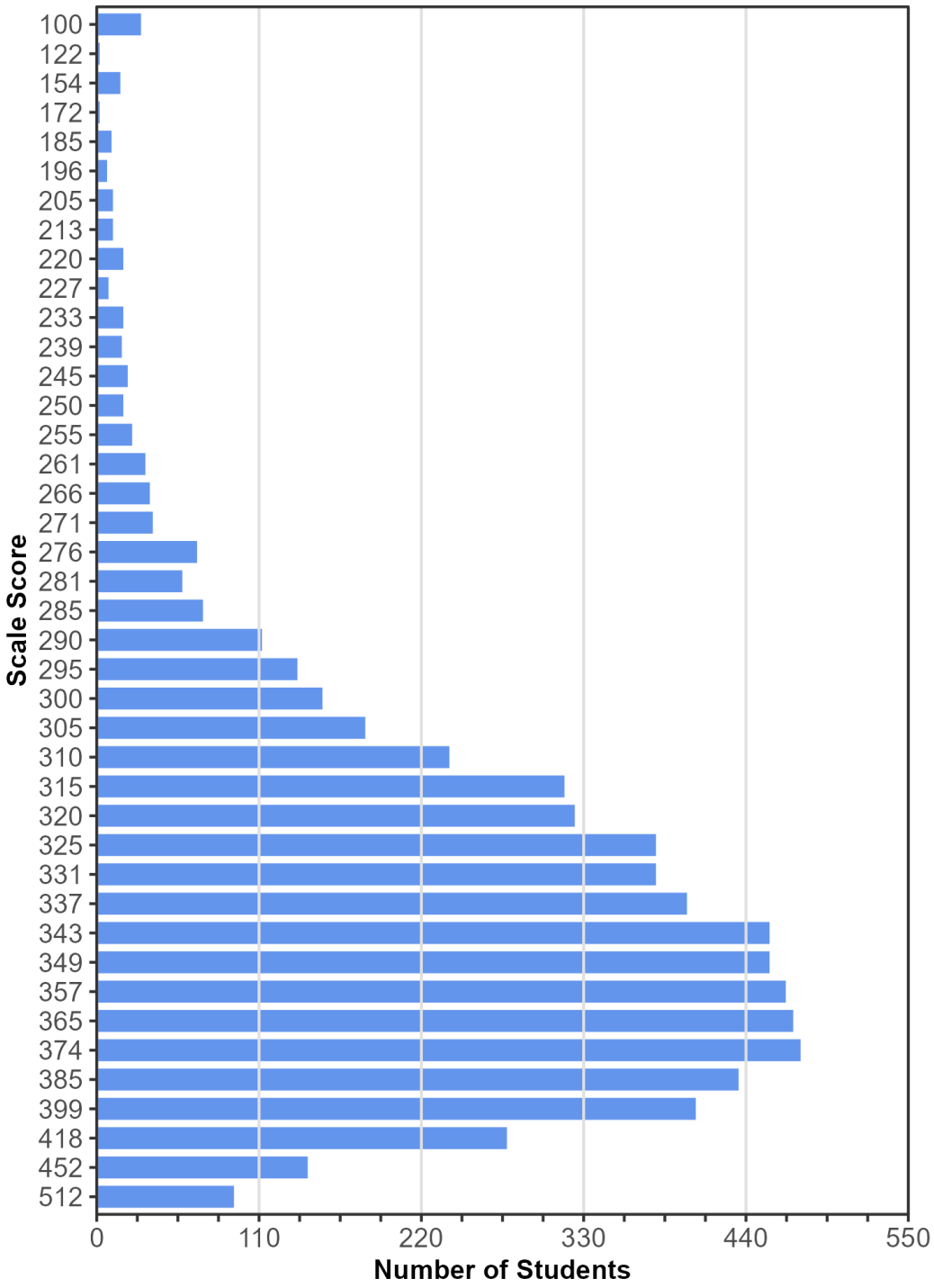
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	71	71	0.96	0.96
141	15	86	0.20	1.16
173	16	102	0.22	1.38
191	12	114	0.16	1.54
204	18	132	0.24	1.79
214	13	145	0.18	1.96
223	25	170	0.34	2.30
230	13	183	0.18	2.48
237	19	202	0.26	2.73
244	15	217	0.20	2.94
250	33	250	0.45	3.38
255	32	282	0.43	3.82
261	38	320	0.51	4.33
266	37	357	0.50	4.83
271	50	407	0.68	5.51
276	48	455	0.65	6.16
281	58	513	0.78	6.94
286	79	592	1.07	8.01
291	91	683	1.23	9.24
295	115	798	1.56	10.80
300	150	948	2.03	12.83
305	165	1113	2.23	15.06
309	179	1292	2.42	17.49
314	197	1489	2.67	20.15
319	263	1752	3.56	23.71
324	311	2063	4.21	27.92
329	384	2447	5.20	33.12
334	407	2854	5.51	38.62
339	466	3320	6.31	44.93
345	479	3799	6.48	51.41
350	516	4315	6.98	58.40
357	498	4813	6.74	65.14
363	490	5303	6.63	71.77
370	495	5798	6.70	78.47
378	448	6246	6.06	84.53
388	366	6612	4.95	89.48
399	303	6915	4.10	93.59
412	224	7139	3.03	96.62
432	115	7254	1.56	98.17
465	85	7339	1.15	99.32
525	50	7389	0.68	100.00

Figure C.6.8. Spring 2023 STAAR Alternate 2 Grade 4 RLA
Frequency Distribution of Scale Scores
All Students



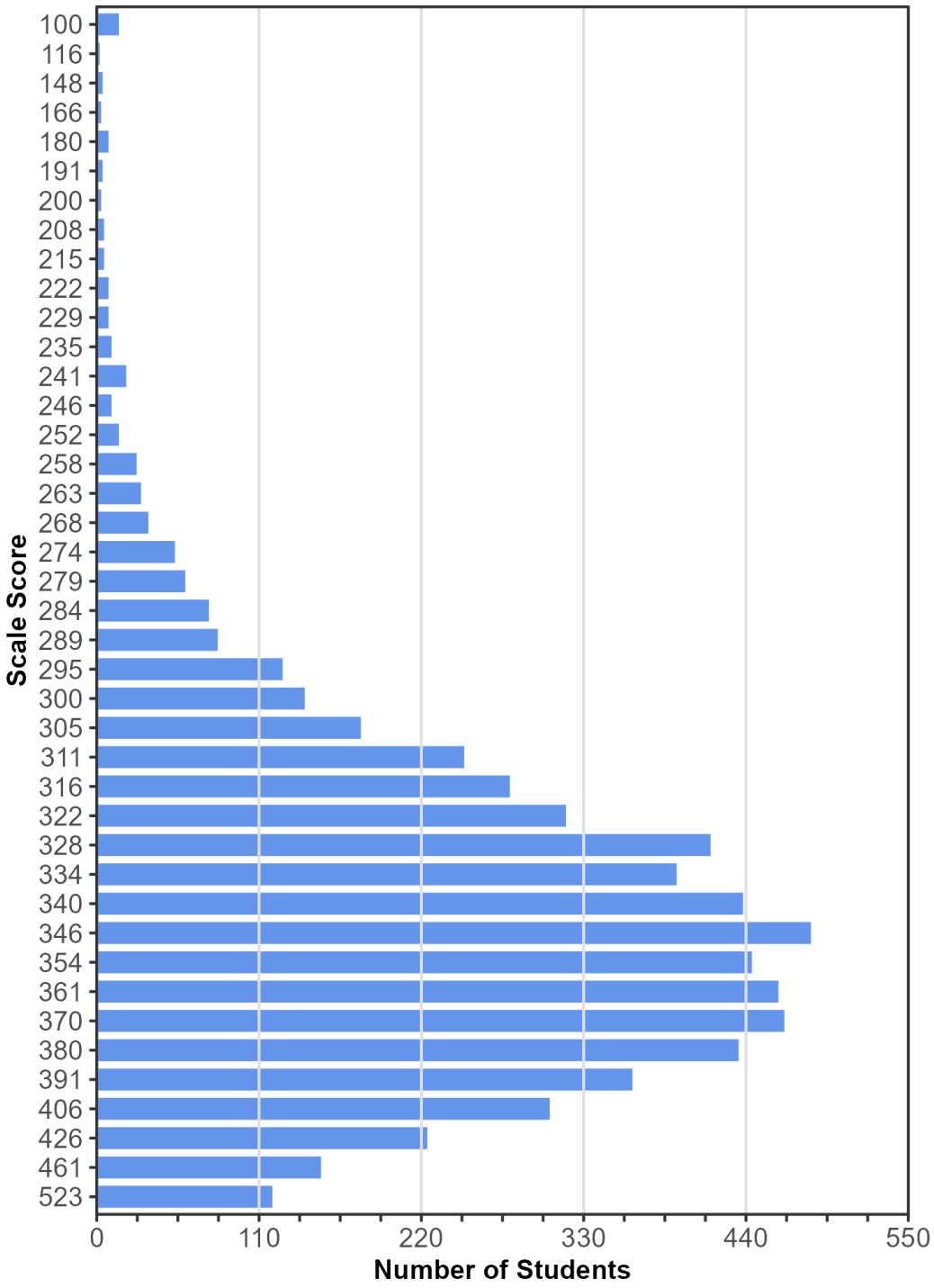
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	38	38	0.52	0.52
122	7	45	0.10	0.62
154	9	54	0.12	0.74
172	4	58	0.05	0.80
186	9	67	0.12	0.92
197	9	76	0.12	1.04
206	14	90	0.19	1.23
215	10	100	0.14	1.37
223	12	112	0.16	1.54
230	13	125	0.18	1.71
236	15	140	0.21	1.92
243	15	155	0.21	2.12
248	12	167	0.16	2.29
254	28	195	0.38	2.67
260	26	221	0.36	3.03
265	42	263	0.58	3.61
270	45	308	0.62	4.22
275	56	364	0.77	4.99
280	64	428	0.88	5.87
285	88	516	1.21	7.07
290	105	621	1.44	8.51
295	140	761	1.92	10.43
300	150	911	2.06	12.49
305	171	1082	2.34	14.83
310	245	1327	3.36	18.19
315	280	1607	3.84	22.03
320	295	1902	4.04	26.07
325	357	2259	4.89	30.97
331	403	2662	5.52	36.49
336	519	3181	7.11	43.61
342	517	3698	7.09	50.69
348	540	4238	7.40	58.09
355	544	4782	7.46	65.55
363	511	5293	7.00	72.56
371	484	5777	6.63	79.19
380	421	6198	5.77	84.96
392	382	6580	5.24	90.20
406	264	6844	3.62	93.82
426	222	7066	3.04	96.86
461	139	7205	1.91	98.77
522	90	7295	1.23	100.00

Figure C.6.9. Spring 2023 STAAR Alternate 2 Grade 5 RLA
Frequency Distribution of Scale Scores
All Students



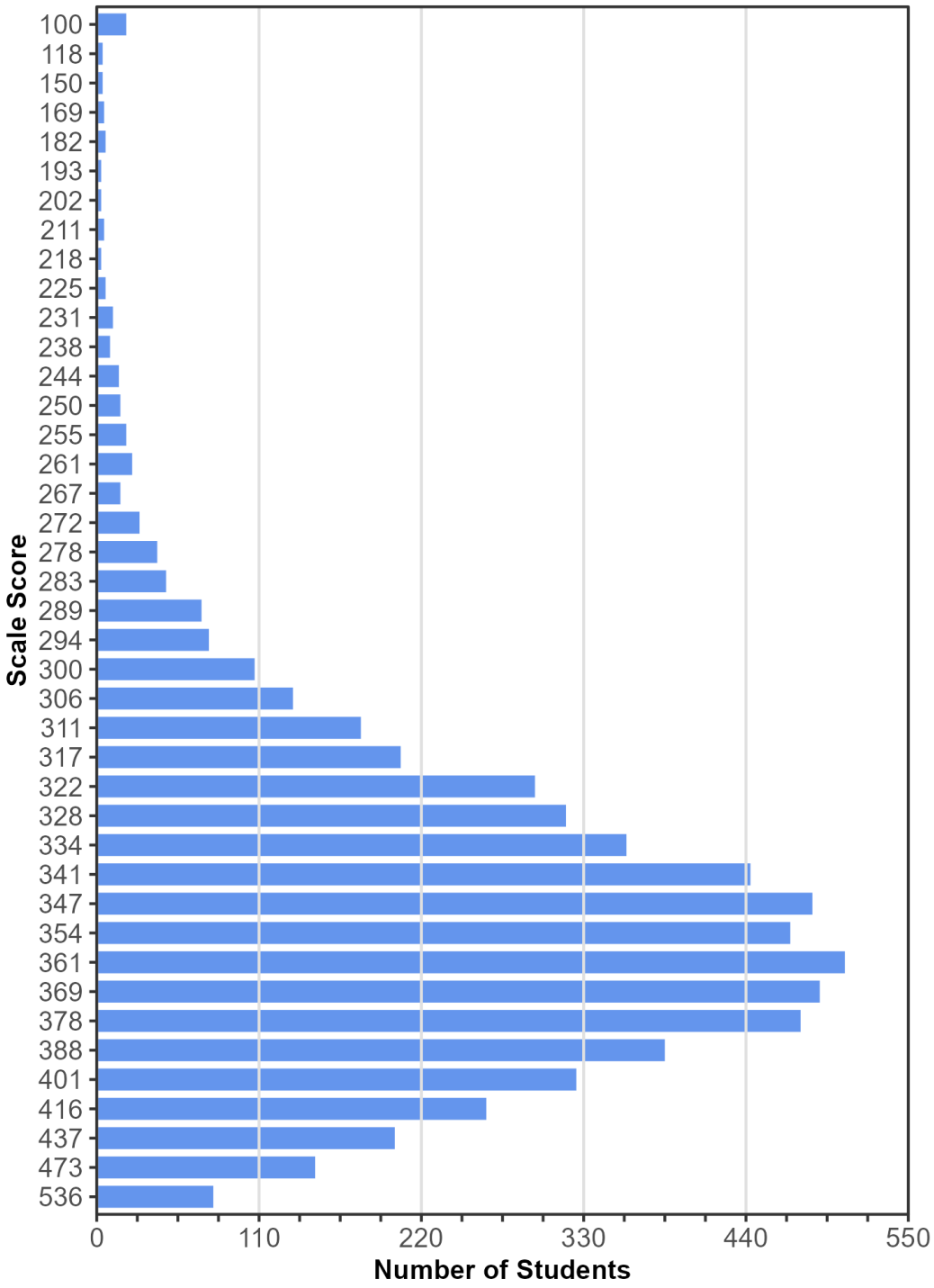
	Freq	Cum Freq	Pct	Cum Pct
100	30	30	0.44	0.44
122	2	32	0.03	0.47
154	16	48	0.23	0.70
172	2	50	0.03	0.73
185	10	60	0.15	0.88
196	7	67	0.10	0.98
205	11	78	0.16	1.14
213	11	89	0.16	1.30
220	18	107	0.26	1.57
227	8	115	0.12	1.69
233	18	133	0.26	1.95
239	17	150	0.25	2.20
245	21	171	0.31	2.51
250	18	189	0.26	2.77
255	24	213	0.35	3.12
261	33	246	0.48	3.61
266	36	282	0.53	4.13
271	38	320	0.56	4.69
276	68	388	1.00	5.69
281	58	446	0.85	6.54
285	72	518	1.06	7.59
290	112	630	1.64	9.23
295	136	766	1.99	11.23
300	153	919	2.24	13.47
305	182	1101	2.67	16.14
310	239	1340	3.50	19.64
315	317	1657	4.65	24.29
320	324	1981	4.75	29.04
325	379	2360	5.56	34.59
331	379	2739	5.56	40.15
337	400	3139	5.86	46.01
343	456	3595	6.68	52.70
349	456	4051	6.68	59.38
357	467	4518	6.85	66.23
365	472	4990	6.92	73.15
374	477	5467	6.99	80.14
385	435	5902	6.38	86.51
399	406	6308	5.95	92.47
418	278	6586	4.08	96.54
452	143	6729	2.10	98.64
512	93	6822	1.36	100.00

Figure C.6.10. Spring 2023 STAAR Alternate 2 Grade 6 RLA
Frequency Distribution of Scale Scores
All Students



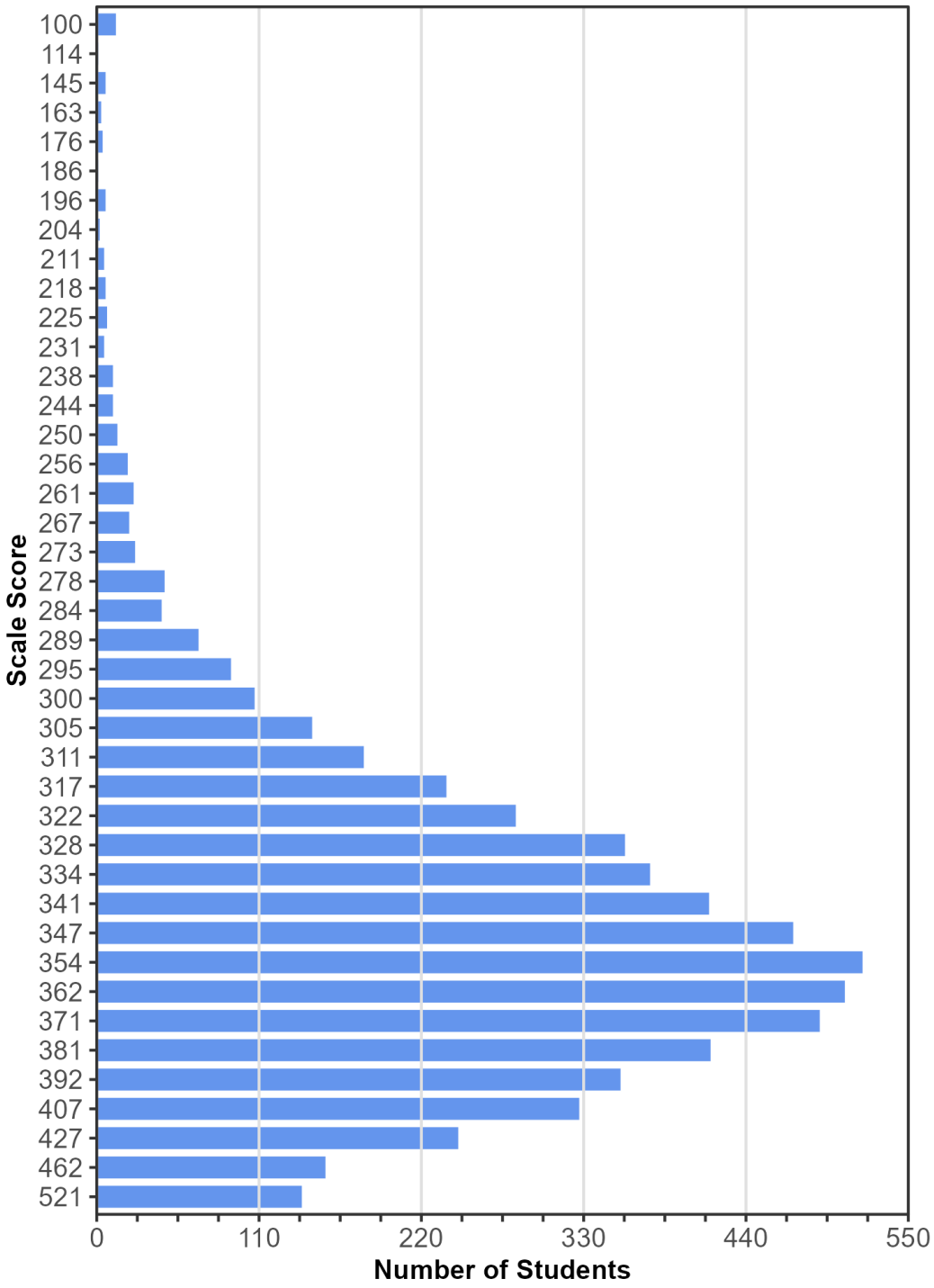
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	15	15	0.23	0.23
116	2	17	0.03	0.26
148	4	21	0.06	0.32
166	3	24	0.05	0.37
180	8	32	0.12	0.49
191	4	36	0.06	0.56
200	3	39	0.05	0.60
208	5	44	0.08	0.68
215	5	49	0.08	0.76
222	8	57	0.12	0.88
229	8	65	0.12	1.00
235	10	75	0.15	1.16
241	20	95	0.31	1.47
246	10	105	0.15	1.62
252	15	120	0.23	1.85
258	27	147	0.42	2.27
263	30	177	0.46	2.73
268	35	212	0.54	3.27
274	53	265	0.82	4.09
279	60	325	0.93	5.02
284	76	401	1.17	6.19
289	82	483	1.27	7.45
295	126	609	1.94	9.40
300	141	750	2.18	11.58
305	179	929	2.76	14.34
311	249	1178	3.84	18.18
316	280	1458	4.32	22.50
322	318	1776	4.91	27.41
328	416	2192	6.42	33.83
334	393	2585	6.07	39.90
340	438	3023	6.76	46.66
346	484	3507	7.47	54.13
354	444	3951	6.85	60.98
361	462	4413	7.13	68.11
370	466	4879	7.19	75.30
380	435	5314	6.71	82.02
391	363	5677	5.60	87.62
406	307	5984	4.74	92.36
426	224	6208	3.46	95.82
461	152	6360	2.35	98.16
523	119	6479	1.84	100.00

Figure C.6.11. Spring 2023 STAAR Alternate 2 Grade 7 RLA
Frequency Distribution of Scale Scores
All Students



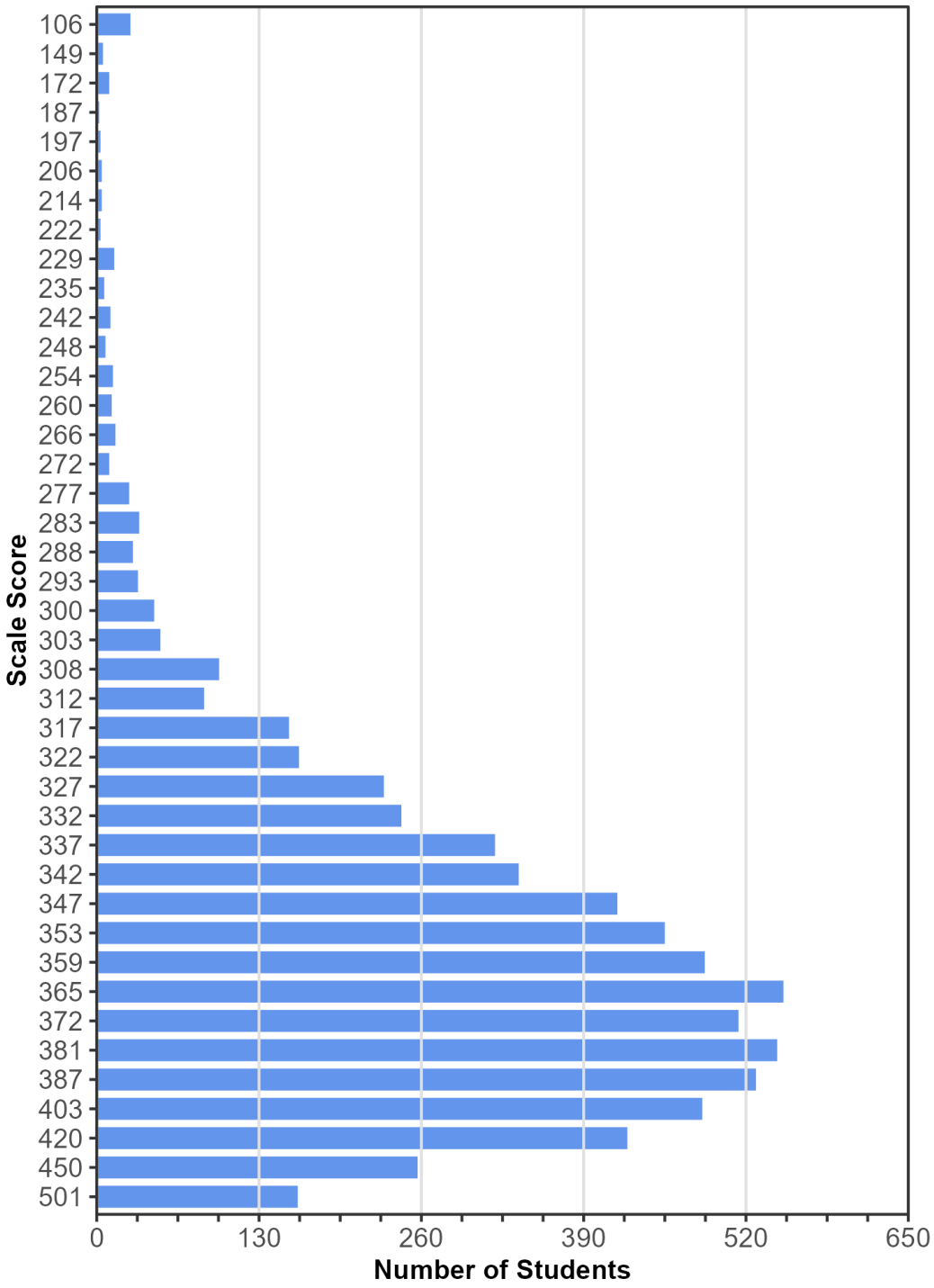
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	20	20	0.32	0.32
118	4	24	0.06	0.38
150	4	28	0.06	0.44
169	5	33	0.08	0.52
182	6	39	0.10	0.62
193	3	42	0.05	0.67
202	3	45	0.05	0.71
211	5	50	0.08	0.79
218	3	53	0.05	0.84
225	6	59	0.10	0.94
231	11	70	0.17	1.11
238	9	79	0.14	1.25
244	15	94	0.24	1.49
250	16	110	0.25	1.74
255	20	130	0.32	2.06
261	24	154	0.38	2.44
267	16	170	0.25	2.69
272	29	199	0.46	3.15
278	41	240	0.65	3.80
283	47	287	0.75	4.55
289	71	358	1.13	5.68
294	76	434	1.20	6.88
300	107	541	1.70	8.58
306	133	674	2.11	10.68
311	179	853	2.84	13.52
317	206	1059	3.27	16.79
322	297	1356	4.71	21.50
328	318	1674	5.04	26.54
334	359	2033	5.69	32.23
341	443	2476	7.02	39.25
347	485	2961	7.69	46.94
354	470	3431	7.45	54.39
361	507	3938	8.04	62.43
369	490	4428	7.77	70.20
378	477	4905	7.56	77.76
388	385	5290	6.10	83.86
401	325	5615	5.15	89.01
416	264	5879	4.19	93.20
437	202	6081	3.20	96.40
473	148	6229	2.35	98.75
536	79	6308	1.25	100.00

Figure C.6.12. Spring 2023 STAAR Alternate 2 Grade 8 RLA
Frequency Distribution of Scale Scores
All Students



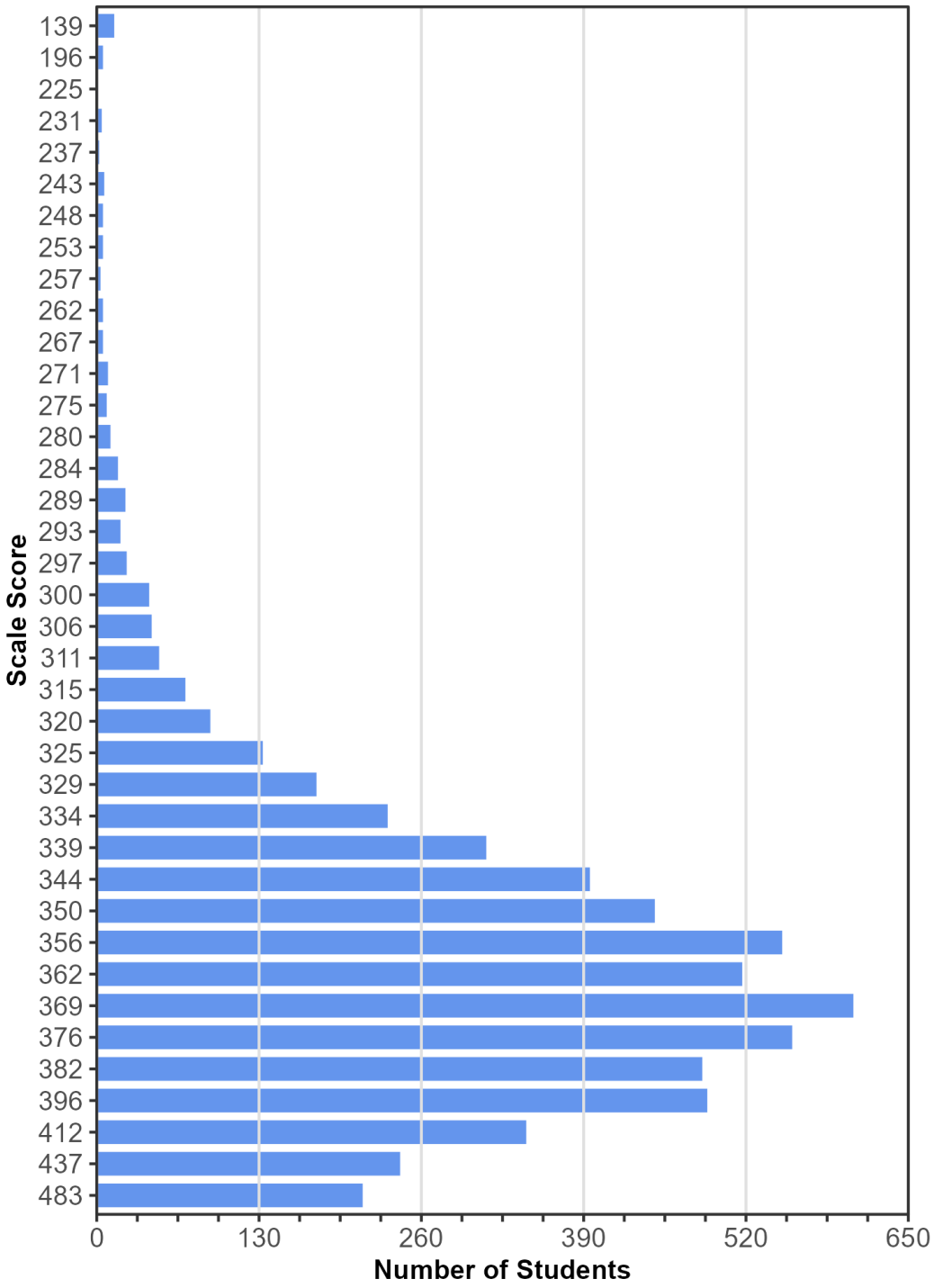
	Freq	Cum Freq	Pct	Cum Pct
100	13	13	0.21	0.21
114	1	14	0.02	0.23
145	6	20	0.10	0.32
163	3	23	0.05	0.37
176	4	27	0.06	0.44
186	1	28	0.02	0.45
196	6	34	0.10	0.55
204	2	36	0.03	0.58
211	5	41	0.08	0.66
218	6	47	0.10	0.76
225	7	54	0.11	0.88
231	5	59	0.08	0.96
238	11	70	0.18	1.14
244	11	81	0.18	1.31
250	14	95	0.23	1.54
256	21	116	0.34	1.88
261	25	141	0.41	2.29
267	22	163	0.36	2.64
273	26	189	0.42	3.06
278	46	235	0.75	3.81
284	44	279	0.71	4.52
289	69	348	1.12	5.64
295	91	439	1.48	7.12
300	107	546	1.74	8.85
305	146	692	2.37	11.22
311	181	873	2.93	14.16
317	237	1110	3.84	18.00
322	284	1394	4.61	22.60
328	358	1752	5.81	28.41
334	375	2127	6.08	34.49
341	415	2542	6.73	41.22
347	472	3014	7.65	48.87
354	519	3533	8.42	57.29
362	507	4040	8.22	65.51
371	490	4530	7.95	73.46
381	416	4946	6.75	80.20
392	355	5301	5.76	85.96
407	327	5628	5.30	91.26
427	245	5873	3.97	95.23
462	155	6028	2.51	97.75
521	139	6167	2.25	100.00

Figure C.6.13. Spring 2023 STAAR Alternate 2 Grade 5 Science
Frequency Distribution of Scale Scores
All Students



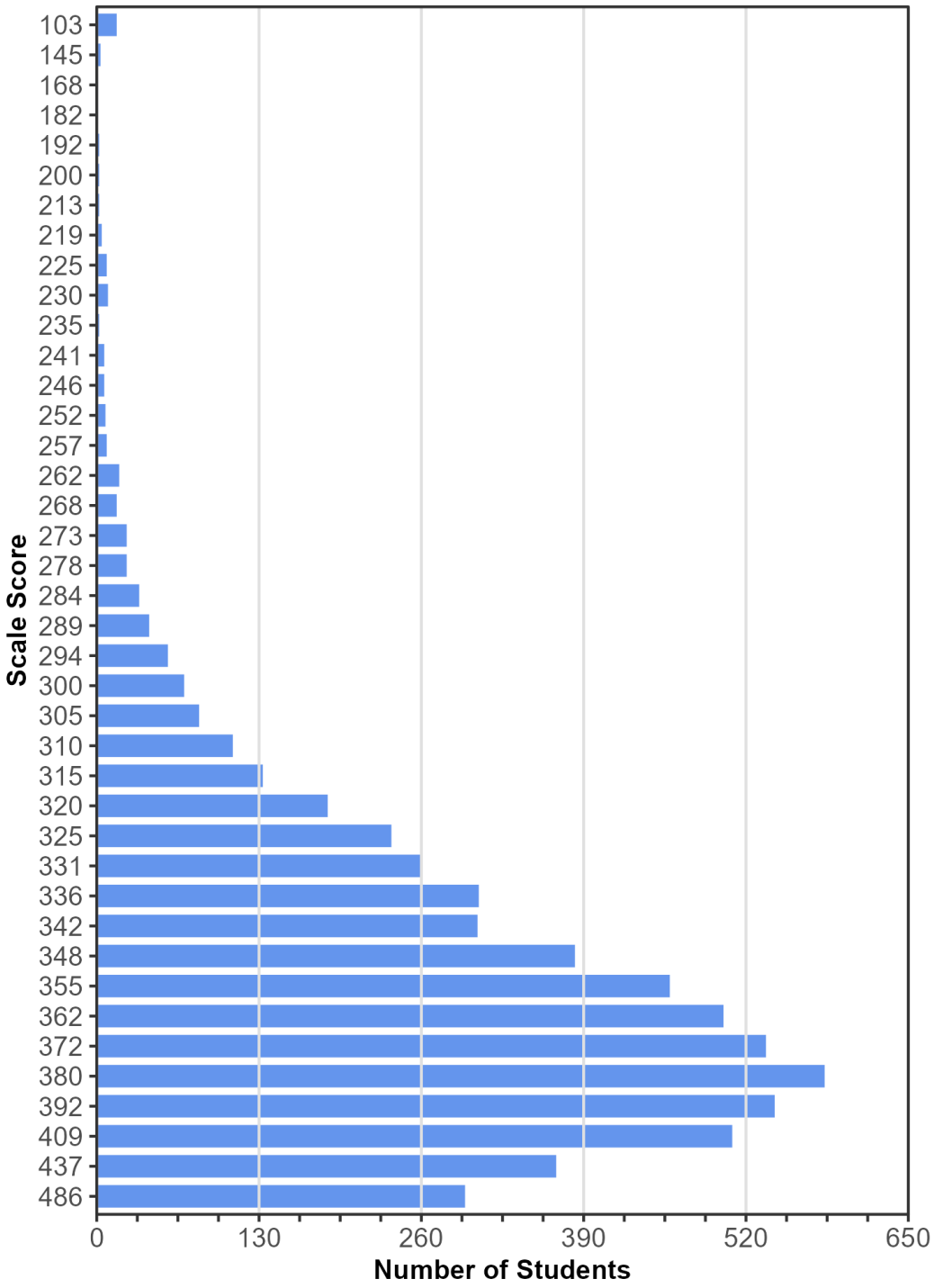
	Freq	Cum Freq	Pct	Cum Pct
106	27	27	0.40	0.40
149	5	32	0.07	0.47
172	10	42	0.15	0.62
187	2	44	0.03	0.65
197	3	47	0.04	0.69
206	4	51	0.06	0.75
214	4	55	0.06	0.81
222	3	58	0.04	0.85
229	14	72	0.21	1.06
235	6	78	0.09	1.14
242	11	89	0.16	1.30
248	7	96	0.10	1.41
254	13	109	0.19	1.60
260	12	121	0.18	1.77
266	15	136	0.22	1.99
272	10	146	0.15	2.14
277	26	172	0.38	2.52
283	34	206	0.50	3.02
288	29	235	0.43	3.45
293	33	268	0.48	3.93
300	46	314	0.67	4.60
303	51	365	0.75	5.35
308	98	463	1.44	6.79
312	86	549	1.26	8.05
317	154	703	2.26	10.31
322	162	865	2.38	12.68
327	230	1095	3.37	16.06
332	244	1339	3.58	19.63
337	319	1658	4.68	24.31
342	338	1996	4.96	29.27
347	417	2413	6.11	35.38
353	455	2868	6.67	42.05
359	487	3355	7.14	49.19
365	550	3905	8.06	57.26
372	514	4419	7.54	64.79
381	545	4964	7.99	72.79
387	528	5492	7.74	80.53
403	485	5977	7.11	87.64
420	425	6402	6.23	93.87
450	257	6659	3.77	97.64
501	161	6820	2.36	100.00

Figure C.6.14. Spring 2023 STAAR Alternate 2 Grade 8 Science
Frequency Distribution of Scale Scores
All Students



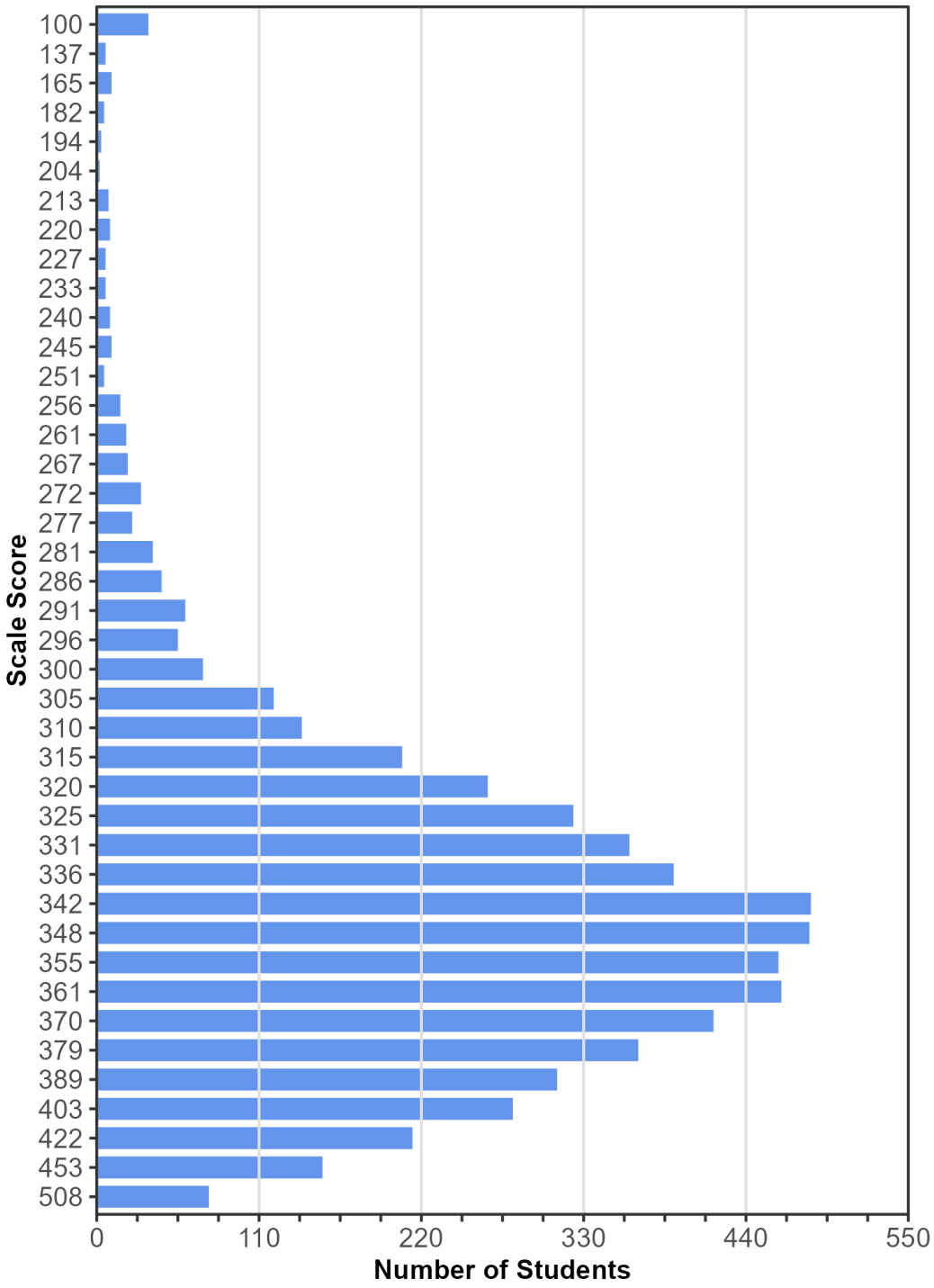
	Cum		Cum
Freq	Freq	Pct	Pct
14	14	0.23	0.23
5	19	0.08	0.31
1	20	0.02	0.32
4	24	0.06	0.39
2	26	0.03	0.42
6	32	0.10	0.52
5	37	0.08	0.60
5	42	0.08	0.68
3	45	0.05	0.73
5	50	0.08	0.81
5	55	0.08	0.89
9	64	0.15	1.04
8	72	0.13	1.17
11	83	0.18	1.35
17	100	0.28	1.62
23	123	0.37	2.00
19	142	0.31	2.30
24	166	0.39	2.69
42	208	0.68	3.37
44	252	0.71	4.09
50	302	0.81	4.90
71	373	1.15	6.05
91	464	1.48	7.53
133	597	2.16	9.69
176	773	2.86	12.54
233	1006	3.78	16.32
312	1318	5.06	21.39
395	1713	6.41	27.79
447	2160	7.25	35.05
549	2709	8.91	43.96
517	3226	8.39	52.34
606	3832	9.83	62.18
557	4389	9.04	71.22
485	4874	7.87	79.08
489	5363	7.93	87.02
344	5707	5.58	92.60
243	5950	3.94	96.54
213	6163	3.46	100.00

Figure C.6.15. Spring 2023 STAAR Alternate 2 Grade 8 Social Studies
Frequency Distribution of Scale Scores
All Students



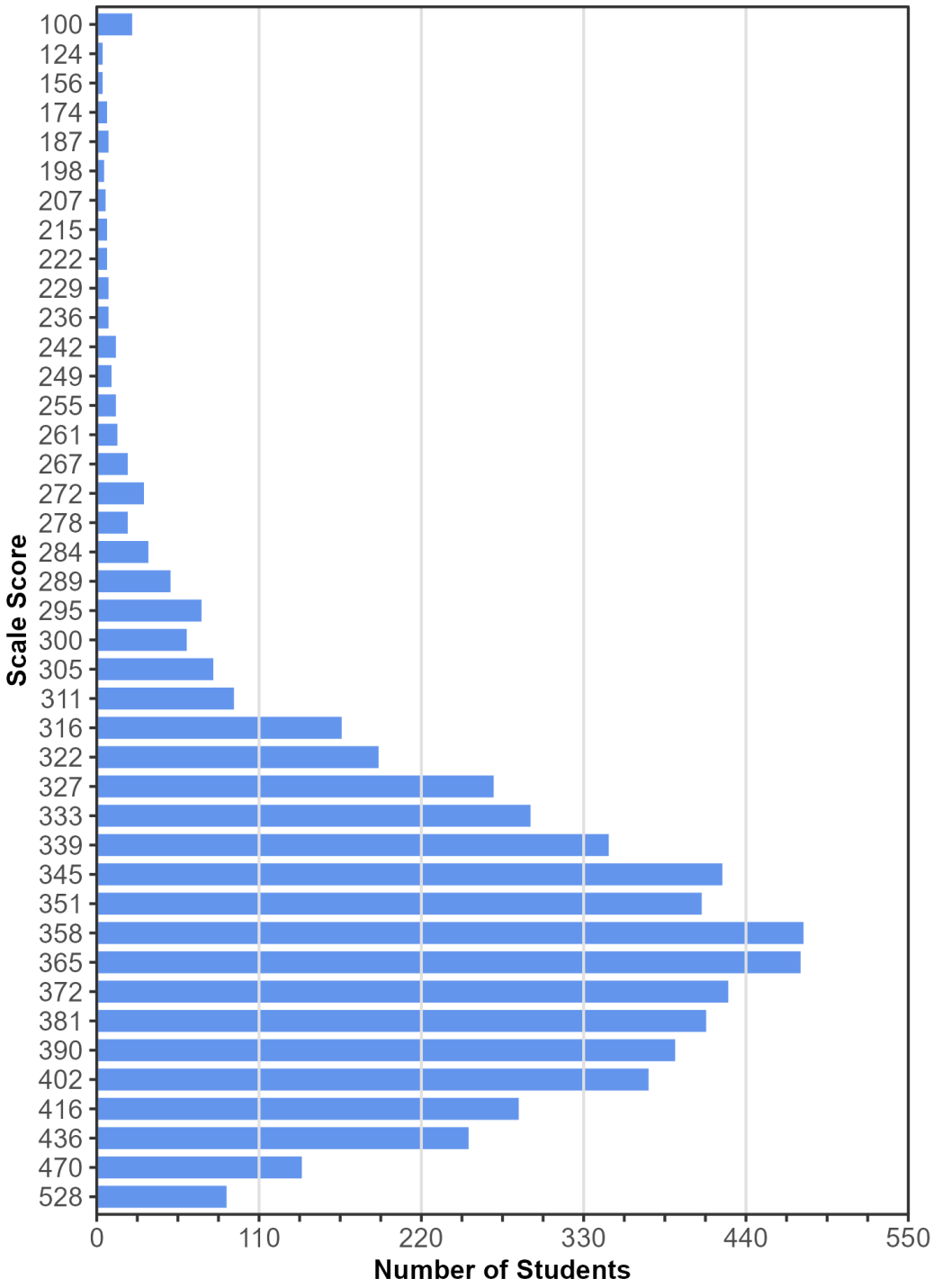
Scale Score	Freq	Cum Freq	Pct	Cum Pct
103	16	16	0.26	0.26
145	3	19	0.05	0.31
168	1	20	0.02	0.32
182	1	21	0.02	0.34
192	2	23	0.03	0.37
200	2	25	0.03	0.41
213	2	27	0.03	0.44
219	4	31	0.06	0.50
225	8	39	0.13	0.63
230	9	48	0.15	0.78
235	2	50	0.03	0.81
241	6	56	0.10	0.91
246	6	62	0.10	1.01
252	7	69	0.11	1.12
257	8	77	0.13	1.25
262	18	95	0.29	1.54
268	16	111	0.26	1.80
273	24	135	0.39	2.19
278	24	159	0.39	2.58
284	34	193	0.55	3.13
289	42	235	0.68	3.82
294	57	292	0.93	4.74
300	70	362	1.14	5.88
305	82	444	1.33	7.21
310	109	553	1.77	8.98
315	133	686	2.16	11.14
320	185	871	3.00	14.15
325	236	1107	3.83	17.98
331	261	1368	4.24	22.22
336	306	1674	4.97	27.19
342	305	1979	4.95	32.14
348	383	2362	6.22	38.36
355	459	2821	7.45	45.82
362	502	3323	8.15	53.97
372	536	3859	8.71	62.68
380	583	4442	9.47	72.15
392	543	4985	8.82	80.96
409	509	5494	8.27	89.23
437	368	5862	5.98	95.21
486	295	6157	4.79	100.00

Figure C.6.16. Spring 2023 STAAR Alternate 2 Algebra I
Frequency Distribution of Scale Scores
All Students



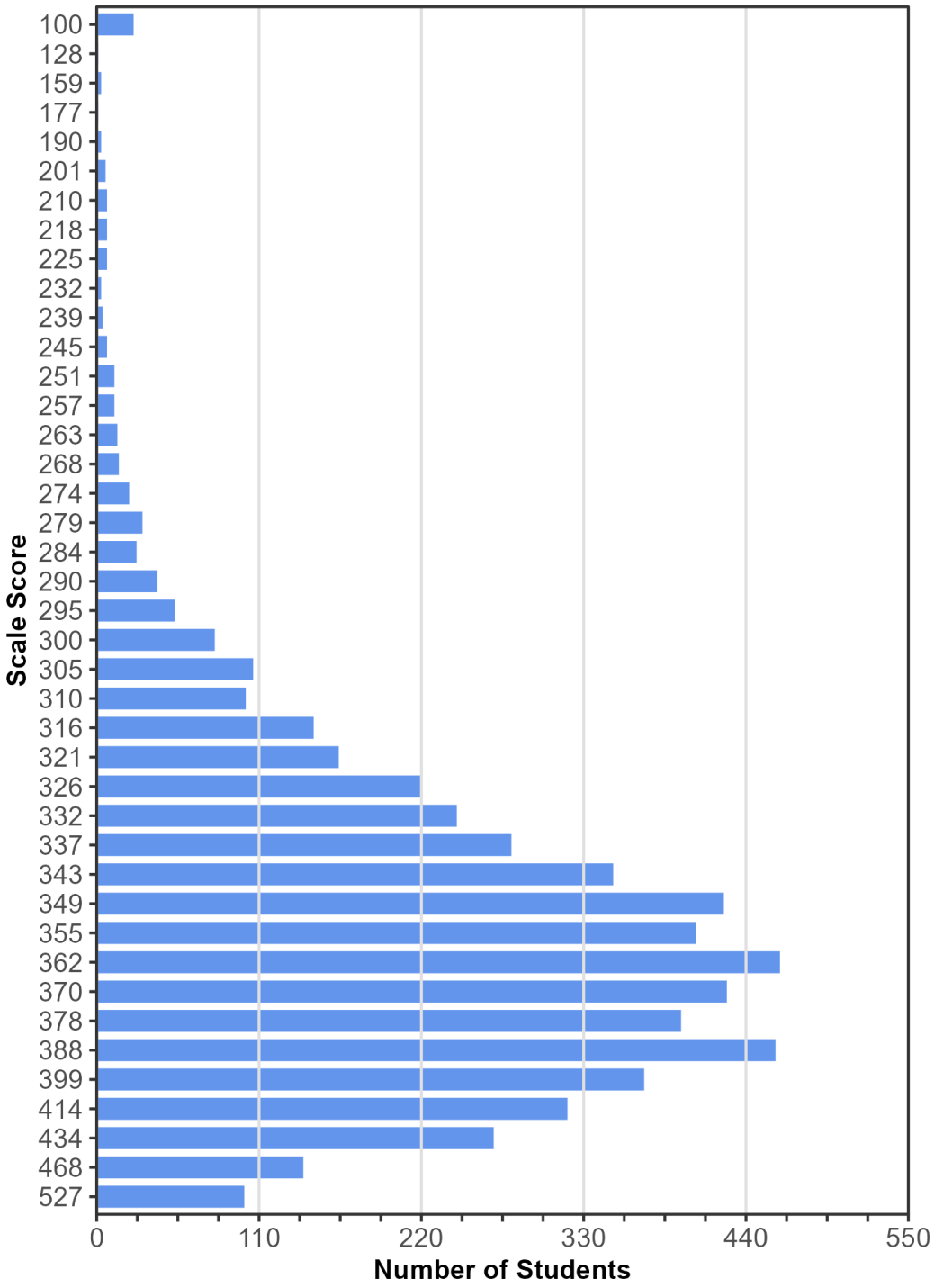
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	35	35	0.58	0.58
137	6	41	0.10	0.68
165	10	51	0.17	0.85
182	5	56	0.08	0.93
194	3	59	0.05	0.98
204	2	61	0.03	1.01
213	8	69	0.13	1.15
220	9	78	0.15	1.30
227	6	84	0.10	1.40
233	6	90	0.10	1.50
240	9	99	0.15	1.65
245	10	109	0.17	1.81
251	5	114	0.08	1.90
256	16	130	0.27	2.16
261	20	150	0.33	2.49
267	21	171	0.35	2.84
272	30	201	0.50	3.34
277	24	225	0.40	3.74
281	38	263	0.63	4.37
286	44	307	0.73	5.10
291	60	367	1.00	6.10
296	55	422	0.91	7.02
300	72	494	1.20	8.21
305	120	614	2.00	10.21
310	139	753	2.31	12.52
315	207	960	3.44	15.96
320	265	1225	4.41	20.37
325	323	1548	5.37	25.74
331	361	1909	6.00	31.74
336	391	2300	6.50	38.24
342	484	2784	8.05	46.28
348	483	3267	8.03	54.31
355	462	3729	7.68	62.00
361	464	4193	7.71	69.71
370	418	4611	6.95	76.66
379	367	4978	6.10	82.76
389	312	5290	5.19	87.95
403	282	5572	4.69	92.64
422	214	5786	3.56	96.19
453	153	5939	2.54	98.74
508	76	6015	1.26	100.00

Figure C.6.17. Spring 2023 STAAR Alternate 2 English I
Frequency Distribution of Scale Scores
All Students



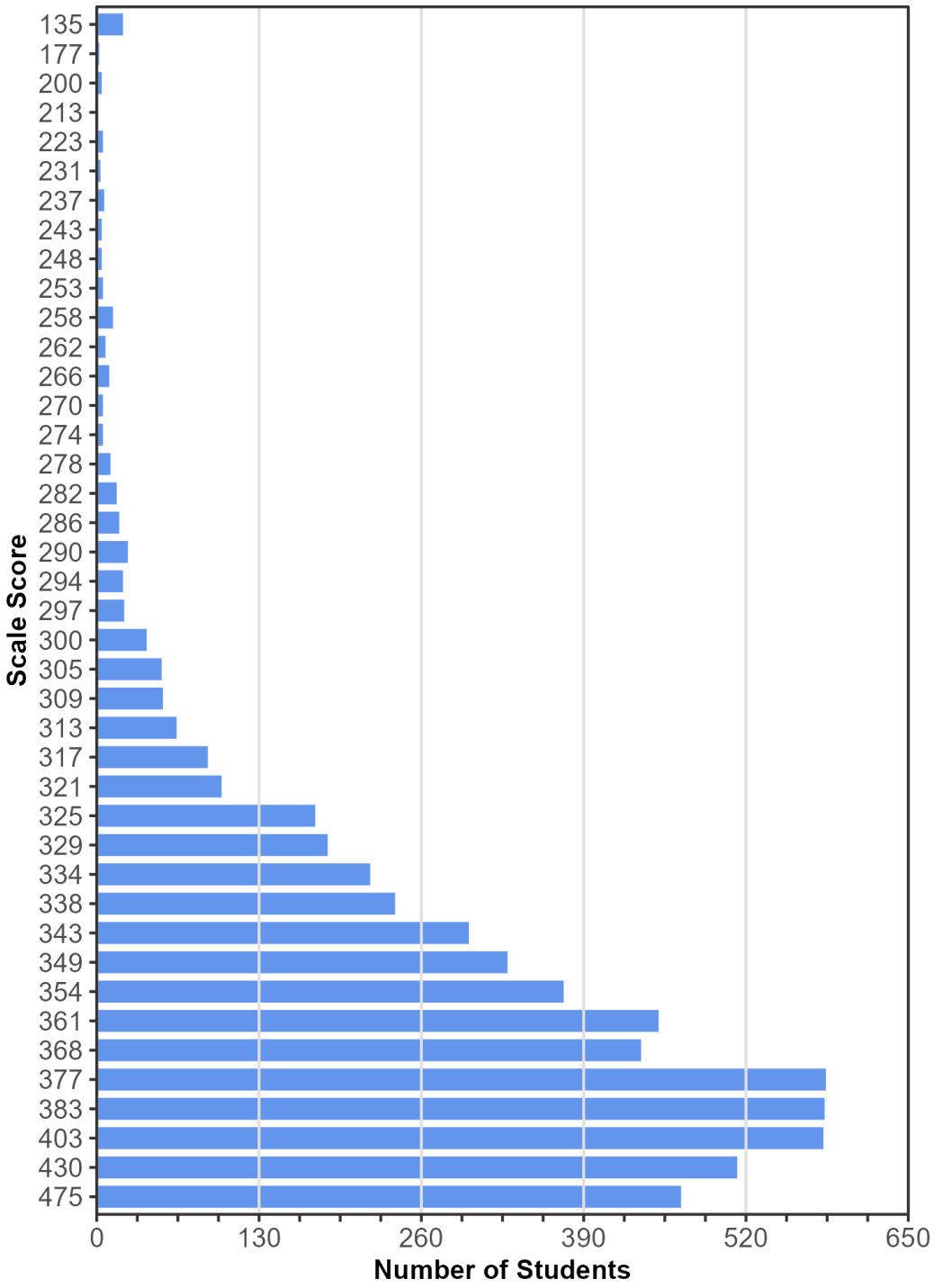
Freq	Cum Freq	Pct	Cum Pct
24	24	0.40	0.40
4	28	0.07	0.46
4	32	0.07	0.53
7	39	0.12	0.65
8	47	0.13	0.78
5	52	0.08	0.86
6	58	0.10	0.96
7	65	0.12	1.08
7	72	0.12	1.19
8	80	0.13	1.33
8	88	0.13	1.46
13	101	0.22	1.67
10	111	0.17	1.84
13	124	0.22	2.06
14	138	0.23	2.29
21	159	0.35	2.64
32	191	0.53	3.17
21	212	0.35	3.52
35	247	0.58	4.10
50	297	0.83	4.93
71	368	1.18	6.10
61	429	1.01	7.11
79	508	1.31	8.42
93	601	1.54	9.97
166	767	2.75	12.72
191	958	3.17	15.89
269	1227	4.46	20.35
294	1521	4.88	25.22
347	1868	5.75	30.98
424	2292	7.03	38.01
410	2702	6.80	44.81
479	3181	7.94	52.75
477	3658	7.91	60.66
428	4086	7.10	67.76
413	4499	6.85	74.61
392	4891	6.50	81.11
374	5265	6.20	87.31
286	5551	4.74	92.06
252	5803	4.18	96.24
139	5942	2.31	98.54
88	6030	1.46	100.00

Figure C.6.18. Spring 2023 STAAR Alternate 2 English II
Frequency Distribution of Scale Scores
All Students



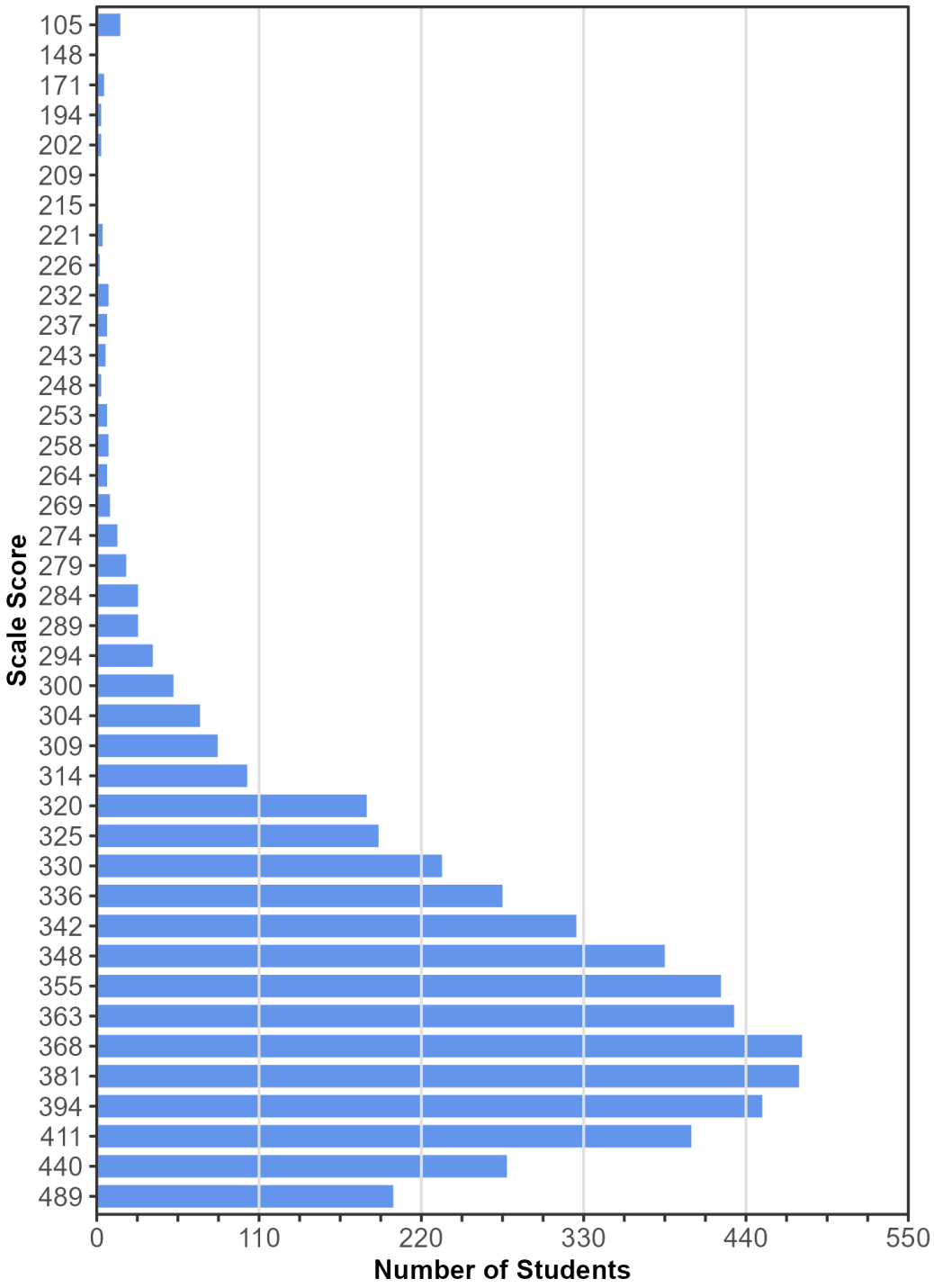
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	25	25	0.43	0.43
128	1	26	0.02	0.45
159	3	29	0.05	0.50
177	1	30	0.02	0.52
190	3	33	0.05	0.57
201	6	39	0.10	0.68
210	7	46	0.12	0.80
218	7	53	0.12	0.92
225	7	60	0.12	1.04
232	3	63	0.05	1.09
239	4	67	0.07	1.16
245	7	74	0.12	1.28
251	12	86	0.21	1.49
257	12	98	0.21	1.70
263	14	112	0.24	1.94
268	15	127	0.26	2.20
274	22	149	0.38	2.58
279	31	180	0.54	3.12
284	27	207	0.47	3.59
290	41	248	0.71	4.30
295	53	301	0.92	5.22
300	80	381	1.39	6.60
305	106	487	1.84	8.44
310	101	588	1.75	10.19
316	147	735	2.55	12.74
321	164	899	2.84	15.58
326	220	1119	3.81	19.39
332	244	1363	4.23	23.62
337	281	1644	4.87	28.49
343	350	1994	6.07	34.56
349	425	2419	7.37	41.92
355	406	2825	7.04	48.96
362	463	3288	8.02	56.98
370	427	3715	7.40	64.38
378	396	4111	6.86	71.25
388	460	4571	7.97	79.22
399	371	4942	6.43	85.65
414	319	5261	5.53	91.18
434	269	5530	4.66	95.84
468	140	5670	2.43	98.27
527	100	5770	1.73	100.00

Figure C.6.19. Spring 2023 STAAR Alternate 2 Biology
Frequency Distribution of Scale Scores
All Students



	Freq	Cum Freq	Pct	Cum Pct
135	21	21	0.35	0.35
177	2	23	0.03	0.38
200	4	27	0.07	0.45
213	1	28	0.02	0.46
223	5	33	0.08	0.55
231	3	36	0.05	0.60
237	6	42	0.10	0.70
243	4	46	0.07	0.76
248	4	50	0.07	0.83
253	5	55	0.08	0.91
258	13	68	0.22	1.13
262	7	75	0.12	1.24
266	10	85	0.17	1.41
270	5	90	0.08	1.49
274	5	95	0.08	1.57
278	11	106	0.18	1.75
282	16	122	0.26	2.02
286	18	140	0.30	2.32
290	25	165	0.41	2.73
294	21	186	0.35	3.08
297	22	208	0.36	3.44
300	40	248	0.66	4.11
305	52	300	0.86	4.97
309	53	353	0.88	5.84
313	64	417	1.06	6.90
317	89	506	1.47	8.38
321	100	606	1.66	10.03
325	175	781	2.90	12.93
329	185	966	3.06	15.99
334	219	1185	3.63	19.62
338	239	1424	3.96	23.57
343	298	1722	4.93	28.51
349	329	2051	5.45	33.95
354	374	2425	6.19	40.14
361	450	2875	7.45	47.59
368	436	3311	7.22	54.81
377	584	3895	9.67	64.48
383	583	4478	9.65	74.13
403	582	5060	9.63	83.76
430	513	5573	8.49	92.25
475	468	6041	7.75	100.00

Figure C.6.20. Spring 2023 STAAR Alternate 2 U.S. History
Frequency Distribution of Scale Scores
All Students



	Freq	Cum Freq	Pct	Cum Pct
105	16	16	0.30	0.30
148	1	17	0.02	0.32
171	5	22	0.10	0.42
194	3	25	0.06	0.48
202	3	28	0.06	0.53
209	1	29	0.02	0.55
215	1	30	0.02	0.57
221	4	34	0.08	0.65
226	2	36	0.04	0.68
232	8	44	0.15	0.84
237	7	51	0.13	0.97
243	6	57	0.11	1.08
248	3	60	0.06	1.14
253	7	67	0.13	1.27
258	8	75	0.15	1.43
264	7	82	0.13	1.56
269	9	91	0.17	1.73
274	14	105	0.27	2.00
279	20	125	0.38	2.38
284	28	153	0.53	2.91
289	28	181	0.53	3.44
294	38	219	0.72	4.16
300	52	271	0.99	5.15
304	70	341	1.33	6.48
309	82	423	1.56	8.04
314	102	525	1.94	9.98
320	183	708	3.48	13.46
325	191	899	3.63	17.09
330	234	1133	4.45	21.54
336	275	1408	5.23	26.77
342	325	1733	6.18	32.95
348	385	2118	7.32	40.27
355	423	2541	8.04	48.31
363	432	2973	8.21	56.52
368	478	3451	9.09	65.61
381	476	3927	9.05	74.66
394	451	4378	8.57	83.23
411	403	4781	7.66	90.89
440	278	5059	5.29	96.18
489	201	5260	3.82	100.00

TECHNICAL DIGEST 2022–2023

Appendix D

TELPAS Statistical Tables and Figures

TELPAS Statistical Tables and Figures

2023 TELPAS Classification Consistency and Accuracy

- Table D.1.1. TELPAS

2023 TELPAS Scale Score Correlations

- Table D.2.1. TELPAS

2023 TELPAS Conditional Standard Error of Measurement for Scale Scores

- Table D.3.1. TELPAS Grades 2–5 Reading
- Table D.3.2. TELPAS Grades 6–12 Reading
- Table D.3.3. TELPAS Grades 2–5 Writing
- Table D.3.4. TELPAS Grades 6–12 Writing
- Table D.3.5. TELPAS Grades 2–12 Listening
- Table D.3.6. TELPAS Grades 2–12 Speaking

2023 TELPAS Mean P-Values

- Table D.4.1. TELPAS Reading and Writing Grade 2 Total Group
- Table D.4.2. TELPAS Reading and Writing Grade 2 Female
- Table D.4.3. TELPAS Reading and Writing Grade 2 Male
- Table D.4.4. TELPAS Reading and Writing Grade 3 Total Group
- Table D.4.5. TELPAS Reading and Writing Grade 3 Female
- Table D.4.6. TELPAS Reading and Writing Grade 3 Male
- Table D.4.7. TELPAS Reading and Writing Grades 4–5 Total Group
- Table D.4.8. TELPAS Reading and Writing Grades 4–5 Female
- Table D.4.9. TELPAS Reading and Writing Grades 4–5 Male
- Table D.4.10. TELPAS Reading and Writing Grades 6–7 Total Group
- Table D.4.11. TELPAS Reading and Writing Grades 6–7 Female
- Table D.4.12. TELPAS Reading and Writing Grades 6–7 Male
- Table D.4.13. TELPAS Reading and Writing Grades 8–9 Total Group
- Table D.4.14. TELPAS Reading and Writing Grades 8–9 Female
- Table D.4.15. TELPAS Reading and Writing Grades 8–9 Male
- Table D.4.16. TELPAS Reading and Writing Grades 10–12 Total Group
- Table D.4.17. TELPAS Reading and Writing Grades 10–12 Female
- Table D.4.18. TELPAS Reading and Writing Grades 10–12 Male
- Table D.4.19. TELPAS Listening and Speaking Grades 2–3 Total Group
- Table D.4.20. TELPAS Listening and Speaking Grades 2–3 Female

- Table D.4.21. TELPAS Listening and Speaking Grades 2–3 Male
- Table D.4.22. TELPAS Listening and Speaking Grades 4–5 Total Group
- Table D.4.23. TELPAS Listening and Speaking Grades 4–5 Female
- Table D.4.24. TELPAS Listening and Speaking Grades 4–5 Male
- Table D.4.25. TELPAS Listening and Speaking Grades 6–8 Total Group
- Table D.4.26. TELPAS Listening and Speaking Grades 6–8 Female
- Table D.4.27. TELPAS Listening and Speaking Grades 6–8 Male
- Table D.4.28. TELPAS Listening and Speaking Grades 9–12 Total Group
- Table D.4.29. TELPAS Listening and Speaking Grades 9–12 Female
- Table D.4.30. TELPAS Listening and Speaking Grades 9–12 Male

2023 TELPAS Scale Score Descriptive Statistics

- Table D.5.1. TELPAS Reading
- Table D.5.2. TELPAS Writing
- Table D.5.3. TELPAS Listening
- Table D.5.4. TELPAS Speaking

2023 TELPAS Frequency Distribution of Scale Scores

- Figure D.5.1. TELPAS Grade 2 Reading
- Figure D.5.2. TELPAS Grade 3 Reading
- Figure D.5.3. TELPAS Grade 4 Reading
- Figure D.5.4. TELPAS Grade 5 Reading
- Figure D.5.5. TELPAS Grade 6 Reading
- Figure D.5.6. TELPAS Grade 7 Reading
- Figure D.5.7. TELPAS Grade 8 Reading
- Figure D.5.8. TELPAS Grade 9 Reading
- Figure D.5.9. TELPAS Grade 10 Reading
- Figure D.5.10. TELPAS Grade 11 Reading
- Figure D.5.11. TELPAS Grade 12 Reading
- Figure D.5.12. TELPAS Grade 2 Writing
- Figure D.5.13. TELPAS Grade 3 Writing
- Figure D.5.14. TELPAS Grade 4 Writing
- Figure D.5.15. TELPAS Grade 5 Writing
- Figure D.5.16. TELPAS Grade 6 Writing
- Figure D.5.17. TELPAS Grade 7 Writing
- Figure D.5.18. TELPAS Grade 8 Writing
- Figure D.5.19. TELPAS Grade 9 Writing
- Figure D.5.20. TELPAS Grade 10 Writing

- Figure D.5.21. TELPAS Grade 11 Writing
- Figure D.5.22. TELPAS Grade 12 Writing
- Figure D.5.23. TELPAS Grade 2 Listening
- Figure D.5.24. TELPAS Grade 3 Listening
- Figure D.5.25. TELPAS Grade 4 Listening
- Figure D.5.26. TELPAS Grade 5 Listening
- Figure D.5.27. TELPAS Grade 6 Listening
- Figure D.5.28. TELPAS Grade 7 Listening
- Figure D.5.29. TELPAS Grade 8 Listening
- Figure D.5.30. TELPAS Grade 9 Listening
- Figure D.5.31. TELPAS Grade 10 Listening
- Figure D.5.32. TELPAS Grade 11 Listening
- Figure D.5.33. TELPAS Grade 12 Listening
- Figure D.5.34. TELPAS Grade 2 Speaking
- Figure D.5.35. TELPAS Grade 3 Speaking
- Figure D.5.36. TELPAS Grade 4 Speaking
- Figure D.5.37. TELPAS Grade 5 Speaking
- Figure D.5.38. TELPAS Grade 6 Speaking
- Figure D.5.39. TELPAS Grade 7 Speaking
- Figure D.5.40. TELPAS Grade 8 Speaking
- Figure D.5.41. TELPAS Grade 9 Speaking
- Figure D.5.42. TELPAS Grade 10 Speaking
- Figure D.5.43. TELPAS Grade 11 Speaking
- Figure D.5.44. TELPAS Grade 12 Speaking

2023 TELPAS Distribution of Proficiency Levels

- Table D.6.1. Since 2018 Redesign—Grades K–6
- Table D.6.2. Since 2018 Redesign—Grades 7–12
- Table D.6.3. Since 2005—Grades K–1
- Table D.6.4. Since 2005—Grades 2–3
- Table D.6.5. Since 2005—Grades 4–5
- Table D.6.6. Since 2005—Grades 6–7
- Table D.6.7. Since 2005—Grades 8–9
- Table D.6.8. Since 2005—Grades 10–11
- Table D.6.9. Since 2005—Grade 12

2023 TELPAS Classification Consistency and Accuracy

Table D.1.1. Spring 2023 TELPAS Assessments
Classification Consistency and Accuracy

Grade	Domain	Decision Consistency[1]	Decision Accuracy[2]
2	Reading	63.4	77.2
3	Reading	63.2	76.7
4–5	Reading	66.3	78.1
6–7	Reading	60.9	74.6
8–9	Reading	65.4	77.8
10–12	Reading	66.7	79.2
2	Writing	65.2	78.7
3	Writing	69.4	81.8
4–5	Writing	67.8	80.5
6–7	Writing	65.5	78.6
8–9	Writing	67.0	80.0
10–12	Writing	62.8	76.5
2–3	Listening	66.1	77.0
4–5	Listening	60.3	73.6
6–8	Listening	62.0	74.4
9–12	Listening	61.6	75.3
2–3	Speaking	74.1	85.0
4–5	Speaking	69.4	81.5
6–8	Speaking	75.3	85.9
9–12	Speaking	74.0	84.5

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

2023 TELPAS
Scale Score Correlations

Table D.2.1. Spring 2023 TELPAS Assessments Scale Score Correlations within Grade

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
2	Reading	Listening	100,025	0.66
	Reading	Speaking	100,025	0.39
	Reading	Writing	98,111	0.76
	Listening	Speaking	100,127	0.49
	Listening	Writing	97,926	0.61
	Speaking	Writing	97,926	0.43
3	Reading	Listening	100,443	0.70
	Reading	Speaking	100,443	0.42
	Reading	Writing	99,650	0.79
	Listening	Speaking	100,525	0.48
	Listening	Writing	99,466	0.66
	Speaking	Writing	99,466	0.48
4	Reading	Listening	101,512	0.76
	Reading	Speaking	101,512	0.43
	Reading	Writing	100,878	0.73
	Listening	Speaking	101,583	0.47
	Listening	Writing	100,728	0.66
	Speaking	Writing	100,728	0.49
5	Reading	Listening	103,005	0.76
	Reading	Speaking	103,005	0.43
	Reading	Writing	102,614	0.72
	Listening	Speaking	103,091	0.46
	Listening	Writing	102,472	0.67
	Speaking	Writing	102,472	0.50
6	Reading	Listening	100,195	0.75
	Reading	Speaking	100,195	0.42
	Reading	Writing	99,806	0.69
	Listening	Speaking	100,335	0.44
	Listening	Writing	99,576	0.61
	Speaking	Writing	99,576	0.46
7	Reading	Listening	98,215	0.76
	Reading	Speaking	98,215	0.43
	Reading	Writing	97,710	0.69
	Listening	Speaking	98,383	0.45
	Listening	Writing	97,447	0.62
	Speaking	Writing	97,447	0.47

(Continued)

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
8	Reading	Listening	95,163	0.76
	Reading	Speaking	95,163	0.44
	Reading	Writing	94,553	0.71
	Listening	Speaking	95,352	0.47
	Listening	Writing	94,274	0.66
	Speaking	Writing	94,274	0.49
9	Reading	Listening	95,561	0.76
	Reading	Speaking	95,561	0.47
	Reading	Writing	93,336	0.72
	Listening	Speaking	96,206	0.50
	Listening	Writing	92,661	0.70
	Speaking	Writing	92,661	0.50
10	Reading	Listening	76,625	0.76
	Reading	Speaking	76,625	0.45
	Reading	Writing	75,348	0.71
	Listening	Speaking	77,195	0.48
	Listening	Writing	74,724	0.66
	Speaking	Writing	74,724	0.49
11	Reading	Listening	54,605	0.75
	Reading	Speaking	54,605	0.42
	Reading	Writing	53,900	0.69
	Listening	Speaking	55,056	0.46
	Listening	Writing	53,384	0.62
	Speaking	Writing	53,384	0.47
12	Reading	Listening	44,991	0.73
	Reading	Speaking	44,991	0.37
	Reading	Writing	44,551	0.67
	Listening	Speaking	45,325	0.42
	Listening	Writing	44,100	0.59
	Speaking	Writing	44,100	0.43

2023 TELPAS
Conditional Standard Error of Measurement for Scale Scores

Table D.3.1. Spring 2023 TELPAS Grades 2–5 Reading Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 2		Grade 3		Grades 4–5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1174		1058		1069	
1	1256	68	1167	90	1177	89
2	1306	49	1232	65	1242	65
3	1336	41	1272	54	1282	54
4	1359	37	1302	48	1312	48
5	1377	34	1326	44	1336	44
6	1393	31	1346	41	1357	41
7	1407	30	1364	39	1376	39
8	1420	29	1381	37	1392	37
9	1432	28	1396	36	1408	36
10	*1443	27	1410	35	1422	35
11	1453	26	1423	34	*1436	34
12	1464	26	*1436	33	1449	33
13	1474	26	1448	33	1462	33
14	1484	26	1460	32	1474	32
15	1493	26	1472	32	1486	32
16	1503	26	1484	32	1498	32
17	1513	26	1495	32	1509	31
18	1523	26	1507	32	1520	31
19	**1533	26	1518	32	**1532	31
20	1544	27	**1530	32	1543	31
21	1555	28	1542	33	1554	32
22	1567	29	1554	33	1566	32
23	1580	30	1566	34	1578	32
24	1594	31	1580	34	1590	33
25	***1610	34	1594	36	***1602	33
26	1628	37	***1608	37	1616	34
27	1651	41	1624	39	1629	35
28	1681	49	1642	41	1644	36
29	1730	68	1662	44	1660	38
30	1813		1686	48	1678	40
31			1715	54	1698	43
32			1755	65	1721	47
33			1820	90	1750	53
34			1928		1789	64
35					1853	88
36					1959	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table D.3.2. Spring 2023 TELPAS Grades 6–12 Reading Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grades 6–7		Grades 8–9		Grades 10–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1088		1166		1149	
1	1189	83	1251	71	1229	66
2	1251	61	1302	51	1278	48
3	1290	51	1334	43	1308	41
4	1318	45	1357	38	1331	36
5	1341	41	1376	34	1350	33
6	1361	38	1392	32	1366	31
7	1378	36	1406	30	1380	30
8	1394	34	1419	29	1394	29
9	1408	33	1431	28	1406	28
10	1421	32	*1442	27	1418	27
11	1434	31	1452	26	*1428	26
12	1446	30	1462	26	1439	26
13	*1457	30	1471	25	1449	25
14	1468	29	1481	25	1459	25
15	1478	29	1489	25	1468	25
16	1488	28	1498	24	1478	24
17	1498	28	1507	24	1487	24
18	1508	28	1515	24	1496	24
19	1518	28	1524	24	1505	24
20	**1528	28	**1532	24	1514	24
21	1538	28	1541	24	1522	24
22	1548	28	1550	25	**1531	24
23	1558	29	1558	25	1540	24
24	1568	29	1567	25	1549	24
25	1579	29	1577	26	1559	25
26	1590	30	1586	26	1568	25
27	***1601	31	1597	27	1578	26
28	1614	32	***1607	28	1589	26
29	1627	33	1619	29	***1600	27
30	1641	34	1631	30	1612	29
31	1657	36	1645	32	1625	30
32	1675	39	1661	34	1640	32
33	1696	43	1679	37	1658	35
34	1722	49	1702	42	1680	40
35	1757	58	1733	51	1709	48
36	1815	81	1783	70	1756	66
37	1913		1868		1836	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table D.3.3. Spring 2023 TELPAS Grades 2–5 Writing Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grade 2		Grade 3		Grades 4–5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
6	1338		1232			
7	1394	44	1301	53		
8	*1432	33	1351	39		
9	1456	27	1380	32	1234	
10	1473	24	*1401	28	1308	58
11	1488	22	1419	26	1346	38
12	1500	22	1435	26	1366	30
13	1513	21	1452	26	1379	26
14	**1525	22	1469	27	1390	24
15	1538	22	1487	28	1400	23
16	1551	22	1506	28	*1409	23
17	1564	22	**1525	28	1418	23
18	1577	22	1543	27	1428	24
19	1589	21	1559	25	1439	26
20	***1600	20	1574	24	1452	28
21	1610	19	1587	23	1466	29
22	1620	19	***1600	23	1481	28
23	1630	19	1612	22	1494	26
24	1640	19	1624	22	1505	25
25	1650	20	1637	23	1515	24
26	1661	21	1650	24	**1525	23
27	1674	23	1664	26	1534	22
28	1690	26	1683	30	1543	22
29	1714	36	1711	40	1551	22
30	1756		1758		1559	22
31					1568	22
32					1576	22
33					1584	21
34					1592	21
35					***1600	21
36					1608	22
37					1616	22
38					1625	23
39					1636	26
40					1649	31
41					1672	44
42					1719	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table D.3.4. Spring 2023 TELPAS Grades 6–12 Writing Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grades 6–7		Grades 8–9		Grades 10–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
9	1268		1178		1250	
10	1329	51	1275	72	1327	61
11	1358	35	1330	44	1361	39
12	1375	29	1355	33	1379	31
13	1387	26	1371	28	1391	28
14	1398	25	1383	26	1402	26
15	1408	24	1393	24	1411	25
16	1418	24	1403	24	1420	24
17	*1429	25	*1413	24	1428	24
18	1440	26	1424	25	1437	24
19	1451	27	1435	26	*1446	25
20	1464	27	1447	28	1454	25
21	1476	27	1461	28	1463	25
22	1487	26	1474	28	1472	25
23	1498	24	1487	26	1481	25
24	1508	23	1498	24	1490	24
25	1517	23	1507	23	1498	24
26	**1525	22	1516	23	1507	24
27	1533	22	**1525	22	1516	25
28	1542	22	1534	23	**1525	25
29	1550	22	1542	23	1534	26
30	1559	23	1552	24	1544	26
31	1567	23	1562	24	1554	26
32	1576	22	1572	24	1564	26
33	1584	22	1581	24	1573	25
34	1592	22	1591	23	1582	25
35	***1600	21	***1600	23	1591	25
36	1608	22	1609	23	***1600	25
37	1616	22	1618	23	1610	27
38	1625	24	1627	24	1621	29
39	1636	27	1638	27	1635	33
40	1651	33	1653	32	1655	42
41	1677	49	1677	46	1696	69
42	1733		1728		1787	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table D.3.5. Spring 2023 TELPAS Grades 2–12 Listening Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grades 2–3		Grades 4–5		Grades 6–8		Grades 9–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1185		1173		1161		1208	
1	1269	69	1253	66	1246	70	1275	56
2	1319	50	1301	48	1298	51	1317	41
3	1350	42	1331	40	1330	43	1344	35
4	1374	38	1353	36	1354	38	1365	32
5	1393	35	1372	33	1374	35	1382	30
6	1410	33	1387	31	1392	33	1397	28
7	1425	31	1401	29	1407	32	1411	27
8	1439	30	1414	28	1422	31	1425	26
9	*1452	29	1426	27	*1435	30	1437	26
10	1465	29	1438	27	1448	29	*1449	25
11	1477	29	1449	26	1460	28	1461	25
12	1489	28	*1459	26	1472	28	1472	25
13	1501	28	1470	26	1483	28	1484	24
14	1512	28	1480	26	1495	28	1494	24
15	1524	28	1491	26	1506	28	1505	24
16	**1536	28	1501	26	1518	28	1516	24
17	1548	29	1512	27	**1530	29	**1527	24
18	1560	29	1523	27	1542	29	1538	25
19	1573	30	**1535	28	1555	30	1550	25
20	1587	31	1548	29	1569	31	1562	26
21	***1602	33	1562	31	1584	32	1575	27
22	1619	35	1577	33	***1600	35	1589	28
23	1638	38	1595	36	1619	38	***1605	31
24	1662	42	***1617	40	1643	42	1625	34
25	1693	50	1647	48	1674	50	1650	41
26	1744	69	1695	66	1724	69	1691	55
27	1828		1775		1808		1758	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table D.3.6. Spring 2023 TELPAS Grades 2–12 Speaking Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Grades 2–3		Grades 4–5		Grades 6–8		Grades 9–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
10	1272		1364		1395		1406	
11	1317	37	1395	26	1420	21	1432	22
12	1345	27	1415	19	1437	16	1448	16
13	1363	24	1427	16	1447	14	1457	13
14	1377	21	1437	15	1456	12	1465	12
15	1390	20	1445	14	*1463	11	1471	11
16	1401	19	1453	13	1469	11	1476	10
17	*1411	19	1459	13	1475	11	1480	10
18	1422	19	1466	12	1481	10	*1485	10
19	1432	19	*1472	12	1486	10	1489	9
20	1442	19	1478	12	1492	10	1493	9
21	1452	19	1484	13	1497	10	1498	10
22	1463	20	1491	13	1503	11	1502	10
23	1475	21	1498	14	1509	11	1507	10
24	1487	21	1506	14	1516	12	1512	11
25	1501	23	1515	16	1523	12	1519	12
26	1517	24	**1527	17	**1531	13	**1527	14
27	**1535	26	1541	20	1540	14	1539	17
28	1555	27	1557	21	1550	14	1555	20
29	1577	27	1575	20	1560	14	1573	19
30	1598	27	1590	19	1570	14	1588	17
31	***1618	27	***1604	18	1580	14	***1600	16
32	1639	27	1617	18	1589	14	1612	16
33	1661	28	1630	18	***1600	15	1625	17
34	1686	32	1645	21	1612	16	1640	19
35	1721	40	1667	27	1629	21	1663	25
36	1771		1699		1655		1694	

Notes:

* Intermediate, ** Advanced, *** Advanced High

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

2023 TELPAS
Mean P-Values

Table D.4.1. Spring 2023 TELPAS Reading and Writing Grade 2 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	30	100,222	12.86	6.21	0.85[(a)]	2.41	42.86
	Understand words and language structures	12	100,222	5.39	2.84	0.71[(a)]	1.52	44.93
	Basic understanding of variety of texts written in English	7	100,222	2.75	1.70	0.53[(a)]	1.17	39.31
	Analyze and evaluate information and ideas in a variety of texts written in English	11	100,222	4.71	2.58	0.68[(a)]	1.45	42.86
WRITING	OVERALL TEST	30	98,111	10.21	4.16	0.88[(b)]	1.42	27.85
	Express ideas in writing and address writing assignments	24	98,111	8.05	2.81	0.84	1.12	11.37
	Use standard grammar, usage, and spelling to edit writing tasks	6	98,111	2.17	1.73	0.74[(a)]	0.87	36.08

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.2. Spring 2023 TELPAS Reading and Writing Grade 2 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	30	48,635	12.99	6.25	0.85[(a)]	2.40	43.29
	Understand words and language structures	12	48,635	5.36	2.81	0.71[(a)]	1.52	44.69
	Basic understanding of variety of texts written in English	7	48,635	2.81	1.73	0.54[(a)]	1.17	40.13
	Analyze and evaluate information and ideas in a variety of texts written in English	11	48,635	4.82	2.59	0.69[(a)]	1.45	43.78
WRITING	OVERALL TEST	30	47,829	10.48	4.32	0.89[(b)]	1.45	28.78
	Express ideas in writing and address writing assignments	24	47,829	8.27	2.97	0.85	1.16	12.60
	Use standard grammar, usage, and spelling to edit writing tasks	6	47,829	2.21	1.73	0.74[(a)]	0.88	36.87

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.3. Spring 2023 TELPAS Reading and Writing Grade 2 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	30	51,393	12.74	6.18	0.85[(a)]	2.41	42.46
	Understand words and language structures	12	51,393	5.42	2.87	0.72[(a)]	1.52	45.17
	Basic understanding of variety of texts written in English	7	51,393	2.70	1.66	0.51[(a)]	1.17	38.54
	Analyze and evaluate information and ideas in a variety of texts written in English	11	51,393	4.62	2.56	0.68[(a)]	1.45	42.01
WRITING	OVERALL TEST	30	50,106	9.96	3.98	0.88[(b)]	1.39	26.97
	Express ideas in writing and address writing assignments	24	50,106	7.84	2.63	0.83	1.08	10.21
	Use standard grammar, usage, and spelling to edit writing tasks	6	50,106	2.12	1.72	0.74[(a)]	0.87	35.34

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.4. Spring 2023 TELPAS Reading and Writing Grade 3 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	34	100,631	18.43	8.06	0.90[(a)]	2.49	54.20
	Understand words and language structures	12	100,631	7.15	3.29	0.82[(a)]	1.41	59.55
	Basic understanding of variety of texts written in English	9	100,631	4.31	2.19	0.63[(a)]	1.33	47.85
	Analyze and evaluate information and ideas in a variety of texts written in English	13	100,631	6.98	3.39	0.79[(a)]	1.55	53.67
WRITING	OVERALL TEST	30	99,650	13.52	5.07	0.91[(b)]	1.56	45.97
	Express ideas in writing and address writing assignments	24	99,650	10.06	3.58	0.87	1.31	22.58
	Use standard grammar, usage, and spelling to edit writing tasks	6	99,650	3.46	1.86	0.80[(a)]	0.84	57.67

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.5. Spring 2023 TELPAS Reading and Writing Grade 3 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	34	48,734	18.65	7.96	0.90[(a)]	2.48	54.86
	Understand words and language structures	12	48,734	7.16	3.22	0.81[(a)]	1.40	59.70
	Basic understanding of variety of texts written in English	9	48,734	4.39	2.18	0.63[(a)]	1.33	48.73
	Analyze and evaluate information and ideas in a variety of texts written in English	13	48,734	7.10	3.38	0.79[(a)]	1.54	54.65
WRITING	OVERALL TEST	30	48,332	14.09	5.21	0.91[(b)]	1.59	47.79
	Express ideas in writing and address writing assignments	24	48,332	10.54	3.74	0.87	1.35	25.23
	Use standard grammar, usage, and spelling to edit writing tasks	6	48,332	3.54	1.84	0.79[(a)]	0.84	59.07

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.6. Spring 2023 TELPAS Reading and Writing Grade 3 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	34	51,732	18.22	8.14	0.91[(a)]	2.49	53.59
	Understand words and language structures	12	51,732	7.13	3.35	0.82[(a)]	1.41	59.42
	Basic understanding of variety of texts written in English	9	51,732	4.23	2.20	0.63[(a)]	1.33	47.03
	Analyze and evaluate information and ideas in a variety of texts written in English	13	51,732	6.86	3.40	0.79[(a)]	1.55	52.76
WRITING	OVERALL TEST	30	51,164	13.00	4.87	0.90[(b)]	1.52	44.27
	Express ideas in writing and address writing assignments	24	51,164	9.61	3.37	0.86	1.26	20.08
	Use standard grammar, usage, and spelling to edit writing tasks	6	51,164	3.38	1.88	0.80[(a)]	0.84	56.36

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.7. Spring 2023 TELPAS Reading and Writing Grades 4–5 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	36	204,820	21.44	8.26	0.91[(a)]	2.48	59.57
	Understand words and language structures	12	204,820	7.80	2.97	0.79[(a)]	1.37	65.03
	Basic understanding of variety of texts written in English	10	204,820	5.60	2.71	0.76[(a)]	1.31	55.97
	Analyze and evaluate information and ideas in a variety of texts written in English	14	204,820	8.04	3.35	0.77[(a)]	1.59	57.45
WRITING	OVERALL TEST	42	203,492	23.52	7.85	0.89[(b)]	2.60	52.04
	Express ideas in writing and address writing assignments	36	203,492	20.07	6.45	0.86	2.43	41.00
	Use standard grammar, usage, and spelling to edit writing tasks	6	203,492	3.45	1.84	0.75[(a)]	0.93	57.56

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.8. Spring 2023 TELPAS Reading and Writing Grades 4–5 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	36	98,925	21.89	8.04	0.91[(a)]	2.47	60.81
	Understand words and language structures	12	98,925	7.93	2.89	0.78[(a)]	1.37	66.07
	Basic understanding of variety of texts written in English	10	98,925	5.75	2.65	0.76[(a)]	1.30	57.54
	Analyze and evaluate information and ideas in a variety of texts written in English	14	98,925	8.21	3.25	0.76[(a)]	1.58	58.65
WRITING	OVERALL TEST	42	98,332	24.67	7.84	0.89[(b)]	2.66	54.69
	Express ideas in writing and address writing assignments	36	98,332	21.09	6.45	0.85	2.49	44.76
	Use standard grammar, usage, and spelling to edit writing tasks	6	98,332	3.58	1.83	0.75[(a)]	0.92	59.66

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.9. Spring 2023 TELPAS Reading and Writing Grades 4–5 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	36	105,639	21.03	8.45	0.91[(a)]	2.49	58.43
	Understand words and language structures	12	105,639	7.69	3.03	0.79[(a)]	1.38	64.09
	Basic understanding of variety of texts written in English	10	105,639	5.45	2.74	0.77[(a)]	1.32	54.54
	Analyze and evaluate information and ideas in a variety of texts written in English	14	105,639	7.89	3.42	0.78[(a)]	1.60	56.35
WRITING	OVERALL TEST	42	104,915	22.46	7.71	0.89[(b)]	2.55	49.58
	Express ideas in writing and address writing assignments	36	104,915	19.13	6.30	0.86	2.38	37.50
	Use standard grammar, usage, and spelling to edit writing tasks	6	104,915	3.34	1.84	0.75[(a)]	0.93	55.62

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.10. Spring 2023 TELPAS Reading and Writing Grades 6–7 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	198,921	21.31	7.48	0.88[(a)]	2.63	57.61
	Understand words and language structures	12	198,921	8.14	2.64	0.73[(a)]	1.38	67.84
	Basic understanding of variety of texts written in English	10	198,921	5.36	2.44	0.68[(a)]	1.39	53.59
	Analyze and evaluate information and ideas in a variety of texts written in English	15	198,921	7.81	3.38	0.74[(a)]	1.74	52.10
WRITING	OVERALL TEST	42	197,516	24.91	7.55	0.87[(b)]	2.72	51.05
	Express ideas in writing and address writing assignments	36	197,516	21.73	6.25	0.84	2.54	47.14
	Use standard grammar, usage, and spelling to edit writing tasks	6	197,516	3.18	1.84	0.71[(a)]	0.98	53.01

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.11. Spring 2023 TELPAS Reading and Writing Grades 6–7 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	95,363	21.94	7.43	0.88[(a)]	2.62	59.31
	Understand words and language structures	12	95,363	8.25	2.58	0.72[(a)]	1.37	68.74
	Basic understanding of variety of texts written in English	10	95,363	5.53	2.43	0.68[(a)]	1.38	55.33
	Analyze and evaluate information and ideas in a variety of texts written in English	15	95,363	8.16	3.38	0.74[(a)]	1.73	54.41
WRITING	OVERALL TEST	42	94,791	26.60	7.43	0.86[(b)]	2.79	54.82
	Express ideas in writing and address writing assignments	36	94,791	23.25	6.15	0.82	2.60	52.77
	Use standard grammar, usage, and spelling to edit writing tasks	6	94,791	3.35	1.83	0.71[(a)]	0.99	55.84

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.12. Spring 2023 TELPAS Reading and Writing Grades 6–7 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	103,249	20.74	7.48	0.88[(a)]	2.64	56.06
	Understand words and language structures	12	103,249	8.04	2.68	0.73[(a)]	1.38	67.04
	Basic understanding of variety of texts written in English	10	103,249	5.20	2.44	0.67[(a)]	1.40	52.02
	Analyze and evaluate information and ideas in a variety of texts written in English	15	103,249	7.50	3.36	0.73[(a)]	1.75	49.98
WRITING	OVERALL TEST	42	102,430	23.35	7.30	0.87[(b)]	2.65	47.61
	Express ideas in writing and address writing assignments	36	102,430	20.33	6.01	0.83	2.46	41.95
	Use standard grammar, usage, and spelling to edit writing tasks	6	102,430	3.03	1.84	0.72[(a)]	0.98	50.43

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.13. Spring 2023 TELPAS Reading and Writing Grades 8–9 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	191,746	21.60	8.22	0.90[(a)]	2.57	58.39
	Understand words and language structures	12	191,746	7.46	2.93	0.77[(a)]	1.42	62.15
	Basic understanding of variety of texts written in English	10	191,746	5.78	2.58	0.73[(a)]	1.33	57.84
	Analyze and evaluate information and ideas in a variety of texts written in English	15	191,746	8.36	3.50	0.77[(a)]	1.68	55.75
WRITING	OVERALL TEST	42	187,889	24.07	7.61	0.88[(b)]	2.59	52.37
	Express ideas in writing and address writing assignments	36	187,889	20.65	6.45	0.86	2.43	43.14
	Use standard grammar, usage, and spelling to edit writing tasks	6	187,889	3.42	1.67	0.70[(a)]	0.91	56.99

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.14. Spring 2023 TELPAS Reading and Writing Grades 8–9 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	89,236	22.49	7.89	0.90[(a)]	2.54	60.79
	Understand words and language structures	12	89,236	7.61	2.82	0.75[(a)]	1.40	63.38
	Basic understanding of variety of texts written in English	10	89,236	6.08	2.48	0.72[(a)]	1.31	60.78
	Analyze and evaluate information and ideas in a variety of texts written in English	15	89,236	8.81	3.37	0.76[(a)]	1.66	58.73
WRITING	OVERALL TEST	42	87,670	25.80	7.53	0.88[(b)]	2.66	56.19
	Express ideas in writing and address writing assignments	36	87,670	22.21	6.38	0.85	2.50	48.91
	Use standard grammar, usage, and spelling to edit writing tasks	6	87,670	3.59	1.64	0.70[(a)]	0.90	59.83

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.15. Spring 2023 TELPAS Reading and Writing Grades 8–9 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	101,885	20.84	8.42	0.90[(a)]	2.60	56.33
	Understand words and language structures	12	101,885	7.33	3.02	0.78[(a)]	1.42	61.12
	Basic understanding of variety of texts written in English	10	101,885	5.53	2.64	0.74[(a)]	1.35	55.31
	Analyze and evaluate information and ideas in a variety of texts written in English	15	101,885	7.98	3.56	0.77[(a)]	1.70	53.17
WRITING	OVERALL TEST	42	99,640	22.56	7.35	0.88[(b)]	2.52	49.05
	Express ideas in writing and address writing assignments	36	99,640	19.29	6.19	0.86	2.35	38.10
	Use standard grammar, usage, and spelling to edit writing tasks	6	99,640	3.27	1.68	0.71[(a)]	0.91	54.53

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.16. Spring 2023 TELPAS Reading and Writing Grades 10–12 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	177,890	21.80	7.82	0.90[(a)]	2.50	58.93
	Understand words and language structures	12	177,890	9.03	2.86	0.83[(a)]	1.17	75.27
	Basic understanding of variety of texts written in English	10	177,890	4.77	2.54	0.71[(a)]	1.38	47.65
	Analyze and evaluate information and ideas in a variety of texts written in English	15	177,890	8.01	3.50	0.77[(a)]	1.68	53.37
WRITING	OVERALL TEST	42	173,799	26.49	7.86	0.85[(b)]	3.01	52.10
	Express ideas in writing and address writing assignments	36	173,799	23.40	6.95	0.83	2.86	53.33
	Use standard grammar, usage, and spelling to edit writing tasks	6	173,799	3.09	1.50	0.62[(a)]	0.92	51.49

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.17. Spring 2023 TELPAS Reading and Writing Grades 10–12 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	82,099	22.26	7.79	0.90[(a)]	2.49	60.16
	Understand words and language structures	12	82,099	9.17	2.82	0.83[(a)]	1.16	76.44
	Basic understanding of variety of texts written in English	10	82,099	4.93	2.55	0.71[(a)]	1.37	49.29
	Analyze and evaluate information and ideas in a variety of texts written in English	15	82,099	8.16	3.48	0.77[(a)]	1.67	54.39
WRITING	OVERALL TEST	42	80,486	27.98	7.55	0.84[(b)]	3.01	54.84
	Express ideas in writing and address writing assignments	36	80,486	24.80	6.63	0.81	2.87	58.51
	Use standard grammar, usage, and spelling to edit writing tasks	6	80,486	3.18	1.49	0.63[(a)]	0.91	53.01

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.18. Spring 2023 TELPAS Reading and Writing Grades 10–12 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
READING	OVERALL TEST	37	94,989	21.42	7.82	0.90[(a)]	2.50	57.88
	Understand words and language structures	12	94,989	8.91	2.88	0.83[(a)]	1.18	74.29
	Basic understanding of variety of texts written in English	10	94,989	4.62	2.52	0.70[(a)]	1.38	46.24
	Analyze and evaluate information and ideas in a variety of texts written in English	15	94,989	7.88	3.50	0.77[(a)]	1.67	52.52
WRITING	OVERALL TEST	42	92,569	25.20	7.89	0.86[(b)]	3.00	49.74
	Express ideas in writing and address writing assignments	36	92,569	22.19	6.99	0.83	2.86	48.84
	Use standard grammar, usage, and spelling to edit writing tasks	6	92,569	3.01	1.50	0.62[(a)]	0.93	50.19

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.19. Spring 2023 TELPAS Listening and Speaking Grades 2–3 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	200,652	18.62	6.05	0.89	2.05	68.15
	Understand spoken words and language structures	5	200,652	3.92	1.31	0.67[(a)]	0.75	78.45
	Basic understanding of spoken English	16	200,652	11.15	3.68	0.82	1.57	68.34
	Analyze and evaluate information in spoken English	6	200,652	3.54	1.72	0.62[(a)]	1.06	59.08
SPEAKING	OVERALL TEST	36	200,652	21.85	6.23	0.92	1.73	49.92
	Provide and summarize information	16	200,652	10.35	2.76	0.81	1.20	56.52
	Share opinions and analyze information	20	200,652	11.50	3.67	0.89	1.22	43.31

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.20. Spring 2023 TELPAS Listening and Speaking Grades 2–3 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	97,282	18.70	5.94	0.88	2.04	68.43
	Understand spoken words and language structures	5	97,282	3.91	1.30	0.66[(a)]	0.76	78.23
	Basic understanding of spoken English	16	97,282	11.26	3.60	0.81	1.57	68.99
	Analyze and evaluate information in spoken English	6	97,282	3.53	1.71	0.62[(a)]	1.06	58.87
SPEAKING	OVERALL TEST	36	97,282	22.30	6.13	0.92	1.75	51.74
	Provide and summarize information	16	97,282	10.55	2.72	0.80	1.22	58.51
	Share opinions and analyze information	20	97,282	11.74	3.63	0.88	1.24	44.96

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.21. Spring 2023 TELPAS Listening and Speaking Grades 2–3 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	103,012	18.55	6.14	0.89	2.05	67.91
	Understand spoken words and language structures	5	103,012	3.93	1.32	0.68[(a)]	0.75	78.69
	Basic understanding of spoken English	16	103,012	11.05	3.74	0.82	1.58	67.76
	Analyze and evaluate information in spoken English	6	103,012	3.56	1.73	0.63[(a)]	1.05	59.30
SPEAKING	OVERALL TEST	36	103,012	21.44	6.28	0.93	1.71	48.25
	Provide and summarize information	16	103,012	10.17	2.79	0.82	1.19	54.69
	Share opinions and analyze information	20	103,012	11.27	3.69	0.89	1.21	41.80

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.22. Spring 2023 TELPAS Listening and Speaking Grades 4–5 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	204,674	19.49	5.90	0.88[(a)]	2.02	72.17
	Understand spoken words and language structures	5	204,674	3.85	1.26	0.58[(a)]	0.82	77.02
	Basic understanding of spoken English	16	204,674	11.49	3.69	0.82[(a)]	1.56	71.82
	Analyze and evaluate information in spoken English	6	204,674	4.14	1.62	0.63[(a)]	0.98	69.07
SPEAKING	OVERALL TEST	36	204,674	24.40	6.05	0.92	1.73	59.55
	Provide and summarize information	16	204,674	11.67	2.58	0.81	1.13	67.53
	Share opinions and analyze information	20	204,674	12.73	3.67	0.88	1.26	51.56

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.23. Spring 2023 TELPAS Listening and Speaking Grades 4–5 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	98,863	19.68	5.80	0.88[(a)]	2.00	72.89
	Understand spoken words and language structures	5	98,863	3.88	1.24	0.58[(a)]	0.80	77.62
	Basic understanding of spoken English	16	98,863	11.58	3.63	0.82[(a)]	1.55	72.41
	Analyze and evaluate information in spoken English	6	98,863	4.22	1.60	0.63[(a)]	0.97	70.26
SPEAKING	OVERALL TEST	36	98,863	24.56	6.01	0.91	1.75	60.09
	Provide and summarize information	16	98,863	11.72	2.57	0.80	1.15	67.91
	Share opinions and analyze information	20	98,863	12.84	3.66	0.88	1.29	52.26

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.24. Spring 2023 TELPAS Listening and Speaking Grades 4–5 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	105,560	19.31	5.99	0.88[(a)]	2.03	71.53
	Understand spoken words and language structures	5	105,560	3.82	1.27	0.58[(a)]	0.83	76.49
	Basic understanding of spoken English	16	105,560	11.41	3.75	0.83[(a)]	1.56	71.31
	Analyze and evaluate information in spoken English	6	105,560	4.08	1.63	0.63[(a)]	0.99	67.99
SPEAKING	OVERALL TEST	36	105,560	24.26	6.07	0.92	1.70	59.07
	Provide and summarize information	16	105,560	11.62	2.59	0.81	1.12	67.21
	Share opinions and analyze information	20	105,560	12.64	3.68	0.89	1.23	50.93

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.25. Spring 2023 TELPAS Listening and Speaking Grades 6–8 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	294,070	18.79	5.38	0.85[(a)]	2.10	69.58
	Understand spoken words and language structures	5	294,070	4.05	1.06	0.48[(a)]	0.76	80.93
	Basic understanding of spoken English	16	294,070	10.75	3.47	0.77[(a)]	1.67	67.20
	Analyze and evaluate information in spoken English	6	294,070	3.99	1.57	0.58[(a)]	1.02	66.48
SPEAKING	OVERALL TEST	36	294,070	23.09	6.87	0.93	1.81	54.66
	Provide and summarize information	16	294,070	11.32	3.08	0.82	1.31	64.17
	Share opinions and analyze information	20	294,070	11.77	4.02	0.91	1.20	45.15

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.26. Spring 2023 TELPAS Listening and Speaking Grades 6–8 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	140,506	18.99	5.31	0.85[(a)]	2.08	70.32
	Understand spoken words and language structures	5	140,506	4.08	1.05	0.48[(a)]	0.76	81.59
	Basic understanding of spoken English	16	140,506	10.83	3.41	0.76[(a)]	1.65	67.66
	Analyze and evaluate information in spoken English	6	140,506	4.08	1.55	0.58[(a)]	1.01	68.02
SPEAKING	OVERALL TEST	36	140,506	23.20	6.92	0.93	1.85	55.10
	Provide and summarize information	16	140,506	11.38	3.08	0.81	1.33	64.69
	Share opinions and analyze information	20	140,506	11.83	4.09	0.91	1.23	45.52

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.27. Spring 2023 TELPAS Listening and Speaking Grades 6–8 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	153,112	18.61	5.44	0.85[(a)]	2.12	68.92
	Understand spoken words and language structures	5	153,112	4.02	1.06	0.48[(a)]	0.77	80.35
	Basic understanding of spoken English	16	153,112	10.69	3.52	0.77[(a)]	1.67	66.79
	Analyze and evaluate information in spoken English	6	153,112	3.91	1.58	0.57[(a)]	1.03	65.09
SPEAKING	OVERALL TEST	36	153,112	22.99	6.81	0.93	1.77	54.28
	Provide and summarize information	16	153,112	11.26	3.08	0.83	1.28	63.72
	Share opinions and analyze information	20	153,112	11.73	3.96	0.91	1.17	44.84

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.28. Spring 2023 TELPAS Listening and Speaking Grades 9–12 Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	273,782	18.07	5.43	0.86[(a)]	2.04	66.92
	Understand spoken words and language structures	5	273,782	3.68	1.11	0.46[(a)]	0.81	73.60
	Basic understanding of spoken English	16	273,782	10.82	3.36	0.79[(a)]	1.55	67.62
	Analyze and evaluate information in spoken English	6	273,782	3.57	1.65	0.60[(a)]	1.04	59.50
SPEAKING	OVERALL TEST	36	273,782	21.84	7.65	0.95	1.72	48.88
	Provide and summarize information	16	273,782	10.14	3.42	0.88	1.20	53.10
	Share opinions and analyze information	20	273,782	11.70	4.40	0.93	1.20	44.66

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.29. Spring 2023 TELPAS Listening and Speaking Grades 9–12 Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	125,889	18.28	5.26	0.85[(a)]	2.02	67.70
	Understand spoken words and language structures	5	125,889	3.68	1.09	0.45[(a)]	0.81	73.57
	Basic understanding of spoken English	16	125,889	10.94	3.25	0.78[(a)]	1.53	68.35
	Analyze and evaluate information in spoken English	6	125,889	3.67	1.62	0.59[(a)]	1.03	61.08
SPEAKING	OVERALL TEST	36	125,889	22.27	7.67	0.95	1.77	50.44
	Provide and summarize information	16	125,889	10.28	3.44	0.87	1.24	54.30
	Share opinions and analyze information	20	125,889	11.99	4.42	0.92	1.23	46.58

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

Table D.4.30. Spring 2023 TELPAS Listening and Speaking Grades 9–12 Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
LISTENING	OVERALL TEST	27	146,612	17.89	5.56	0.86[(a)]	2.05	66.27
	Understand spoken words and language structures	5	146,612	3.68	1.13	0.48[(a)]	0.82	73.65
	Basic understanding of spoken English	16	146,612	10.72	3.45	0.80[(a)]	1.56	67.01
	Analyze and evaluate information in spoken English	6	146,612	3.49	1.67	0.61[(a)]	1.05	58.16
SPEAKING	OVERALL TEST	36	146,612	21.48	7.60	0.95	1.68	47.60
	Provide and summarize information	16	146,612	10.02	3.40	0.88	1.17	52.13
	Share opinions and analyze information	20	146,612	11.46	4.37	0.93	1.17	43.07

Notes:
1. Total number of Score Points
2. Coefficient Alpha computed if there are any polytomous items;
(a) KR-20 is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items

2023 TELPAS
Scale Score
Descriptive Statistics

Glossary

This glossary provides definitions for the statistical terms that appear in the tables and graphs in this section. Definitions of statistical terms and concepts in the other sections are given in Chapter 3 and Chapter 6.

Descriptive Statistics

Mean The mean is a measure of central tendency. It is the average score for the assessment. It is computed by summing the scores of all students and dividing the sum by the total number of students (*N*).

Median The median is another measure of central tendency. It is the score at the middle of the frequency distribution for the assessment. It is computed by finding the score at which there are the same number of scores above as there are below.

Mode The mode is another measure of central tendency. It is the most frequently obtained score for the assessment. It is determined by computing the frequency distribution and finding the score point with the highest frequency (*n*-count).

Range The range is a measure of statistical dispersion (variability or spread). It is the difference between the lowest and highest scores obtained by students on the assessment. It is computed by subtracting the lowest score from the highest score.

Interquartile Range The interquartile range is another measure of statistical dispersion (variability or spread). It is the difference between the first and third quartiles (or 25th and 75th percentiles) of the score distribution for the assessment. It is computed by subtracting the score at the first quartile (the point that splits the lowest 25 percent of the scores) from the score at the third quartile (the point that splits the highest 25 percent of the scores).

Standard Deviation The standard deviation is another measure of statistical dispersion (variability or spread). It is an indicator of the degree of score variation around the mean. It is computed using the following formula:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, and N is the total number of students who took the assessment.

Variance The variance is another measure of statistical dispersion (variability or spread) around the mean. It is computed as the square of the standard deviation.

Skewness The skewness is an indicator of the shape of the score distribution. It measures the extent to which the score distribution "leans" to one side of the mean. A positive skewness indicates that the score

distribution leans below the mean. A negative skewness indicates that the score distribution leans above the mean. A skewness of zero indicates that the score distribution is symmetric around the mean. It is computed using the following formula:

$$Skewness = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^3,$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Kurtosis The kurtosis is another indicator of the shape of the score distribution. It measures the "peakedness" of the score distribution. A positive kurtosis is referred to as *leptokurtic*, meaning that the distribution has a more acute peak around the mean and fatter tails. A negative kurtosis is called *platykurtic*, meaning the distribution has a lower, wider peak around the mean and thinner tails. It is computed using the following formula:

$$Kurtosis = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Frequency Distributions

Frequency (Freq) The frequency is the number of students who obtained the particular score point on the assessment.

Cumulative Frequency (CumFreq) The cumulative frequency is the number of students who obtained a score that is less than or equal to the particular score point on the assessment.

Percentage (Pct) The percentage is the percentage of students who obtained the particular score point on the assess-ment. It is computed as: Pct = Freq ÷ N × 100.

Cumulative Percentage (CumPct) The cumulative percentage is the percentage of students who obtained a score that is less than or equal to the particular score point on the assessment. It is computed as: CumPct = CumFreq ÷ N × 100.

Table D.5.1. Spring 2023 TELPAS Assessments Reading Scale Score Descriptive Statistics (Online Only)

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Reading	100,222	1471.34	1453	1420	639	93	73.17	5353.32	0.85	1.23
Grade 3	Reading	100,631	1519.97	1507	1410	870	171	117.77	13869.92	0.55	0.15
Grade 4	Reading	101,667	1550.91	1543	1498	890	154	115.70	13387.46	0.65	0.52
Grade 5	Reading	103,153	1596.51	1590	1721	890	180	129.36	16732.97	0.49	0.07
Grade 6	Reading	100,432	1540.35	1538	1538	825	123	87.31	7622.17	0.41	0.46
Grade 7	Reading	98,489	1553.73	1548	1558	825	126	92.61	8576.99	0.38	0.38
Grade 8	Reading	95,450	1557.60	1558	1631	702	130	83.85	7030.38	0.12	-0.20
Grade 9	Reading	96,296	1547.83	1550	1619	702	126	87.01	7570.58	0.14	-0.18
Grade 10	Reading	77,284	1531.61	1522	1505	687	121	87.49	7654.15	0.39	0.49
Grade 11	Reading	55,145	1539.92	1531	1514	687	122	87.36	7632.08	0.32	0.61
Grade 12	Reading	45,461	1537.40	1531	1505	687	111	83.37	6950.75	0.32	0.81

Table D.5.2. Spring 2023 TELPAS Assessments Writing Scale Score Descriptive Statistics (Online Only)

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Writing	98,111	1451.39	1456	1394	418	119	78.35	6139.02	0.18	-0.87
Grade 3	Writing	99,650	1449.00	1469	1301	526	145	105.25	11077.64	-0.36	-0.67
Grade 4	Writing	100,878	1468.04	1494	1234	485	125	99.26	9853.22	-0.79	0.05
Grade 5	Writing	102,614	1499.91	1525	1525	485	107	93.50	8741.92	-1.06	1.04
Grade 6	Writing	99,806	1499.24	1517	1525	465	95	84.39	7122.47	-0.99	1.01
Grade 7	Writing	97,710	1511.45	1525	1550	465	91	84.17	7084.40	-1.02	1.31
Grade 8	Writing	94,553	1488.96	1507	1498	550	105	94.76	8979.29	-1.16	1.95
Grade 9	Writing	93,336	1477.09	1498	1498	550	107	104.06	10828.10	-1.06	1.15
Grade 10	Writing	75,348	1506.37	1516	1554	537	101	86.80	7533.57	-0.64	1.21
Grade 11	Writing	53,900	1515.95	1525	1554	537	101	81.73	6680.50	-0.56	1.44
Grade 12	Writing	44,551	1513.91	1525	1554	537	92	78.30	6131.45	-0.51	1.39

Table D.5.3. Spring 2023 TELPAS Assessments Listening Scale Score Descriptive Statistics (Online Only)

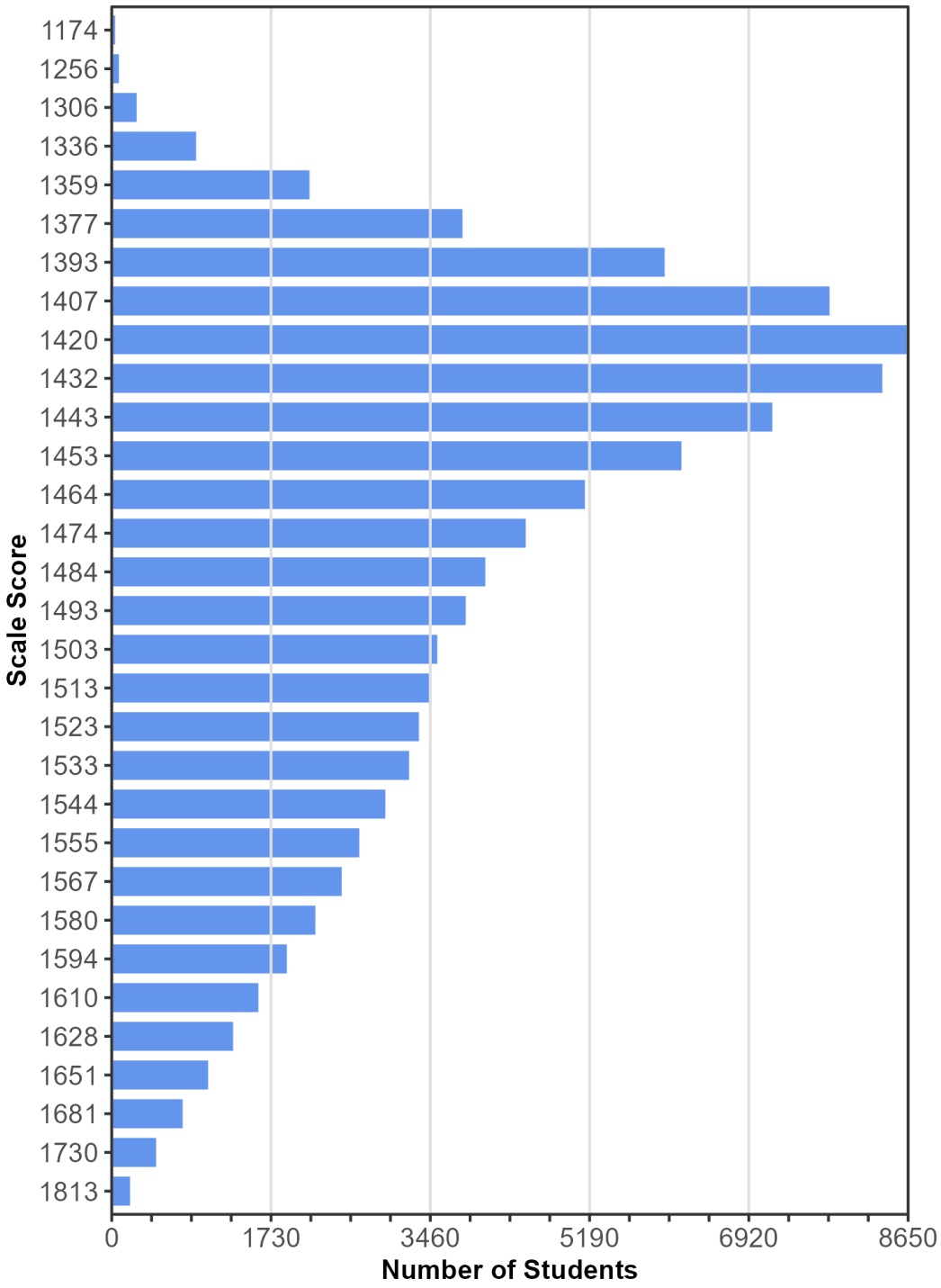
Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Listening	100,127	1557.47	1548	1638	643	130	91.42	8358.15	0.36	0.19
Grade 3	Listening	100,525	1618.90	1619	1693	643	145	101.95	10394.65	-0.01	-0.26
Grade 4	Listening	101,583	1547.87	1548	1617	602	104	89.45	8001.52	0.14	-0.05
Grade 5	Listening	103,091	1577.22	1577	1647	602	135	94.43	8916.69	-0.02	-0.19
Grade 6	Listening	100,335	1557.98	1555	1600	647	94	82.67	6833.77	0.37	0.48
Grade 7	Listening	98,383	1567.45	1569	1619	647	113	88.49	7830.75	0.32	0.27
Grade 8	Listening	95,352	1580.18	1584	1643	647	125	93.61	8762.99	0.22	0.07
Grade 9	Listening	96,206	1538.75	1538	1562	550	95	76.15	5798.83	-0.09	0.52
Grade 10	Listening	77,195	1548.21	1550	1589	550	84	74.73	5583.91	-0.04	0.55
Grade 11	Listening	55,056	1555.16	1550	1589	550	100	74.31	5521.76	-0.06	0.80
Grade 12	Listening	45,325	1552.33	1550	1589	550	84	71.37	5092.99	-0.03	1.03

Table D.5.4. Spring 2023 TELPAS Assessments Speaking Scale Score Descriptive Statistics (Online Only)

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Speaking	100,127	1450.60	1452	1272	499	100	88.80	7885.93	-0.10	0.00
Grade 3	Speaking	100,525	1485.89	1487	1535	499	123	91.04	8288.75	-0.33	0.24
Grade 4	Speaking	101,583	1520.26	1527	1557	335	73	63.67	4053.41	-0.24	0.35
Grade 5	Speaking	103,091	1525.80	1527	1557	335	91	66.18	4379.57	-0.23	0.25
Grade 6	Speaking	100,335	1518.38	1523	1540	260	64	53.15	2825.12	-0.16	0.09
Grade 7	Speaking	98,383	1513.17	1516	1550	260	69	56.00	3136.09	-0.13	-0.12
Grade 8	Speaking	95,352	1513.19	1516	1395	260	69	58.24	3392.23	-0.16	-0.24
Grade 9	Speaking	96,206	1504.73	1507	1406	288	98	63.64	4050.24	0.04	-0.62
Grade 10	Speaking	77,195	1510.52	1512	1406	288	90	64.80	4198.60	-0.05	-0.62
Grade 11	Speaking	55,056	1514.78	1519	1406	288	84	66.10	4368.62	-0.08	-0.62
Grade 12	Speaking	45,325	1510.39	1512	1406	288	90	65.81	4331.57	-0.04	-0.65

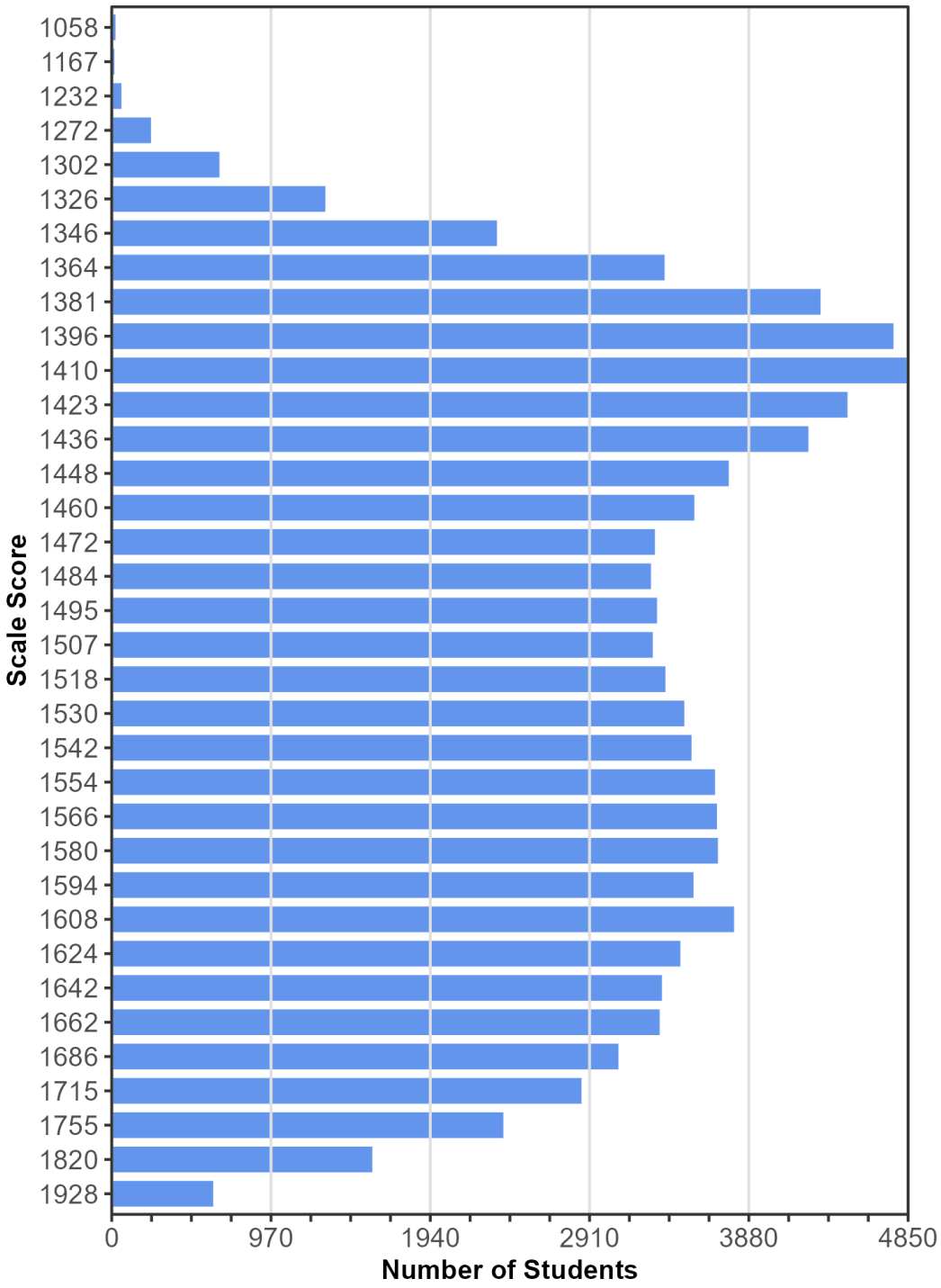
2023 TELPAS
Frequency Distribution of Scale Scores

Figure D.5.1. Spring 2023 TELPAS Grade 2 Reading
Frequency Distribution of Scale Scores
(Online Only)



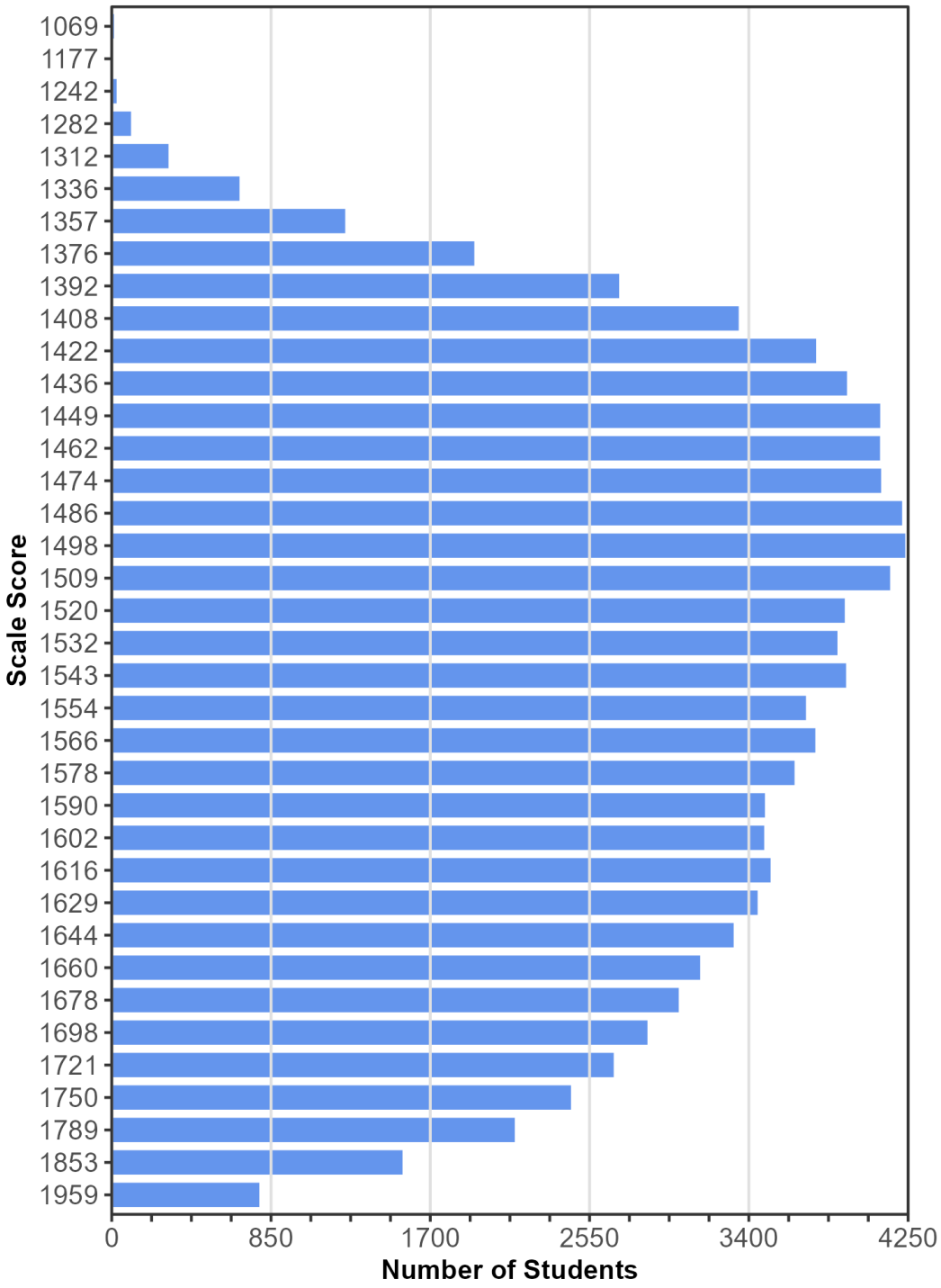
	Freq	Cum Freq	Pct	Cum Pct
1174	37	37	0.04	0.04
1256	78	115	0.08	0.11
1306	271	386	0.27	0.39
1336	917	1303	0.91	1.30
1359	2148	3451	2.14	3.44
1377	3810	7261	3.80	7.24
1393	6006	13267	5.99	13.24
1407	7797	21064	7.78	21.02
1420	8641	29705	8.62	29.64
1432	8370	38075	8.35	37.99
1443	7175	45250	7.16	45.15
1453	6188	51438	6.17	51.32
1464	5140	56578	5.13	56.45
1474	4497	61075	4.49	60.94
1484	4058	65133	4.05	64.99
1493	3845	68978	3.84	68.83
1503	3536	72514	3.53	72.35
1513	3459	75973	3.45	75.80
1523	3337	79310	3.33	79.13
1533	3230	82540	3.22	82.36
1544	2972	85512	2.97	85.32
1555	2689	88201	2.68	88.01
1567	2498	90699	2.49	90.50
1580	2212	92911	2.21	92.71
1594	1902	94813	1.90	94.60
1610	1593	96406	1.59	96.19
1628	1318	97724	1.32	97.51
1651	1047	98771	1.04	98.55
1681	770	99541	0.77	99.32
1730	482	100023	0.48	99.80
1813	199	100222	0.20	100.00

Figure D.5.2. Spring 2023 TELPAS Grade 3 Reading
Frequency Distribution of Scale Scores
(Online Only)



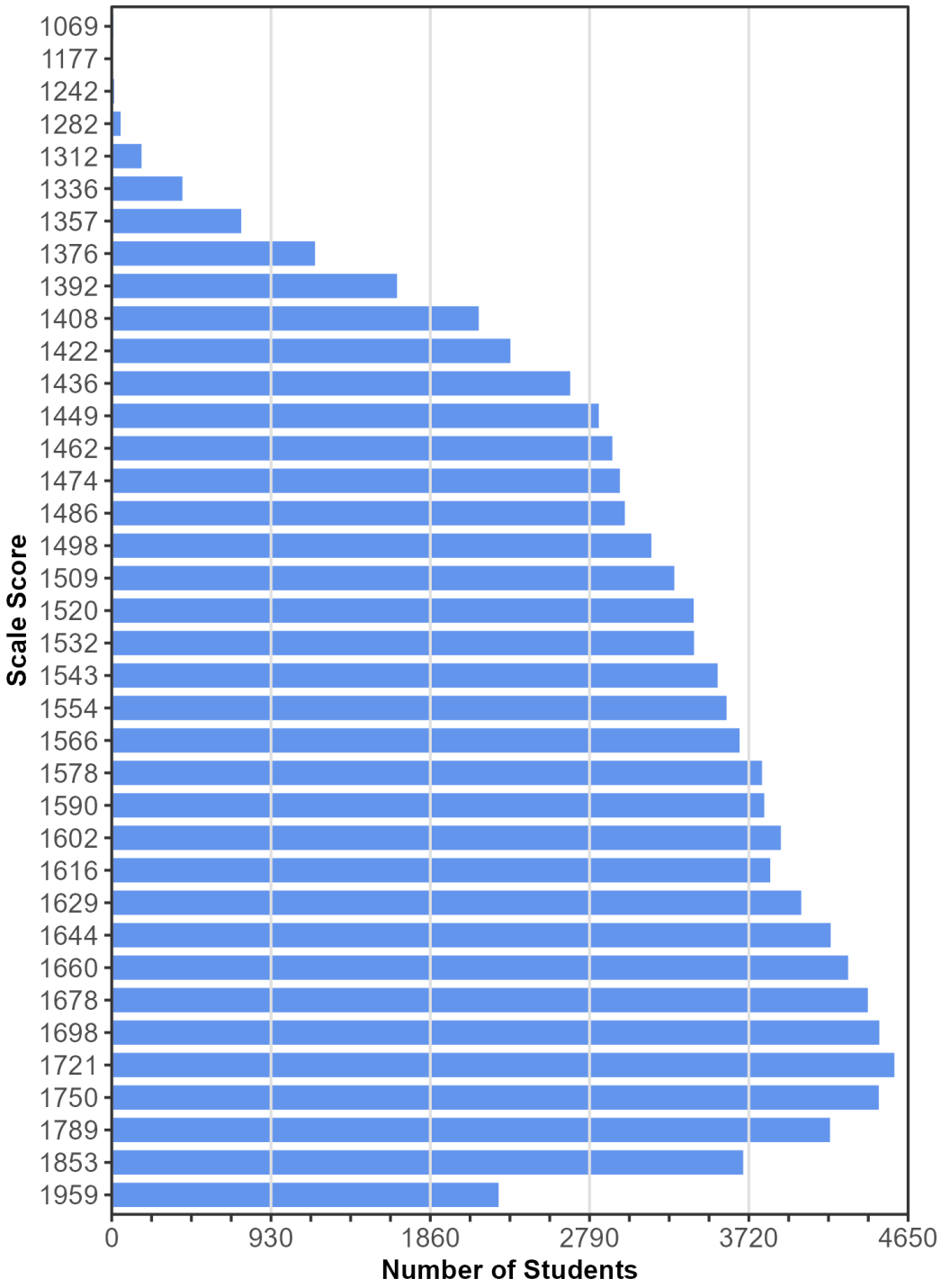
	Freq	Cum Freq	Pct	Cum Pct
1058	23	23	0.02	0.02
1167	16	39	0.02	0.04
1232	59	98	0.06	0.10
1272	239	337	0.24	0.33
1302	656	993	0.65	0.99
1326	1301	2294	1.29	2.28
1346	2346	4640	2.33	4.61
1364	3367	8007	3.35	7.96
1381	4317	12324	4.29	12.25
1396	4761	17085	4.73	16.98
1410	4845	21930	4.81	21.79
1423	4481	26411	4.45	26.25
1436	4243	30654	4.22	30.46
1448	3758	34412	3.73	34.20
1460	3548	37960	3.53	37.72
1472	3308	41268	3.29	41.01
1484	3284	44552	3.26	44.27
1495	3321	47873	3.30	47.57
1507	3295	51168	3.27	50.85
1518	3372	54540	3.35	54.20
1530	3487	58027	3.47	57.66
1542	3531	61558	3.51	61.17
1554	3674	65232	3.65	64.82
1566	3686	68918	3.66	68.49
1580	3692	72610	3.67	72.15
1594	3543	76153	3.52	75.68
1608	3790	79943	3.77	79.44
1624	3463	83406	3.44	82.88
1642	3351	86757	3.33	86.21
1662	3337	90094	3.32	89.53
1686	3086	93180	3.07	92.60
1715	2861	96041	2.84	95.44
1755	2385	98426	2.37	97.81
1820	1587	100013	1.58	99.39
1928	618	100631	0.61	100.00

Figure D.5.3. Spring 2023 TELPAS Grade 4 Reading
Frequency Distribution of Scale Scores
(Online Only)



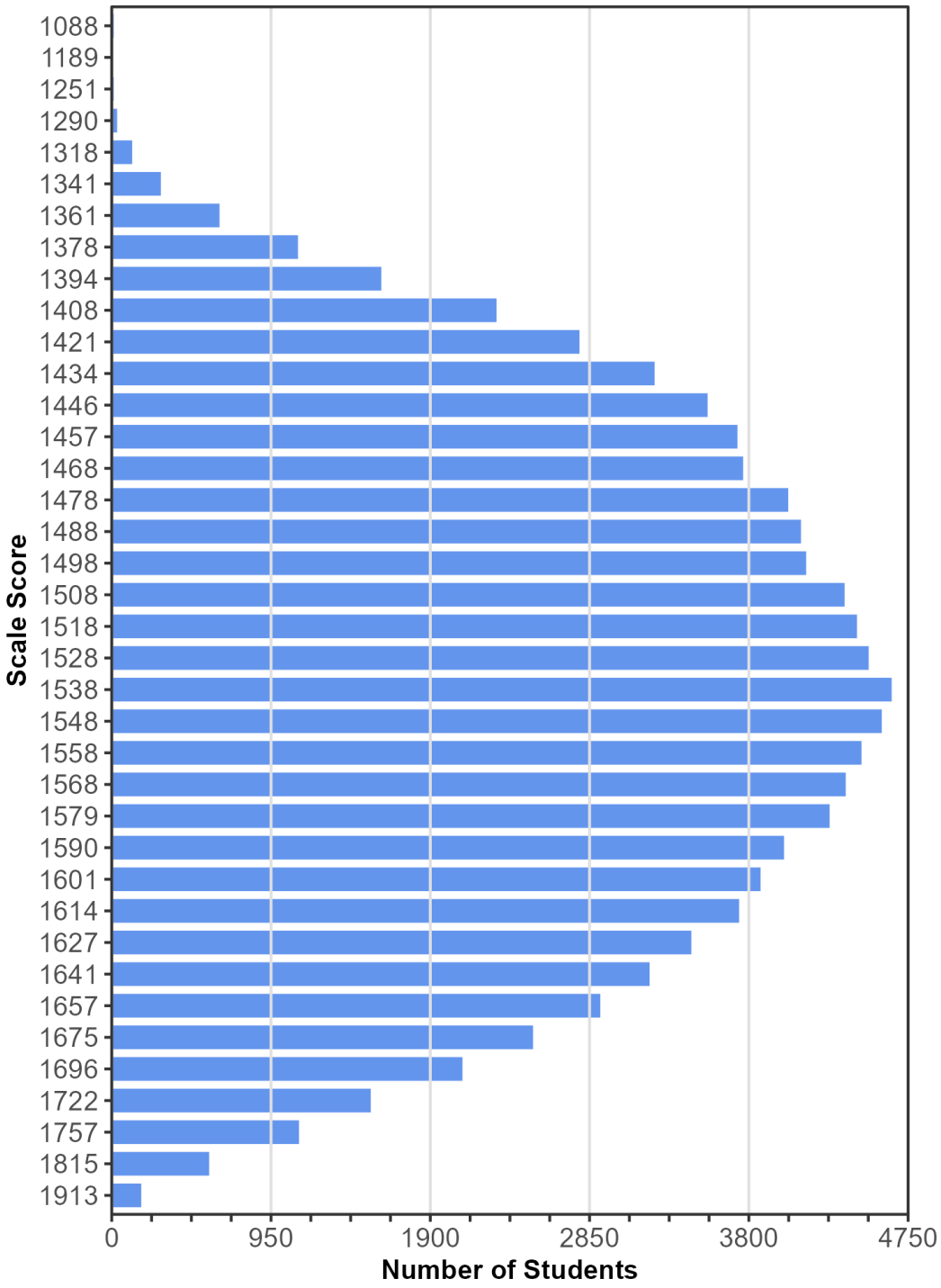
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1069	12	12	0.01	0.01
1177	5	17	0.00	0.02
1242	26	43	0.03	0.04
1282	103	146	0.10	0.14
1312	303	449	0.30	0.44
1336	682	1131	0.67	1.11
1357	1246	2377	1.23	2.34
1376	1935	4312	1.90	4.24
1392	2708	7020	2.66	6.90
1408	3346	10366	3.29	10.20
1422	3759	14125	3.70	13.89
1436	3924	18049	3.86	17.75
1449	4101	22150	4.03	21.79
1462	4100	26250	4.03	25.82
1474	4106	30356	4.04	29.86
1486	4218	34574	4.15	34.01
1498	4235	38809	4.17	38.17
1509	4154	42963	4.09	42.26
1520	3912	46875	3.85	46.11
1532	3873	50748	3.81	49.92
1543	3919	54667	3.85	53.77
1554	3705	58372	3.64	57.41
1566	3755	62127	3.69	61.11
1578	3644	65771	3.58	64.69
1590	3486	69257	3.43	68.12
1602	3482	72739	3.42	71.55
1616	3516	76255	3.46	75.00
1629	3447	79702	3.39	78.40
1644	3319	83021	3.26	81.66
1660	3140	86161	3.09	84.75
1678	3026	89187	2.98	87.72
1698	2859	92046	2.81	90.54
1721	2679	94725	2.64	93.17
1750	2451	97176	2.41	95.58
1789	2151	99327	2.12	97.70
1853	1552	100879	1.53	99.22
1959	788	101667	0.78	100.00

Figure D.5.4. Spring 2023 TELPAS Grade 5 Reading
Frequency Distribution of Scale Scores
(Online Only)



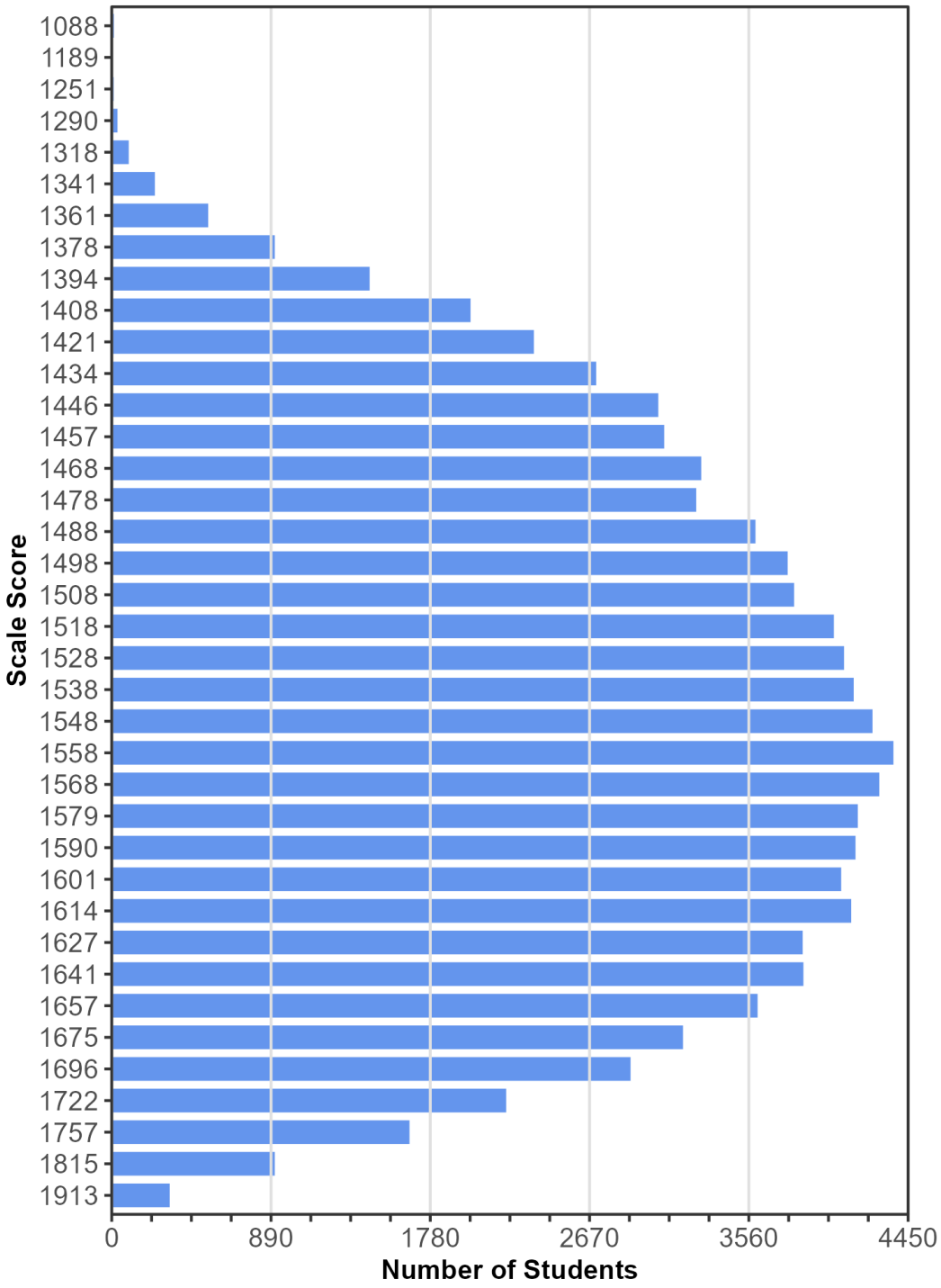
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1069	10	10	0.01	0.01
1177	6	16	0.01	0.02
1242	14	30	0.01	0.03
1282	52	82	0.05	0.08
1312	174	256	0.17	0.25
1336	413	669	0.40	0.65
1357	756	1425	0.73	1.38
1376	1187	2612	1.15	2.53
1392	1666	4278	1.62	4.15
1408	2143	6421	2.08	6.22
1422	2328	8749	2.26	8.48
1436	2677	11426	2.60	11.08
1449	2844	14270	2.76	13.83
1462	2923	17193	2.83	16.67
1474	2967	20160	2.88	19.54
1486	2996	23156	2.90	22.45
1498	3151	26307	3.05	25.50
1509	3285	29592	3.18	28.69
1520	3398	32990	3.29	31.98
1532	3400	36390	3.30	35.28
1543	3538	39928	3.43	38.71
1554	3590	43518	3.48	42.19
1566	3666	47184	3.55	45.74
1578	3797	50981	3.68	49.42
1590	3810	54791	3.69	53.12
1602	3907	58698	3.79	56.90
1616	3845	62543	3.73	60.63
1629	4026	66569	3.90	64.53
1644	4198	70767	4.07	68.60
1660	4300	75067	4.17	72.77
1678	4415	79482	4.28	77.05
1698	4482	83964	4.35	81.40
1721	4570	88534	4.43	85.83
1750	4479	93013	4.34	90.17
1789	4194	97207	4.07	94.24
1853	3687	100894	3.57	97.81
1959	2259	103153	2.19	100.00

Figure D.5.5. Spring 2023 TELPAS Grade 6 Reading
Frequency Distribution of Scale Scores
(Online Only)



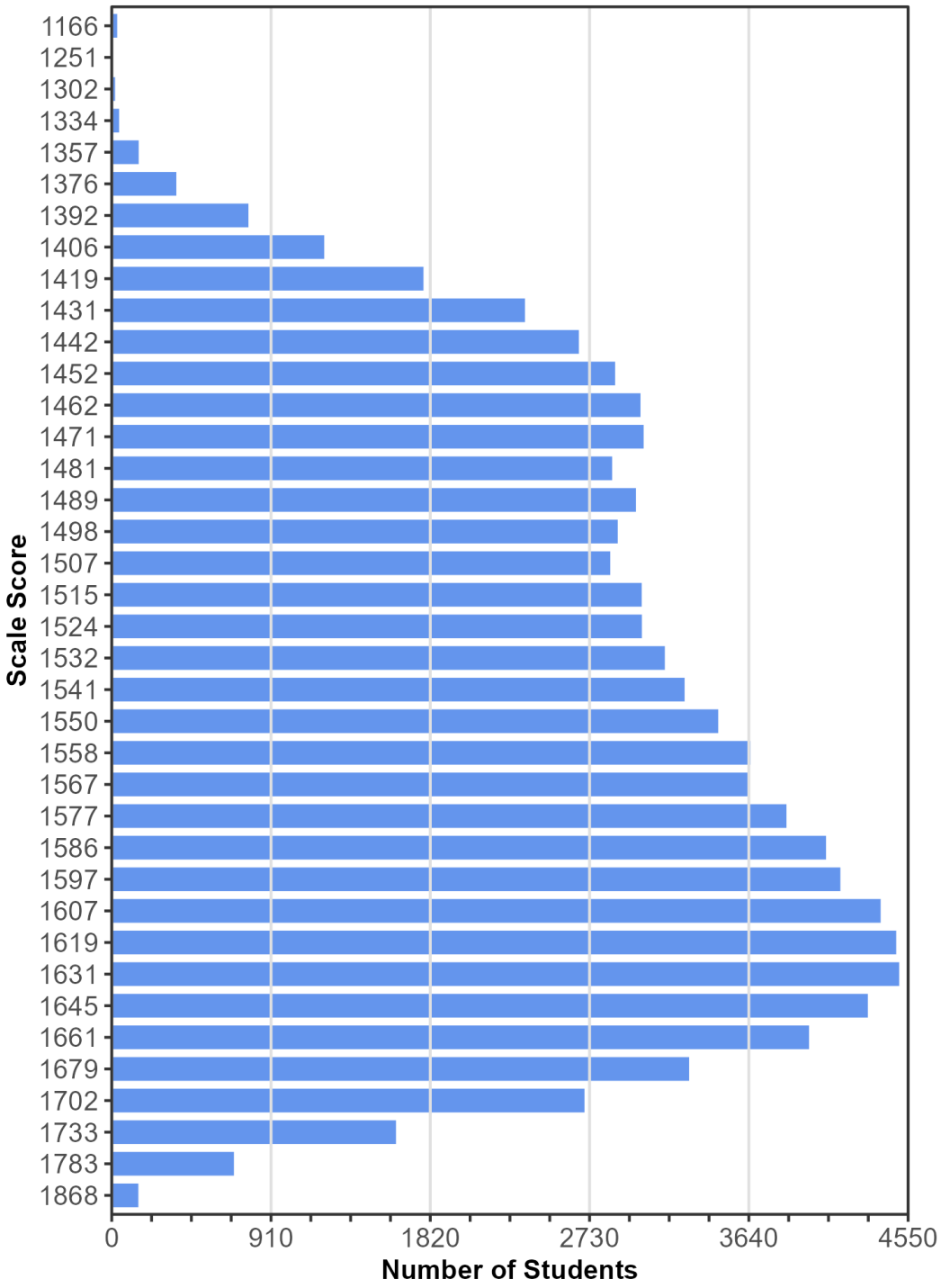
Freq	Cum Freq	Pct	Cum Pct
12	12	0.01	0.01
3	15	0.00	0.01
12	27	0.01	0.03
32	59	0.03	0.06
122	181	0.12	0.18
293	474	0.29	0.47
643	1117	0.64	1.11
1111	2228	1.11	2.22
1608	3836	1.60	3.82
2295	6131	2.29	6.10
2790	8921	2.78	8.88
3238	12159	3.22	12.11
3554	15713	3.54	15.65
3732	19445	3.72	19.36
3766	23211	3.75	23.11
4035	27246	4.02	27.13
4111	31357	4.09	31.22
4142	35499	4.12	35.35
4371	39870	4.35	39.70
4445	44315	4.43	44.12
4515	48830	4.50	48.62
4652	53482	4.63	53.25
4593	58075	4.57	57.83
4472	62547	4.45	62.28
4378	66925	4.36	66.64
4282	71207	4.26	70.90
4010	75217	3.99	74.89
3870	79087	3.85	78.75
3742	82829	3.73	82.47
3457	86286	3.44	85.91
3208	89494	3.19	89.11
2914	92408	2.90	92.01
2513	94921	2.50	94.51
2092	97013	2.08	96.60
1545	98558	1.54	98.13
1117	99675	1.11	99.25
581	100256	0.58	99.82
176	100432	0.18	100.00

Figure D.5.6. Spring 2023 TELPAS Grade 7 Reading
Frequency Distribution of Scale Scores
(Online Only)



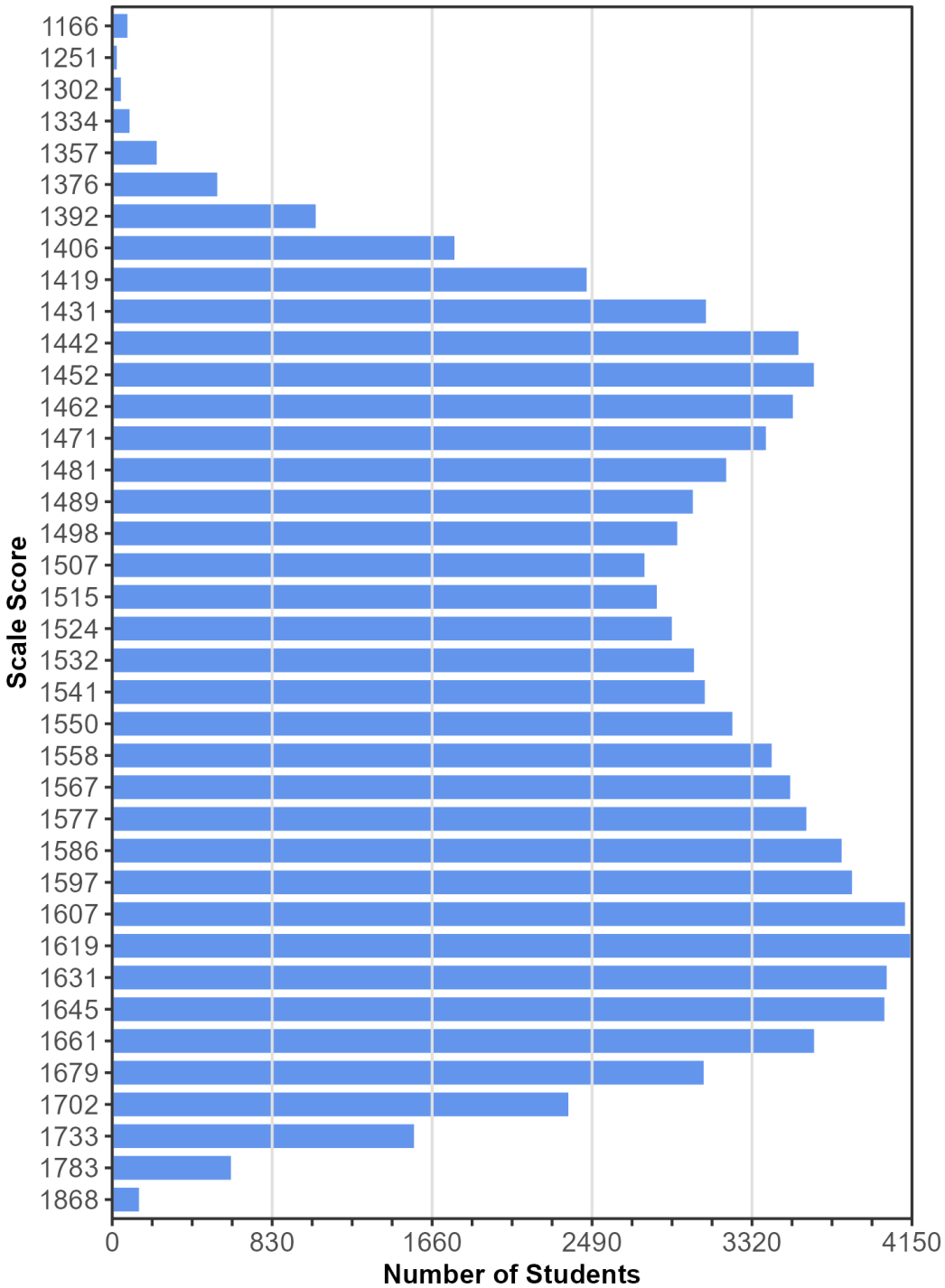
Freq	Cum Freq	Pct	Cum Pct
12	12	0.01	0.01
5	17	0.01	0.02
11	28	0.01	0.03
32	60	0.03	0.06
95	155	0.10	0.16
241	396	0.24	0.40
539	935	0.55	0.95
911	1846	0.92	1.87
1441	3287	1.46	3.34
2005	5292	2.04	5.37
2359	7651	2.40	7.77
2707	10358	2.75	10.52
3054	13412	3.10	13.62
3087	16499	3.13	16.75
3294	19793	3.34	20.10
3266	23059	3.32	23.41
3597	26656	3.65	27.06
3777	30433	3.83	30.90
3813	34246	3.87	34.77
4035	38281	4.10	38.87
4092	42373	4.15	43.02
4146	46519	4.21	47.23
4251	50770	4.32	51.55
4368	55138	4.44	55.98
4289	59427	4.35	60.34
4169	63596	4.23	64.57
4156	67752	4.22	68.79
4076	71828	4.14	72.93
4132	75960	4.20	77.13
3861	79821	3.92	81.05
3865	83686	3.92	84.97
3609	87295	3.66	88.63
3192	90487	3.24	91.88
2899	93386	2.94	94.82
2204	95590	2.24	97.06
1664	97254	1.69	98.75
911	98165	0.92	99.67
324	98489	0.33	100.00

Figure D.5.7. Spring 2023 TELPAS Grade 8 Reading
Frequency Distribution of Scale Scores
(Online Only)



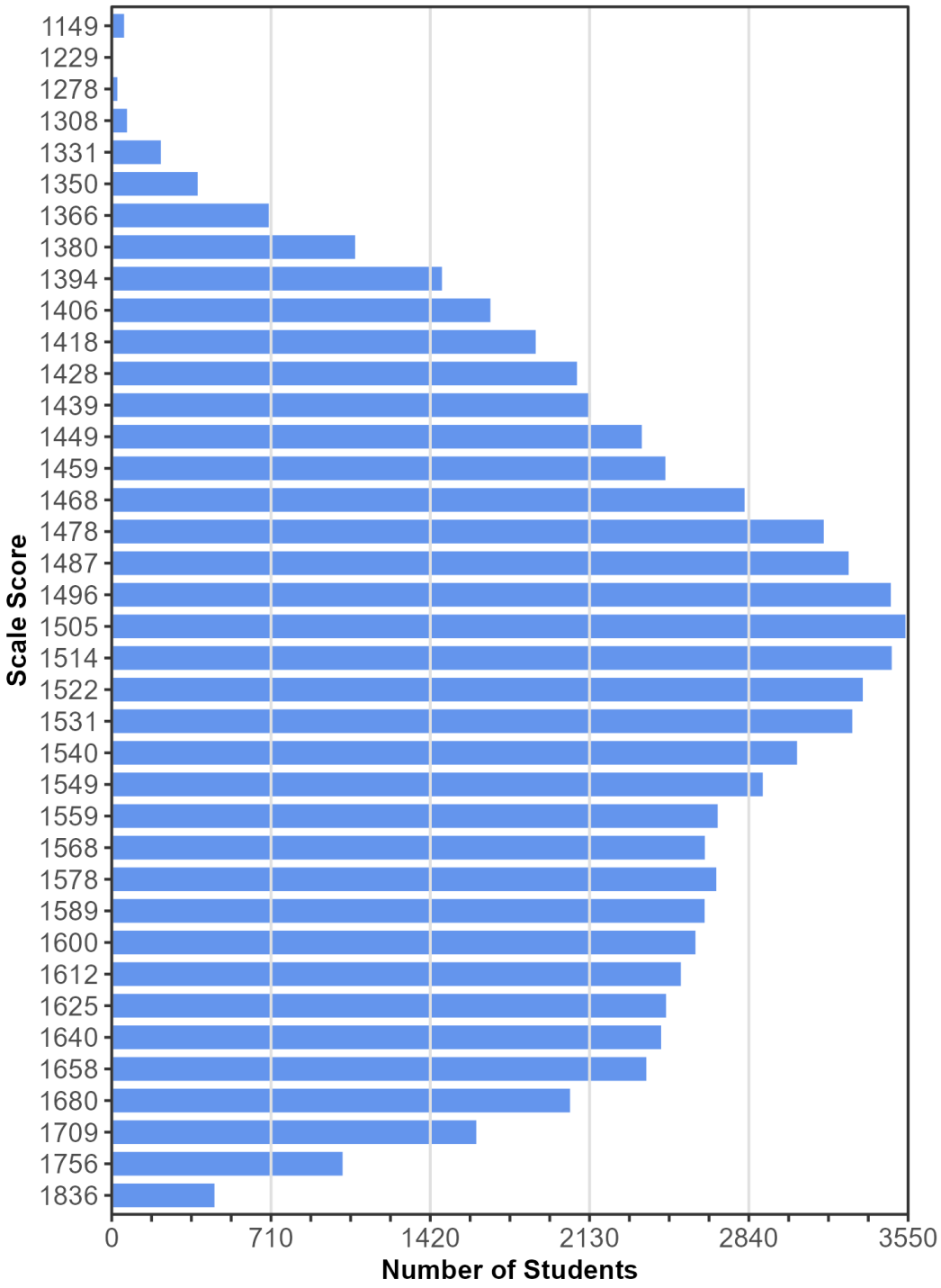
Freq	Cum Freq	Pct	Cum Pct
31	31	0.03	0.03
6	37	0.01	0.04
19	56	0.02	0.06
42	98	0.04	0.10
154	252	0.16	0.26
369	621	0.39	0.65
781	1402	0.82	1.47
1214	2616	1.27	2.74
1781	4397	1.87	4.61
2361	6758	2.47	7.08
2669	9427	2.80	9.88
2876	12303	3.01	12.89
3021	15324	3.17	16.05
3039	18363	3.18	19.24
2859	21222	3.00	22.23
2995	24217	3.14	25.37
2891	27108	3.03	28.40
2848	29956	2.98	31.38
3028	32984	3.17	34.56
3029	36013	3.17	37.73
3160	39173	3.31	41.04
3273	42446	3.43	44.47
3465	45911	3.63	48.10
3649	49560	3.82	51.92
3639	53199	3.81	55.73
3855	57054	4.04	59.77
4081	61135	4.28	64.05
4163	65298	4.36	68.41
4393	69691	4.60	73.01
4482	74173	4.70	77.71
4499	78672	4.71	82.42
4320	82992	4.53	86.95
3984	86976	4.17	91.12
3299	90275	3.46	94.58
2701	92976	2.83	97.41
1624	94600	1.70	99.11
698	95298	0.73	99.84
152	95450	0.16	100.00

Figure D.5.8. Spring 2023 TELPAS Grade 9 Reading
Frequency Distribution of Scale Scores
(Online Only)



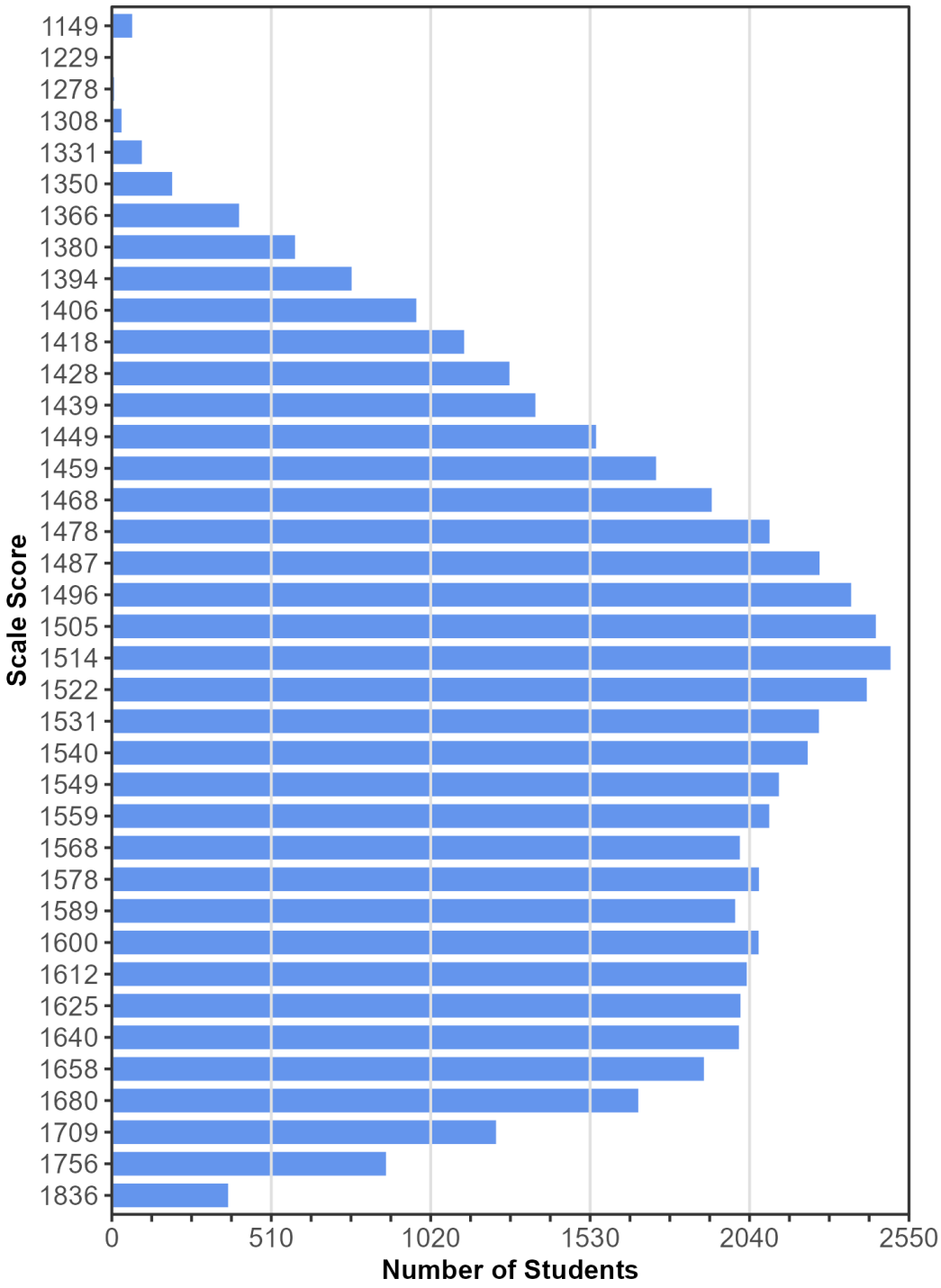
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1166	79	79	0.08	0.08
1251	24	103	0.02	0.11
1302	45	148	0.05	0.15
1334	90	238	0.09	0.25
1357	231	469	0.24	0.49
1376	545	1014	0.57	1.05
1392	1056	2070	1.10	2.15
1406	1776	3846	1.84	3.99
1419	2462	6308	2.56	6.55
1431	3081	9389	3.20	9.75
1442	3561	12950	3.70	13.45
1452	3641	16591	3.78	17.23
1462	3532	20123	3.67	20.90
1471	3392	23515	3.52	24.42
1481	3186	26701	3.31	27.73
1489	3013	29714	3.13	30.86
1498	2932	32646	3.04	33.90
1507	2762	35408	2.87	36.77
1515	2826	38234	2.93	39.70
1524	2904	41138	3.02	42.72
1532	3019	44157	3.14	45.86
1541	3075	47232	3.19	49.05
1550	3218	50450	3.34	52.39
1558	3422	53872	3.55	55.94
1567	3518	57390	3.65	59.60
1577	3602	60992	3.74	63.34
1586	3785	64777	3.93	67.27
1597	3839	68616	3.99	71.26
1607	4114	72730	4.27	75.53
1619	4141	76871	4.30	79.83
1631	4019	80890	4.17	84.00
1645	4007	84897	4.16	88.16
1661	3642	88539	3.78	91.94
1679	3069	91608	3.19	95.13
1702	2367	93975	2.46	97.59
1733	1566	95541	1.63	99.22
1783	616	96157	0.64	99.86
1868	139	96296	0.14	100.00

Figure D.5.9. Spring 2023 TELPAS Grade 10 Reading
Frequency Distribution of Scale Scores
(Online Only)



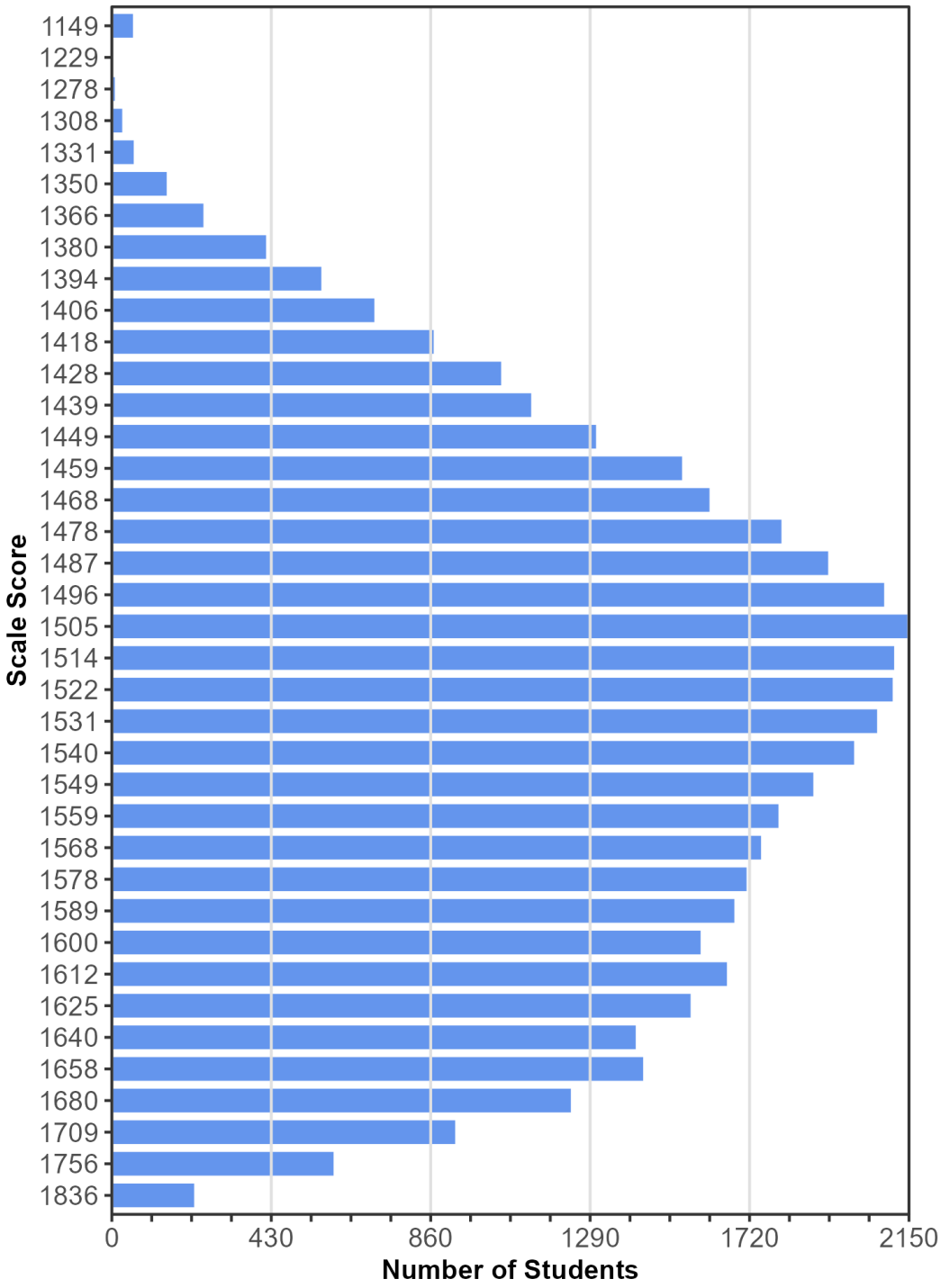
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1149	55	55	0.07	0.07
1229	5	60	0.01	0.08
1278	25	85	0.03	0.11
1308	68	153	0.09	0.20
1331	219	372	0.28	0.48
1350	383	755	0.50	0.98
1366	700	1455	0.91	1.88
1380	1085	2540	1.40	3.29
1394	1472	4012	1.90	5.19
1406	1688	5700	2.18	7.38
1418	1890	7590	2.45	9.82
1428	2074	9664	2.68	12.50
1439	2135	11799	2.76	15.27
1449	2363	14162	3.06	18.32
1459	2468	16630	3.19	21.52
1468	2821	19451	3.65	25.17
1478	3174	22625	4.11	29.28
1487	3285	25910	4.25	33.53
1496	3473	29383	4.49	38.02
1505	3538	32921	4.58	42.60
1514	3477	36398	4.50	47.10
1522	3348	39746	4.33	51.43
1531	3301	43047	4.27	55.70
1540	3055	46102	3.95	59.65
1549	2902	49004	3.75	63.41
1559	2701	51705	3.49	66.90
1568	2644	54349	3.42	70.32
1578	2695	57044	3.49	73.81
1589	2643	59687	3.42	77.23
1600	2602	62289	3.37	80.60
1612	2537	64826	3.28	83.88
1625	2471	67297	3.20	87.08
1640	2449	69746	3.17	90.25
1658	2383	72129	3.08	93.33
1680	2043	74172	2.64	95.97
1709	1625	75797	2.10	98.08
1756	1029	76826	1.33	99.41
1836	458	77284	0.59	100.00

Figure D.5.10. Spring 2023 TELPAS Grade 11 Reading
Frequency Distribution of Scale Scores
(Online Only)



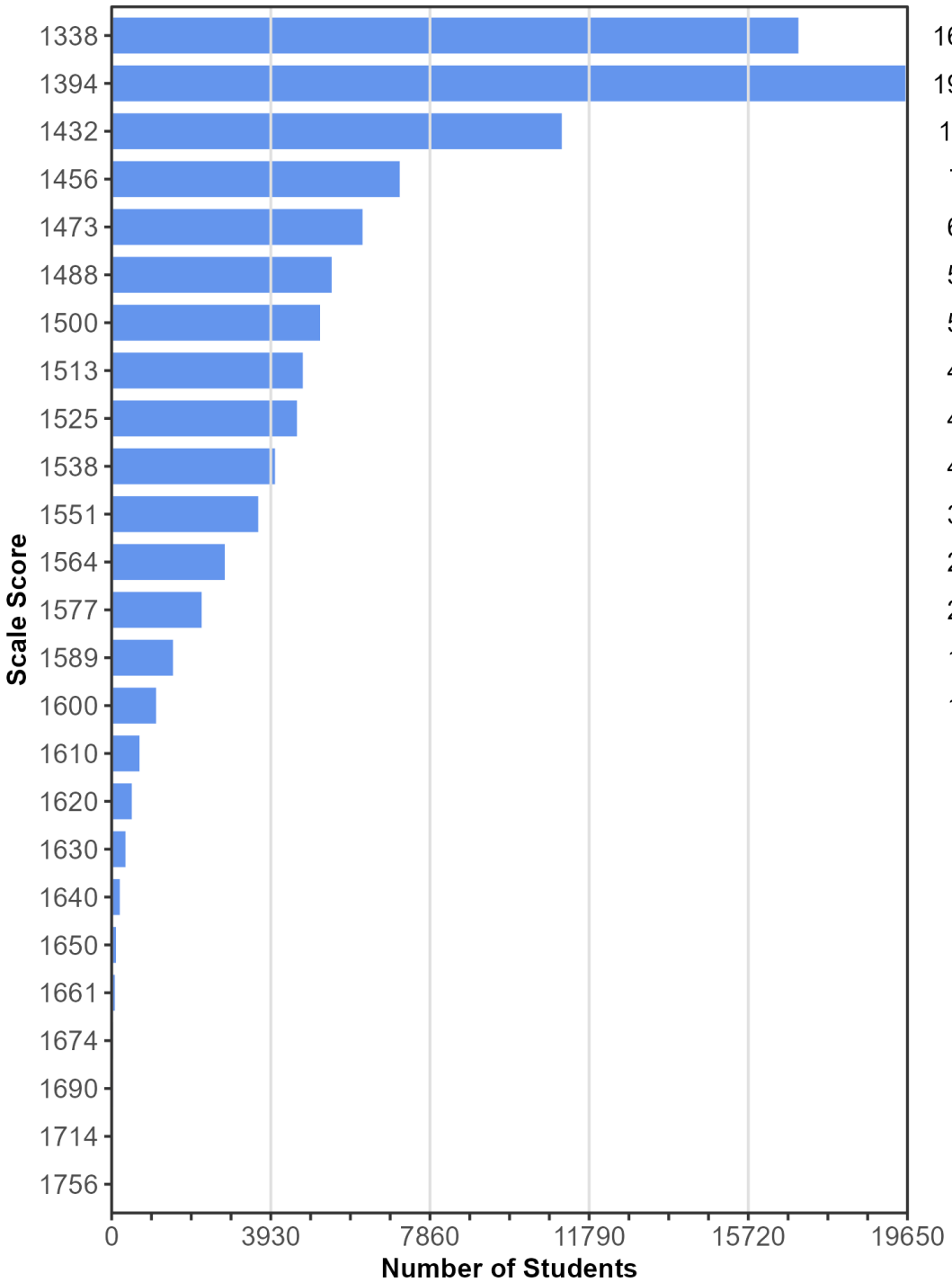
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1149	65	65	0.12	0.12
1229	2	67	0.00	0.12
1278	7	74	0.01	0.13
1308	31	105	0.06	0.19
1331	96	201	0.17	0.36
1350	193	394	0.35	0.71
1366	407	801	0.74	1.45
1380	586	1387	1.06	2.52
1394	767	2154	1.39	3.91
1406	974	3128	1.77	5.67
1418	1127	4255	2.04	7.72
1428	1272	5527	2.31	10.02
1439	1355	6882	2.46	12.48
1449	1549	8431	2.81	15.29
1459	1741	10172	3.16	18.45
1468	1919	12091	3.48	21.93
1478	2104	14195	3.82	25.74
1487	2264	16459	4.11	29.85
1496	2365	18824	4.29	34.14
1505	2444	21268	4.43	38.57
1514	2491	23759	4.52	43.08
1522	2415	26174	4.38	47.46
1531	2262	28436	4.10	51.57
1540	2226	30662	4.04	55.60
1549	2134	32796	3.87	59.47
1559	2103	34899	3.81	63.29
1568	2009	36908	3.64	66.93
1578	2070	38978	3.75	70.68
1589	1994	40972	3.62	74.30
1600	2069	43041	3.75	78.05
1612	2031	45072	3.68	81.73
1625	2011	47083	3.65	85.38
1640	2006	49089	3.64	89.02
1658	1894	50983	3.43	92.45
1680	1684	52667	3.05	95.51
1709	1229	53896	2.23	97.74
1756	877	54773	1.59	99.33
1836	372	55145	0.67	100.00

Figure D.5.11. Spring 2023 TELPAS Grade 12 Reading
Frequency Distribution of Scale Scores
(Online Only)



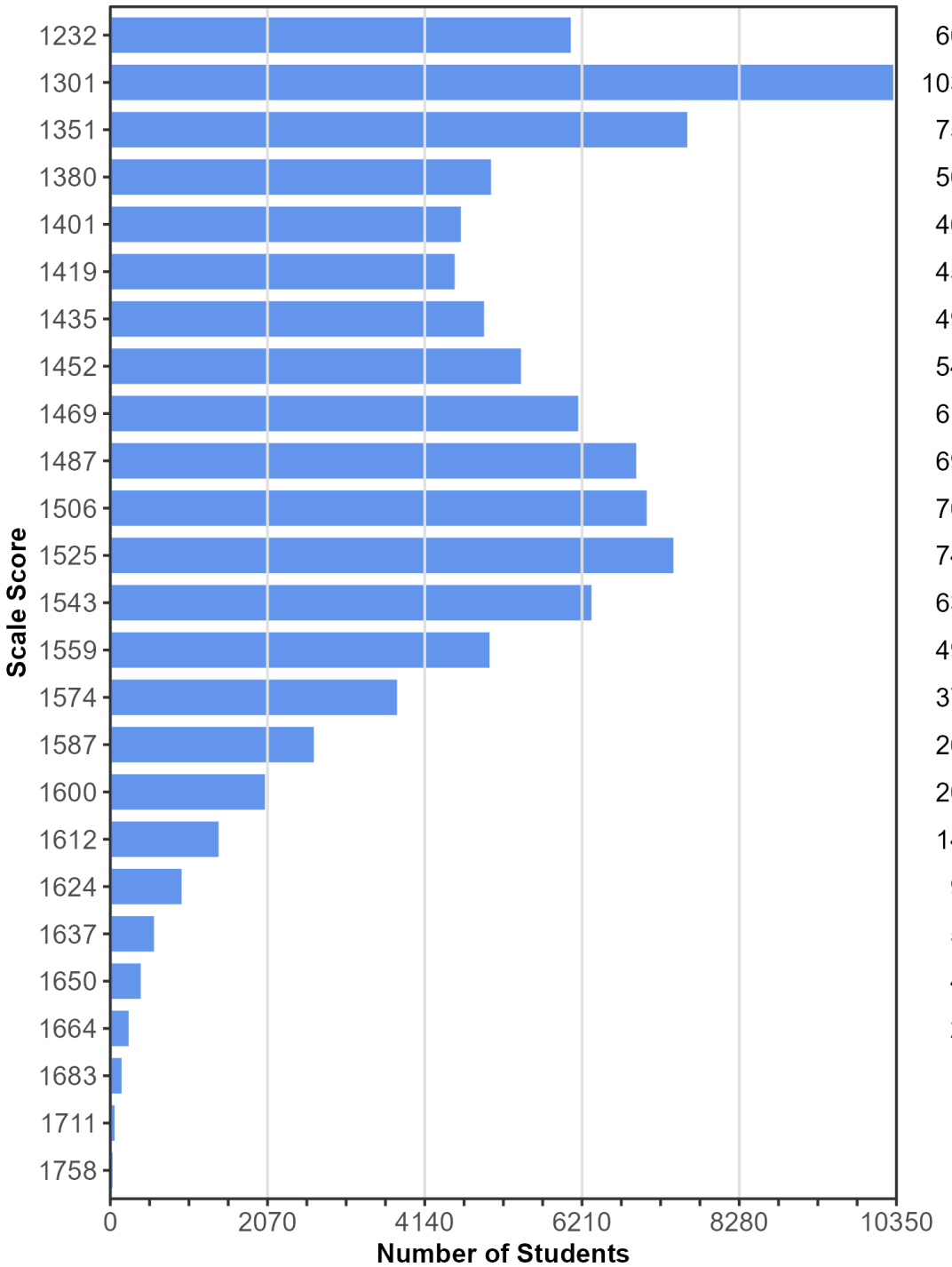
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1149	57	57	0.13	0.13
1229	1	58	0.00	0.13
1278	8	66	0.02	0.15
1308	28	94	0.06	0.21
1331	59	153	0.13	0.34
1350	148	301	0.33	0.66
1366	247	548	0.54	1.21
1380	416	964	0.92	2.12
1394	565	1529	1.24	3.36
1406	708	2237	1.56	4.92
1418	868	3105	1.91	6.83
1428	1050	4155	2.31	9.14
1439	1131	5286	2.49	11.63
1449	1306	6592	2.87	14.50
1459	1538	8130	3.38	17.88
1468	1612	9742	3.55	21.43
1478	1806	11548	3.97	25.40
1487	1932	13480	4.25	29.65
1496	2083	15563	4.58	34.23
1505	2146	17709	4.72	38.95
1514	2110	19819	4.64	43.60
1522	2106	21925	4.63	48.23
1531	2064	23989	4.54	52.77
1540	2002	25991	4.40	57.17
1549	1892	27883	4.16	61.33
1559	1798	29681	3.96	65.29
1568	1751	31432	3.85	69.14
1578	1712	33144	3.77	72.91
1589	1679	34823	3.69	76.60
1600	1588	36411	3.49	80.09
1612	1659	38070	3.65	83.74
1625	1561	39631	3.43	87.18
1640	1413	41044	3.11	90.28
1658	1433	42477	3.15	93.44
1680	1238	43715	2.72	96.16
1709	926	44641	2.04	98.20
1756	598	45239	1.32	99.51
1836	222	45461	0.49	100.00

Figure D.5.12. Spring 2023 TELPAS Grade 2 Writing
Frequency Distribution of Scale Scores
(Online Only)



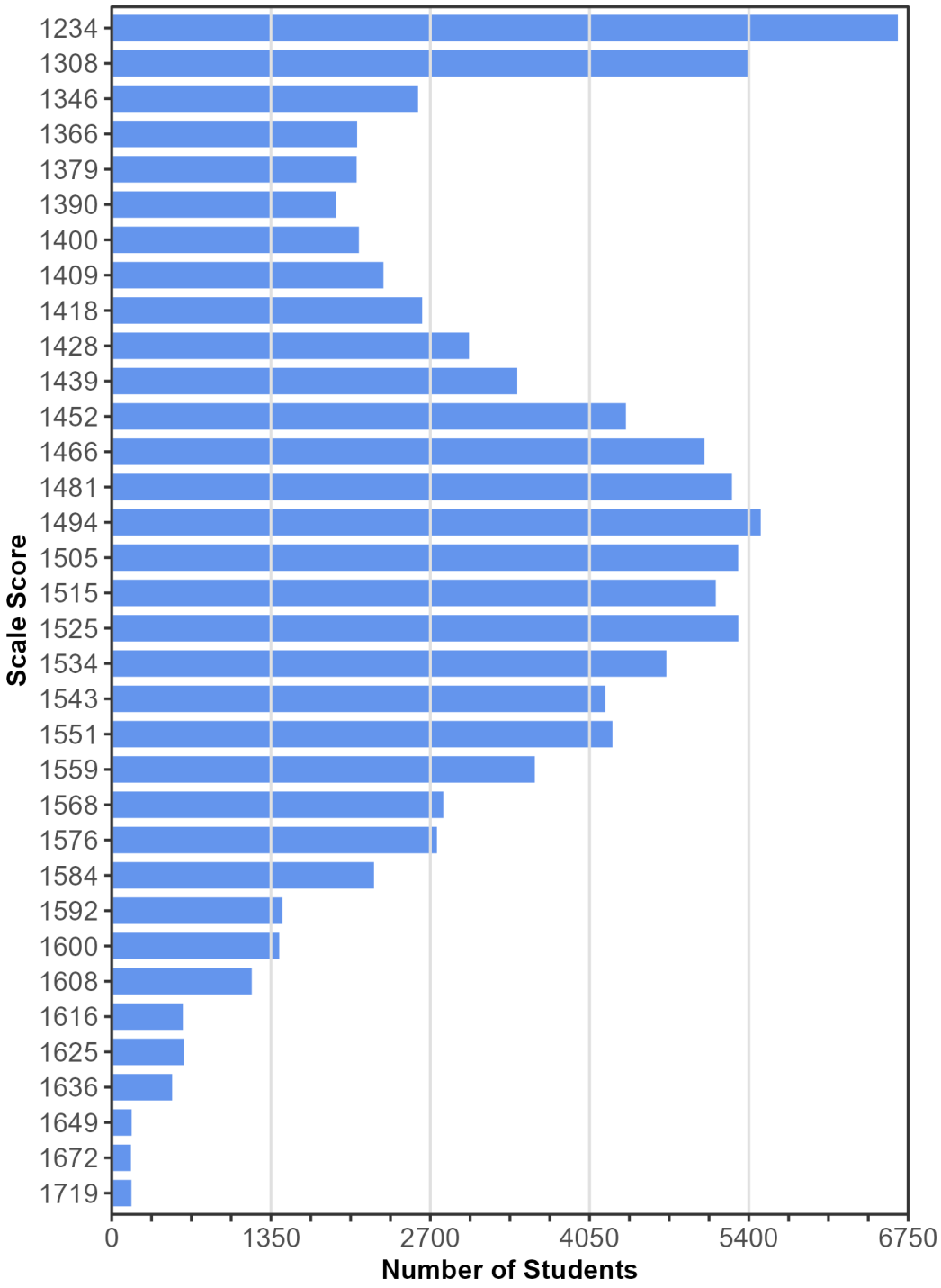
	Freq	Cum Freq	Pct	Cum Pct
1338	16958	16958	17.28	17.28
1394	19599	36557	19.98	37.26
1432	11115	47672	11.33	48.59
1456	7113	54785	7.25	55.84
1473	6195	60980	6.31	62.15
1488	5432	66412	5.54	67.69
1500	5146	71558	5.25	72.94
1513	4720	76278	4.81	77.75
1525	4576	80854	4.66	82.41
1538	4033	84887	4.11	86.52
1551	3619	88506	3.69	90.21
1564	2794	91300	2.85	93.06
1577	2222	93522	2.26	95.32
1589	1515	95037	1.54	96.87
1600	1097	96134	1.12	97.98
1610	687	96821	0.70	98.69
1620	497	97318	0.51	99.19
1630	341	97659	0.35	99.54
1640	201	97860	0.20	99.74
1650	107	97967	0.11	99.85
1661	78	98045	0.08	99.93
1674	31	98076	0.03	99.96
1690	28	98104	0.03	99.99
1714	4	98108	0.00	100.00
1756	3	98111	0.00	100.00

Figure D.5.13. Spring 2023 TELPAS Grade 3 Writing
Frequency Distribution of Scale Scores
(Online Only)



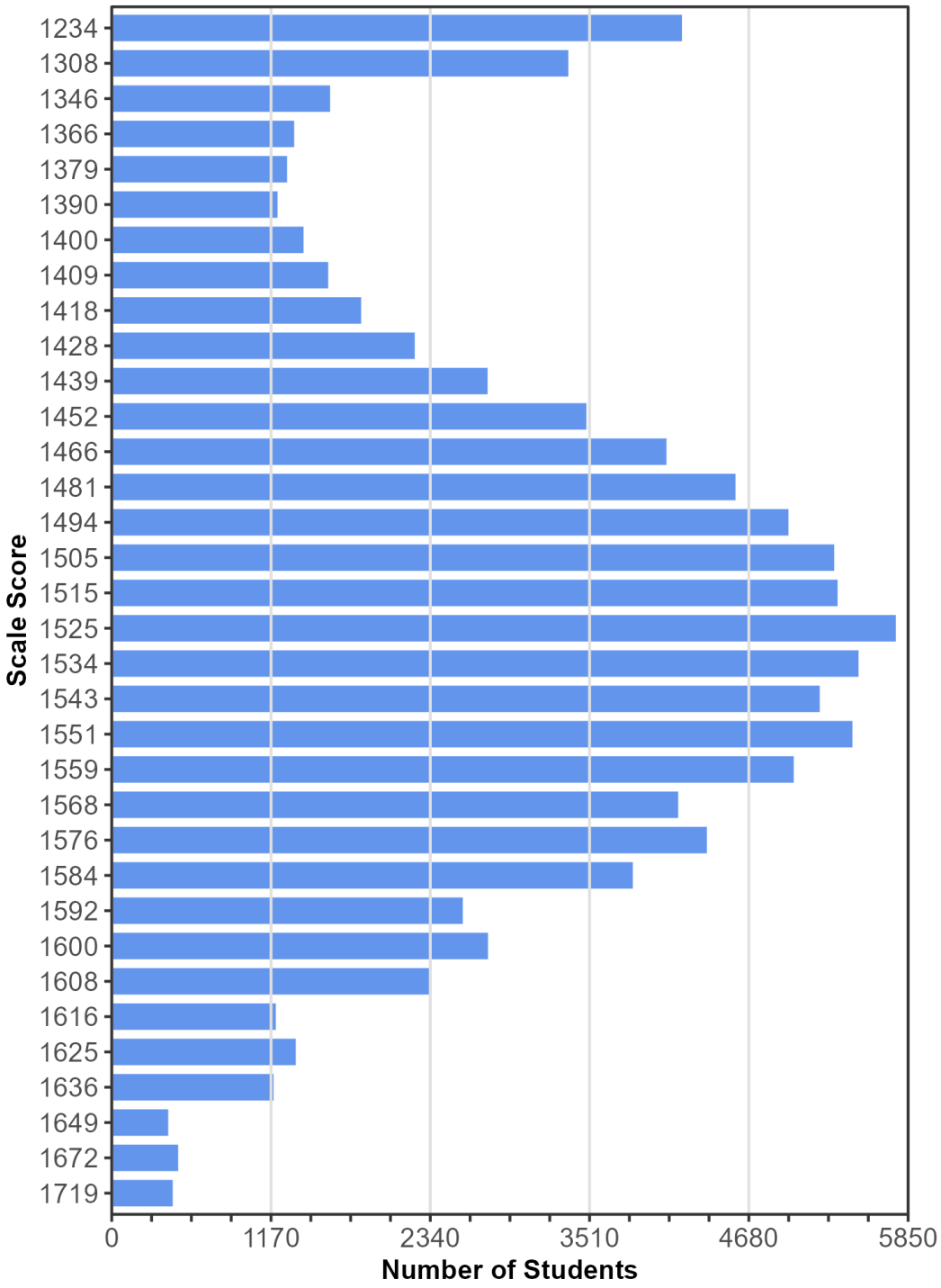
	Freq	Cum Freq	Pct	Cum Pct
1232	6062	6062	6.08	6.08
1301	10308	16370	10.34	16.43
1351	7596	23966	7.62	24.05
1380	5012	28978	5.03	29.08
1401	4615	33593	4.63	33.71
1419	4534	38127	4.55	38.26
1435	4920	43047	4.94	43.20
1452	5407	48454	5.43	48.62
1469	6161	54615	6.18	54.81
1487	6924	61539	6.95	61.76
1506	7063	68602	7.09	68.84
1525	7414	76016	7.44	76.28
1543	6335	82351	6.36	82.64
1559	4993	87344	5.01	87.65
1574	3775	91119	3.79	91.44
1587	2680	93799	2.69	94.13
1600	2036	95835	2.04	96.17
1612	1426	97261	1.43	97.60
1624	939	98200	0.94	98.54
1637	575	98775	0.58	99.12
1650	401	99176	0.40	99.52
1664	241	99417	0.24	99.77
1683	148	99565	0.15	99.91
1711	56	99621	0.06	99.97
1758	29	99650	0.03	100.00

Figure D.5.14. Spring 2023 TELPAS Grade 4 Writing
Frequency Distribution of Scale Scores
(Online Only)



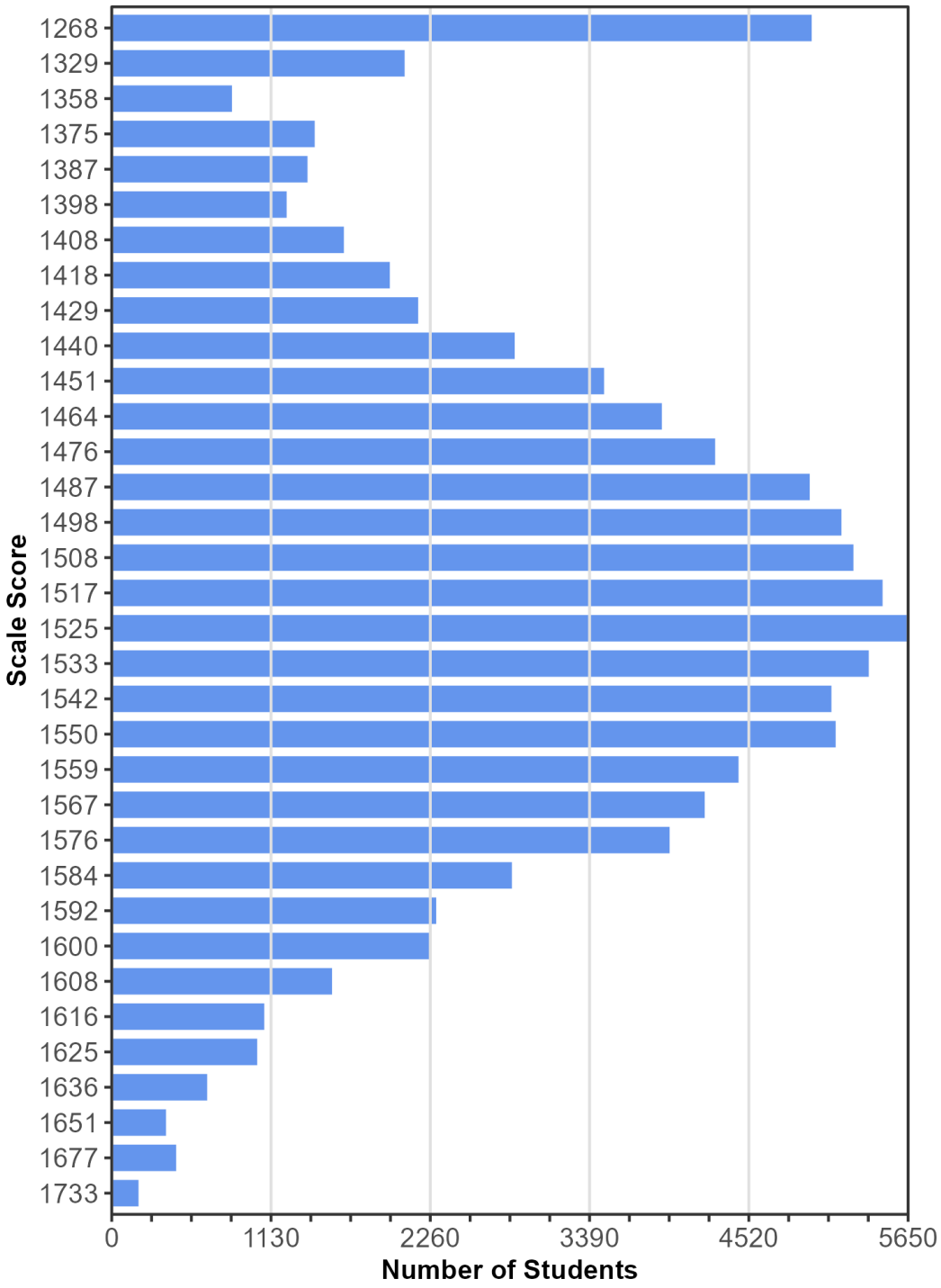
	Freq	Cum Freq	Pct	Cum Pct
1234	6663	6663	6.61	6.61
1308	5404	12067	5.36	11.96
1346	2596	14663	2.57	14.54
1366	2080	16743	2.06	16.60
1379	2076	18819	2.06	18.66
1390	1903	20722	1.89	20.54
1400	2095	22817	2.08	22.62
1409	2303	25120	2.28	24.90
1418	2631	27751	2.61	27.51
1428	3028	30779	3.00	30.51
1439	3437	34216	3.41	33.92
1452	4358	38574	4.32	38.24
1466	5023	43597	4.98	43.22
1481	5257	48854	5.21	48.43
1494	5501	54355	5.45	53.88
1505	5310	59665	5.26	59.15
1515	5120	64785	5.08	64.22
1525	5311	70096	5.26	69.49
1534	4701	74797	4.66	74.15
1543	4185	78982	4.15	78.29
1551	4244	83226	4.21	82.50
1559	3586	86812	3.55	86.06
1568	2810	89622	2.79	88.84
1576	2756	92378	2.73	91.57
1584	2223	94601	2.20	93.78
1592	1446	96047	1.43	95.21
1600	1420	97467	1.41	96.62
1608	1187	98654	1.18	97.80
1616	603	99257	0.60	98.39
1625	610	99867	0.60	99.00
1636	512	100379	0.51	99.51
1649	169	100548	0.17	99.67
1672	163	100711	0.16	99.83
1719	167	100878	0.17	100.00

Figure D.5.15. Spring 2023 TELPAS Grade 5 Writing
Frequency Distribution of Scale Scores
(Online Only)



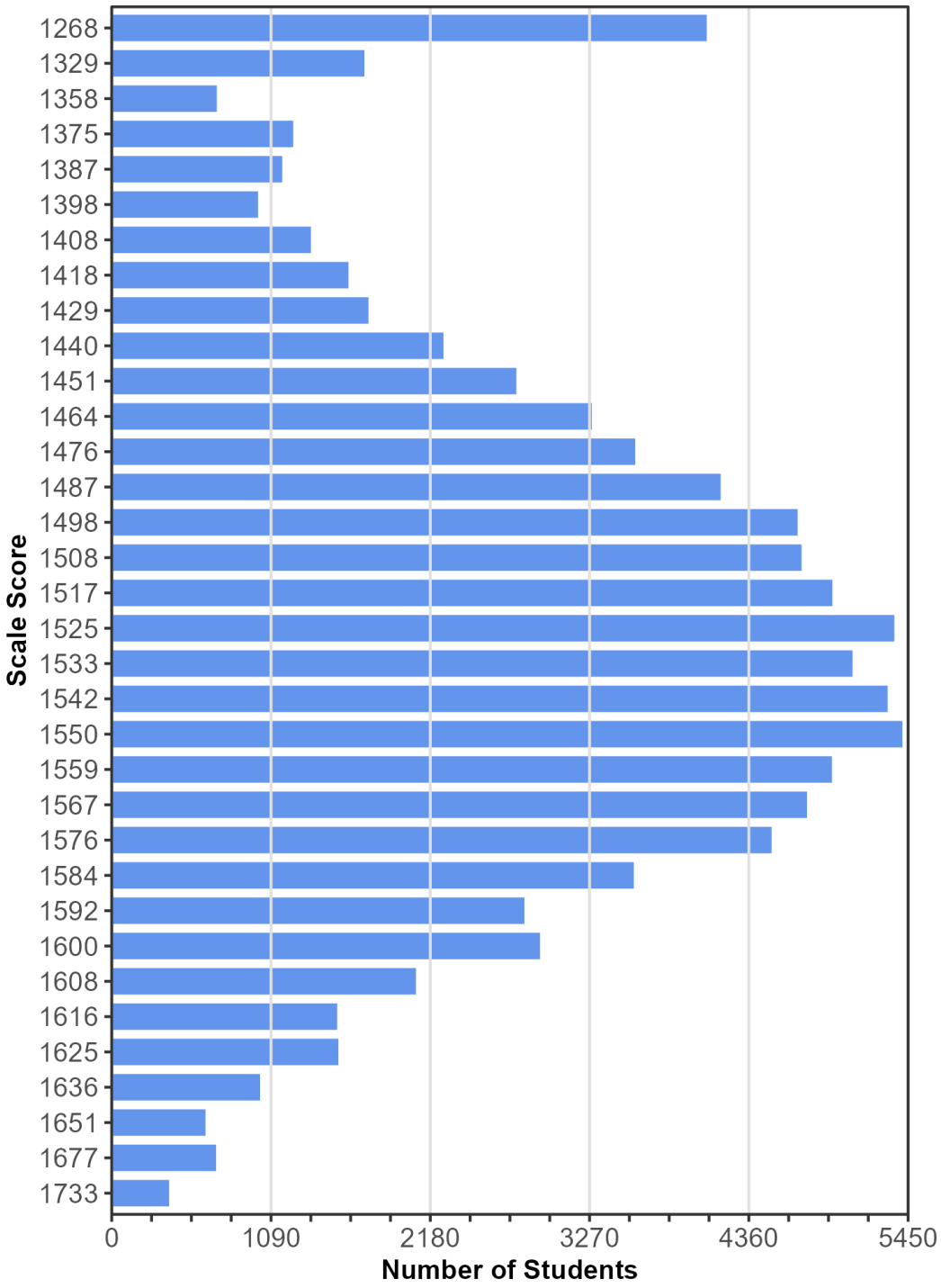
	Freq	Cum Freq	Pct	Cum Pct
1234	4189	4189	4.08	4.08
1308	3354	7543	3.27	7.35
1346	1604	9147	1.56	8.91
1366	1340	10487	1.31	10.22
1379	1288	11775	1.26	11.48
1390	1218	12993	1.19	12.66
1400	1409	14402	1.37	14.04
1409	1590	15992	1.55	15.58
1418	1832	17824	1.79	17.37
1428	2226	20050	2.17	19.54
1439	2761	22811	2.69	22.23
1452	3487	26298	3.40	25.63
1466	4075	30373	3.97	29.60
1481	4582	34955	4.47	34.06
1494	4971	39926	4.84	38.91
1505	5307	45233	5.17	44.08
1515	5332	50565	5.20	49.28
1525	5760	56325	5.61	54.89
1534	5486	61811	5.35	60.24
1543	5201	67012	5.07	65.30
1551	5441	72453	5.30	70.61
1559	5010	77463	4.88	75.49
1568	4161	81624	4.06	79.54
1576	4372	85996	4.26	83.81
1584	3828	89824	3.73	87.54
1592	2579	92403	2.51	90.05
1600	2764	95167	2.69	92.74
1608	2350	97517	2.29	95.03
1616	1205	98722	1.17	96.21
1625	1352	100074	1.32	97.52
1636	1189	101263	1.16	98.68
1649	415	101678	0.40	99.09
1672	488	102166	0.48	99.56
1719	448	102614	0.44	100.00

Figure D.5.16. Spring 2023 TELPAS Grade 6 Writing
Frequency Distribution of Scale Scores
(Online Only)



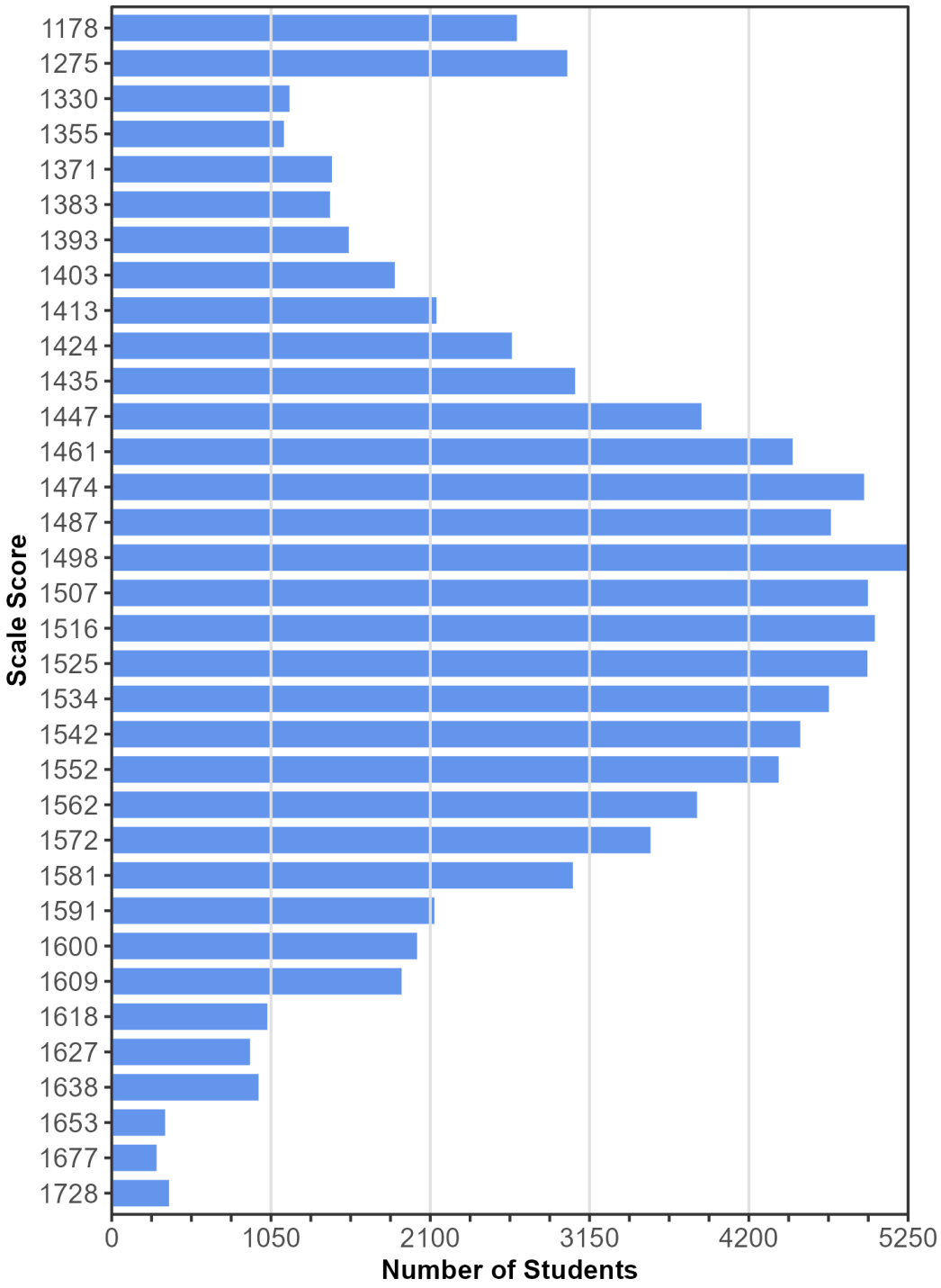
	Freq	Cum Freq	Pct	Cum Pct
1268	4966	4966	4.98	4.98
1329	2078	7044	2.08	7.06
1358	853	7897	0.85	7.91
1375	1440	9337	1.44	9.36
1387	1389	10726	1.39	10.75
1398	1241	11967	1.24	11.99
1408	1647	13614	1.65	13.64
1418	1973	15587	1.98	15.62
1429	2174	17761	2.18	17.80
1440	2859	20620	2.86	20.66
1451	3493	24113	3.50	24.16
1464	3903	28016	3.91	28.07
1476	4281	32297	4.29	32.36
1487	4952	37249	4.96	37.32
1498	5177	42426	5.19	42.51
1508	5262	47688	5.27	47.78
1517	5469	53157	5.48	53.26
1525	5641	58798	5.65	58.91
1533	5371	64169	5.38	64.29
1542	5106	69275	5.12	69.41
1550	5136	74411	5.15	74.56
1559	4447	78858	4.46	79.01
1567	4207	83065	4.22	83.23
1576	3958	87023	3.97	87.19
1584	2839	89862	2.84	90.04
1592	2302	92164	2.31	92.34
1600	2256	94420	2.26	94.60
1608	1563	95983	1.57	96.17
1616	1082	97065	1.08	97.25
1625	1032	98097	1.03	98.29
1636	677	98774	0.68	98.97
1651	385	99159	0.39	99.35
1677	457	99616	0.46	99.81
1733	190	99806	0.19	100.00

Figure D.5.17. Spring 2023 TELPAS Grade 7 Writing
Frequency Distribution of Scale Scores
(Online Only)



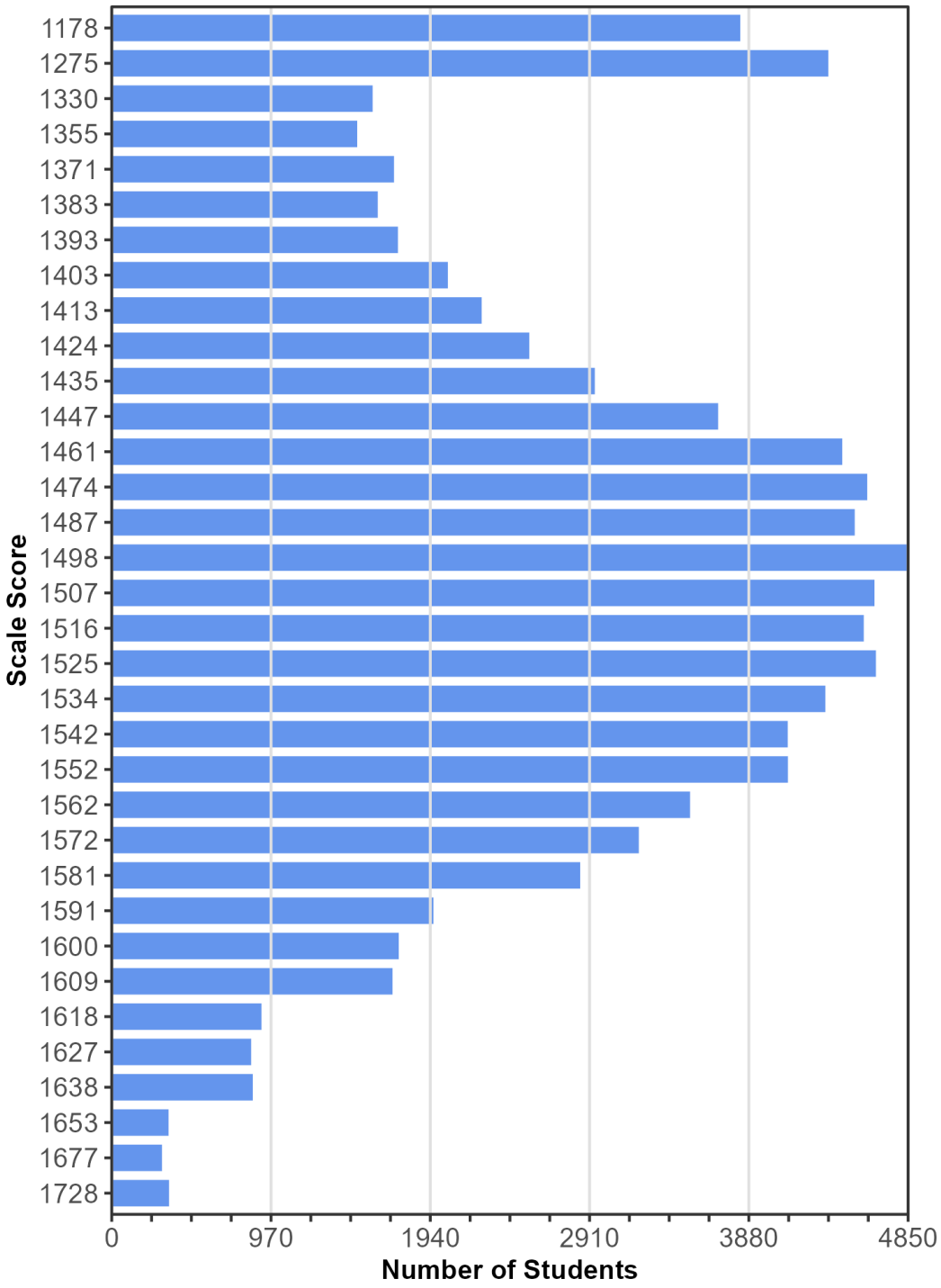
	Freq	Cum Freq	Pct	Cum Pct
1268	4072	4072	4.17	4.17
1329	1729	5801	1.77	5.94
1358	719	6520	0.74	6.67
1375	1242	7762	1.27	7.94
1387	1167	8929	1.19	9.14
1398	1001	9930	1.02	10.16
1408	1363	11293	1.39	11.56
1418	1620	12913	1.66	13.22
1429	1757	14670	1.80	15.01
1440	2270	16940	2.32	17.34
1451	2769	19709	2.83	20.17
1464	3286	22995	3.36	23.53
1476	3583	26578	3.67	27.20
1487	4167	30745	4.26	31.47
1498	4694	35439	4.80	36.27
1508	4721	40160	4.83	41.10
1517	4932	45092	5.05	46.15
1525	5356	50448	5.48	51.63
1533	5070	55518	5.19	56.82
1542	5310	60828	5.43	62.25
1550	5411	66239	5.54	67.79
1559	4929	71168	5.04	72.84
1567	4758	75926	4.87	77.71
1576	4516	80442	4.62	82.33
1584	3573	84015	3.66	85.98
1592	2824	86839	2.89	88.87
1600	2931	89770	3.00	91.87
1608	2082	91852	2.13	94.00
1616	1543	93395	1.58	95.58
1625	1551	94946	1.59	97.17
1636	1015	95961	1.04	98.21
1651	642	96603	0.66	98.87
1677	714	97317	0.73	99.60
1733	393	97710	0.40	100.00

Figure D.5.18. Spring 2023 TELPAS Grade 8 Writing
Frequency Distribution of Scale Scores
(Online Only)



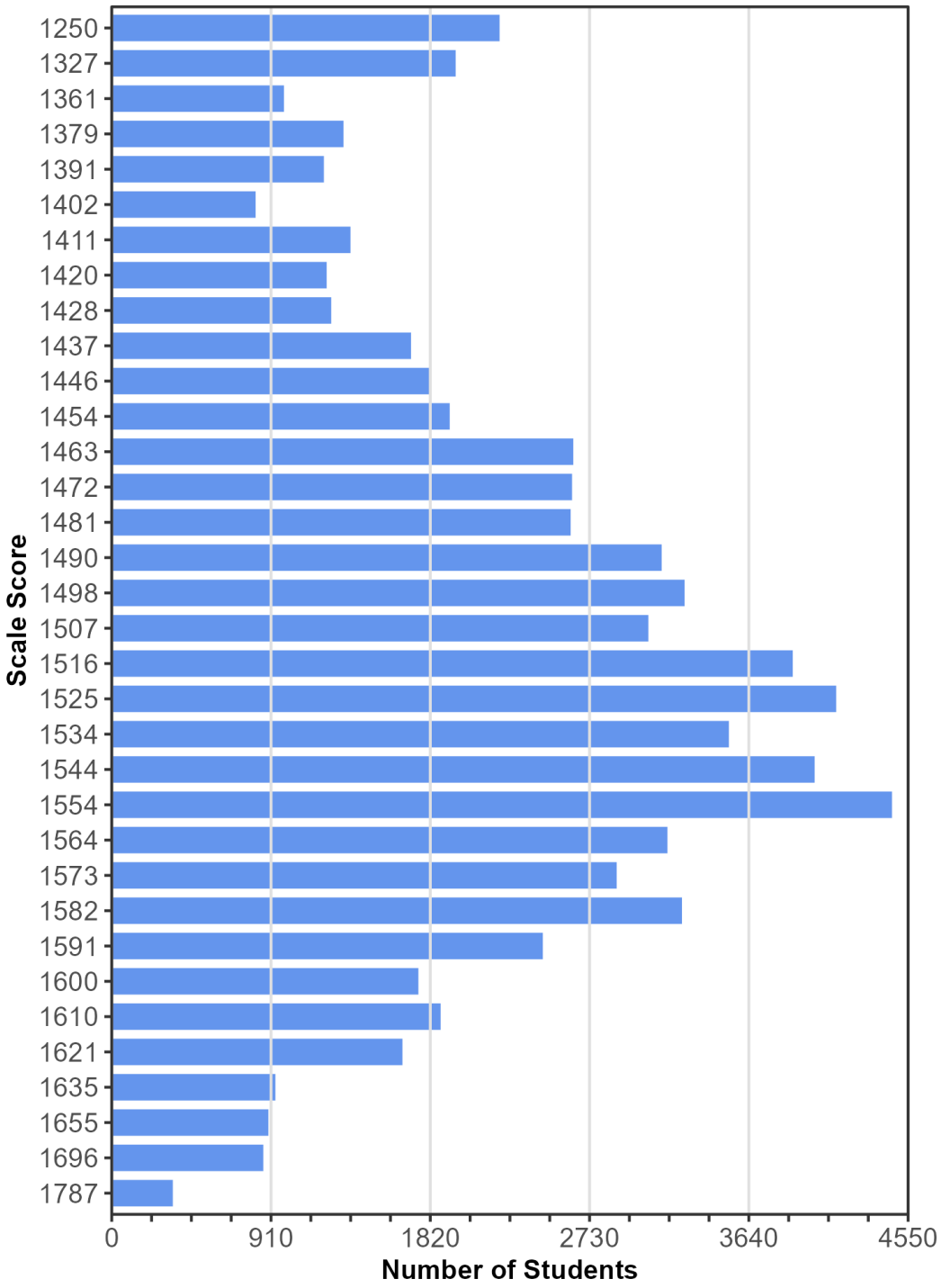
	Freq	Cum Freq	Pct	Cum Pct
1178	2671	2671	2.82	2.82
1275	3003	5674	3.18	6.00
1330	1172	6846	1.24	7.24
1355	1135	7981	1.20	8.44
1371	1452	9433	1.54	9.98
1383	1439	10872	1.52	11.50
1393	1563	12435	1.65	13.15
1403	1866	14301	1.97	15.12
1413	2141	16442	2.26	17.39
1424	2638	19080	2.79	20.18
1435	3055	22135	3.23	23.41
1447	3888	26023	4.11	27.52
1461	4489	30512	4.75	32.27
1474	4960	35472	5.25	37.52
1487	4741	40213	5.01	42.53
1498	5244	45457	5.55	48.08
1507	4985	50442	5.27	53.35
1516	5031	55473	5.32	58.67
1525	4982	60455	5.27	63.94
1534	4728	65183	5.00	68.94
1542	4539	69722	4.80	73.74
1552	4397	74119	4.65	78.39
1562	3859	77978	4.08	82.47
1572	3552	81530	3.76	86.23
1581	3040	84570	3.22	89.44
1591	2127	86697	2.25	91.69
1600	2013	88710	2.13	93.82
1609	1911	90621	2.02	95.84
1618	1026	91647	1.09	96.93
1627	912	92559	0.96	97.89
1638	968	93527	1.02	98.91
1653	352	93879	0.37	99.29
1677	296	94175	0.31	99.60
1728	378	94553	0.40	100.00

Figure D.5.19. Spring 2023 TELPAS Grade 9 Writing
Frequency Distribution of Scale Scores
(Online Only)



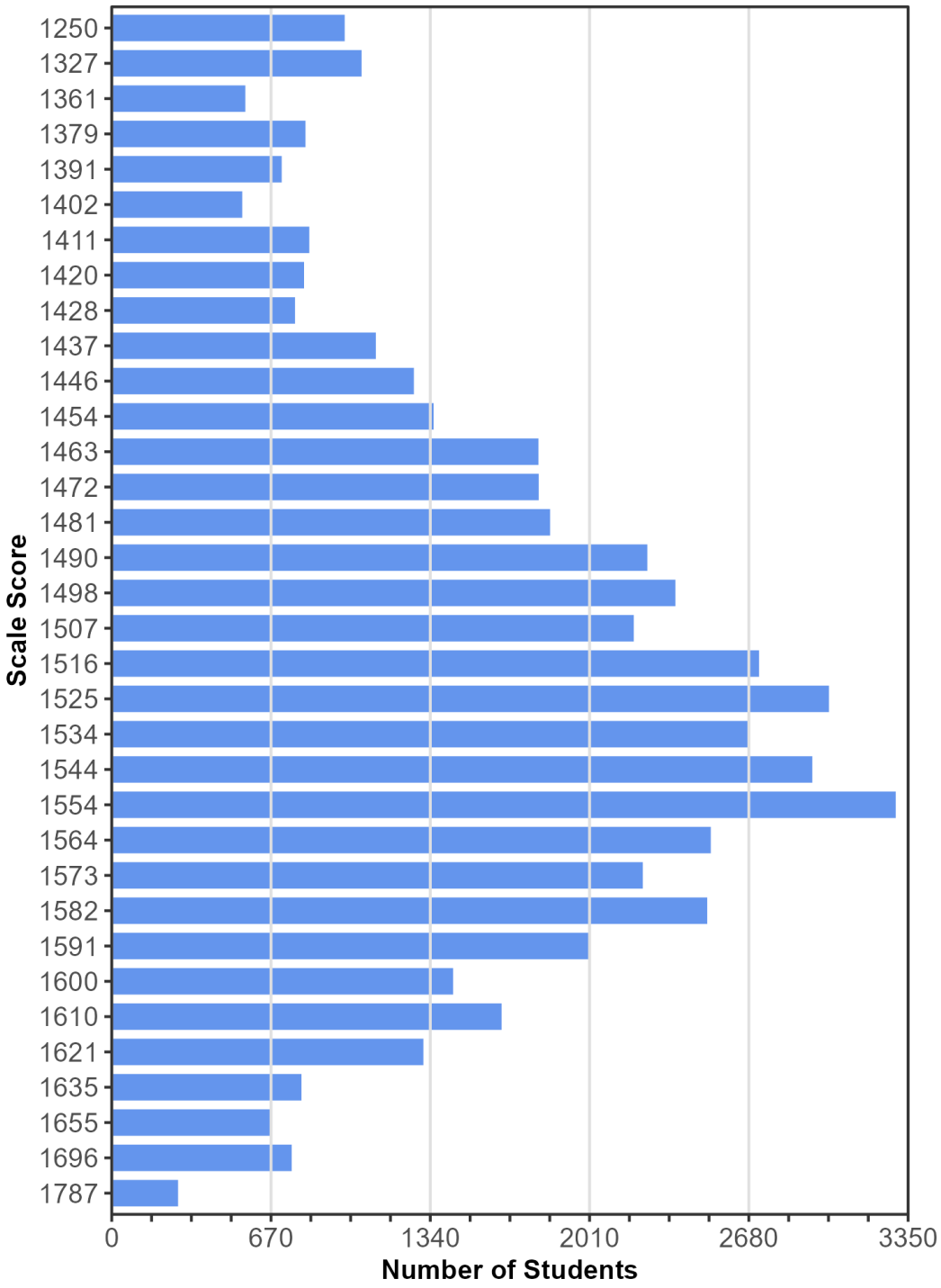
	Freq	Cum Freq	Pct	Cum Pct
1178	3828	3828	4.10	4.10
1275	4364	8192	4.68	8.78
1330	1589	9781	1.70	10.48
1355	1495	11276	1.60	12.08
1371	1719	12995	1.84	13.92
1383	1620	14615	1.74	15.66
1393	1743	16358	1.87	17.53
1403	2047	18405	2.19	19.72
1413	2253	20658	2.41	22.13
1424	2543	23201	2.72	24.86
1435	2942	26143	3.15	28.01
1447	3693	29836	3.96	31.97
1461	4449	34285	4.77	36.73
1474	4601	38886	4.93	41.66
1487	4525	43411	4.85	46.51
1498	4842	48253	5.19	51.70
1507	4645	52898	4.98	56.67
1516	4580	57478	4.91	61.58
1525	4654	62132	4.99	66.57
1534	4346	66478	4.66	71.22
1542	4117	70595	4.41	75.64
1552	4118	74713	4.41	80.05
1562	3522	78235	3.77	83.82
1572	3210	81445	3.44	87.26
1581	2853	84298	3.06	90.32
1591	1959	86257	2.10	92.42
1600	1748	88005	1.87	94.29
1609	1710	89715	1.83	96.12
1618	912	90627	0.98	97.10
1627	849	91476	0.91	98.01
1638	859	92335	0.92	98.93
1653	346	92681	0.37	99.30
1677	306	92987	0.33	99.63
1728	349	93336	0.37	100.00

Figure D.5.20. Spring 2023 TELPAS Grade 10 Writing
Frequency Distribution of Scale Scores
(Online Only)



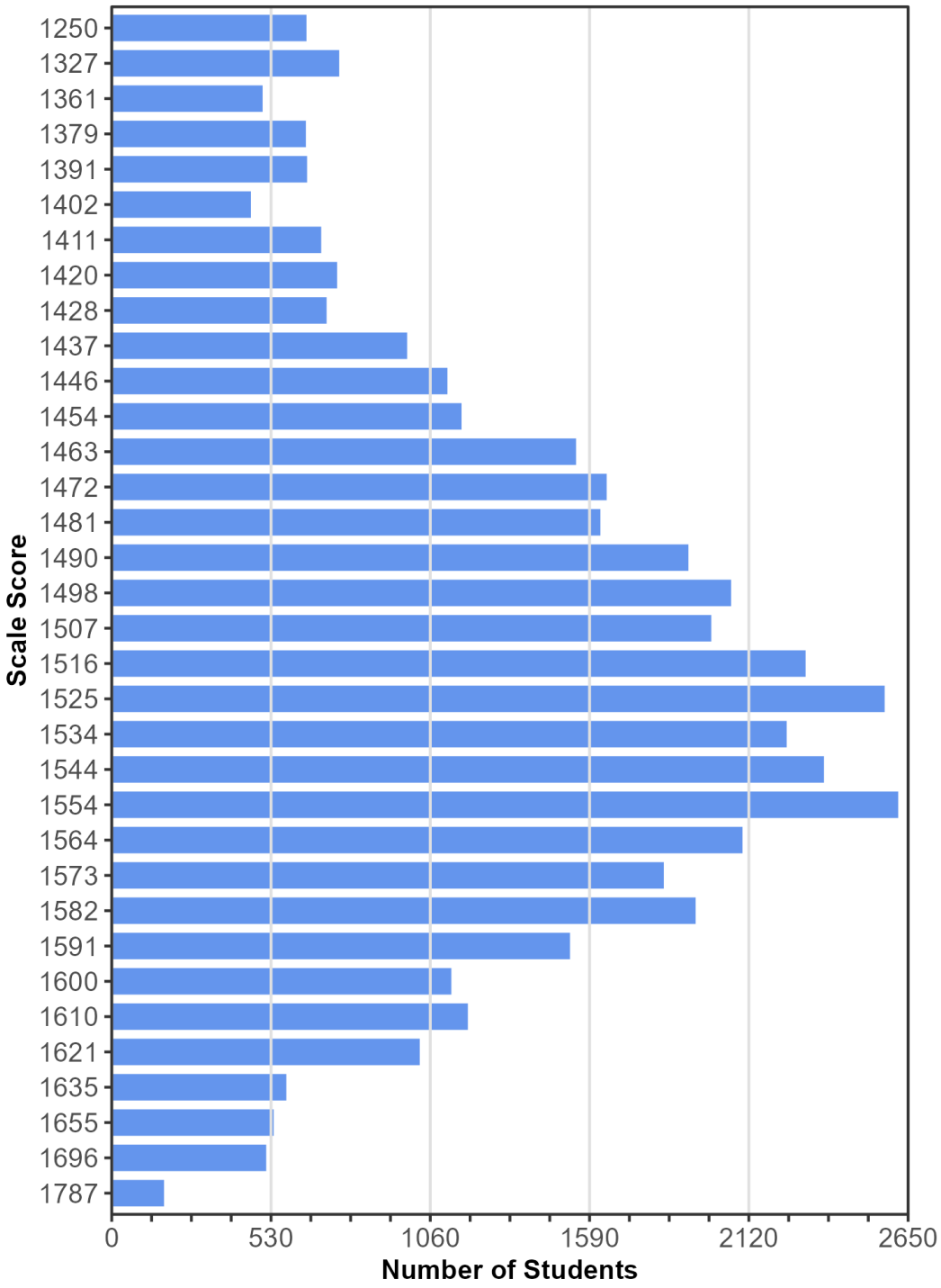
	Freq	Cum Freq	Pct	Cum Pct
1250	2216	2216	2.94	2.94
1327	1965	4181	2.61	5.55
1361	984	5165	1.31	6.85
1379	1324	6489	1.76	8.61
1391	1212	7701	1.61	10.22
1402	822	8523	1.09	11.31
1411	1364	9887	1.81	13.12
1420	1228	11115	1.63	14.75
1428	1254	12369	1.66	16.42
1437	1710	14079	2.27	18.69
1446	1820	15899	2.42	21.10
1454	1931	17830	2.56	23.66
1463	2637	20467	3.50	27.16
1472	2630	23097	3.49	30.65
1481	2622	25719	3.48	34.13
1490	3142	28861	4.17	38.30
1498	3273	32134	4.34	42.65
1507	3066	35200	4.07	46.72
1516	3891	39091	5.16	51.88
1525	4139	43230	5.49	57.37
1534	3526	46756	4.68	62.05
1544	4016	50772	5.33	67.38
1554	4458	55230	5.92	73.30
1564	3175	58405	4.21	77.51
1573	2885	61290	3.83	81.34
1582	3258	64548	4.32	85.67
1591	2463	67011	3.27	88.94
1600	1752	68763	2.33	91.26
1610	1879	70642	2.49	93.75
1621	1661	72303	2.20	95.96
1635	935	73238	1.24	97.20
1655	895	74133	1.19	98.39
1696	866	74999	1.15	99.54
1787	349	75348	0.46	100.00

Figure D.5.21. Spring 2023 TELPAS Grade 11 Writing
Frequency Distribution of Scale Scores
(Online Only)



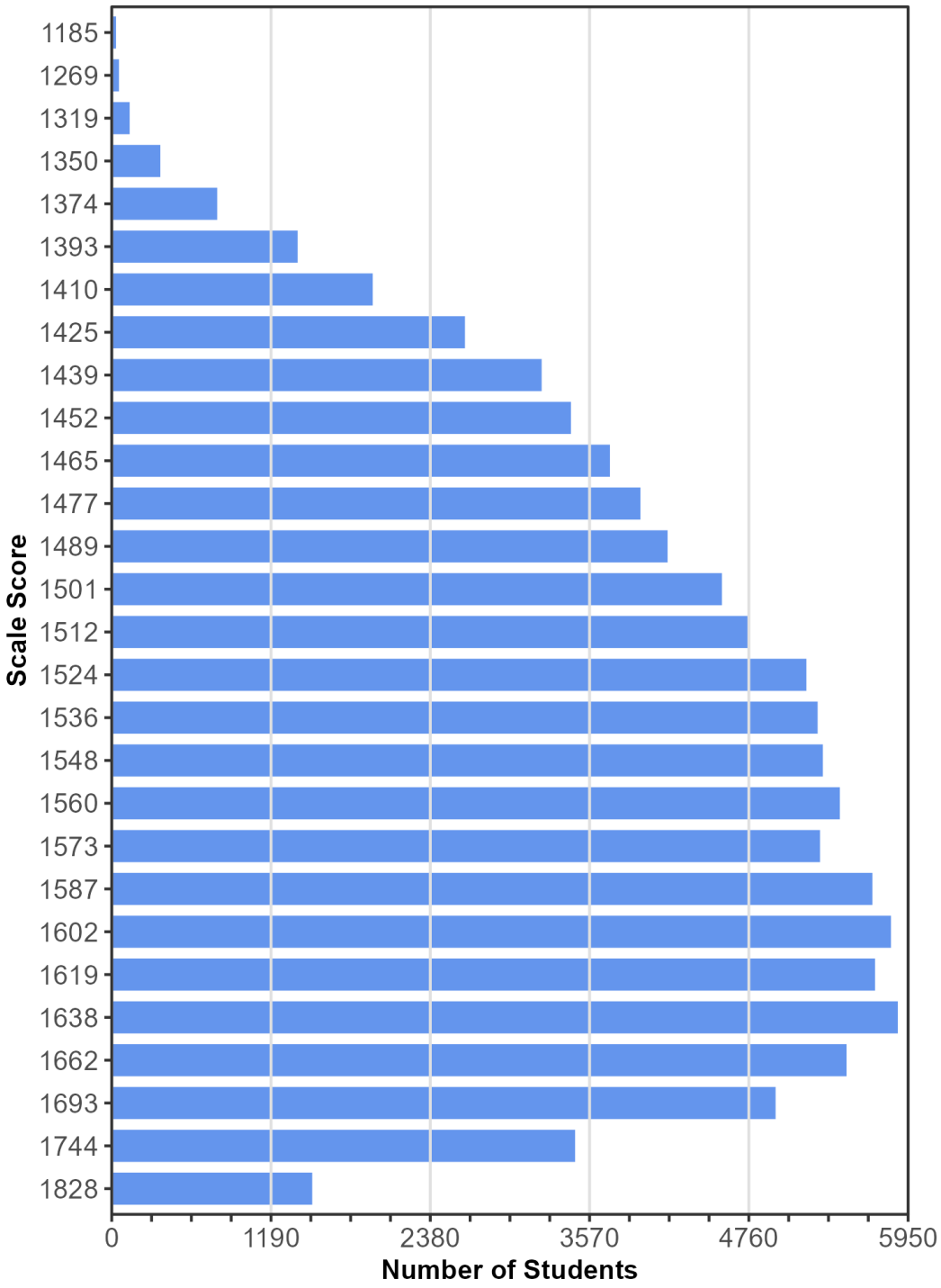
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1250	980	980	1.82	1.82
1327	1051	2031	1.95	3.77
1361	562	2593	1.04	4.81
1379	815	3408	1.51	6.32
1391	715	4123	1.33	7.65
1402	549	4672	1.02	8.67
1411	831	5503	1.54	10.21
1420	809	6312	1.50	11.71
1428	771	7083	1.43	13.14
1437	1111	8194	2.06	15.20
1446	1271	9465	2.36	17.56
1454	1354	10819	2.51	20.07
1463	1795	12614	3.33	23.40
1472	1796	14410	3.33	26.73
1481	1844	16254	3.42	30.16
1490	2253	18507	4.18	34.34
1498	2371	20878	4.40	38.73
1507	2196	23074	4.07	42.81
1516	2723	25797	5.05	47.86
1525	3017	28814	5.60	53.46
1534	2680	31494	4.97	58.43
1544	2947	34441	5.47	63.90
1554	3298	37739	6.12	70.02
1564	2520	40259	4.68	74.69
1573	2234	42493	4.14	78.84
1582	2505	44998	4.65	83.48
1591	2006	47004	3.72	87.21
1600	1436	48440	2.66	89.87
1610	1640	50080	3.04	92.91
1621	1311	51391	2.43	95.35
1635	798	52189	1.48	96.83
1655	675	52864	1.25	98.08
1696	757	53621	1.40	99.48
1787	279	53900	0.52	100.00

Figure D.5.22. Spring 2023 TELPAS Grade 12 Writing
Frequency Distribution of Scale Scores
(Online Only)



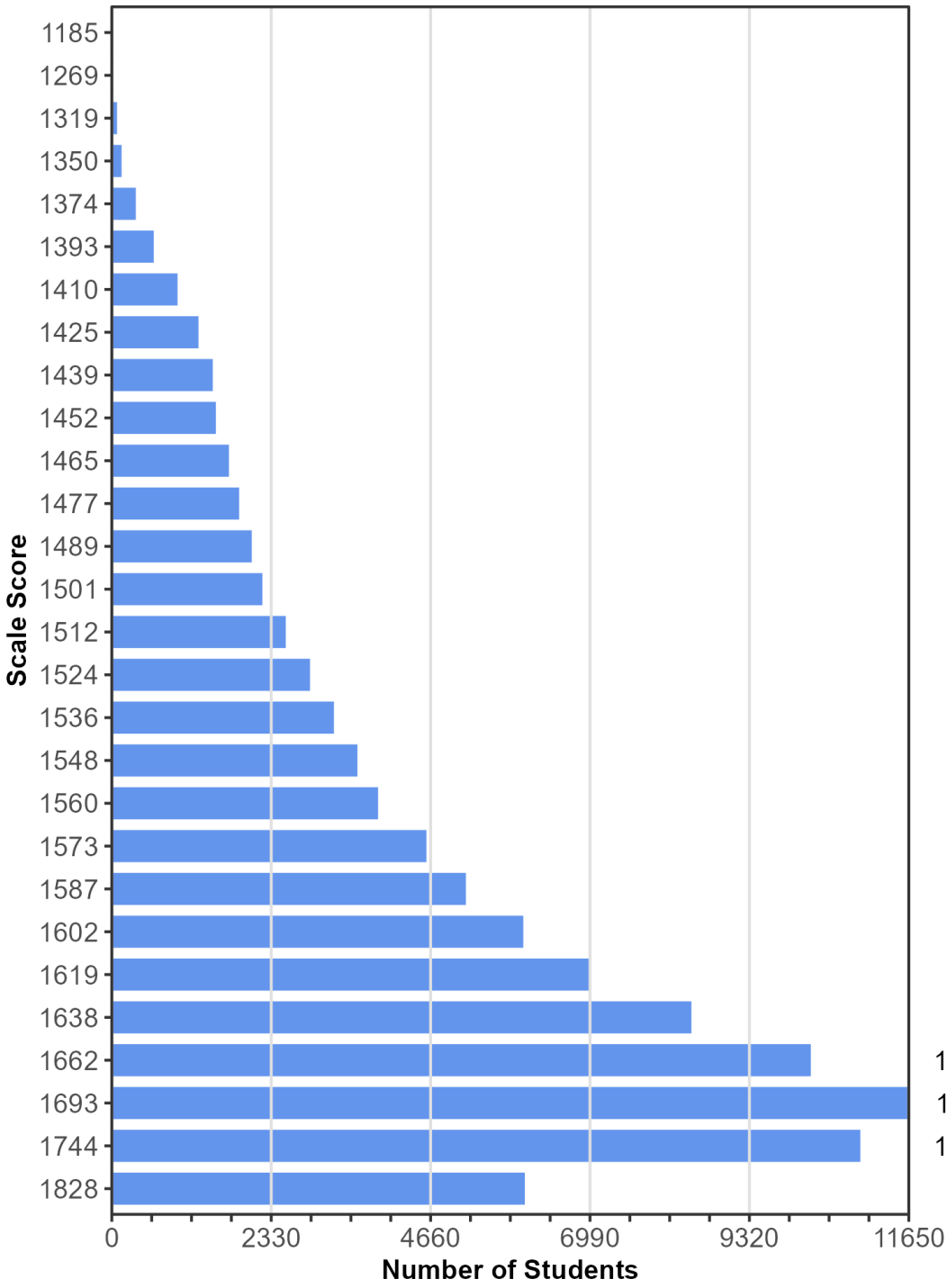
	Freq	Cum Freq	Pct	Cum Pct
1250	648	648	1.45	1.45
1327	757	1405	1.70	3.15
1361	502	1907	1.13	4.28
1379	646	2553	1.45	5.73
1391	650	3203	1.46	7.19
1402	463	3666	1.04	8.23
1411	697	4363	1.56	9.79
1420	750	5113	1.68	11.48
1428	715	5828	1.60	13.08
1437	983	6811	2.21	15.29
1446	1117	7928	2.51	17.80
1454	1164	9092	2.61	20.41
1463	1545	10637	3.47	23.88
1472	1647	12284	3.70	27.57
1481	1626	13910	3.65	31.22
1490	1919	15829	4.31	35.53
1498	2061	17890	4.63	40.16
1507	1995	19885	4.48	44.63
1516	2309	22194	5.18	49.82
1525	2572	24766	5.77	55.59
1534	2246	27012	5.04	60.63
1544	2370	29382	5.32	65.95
1554	2617	31999	5.87	71.83
1564	2099	34098	4.71	76.54
1573	1837	35935	4.12	80.66
1582	1943	37878	4.36	85.02
1591	1525	39403	3.42	88.44
1600	1130	40533	2.54	90.98
1610	1185	41718	2.66	93.64
1621	1025	42743	2.30	95.94
1635	581	43324	1.30	97.25
1655	539	43863	1.21	98.46
1696	514	44377	1.15	99.61
1787	174	44551	0.39	100.00

Figure D.5.23. Spring 2023 TELPAS Grade 2 Listening
Frequency Distribution of Scale Scores
(Online Only)



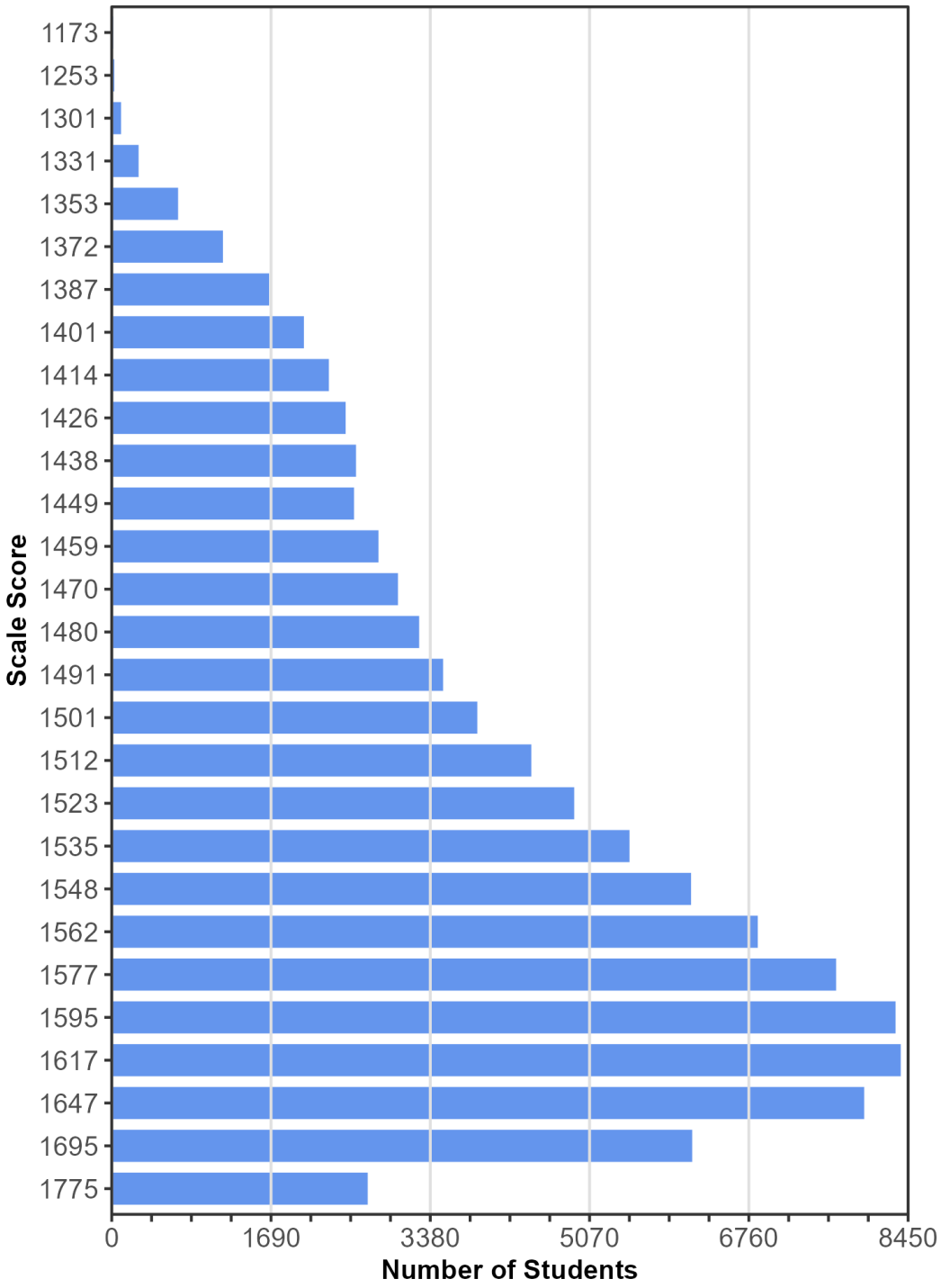
	Freq	Cum Freq	Pct	Cum Pct
1185	32	32	0.03	0.03
1269	54	86	0.05	0.09
1319	134	220	0.13	0.22
1350	363	583	0.36	0.58
1374	788	1371	0.79	1.37
1393	1389	2760	1.39	2.76
1410	1950	4710	1.95	4.70
1425	2639	7349	2.64	7.34
1439	3212	10561	3.21	10.55
1452	3431	13992	3.43	13.97
1465	3722	17714	3.72	17.69
1477	3950	21664	3.94	21.64
1489	4153	25817	4.15	25.78
1501	4559	30376	4.55	30.34
1512	4751	35127	4.74	35.08
1524	5190	40317	5.18	40.27
1536	5274	45591	5.27	45.53
1548	5313	50904	5.31	50.84
1560	5440	56344	5.43	56.27
1573	5292	61636	5.29	61.56
1587	5683	67319	5.68	67.23
1602	5822	73141	5.81	73.05
1619	5703	78844	5.70	78.74
1638	5873	84717	5.87	84.61
1662	5490	90207	5.48	90.09
1693	4960	95167	4.95	95.05
1744	3462	98629	3.46	98.50
1828	1498	100127	1.50	100.00

Figure D.5.24. Spring 2023 TELPAS Grade 3 Listening
Frequency Distribution of Scale Scores
(Online Only)



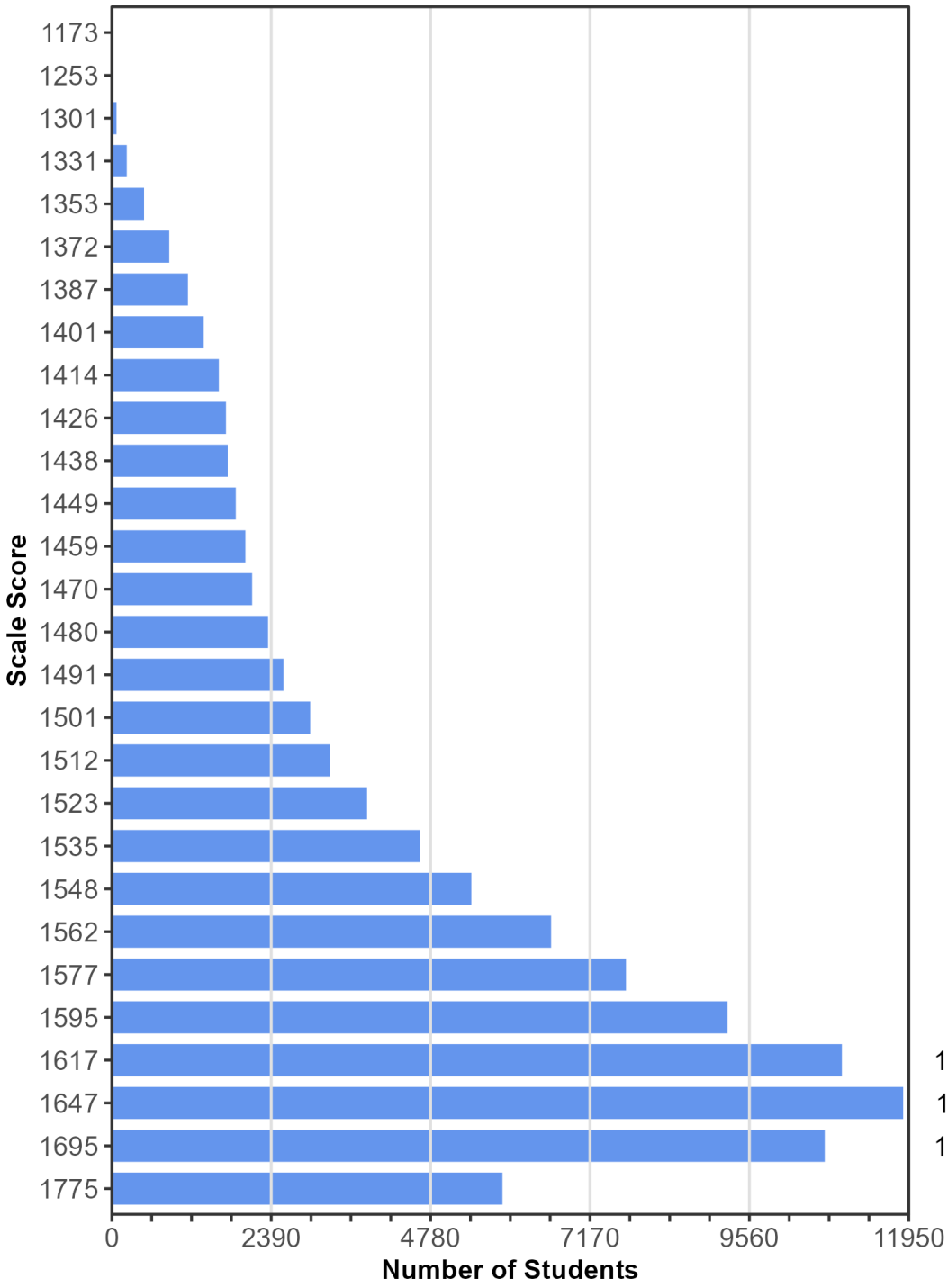
	Freq	Cum Freq	Pct	Cum Pct
1185	18	18	0.02	0.02
1269	19	37	0.02	0.04
1319	76	113	0.08	0.11
1350	143	256	0.14	0.25
1374	351	607	0.35	0.60
1393	613	1220	0.61	1.21
1410	961	2181	0.96	2.17
1425	1267	3448	1.26	3.43
1439	1476	4924	1.47	4.90
1452	1521	6445	1.51	6.41
1465	1712	8157	1.70	8.11
1477	1861	10018	1.85	9.97
1489	2046	12064	2.04	12.00
1501	2201	14265	2.19	14.19
1512	2543	16808	2.53	16.72
1524	2897	19705	2.88	19.60
1536	3247	22952	3.23	22.83
1548	3590	26542	3.57	26.40
1560	3892	30434	3.87	30.28
1573	4599	35033	4.57	34.85
1587	5176	40209	5.15	40.00
1602	6015	46224	5.98	45.98
1619	6998	53222	6.96	52.94
1638	8472	61694	8.43	61.37
1662	10218	71912	10.16	71.54
1693	11634	83546	11.57	83.11
1744	10942	94488	10.88	93.99
1828	6037	100525	6.01	100.00

Figure D.5.25. Spring 2023 TELPAS Grade 4 Listening
Frequency Distribution of Scale Scores
(Online Only)



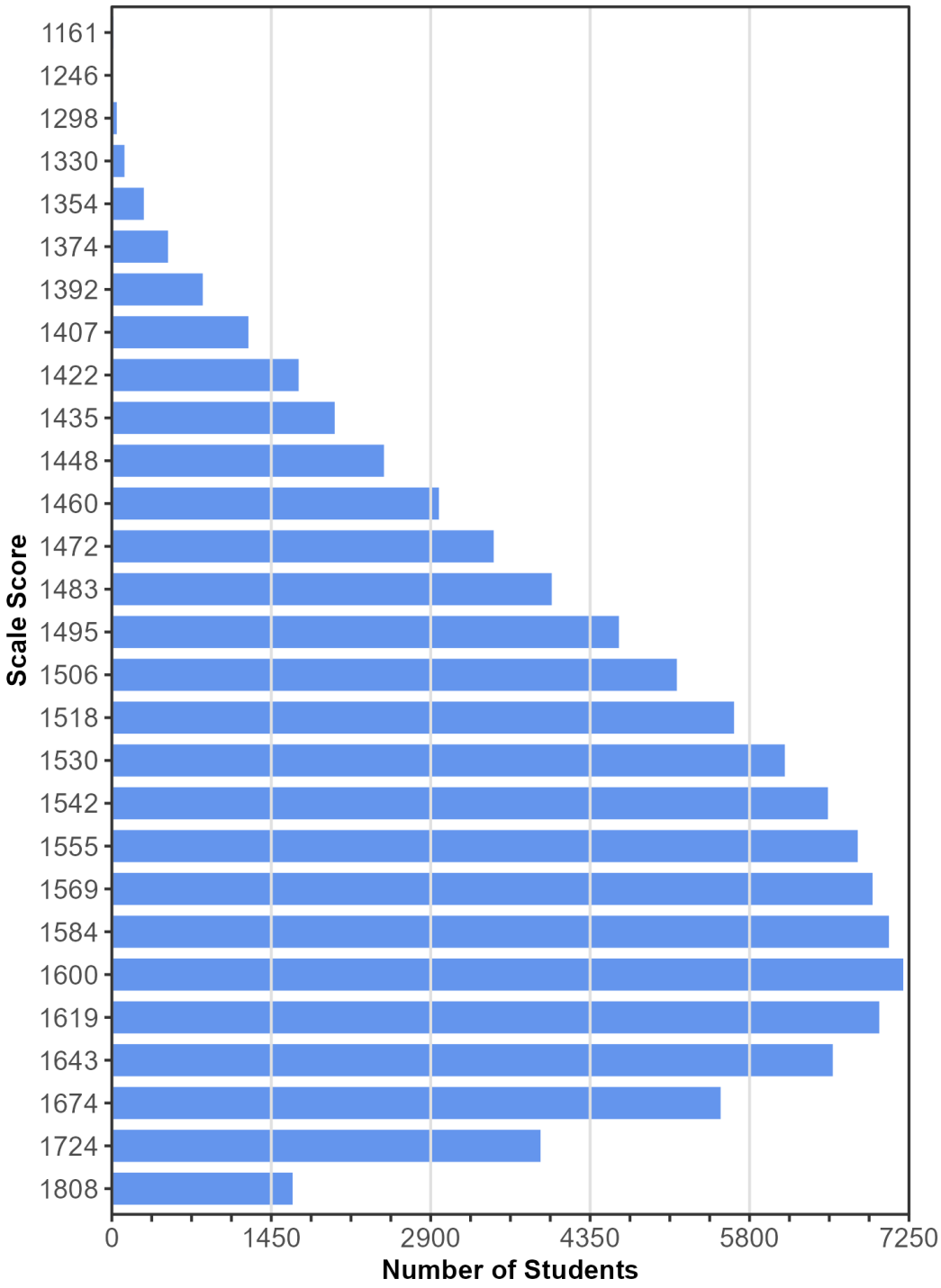
	Freq	Cum Freq	Pct	Cum Pct
1173	17	17	0.02	0.02
1253	27	44	0.03	0.04
1301	99	143	0.10	0.14
1331	285	428	0.28	0.42
1353	704	1132	0.69	1.11
1372	1180	2312	1.16	2.28
1387	1670	3982	1.64	3.92
1401	2039	6021	2.01	5.93
1414	2304	8325	2.27	8.20
1426	2482	10807	2.44	10.64
1438	2592	13399	2.55	13.19
1449	2571	15970	2.53	15.72
1459	2830	18800	2.79	18.51
1470	3037	21837	2.99	21.50
1480	3262	25099	3.21	24.71
1491	3516	28615	3.46	28.17
1501	3879	32494	3.82	31.99
1512	4452	36946	4.38	36.37
1523	4907	41853	4.83	41.20
1535	5494	47347	5.41	46.61
1548	6147	53494	6.05	52.66
1562	6855	60349	6.75	59.41
1577	7686	68035	7.57	66.97
1595	8317	76352	8.19	75.16
1617	8372	84724	8.24	83.40
1647	7984	92708	7.86	91.26
1695	6159	98867	6.06	97.33
1775	2716	101583	2.67	100.00

Figure D.5.26. Spring 2023 TELPAS Grade 5 Listening
Frequency Distribution of Scale Scores
(Online Only)



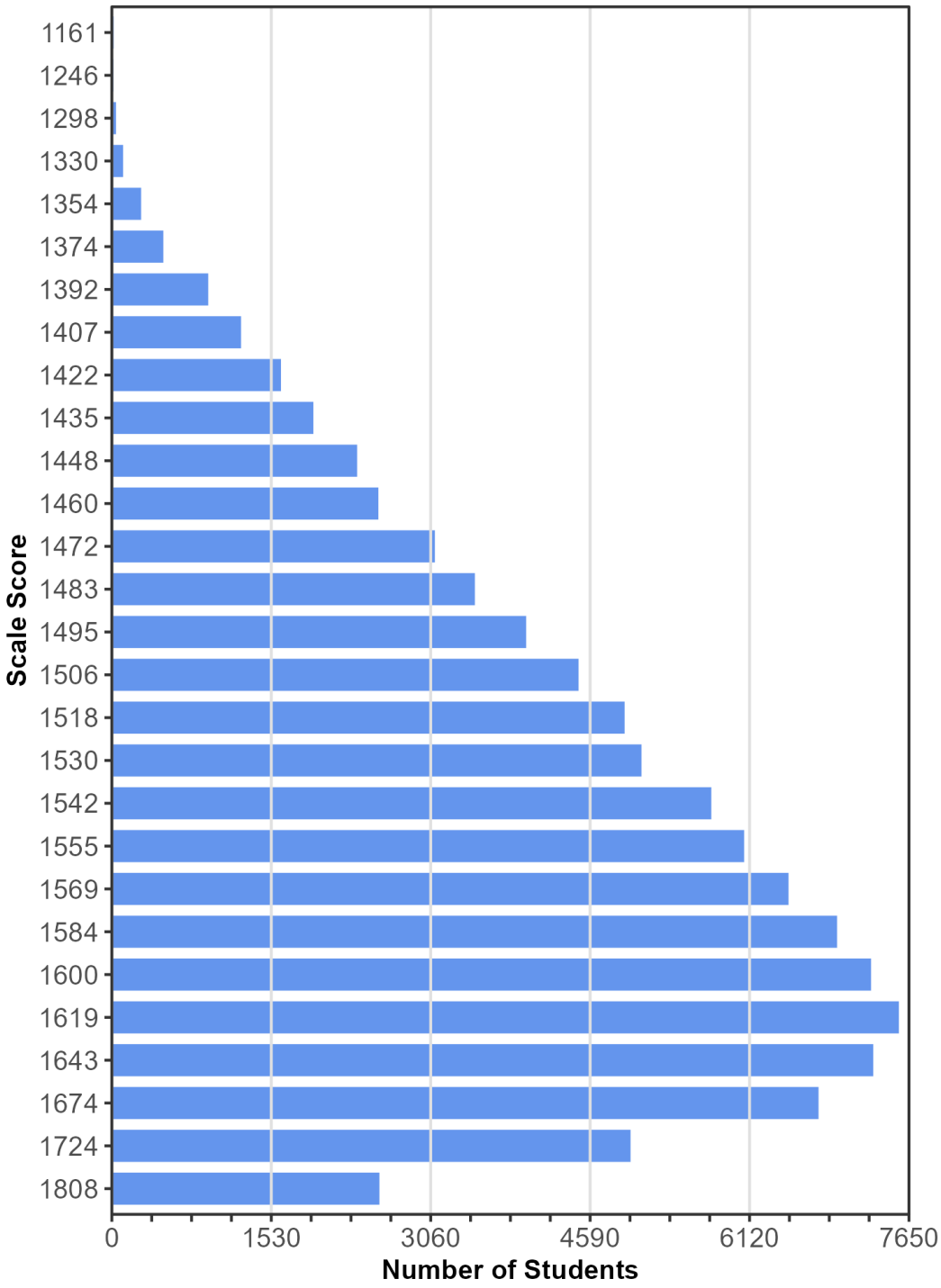
	Freq	Cum Freq	Pct	Cum Pct
1173	12	12	0.01	0.01
1253	18	30	0.02	0.03
1301	70	100	0.07	0.10
1331	224	324	0.22	0.31
1353	483	807	0.47	0.78
1372	861	1668	0.84	1.62
1387	1141	2809	1.11	2.72
1401	1378	4187	1.34	4.06
1414	1605	5792	1.56	5.62
1426	1712	7504	1.66	7.28
1438	1738	9242	1.69	8.96
1449	1859	11101	1.80	10.77
1459	2003	13104	1.94	12.71
1470	2104	15208	2.04	14.75
1480	2343	17551	2.27	17.02
1491	2573	20124	2.50	19.52
1501	2975	23099	2.89	22.41
1512	3268	26367	3.17	25.58
1523	3827	30194	3.71	29.29
1535	4618	34812	4.48	33.77
1548	5391	40203	5.23	39.00
1562	6587	46790	6.39	45.39
1577	7710	54500	7.48	52.87
1595	9231	63731	8.95	61.82
1617	10947	74678	10.62	72.44
1647	11866	86544	11.51	83.95
1695	10690	97234	10.37	94.32
1775	5857	103091	5.68	100.00

Figure D.5.27. Spring 2023 TELPAS Grade 6 Listening
Frequency Distribution of Scale Scores
(Online Only)



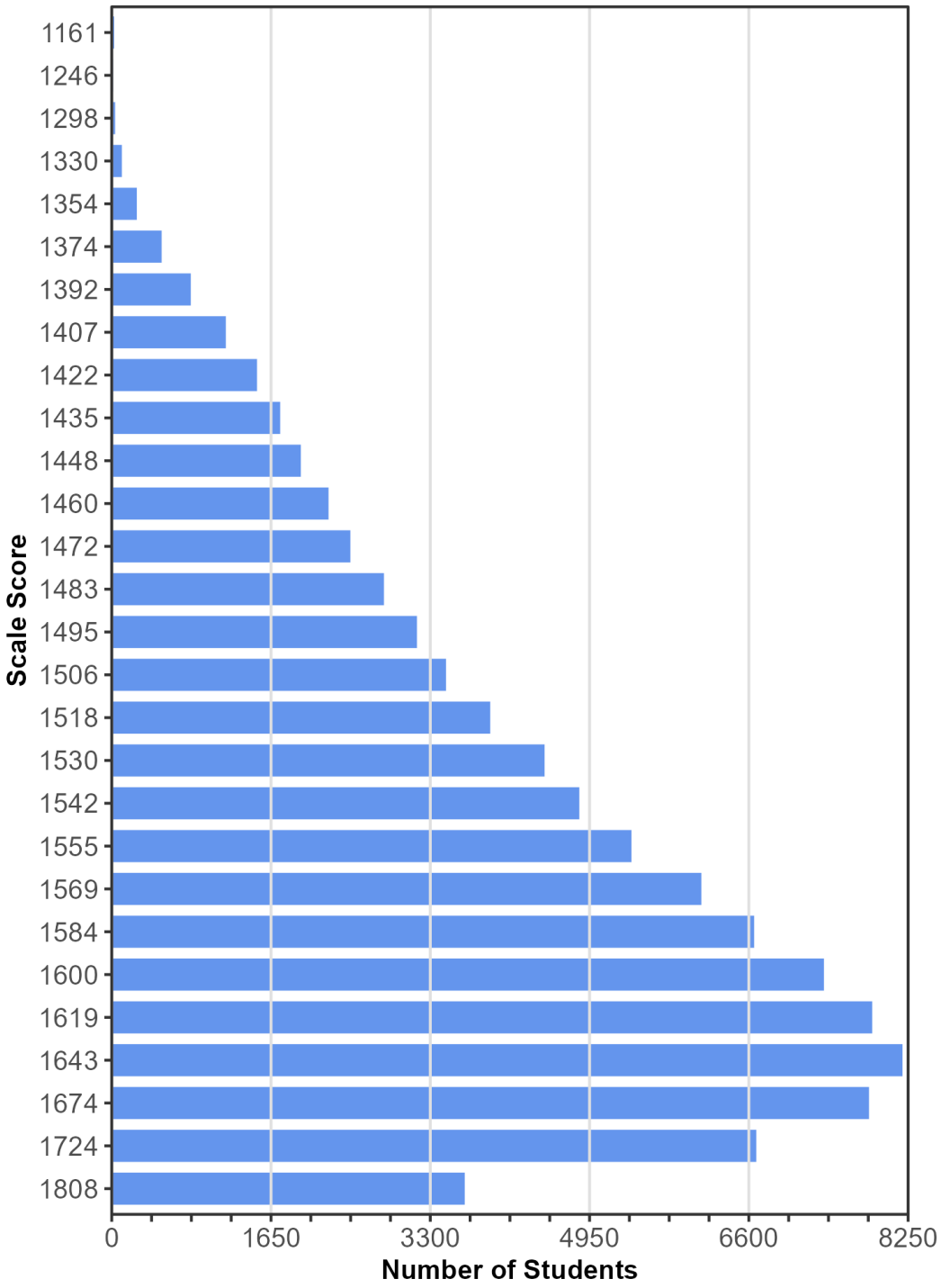
	Freq	Cum Freq	Pct	Cum Pct
1161	14	14	0.01	0.01
1246	10	24	0.01	0.02
1298	45	69	0.04	0.07
1330	115	184	0.11	0.18
1354	291	475	0.29	0.47
1374	511	986	0.51	0.98
1392	827	1813	0.82	1.81
1407	1242	3055	1.24	3.04
1422	1699	4754	1.69	4.74
1435	2027	6781	2.02	6.76
1448	2475	9256	2.47	9.23
1460	2975	12231	2.97	12.19
1472	3473	15704	3.46	15.65
1483	4001	19705	3.99	19.64
1495	4612	24317	4.60	24.24
1506	5139	29456	5.12	29.36
1518	5659	35115	5.64	35.00
1530	6121	41236	6.10	41.10
1542	6513	47749	6.49	47.59
1555	6784	54533	6.76	54.35
1569	6919	61452	6.90	61.25
1584	7068	68520	7.04	68.29
1600	7198	75718	7.17	75.47
1619	6980	82698	6.96	82.42
1643	6557	89255	6.54	88.96
1674	5537	94792	5.52	94.48
1724	3899	98691	3.89	98.36
1808	1644	100335	1.64	100.00

Figure D.5.28. Spring 2023 TELPAS Grade 7 Listening
Frequency Distribution of Scale Scores
(Online Only)



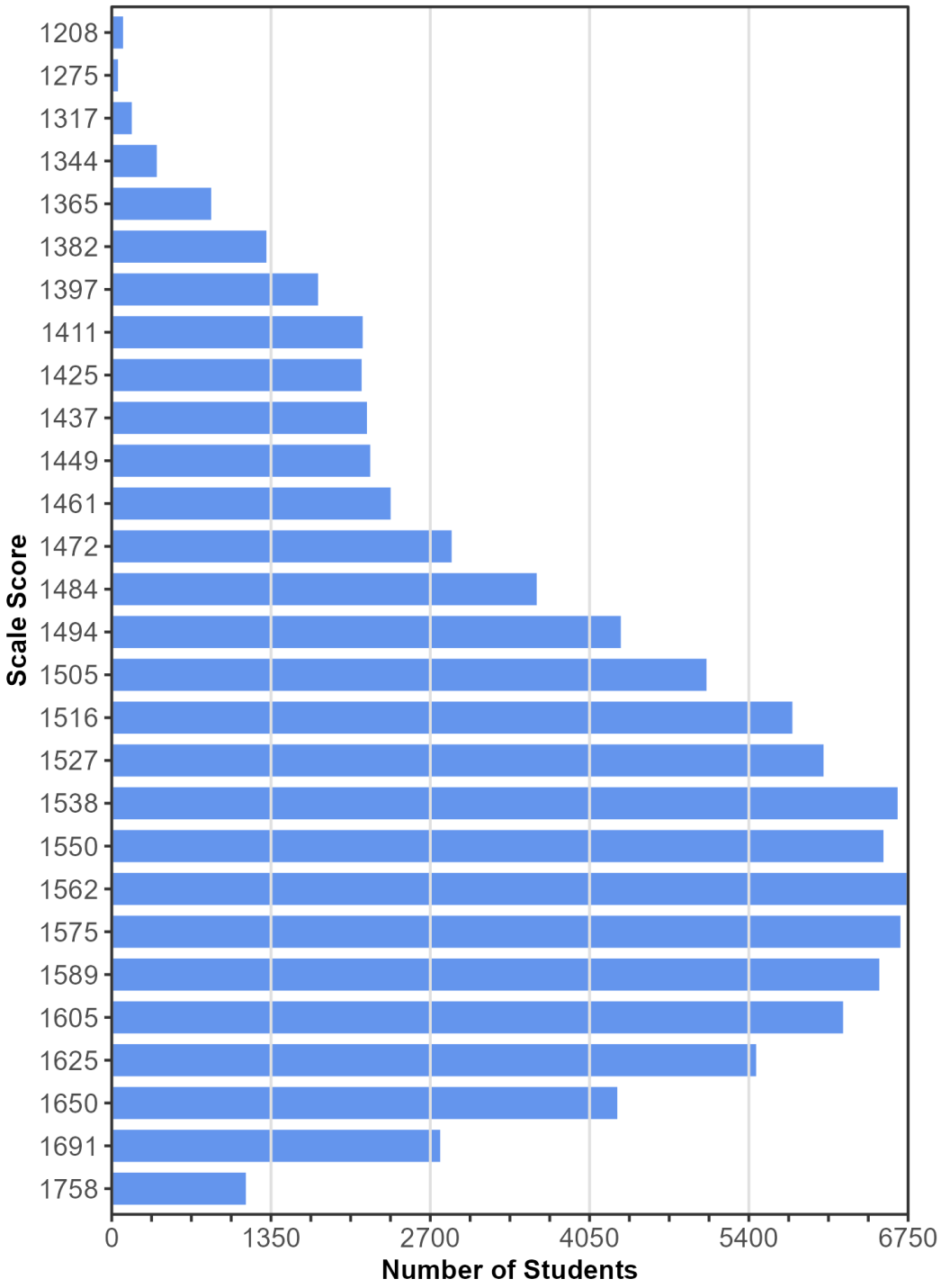
	Freq	Cum Freq	Pct	Cum Pct
1161	18	18	0.02	0.02
1246	16	34	0.02	0.03
1298	40	74	0.04	0.08
1330	108	182	0.11	0.18
1354	281	463	0.29	0.47
1374	494	957	0.50	0.97
1392	925	1882	0.94	1.91
1407	1240	3122	1.26	3.17
1422	1623	4745	1.65	4.82
1435	1934	6679	1.97	6.79
1448	2354	9033	2.39	9.18
1460	2558	11591	2.60	11.78
1472	3100	14691	3.15	14.93
1483	3484	18175	3.54	18.47
1495	3976	22151	4.04	22.52
1506	4479	26630	4.55	27.07
1518	4921	31551	5.00	32.07
1530	5083	36634	5.17	37.24
1542	5753	42387	5.85	43.08
1555	6068	48455	6.17	49.25
1569	6494	54949	6.60	55.85
1584	6960	61909	7.07	62.93
1600	7286	69195	7.41	70.33
1619	7553	76748	7.68	78.01
1643	7307	84055	7.43	85.44
1674	6782	90837	6.89	92.33
1724	4978	95815	5.06	97.39
1808	2568	98383	2.61	100.00

Figure D.5.29. Spring 2023 TELPAS Grade 8 Listening
Frequency Distribution of Scale Scores
(Online Only)



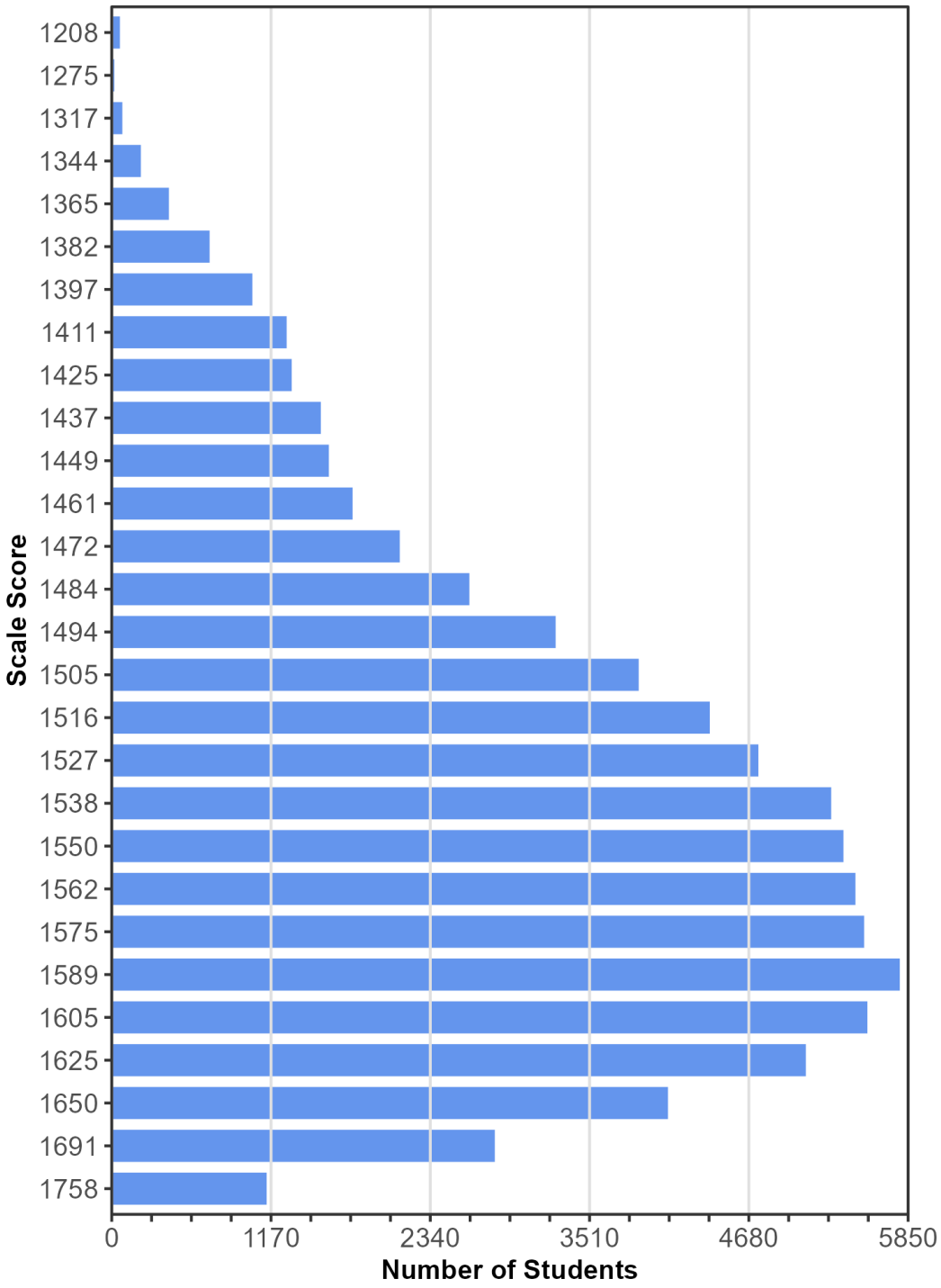
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1161	25	25	0.03	0.03
1246	14	39	0.01	0.04
1298	35	74	0.04	0.08
1330	105	179	0.11	0.19
1354	260	439	0.27	0.46
1374	517	956	0.54	1.00
1392	819	1775	0.86	1.86
1407	1182	2957	1.24	3.10
1422	1505	4462	1.58	4.68
1435	1745	6207	1.83	6.51
1448	1959	8166	2.05	8.56
1460	2246	10412	2.36	10.92
1472	2472	12884	2.59	13.51
1483	2820	15704	2.96	16.47
1495	3162	18866	3.32	19.79
1506	3463	22329	3.63	23.42
1518	3921	26250	4.11	27.53
1530	4484	30734	4.70	32.23
1542	4844	35578	5.08	37.31
1555	5384	40962	5.65	42.96
1569	6108	47070	6.41	49.36
1584	6655	53725	6.98	56.34
1600	7378	61103	7.74	64.08
1619	7878	68981	8.26	72.34
1643	8192	77173	8.59	80.93
1674	7845	85018	8.23	89.16
1724	6677	91695	7.00	96.16
1808	3657	95352	3.84	100.00

Figure D.5.30. Spring 2023 TELPAS Grade 9 Listening
Frequency Distribution of Scale Scores
(Online Only)



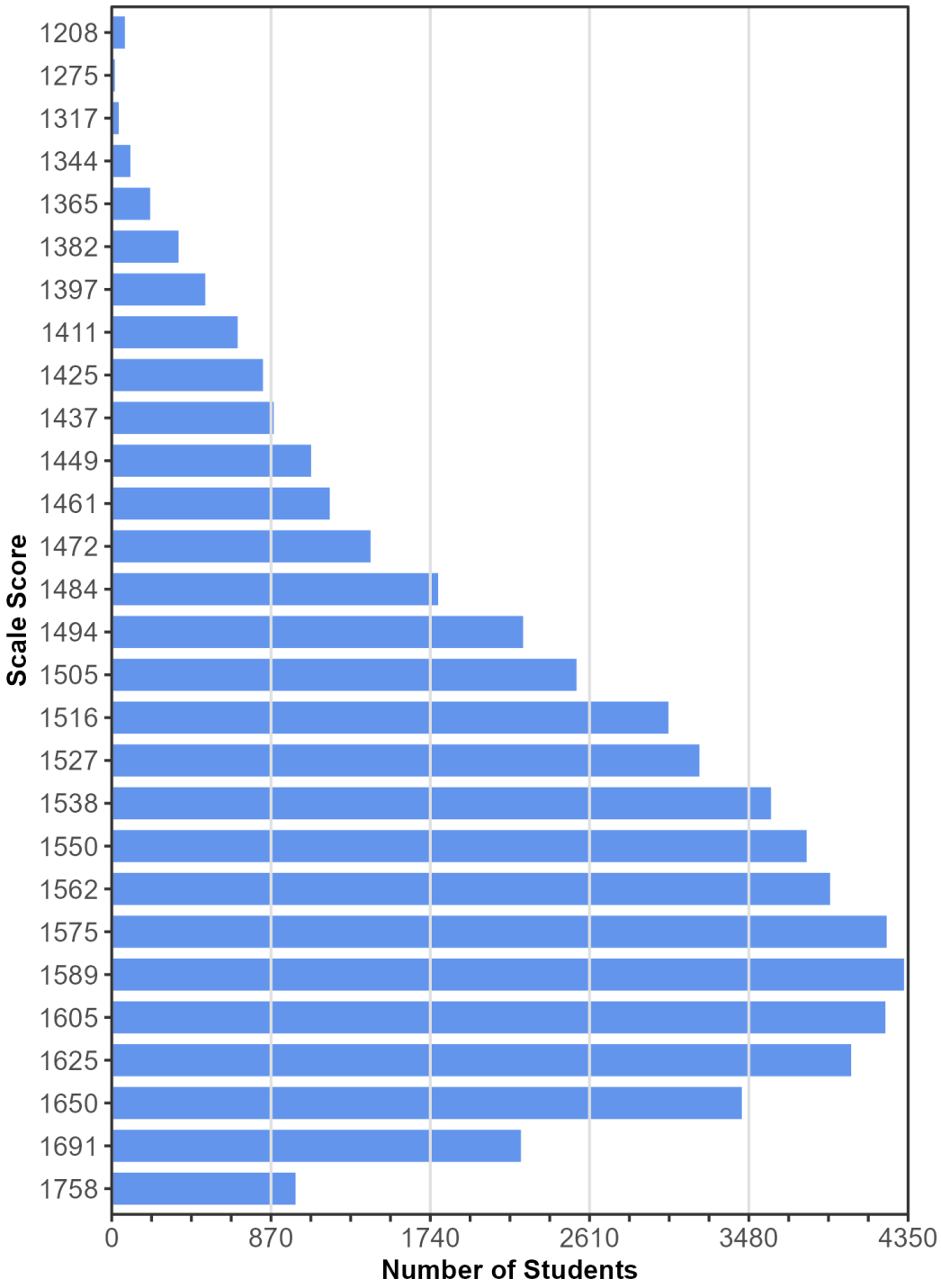
	Freq	Cum Freq	Pct	Cum Pct
1208	96	96	0.10	0.10
1275	53	149	0.06	0.15
1317	170	319	0.18	0.33
1344	382	701	0.40	0.73
1365	843	1544	0.88	1.60
1382	1311	2855	1.36	2.97
1397	1749	4604	1.82	4.79
1411	2127	6731	2.21	7.00
1425	2118	8849	2.20	9.20
1437	2163	11012	2.25	11.45
1449	2191	13203	2.28	13.72
1461	2364	15567	2.46	16.18
1472	2881	18448	2.99	19.18
1484	3602	22050	3.74	22.92
1494	4315	26365	4.49	27.40
1505	5041	31406	5.24	32.64
1516	5769	37175	6.00	38.64
1527	6033	43208	6.27	44.91
1538	6661	49869	6.92	51.84
1550	6541	56410	6.80	58.63
1562	6738	63148	7.00	65.64
1575	6685	69833	6.95	72.59
1589	6506	76339	6.76	79.35
1605	6199	82538	6.44	85.79
1625	5462	88000	5.68	91.47
1650	4285	92285	4.45	95.92
1691	2784	95069	2.89	98.82
1758	1137	96206	1.18	100.00

Figure D.5.31. Spring 2023 TELPAS Grade 10 Listening
Frequency Distribution of Scale Scores
(Online Only)



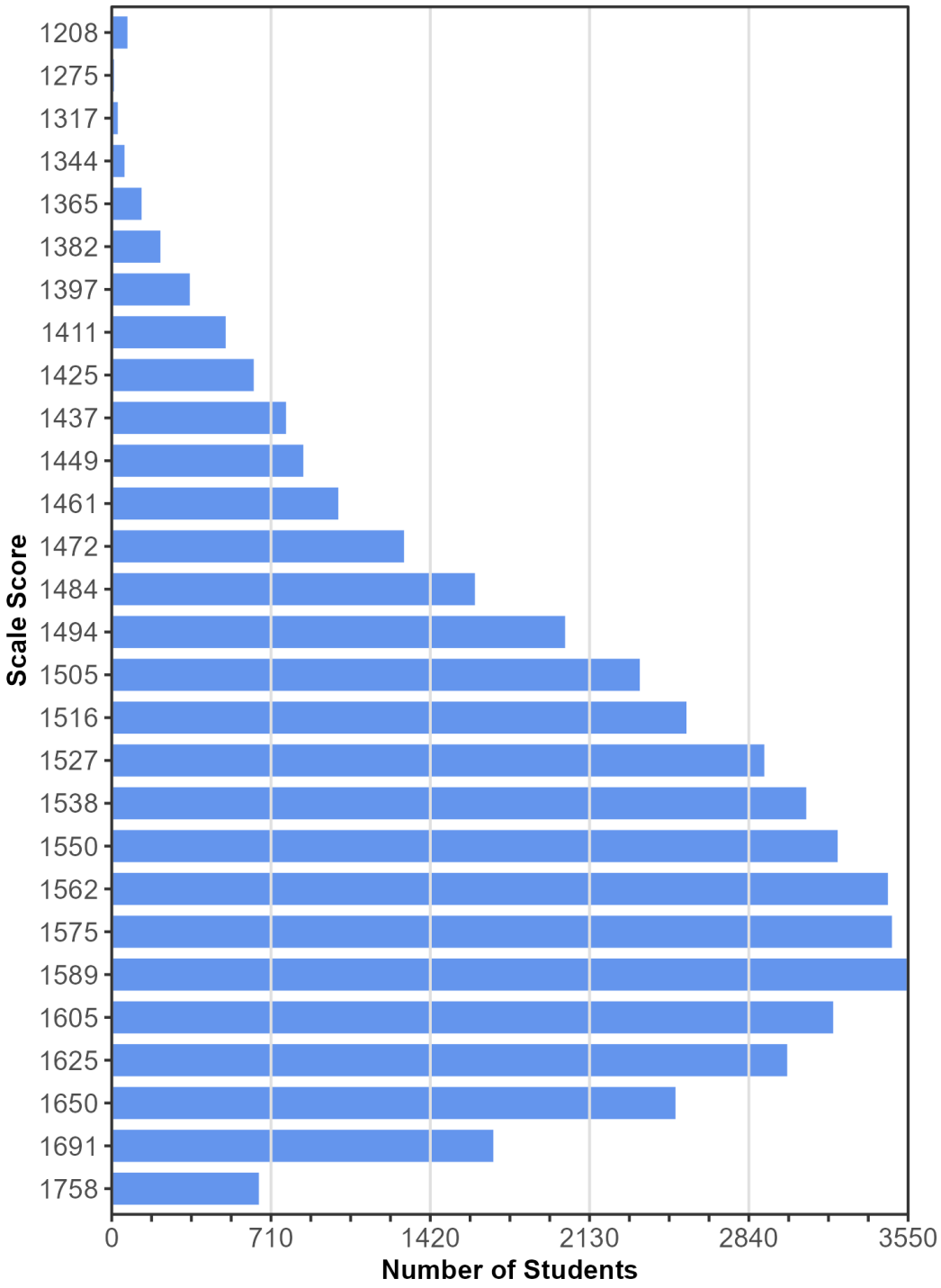
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1208	60	60	0.08	0.08
1275	19	79	0.02	0.10
1317	78	157	0.10	0.20
1344	214	371	0.28	0.48
1365	420	791	0.54	1.02
1382	719	1510	0.93	1.96
1397	1033	2543	1.34	3.29
1411	1285	3828	1.66	4.96
1425	1322	5150	1.71	6.67
1437	1536	6686	1.99	8.66
1449	1595	8281	2.07	10.73
1461	1769	10050	2.29	13.02
1472	2116	12166	2.74	15.76
1484	2627	14793	3.40	19.16
1494	3261	18054	4.22	23.39
1505	3871	21925	5.01	28.40
1516	4393	26318	5.69	34.09
1527	4750	31068	6.15	40.25
1538	5285	36353	6.85	47.09
1550	5375	41728	6.96	54.06
1562	5463	47191	7.08	61.13
1575	5527	52718	7.16	68.29
1589	5789	58507	7.50	75.79
1605	5551	64058	7.19	82.98
1625	5099	69157	6.61	89.59
1650	4086	73243	5.29	94.88
1691	2814	76057	3.65	98.53
1758	1138	77195	1.47	100.00

Figure D.5.32. Spring 2023 TELPAS Grade 11 Listening
Frequency Distribution of Scale Scores
(Online Only)



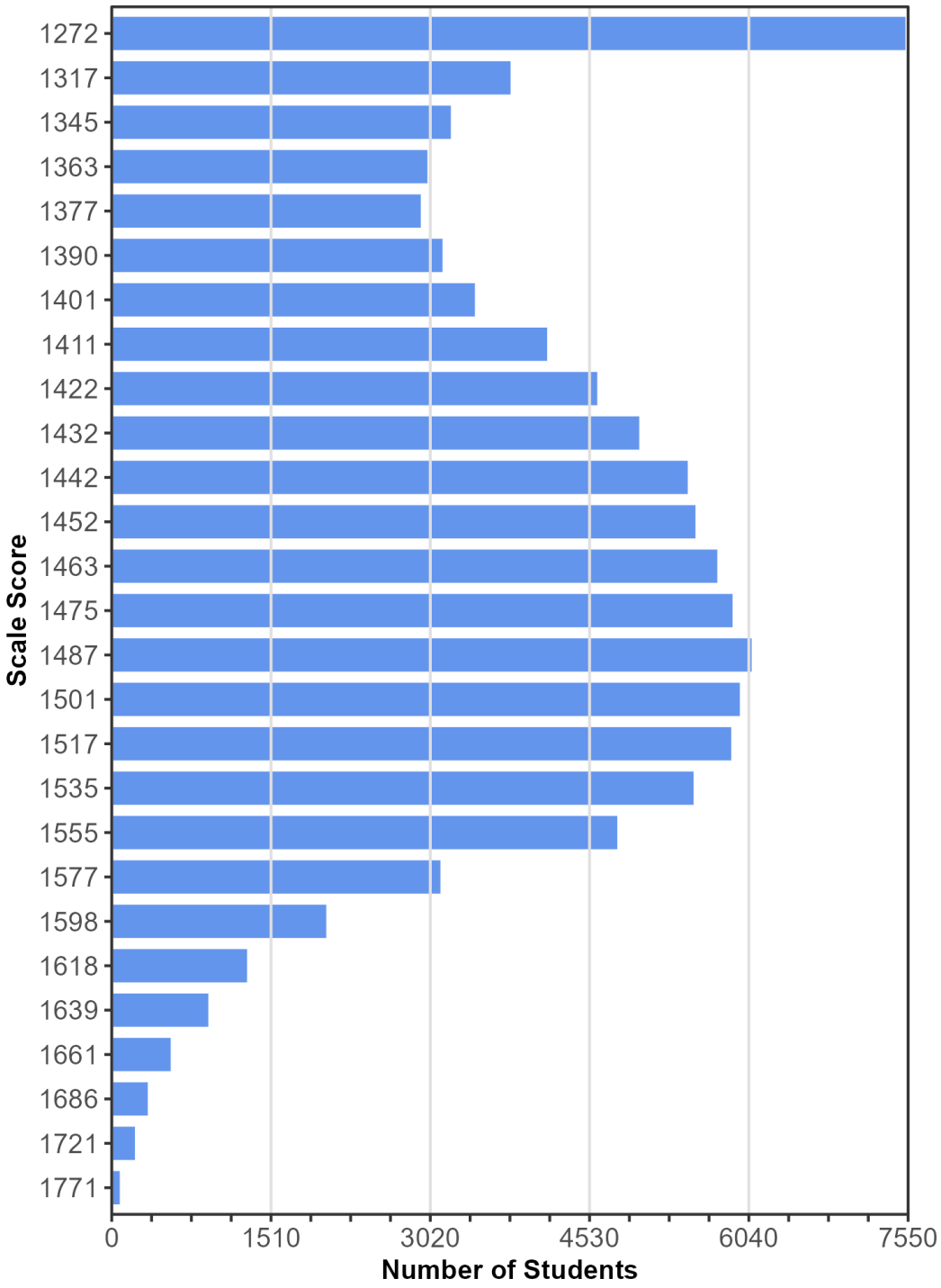
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1208	72	72	0.13	0.13
1275	17	89	0.03	0.16
1317	38	127	0.07	0.23
1344	102	229	0.19	0.42
1365	210	439	0.38	0.80
1382	365	804	0.66	1.46
1397	511	1315	0.93	2.39
1411	688	2003	1.25	3.64
1425	826	2829	1.50	5.14
1437	885	3714	1.61	6.75
1449	1089	4803	1.98	8.72
1461	1191	5994	2.16	10.89
1472	1414	7408	2.57	13.46
1484	1783	9191	3.24	16.69
1494	2247	11438	4.08	20.78
1505	2539	13977	4.61	25.39
1516	3041	17018	5.52	30.91
1527	3210	20228	5.83	36.74
1538	3601	23829	6.54	43.28
1550	3796	27625	6.89	50.18
1562	3924	31549	7.13	57.30
1575	4233	35782	7.69	64.99
1589	4328	40110	7.86	72.85
1605	4226	44336	7.68	80.53
1625	4039	48375	7.34	87.87
1650	3442	51817	6.25	94.12
1691	2235	54052	4.06	98.18
1758	1004	55056	1.82	100.00

Figure D.5.33. Spring 2023 TELPAS Grade 12 Listening
Frequency Distribution of Scale Scores
(Online Only)



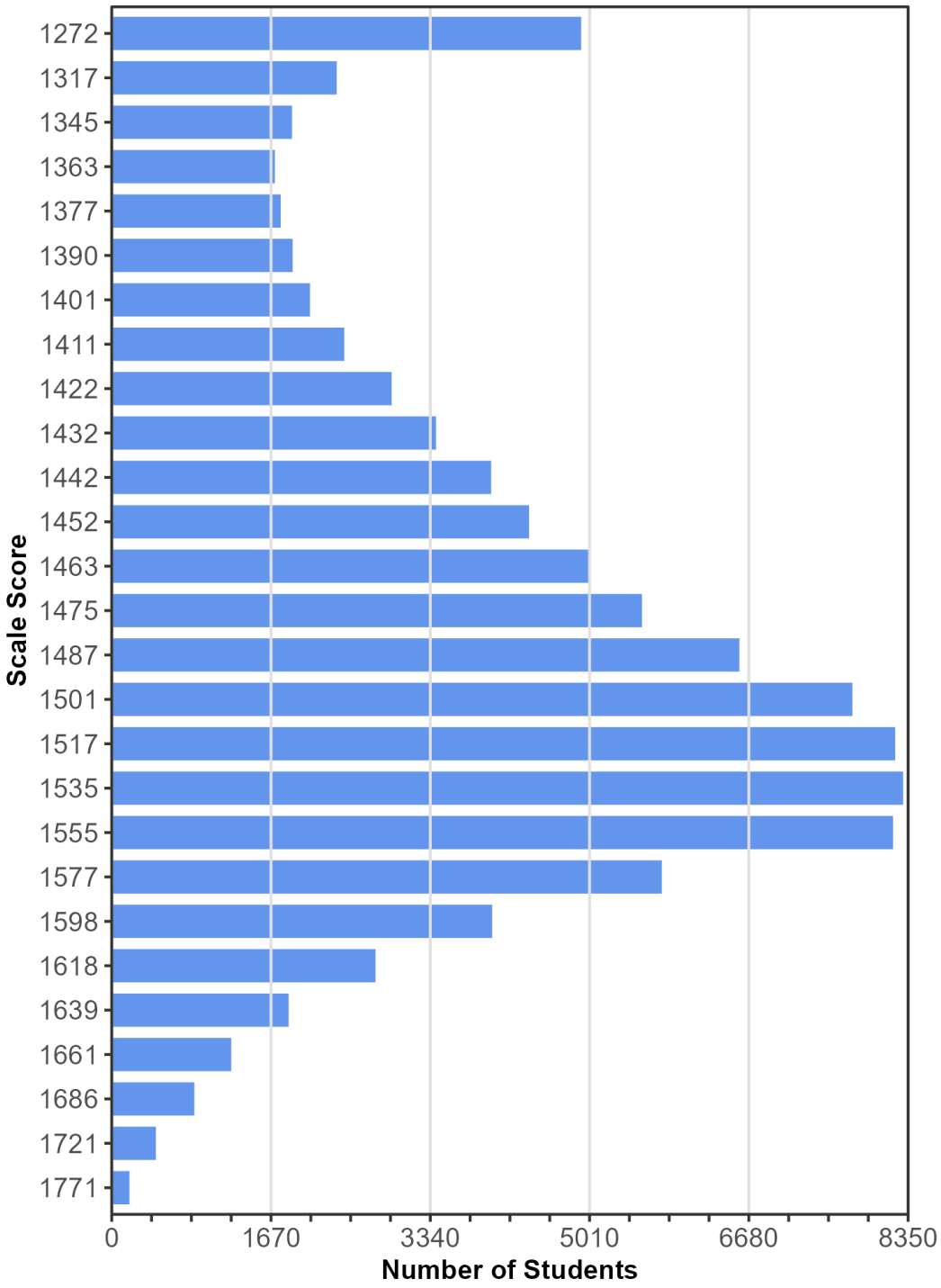
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1208	70	70	0.15	0.15
1275	10	80	0.02	0.18
1317	27	107	0.06	0.24
1344	57	164	0.13	0.36
1365	133	297	0.29	0.66
1382	217	514	0.48	1.13
1397	348	862	0.77	1.90
1411	508	1370	1.12	3.02
1425	633	2003	1.40	4.42
1437	777	2780	1.71	6.13
1449	854	3634	1.88	8.02
1461	1010	4644	2.23	10.25
1472	1303	5947	2.87	13.12
1484	1619	7566	3.57	16.69
1494	2021	9587	4.46	21.15
1505	2354	11941	5.19	26.35
1516	2562	14503	5.65	32.00
1527	2909	17412	6.42	38.42
1538	3096	20508	6.83	45.25
1550	3236	23744	7.14	52.39
1562	3460	27204	7.63	60.02
1575	3478	30682	7.67	67.69
1589	3546	34228	7.82	75.52
1605	3216	37444	7.10	82.61
1625	3011	40455	6.64	89.26
1650	2513	42968	5.54	94.80
1691	1701	44669	3.75	98.55
1758	656	45325	1.45	100.00

Figure D.5.34. Spring 2023 TELPAS Grade 2 Speaking
Frequency Distribution of Scale Scores
(Online Only)



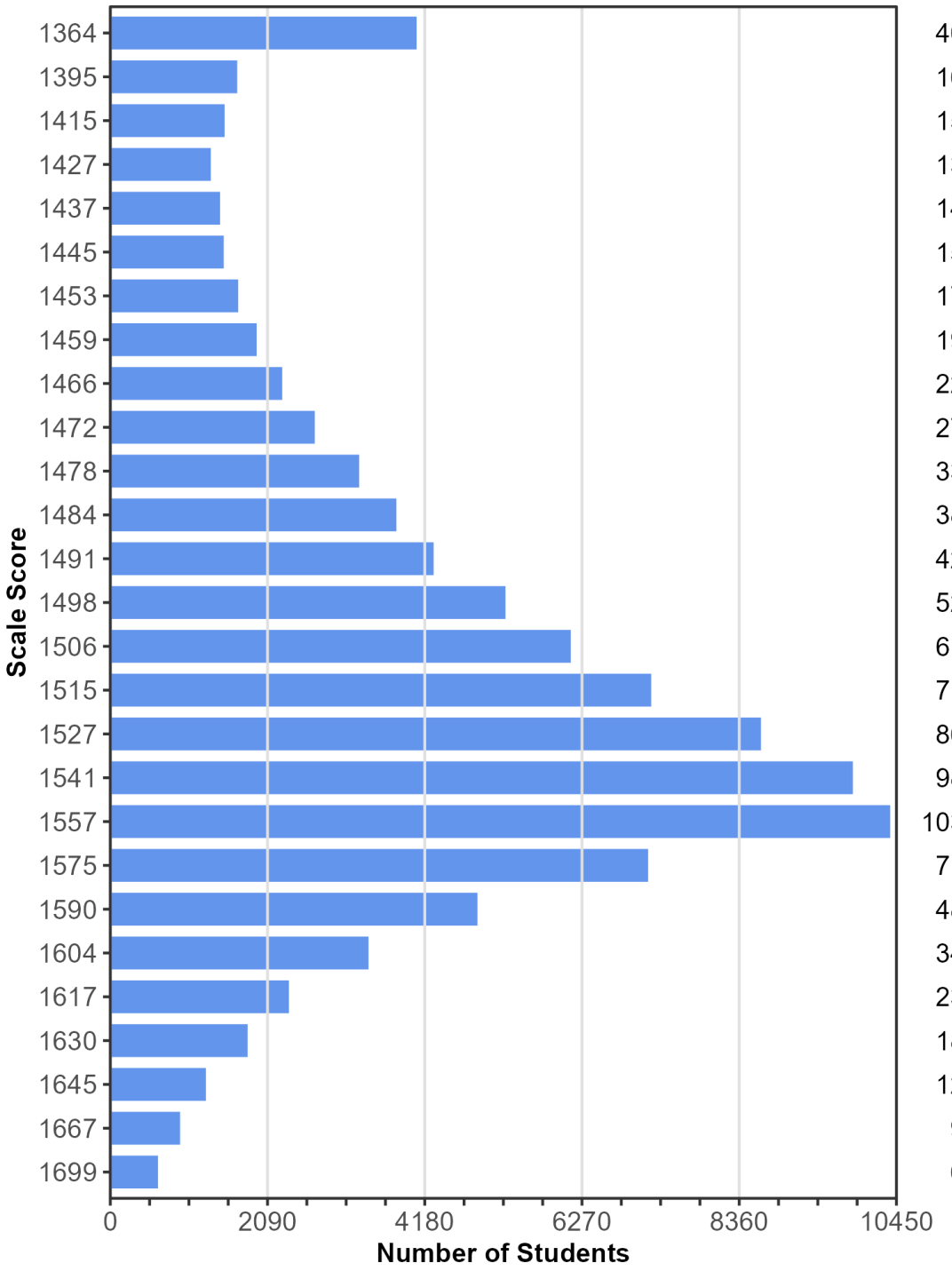
	Freq	Cum Freq	Pct	Cum Pct
1272	7525	7525	7.52	7.52
1317	3781	11306	3.78	11.29
1345	3215	14521	3.21	14.50
1363	2993	17514	2.99	17.49
1377	2930	20444	2.93	20.42
1390	3136	23580	3.13	23.55
1401	3443	27023	3.44	26.99
1411	4128	31151	4.12	31.11
1422	4603	35754	4.60	35.71
1432	5001	40755	4.99	40.70
1442	5461	46216	5.45	46.16
1452	5534	51750	5.53	51.68
1463	5741	57491	5.73	57.42
1475	5886	63377	5.88	63.30
1487	6067	69444	6.06	69.36
1501	5955	75399	5.95	75.30
1517	5873	81272	5.87	81.17
1535	5516	86788	5.51	86.68
1555	4793	91581	4.79	91.46
1577	3116	94697	3.11	94.58
1598	2034	96731	2.03	96.61
1618	1283	98014	1.28	97.89
1639	916	98930	0.91	98.80
1661	559	99489	0.56	99.36
1686	342	99831	0.34	99.70
1721	220	100051	0.22	99.92
1771	76	100127	0.08	100.00

Figure D.5.35. Spring 2023 TELPAS Grade 3 Speaking
Frequency Distribution of Scale Scores
(Online Only)



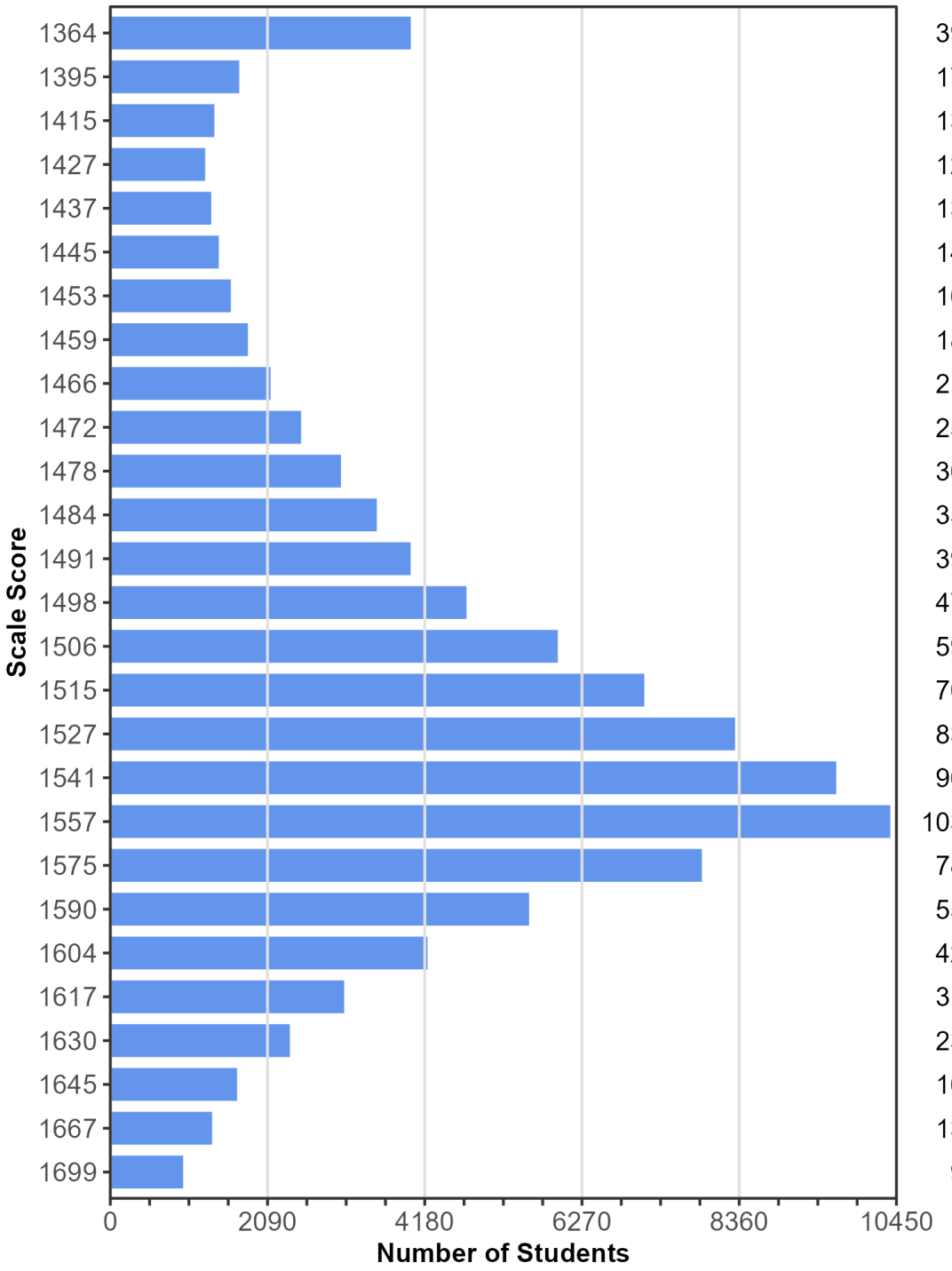
	Freq	Cum Freq	Pct	Cum Pct
1272	4922	4922	4.90	4.90
1317	2359	7281	2.35	7.24
1345	1890	9171	1.88	9.12
1363	1711	10882	1.70	10.83
1377	1772	12654	1.76	12.59
1390	1897	14551	1.89	14.48
1401	2079	16630	2.07	16.54
1411	2438	19068	2.43	18.97
1422	2935	22003	2.92	21.89
1432	3400	25403	3.38	25.27
1442	3978	29381	3.96	29.23
1452	4376	33757	4.35	33.58
1463	5010	38767	4.98	38.56
1475	5560	44327	5.53	44.10
1487	6581	50908	6.55	50.64
1501	7766	58674	7.73	58.37
1517	8215	66889	8.17	66.54
1535	8298	75187	8.25	74.79
1555	8192	83379	8.15	82.94
1577	5768	89147	5.74	88.68
1598	3989	93136	3.97	92.65
1618	2766	95902	2.75	95.40
1639	1855	97757	1.85	97.25
1661	1253	99010	1.25	98.49
1686	866	99876	0.86	99.35
1721	463	100339	0.46	99.81
1771	186	100525	0.19	100.00

Figure D.5.36. Spring 2023 TELPAS Grade 4 Speaking
Frequency Distribution of Scale Scores
(Online Only)



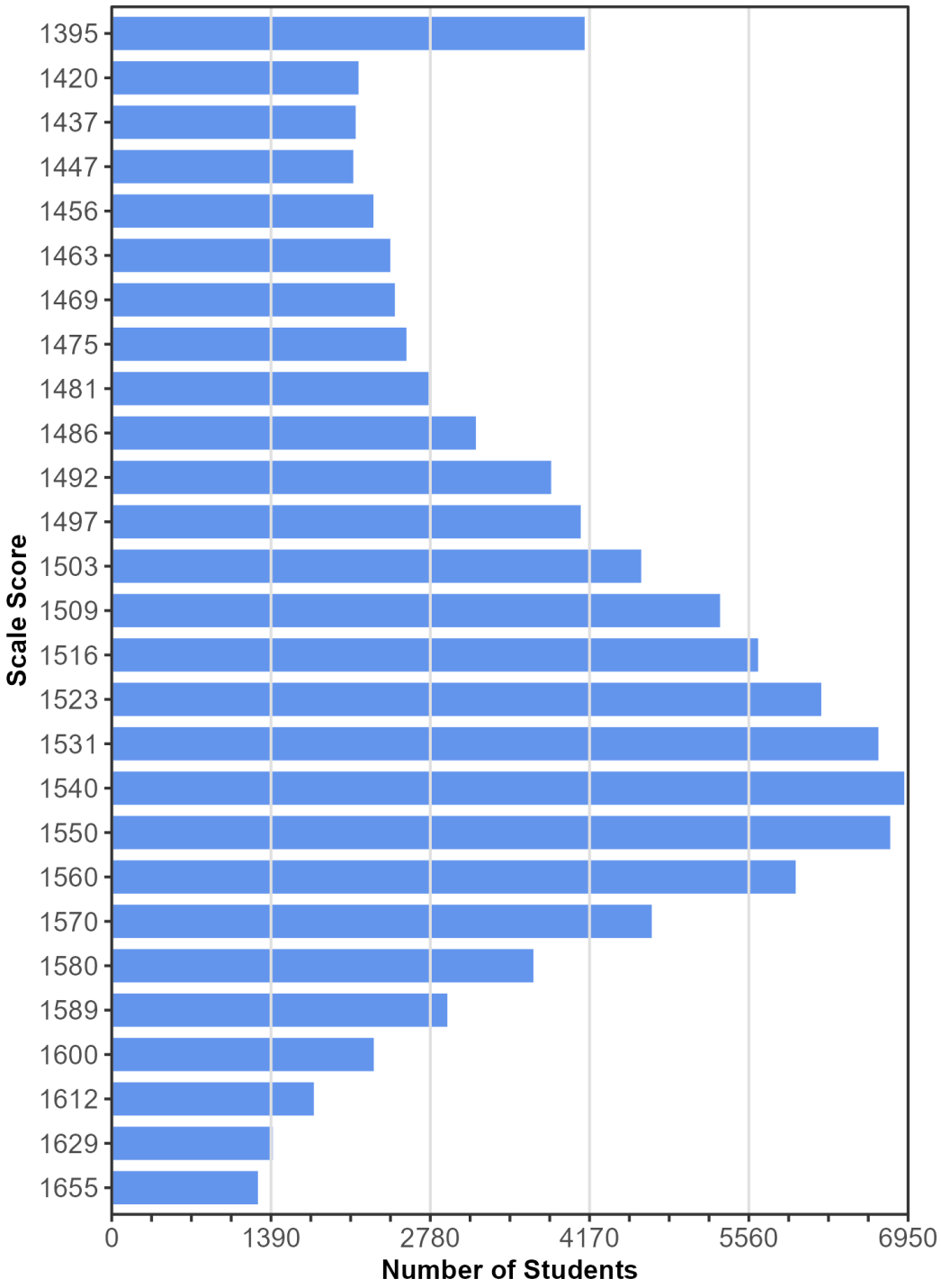
	Freq	Cum Freq	Pct	Cum Pct
1364	4072	4072	4.01	4.01
1395	1687	5759	1.66	5.67
1415	1519	7278	1.50	7.16
1427	1336	8614	1.32	8.48
1437	1459	10073	1.44	9.92
1445	1508	11581	1.48	11.40
1453	1700	13281	1.67	13.07
1459	1945	15226	1.91	14.99
1466	2285	17511	2.25	17.24
1472	2718	20229	2.68	19.91
1478	3308	23537	3.26	23.17
1484	3803	27340	3.74	26.91
1491	4298	31638	4.23	31.14
1498	5253	36891	5.17	36.32
1506	6121	43012	6.03	42.34
1515	7190	50202	7.08	49.42
1527	8649	58851	8.51	57.93
1541	9871	68722	9.72	67.65
1557	10367	79089	10.21	77.86
1575	7149	86238	7.04	84.89
1590	4881	91119	4.80	89.70
1604	3432	94551	3.38	93.08
1617	2374	96925	2.34	95.41
1630	1825	98750	1.80	97.21
1645	1271	100021	1.25	98.46
1667	928	100949	0.91	99.38
1699	634	101583	0.62	100.00

Figure D.5.37. Spring 2023 TELPAS Grade 5 Speaking
Frequency Distribution of Scale Scores
(Online Only)



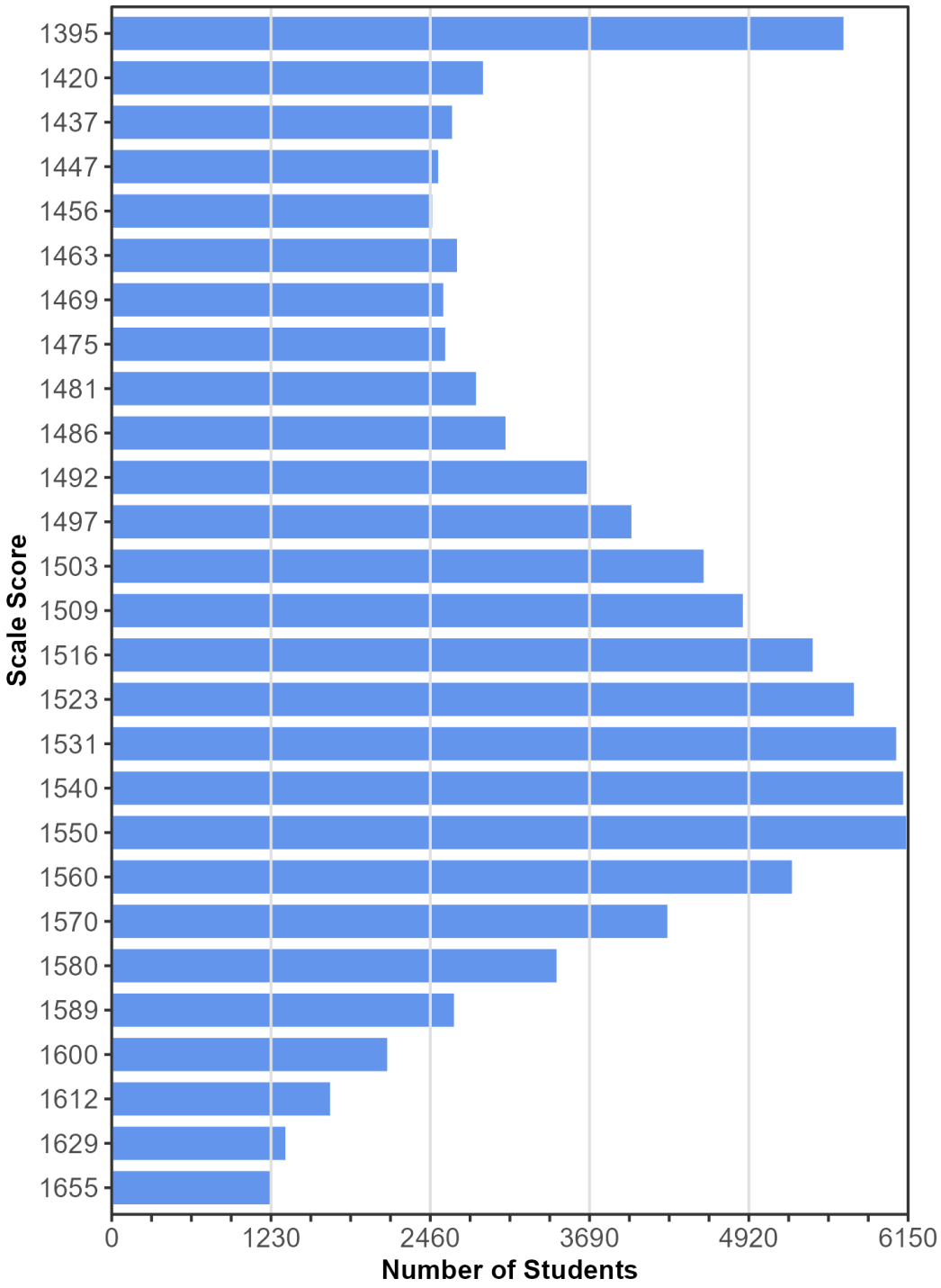
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1364	3994	3994	3.87	3.87
1395	1714	5708	1.66	5.54
1415	1383	7091	1.34	6.88
1427	1261	8352	1.22	8.10
1437	1342	9694	1.30	9.40
1445	1442	11136	1.40	10.80
1453	1603	12739	1.55	12.36
1459	1829	14568	1.77	14.13
1466	2130	16698	2.07	16.20
1472	2536	19234	2.46	18.66
1478	3066	22300	2.97	21.63
1484	3542	25842	3.44	25.07
1491	3992	29834	3.87	28.94
1498	4735	34569	4.59	33.53
1506	5949	40518	5.77	39.30
1515	7099	47617	6.89	46.19
1527	8305	55922	8.06	54.25
1541	9650	65572	9.36	63.61
1557	10369	75941	10.06	73.66
1575	7864	83805	7.63	81.29
1590	5567	89372	5.40	86.69
1604	4216	93588	4.09	90.78
1617	3109	96697	3.02	93.80
1630	2387	99084	2.32	96.11
1645	1685	100769	1.63	97.75
1667	1353	102122	1.31	99.06
1699	969	103091	0.94	100.00

Figure D.5.38. Spring 2023 TELPAS Grade 6 Speaking
Frequency Distribution of Scale Scores
(Online Only)



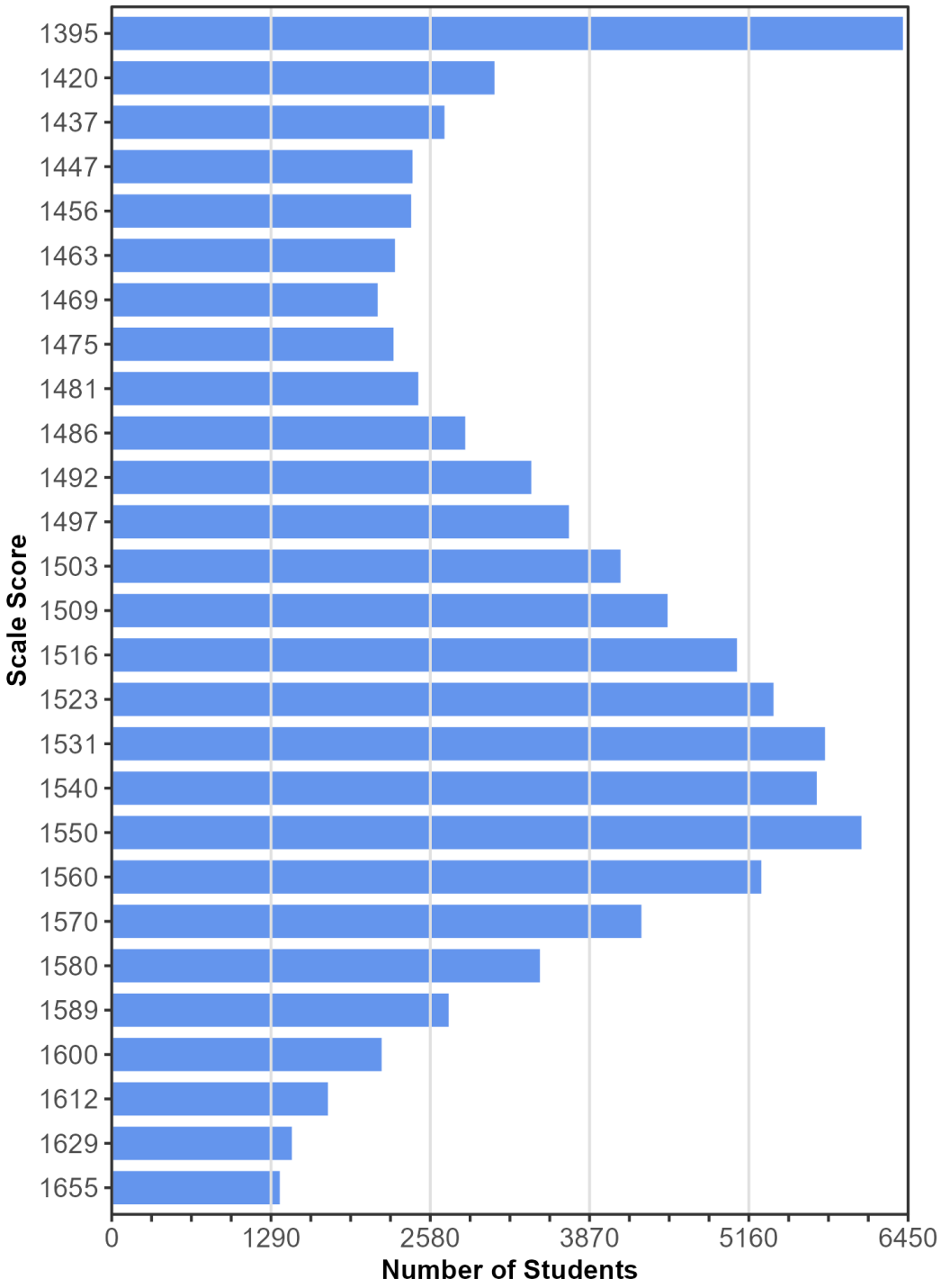
	Freq	Cum Freq	Pct	Cum Pct
1395	4127	4127	4.11	4.11
1420	2154	6281	2.15	6.26
1437	2129	8410	2.12	8.38
1447	2108	10518	2.10	10.48
1456	2284	12802	2.28	12.76
1463	2431	15233	2.42	15.18
1469	2471	17704	2.46	17.64
1475	2572	20276	2.56	20.21
1481	2765	23041	2.76	22.96
1486	3178	26219	3.17	26.13
1492	3835	30054	3.82	29.95
1497	4093	34147	4.08	34.03
1503	4621	38768	4.61	38.64
1509	5309	44077	5.29	43.93
1516	5641	49718	5.62	49.55
1523	6192	55910	6.17	55.72
1531	6691	62601	6.67	62.39
1540	6917	69518	6.89	69.29
1550	6794	76312	6.77	76.06
1560	5969	82281	5.95	82.01
1570	4713	86994	4.70	86.70
1580	3680	90674	3.67	90.37
1589	2929	93603	2.92	93.29
1600	2287	95890	2.28	95.57
1612	1764	97654	1.76	97.33
1629	1405	99059	1.40	98.73
1655	1276	100335	1.27	100.00

Figure D.5.39. Spring 2023 TELPAS Grade 7 Speaking
Frequency Distribution of Scale Scores
(Online Only)



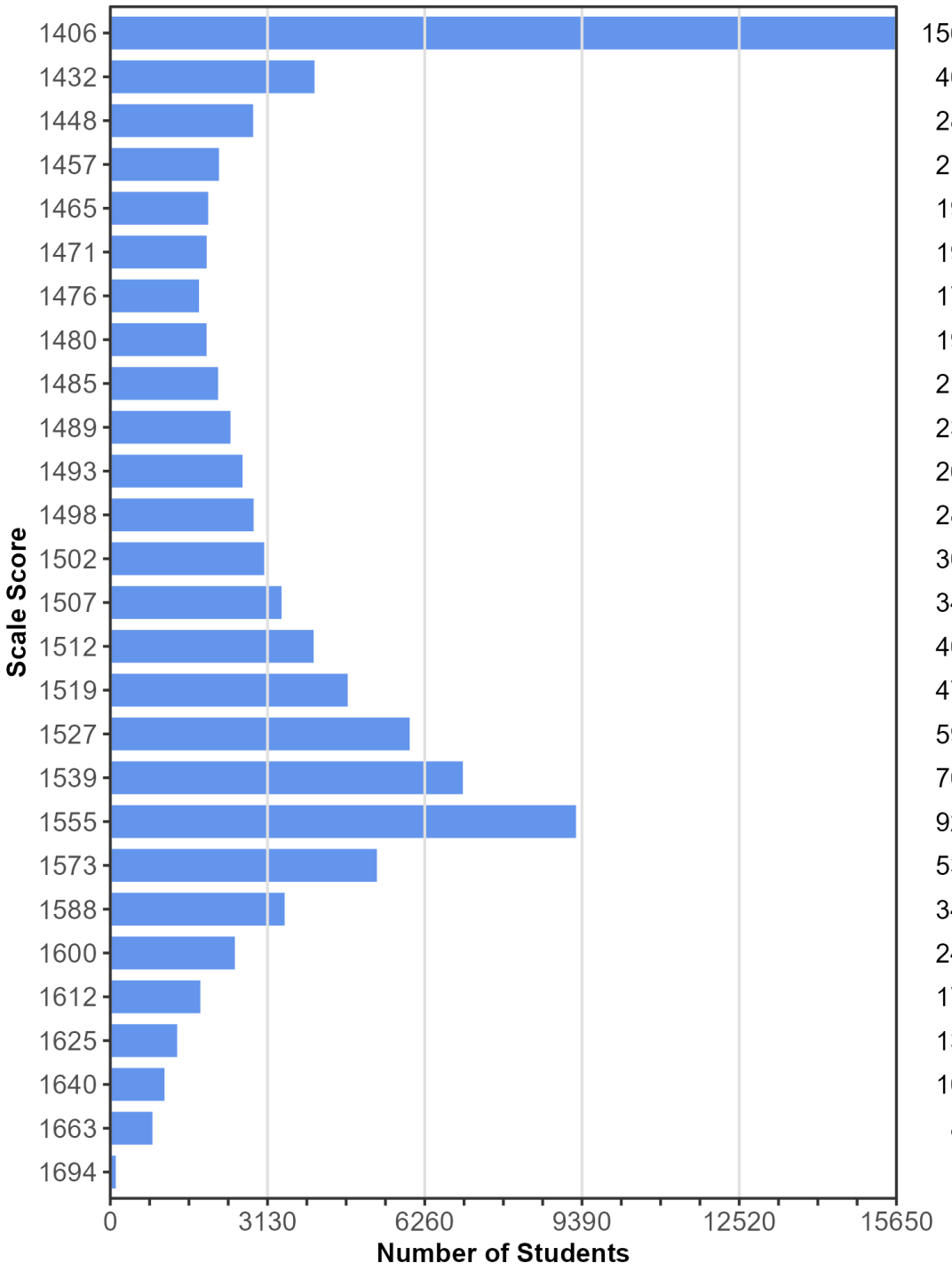
	Freq	Cum Freq	Pct	Cum Pct
1395	5651	5651	5.74	5.74
1420	2867	8518	2.91	8.66
1437	2628	11146	2.67	11.33
1447	2521	13667	2.56	13.89
1456	2474	16141	2.51	16.41
1463	2666	18807	2.71	19.12
1469	2560	21367	2.60	21.72
1475	2575	23942	2.62	24.34
1481	2813	26755	2.86	27.19
1486	3041	29796	3.09	30.29
1492	3669	33465	3.73	34.02
1497	4013	37478	4.08	38.09
1503	4571	42049	4.65	42.74
1509	4874	46923	4.95	47.69
1516	5413	52336	5.50	53.20
1523	5731	58067	5.83	59.02
1531	6058	64125	6.16	65.18
1540	6112	70237	6.21	71.39
1550	6135	76372	6.24	77.63
1560	5253	81625	5.34	82.97
1570	4291	85916	4.36	87.33
1580	3435	89351	3.49	90.82
1589	2643	91994	2.69	93.51
1600	2127	94121	2.16	95.67
1612	1686	95807	1.71	97.38
1629	1341	97148	1.36	98.74
1655	1235	98383	1.26	100.00

Figure D.5.40. Spring 2023 TELPAS Grade 8 Speaking
Frequency Distribution of Scale Scores
(Online Only)



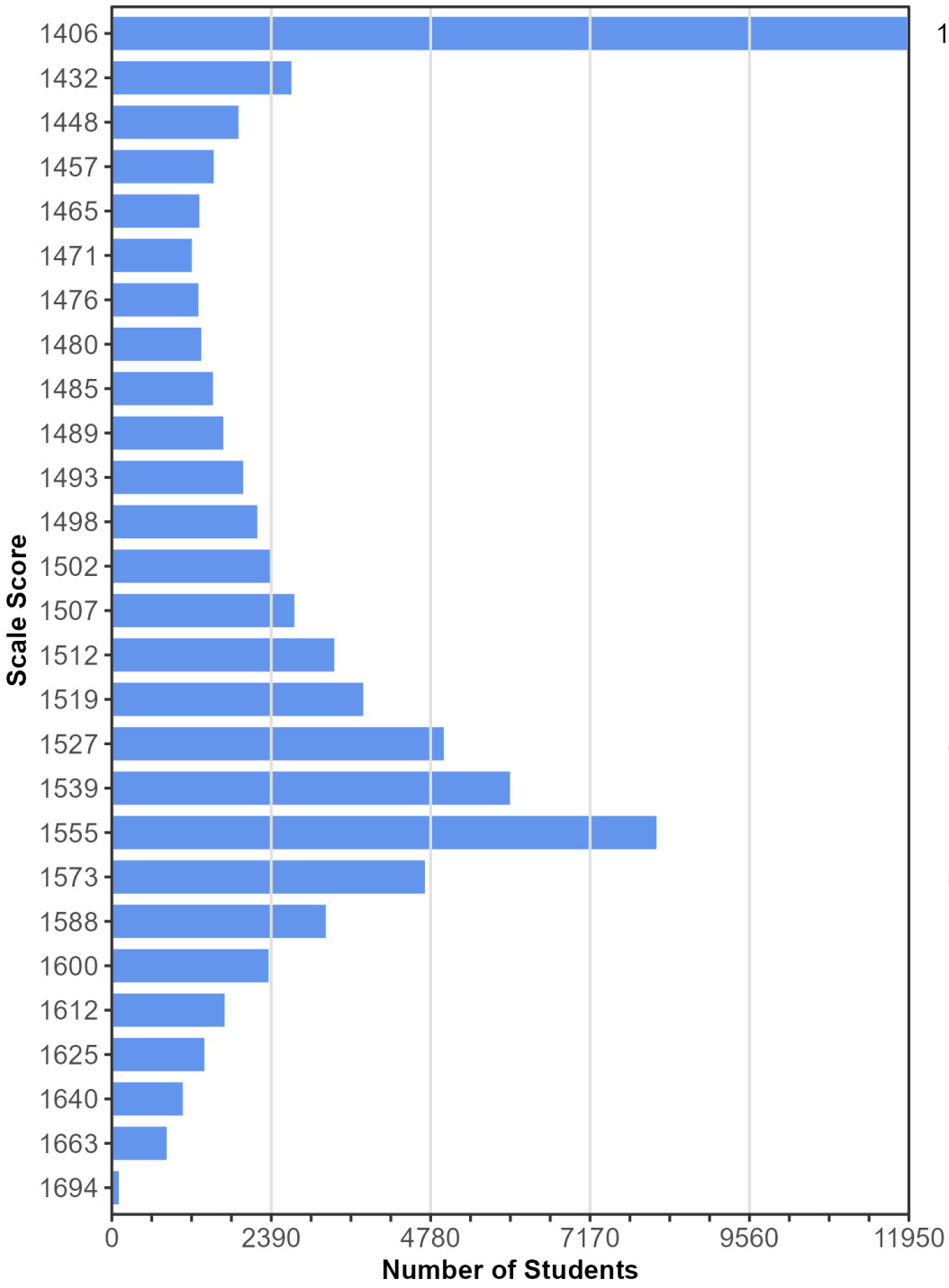
	Freq	Cum Freq	Pct	Cum Pct
1395	6408	6408	6.72	6.72
1420	3100	9508	3.25	9.97
1437	2695	12203	2.83	12.80
1447	2436	14639	2.55	15.35
1456	2425	17064	2.54	17.90
1463	2294	19358	2.41	20.30
1469	2154	21512	2.26	22.56
1475	2282	23794	2.39	24.95
1481	2483	26277	2.60	27.56
1486	2863	29140	3.00	30.56
1492	3398	32538	3.56	34.12
1497	3703	36241	3.88	38.01
1503	4121	40362	4.32	42.33
1509	4502	44864	4.72	47.05
1516	5064	49928	5.31	52.36
1523	5360	55288	5.62	57.98
1531	5777	61065	6.06	64.04
1540	5710	66775	5.99	70.03
1550	6072	72847	6.37	76.40
1560	5261	78108	5.52	81.92
1570	4290	82398	4.50	86.41
1580	3468	85866	3.64	90.05
1589	2729	88595	2.86	92.91
1600	2186	90781	2.29	95.21
1612	1751	92532	1.84	97.04
1629	1459	93991	1.53	98.57
1655	1361	95352	1.43	100.00

Figure D.5.41. Spring 2023 TELPAS Grade 9 Speaking
Frequency Distribution of Scale Scores
(Online Only)



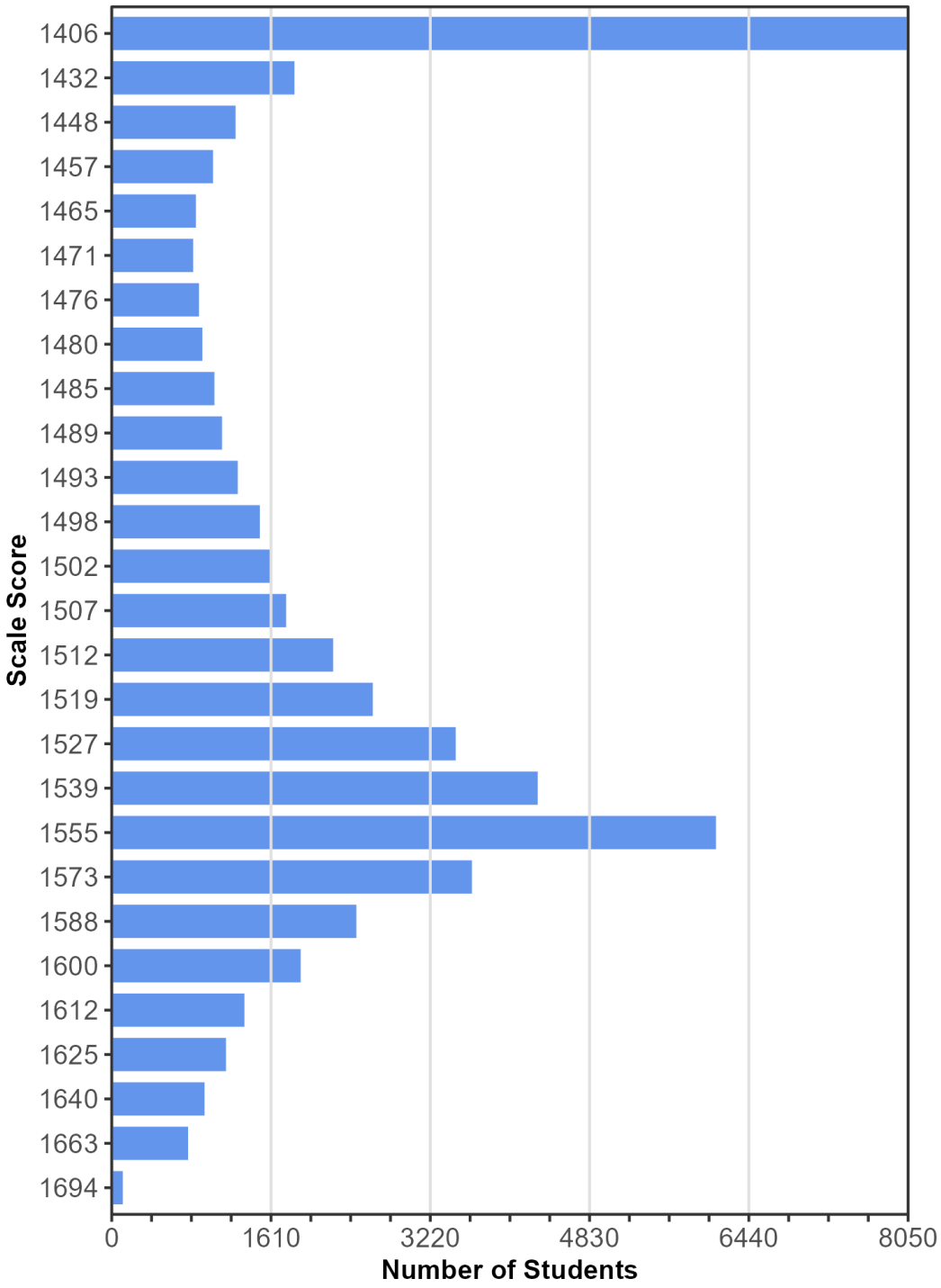
	Freq	Cum Freq	Pct	Cum Pct
1406	15647	15647	16.26	16.26
1432	4066	19713	4.23	20.49
1448	2844	22557	2.96	23.45
1457	2164	24721	2.25	25.70
1465	1949	26670	2.03	27.72
1471	1919	28589	1.99	29.72
1476	1766	30355	1.84	31.55
1480	1917	32272	1.99	33.54
1485	2146	34418	2.23	35.78
1489	2393	36811	2.49	38.26
1493	2633	39444	2.74	41.00
1498	2853	42297	2.97	43.97
1502	3064	45361	3.18	47.15
1507	3409	48770	3.54	50.69
1512	4048	52818	4.21	54.90
1519	4727	57545	4.91	59.81
1527	5960	63505	6.20	66.01
1539	7019	70524	7.30	73.31
1555	9272	79796	9.64	82.94
1573	5307	85103	5.52	88.46
1588	3471	88574	3.61	92.07
1600	2480	91054	2.58	94.64
1612	1793	92847	1.86	96.51
1625	1331	94178	1.38	97.89
1640	1079	95257	1.12	99.01
1663	840	96097	0.87	99.89
1694	109	96206	0.11	100.00

Figure D.5.42. Spring 2023 TELPAS Grade 10 Speaking
Frequency Distribution of Scale Scores
(Online Only)



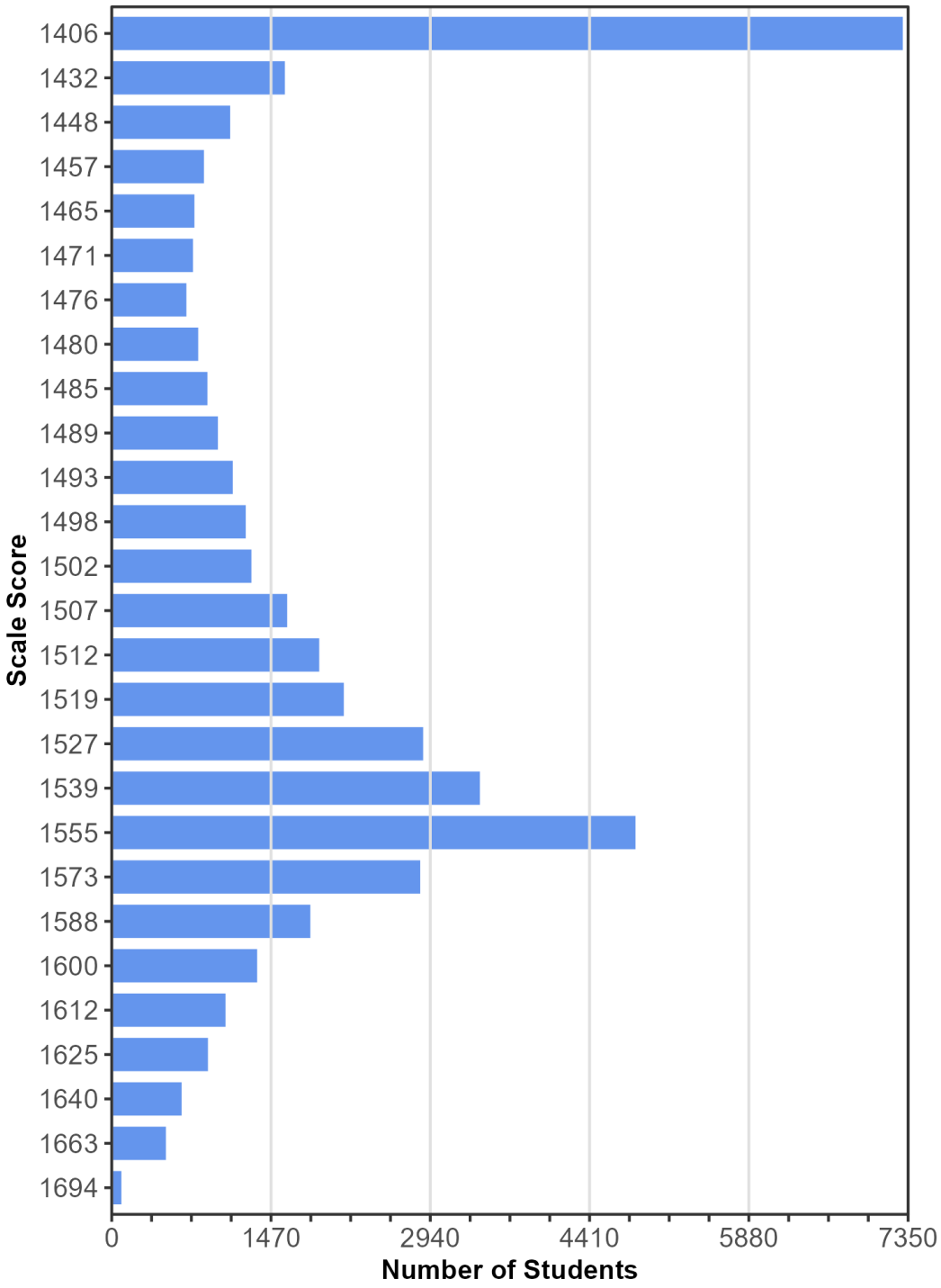
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1406	11922	11922	15.44	15.44
1432	2693	14615	3.49	18.93
1448	1899	16514	2.46	21.39
1457	1527	18041	1.98	23.37
1465	1312	19353	1.70	25.07
1471	1200	20553	1.55	26.62
1476	1298	21851	1.68	28.31
1480	1342	23193	1.74	30.04
1485	1517	24710	1.97	32.01
1489	1672	26382	2.17	34.18
1493	1970	28352	2.55	36.73
1498	2182	30534	2.83	39.55
1502	2384	32918	3.09	42.64
1507	2737	35655	3.55	46.19
1512	3335	38990	4.32	50.51
1519	3769	42759	4.88	55.39
1527	4977	47736	6.45	61.84
1539	5970	53706	7.73	69.57
1555	8166	61872	10.58	80.15
1573	4694	66566	6.08	86.23
1588	3208	69774	4.16	90.39
1600	2350	72124	3.04	93.43
1612	1691	73815	2.19	95.62
1625	1387	75202	1.80	97.42
1640	1064	76266	1.38	98.80
1663	823	77089	1.07	99.86
1694	106	77195	0.14	100.00

Figure D.5.43. Spring 2023 TELPAS Grade 11 Speaking
Frequency Distribution of Scale Scores
(Online Only)



Scale Score	Freq	Cum Freq	Pct	Cum Pct
1406	8049	8049	14.62	14.62
1432	1847	9896	3.35	17.97
1448	1252	11148	2.27	20.25
1457	1024	12172	1.86	22.11
1465	851	13023	1.55	23.65
1471	823	13846	1.49	25.15
1476	882	14728	1.60	26.75
1480	916	15644	1.66	28.41
1485	1038	16682	1.89	30.30
1489	1115	17797	2.03	32.33
1493	1274	19071	2.31	34.64
1498	1497	20568	2.72	37.36
1502	1615	22183	2.93	40.29
1507	1763	23946	3.20	43.49
1512	2238	26184	4.06	47.56
1519	2639	28823	4.79	52.35
1527	3477	32300	6.32	58.67
1539	4306	36606	7.82	66.49
1555	6108	42714	11.09	77.58
1573	3641	46355	6.61	84.20
1588	2473	48828	4.49	88.69
1600	1910	50738	3.47	92.16
1612	1341	52079	2.44	94.59
1625	1155	53234	2.10	96.69
1640	938	54172	1.70	98.39
1663	772	54944	1.40	99.80
1694	112	55056	0.20	100.00

Figure D.5.44. Spring 2023 TELPAS Grade 12 Speaking
Frequency Distribution of Scale Scores
(Online Only)



	Freq	Cum Freq	Pct	Cum Pct
1406	7301	7301	16.11	16.11
1432	1598	8899	3.53	19.63
1448	1093	9992	2.41	22.05
1457	851	10843	1.88	23.92
1465	763	11606	1.68	25.61
1471	750	12356	1.65	27.26
1476	689	13045	1.52	28.78
1480	798	13843	1.76	30.54
1485	883	14726	1.95	32.49
1489	980	15706	2.16	34.65
1493	1117	16823	2.46	37.12
1498	1237	18060	2.73	39.85
1502	1289	19349	2.84	42.69
1507	1619	20968	3.57	46.26
1512	1915	22883	4.23	50.49
1519	2142	25025	4.73	55.21
1527	2874	27899	6.34	61.55
1539	3398	31297	7.50	69.05
1555	4834	36131	10.67	79.72
1573	2847	38978	6.28	86.00
1588	1833	40811	4.04	90.04
1600	1342	42153	2.96	93.00
1612	1050	43203	2.32	95.32
1625	888	44091	1.96	97.28
1640	645	44736	1.42	98.70
1663	500	45236	1.10	99.80
1694	89	45325	0.20	100.00

2023 TELPAS Distribution of Proficiency Levels

Table D.6.1. Distribution of TELPAS Proficiency Levels by Administration Since 2018 Redesign - Grades K–6

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
K	2023	96717	42	31	18	9	96519	50	27	15	8	96425	65	18	11	7	96449	67	18	9	6	96183	49	31	13	7
	2022	95188	42	31	18	10	95119	49	27	16	8	95045	66	17	10	7	95060	69	17	9	5	95061	48	32	13	7
	2021	91494	42	32	18	8	91420	49	29	16	7	91352	66	18	10	6	91244	68	18	9	5	91243	48	33	13	6
	2020	63415	40	33	18	9	63359	47	30	16	7	63267	64	20	11	6	63119	68	20	8	4	63065	47	35	13	6
	2019	96646	36	34	20	10	96578	44	31	17	8	96485	61	20	12	8	96443	64	21	10	5	96506	43	36	15	7
	2018	98802	36	34	20	10	98702	44	31	17	8	98560	60	21	12	7	98548	64	21	10	5	98634	43	35	15	7
1	2023	103723	21	34	26	19	103510	30	32	22	16	103341	44	27	16	14	103338	48	28	14	11	103187	28	37	21	14
	2022	100132	20	34	27	19	100028	28	33	24	16	99945	44	27	16	13	99940	48	28	14	10	99947	26	38	22	14
	2021	97267	19	35	28	18	97191	26	34	25	15	97110	43	28	17	12	97032	46	29	15	9	97043	25	39	23	13
	2020	68737	16	34	29	21	68689	23	34	26	18	68590	36	31	18	15	68466	40	32	16	12	68431	21	38	25	16
	2019	104690	13	34	31	22	104627	20	35	27	18	104473	33	31	20	16	104478	37	33	18	12	104571	18	39	26	17
	2018	106800	13	33	31	23	106725	20	35	27	18	106546	32	31	21	16	106551	36	34	19	12	106682	18	38	27	17
2	2023	100204	11	30	27	33	100196	27	54	15	3	100259	38	41	15	5	100251	38	40	19	3	100158	17	55	25	3
	2022	96840	9	24	41	27	96835	22	56	20	3	96891	36	36	21	8	96403	32	36	22	11	96048	13	51	31	5
	2021	87172	4	24	38	33	87170	22	60	16	2	87140	34	34	22	10	93529	28	39	23	9	85042	11	51	34	4
	2020	72327	6	24	42	28	72324	18	54	19	10	80076	22	42	24	11	49161	23	38	26	13	40821	8	46	37	7
	2019	103082	7	24	32	37	103073	13	50	28	10	103092	23	37	27	13	102483	20	38	27	14	102389	7	44	39	10
	2018	105966	5	30	41	24	105905	8	52	34	7	105681	18	41	26	14	105275	20	38	27	15	105333	4	45	42	9
3	2023	100597	5	15	20	60	100591	17	50	26	7	100651	26	28	21	24	100650	29	40	25	6	100563	10	40	38	12
	2022	98822	4	11	33	52	98817	14	49	31	6	98876	20	32	23	24	98362	19	35	29	17	98031	6	37	42	15
	2021	88856	2	11	29	59	88854	12	55	26	6	88891	24	31	18	27	95625	16	37	31	16	86792	5	39	43	13
	2020	74431	3	11	32	54	74428	10	44	27	19	82031	18	26	23	33	48948	14	32	32	22	41270	5	29	44	22
	2019	103773	4	12	25	60	103768	8	41	34	17	103787	14	30	25	30	103232	12	34	33	21	103124	3	30	45	22
	2018	106617	2	15	36	47	106581	5	43	40	11	106296	12	32	26	30	105846	12	33	33	22	105987	2	30	48	20
4	2023	101659	16	25	34	25	101649	17	32	40	10	101693	14	32	22	32	101692	23	41	31	5	101633	11	36	39	14
	2022	100355	15	28	37	21	100353	15	37	42	6	100375	13	36	24	27	99884	11	28	34	28	99607	7	34	42	16
	2021	86990	12	31	38	19	86988	14	37	44	5	87046	10	39	23	28	93711	10	30	35	25	84837	5	36	44	15
	2020	75670	11	29	35	25	75665	8	35	48	9	81907	13	29	24	34	48764	8	25	35	32	42546	5	29	45	22
	2019	100655	13	35	38	15	100652	10	33	37	20	100691	11	34	26	29	100124	8	26	36	30	100025	5	33	43	19
	2018	96904	8	33	42	17	96860	6	33	51	10	96542	7	32	35	26	96118	8	26	36	30	96285	3	29	50	18
5	2023	103145	11	19	33	38	103142	16	30	41	13	103176	8	23	21	47	103176	14	35	41	10	103100	8	27	41	24
	2022	98982	10	21	37	31	98979	14	35	44	7	99031	9	27	25	40	98489	8	21	34	37	98216	5	27	44	24
	2021	84508	8	22	38	31	84505	14	34	46	6	84611	7	28	23	43	90745	6	24	36	34	82234	4	27	45	24
	2020	74409	7	20	34	39	74410	9	33	47	11	79738	9	22	21	48	48255	6	20	34	40	42024	4	20	44	31
	2019	91372	8	26	41	25	91367	12	30	35	23	91365	7	25	26	43	90878	6	21	35	38	90784	4	25	43	29
	2018	88040	6	23	43	28	88016	7	30	51	13	87750	5	23	34	39	87370	6	21	36	38	87445	2	21	49	28
6	2023	100388	5	30	33	32	100383	13	43	38	7	100455	16	28	31	25	100455	16	37	39	8	100287	6	35	44	14
	2022	94690	5	26	31	38	94688	11	45	42	3	94763	15	31	27	26	94162	7	21	35	37	93685	4	31	46	19
	2021	76394	5	24	34	37	76391	14	47	35	3	76468	15	28	29	28	79496	5	24	37	34	70383	3	31	48	17
	2020	59588	4	18	33	45	59587	11	43	40	6	65722	13	30	31	26	38114	6	20	35	39	30781	3	24	49	23
	2019	79561	3	19	42	35	79560	14	52	28	6	79624	14	32	30	24	79158	6	22	37	36	78995	3	31	51	15
	2018	70933	3	22	43	32	70910	6	45	44	5	70776	10	32	34	24	70378	6	22	37	36	70312	2	27	53	19

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.2. Distribution of TELPAS Proficiency Levels by Administration Since 2018 Redesign - Grades 7–12

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
7	2023	98437	5	27	31	37	98432	16	43	34	7	98513	14	25	30	31	98514	14	33	42	11	98304	6	33	44	17
	2022	92519	5	21	29	46	92516	14	46	38	3	92640	14	28	26	33	92026	6	20	35	39	91378	4	29	46	21
	2021	65716	5	21	33	41	65714	17	49	31	3	65900	14	25	29	32	68746	5	22	37	36	59926	4	31	48	17
	2020	55073	4	15	28	53	55073	15	44	36	5	61736	13	27	30	30	35106	5	18	35	41	28525	4	25	48	24
	2019	68708	4	18	38	40	68707	19	51	25	5	68765	14	28	28	30	68232	6	21	36	37	68059	4	33	49	15
	2018	63392	4	20	41	35	63382	9	46	41	4	63295	10	30	33	28	62882	6	21	37	37	62784	2	26	52	19
8	2023	95401	5	23	29	44	95395	18	40	35	7	95483	7	31	31	32	95480	16	43	33	8	95248	5	35	43	17
	2022	81952	5	19	26	51	81950	15	45	37	3	82034	10	35	32	23	81430	5	17	34	43	80805	4	28	47	21
	2021	59699	5	19	29	47	59697	19	49	30	3	59686	11	34	32	23	63182	4	19	36	41	54427	4	30	49	17
	2020	47523	4	13	23	60	47521	16	43	35	6	52952	6	37	33	24	31222	5	17	33	45	25169	3	25	48	24
	2019	61716	4	16	35	45	61713	20	50	25	5	61777	10	35	35	20	61194	5	18	35	42	61043	3	32	50	15
	2018	54294	4	18	38	40	54287	10	43	42	5	54274	7	39	37	17	53857	5	18	35	41	53661	2	26	53	19
9	2023	96472	11	27	41	21	96467	34	26	32	8	96465	10	33	29	29	96466	22	41	30	7	95757	11	40	36	13
	2022	84571	14	25	40	20	84570	33	32	31	4	84774	15	35	30	21	83291	10	23	34	33	80554	9	39	38	14
	2021	51951	8	24	47	20	51947	32	28	33	7	52041	13	33	30	25	52767	6	24	37	33	44717	6	37	42	16
	2020	51988	13	28	34	25	51987	37	30	27	5	54275	9	40	30	21	28957	9	22	34	35	25309	7	42	36	15
	2019	59413	14	33	35	19	59411	35	31	27	6	59437	14	36	32	18	58081	9	24	36	32	57743	9	43	36	12
	2018	50419	10	31	43	15	50411	25	29	40	5	50356	10	42	33	15	49793	10	25	36	30	49155	6	39	43	12
10	2023	77503	9	25	42	24	77502	30	25	35	10	77622	10	42	26	23	77616	20	33	36	11	76881	9	40	36	15
	2022	59135	10	24	42	24	59133	30	29	35	6	59276	12	42	27	20	58697	5	21	36	37	56754	6	39	40	15
	2021	43063	8	24	47	21	43063	31	27	34	8	43139	13	39	27	20	44943	4	22	37	36	36932	6	37	42	16
	2020	37263	9	26	34	30	37263	34	29	30	7	39378	12	43	27	18	21200	5	22	36	37	18527	5	42	37	16
	2019	42307	11	32	36	21	42305	32	31	29	8	42349	13	43	27	17	41574	5	24	38	33	41419	6	44	37	13
	2018	37588	8	31	44	17	37583	21	28	44	7	37526	10	43	35	12	37414	6	25	36	32	36927	4	37	46	12
11	2023	55395	7	24	42	27	55394	29	24	36	11	55459	8	40	27	26	55460	17	32	38	12	54907	7	39	38	17
	2022	48788	9	23	44	24	48788	29	29	36	6	48977	10	41	28	21	48568	4	20	36	41	46816	5	38	41	16
	2021	30728	7	23	48	22	30728	29	26	36	10	30764	11	40	28	21	32516	3	20	39	39	26291	4	36	43	17
	2020	25912	7	25	36	31	25910	32	28	32	8	27759	10	42	29	20	15365	3	20	37	40	13325	3	41	38	18
	2019	31768	10	32	36	23	31766	29	30	31	10	31819	11	41	28	20	21285	3	22	39	36	31130	4	42	38	16
	2018	28098	5	27	47	21	28088	17	25	49	9	28058	7	39	39	15	28046	3	21	38	38	27709	2	31	50	16
12	2023	45346	6	26	44	24	45346	31	25	35	10	45476	7	41	28	23	45476	16	34	38	11	45021	6	42	38	14
	2022	36047	8	25	45	22	36045	30	29	35	6	36226	10	42	29	19	36155	2	17	37	44	34588	3	40	41	15
	2021	22276	6	25	50	20	22274	29	26	36	9	22326	10	42	29	19	24037	2	17	38	43	19121	3	37	44	16
	2020	20377	7	28	36	29	20376	32	27	33	8	21867	9	44	28	19	11210	2	18	38	42	9564	3	42	39	16
	2019	25315	9	33	37	21	25315	28	29	32	10	25307	10	41	29	19	24871	2	19	39	40	24773	3	41	40	16
	2018	21431	4	27	49	21	21429	16	23	51	10	21387	6	38	39	16	21422	2	18	37	43	21095	2	29	52	18

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.3. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades K–1

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
K	2023	96717	42	31	18	9	96519	50	27	15	8	96425	65	18	11	7	96449	67	18	9	6	96183	49	31	13	7
	2022	95188	42	31	18	10	95119	49	27	16	8	95045	66	17	10	7	95060	69	17	9	5	95061	48	32	13	7
	2021	91494	42	32	18	8	91420	49	29	16	7	91352	66	18	10	6	91244	68	18	9	5	91243	48	33	13	6
	2020	63415	40	33	18	9	63359	47	30	16	7	63267	64	20	11	6	63119	68	20	8	4	63065	47	35	13	6
	2019	96646	36	34	20	10	96578	44	31	17	8	96485	61	20	12	8	96443	64	21	10	5	96506	43	36	15	7
	2018	98802	36	34	20	10	98702	44	31	17	8	98560	60	21	12	7	98548	64	21	10	5	98634	43	35	15	7
	2017	102005	37	35	20	9	101932	45	31	16	7	101759	61	21	12	7	101757	64	22	10	5	101689	58	23	12	6
	2016	103696	37	34	19	9	103604	46	30	16	7	103206	60	21	12	6	103001	64	22	10	5	102778	58	23	13	6
	2015	108256	39	34	18	9	108177	48	30	15	7	108042	62	20	11	6	108016	65	21	9	5	107987	59	23	12	6
	2014	108586	40	33	18	9	108500	49	29	15	7	108388	62	20	11	7	108377	65	20	9	5	108348	59	23	12	6
	2013	108411	40	32	18	9	108341	49	29	15	7	108185	62	20	11	7	108180	65	20	9	5	108143	62	20	11	6
	2012	106179	40	33	18	9	106122	49	28	15	7	106007	62	20	12	7	106002	65	20	10	5	105976	62	20	12	7
	2011	105706	40	33	18	9	105632	49	29	15	7	105530	62	20	11	7	105434	65	20	9	5	105401	62	20	11	6
	2010	104027	40	32	18	10	103966	49	28	16	8	103838	62	19	12	7	103721	66	19	10	5	103685	62	19	12	7
	2009	103037	42	31	18	9	102964	50	27	15	7	102556	63	19	11	7	102718	67	19	9	5	102390	63	19	11	7
	2008	100840	43	31	17	9	100803	51	27	14	7	100678	64	19	11	7	100625	67	18	9	5	100518	64	19	11	6
	2007	97932	45	31	16	7	97893	54	27	13	6	97738	67	18	10	5	97491	70	18	8	4	97426	67	18	10	5
	2006	94307	50	30	14	6	94264	57	26	12	5	94137	70	17	9	5	93526	73	17	7	3	93434	70	17	9	4
	2005	89002	48	32	14	6	88946	57	27	12	4	88831	69	18	9	4	86831	72	17	7	3	86756	69	18	9	4
1	2023	103723	21	34	26	19	103510	30	32	22	16	103341	44	27	16	14	103338	48	28	14	11	103187	28	37	21	14
	2022	100132	20	34	27	19	100028	28	33	24	16	99945	44	27	16	13	99940	48	28	14	10	99947	26	38	22	14
	2021	97267	19	35	28	18	97191	26	34	25	15	97110	43	28	17	12	97032	46	29	15	9	97043	25	39	23	13
	2020	68737	16	34	29	21	68689	23	34	26	18	68590	36	31	18	15	68466	40	32	16	12	68431	21	38	25	16
	2019	104690	13	34	31	22	104627	20	35	27	18	104473	33	31	20	16	104478	37	33	18	12	104571	18	39	26	17
	2018	106800	13	33	31	23	106725	20	35	27	18	106546	32	31	21	16	106551	36	34	19	12	106682	18	38	27	17
	2017	110611	13	34	31	21	110513	20	36	27	17	110254	31	32	21	16	110239	36	34	18	12	110169	29	34	22	15
	2016	115126	13	35	31	21	115039	20	36	27	17	114477	31	32	21	16	114454	35	35	18	11	114193	28	34	22	15
	2015	115881	14	35	31	20	115779	22	36	26	16	115566	32	32	21	15	115548	37	34	18	11	115522	29	34	22	15
	2014	115003	14	35	31	20	114908	22	36	26	16	114703	32	32	21	15	114682	36	35	18	11	114646	29	34	23	15
	2013	112194	14	35	31	20	112102	22	36	26	16	111919	32	32	21	15	111893	37	34	18	11	111858	32	32	22	15
	2012	110885	15	36	30	20	110788	23	36	26	16	110608	32	32	21	15	110591	37	34	18	11	110562	32	32	21	15
	2011	109990	15	35	30	20	109907	23	35	25	16	109727	32	31	21	16	109662	38	33	18	11	109626	32	31	21	16
	2010	109134	17	34	30	20	109054	25	34	25	16	108870	34	29	21	16	108736	40	32	18	11	108694	34	29	21	15
	2009	107003	19	34	28	18	106947	28	33	24	15	106613	36	28	20	15	106714	42	30	17	10	106466	36	28	21	15
	2008	105515	21	35	28	17	105464	29	34	24	14	105235	37	28	20	14	105181	43	30	17	10	105049	37	29	20	14
	2007	101437	24	35	26	14	101398	33	33	22	11	101209	42	27	19	12	101063	48	29	15	8	100971	42	27	19	12
	2006	96333	28	37	24	11	96303	37	34	20	9	96197	45	27	17	10	95704	53	28	13	6	95585	45	27	17	10
	2005	92341	29	38	23	9	92308	38	35	19	7	92134	47	28	17	9	92206	53	30	13	5	92040	47	28	17	9

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.4. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades 2–3

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
2	2023	100204	11	30	27	33	100196	27	54	15	3	100259	38	41	15	5	100251	38	40	19	3	100158	17	55	25	3
	2022	96840	9	24	41	27	96835	22	56	20	3	96891	36	36	21	8	96403	32	36	22	11	96048	13	51	31	5
	2021	87172	4	24	38	33	87170	22	60	16	2	87140	34	34	22	10	93529	28	39	23	9	85042	11	51	34	4
	2020	72327	6	24	42	28	72324	18	54	19	10	80076	22	42	24	11	49161	23	38	26	13	40821	8	46	37	7
	2019	103082	7	24	32	37	103073	13	50	28	10	103092	23	37	27	13	102483	20	38	27	14	102389	7	44	39	10
	2018	105966	5	30	41	24	105905	8	52	34	7	105681	18	41	26	14	105275	20	38	27	15	105333	4	45	42	9
	2017	111398	6	22	37	36	111287	10	28	35	28	111225	20	37	24	19	110855	18	37	29	17	110581	12	36	32	20
	2016	112211	6	22	37	36	112102	10	28	35	28	111931	19	40	25	16	111607	18	37	28	17	111301	12	37	33	19
	2015	111676	6	22	37	35	111579	10	28	34	27	111378	18	39	25	18	111176	18	38	28	16	110898	11	36	33	20
	2014	108908	6	22	37	35	108808	9	29	34	27	108796	20	34	29	17	108423	18	38	28	16	108179	12	34	35	19
	2013	106743	6	23	37	34	106649	9	29	34	27	106663	11	26	29	34	106291	18	38	28	16	106071	11	27	30	32
	2012	104783	6	23	37	34	104690	10	29	34	27	104697	10	26	29	35	104275	18	38	27	17	104006	10	27	31	33
	2011	105358	6	24	36	34	105262	10	29	34	27	105043	8	27	30	34	104818	19	37	28	17	104434	8	27	32	32
	2010	103418	8	25	35	32	103351	13	29	32	25	102852	11	28	27	34	102938	22	36	26	16	102369	11	28	29	32
	2009	102076	9	26	35	30	102016	14	30	32	24	101619	12	30	28	30	101586	23	35	25	16	101151	12	31	30	28
	2008	99042	11	29	34	26	98992	16	31	31	21	98555	14	29	27	29	98499	26	36	24	14	97445	14	30	30	27
	2007	93409	12	30	35	23	93359	18	32	31	19	91953	23	32	27	18	92971	27	37	24	12	91648	23	32	27	18
	2006	89751	14	31	34	21	89697	20	33	30	17	88727	25	33	26	16	89347	31	37	22	10	88424	25	33	26	16
	2005	87330	18	37	31	13	87297	25	37	27	11	87198	31	35	23	10	87138	41	37	17	6	87072	31	35	23	10
3	2023	100597	5	15	20	60	100591	17	50	26	7	100651	26	28	21	24	100650	29	40	25	6	100563	10	40	38	12
	2022	98822	4	11	33	52	98817	14	49	31	6	98876	20	32	23	24	98362	19	35	29	17	98031	6	37	42	15
	2021	88856	2	11	29	59	88854	12	55	26	6	88891	24	31	18	27	95625	16	37	31	16	86792	5	39	43	13
	2020	74431	3	11	32	54	74428	10	44	27	19	82031	18	26	23	33	48948	14	32	32	22	41270	5	29	44	22
	2019	103773	4	12	25	60	103768	8	41	34	17	103787	14	30	25	30	103232	12	34	33	21	103124	3	30	45	22
	2018	106617	2	15	36	47	106581	5	43	40	11	106296	12	32	26	30	105846	12	33	33	22	105987	2	30	48	20
	2017	108332	4	14	32	51	108237	6	20	34	41	108022	17	26	29	28	107650	11	31	34	25	107409	8	25	35	31
	2016	108054	4	13	32	50	107954	6	20	35	40	107804	14	26	33	28	107470	11	32	34	24	107216	7	24	37	31
	2015	105930	4	14	33	50	105827	6	20	35	40	105624	17	25	29	29	105408	11	31	34	24	105183	8	25	35	31
	2014	102961	3	13	33	51	102874	5	20	35	40	102835	16	24	30	30	102474	10	31	34	25	102241	7	25	36	32
	2013	100251	3	14	33	50	100182	5	20	35	39	100051	9	17	23	51	99752	10	31	34	24	99561	9	17	25	49
	2012	99399	3	15	34	48	99308	5	21	35	39	99482	9	16	27	47	98942	10	32	34	24	98699	9	16	29	46
	2011	99189	4	16	33	47	99104	6	21	35	38	98894	9	19	27	45	98715	11	32	33	24	98308	9	19	29	44
	2010	97674	5	19	35	41	97622	8	24	34	33	97182	9	17	26	47	97209	13	34	32	21	96769	9	17	28	45
	2009	94885	6	20	34	39	94834	9	25	34	32	94479	10	20	28	42	94490	15	34	30	21	94078	10	21	29	40
	2008	90086	7	22	35	36	90036	11	26	34	29	89459	12	19	27	42	89541	17	35	29	19	88661	11	20	29	40
	2007	84705	8	25	37	30	84658	12	29	35	24	84266	12	18	21	50	84218	19	37	29	15	83547	11	19	26	45
	2006	82732	10	28	36	26	82703	15	31	33	21	82317	13	20	25	41	82346	23	39	26	12	81595	13	21	29	37
	2005	77742	15	34	35	16	77716	20	36	31	12	77272	14	20	25	42	77600	35	40	19	6	76774	14	20	28	39

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.5. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades 4–5

			Listening Proficiency Level (%)					Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)			
Grade	**Year**	**Number of Students**	**B**	**I**	**A**	**H**	**Number of Students**	**B**	**I**	**A**	**H**	**Number of Students**	**B**	**I**	**A**	**H**	**Number of Students**	**B**	**I**	**A**	**H**	**Number of Students**	**B**	**I**	**A**	**H**
4	2023	101659	16	25	34	25	101649	17	32	40	10	101693	14	32	22	32	101692	23	41	31	5	101633	11	36	39	14
	2022	100355	15	28	37	21	100353	15	37	42	6	100375	13	36	24	27	99884	11	28	34	28	99607	7	34	42	16
	2021	86990	12	31	38	19	86988	14	37	44	5	87046	10	39	23	28	93711	10	30	35	25	84837	5	36	44	15
	2020	75670	11	29	35	25	75665	8	35	48	9	81907	13	29	24	34	48764	8	25	35	32	42546	5	29	45	22
	2019	100655	13	35	38	15	100652	10	33	37	20	100691	11	34	26	29	100124	8	26	36	30	100025	5	33	43	19
	2018	96904	8	33	42	17	96860	6	33	51	10	96542	7	32	35	26	96118	8	26	36	30	96285	3	29	50	18
	2017	98325	4	10	26	61	98240	5	14	31	50	98062	13	31	40	17	97769	8	24	36	33	97522	6	22	42	31
	2016	95284	3	9	26	61	95200	5	14	32	50	95026	12	28	42	18	94771	7	25	36	32	94455	5	21	42	31
	2015	93341	3	9	26	61	93256	5	14	31	50	93101	12	30	41	17	92899	7	24	35	34	92649	5	21	42	31
	2014	87018	3	9	27	61	86953	4	14	33	49	86756	13	29	40	18	86550	7	23	37	33	86301	5	21	43	31
	2013	85437	3	10	28	59	85364	4	14	34	48	85270	7	18	27	48	85012	6	24	37	33	84807	6	19	28	47
	2012	81536	3	11	30	57	81472	4	16	34	46	81702	6	16	29	49	81145	7	25	36	32	80905	6	17	30	48
	2011	81998	3	11	30	55	81939	5	16	34	45	81775	5	16	26	53	81578	8	24	35	33	81255	5	16	27	52
	2010	78574	4	14	33	49	78524	6	19	35	40	78193	6	18	28	48	78161	9	27	36	28	77763	6	18	30	46
	2009	75397	5	15	33	47	75362	7	20	34	39	75068	7	17	26	50	75000	10	27	34	29	74713	7	18	28	48
	2008	69197	6	17	35	42	69157	8	22	35	35	68669	8	18	26	48	68806	12	29	33	25	67980	8	18	28	46
	2007	58398	8	22	38	33	58347	11	26	36	26	58080	8	13	38	40	57986	16	34	32	19	57484	8	14	41	38
	2006	55265	10	25	38	28	55254	14	29	35	22	54884	12	16	40	32	54925	20	35	29	15	54325	12	16	42	30
	2005	52422	14	33	37	16	52404	19	36	33	12	52021	13	16	40	31	52292	31	40	22	7	51638	13	16	41	30
5	2023	103145	11	19	33	38	103142	16	30	41	13	103176	8	23	21	47	103176	14	35	41	10	103100	8	27	41	24
	2022	98982	10	21	37	31	98979	14	35	44	7	99031	9	27	25	40	98489	8	21	34	37	98216	5	27	44	24
	2021	84508	8	22	38	31	84505	14	34	46	6	84611	7	28	23	43	90745	6	24	36	34	82234	4	27	45	24
	2020	74409	7	20	34	39	74410	9	33	47	11	79738	9	22	21	48	48255	6	20	34	40	42024	4	20	44	31
	2019	91372	8	26	41	25	91367	12	30	35	23	91365	7	25	26	43	90878	6	21	35	38	90784	4	25	43	29
	2018	88040	6	23	43	28	88016	7	30	51	13	87750	5	23	34	39	87370	6	21	36	38	87445	2	21	49	28
	2017	85788	3	7	21	68	85707	5	10	27	58	85621	8	20	40	31	85345	6	19	34	41	85111	5	14	37	44
	2016	82864	3	7	22	68	82788	4	10	28	58	82718	8	19	42	32	82426	6	19	35	40	82183	4	14	39	44
	2015	78434	3	7	23	67	78392	4	11	28	57	78277	8	20	41	30	78068	6	19	35	40	77871	4	14	39	43
	2014	72435	3	7	24	66	72365	4	11	30	56	72394	8	19	42	31	72064	5	19	37	39	71886	4	14	40	43
	2013	68832	3	8	25	65	68770	4	11	30	55	68751	5	10	19	66	68469	5	19	37	39	68293	4	10	20	65
	2012	59323	3	10	27	61	59272	4	14	32	50	59486	5	11	23	61	58991	6	22	37	36	58739	4	11	25	60
	2011	57878	4	10	28	58	57827	5	15	33	47	57667	4	11	21	64	57510	7	23	37	34	57231	4	11	23	63
	2010	54794	4	13	31	52	54760	6	18	35	42	54574	5	11	21	62	54477	8	27	37	29	54247	5	12	24	60
	2009	53401	5	14	33	48	53364	7	19	35	40	53150	6	12	20	63	53086	9	27	36	28	52868	6	12	23	60
	2008	49874	6	17	35	42	49848	9	21	35	34	49418	7	13	21	60	49500	11	30	35	23	48843	7	13	24	56
	2007	47558	8	21	37	34	47525	11	25	37	27	47225	7	10	31	53	47228	14	34	34	18	46735	6	11	35	48
	2006	47509	10	24	39	28	47473	13	28	37	22	47209	10	11	33	46	47186	17	38	31	14	46661	9	12	38	41
	2005	43104	14	33	37	16	43094	18	36	33	12	42719	11	12	34	43	43029	28	44	22	6	42348	10	12	37	41

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.6. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades 6–7

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
6	2023	100388	5	30	33	32	100383	13	43	38	7	100455	16	28	31	25	100455	16	37	39	8	100287	6	35	44	14
	2022	94690	5	26	31	38	94688	11	45	42	3	94763	15	31	27	26	94162	7	21	35	37	93685	4	31	46	19
	2021	76394	5	24	34	37	76391	14	47	35	3	76468	15	28	29	28	79496	5	24	37	34	70383	3	31	48	17
	2020	59588	4	18	33	45	59587	11	43	40	6	65722	13	30	31	26	38114	6	20	35	39	30781	3	24	49	23
	2019	79561	3	19	42	35	79560	14	52	28	6	79624	14	32	30	24	79158	6	22	37	36	78995	3	31	51	15
	2018	70933	3	22	43	32	70910	6	45	44	5	70776	10	32	34	24	70378	6	22	37	36	70312	2	27	53	19
	2017	71296	4	9	23	64	71260	5	11	27	57	71259	9	32	41	18	70974	6	19	36	39	70713	5	17	44	34
	2016	66118	3	9	23	65	66095	5	11	28	57	66134	9	27	50	14	65857	5	19	36	40	65594	4	16	45	35
	2015	61311	3	9	24	64	61272	5	11	28	56	61228	8	32	43	18	61016	5	19	37	39	60811	4	17	45	34
	2014	52847	3	9	26	62	52802	4	12	30	54	52889	11	30	43	16	52552	5	20	39	36	52355	4	18	47	31
	2013	47211	3	11	28	57	47182	5	14	33	48	47470	4	14	32	50	46991	6	23	40	31	46770	4	14	34	49
	2012	44423	3	11	30	56	44387	5	14	33	48	44340	4	11	32	53	44205	5	23	39	33	43935	3	11	34	52
	2011	43752	4	12	31	53	43726	5	15	33	46	43693	4	12	29	55	43483	6	23	38	32	43233	4	12	31	54
	2010	42235	5	14	32	49	42209	6	17	34	43	42020	5	12	33	50	41995	7	26	37	29	41697	5	13	35	48
	2009	40302	5	16	34	44	40279	7	20	35	38	40137	6	14	29	51	39983	9	28	36	28	39581	5	14	32	49
	2008	39038	7	20	35	38	39023	9	23	35	33	39098	7	15	29	49	38754	11	32	35	22	38166	6	16	32	46
	2007	34666	9	25	38	27	34645	12	28	37	23	34408	6	9	45	40	34427	14	38	34	14	33909	6	11	48	36
	2006	32766	12	29	37	22	32751	16	31	34	18	32451	9	11	44	35	32469	19	42	29	10	31968	9	13	48	30
	2005	37271	15	35	36	14	37279	19	37	33	11	36987	9	12	48	31	37269	27	46	22	5	36629	9	12	50	29
7	2023	98437	5	27	31	37	98432	16	43	34	7	98513	14	25	30	31	98514	14	33	42	11	98304	6	33	44	17
	2022	92519	5	21	29	46	92516	14	46	38	3	92640	14	28	26	33	92026	6	20	35	39	91378	4	29	46	21
	2021	65716	5	21	33	41	65714	17	49	31	3	65900	14	25	29	32	68746	5	22	37	36	59926	4	31	48	17
	2020	55073	4	15	28	53	55073	15	44	36	5	61736	13	27	30	30	35106	5	18	35	41	28525	4	25	48	24
	2019	68708	4	18	38	40	68707	19	51	25	5	68765	14	28	28	30	68232	6	21	36	37	68059	4	33	49	15
	2018	63392	4	20	41	35	63382	9	46	41	4	63295	10	30	33	28	62882	6	21	37	37	62784	2	26	52	19
	2017	60180	5	10	24	62	60151	6	11	27	56	60170	9	29	41	21	59896	6	19	37	38	59647	5	16	43	36
	2016	54017	4	10	25	62	53979	6	12	28	54	54130	9	24	51	16	53765	6	19	37	38	53501	5	15	45	35
	2015	48594	5	10	25	60	48563	6	12	29	53	48601	9	29	41	21	48307	6	19	38	37	48090	5	16	44	35
	2014	43971	4	10	27	59	43952	6	13	30	51	44122	11	27	42	20	43767	6	20	38	36	43539	5	17	45	33
	2013	37967	4	11	29	56	37952	6	13	32	49	38167	4	11	24	61	37760	6	21	40	34	37541	4	11	26	59
	2012	36793	4	11	29	56	36784	5	14	31	49	36661	4	10	27	59	36555	6	20	38	36	36312	3	11	28	58
	2011	36807	5	13	30	53	36785	6	15	32	46	36709	4	11	25	60	36556	7	22	38	33	36293	4	11	27	58
	2010	34271	5	15	31	49	34253	7	17	33	43	34223	5	13	29	53	34036	8	24	38	30	33815	5	13	31	51
	2009	33676	6	17	33	44	33668	8	20	33	38	33617	6	13	24	57	33410	9	26	37	28	33088	6	13	26	55
	2008	31655	7	20	35	38	31644	10	23	34	33	31380	7	16	27	50	31364	11	31	35	23	30761	7	16	30	47
	2007	27502	10	25	37	27	27499	13	28	36	23	27198	6	11	44	38	27289	15	37	33	14	26768	6	12	47	35
	2006	31002	11	28	37	23	30996	15	30	35	20	30610	9	12	45	34	30733	18	41	31	10	30150	9	13	48	30
	2005	29074	15	35	35	14	29072	19	37	32	11	28719	10	13	46	31	28994	27	46	22	5	28390	9	13	48	29

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.7. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades 8–9

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
8	2023	95401	5	23	29	44	95395	18	40	35	7	95483	7	31	31	32	95480	16	43	33	8	95248	5	35	43	17
	2022	81952	5	19	26	51	81950	15	45	37	3	82034	10	35	32	23	81430	5	17	34	43	80805	4	28	47	21
	2021	59699	5	19	29	47	59697	19	49	30	3	59686	11	34	32	23	63182	4	19	36	41	54427	4	30	49	17
	2020	47523	4	13	23	60	47521	16	43	35	6	52952	6	37	33	24	31222	5	17	33	45	25169	3	25	48	24
	2019	61716	4	16	35	45	61713	20	50	25	5	61777	10	35	35	20	61194	5	18	35	42	61043	3	32	50	15
	2018	54294	4	18	38	40	54287	10	43	42	5	54274	7	39	37	17	53857	5	18	35	41	53661	2	26	53	19
	2017	51580	4	9	22	65	51552	6	11	25	58	51580	10	25	49	16	51340	6	17	34	43	51039	5	14	43	38
	2016	45266	5	10	23	62	45234	7	12	26	55	45423	10	30	47	13	45082	6	18	35	42	44754	5	16	45	34
	2015	42139	4	10	24	62	42128	7	12	27	54	42187	11	28	49	13	41938	6	17	35	41	41704	5	16	45	34
	2014	34985	4	9	25	62	34972	6	12	27	55	35099	10	25	49	16	34796	6	17	36	41	34541	5	14	45	36
	2013	30291	4	10	27	58	30281	6	13	30	51	30504	7	9	24	60	30122	6	19	39	37	29874	6	10	26	58
	2012	28711	4	12	29	54	28696	6	16	31	47	28549	7	10	25	58	28449	6	21	37	37	28164	7	10	27	56
	2011	27778	5	14	30	51	27751	7	17	32	44	27615	7	12	26	55	27479	7	23	37	34	27168	6	12	28	54
	2010	28275	5	15	31	48	28259	8	18	32	42	28050	8	11	26	55	28014	8	24	36	32	27681	7	12	28	53
	2009	27438	6	17	32	45	27418	9	20	32	39	27094	8	14	28	50	27115	9	26	35	30	26596	8	14	30	48
	2008	25861	7	20	34	39	25836	11	23	33	33	25511	9	15	28	47	25563	11	31	35	24	24817	9	16	31	45
	2007	29347	8	24	38	31	29333	11	27	36	26	28959	4	9	41	46	29083	12	36	36	16	28480	4	10	44	42
	2006	27180	10	27	38	25	27168	14	30	36	21	26853	8	11	40	41	26897	15	40	33	12	26363	7	12	44	37
	2005	25221	15	35	36	15	25221	20	36	32	12	24759	9	12	42	37	25137	26	45	23	6	24475	9	12	45	35
9	2023	96472	11	27	41	21	96467	34	26	32	8	96465	10	33	29	29	96466	22	41	30	7	95757	11	40	36	13
	2022	84571	14	25	40	20	84570	33	32	31	4	84774	15	35	30	21	83291	10	23	34	33	80554	9	39	38	14
	2021	51951	8	24	47	20	51947	32	28	33	7	52041	13	33	30	25	52767	6	24	37	33	44717	6	37	42	16
	2020	51988	13	28	34	25	51987	37	30	27	5	54275	9	40	30	21	28957	9	22	34	35	25309	7	42	36	15
	2019	59413	14	33	35	19	59411	35	31	27	6	59437	14	36	32	18	58081	9	24	36	32	57743	9	43	36	12
	2018	50419	10	31	43	15	50411	25	29	40	5	50356	10	42	33	15	49793	10	25	36	30	49155	6	39	43	12
	2017	48999	11	16	26	47	48973	15	17	26	42	49271	19	28	40	14	48684	13	24	33	29	48057	12	22	41	26
	2016	43835	10	15	27	48	43824	13	17	27	43	44015	16	32	40	12	43644	11	23	35	31	42842	9	22	43	25
	2015	38954	11	16	27	46	38940	15	17	27	40	38876	18	27	42	13	38714	13	24	34	28	38104	11	22	42	26
	2014	33788	11	16	27	46	33775	14	18	28	40	33903	18	26	40	17	33488	12	25	35	28	32939	11	20	42	27
	2013	30798	10	17	29	44	30786	13	20	29	39	30870	12	12	24	52	30418	12	26	35	28	29956	11	13	27	50
	2012	29250	9	18	31	43	29232	13	21	30	37	29045	11	13	25	51	28908	11	27	34	27	28343	10	14	27	49
	2011	31603	10	20	31	39	31586	14	22	30	33	31376	11	16	27	47	31217	13	30	33	25	30672	9	17	29	44
	2010	30392	11	22	32	35	30381	16	24	30	31	30307	13	17	27	43	30002	14	31	32	23	29554	12	18	29	41
	2009	30829	11	23	33	32	30815	16	25	32	27	30575	12	18	28	42	30220	15	34	43	20	29652	11	19	30	39
	2008	35026	10	24	35	31	35011	15	26	32	27	33924	11	18	29	42	34489	14	34	33	20	32846	10	19	33	38
	2007	33777	13	28	34	25	33764	18	29	32	22	32834	15	17	45	23	33177	17	37	32	15	31894	14	19	46	21
	2006	33457	16	30	34	20	33445	21	32	31	16	32480	19	16	47	18	32836	20	41	29	10	31619	18	17	48	16
	2005	32336	23	36	30	11	32318	29	36	26	9	31188	20	20	44	16	32044	35	43	18	4	30687	20	20	45	15

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.8. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grades 10–11

Grade	Year	Listening Proficiency Level (%)				Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
10	2023	77503	9	25	42	24	77502	30	25	35	10	77622	10	42	26	23	77616	20	33	36	11	76881	9	40	36	15
	2022	59135	10	24	42	24	59133	30	29	35	6	59276	12	42	27	20	58697	5	21	36	37	56754	6	39	40	15
	2021	43063	8	24	47	21	43063	31	27	34	8	43139	13	39	27	20	44943	4	22	37	36	36932	6	37	42	16
	2020	37263	9	26	34	30	37263	34	29	30	7	39378	12	43	27	18	21200	5	22	36	37	18527	5	42	37	16
	2019	42307	11	32	36	21	42305	32	31	29	8	42349	13	43	27	17	41574	5	24	38	33	41419	6	44	37	13
	2018	37588	8	31	44	17	37583	21	28	44	7	37526	10	43	35	12	37414	6	25	36	32	36927	4	37	46	12
	2017	35171	5	15	28	52	35155	8	18	29	45	35141	13	27	45	15	34984	6	24	36	34	34582	6	21	44	30
	2016	30347	5	15	29	51	30339	8	18	29	45	30306	11	28	46	15	30160	6	24	36	34	29763	5	20	45	30
	2015	25866	5	15	29	50	25853	9	19	29	43	25685	12	27	46	16	25690	6	25	36	33	25370	5	21	45	29
	2014	23570	4	14	30	52	23562	7	18	31	44	23432	13	26	45	16	23419	5	23	38	34	23138	5	20	45	30
	2013	20689	4	15	29	53	20682	6	18	30	46	20555	5	15	27	52	20533	4	24	36	36	20275	4	16	30	50
	2012	19440	4	15	32	49	19432	7	20	32	41	19250	6	16	27	51	19256	5	25	39	32	18994	5	17	29	49
	2011	18519	4	17	34	46	18501	6	22	33	39	18347	6	17	30	47	18349	5	27	38	30	18048	5	17	32	45
	2010	18785	4	18	36	42	18779	7	23	34	35	18465	5	17	31	47	18591	6	28	39	27	18062	4	17	34	45
	2009	20753	4	18	36	41	20739	7	23	35	35	20311	6	17	30	47	20514	6	29	39	27	19961	5	17	33	45
	2008	20460	5	20	37	38	20446	8	24	35	33	20321	7	18	30	45	20231	6	31	38	25	19512	6	18	33	43
	2007	18943	5	26	39	30	18926	10	30	36	25	18433	8	14	47	31	18706	8	36	37	19	18058	8	15	49	29
	2006	18662	8	28	40	25	18661	12	32	36	20	18129	9	13	50	28	18437	10	41	36	14	17734	9	14	52	25
	2005	19040	12	36	38	14	19035	17	39	32	12	18322	10	16	50	25	18902	22	48	24	6	18119	10	16	51	23
11	2023	55395	7	24	42	27	55394	29	24	36	11	55459	8	40	27	26	55460	17	32	38	12	54907	7	39	38	17
	2022	48788	9	23	44	24	48788	29	29	36	6	48977	10	41	28	21	48568	4	20	36	41	46816	5	38	41	16
	2021	30728	7	23	48	22	30728	29	26	36	10	30764	11	40	28	21	32516	3	20	39	39	26291	4	36	43	17
	2020	25912	7	25	36	31	25910	32	28	32	8	27759	10	42	29	20	15365	3	20	37	40	13325	3	41	38	18
	2019	31768	10	32	36	23	31766	29	30	31	10	31819	11	41	28	20	21285	3	22	39	36	31130	4	42	38	16
	2018	28098	5	27	47	21	28088	17	25	49	9	28058	7	39	39	15	28046	3	21	38	38	27709	2	31	50	16
	2017	24693	3	12	28	57	24687	5	17	29	49	24647	9	24	48	19	24574	3	21	37	39	24274	3	17	45	35
	2016	20792	3	13	29	56	20781	5	17	30	48	20664	7	25	48	20	20642	3	21	37	39	20333	2	16	46	36
	2015	19197	2	12	30	56	19190	5	17	31	47	19107	8	23	50	19	19081	3	20	38	38	18834	3	16	46	35
	2014	17972	2	11	29	58	17963	4	16	31	49	17841	8	21	46	24	17859	3	20	37	41	17664	2	16	43	39
	2013	15042	3	14	32	51	15039	5	20	32	42	14925	4	16	27	53	14919	4	24	39	34	14707	3	16	29	51
	2012	14187	2	14	32	52	14184	5	20	33	42	13977	5	15	27	54	14042	3	24	38	35	13771	4	15	29	52
	2011	14270	3	14	34	49	14261	5	21	34	40	14085	5	15	29	51	14093	4	24	40	33	13822	4	15	31	49
	2010	16023	3	15	35	47	16016	4	20	36	40	15936	3	13	31	53	15838	3	24	40	32	15601	2	14	33	50
	2009	15026	3	17	37	43	15024	5	22	37	36	14798	5	14	28	53	14802	4	28	40	29	14516	4	15	31	51
	2008	14102	3	18	38	41	14081	6	23	37	34	13625	5	15	30	51	13913	4	28	41	27	13230	4	15	33	48
	2007	13485	4	21	41	33	13467	7	27	38	28	12965	6	11	45	38	13284	6	33	40	21	12685	5	11	48	36
	2006	13947	5	24	41	29	13940	9	30	38	24	13387	6	9	49	35	13732	8	36	38	18	13112	6	10	52	32
	2005	14117	9	33	41	18	14114	13	38	35	14	13571	6	12	49	33	14041	16	47	28	8	13383	6	12	50	32

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

Table D.6.9. Distribution of TELPAS Proficiency Levels by Administration Since 2005 - Grade 12

Grade	Year	Listening Proficiency Level (%)					Speaking Proficiency Level (%)					Reading Proficiency Level (%)					Writing Proficiency Level (%)					Composite Rating (%)				
		Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H	Number of Students	B	I	A	H
12	2023	45346	6	26	44	24	45346	31	25	35	10	45476	7	41	28	23	45476	16	34	38	11	45021	6	42	38	14
	2022	36047	8	25	45	22	36045	30	29	35	6	36226	10	42	29	19	36155	2	17	37	44	34588	3	40	41	15
	2021	22276	6	25	50	20	22274	29	26	36	9	22326	10	42	29	19	24037	2	17	38	43	19121	3	37	44	16
	2020	20377	7	28	36	29	20376	32	27	33	8	21867	9	44	28	19	11210	2	18	38	42	9564	3	42	39	16
	2019	25315	9	33	37	21	25315	28	29	32	10	25307	10	41	29	19	24871	2	19	39	40	24773	3	41	40	16
	2018	21431	4	27	49	21	21429	16	23	51	10	21387	6	38	39	16	21422	2	18	37	43	21095	2	29	52	18
	2017	17993	1	9	27	62	17996	3	14	31	52	17859	8	23	49	20	17854	2	18	37	43	17569	2	15	45	39
	2016	16092	2	9	27	62	16084	3	14	30	53	15982	6	24	49	21	15956	2	17	37	45	15969	2	14	45	40
	2015	14667	2	9	27	63	14658	3	13	30	53	14464	7	22	50	21	14499	2	16	37	45	14222	2	13	45	40
	2014	9863	3	11	32	55	9853	4	17	35	43	9683	11	27	43	19	9723	3	21	40	35	9558	3	19	47	31
	2013	9432	2	12	31	55	9421	4	18	34	44	9259	4	16	30	50	9290	3	21	39	36	9076	3	17	31	49
	2012	9240	2	11	32	55	9230	3	17	35	45	9034	5	15	28	52	8975	3	21	41	36	8728	4	16	30	50
	2011	10339	2	12	32	54	10322	3	17	35	45	10303	6	14	30	50	10046	2	21	40	36	9822	3	16	32	49
	2010	9529	2	13	35	50	9524	3	18	37	41	9469	5	14	33	49	9318	3	23	41	34	9114	3	15	35	48
	2009	9898	2	13	37	47	9879	4	20	38	39	9702	5	13	29	53	9651	3	25	40	31	9386	3	14	31	51
	2008	9632	3	16	38	43	9605	5	22	38	35	9164	4	14	31	51	9414	4	27	41	29	8806	4	15	33	48
	2007	9289	3	19	40	38	9267	5	24	39	32	8695	5	10	46	39	9065	4	29	43	24	8488	4	11	48	37
	2006	9016	4	22	42	32	9006	6	28	40	26	8444	5	8	52	35	8787	6	35	40	20	8253	5	9	55	32
	2005	9414	6	30	43	22	9409	10	34	39	18	8786	5	10	51	34	9297	13	44	32	11	8626	4	10	52	33

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to redesigns of the TELPAS administration. Refer to the Evidence Based on Consequences of Testing section of Chapter 6, “TELPAS”.

TECHNICAL DIGEST 2022–2023

Appendix E

TELPAS Alternate Statistical Tables and Figures

TELPAS Alternate Statistical Tables and Figures

2023 TELPAS Alternate Classification Consistency and Accuracy

- Table E.1.1. TELPAS Alternate

2023 TELPAS Alternate Scale Score Correlations

- Table E.2.1. TELPAS Alternate

2023 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

- Table E.3.1. TELPAS Alternate

2023 TELPAS Alternate Mean P-Values

- Table E.4.1. TELPAS Alternate Total Group
- Table E.4.2. TELPAS Alternate Female
- Table E.4.3. TELPAS Alternate Male

2023 TELPAS Alternate Scale Score Descriptive Statistics

- Table E.5.1. TELPAS Alternate

2023 TELPAS Alternate Frequency Distribution of Scale Scores

- Figure E.5.1. TELPAS Alternate Listening
- Figure E.5.2. TELPAS Alternate Speaking
- Figure E.5.3. TELPAS Alternate Reading
- Figure E.5.4. TELPAS Alternate Writing

2023 TELPAS Alternate Classification Consistency and Accuracy

Table E.1.1. Spring 2023 TELPAS Alternate Assessments
Classification Consistency and Accuracy

Domain	Decision Consistency[1]	Decision Accuracy[2]
Listening	73.8	84.9
Speaking	78.6	88.0
Reading	75.4	85.9
Writing	76.3	86.5

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

2023 TELPAS Alternate Scale Score Correlations

Table E.2.1. Spring 2023 TELPAS Alternate Assessments Scale Score Correlations

TELPAS Alternate Domain	TELPAS Alternate Domain	N	Correlation
Listening	Speaking	11,340	0.88
Listening	Reading	11,340	0.88
Listening	Writing	11,340	0.85
Speaking	Reading	11,340	0.83
Speaking	Writing	11,340	0.81
Reading	Writing	11,340	0.93

2023 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

Table E.3.1. Spring 2023 TELPAS Alternate
Conditional Standard Error of Measurement (CSEM) for Scale Score (SS)

	Listening		Speaking		Reading		Writing	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
10	609		600		601		612	
11	638	24	627	23	630	23	640	23
12	657	18	645	17	648	18	660	18
13	669	15	657	15	660	15	674	15
14	678	14	666	14	670	14	684	14
15	686	13	675	13	679	14	693	13
16	693	12	*682	12	687	13	*701	12
17	*699	12	690	12	696	13	708	12
18	706	12	697	12	*704	13	715	12
19	712	12	704	12	712	14	722	11
20	717	11	711	12	720	13	728	11
21	723	11	718	12	728	13	734	11
22	729	11	725	12	736	13	739	11
23	734	11	732	12	743	12	745	10
24	740	11	738	11	**750	12	**750	10
25	745	11	744	11	756	12	755	10
26	**750	11	**750	11	762	11	760	10
27	755	11	756	11	768	11	766	10
28	760	11	761	11	773	11	771	10
29	765	11	766	11	779	11	776	11
30	770	11	772	11	784	11	782	11
31	775	11	777	11	789	11	788	11
32	780	11	783	11	795	11	794	11
33	785	11	788	11	***800	11	***800	11
34	790	11	794	11	805	11	806	11
35	795	11	***800	11	810	11	813	11
36	***800	11	806	12	816	11	819	12
37	805	11	813	12	821	11	826	12
38	811	11	819	12	826	11	832	12
39	816	11	827	12	832	11	839	12
40	822	12	834	12	837	11	846	12
41	829	12	841	13	843	11	****853	12
42	835	12	849	13	****849	12	860	12
43	842	13	856	13	855	12	867	12
44	850	13	****864	13	862	13	875	12
45	****858	14	872	13	870	13	882	13
46	867	15	881	14	878	14	891	13
47	877	16	892	15	888	15	900	14
48	889	18	905	18	901	18	911	17
49	908	24	924	23	919	24	929	22
50	938		952		948		955	

Notes:

* Imitation, ** Early Independence, *** Developing Independence, **** Basic Fluency

Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

2023 TELPAS Alternate
Mean P-Values

Table E.4.1. Spring 2023 TELPAS Alternate Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
LISTENING	OVERALL TEST	50	11,340	32.08	11.77	0.97	1.89	55.20
	Understand spoken words and language structures	20	11,340	12.73	5.05	0.94	1.20	54.57
	Basic understanding of spoken English	30	11,340	19.35	6.93	0.96	1.39	55.62
SPEAKING	OVERALL TEST	50	11,340	30.71	12.92	0.98	1.64	51.77
	Provide and summarize information	35	11,340	21.42	9.19	0.98	1.38	51.50
	Share opinions and analyze information	15	11,340	9.29	3.84	0.95	0.87	52.42
READING	OVERALL TEST	50	11,340	27.66	12.02	0.98	1.74	44.16
	Understand words and language structures	25	11,340	14.46	6.30	0.97	1.14	47.32
	Basic understanding of a variety of texts written in English	15	11,340	7.99	3.64	0.94	0.90	41.62
	Analyze and evaluate information and ideas in a variety of texts written in English	10	11,340	5.21	2.47	0.94	0.62	40.07
WRITING	OVERALL TEST	50	11,340	24.44	11.41	0.98	1.68	36.11
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	11,340	10.75	4.90	0.96	1.03	42.17
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	11,340	13.70	6.77	0.97	1.15	32.07

Notes:
1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

Table E.4.2. Spring 2023 TELPAS Alternate Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
LISTENING	OVERALL TEST	50	3,787	32.18	11.77	0.97	1.87	55.46
	Understand spoken words and language structures	20	3,787	12.74	5.05	0.94	1.18	54.65
	Basic understanding of spoken English	30	3,787	19.44	6.93	0.96	1.38	56.00
SPEAKING	OVERALL TEST	50	3,787	31.12	12.86	0.98	1.63	52.81
	Provide and summarize information	35	3,787	21.70	9.14	0.98	1.37	52.48
	Share opinions and analyze information	15	3,787	9.43	3.83	0.95	0.85	53.58
READING	OVERALL TEST	50	3,787	27.57	12.08	0.98	1.68	43.92
	Understand words and language structures	25	3,787	14.32	6.30	0.97	1.11	46.62
	Basic understanding of a variety of texts written in English	15	3,787	8.01	3.66	0.94	0.88	41.78
	Analyze and evaluate information and ideas in a variety of texts written in English	10	3,787	5.23	2.47	0.94	0.62	40.36
WRITING	OVERALL TEST	50	3,787	24.58	11.56	0.98	1.63	36.46
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	3,787	10.70	4.91	0.96	1.01	41.89
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	3,787	13.88	6.88	0.97	1.14	32.84

Notes:
1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

Table E.4.3. Spring 2023 TELPAS Alternate Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
LISTENING	OVERALL TEST	50	7,541	32.03	11.77	0.97	1.90	55.08
	Understand spoken words and language structures	20	7,541	12.73	5.06	0.94	1.20	54.53
	Basic understanding of spoken English	30	7,541	19.30	6.92	0.96	1.39	55.44
SPEAKING	OVERALL TEST	50	7,541	30.50	12.95	0.98	1.65	51.25
	Provide and summarize information	35	7,541	21.28	9.22	0.98	1.38	51.00
	Share opinions and analyze information	15	7,541	9.22	3.84	0.95	0.88	51.84
READING	OVERALL TEST	50	7,541	27.71	11.98	0.98	1.77	44.29
	Understand words and language structures	25	7,541	14.53	6.30	0.97	1.16	47.67
	Basic understanding of a variety of texts written in English	15	7,541	7.99	3.63	0.94	0.91	41.55
	Analyze and evaluate information and ideas in a variety of texts written in English	10	7,541	5.19	2.46	0.94	0.63	39.93
WRITING	OVERALL TEST	50	7,541	24.37	11.33	0.98	1.70	35.93
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	7,541	10.77	4.90	0.95	1.04	42.31
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	7,541	13.60	6.71	0.97	1.15	31.68

Notes:

1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

2023 TELPAS Alternate Scale Score Descriptive Statistics

Glossary

This glossary provides definitions for the statistical terms that appear in the tables and graphs in this section. Definitions of statistical terms and concepts in the other sections are given in Chapter 3 and Chapter 7.

Descriptive Statistics

Mean The mean is a measure of central tendency. It is the average score for the assessment. It is computed by summing the scores of all students and dividing the sum by the total number of students (*N*).

Median The median is another measure of central tendency. It is the score at the middle of the frequency distribution for the assessment. It is computed by finding the score at which there are the same number of scores above as there are below.

Mode The mode is another measure of central tendency. It is the most frequently obtained score for the assessment. It is determined by computing the frequency distribution and finding the score point with the highest frequency (*n*-count).

Range The range is a measure of statistical dispersion (variability or spread). It is the difference between the lowest and highest scores obtained by students on the assessment. It is computed by subtracting the lowest score from the highest score.

Interquartile Range The interquartile range is another measure of statistical dispersion (variability or spread). It is the difference between the first and third quartiles (or 25th and 75th percentiles) of the score distribution for the assessment. It is computed by subtracting the score at the first quartile (the point that splits the lowest 25 percent of the scores) from the score at the third quartile (the point that splits the highest 25 percent of the scores).

Standard Deviation The standard deviation is another measure of statistical dispersion (variability or spread). It is an indicator of the degree of score variation around the mean. It is computed using the following formula:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, and N is the total number of students who took the assessment.

Variance The variance is another measure of statistical dispersion (variability or spread) around the mean. It is computed as the square of the standard deviation.

Skewness The skewness is an indicator of the shape of the score distribution. It measures the extent to which the score distribution "leans" to one side of the mean. A positive skewness indicates that the score

distribution leans below the mean. A negative skewness indicates that the score distribution leans above the mean. A skewness of zero indicates that the score distribution is symmetric around the mean. It is computed using the following formula:

$$Skewness = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^3,$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Kurtosis The kurtosis is another indicator of the shape of the score distribution. It measures the "peakedness" of the score distribution. A positive kurtosis is referred to as *leptokurtic*, meaning that the distribution has a more acute peak around the mean and fatter tails. A negative kurtosis is called *platykurtic*, meaning the distribution has a lower, wider peak around the mean and thinner tails. It is computed using the following formula:

$$Kurtosis = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}}{S_X} \right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)},$$

where X_i is the score for student i, $\overline{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Frequency Distributions

Frequency (Freq) The frequency is the number of students who obtained the particular score point on the assessment.

Cumulative Frequency (CumFreq) The cumulative frequency is the number of students who obtained a score that is less than or equal to the particular score point on the assessment.

Percentage (Pct) The percentage is the percentage of students who obtained the particular score point on the assess-ment. It is computed as: Pct = Freq ÷ N × 100.

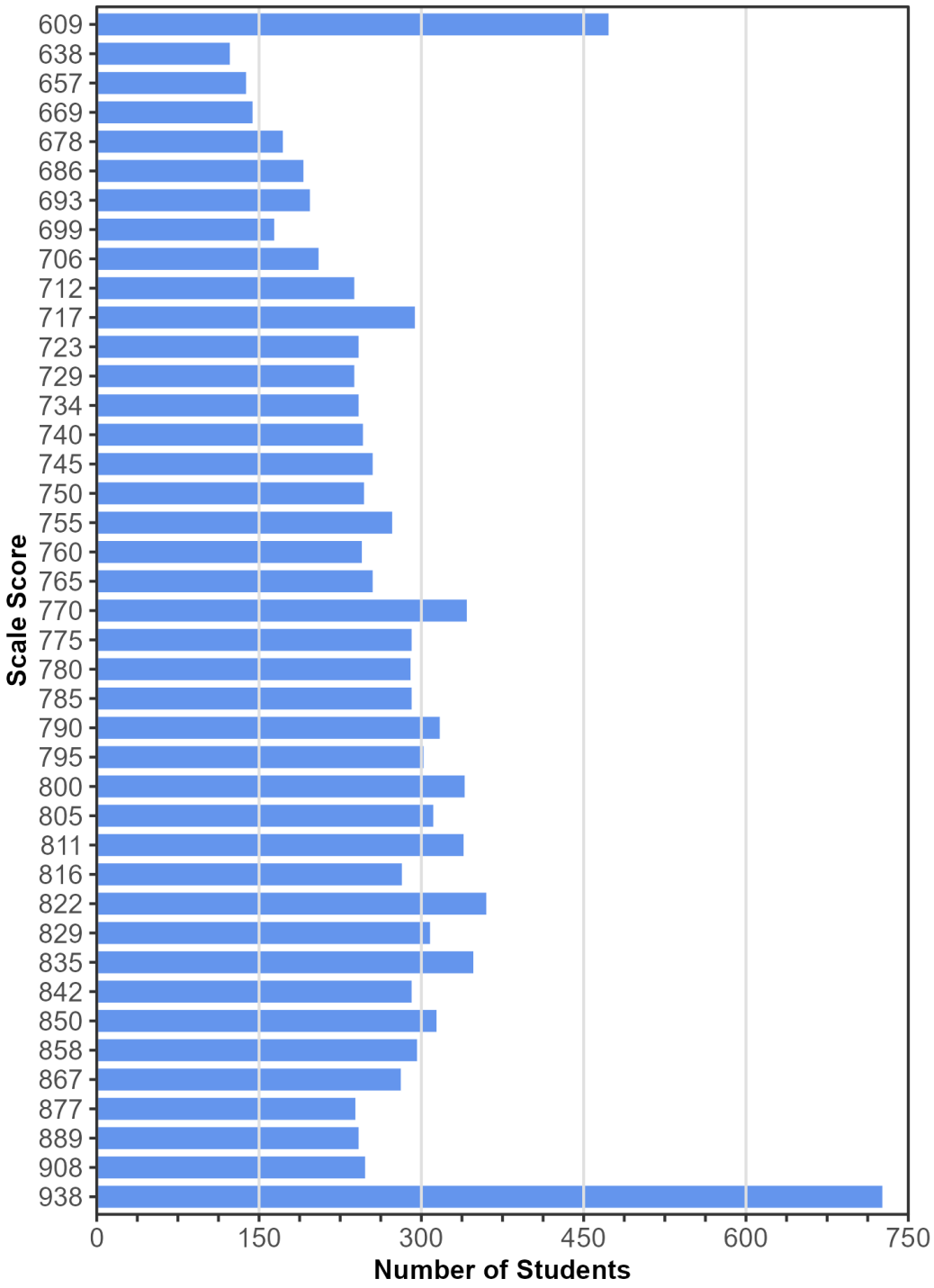
Cumulative Percentage (CumPct) The cumulative percentage is the percentage of students who obtained a score that is less than or equal to the particular score point on the assessment. It is computed as: CumPct = CumFreq ÷ N × 100.

Table E.5.1. Spring 2023 TELPAS Alternate Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
TELPAS Alternate Listening	11,340	784.10	785	938	329	101	78.62	6180.55	-0.07	-0.22
TELPAS Alternate Speaking	11,340	777.71	777	600	352	138	97.99	9601.86	-0.06	-0.75
TELPAS Alternate Reading	11,340	761.57	768	601	347	117	87.87	7720.66	-0.04	-0.54
TELPAS Alternate Writing	11,340	743.89	745	612	343	116	84.05	7064.97	0.18	-0.47

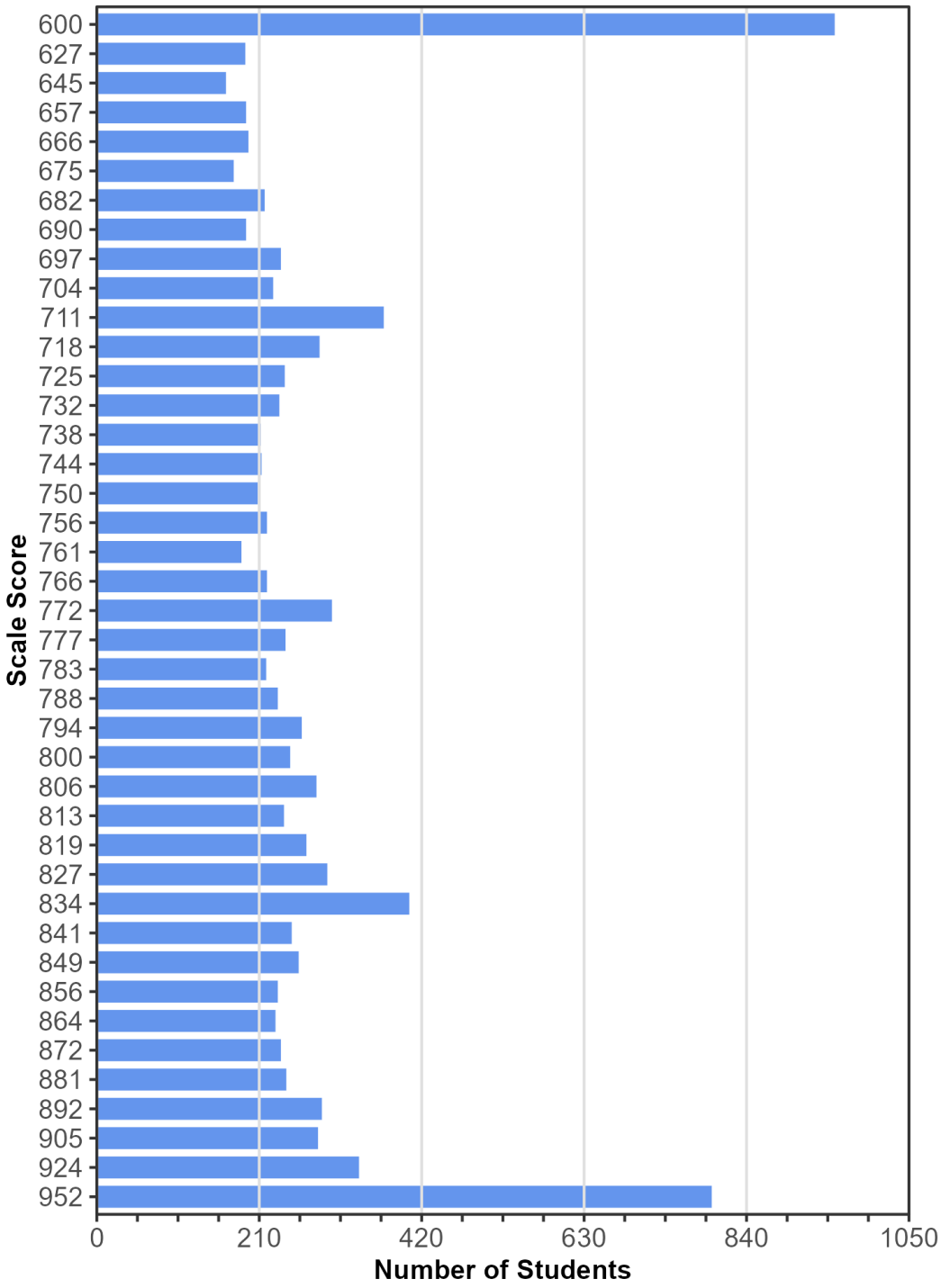
2023 TELPAS Alternate Frequency Distribution of Scale Scores

Figure E.5.1. Spring 2023 TELPAS Alternate Listening
Frequency Distribution of Scale Scores
All Students



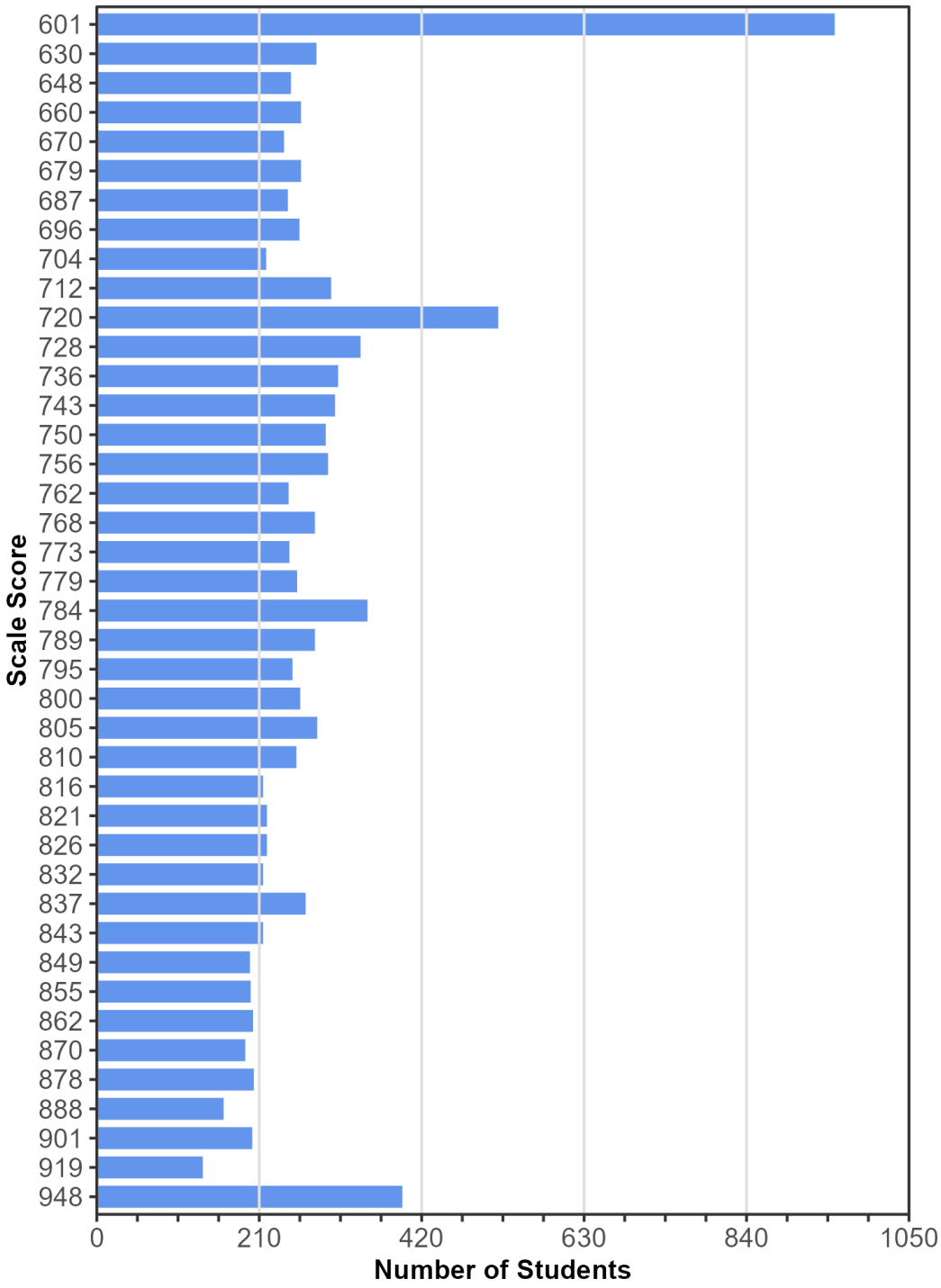
	Freq	Cum Freq	Pct	Cum Pct
609	473	473	4.17	4.17
638	123	596	1.08	5.26
657	138	734	1.22	6.47
669	144	878	1.27	7.74
678	172	1050	1.52	9.26
686	191	1241	1.68	10.94
693	197	1438	1.74	12.68
699	164	1602	1.45	14.13
706	205	1807	1.81	15.93
712	238	2045	2.10	18.03
717	294	2339	2.59	20.63
723	242	2581	2.13	22.76
729	238	2819	2.10	24.86
734	242	3061	2.13	26.99
740	246	3307	2.17	29.16
745	255	3562	2.25	31.41
750	247	3809	2.18	33.59
755	273	4082	2.41	36.00
760	245	4327	2.16	38.16
765	255	4582	2.25	40.41
770	342	4924	3.02	43.42
775	291	5215	2.57	45.99
780	290	5505	2.56	48.54
785	291	5796	2.57	51.11
790	317	6113	2.80	53.91
795	302	6415	2.66	56.57
800	340	6755	3.00	59.57
805	311	7066	2.74	62.31
811	339	7405	2.99	65.30
816	282	7687	2.49	67.79
822	360	8047	3.17	70.96
829	308	8355	2.72	73.68
835	348	8703	3.07	76.75
842	291	8994	2.57	79.31
850	314	9308	2.77	82.08
858	296	9604	2.61	84.69
867	281	9885	2.48	87.17
877	239	10124	2.11	89.28
889	242	10366	2.13	91.41
908	248	10614	2.19	93.60
938	726	11340	6.40	100.00

Figure E.5.2. Spring 2023 TELPAS Alternate Speaking
Frequency Distribution of Scale Scores
All Students



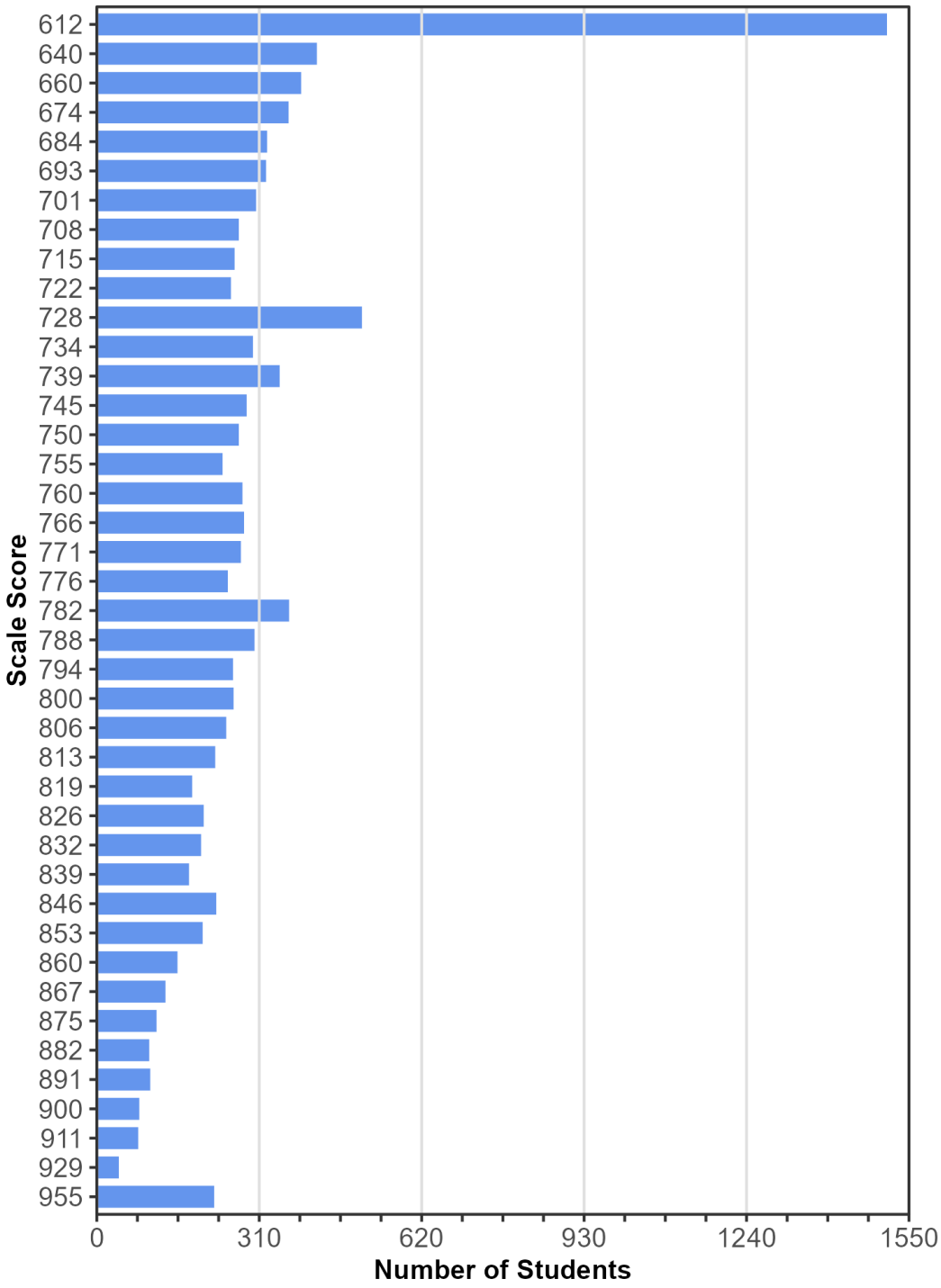
Scale Score	Freq	Cum Freq	Pct	Cum Pct
600	954	954	8.41	8.41
627	192	1146	1.69	10.11
645	167	1313	1.47	11.58
657	193	1506	1.70	13.28
666	196	1702	1.73	15.01
675	177	1879	1.56	16.57
682	217	2096	1.91	18.48
690	193	2289	1.70	20.19
697	238	2527	2.10	22.28
704	228	2755	2.01	24.29
711	371	3126	3.27	27.57
718	288	3414	2.54	30.11
725	243	3657	2.14	32.25
732	236	3893	2.08	34.33
738	212	4105	1.87	36.20
744	213	4318	1.88	38.08
750	208	4526	1.83	39.91
756	220	4746	1.94	41.85
761	187	4933	1.65	43.50
766	220	5153	1.94	45.44
772	304	5457	2.68	48.12
777	244	5701	2.15	50.27
783	219	5920	1.93	52.20
788	234	6154	2.06	54.27
794	265	6419	2.34	56.60
800	250	6669	2.20	58.81
806	284	6953	2.50	61.31
813	242	7195	2.13	63.45
819	271	7466	2.39	65.84
827	298	7764	2.63	68.47
834	404	8168	3.56	72.03
841	252	8420	2.22	74.25
849	261	8681	2.30	76.55
856	234	8915	2.06	78.62
864	231	9146	2.04	80.65
872	238	9384	2.10	82.75
881	245	9629	2.16	84.91
892	291	9920	2.57	87.48
905	286	10206	2.52	90.00
924	339	10545	2.99	92.99
952	795	11340	7.01	100.00

Figure E.5.3. Spring 2023 TELPAS Alternate Reading
Frequency Distribution of Scale Scores
All Students



	Freq	Cum Freq	Pct	Cum Pct
601	954	954	8.41	8.41
630	284	1238	2.50	10.92
648	251	1489	2.21	13.13
660	264	1753	2.33	15.46
670	242	1995	2.13	17.59
679	264	2259	2.33	19.92
687	247	2506	2.18	22.10
696	262	2768	2.31	24.41
704	219	2987	1.93	26.34
712	303	3290	2.67	29.01
720	519	3809	4.58	33.59
728	341	4150	3.01	36.60
736	312	4462	2.75	39.35
743	308	4770	2.72	42.06
750	296	5066	2.61	44.67
756	299	5365	2.64	47.31
762	248	5613	2.19	49.50
768	282	5895	2.49	51.98
773	249	6144	2.20	54.18
779	259	6403	2.28	56.46
784	350	6753	3.09	59.55
789	282	7035	2.49	62.04
795	253	7288	2.23	64.27
800	263	7551	2.32	66.59
805	285	7836	2.51	69.10
810	258	8094	2.28	71.38
816	215	8309	1.90	73.27
821	220	8529	1.94	75.21
826	220	8749	1.94	77.15
832	215	8964	1.90	79.05
837	270	9234	2.38	81.43
843	215	9449	1.90	83.32
849	198	9647	1.75	85.07
855	199	9846	1.75	86.83
862	202	10048	1.78	88.61
870	192	10240	1.69	90.30
878	203	10443	1.79	92.09
888	164	10607	1.45	93.54
901	201	10808	1.77	95.31
919	137	10945	1.21	96.52
948	395	11340	3.48	100.00

Figure E.5.4. Spring 2023 TELPAS Alternate Writing
Frequency Distribution of Scale Scores
All Students



	Freq	Cum Freq	Pct	Cum Pct
612	1508	1508	13.30	13.30
640	420	1928	3.70	17.00
660	390	2318	3.44	20.44
674	366	2684	3.23	23.67
684	325	3009	2.87	26.53
693	323	3332	2.85	29.38
701	304	3636	2.68	32.06
708	271	3907	2.39	34.45
715	263	4170	2.32	36.77
722	256	4426	2.26	39.03
728	506	4932	4.46	43.49
734	298	5230	2.63	46.12
739	349	5579	3.08	49.20
745	286	5865	2.52	51.72
750	271	6136	2.39	54.11
755	240	6376	2.12	56.23
760	278	6654	2.45	58.68
766	281	6935	2.48	61.16
771	275	7210	2.43	63.58
776	250	7460	2.20	65.78
782	367	7827	3.24	69.02
788	301	8128	2.65	71.68
794	260	8388	2.29	73.97
800	261	8649	2.30	76.27
806	247	8896	2.18	78.45
813	226	9122	1.99	80.44
819	182	9304	1.60	82.05
826	204	9508	1.80	83.84
832	199	9707	1.75	85.60
839	176	9883	1.55	87.15
846	228	10111	2.01	89.16
853	202	10313	1.78	90.94
860	154	10467	1.36	92.30
867	131	10598	1.16	93.46
875	114	10712	1.01	94.46
882	100	10812	0.88	95.34
891	102	10914	0.90	96.24
900	81	10995	0.71	96.96
911	79	11074	0.70	97.65
929	42	11116	0.37	98.02
955	224	11340	1.98	100.00