TECHNICAL DIGEST

2023–2024 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

Appendix A Resources, Glossary, and References

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 2023–2024

Chapter 1

Historical Overview of Assessment in Texas

Chapter 1

Overview

Assessment Timeline

Changes to the Assessment Program Over Time

Overview

Texas has a history of student assessment dating back to 1979, when the Texas Legislature required the first statewide testing program. 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 a timeline detailing the administration and content of specific assessments and a description of legislative, policy, and curriculum changes to the assessment program over time.

Assessment Timeline

—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 for 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 for 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 assessments 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 for 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 for grades 3–11; in reading for grades 3–9; in writing for grades 4 and 7; in English Language Arts (ELA) for grades 10 and 11; in science for grades 5, 10, and 11; and in social studies for grades 8, 10, and 11. Spanish versions of TAKS were administered for students 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 for kindergarten through grade 12 and in reading for 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.

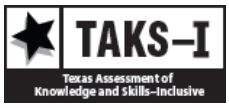
State-Developed Alternative Assessment II (SDAA II)

SDAA was replaced with SDAA II in spring 2005 to better align the alternate assessment to TAKS. For students who received special education services, SDAA II was available in mathematics for kindergarten through grade 10, in reading for kindergarten through grade 9, in writing for kindergarten through grade 9, and in ELA for grade 10. The final administration of SDAA II occurred 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, the TAKS grade 8 science assessment 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 assessments in science for grades 5, 8, 10, and 11; in science in Spanish for grade 5; in social studies for grades 8, 10, and 11; and in mathematics and ELA for grade 11. The final administration of TAKS–I was in spring 2007.

—2006–2007

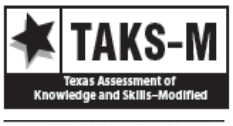
LAT administrations of TAKS in reading for grades 3–8 and in ELA for grade 10 were implemented in spring 2007 for eligible EB students.

—2007–2008

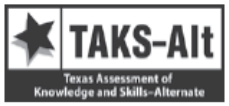
LAT administrations of TAKS in science for grades 5, 8, and 10 were implemented in spring 2008 for eligible EB students.

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 in mathematics for grades 3–11; in reading for grades 3–9; in writing for grades 4 and 7; in ELA for grades 10 and 11; in science for grades 5, 8, 10, and 11; and in social studies for grades 8, 10, and 11. The final administration of TAKS (Accommodated) was in spring 2011.



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

ELA for grades 10 and 11; science for grades 5, 8, 10, and 11; and social studies for grades 8, 10, and 11. The final administration of TAKS–Alt occurred 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

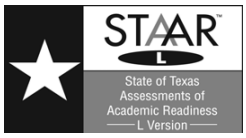
EOC assessments in Algebra II and English I were first administered on a voluntary basis.

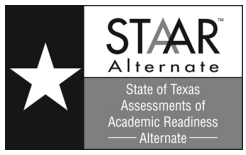
—2011–2012

STAAR replaced TAKS as the state academic assessment program beginning in spring 2012. STAAR included mathematics and reading for grades 3–8, writing for grades 4 and 7, science for grades 5 and 8, and social studies for 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 to students 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 had 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 in grades 3–8 and included 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 for grades 3–8, writing for grades 4 and 7, science for grades 5 and 8, and social studies for 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 for grades 3–8, writing for grades 4 and 7, science for grades 5 and 8, social studies for 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.

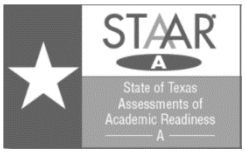
STAAR Modified EOC assessments in Algebra I, Geometry, English I reading, English I writing, English II reading, English II writing, Biology, World Geography, and World History were added to the testing program.

—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 in mathematics and reading for grades 3–8, in writing for grades 4 and 7, in science for grades 5 and 8, in social studies for grade 8, and as 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 in mathematics and reading for grades 3–8, in science for grades 5 and 8, in social studies for 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 based on each student's needs, including content supports and language and vocabulary supports.

—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, in the absence of STAAR,

that school systems and parents could choose to administer to evaluate the academic progress students made.

—2020–2021

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

—2021–2022

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

—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.

—2023–2024

Beginning with the December 2023 STAAR administration, student responses to constructed-response questions were scored using a hybrid scoring model in which the responses were initially scored by an automated scoring engine (ASE). At least 25 percent of student responses were then routed to human scorers, whose scores were also used throughout the scoring window to monitor scores generated by the ASE. The hybrid scoring model was also used for the spring 2024 STAAR and TELPAS administrations.

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 have 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 these 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 with 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 and STAAR Alternate 2 were not administered and specific STAAR EOC assessment requirements for graduation were waived for students enrolled in and completing the corresponding course. Since 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 tests by request is limited to 3 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. Holy days are 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 2023–2024

Chapter 2

Building a High-Quality Assessment System

Chapter 2

Test Development Activities

Groups Involved

Item Development and Review

Pilot Testing

Field Testing and Data Review

Test Construction

Security

Performance Tasks

Quality Control Procedures for Scoring and Reporting

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 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 assessment 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.

Step	Process
13	A final blueprint for each test is developed to establish the number of questions on the assessment and the number of test questions measuring each reporting category.
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*	Assessments 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. Assessment results are reported at the student, campus, district, regional, and state levels.
18	In accordance with state law, the Texas Assessment Program releases assessments 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 each state assessment.
20	A technical digest is developed and published annually to provide verified technical information about the assessments.

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

Groups Involved

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

TEA Assessment Development Division

The Assessment Development Division is composed of content specialists and psychometricians. The content specialists provide subject-specific 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.

TEA Student Assessment Division

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, which 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 wide-ranging 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 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 item development role for STAAR (including STAAR Spanish), STAAR Alternate 2, and TELPAS, 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 assessments, large-scale criterion-referenced measurements, and English language proficiency assessments. 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 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.

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 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 content specialists 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 and 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 assessment. • 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 students at the grade being tested. • 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 represents cultural, ethnic, social, and political diversity fairly.

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 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. 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 assessments 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 assessments.) TEA also periodically conducts standalone field tests of new items (e.g., extended constructed-response [ECR] items) by administering them to a purposefully selected representative student sample. Refer to Chapter 4, "STAAR," for detailed information about standalone field testing.

Typically, for STAAR grades 3–8 assessments, 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/subjects and courses assessed.

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

TELPAS Alternate does not include field-test questions, as the Observable Behaviors are the same each year.

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 and fit of each test item; and
- generalized Mantel-Haenszel (MH) statistics to identify, by gender and ethnicity, unexpected differences in group performance on any single item, using MH-delta for dichotomous items and standardized mean differences (SMD) for polytomous items.

Data Review Procedures

After field testing, Pearson provides TEA content specialists statistical data for each item. TEA content specialists review each flagged item and 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 bias on gender and ethnicity. Differential item functioning (DIF) flags indicate the potential for bias. Content specialists review items flagged for DIF, make a final determination on item bias, and 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 assessments for content coverage and balance.

Test Construction

Each grade/subject or course assessment is based on a specific test blueprint that guides how each assessment is constructed. Test blueprints delineate the number of items and points from each reporting category that will appear on a given assessment.

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 assessed 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.

Assessments 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, which allows for the comparison of the difficulty of each item with that of all other items in the bank. Items are selected to meet sound content and test construction practices, ensuring that assessments are approximately comparable in difficulty from one administration to the next. Refer to Chapter 3, "Standard Technical Processes," for detailed information about Rasch scaling.

Assessments 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 and STAAR Spanish, blueprints indicate the number of dichotomous and polytomous items and the number of ECR items and list a specific number of readiness and supporting standards. Items that assess 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 STAAR and 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 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 and 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. Because it is a holistic inventory that contains the Observable Behaviors, TELPAS Alternate does not include embedded accommodations.

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

Because TELPAS Alternate is a holistic inventory that contains the same Observable Behaviors every year, 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 manuals available on the Test Administration Resources webpage provide guidelines on training testing personnel, administering assessments, 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 procedures, 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 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 assessment results invalidated.

Incident Tracking

TEA regularly monitors and tracks testing irregularities and reviews all incidents reported by 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,
- 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.

Performance Tasks

STAAR, STAAR Spanish, and TELPAS include performance tasks 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 and STAAR Spanish RLA assessments use ECR questions to measure a student's ability to synthesize the component skills of writing; that is, the ECR task requires the student to express ideas effectively in writing for a specified purpose. RLA assessments include short constructed-response (SCR) questions as well as an ECR question at every grade level. Science and social studies assessments include SCR questions.

TELPAS grades 2–12 speaking assessments consist of prompts that elicit student speaking responses recorded using a headset with a microphone. TELPAS grades 2–12 writing assessments include constructed-response and sentence-rewrite items.

Scoring Rubrics

ECRs for STAAR and STAAR Spanish are evaluated using a holistic scoring process, meaning that the student response is evaluated as a whole according to preestablished criteria. These criteria are explained in detail in the scoring rubrics for each type of writing and are used to determine the effectiveness, and thus the score, of each response.

The STAAR 5-point rubric for ECRs includes two main components: 1) development and organization of ideas and 2) conventions. A student response may receive from 0 to 3 points for development and organization of ideas and up to 2 points for the use of writing conventions. ECRs receive two independent sets of scores, and the scores are combined for a final score; therefore, a student may receive up to 10 points for his or her response. The STAAR writing rubrics for ECRs can be found on the STAAR RLA Resources webpage.

SCRs in the STAAR reading domain are scored using 2-point item-specific rubrics, and SCRs in the STAAR writing domain are scored using a 1-point rubric. Rubrics for STAAR RLA SCRs are included in the constructed-response scoring guides found on the STAAR RLA Resources webpage. STAAR science and social studies SCRs are scored using 2-point item-specific rubrics, which are included in the constructed-response scoring guides found on the respective resources webpages.

TELPAS writing items are also evaluated using a holistic scoring process. Sentence-rewrite items receive a score of 0 or 1 based on the criteria defined in the rubric. A 4-point writing rubric is used to evaluate constructed responses at grades 2 and 3. Student responses receive two independent scores, and those scores are combined to calculate the student's score (from 2 to 8). For grades 4–12, 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 combined to calculate a score of 3 to 12 points. The scoring rubrics for TELPAS writing items can be found on the TELPAS Resources webpage.

TELPAS speaking responses are scored according to a 2-point or 4-point rubric, depending on the item type. The scoring rubrics for TELPAS speaking items can also be found on the TELPAS Resources webpage.

Hybrid Scoring Model

STAAR ECRs and SCRs, TELPAS grades 4–12 constructed responses and sentence rewrites, and TELPAS speaking responses are scored using a hybrid scoring model. In the hybrid scoring model, responses are initially scored by an ASE, and 25 percent are then routed to human scorers to ensure continued validity, reliability, and calibration of the assessment scoring process. Data from these two methods are continuously compared throughout the process.

For STAAR ECRs, based on slight variations in programming, the ASE assigns two independent scores to each student response, and the scores are then summed. Scores can be exact (score 1 and score 2 agree) or adjacent (score 1 and score 2 differ by no more than 1 point). In instances in which scores are discrepant (i.e., scores differ by more than 1 point), the student response is routed to a resolution queue to be resolved by human scoring. A supervisor or scoring director reviews the response and assigns a resolution score. This score is doubled and becomes the final reported score.

STAAR constructed responses that are blank and those with indecipherable language are assigned a condition code by the ASE. Other types of nonscorable responses are assigned a condition code by the ASE and then sent to a human supervisor or scoring director to confirm that the response is nonscorable. The supervisor or scoring director thoroughly reviews the response, and if it is determined to be scorable, it is assigned a score and routed to a second content scoring leader. If the supervisor or scoring director determines that the response is nonscorable, a nonscorable code is applied, and the response is routed to another content scoring leader for confirmation. While the response is under review, it is held in a review queue that prevents it from being distributed to other scorers.

Throughout scoring, TEA staff members are consulted on responses that are highly unusual or require a policy decision from TEA. Additionally, scores generated by human scorers are used throughout the scoring window to monitor scores assigned by the ASE. This process is similar to a second reader in human scoring models. More information on the hybrid scoring model used for STAAR and TELPAS is available on the STAAR Resources and TELPAS Resources webpages, respectively.

STAAR Spanish ECRs and SCRs and TELPAS grades 2 and 3 writing responses are completely scored by humans.

Human Scoring

All scorers for the Texas Assessment Program have at least a four-year college degree and 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 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. Program managers monitor all aspects of scoring for STAAR, STAAR Spanish, and TELPAS; specify the configuration of training materials; and oversee the schedule and process for performing the work.

Scorer Training Process

All scorers who work on STAAR and TELPAS performance task scoring projects receive extensive training through Pearson's online modules. This training covers the materials associated with the performance tasks for each assessment and includes orientation in the secure web-based scoring system. Scorers receive training on 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.

Training Materials

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

Scoring Process

Pearson's OSCAR Scoring System

For performance tasks scored by humans, scorers use Pearson's OSCAR scoring system to access, review, and assign scores for STAAR, STAAR Spanish, and TELPAS writing responses as well as for TELPAS speaking responses. Responses are grouped by grade/subject, course, and domain and are stored on the OSCAR server. All responses from a particular student and assessment 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 OSCAR.

Once scorers have successfully completed training and qualification, they have secure access to students' responses in OSCAR. As scorers score the responses, more responses are routed into their scoring queues. Each scorer independently reads or listens to a response and selects a score. Scoring supervisors, scoring directors, and content specialists can identify which scorer scores each response and the system provides numerous tools and reports to help supervisory staff monitor scoring. The rubric and training can also be reinforced through qualification sets delivered regularly or when needed to address a scoring issue.

In addition, Pearson uses a distributed scoring model for STAAR, STAAR Spanish, and TELPAS to allow scorers to participate in the scoring process from any location within the United States and Puerto Rico, provided they qualify and meet strict requirements. Distributed scoring 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.

Anchor Sets

In addition to field-test scoring performed by scorers, TEA and Pearson staff members independently score samples of field-test responses that will be used on 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 sets for use in scorer training and qualification. Educators assist in the review and make recommendations to reach a consensus on the scores. TEA staff review and approve all scoring guides and training sets.

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 the validity responses. Scorers failing to meet the standard for validity after remediation are dismissed from the project, and the responses they have scored are returned to the queue to be scored again.

Rescores and Appeals

If a district has questions about the score assigned to a response, a rescore can be requested in the Test Information Distribution Engine (TIDE) during specific windows. If the score changes, CAI provides rescore results by posting an updated student report card to the TIDE Secure File Center 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 outcome of the appeal and whether the score was changed.

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. Refer to Chapter 3, “Standard Technical Processes,” for more information on reliability and validity.

It is important to note that prior to implementing the hybrid scoring model for the 2023–2024 school year, TEA and CAI conducted a study to determine the reliability of the ASE. For the Spring 2023 Hybrid Scoring Study, STAAR constructed-response items administered in spring 2023 and scored by humans were reevaluated using the ASE. The results of the study show that the ASE met the performance criteria and is comparable to human scoring. The study is available on the Assessment Reports and Studies webpage.

Quality Control Procedures for Scoring and Reporting

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 to determine accelerated instruction requirements, reclassification of EB students, and eligibility for high school 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 process to review and confirm the validity of the scores and reports produced by its testing contractors.

Data and Report Processing

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

CAI deploys internal systems that support the QC process, including online test delivery, scoring, and reporting. QC in this environment verifies that the systems are properly configured and that they operate in accordance with program requirements. Once the proper system configuration has been completed and approved, the systems are moved into production.

Prior to reporting results, TEA and CAI engage in an extensive and comprehensive QC process to verify the accuracy of reports for each test administration. The joint QC process involves a

complete scoring and reporting test run. Reports are not sent to districts until all discrepancies in the QC data set are resolved and the reports generated by TEA and CAI match.

The QC process is as follows:

1. Prepare a test design for each test administration. A test design is a set of specific instructions for preparing fictitious students 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.
2. Prepare, enter, and submit online test cases for fictitious students.
    - 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. Include additions or changes necessary to reflect new reporting policies or new conditions that should be tested.
    - Create a registration file that follows the registration layout and includes all fictitious students to be assessed.
    - Create an appropriate test attributes upload file and an additional fields upload file based on accommodations, test language, and test format fields.
    - Register fictitious students for assessments using the contractor's user acceptance testing (UAT) environment.
    - Verify that the information on the test ticket matches the online assessment that is generated upon logging in.
    - Create spreadsheets for each assessment that consist of a list of students and the answer selections that should be entered for the given assessment.
    - Take assessments by selecting the answers as outlined in the spreadsheet of test cases.
    - Submit all assessments for scoring.
    - Ensure that responses, score codes, accommodations, test information attributes, and demographic information are appropriately recorded and reported.
3. Receive raw score to scale score (RSSS) conversion tables from CAI that show the scale score corresponding to each raw score point for an assessment.

- Verify, approve, and incorporate the conversion tables into computer programs that produce the student and district and campus files and reports. Refer to the Equating section in Chapter 3, “Standard Technical Processes,” for more information about RSSS conversion tables.
- Repeat the verification step using the post-equated conversion tables.

4. Create a student-level data file with data from the simulated processing of test-case student responses and merge it with historical data to create the district data file.
    - 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 contractor's 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 assessments administered.
6. Compare TEA student-level data files to CAI student-level data files.
    - For each record in the data file:
        - compare every variable in CAI 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 data files match TEA data files.
7. Produce reports.
    - Receive electronically from CAI all standard reports for three fictitious districts for the assessments administered.
    - Receive assurance from CAI that printed copies of student report cards for spring administrations and confidential student labels for all administrations meet specifications.
    - TEA and CAI generate reports using independently produced computer programs.
8. Verify reports.
    - Compare TEA reports to CAI 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 QC process is complete.
10. Verify data in the Family and Research portals by selecting a random sample of student records for each test case and executing a QC checklist confirming the following functionalities:
    - Reports are accessible by portal access code, first name, and date of birth.
    - Reports populate as expected for each test case spanning the administration type, grades, subjects, and courses.
    - Data for the selected student records appear correctly, including but not limited to student identifying information, demographic information, and both current and historical test result information.
    - All functionality within the portals 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 Google Chrome and Microsoft Edge.
    - Test cases run correctly on mobile devices such as iPads.
11. Verify the Centralized Reporting System (CRS) by executing a QC 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 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 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."

TECHNICAL DIGEST 2023–2024

Chapter 3

Standard Technical Processes

Overview
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.

Performance Standards

A critical aspect of any statewide testing program is the establishment of performance or proficiency 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:

- Masters Grade Level
- Meets Grade Level
- Approaches Grade Level

- Did Not Meet Grade Level

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

- Level III: Accomplished Academic Performance
- Level II: Satisfactory Academic Performance
- Level I: Developing 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 or proficiency 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 performance level descriptors 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 reset 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 performance level descriptors, 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 and used item mapping with external data methodology (Ferrara, Lewis, Mercado, D'Brot, Barth, & Egan, 2011; Phillips, 2012), 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 2023 as the assessment transitioned to a standardized online assessment. The standard-setting methodology used was a modification of the 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. A test-centered, criterion-referenced method was used to guide panelists in establishing the proficiency levels for each domain. 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 performance or 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, 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 DIF, are conducted using the 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.

It should be noted 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 and 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 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, 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, 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:

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

Because each assessment has a unique combination of selected standards (e.g., Meets Grade Level and Approaches Grade Level for STAAR) and Rasch-based proficiency estimates at those standards, scaling constants *A* and *B* are unique across assessments. 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.

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 STAAR Alternate 2 assessment's horizontal scale is derived using its own sample group, σ_θ varies across assessments. The STAAR Alternate 2 Level II: Satisfactory performance standards are 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 such that the scale score for Approaches Grade Level in the base grade (e.g., STAAR grade 3 mathematics) is fixed, 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, STAAR grades 3–8 RLA, and STAAR 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 Approaches Grade Level cut score for the base grade in the vertical scale, and $\theta_{Approaches}$ represents the approved Approaches Grade Level performance standard in Rasch units for the base grade 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.

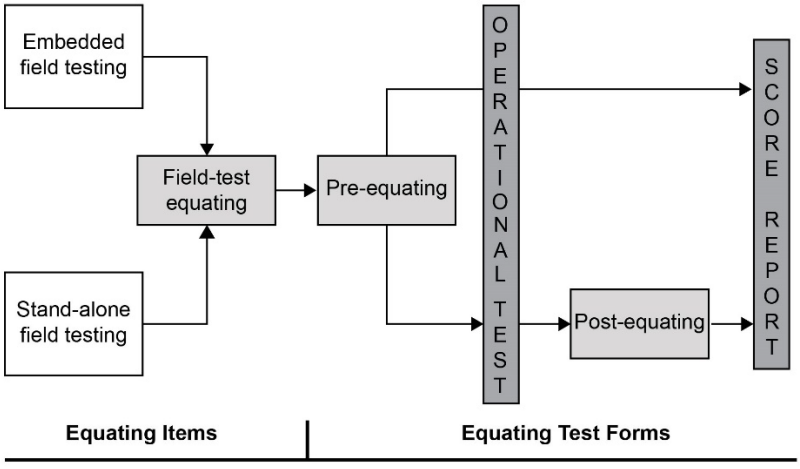
Types of Equating

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

- pre-equating test forms that are under construction
- post-equating operational test forms after administration
- 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 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 that have been equated to the Rasch scale are selected 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

When changes in item presentation (e.g., position, formatting) or instructional practice have occurred since an item was field tested, those changes might impact the estimated difficulty of the item, and post-equating might be preferred. STAAR, STAAR Spanish, STAAR Alternate 2, and TELPAS grades 2–12 are post-equated.

Post-equating for STAAR and STAAR Spanish employs a fixed common-items parameter approach where all machine-scored items are anchored to their bank value, and handscored or open-ended items are calibrated and equated to the bank scale in a single step. Displacement values are reviewed for anchored items. An item's displacement statistic can be used to identify questions with difficulties that have shifted relative to the difficulties of the other questions on the form (Linacre, 2022). Items with a displacement greater than 1 or less than -1 are reviewed by content specialists. If there is no content reason for flagging, then no action is taken, and the

item remains as is in the anchor file. If content specialists determine an edit was made between field testing and operational use or another contextual reason is determined for a change in the item performance, the item is removed from the anchor file and freely calibrated. Content changes include changes to the item stem, answer options, stimulus text, graphics, punctuation, bolding, underlining, font, or accessibility features (e.g., text-to-speech [TTS]), as well as new responses to technology-enhanced items. Contextual changes may include current events.

Although STAAR grades 3–8 mathematics and Algebra I assessments have no handscored items, items for these assessments are reviewed for displacement. If no items are flagged for displacement or no content reasons are determined for flagged items, these assessments are pre-equated. Otherwise, the flagged items are dropped from anchoring, and those item parameters are post-equated.

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

1. Tests are assembled and evaluated using Rasch-based difficulty targets.
2. Using the calibration samples, fixed common-item parameter post-equating is completed.
3. Item parameters for all items are determined.
4. Using these final item parameters, RSSS conversion tables are produced.

Post-equating for STAAR Alternate 2 and TELPAS grades 2–12 employs a mean-mean equating method with a conventional common-item nonequivalent groups design whereby an equating constant is calculated and used to transform the Rasch difficulty obtained from the current calibration to the underlying Rasch difficulty scale established by the original test form in the scaling year. This equating constant is defined as:

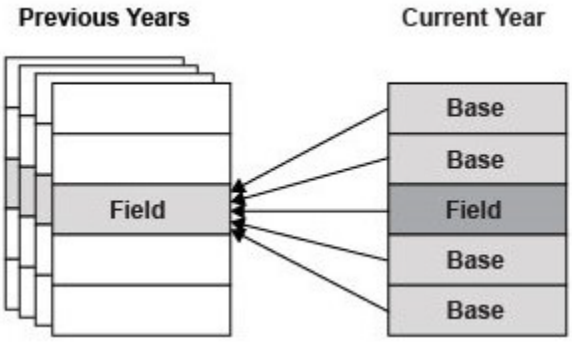
$$t_{a,b} = \frac{\sum_{i=1}^{k}(d_{i,a} - d_{i,b})}{k}, \quad (10)$$

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.

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 3.2 illustrates the source of the equating items for STAAR Alternate 2 and TELPAS. The arrows in Figure 3.2 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 3.2. 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 post-equating procedure for STAAR Alternate 2 and TELPAS grades 2–12 involves the following steps:

1. Tests are assembled and evaluated using Rasch-based difficulty targets.
2. Rasch item difficulty calibrations are conducted using the calibration samples.
3. 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.
4. The post-equating constant is applied to the Rasch difficulty estimates for the operational test items, and RSSS conversion tables are produced.

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

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, STAAR Spanish, STAAR Alternate 2, and TELPAS grades 2–12 use embedded field-test designs to collect data on field-test items. A stand-alone field test is occasionally conducted for STAAR RLA ECR items.

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 and 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 a mean-mean equating method with common item non-equivalent group design 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 (KR_{20}) 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 on 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:

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

where α is a lower-bound estimate of the true reliability, k is the number of items on 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, item subsets are multiple choice and 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 the multiple-choice component and non-multiple-choice components, 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 ECR 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, each student response receives two independent sets of scores. If the two scores differ by more than one point, the response is routed to a resolution queue to be resolved by a supervisor or scoring director. 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 scores is considered 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.

Estimates of classification accuracy and consistency are calculated using Rudner's (2000, 2005) methodology and its extensions by Li (2006). The classification accuracy and consistency indices used were developed for IRT models. The basic idea is to estimate the probability of classifying each score into each performance level conditional on each test score based on an IRT model.

The following notation is used throughout the calculations. For an assessment with four performance levels, $\text{cut}_{1|2}$, $\text{cut}_{2|3}$, and $\text{cut}_{3|4}$ denote the performance level cut scores in the Rasch student proficiency metric (θ), and $\text{cut}_{0|1}$ and $\text{cut}_{4|5}$ are set to $-\infty$ and $+\infty$, respectively. Then the probability that θ is classified as performance level *k*, where $k \in \{1,2,3,4\}$, is calculated as:

$$f\left(k, \hat{\theta}, \hat{\theta}_{\text{SE}}\right) = \Phi\left(\frac{\text{cut}_{k|(k+1)} - \hat{\theta}}{\hat{\theta}_{\text{SE}}}\right) - \Phi\left(\frac{\text{cut}_{(k-1)|k} - \hat{\theta}}{\hat{\theta}_{\text{SE}}}\right), \tag{16}$$

where Φ is a cumulative normal distribution function.

Classification consistency is calculated in two steps. First, for each student *j*, the probability of the classification being consistent is calculated as:

$$\text{consistency}_j = \sum_{k=1}^{4} f\left(k, \hat{\theta}_j, \hat{\theta}_{\text{SE}\,j}\right)^2, \tag{17}$$

where the function *f* is defined. Second, after the consistency is calculated for all students, the average is taken. This value is the overall classification consistency.

Classification accuracy is also calculated in two steps. First, for each student *j*, the probability of the classification being accurate is calculated as:

$$\text{accuracy}_j = f\left(k, \hat{\theta}_j, \hat{\theta}_{\text{SE}\,j}\right), \tag{18}$$

where *k* is the performance level to which $\hat{\theta}_j$ is classified. Second, after the accuracy is calculated for all students, the average is taken. This value is the overall classification accuracy.

Note that each STAAR EOC assessment has three different cut scores for Approaches Grade Level: 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. 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. 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 or proficiency, 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:

1. Items are developed based on the TEKS curriculum standards and item guidelines.
2. Items are reviewed for appropriateness of item content and difficulty, for alignment to the TEKS, and to eliminate potential bias.
3. Data on field-test items are collected and reviewed to determine appropriateness for

inclusion on a test.

4. Tests are built to pre-defined criteria.
5. 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 ECR 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, SCR, and ECR. 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 assessments 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 assessment, 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 when the items are field-tested. 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, and TELPAS 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. STAAR Spanish and TELPAS grades 2–3 writing constructed-response items are human scored during operational use, with the human rater implementing the rubrics. STAAR constructed-response items, TELPAS speaking responses, and TELPAS grades 4–12 writing constructed responses and sentence rewrites undergo a hybrid scoring process during operational use. 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 an assessment is designed to measure a single construct, the internal components of the assessment 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 assessment. For example, a study conducted on the structural equivalence of transadapted assessments (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 assessments.

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 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 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], Lexile® measures, Quantile® measures, 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 and proficiency levels.

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.

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 On Track or Not 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 2016 with the mathematics and reading assessments. STAAR Alternate 2 progress measures are calculated and reported for mathematics and RLA. On-track measures are also available for STAAR Alternate 2 mathematics and RLA.

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 2023–2024

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 includes 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 subjects.

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. STAAR Spanish 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 for RLA.

STAAR Interim Assessments

STAAR Interim Assessments are optional online assessments aligned to the TEKS that help educators 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 in grades 3–8 and specific high school courses 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.

STAAR testing was suspended for spring and summer 2020 due to the impact of the COVID-19 pandemic, and a STAAR EOC assessment waiver reduced the number of EOC assessments that certain students were 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 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 and 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, 2022, and 2024. 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 developed in Spanish, which allows the Spanish RLA curriculum to be assessed in a more authentic and meaningful manner. Items for STAAR Spanish mathematics and science are transadapted, which 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 (e.g., 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 reviews 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 in the Accommodations section 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 classroom instruction; however, a student cannot be required to use them during STAAR. District and campus testing 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. Forms must be submitted by the posted deadlines. Late requests might 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 2023–2024

Assessment	Assessments Administered
STAAR Grade 3 Mathematics	373,259
STAAR Grade 3 RLA	359,823
STAAR Spanish Grade 3 Mathematics	20,622
STAAR Spanish Grade 3 RLA	34,269
STAAR Grade 4 Mathematics	377,761
STAAR Grade 4 RLA	368,567
STAAR Spanish Grade 4 Mathematics	15,002
STAAR Spanish Grade 4 RLA	25,343
STAAR Grade 5 Mathematics	380,596
STAAR Grade 5 RLA	375,573
STAAR Grade 5 Science	381,003
STAAR Spanish Grade 5 Mathematics	11,560
STAAR Spanish Grade 5 RLA	18,548

Assessment	Assessments Administered
STAAR Spanish Grade 5 Science	13,049
STAAR Grade 6 Mathematics	387,459
STAAR Grade 6 RLA	394,035
STAAR Grade 7 Mathematics	324,116
STAAR Grade 7 RLA	397,573
STAAR Grade 8 Mathematics	356,728
STAAR Grade 8 RLA	402,885
STAAR Grade 8 Science	398,010
STAAR Grade 8 Social Studies	405,806
STAAR Algebra I	503,449
STAAR English I	666,003
STAAR English II	592,216
STAAR Biology	500,517
STAAR U.S. History	426,891

NOTE: For STAAR EOC assessments, the table includes the sum 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 and starting test sessions, and the Student Interface, which allows students to participate in a secure online environment for testing using the Secure Browser application. 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 1.3 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

Out-of-district (OOD) testing allows students who will be away from their home districts for a scheduled administration to take a STAAR EOC assessment at a participating district. For example, a student from Houston who spends the summer in Dallas could register to test in Dallas in June. 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

Out-of-school (OOS) testing allows examinees who are no longer enrolled in school but who still need to test to take a STAAR EOC assessment during a scheduled 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

The following are the performance level categories and policy definitions for STAAR and STAAR Spanish.

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.

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.

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.

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.

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. 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 were composed of 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)
STAAR Grade 3 Mathematics	1360	1471	1600
STAAR Grade 4 Mathematics	1462	1557	1690
STAAR Grade 5 Mathematics	1515	1634	1776
STAAR Grade 6 Mathematics	1616	1745	1889
STAAR Grade 7 Mathematics	1703	1793	1965
STAAR Grade 8 Mathematics	1754	1859	2009
STAAR Grade 3 RLA	1345	1467	1596
STAAR Grade 4 RLA	1414	1552	1663
STAAR Grade 5 RLA	1475	1592	1700
STAAR Grade 6 RLA	1535	1634	1749
STAAR Grade 7 RLA	1564	1669	1771
STAAR Grade 8 RLA	1592	1698	1803
STAAR Grade 5 Science	3550	4000	4380
STAAR Grade 8 Science	3550	4000	4619
STAAR Grade 8 Social Studies	3550	4000	4352
STAAR Spanish Grade 3 Mathematics	1360	1471	1600
STAAR Spanish Grade 4 Mathematics	1462	1557	1690
STAAR Spanish Grade 5 Mathematics	1515	1634	1776
STAAR Spanish Grade 3 RLA	1318	1447	1515
STAAR Spanish Grade 4 RLA	1408	1488	1581
STAAR Spanish Grade 5 RLA	1431	1556	1662
STAAR Spanish Grade 5 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)
STAAR Algebra I	3489	3541	3550	4000	4345
STAAR English I	3775	3775	3775	4000	4606
STAAR English II	3766	3775	3775	4000	4734
STAAR Biology	3516	3550	3550	4000	4531
STAAR 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 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, Lexile measures, Quantile measures, 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 Masters Grade Level, Meets Grade Level, Approaches Grade Level, or Did Not Meet 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 OOS examinees) 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 is available on the Literacy and Lexile Measures and The Quantile Framework for Mathematics pages of 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. More information about the EL performance measure can be found on the Progress Measures webpage.

Assessment Reports

TEA provides reports of student performance on STAAR and 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 and 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 Assessment Results section 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 Masters Grade Level, Meets Grade Level, and Approaches 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 highest percentage achieving the Meets Grade Level performance standard, and which group has the lowest percentage meeting the Approaches 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. Because the tests are designed to measure content areas within the required state curriculum, 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 levels of achievement, progress measures describe students' levels of 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. Progress on STAAR 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 Accelerated, Expected, and Limited in relation to progress targets. The progress targets define the expectation of annual progress for each grade and content area and are grounded in the STAAR performance standards and the goals of having all students achieve Meets Grade Level or above and having high-performing students maintain Masters Grade Level performance.

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. STAAR progress measure results are available for grades 4–8

mathematics, grades 4–8 RLA, Algebra I, English I, and English II and are summarized in Appendix B.

STAAR On-Track Measure

Although 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. STAAR on-track measure results from the spring administration are summarized in Appendix B.

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 and STAAR Spanish use 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

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

Table 4.5. Horizontal Scaling Constants for STAAR EOC Assessments

Assessment	*A*	*B*
STAAR Algebra I	460.7351	3919.0028
STAAR English I	429.3074	3845.4064
STAAR English II	444.4006	3852.8590
STAAR Biology	435.9620	4042.0267
STAAR 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 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 versions. Separate vertical scales were established for the English and the Spanish versions of grades 3–5 RLA.

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

Grade	Language	Content Area	*A*	*B*	V_g
3	English and Spanish	Mathematics	130.0052	1454.3188	0
4	English and Spanish	Mathematics	130.0052	1454.3188	0.5911
5	English and 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 test administrations can be compared. TEA uses pre-equating for all STAAR assessments during test construction. STAAR assessments without constructed-response items are reviewed for item stability. If all items pass statistical and content review checks, these assessments are scored using pre-equating. STAAR assessments that include constructed-response items are post-equated to obtain item parameters for the constructed-response items. These assessments are then scored using the resulting post-equating.

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 and STAAR Spanish through an embedded-field-test design for all tests. In some years, stand-alone field tests are conducted for STAAR and STAAR Spanish. 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 assessment 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 and 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 assessments. 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 quality 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 CCRS, which specify the knowledge and skills necessary to succeed in entry-level community college and university courses, 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. Composed of Texas educators, item review committees are convened for all STAAR assessments and revise and edit items, as appropriate, prior to field testing. 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 and 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. Content experts then review student responses to resolve scoring discrepancies or uncertainties.

For SCR and ECR questions, rubrics are developed to evaluate student responses. All rubrics for STAAR are validated by educator committees and content specialists. 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 requested by district personnel.

The ASE used in the hybrid scoring model for STAAR constructed responses is trained for each item, and trait where applicable, using the human-scored data. STAAR constructed responses are initially scored by the ASE, and for STAAR ECRs the ASE assigns two independent scores for each dimension. Approximately 25 percent of STAAR constructed responses scored by the ASE are routed for human scoring, including a random sample of 10 percent of responses to monitor performance and responses routed for verification scoring of unusual or difficult-to-score responses. For any response routed for human scoring, the human score is used when determining a student's score. All STAAR Spanish SCRs and ECRs are scored entirely by humans.

Score reliability for STAAR and STAAR Spanish constructed responses is generated and evaluated in terms of scorer agreement rates and the commonly used kappa with quadratic weights (Fleiss & Cohen, 1973). For STAAR, the exact agreement rate, adjacent agreement rate, and total agreement rate (exact and adjacent) between both the ASE and human scores are computed and evaluated (refer to Tables 4.7 and 4.9). For STAAR Spanish, these rates are computed between the two human scorers (refer to Tables 4.8 and 4.10). For more information on hybrid scoring and reliability, refer to the Spring 2023 Hybrid Scoring Study, which is available on the Assessment Reports and Studies webpage.

Table 4.7. Summary of Score Agreement (Reliability) for Spring 2024 STAAR RLA ECRs

ECR Trait	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Grade 3 RLA					
Ideas	29,625	72%	27%	99%	0.78
Conventions	29,625	70%	29%	99%	0.72
STAAR Grade 4 RLA					
Ideas	31,955	63%	34%	97%	0.78
Conventions	31,955	67%	32%	98%	0.70
STAAR Grade 5 RLA					
Ideas	34,128	68%	28%	97%	0.79
Conventions	34,128	72%	25%	97%	0.73
STAAR Grade 6 RLA					
Ideas	35,671	66%	31%	97%	0.83
Conventions	35,671	69%	28%	97%	0.72
STAAR Grade 7 RLA					
Ideas	36,722	68%	31%	99%	0.82
Conventions	36,722	67%	31%	98%	0.72

ECR Trait	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Grade 8 RLA					
Ideas	36,407	67%	31%	99%	0.83
Conventions	36,407	71%	28%	99%	0.77
STAAR English I					
Ideas	42,085	71%	28%	99%	0.86
Conventions	42,085	73%	26%	99%	0.79
STAAR English II					
Ideas	41,233	63%	34%	97%	0.80
Conventions	41,233	69%	29%	97%	0.72

Table 4.8. Summary of Score Agreement (Reliability) for Spring 2024 STAAR Spanish RLA ECRs

ECR Trait	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Spanish Grade 3 RLA					
Ideas	34,134	72%	24%	96%	0.76
Conventions	34,134	76%	23%	99%	0.73
STAAR Spanish Grade 4 RLA					
Ideas	25,248	70%	26%	96%	0.66
Conventions	25,248	76%	23%	97%	0.62
STAAR Spanish Grade 5 RLA					
Ideas	18,492	66%	33%	97%	0.74
Conventions	18,492	70%	27%	97%	0.70

Table 4.9. Summary of Scorer Agreement (Reliability) for Spring 2024 STAAR SCRs

SCR Indicator	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Grade 3 RLA					
Reading SCR	34,095	80%	20%	100%	0.81
Writing SCR	33,369	97%	3%	100%	0.94
STAAR Grade 4 RLA					
Reading SCR	36,375	72%	27%	99%	0.74
Writing SCR	34,987	97%	3%	100%	0.94
STAAR Grade 5 RLA					
Reading SCR	37,146	75%	24%	99%	0.77
Writing SCR	36,453	93%	7%	100%	0.86
STAAR Grade 6 RLA					
Reading SCR	38,719	81%	18%	100%	0.84
Writing SCR	38,352	95%	5%	100%	0.88
STAAR Grade 7 RLA					
Reading SCR	39,107	78%	22%	99%	0.80
Writing SCR	38,489	98%	2%	100%	0.96
STAAR Grade 8 RLA					
Reading SCR	39,187	81%	19%	100%	0.82
Writing SCR	39,229	94%	6%	100%	0.86
STAAR English I					
Reading SCR	47,152	82%	18%	100%	0.82
Writing SCR	45,710	98%	2%	100%	0.96
STAAR English II					
Reading SCR	44,815	79%	21%	100%	0.81
Writing SCR	44,157	94%	6%	100%	0.87
STAAR Grade 5 Science					
SCR 1	36,890	87%	12%	100%	0.89
STAAR Grade 8 Science					
SCR 1	37,398	87%	13%	100%	0.85
SCR 2	38,470	87%	12%	100%	0.88

SCR Indicator	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Biology					
SCR 1	39,380	97%	3%	100%	0.97
SCR 2	39,558	94%	5%	100%	0.96
STAAR Grade 8 Social Studies					
SCR 1	37,888	93%	7%	99%	0.93
SCR 2	38,038	83%	17%	100%	0.80
STAAR U.S. History					
SCR 1	37,471	85%	14%	99%	0.88
SCR 2	34,054	87%	13%	100%	0.84

Table 4.10. Summary of Scorer Agreement (Reliability) for Spring 2024 STAAR Spanish SCRs

SCR Indicator	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Spanish Grade 3 RLA					
Reading SCR	34,043	95%	4%	99%	0.88
STAAR Spanish Grade 4 RLA					
Reading SCR	25,228	94%	5%	99%	0.86
STAAR Spanish Grade 5 RLA					
Reading SCR	18,511	81%	18%	99%	0.75
Writing SCR	18,487	93%	6%	99%	0.87
STAAR Spanish Grade 5 Science					
Reading SCR	13,017	95%	4%	99%	0.89

Validity of human scoring 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 scorer'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 rates in Tables 4.11 and 4.12 are expressed in terms of exact agreement between the score assigned by a given scorer and the true score approved by scoring leaders.

Table 4.11. Summary of Validity Results for Spring 2024 STAAR RLA ECRs

ECR Trait	Exact Agreement Rate
STAAR Grade 3 RLA	
Ideas	85%
Conventions	83%
STAAR Grade 4 RLA	
Ideas	84%
Conventions	87%
STAAR Grade 5 RLA	
Ideas	81%
Conventions	83%
STAAR Grade 6 RLA	
Ideas	81%
Conventions	81%
STAAR Grade 7 RLA	
Ideas	79%
Conventions	76%
STAAR Grade 8 RLA	
Ideas	81%
Conventions	83%
STAAR English I	
Ideas	85%
Conventions	86%
STAAR English II	
Ideas	86%
Conventions	88%
STAAR Spanish Grade 3 RLA	
Ideas	79%
Conventions	78%
STAAR Spanish Grade 4 RLA	
Ideas	78%
Conventions	81%
STAAR Spanish Grade 5 RLA	
Ideas	73%
Conventions	78%

Table 4.12. Summary of Validity Results for Spring 2024 STAAR SCRs

SCR Indicator	Exact Agreement Rate
STAAR Grade 3 RLA, Reading SCR	94%
STAAR Grade 3 RLA, Writing SCR	99%
STAAR Grade 4 RLA, Reading SCR	89%
STAAR Grade 4 RLA, Writing SCR	99%
STAAR Grade 5 RLA, Reading SCR	83%
STAAR Grade 5 RLA, Writing SCR	97%
STAAR Grade 6 RLA, Reading SCR	92%
STAAR Grade 6 RLA, Writing SCR	98%
STAAR Grade 7 RLA, Reading SCR	94%
STAAR Grade 7 RLA, Writing SCR	99%
STAAR Grade 8 RLA, Reading SCR	93%
STAAR Grade 8 RLA, Writing SCR	99%
STAAR English I, Reading SCR	93%
STAAR English I, Writing SCR	98%
STAAR English II, Reading SCR	94%
STAAR English II, Writing SCR	95%
STAAR Grade 5 Science, SCR 1	96%
STAAR Grade 8 Science, SCR 1	91%
STAAR Grade 8 Science, SCR 2	91%
STAAR Biology, SCR 1	97%
STAAR Biology, SCR 2	98%
STAAR Grade 8 Social Studies, SCR 1	89%
STAAR Grade 8 Social Studies, SCR 2	81%
STAAR U.S. History, SCR 1	98%
STAAR U.S. History, SCR 2	93%
STAAR Spanish Grade 3 RLA, Writing SCR	99%
STAAR Spanish Grade 4 RLA, Writing SCR	94%
STAAR Spanish Grade 5 RLA, Reading SCR	96%
STAAR Spanish Grade 5 RLA, Writing SCR	96%
STAAR Spanish Grade 5 Science, SCR 1	98%

Evidence Based on Internal Structure

TEA collects evidence that shows the relationship of students' responses between items, within reporting categories of items, and within full tests to verify that the elements of an assessment conform to the intended test construct. TEA also conducts internal consistency studies to gather evidence based on internal structure. The internal consistency of STAAR is evaluated using KR_{20} 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 and 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, Lexile measures, Quantile measures, STAAR Interim Assessments)
- college students taking STAAR studies, which link performance on STAAR EOC assessments to college course grades

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 on a regular basis after a program has been in place for some time.

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 an indicator of a student's level of proficiency in a content area or specific course.
- Performance on STAAR is an indicator of a student's readiness for the next grade level or course in the same content area.
- Performance on STAAR is an indicator of a student's possible need for academic intervention.
- Performance on STAAR across years provides an 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 an indicator of overall student proficiency at a district or campus.
- STAAR performance results can be aggregated to provide an 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 an 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 2023–2024

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. Implemented in the 2014–2015 school year, STAAR Alternate 2 fulfills ESSA, which requires that all students be assessed in specific grades and subjects throughout their academic careers, and IDEA, which 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 specific to each item on STAAR Alternate 2.

STAAR Alternate 2 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.

With the redesign of STAAR, STAAR Alternate 2 reading and writing assessments were combined into an RLA assessment for each grade to mirror STAAR. 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 may be 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. In accordance with the requirements of ESSA, the STAAR Alternate 2 participation requirements were revised in 2023–2024 to clarify the scope of students that are assessed with STAAR Alternate 2. 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 must review 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, "Building a High-Quality Assessment System," 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 or strand statements act as a bridge between grade-level content standards and STAAR Alternate 2 prerequisite skills. Specific statements are selected each year and provided to educators in the fall, allowing time for planning instruction and developing 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 and presentation of the state assessments according to his or her individual response modes. 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 students with disabilities to enable them 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 they 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, 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 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 components of the Texas Assessment Program. This online resource contains preparation and administration procedures for each state-required assessment, and the version for the new school year 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 Alternate 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. STAAR Alternate 2 is administered during a five-week window in the spring, and retest opportunities are not offered. The number of students who took each STAAR Alternate 2 assessment is shown in Table 5.1.

Table 5.1. STAAR Alternate 2 Assessments Administered in 2023–2024

Assessment	Assessments Administered
Grade 3 Mathematics	6,620
Grade 3 RLA	6,623
Grade 4 Mathematics	6,491
Grade 4 RLA	6,489
Grade 5 Mathematics	6,033
Grade 5 RLA	6,032
Grade 5 Science	6,033
Grade 6 Mathematics	5,586
Grade 6 RLA	5,585
Grade 7 Mathematics	5,227
Grade 7 RLA	5,230
Grade 8 Mathematics	4,984
Grade 8 RLA	4,985
Grade 8 Science	4,982
Grade 8 Social Studies	4,988
Algebra I	5,115
English I	5,117
English II	4,683
Biology	5,029
U.S. History	4,475

Each STAAR Alternate 2 test question measures a targeted prerequisite skill. A cluster of four 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 students 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 supports or 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 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.

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 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.

Standard Setting

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 were composed of 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.

Initial performance standards for all STAAR Alternate 2 assessments were established in 2015, and performance standards for STAAR Alternate 2 RLA were reset in 2023 with the redesign of the assessments. The current performance standards for STAAR Alternate 2 are provided in Table 5.2.

Table 5.2. STAAR Alternate 2 Performance Standards

Assessment	Level II: Satisfactory	Level III: Accomplished
Grade 3 Mathematics	300	375
Grade 4 Mathematics	300	387
Grade 5 Mathematics	300	379
Grade 6 Mathematics	300	373
Grade 7 Mathematics	300	375
Grade 8 Mathematics	300	365

Assessment	Level II: Satisfactory	Level III: Accomplished
Grade 3 RLA	300	388
Grade 4 RLA	300	380
Grade 5 RLA	300	374
Grade 6 RLA	300	370
Grade 7 RLA	300	378
Grade 8 RLA	300	371
Grade 5 Science	300	387
Grade 8 Science	300	382
Grade 8 Social Studies	300	372
Algebra I	300	361
English I	300	365
English II	300	370
Biology	300	383
U.S. History	300	368

Refer to the STAAR Alternate 2 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 test 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 III: Accomplished Academic

Performance, Level II: Satisfactory Academic Performance, or Level I: Developing 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 Assessment Results section 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 III, Level II, and Level I 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 classifying the student's scores from the previous year and the current year in terms of the stage of performance achieved and then 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. STAAR Alternate 2 progress measures are available for grades 4–8 mathematics, grades 4–8 RLA, Algebra I, English I, and English II, and are summarized in Appendix C.

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

Content 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
RLA	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

Content Area	Grade/Course	*A*	*B*
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 are followed while developing STAAR Alternate 2 assessments for the Texas Assessment Program, supporting 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 quality 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, curriculum framework documents, 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 or strand 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 to document 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, as described in the Performance Standards section of this chapter. 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 an indicator of a student's level of proficiency in a content area or specific course.
- Performance on STAAR Alternate 2 is an indicator of a student's readiness for the next grade or course in the same content area.
- Performance on STAAR Alternate 2 is an indicator of a student's possible need for academic intervention.
- Performance on STAAR Alternate 2 across years provides an 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 an indicator of overall student proficiency at a district or campus.
- STAAR Alternate 2 performance results can be aggregated to provide an 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 an 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. 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 2023–2024

Chapter 6

Texas English Language Proficiency Assessment System

Overview

Participation Requirements

Test Development

Training

Test Administration

Proficiency Standards

Scores and Reports

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

Participation Requirements

All EB students in kindergarten through grade 12 are required to participate in TELPAS unless they meet 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 circumstances 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 one or more TELPAS language domains for reasons associated with the student's particular disability. Participation decisions must be considered individually for each student on a domain-by-domain basis. The reason for not assessing the student must 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. 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. 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.

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 Grade 3
Grades 4–5	Grades 4–5
Grades 6–8	Grades 6–7 Grades 8–9
Grades 9–12	Grades 10–12

TEA developed the TELPAS holistically rated components in collaboration with test development experts, bilingual and ESL consultants, and focus groups including teachers, bilingual and ESL directors, assessment directors, campus administrators, and university professors. Like the TELPAS grades 2–12 online assessments, the holistically rated 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. The educator guide is designed to familiarize educators with TELPAS and describes the integral relationship between TELPAS and the ELPS, including explanatory information on the TELPAS language domains as well as examples of classroom instruction and annotated test question descriptions. The educator guide also includes information on test design, training, and test results.

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 English language proficiency. 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 the version for the new school year is available prior to the annual ESC training.

For holistically rated TELPAS assessments, 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 proficiency level descriptors (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. Raters 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 online 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 calibration set, the rater attempts a second and final 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 is indicated in Table 6.2.

Table 6.2. TELPAS Assessments Administered in 2023–2024

Grade	Listening	Speaking	Reading	Writing
Kindergarten	101,812	101,560	101,599	101,557
Grade 1	107,163	106,961	106,941	106,868
Grade 2	105,121	105,117	105,181	105,176
Grade 3	105,764	105,750	105,806	105,799
Grade 4	105,945	105,933	106,017	106,013

Grade	Listening	Speaking	Reading	Writing
Grade 5	106,555	106,546	106,609	106,605
Grade 6	105,561	105,559	105,582	105,578
Grade 7	104,909	104,904	104,997	104,993
Grade 8	101,475	101,470	101,565	101,559
Grade 9	109,474	109,472	109,615	109,608
Grade 10	86,285	86,284	86,472	86,465
Grade 11	68,599	68,598	68,714	68,711
Grade 12	51,160	51,158	51,248	51,249
Total	1,259,823	1,259,312	1,260,346	1,260,181

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, EB students in grades 2–12 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 and starting test sessions, and the Student Interface, which allows students to participate in a secure online environment for testing using the Secure Browser application. 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 1.3 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.

Holistically Rated Assessments

A holistic rating process is used for kindergarten and grade 1 for all four language domains. Raters are specially trained to use the ELPS PLDs as holistic rating rubrics to 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.

Proficiency Standards

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

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.

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

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.

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 descriptor criteria in Table 6.5. All criteria listed for a particular rating must be met for a student to receive that rating.

Table 6.5. TELPAS Composite Rating Descriptors

A student will receive a composite rating of	if he or she achieves
Beginning	• a composite score that fails to meet the Intermediate requirements
Intermediate	• a 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
Advanced	• a 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
Advanced High	• a composite score of 3.5 or higher, and • a minimum proficiency level of Advanced in all domains

Figure 6.1 provides a student example to show how composite results are generated.

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 table 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 assessed 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 × 0.33) + (Speaking × 0.33) + (Reading × 0.33)

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

Composite Score = (2 × 0.33) + (2 × 0.33) + (1 × 0.33) = 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.

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 is provided on the summary reports for each assessed grade that contain information about every student for whom a TELPAS record was submitted.

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 Assessment Results section 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, 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.

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.

Scales 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 for TELPAS are provided in Table 6.6.

Table 6.6. 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
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. Refer to the Description of Scores section for more information.

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

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

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.

Tables 6.7 and 6.8 show the median response time per item, 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 45-Second 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	6.4	6.4	5.2	3.4
Advanced	12.8	11.9	14.9	7.7
Advanced High	27.0	21.4	32.2	18.4

Table 6.8. TELPAS Speaking Median Response Time per 90-Second 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	17.7	18.4	9.8	12.2
Advanced	34.6	37.1	24.8	28.3
Advanced High	58.5	57.6	52.7	52.6

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 are estimated based on their internal consistency values on the continuous scale. The results of these analyses, presented in Table 6.9, show that the weighted TELPAS composite scores have reliability estimates of at least 0.921.

Table 6.9. TELPAS Composite Score Reliability Estimates

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 2 (*n* = 102,399)	Listening	2.800	0.936	0.774	0.921
	Speaking	1.846	0.777	0.834	
	Reading	1.908	0.882	0.766	
	Writing	1.785	0.799	0.818	
Grade 3 (*n* = 104,636)	Listening	3.277	0.882	0.795	0.933
	Speaking	2.154	0.888	0.843	
	Reading	2.306	1.103	0.834	
	Writing	1.931	0.838	0.815	
Grade 4 (*n* = 105,087)	Listening	2.652	1.042	0.799	0.933
	Speaking	2.357	0.874	0.828	
	Reading	2.709	0.994	0.809	
	Writing	2.170	0.915	0.836	
Grade 5 (*n* = 105,804)	Listening	2.937	1.042	0.819	0.939
	Speaking	2.423	0.914	0.836	
	Reading	3.008	0.992	0.829	
	Writing	2.472	0.966	0.833	
Grade 6 (*n* = 104,669)	Listening	2.907	0.886	0.755	0.925
	Speaking	2.338	0.757	0.808	
	Reading	2.667	1.064	0.823	
	Writing	2.376	0.860	0.803	
Grade 7 (*n* = 103,792)	Listening	2.993	0.894	0.768	0.930
	Speaking	2.261	0.782	0.822	
	Reading	2.788	1.070	0.836	
	Writing	2.500	0.888	0.809	
Grade 8 (*n* = 99,958)	Listening	3.099	0.893	0.778	0.924
	Speaking	2.263	0.801	0.834	
	Reading	2.843	0.938	0.815	
	Writing	2.244	0.802	0.767	
Grade 9 (*n* = 105,052)	Listening	2.735	0.952	0.813	0.938
	Speaking	2.198	0.990	0.900	
	Reading	2.804	0.956	0.826	
	Writing	2.194	0.817	0.771	

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 10 (*n* = 82,937)	Listening	2.815	0.931	0.790	0.938
	Speaking	2.246	0.997	0.902	
	Reading	2.551	0.950	0.817	
	Writing	2.465	1.038	0.815	
Grade 11 (*n* = 65,790)	Listening	2.859	0.905	0.782	0.936
	Speaking	2.291	1.009	0.902	
	Reading	2.619	0.949	0.814	
	Writing	2.541	1.029	0.816	
Grade 12 (*n* = 49,320)	Listening	2.800	0.901	0.781	0.930
	Speaking	2.195	1.008	0.902	
	Reading	2.555	0.924	0.812	
	Writing	2.474	1.005	0.790	

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 are followed while developing TELPAS assessments for the Texas Assessment Program, supporting 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 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 6.10.

Table 6.10. 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

Proficiency Level	Degree of Linguistic Accommodation Applied to Stimulus and Item Development	
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.

Holistically Rated Assessments

TELPAS holistically rated assessments are aligned with the ELPS and 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 that these trainings 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 by item type. A variety of item types (e.g., multiple-choice, fill in the blank, drag and drop, hot spots) and response interactions are used on TELPAS 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."

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 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 writing constructed-response and sentence-rewrite items and speaking constructed-response items, rubrics are used by human scorers 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 requested by district personnel.

The ASE used in the hybrid scoring model for TELPAS grades 4–12 writing is trained for each item and dimension using the human-scored data. Responses are initially scored by the ASE, and approximately 25 percent are routed for human scoring to monitor performance and verify scoring of unusual or difficult-to-score responses. For any response routed for human scoring, the human score is used when determining a student's score. TELPAS writing constructed responses and sentence-rewrite items in grades 2 and 3 are scored completely by humans.

Score reliability for TELPAS grades 2–12 writing constructed-response and sentence-rewrite items is generated and evaluated in terms of scorer agreement rates and the commonly used

quadratic weighted kappa (Fleiss & Cohen, 1973). For grades 2 and 3, the agreement rates between the human scorers are computed and evaluated on a 100 percent reliability read for the constructed-response items and a 5 percent reliability read for the sentence-rewrite items. For grades 4–12, the agreement rates between the ASE and human scores are computed and evaluated on a 10 percent random sample. Tables 6.11–6.13 provide reliability information for TELPAS writing constructed-response and sentence-rewrite items.

Table 6.11. Summary of Scorer Agreement (Reliability) for TELPAS Grades 2 and 3 Constructed Responses

Item	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Grade 2					
Item 1	103,544	88%	12%	100%	0.82
Item 2	103,472	86%	14%	100%	0.82
Item 3	103,431	85%	15%	100%	0.81
Grade 3					
Item 1	105,252	78%	21%	99%	0.78
Item 2	105,350	75%	24%	99%	0.81
Item 3	105,134	80%	20%	100%	0.80

Table 6.12. Summary of Scorer Agreement (Reliability) for TELPAS Grades 4–12 Constructed Responses

Item	Dimension	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
			Exact	Adjacent	Exact + Adjacent	
Grades 4–5						
Item 1	Completeness	20,400	79%	21%	100%	0.89
	Usage	20,400	78%	21%	100%	0.87
	Vocabulary	20,400	77%	23%	100%	0.86
Item 2	Completeness	20,053	77%	23%	100%	0.85
	Usage	20,053	77%	23%	100%	0.83
	Vocabulary	20,053	79%	21%	100%	0.85
Item 3	Completeness	20,664	75%	25%	100%	0.86
	Usage	20,664	75%	24%	100%	0.84
	Vocabulary	20,664	75%	25%	100%	0.85

Item	Dimension	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
			Exact	Adjacent	Exact + Adjacent	
Grades 6–7						
Item 1	Completeness	20,057	81%	19%	100%	0.88
	Usage	20,057	78%	21%	100%	0.85
	Vocabulary	20,057	80%	20%	100%	0.87
Item 2	Completeness	20,466	79%	21%	100%	0.85
	Usage	20,466	77%	22%	100%	0.83
	Vocabulary	20,466	79%	21%	100%	0.85
Item 3	Completeness	20,501	81%	19%	100%	0.90
	Usage	20,501	77%	23%	100%	0.86
	Vocabulary	20,501	76%	24%	100%	0.87
Grades 8–9						
Item 1	Completeness	20,457	81%	19%	100%	0.85
	Usage	20,457	81%	19%	100%	0.83
	Vocabulary	20,457	79%	21%	100%	0.82
Item 2	Completeness	19,860	78%	22%	100%	0.87
	Usage	19,860	76%	24%	100%	0.84
	Vocabulary	19,860	79%	21%	100%	0.87
Item 3	Completeness	19,678	79%	21%	100%	0.86
	Usage	19,678	77%	23%	100%	0.83
	Vocabulary	19,678	76%	24%	100%	0.83
Grades 10–12						
Item 1	Completeness	19,205	80%	20%	100%	0.88
	Usage	19,205	81%	19%	100%	0.87
	Vocabulary	19,205	84%	16%	100%	0.90
Item 2	Completeness	19,617	83%	17%	100%	0.90
	Usage	19,617	80%	19%	100%	0.87
	Vocabulary	19,617	80%	20%	100%	0.87
Item 3	Completeness	19,470	75%	25%	100%	0.86
	Usage	19,470	75%	25%	100%	0.84
	Vocabulary	19,470	76%	24%	100%	0.85

Table 6.13. Summary of Scorer Agreement (Reliability) for TELPAS Sentence-Rewrite Items

Grade or Grade Band	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
Grade 2	5,189	98%	2%	100%	0.96
Grade 3	5,270	97%	3%	100%	0.93
Grades 4–5	21,203	98%	2%	100%	0.95
Grades 6–7	20,589	96%	4%	100%	0.89
Grades 8–9	20,994	97%	3%	100%	0.94
Grades 10–12	20,481	97%	3%	100%	0.93

In the hybrid scoring model used for TELPAS speaking, student responses 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 the ASE and human scores. A random sample of 15 percent of student responses per grade band is selected for human scoring. The grade band correlations between the total raw scores on the human-scored samples and the total raw scores on the ASE-scored samples are presented in Table 6.14.

Table 6.14. TELPAS Speaking Inter-Rater Correlations

Grade Band	Number of Responses	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

Validity of human scoring 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.

Holistically Rated 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 KR_{20} 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.

Holistically Rated Assessments

Evidence of the validity of TELPAS holistically rated 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.

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 to document 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 an indicator of a student's level of proficiency in learning English.
- Proficiency on TELPAS is an indicator of a student's possible need for academic intervention.
- Proficiency on TELPAS across years provides an 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. Distribution of TELPAS

proficiency levels by administration since 2012 are provided in Appendix D, but comparisons in performance are only appropriate across certain years due to redesign of the TELPAS assessments. For example, TELPAS writing results for all grades can be compared for the periods of 2012–2022 and 2023–present. For grades 2–12 listening and speaking, results can be compared for the periods of 2012–2017 and 2018–present. For reading, results can be compared for the periods of 2012–2013, 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. When a discontinuity occurred in the performance level of any domain, it also appeared in the composite scores and ratings.

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 2023–2024

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 EB/EL in PEIMS whose parents have declined bilingual or ESL program services.

EB students in kindergarten and grade 1 are assessed holistically in all four language domains using TELPAS. 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 the ARD committee, in conjunction with the LPAC, must review carefully 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 must review 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. The educator guide is designed to familiarize educators with TELPAS Alternate and describes the integral relationship between TELPAS Alternate and the ELPS, including explanatory information on the TELPAS Alternate language domains as well as a sample of Observable Behaviors for reading. The educator guide also includes information on test design, training, and test results.

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 English language proficiency. 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 the version for the new school year is available prior to the annual ESC training.

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 the same six-week testing window as TELPAS. 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. Some EB students use ASL, braille, or another method of communication as a substitute for traditional English in one or more domains. 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 is indicated in Table 7.1.

Table 7.1. TELPAS Alternate Assessments Administered in 2023–2024

Grade	Assessments Administered
Grade 2	1,899
Grade 3	1,588
Grade 4	1,452
Grade 5	1,198
Grade 6	987
Grade 7	854
Grade 8	732
Grade 9	705
Grade 10	507
Grade 11	362
Grade 12	322

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. The current proficiency standard ranges for TELPAS Alternate are provided in Table 7.2.

Table 7.2. TELPAS Alternate Proficiency Standards

Domain	Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
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 descriptor criteria in Table 7.4. All criteria listed for a particular rating must be met for a student to receive that rating.

Table 7.4. TELPAS Alternate Composite Rating Descriptors

A student will receive a composite rating of	if he or she achieves
Awareness	• a composite score that fails to meet the Imitation requirements
Imitation	• a composite score of 1.5 or higher, and • 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, and • 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, and • 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, and • a minimum proficiency level of Early Independence in all domains

Figure 7.1 provides a student example to show how composite results are generated.

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 table 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 Assessment Results section 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, 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 lowest obtainable scale scores of 600 and 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 items. Values for the horizontal scaling constants for TELPAS Alternate 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. Refer to the Description of Scores section for more information.

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. 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 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	2.907	1.284	0.929	0.984
Speaking	2.816	1.394	0.950	
Reading	2.315	1.285	0.945	
Writing	2.593	1.388	0.947	

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 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 cognitive 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 comprehensive 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 performance on 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 assessments 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. TELPAS Reading Proficiency Compared to STAAR Alternate 2 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
Grade 3	Awareness	572	312	69
	Imitation	378	346	92
	Early Independence	340	357	96
	Developing Independence	189	362	98
	Basic Fluency	77	369	99
Grade 4	Awareness	376	308	69
	Imitation	327	337	90
	Early Independence	340	345	94
	Developing Independence	255	354	96
	Basic Fluency	133	365	99
Grade 5	Awareness	274	307	65
	Imitation	217	333	84
	Early Independence	257	344	89
	Developing Independence	240	357	95
	Basic Fluency	190	363	98
Grade 6	Awareness	252	314	77
	Imitation	176	341	89
	Early Independence	192	345	94
	Developing Independence	166	355	95
	Basic Fluency	185	374	99
Grade 7	Awareness	209	317	69
	Imitation	126	342	89
	Early Independence	178	354	96
	Developing Independence	172	365	97
	Basic Fluency	147	380	98
Grade 8	Awareness	150	318	71
	Imitation	119	335	84
	Early Independence	137	352	94
	Developing Independence	149	362	97
	Basic Fluency	152	366	97

Grade or Course	TELPAS Alternate Reading Proficiency Level	Number of Students	STAAR Alternate 2 RLA Average Scale Score*	STAAR Alternate 2 Passing Rate
English I	Awareness	192	333	85
	Imitation	136	348	93
	Early Independence	133	360	98
	Developing Independence	121	366	98
	Basic Fluency	112	370	100
English II	Awareness	111	331	77
	Imitation	84	339	88
	Early Independence	99	357	95
	Developing Independence	105	371	98
	Basic Fluency	97	382	100

* A scale score of 300 is necessary to meet the STAAR Alternate 2 Level II: Satisfactory performance level.

Table 7.7. TELPAS Writing Proficiency Compared to STAAR Alternate 2 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
Grade 3	Awareness	684	317	72
	Imitation	434	353	94
	Early Independence	310	358	98
	Developing Independence	105	364	96
	Basic Fluency	23	373	96
Grade 4	Awareness	472	313	72
	Imitation	388	339	92
	Early Independence	347	351	95
	Developing Independence	163	360	98
	Basic Fluency	61	367	100
Grade 5	Awareness	334	309	67
	Imitation	279	339	87
	Early Independence	279	350	93
	Developing Independence	186	361	97
	Basic Fluency	100	367	99

Grade or Course	TELPAS Alternate Writing Proficiency Level	Number of Students	STAAR Alternate 2 RLA Average Scale Score*	STAAR Alternate 2 Passing Rate
Grade 6	Awareness	311	318	79
	Imitation	199	342	91
	Early Independence	207	348	95
	Developing Independence	153	366	96
	Basic Fluency	101	378	100
Grade 7	Awareness	236	317	72
	Imitation	156	349	89
	Early Independence	191	358	96
	Developing Independence	137	371	99
	Basic Fluency	112	381	98
Grade 8	Awareness	177	318	72
	Imitation	129	343	90
	Early Independence	167	356	95
	Developing Independence	126	361	96
	Basic Fluency	108	370	96
English I	Awareness	211	330	84
	Imitation	148	358	97
	Early Independence	136	361	99
	Developing Independence	104	362	98
	Basic Fluency	95	373	100
English II	Awareness	116	327	77
	Imitation	111	347	91
	Early Independence	95	362	95
	Developing Independence	90	371	100
	Basic Fluency	84	385	100

* 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 to document 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 an indicator of a student's level of proficiency in learning English.
- Proficiency on TELPAS Alternate is an indicator of a student's possible need for academic intervention.
- Proficiency on TELPAS Alternate across years provides an 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 based on the policy definitions for district- or campus-level performance:

- 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. TELPAS Alternate has been administered since 2019, and 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 2023–2024

Appendix A

Resources, Glossary, and References

Resources

Glossary of Statistical Terminology

References

Resources

- Student Assessment Overview
- STAAR Resources
- STAAR Alternate 2 Resources
- TELPAS Resources
- TELPAS Alternate Resources
- *District and Campus Coordinator Resources*
- Accommodation Resources
- Test Administration Resources
- Assessments for Special Populations
- Student Assessment Results
- Assessment Reports and Studies
- Research Portal
- TexasAssessment.gov
- Assessment Resources for Testing Personnel
- Assessment Resources for Educators
- Assessment Resources for Students and Families

Glossary of Statistical Terminology

This glossary provides definitions for the statistical terms that appear in the tables and graphs in the scale score descriptive statistics tables and frequency distribution of scale scores figures in appendices B, C, D, and E.

Descriptive Statistics

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).

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:

$$\text{Kurtosis} = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{X_i - \bar{X}}{S_X} \right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)},$$

where X_i is the score for student i, $\bar{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

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.

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:

$$\text{Skewness} = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \bar{X}}{S_X} \right)^3,$$

where X_i is the score for student i, $\bar{X}$ is the mean score, S_X is the standard deviation, and N is the total number of students who took the assessment.

Standard Deviation (SD)—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:

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N-1}},$$

where X_i is the score for student i, $\bar{X}$ 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.

Frequency Distributions

Frequency (Freq)—The frequency is the number of students who obtained a 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 a particular score point on the assessment.

Percentage (Pct)—The percentage is the percentage of students who obtained a particular score point on the assessment. 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 a particular score point on the assessment. It is computed as: CumPct = CumFreq ÷ N × 100.

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.

Li, S. (2006). Evaluating the consistency and accuracy of proficiency classifications using item response theory. (Publication No. 5761) [Doctoral dissertation, University of Massachusetts]. Scholarworks UMass. https://scholarworks.umass.edu/dissertations_1/5761

Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.

Linacre, J. M. (2022). *Winsteps® Rasch measurement computer program*. 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 2023–2024

Appendix B

STAAR Statistical Tables and Figures

List of Tables

Table B.1.1. Spring 2024 STAAR Grades 3–8 Classification Consistency and Accuracy

Table B.1.2. Spring 2024 STAAR Spanish Grades 3–5 Classification Consistency and Accuracy

Table B.1.3. Spring 2024 STAAR EOC Assessments Classification Consistency and Accuracy

Table B.2.1. Spring 2024 STAAR Grades 3–8 Scale Score Correlations

Table B.2.2. Spring 2024 STAAR Spanish Grades 3–5 Scale Score Correlations

Table B.2.3. Spring 2024 STAAR EOC Assessments Scale Score Correlations

Table B.3.1. Spring 2024 STAAR Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.2. Spring 2024 STAAR Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.3. Spring 2024 STAAR Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.4. Spring 2024 STAAR Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.5. Spring 2024 STAAR Grade 5 and Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

Table B.3.6. Spring 2024 STAAR Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

Table B.3.7. Spring 2024 STAAR Spanish Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.8. Spring 2024 STAAR Spanish Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.9. Spring 2024 STAAR Spanish Grade 5 Science Conditional Standard Error of Measurement for Scale Scores

Table B.3.10. Spring 2024 STAAR EOC Assessments Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.11. Spring 2024 STAAR EOC Assessments RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.12. Spring 2024 STAAR EOC Assessments Science Conditional Standard Error of Measurement for Scale Scores

Table B.3.13. Spring 2024 STAAR EOC Assessments Social Studies Conditional Standard Error of Measurement for Scale Scores

Table B.4.1. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.2. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.3. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.4. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.5. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.6. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.7. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.8. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.9. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.10. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.11. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.12. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.13. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.14. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.15. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.16. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.17. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.18. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.19. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.20. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.21. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.22. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.23. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.24. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.25. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.26. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.27. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.28. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.29. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.30. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.31. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.32. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.33. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.34. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.35. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.36. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.37. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.38. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.39. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.40. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.41. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.42. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.43. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.44. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.45. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.46. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.47. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.48. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.49. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.50. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.51. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.52. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.53. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.54. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.55. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.56. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.57. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.58. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.59. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.60. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.61. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.62. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.63. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.64. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.65. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.66. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.67. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.68. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.69. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.70. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.71. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.72. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.73. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.74. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.75. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.76. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.77. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.78. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.79. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.80. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.81. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.82. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.83. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.84. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.85. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.86. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.87. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.88. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.89. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.90. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.91. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.92. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.93. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.94. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.95. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.96. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.97. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.98. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.99. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.100. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.101. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.102. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.103. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.104. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.105. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.106. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.107. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.108. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.109. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.110. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.111. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.112. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.5.1. Spring 2024 STAAR Grade 4 Mathematics Progress Measure

Table B.5.2. Spring 2024 STAAR Grade 4 RLA Progress Measure

Table B.5.3. Spring 2024 STAAR Grade 5 Mathematics Progress Measure

Table B.5.4. Spring 2024 STAAR Grade 5 RLA Progress Measure

Table B.5.5. Spring 2024 STAAR Grade 6 Mathematics Progress Measure

Table B.5.6. Spring 2024 STAAR Grade 6 RLA Progress Measure

Table B.5.7. Spring 2024 STAAR Grade 7 Mathematics Progress Measure

Table B.5.8. Spring 2024 STAAR Grade 7 RLA Progress Measure

Table B.5.9. Spring 2024 STAAR Grade 8 Mathematics Progress Measure

Table B.5.10. Spring 2024 STAAR Grade 8 RLA Progress Measure

Table B.5.11. Spring 2024 STAAR Spanish Grade 4 Mathematics Progress Measure

Table B.5.12. Spring 2024 STAAR Spanish Grade 4 RLA Progress Measure

Table B.5.13. Spring 2024 STAAR Spanish Grade 5 Mathematics Progress Measure

Table B.5.14. Spring 2024 STAAR Spanish Grade 5 RLA Progress Measure

Table B.5.15. Spring 2024 STAAR EOC Assessments Algebra I Progress Measure

Table B.5.16. Spring 2024 STAAR EOC Assessments English I Progress Measure

Table B.5.17. Spring 2024 STAAR EOC Assessments English II Progress Measure

Table B.6.1. Spring 2024 STAAR Grades 3–8 and English I On-Track Measure

Table B.7.1. Spring 2024 STAAR Grades 3–8 Scale Score Descriptive Statistics

Table B.7.2. Spring 2024 STAAR Spanish Grades 3–5 Scale Score Descriptive Statistics

Table B.7.3. Spring 2024 STAAR EOC Assessments Scale Score Descriptive Statistics

Table B.8.1. Spring 2024 STAAR Grades 3–8 Mathematics English Learner Performance Measure

Table B.8.2. Spring 2024 STAAR Grades 3–8 RLA English Learner Performance Measure

Table B.8.3. Spring 2024 STAAR Grade 5 and Grade 8 Science English Learner Performance Measure

Table B.8.4. Spring 2024 STAAR Grade 8 Social Studies English Learner Performance Measure

Table B.8.5. Spring 2024 STAAR EOC Assessments English Learner Performance Measure

List of Figures

Figure B.7.1. Spring 2024 STAAR Grade 3 Mathematics Frequency Distribution of Scale Scores
Figure B.7.2. Spring 2024 STAAR Grade 4 Mathematics Frequency Distribution of Scale Scores
Figure B.7.3. Spring 2024 STAAR Grade 5 Mathematics Frequency Distribution of Scale Scores
Figure B.7.4. Spring 2024 STAAR Grade 6 Mathematics Frequency Distribution of Scale Scores
Figure B.7.5. Spring 2024 STAAR Grade 7 Mathematics Frequency Distribution of Scale Scores
Figure B.7.6. Spring 2024 STAAR Grade 8 Mathematics Frequency Distribution of Scale Scores
Figure B.7.7. Spring 2024 STAAR Grade 3 RLA Frequency Distribution of Scale Scores
Figure B.7.8. Spring 2024 STAAR Grade 4 RLA Frequency Distribution of Scale Scores
Figure B.7.9. Spring 2024 STAAR Grade 5 RLA Frequency Distribution of Scale Scores
Figure B.7.10. Spring 2024 STAAR Grade 6 RLA Frequency Distribution of Scale Scores
Figure B.7.11. Spring 2024 STAAR Grade 7 RLA Frequency Distribution of Scale Scores
Figure B.7.12. Spring 2024 STAAR Grade 8 RLA Frequency Distribution of Scale Scores
Figure B.7.13. Spring 2024 STAAR Grade 5 Science Frequency Distribution of Scale Scores
Figure B.7.14. Spring 2024 STAAR Grade 8 Science Frequency Distribution of Scale Scores
Figure B.7.15. Spring 2024 STAAR Grade 8 Social Studies Frequency Distribution of Scale Scores
Figure B.7.16. Spring 2024 STAAR Spanish Grade 3 Mathematics Frequency Distribution of Scale Scores
Figure B.7.17. Spring 2024 STAAR Spanish Grade 4 Mathematics Frequency Distribution of Scale Scores
Figure B.7.18. Spring 2024 STAAR Spanish Grade 5 Mathematics Frequency Distribution of Scale Scores
Figure B.7.19. Spring 2024 STAAR Spanish Grade 3 RLA Frequency Distribution of Scale Scores
Figure B.7.20. Spring 2024 STAAR Spanish Grade 4 RLA Frequency Distribution of Scale Scores
Figure B.7.21. Spring 2024 STAAR Spanish Grade 5 RLA Frequency Distribution of Scale Scores
Figure B.7.22. Spring 2024 STAAR Spanish Grade 5 Science Frequency Distribution of Scale Scores
Figure B.7.23. Spring 2024 STAAR EOC Assessments Algebra I Frequency Distribution of Scale Scores
Figure B.7.24. Spring 2024 STAAR EOC Assessments English I Frequency Distribution of Scale Scores
Figure B.7.25. Spring 2024 STAAR EOC Assessments English II Frequency Distribution of Scale Scores
Figure B.7.26. Spring 2024 STAAR EOC Assessments Biology Frequency Distribution of Scale Scores
Figure B.7.27. Spring 2024 STAAR EOC Assessments U.S. History Frequency Distribution of Scale Scores

Spring 2024 STAAR Classification Consistency and Accuracy

Table B.1.1. Spring 2024 STAAR Grades 3–8
Classification Consistency and Accuracy

	Decision Consistency				Decision Accuracy			
Grade	**Mathematics**	**RLA**	**Science**	**Social Studies**	**Mathematics**	**RLA**	**Science**	**Social Studies**
3	67.6	70.8			76.3	78.9		
4	70.6	70.2			78.6	78.5		
5	69.8	70.7	67.7		78.1	78.7	75.9	
6	69.9	70.7			78.3	78.7		
7	73.9	71.9			81.0	79.5		
8	69.2	70.0	69.3	71.7	77.6	78.3	77.7	79.2

Table B.1.2. Spring 2024 STAAR Spanish Grades 3–5
Classification Consistency and Accuracy

	Decision Consistency			Decision Accuracy		
Grade	**Mathematics**	**RLA**	**Science**	**Mathematics**	**RLA**	**Science**
3	72.0	76.4		79.7	82.7	
4	76.5	72.5		83.1	79.4	
5	73.7	71.4	81.9	81.1	79.3	86.9

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 2024 STAAR EOC Assessments
Classification Consistency and Accuracy

Test	Design Consistency	Design Accuracy
Algebra I	71.2	79.1
English I	78.2	84.2
English II	77.3	83.4
Biology	71.5	79.4
U.S. History	74.8	82.0

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 2024 STAAR
Scale Score Correlations

Table B.2.1. Spring 2024 STAAR Grades 3–8 Scale Score Correlations

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	357,595	0.77
4	Mathematics	RLA	365,361	0.75
5	Mathematics	RLA	370,614	0.74
	Mathematics	Science	375,191	0.76
	RLA	Science	373,900	0.76
6	Mathematics	RLA	382,268	0.74
7	Mathematics	RLA	315,486	0.72
8	Mathematics	RLA	277,626	0.68
	Mathematics	Science	282,599	0.75
	Mathematics	Social Studies	278,115	0.67
	RLA	Science	383,797	0.79
	RLA	Social Studies	398,175	0.79
	Science	Social Studies	386,928	0.81

Table B.2.2. Spring 2024 STAAR Spanish Grades 3–5 Scale Score Correlations

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	19,874	0.68
4	Mathematics	RLA	14,333	0.66
5	Mathematics	RLA	11,058	0.65
	Mathematics	Science	10,011	0.62
	RLA	Science	12,435	0.65

Table B.2.3. Spring 2024 STAAR EOC Assessments Scale Score Correlations

STAAR	STAAR	N	Correlation
Algebra I	English I	303,544	0.67
Algebra I	English II	38,769	0.50
Algebra I	Biology	264,326	0.71
Algebra I	U.S. History	14,945	0.50
English I	English II	61,951	0.64
English I	Biology	369,858	0.80
English I	U.S. History	25,062	0.61
English II	Biology	64,367	0.75
English II	U.S. History	70,653	0.77
Biology	U.S. History	15,868	0.76

Spring 2024 STAAR
Conditional Standard Error of Measurement for Scale Scores

Table B.3.1. Spring 2024 STAAR Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	860		910		1000	
1	934	133	1025	133	1087	133
2	1031	96	1121	96	1182	96
3	1089	80	1179	79	1240	79
4	1132	70	1221	70	1283	70
5	1166	64	1255	63	1317	63
6	1195	59	1284	59	1345	59
7	1221	56	1309	55	1370	55
8	1244	53	1331	53	1392	52
9	1265	51	1351	50	1412	50
10	1284	50	1370	49	1431	48
11	1303	48	1388	47	1448	47
12	1321	48	1405	46	1464	46
13	1338	47	1421	45	1480	45
14	1354	46	1437	45	1495	44
15	*1360	46	1452	44	1509	43
16	1387	46	*1462	44	*1515	43
17	1403	46	1481	43	1538	43
18	1419	46	1496	43	1552	43
19	1435	46	1510	43	1566	42
20	1451	46	1524	43	1580	43
21	1467	46	1539	43	1594	43
22	**1471	47	1553	43	1608	43
23	1501	48	**1557	44	1622	43
24	1519	48	1583	44	**1634	43
25	1537	50	1598	45	1651	44
26	1557	51	1613	45	1666	44
27	1577	52	1629	46	1681	45
28	1599	54	1646	47	1696	45
29	***1600	56	1664	49	1713	46
30	1647	59	1683	50	1729	47
31	1675	62	***1690	52	1747	48
32	1707	66	1724	54	1765	50

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1744	73	1748	57	***1776	51
34	1789	82	1775	61	1806	53
35	1850	97	1805	65	1829	56
36	1948	134	1841	71	1854	60
37	2070		1885	81	1884	64
38			1944	97	1919	71
39			2042	133	1962	80
40			2130		2021	96
41					2117	133
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 2024 STAAR Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1070		1150		1240	
1	1199	133	1237	131	1316	131
2	1295	96	1328	93	1409	94
3	1354	80	1383	77	1465	78
4	1398	71	1422	67	1506	69
5	1432	64	1454	61	1539	62
6	1462	60	1481	57	1566	58
7	1487	56	1504	54	1590	54
8	1510	53	1525	51	1612	52
9	1531	51	1545	50	1632	49
10	1550	49	1563	48	1650	48
11	1568	47	1580	47	1667	46
12	1584	46	1597	46	1683	45
13	1600	45	1613	45	1698	44
14	1615	44	1628	44	1712	43
15	*1616	43	1643	43	1726	42
16	1644	42	1657	43	1739	41
17	1657	42	1671	42	1752	41
18	1671	42	1685	42	*1754	40
19	1684	41	1698	42	1777	40
20	1697	41	*1703	41	1789	40
21	1710	41	1725	41	1801	39
22	1723	41	1738	41	1813	39
23	1736	41	1751	41	1825	39
24	**1745	41	1764	41	1837	39
25	1762	42	1777	41	1848	39
26	1776	42	1790	42	**1859	39
27	1790	43	**1793	42	1872	39
28	1804	43	1817	42	1884	40
29	1818	44	1831	43	1896	40
30	1833	45	1845	43	1909	40
31	1849	46	1860	44	1922	41
32	1866	47	1875	45	1935	42

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1883	49	1891	46	1948	42
34	***1889	51	1907	47	1962	43
35	1923	53	1925	48	1977	44
36	1946	56	1943	50	1992	45
37	1971	59	1963	51	***2009	47
38	2001	64	***1965	54	2026	48
39	2035	71	2007	56	2044	50
40	2079	80	2033	60	2064	52
41	2138	96	2063	65	2086	55
42	2235	133	2098	71	2111	59
43	2350		2141	80	2140	63
44			2200	96	2174	70
45			2297	133	2216	79
46			2400		2274	95
47					2369	132
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 2024 STAAR Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	720		820		830	
1	842	146	933	147	931	147
2	946	105	1038	106	1037	106
3	1010	87	1102	88	1101	88
4	1056	77	1149	77	1149	78
5	1094	70	1187	70	1187	71
6	1125	65	1219	65	1220	66
7	1153	61	1247	62	1248	62
8	1178	58	1272	59	1274	59
9	1200	56	1295	56	1297	57
10	1221	54	1316	54	1319	55
11	1240	52	1335	52	1340	53
12	1259	51	1354	51	1359	52
13	1276	49	1371	49	1377	51
14	1292	48	1388	48	1395	50
15	1308	47	1403	47	1412	49
16	1323	46	*1414	46	1428	48
17	1338	46	1433	45	1444	48
18	*1345	45	1447	44	1460	47
19	1366	44	1460	43	*1475	47
20	1379	43	1472	42	1491	46
21	1392	43	1484	41	1505	46
22	1405	42	1496	40	1520	46
23	1417	42	1507	40	1535	45
24	1429	41	1517	39	1549	45
25	1441	41	1528	38	1563	44
26	1452	41	1538	38	1576	44
27	1464	41	1548	38	1589	43
28	**1467	41	**1552	38	**1592	42
29	1487	41	1568	38	1615	42
30	1498	41	1579	39	1626	41
31	1510	41	1589	39	1638	40
32	1522	42	1600	40	1649	39

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1535	42	1611	40	1659	39
34	1547	43	1622	41	1670	39
35	1560	43	1634	42	1681	39
36	1573	44	1647	43	1691	40
37	1587	45	1660	44	***1700	41
38	***1596	45	***1663	45	1714	42
39	1616	46	1688	46	1727	43
40	1631	48	1704	48	1740	45
41	1647	49	1720	49	1755	47
42	1665	51	1737	51	1771	49
43	1683	53	1756	53	1788	51
44	1703	55	1777	56	1808	54
45	1726	58	1800	59	1829	57
46	1751	62	1826	63	1854	62
47	1780	67	1856	68	1883	67
48	1814	74	1892	76	1917	74
49	1857	84	1937	86	1960	85
50	1917	102	2000	105	2021	103
51	2016	143	2104	146	2123	145
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 2024 STAAR Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	880		890		980	
1	993	147	978	147	1072	147
2	1100	107	1083	106	1179	106
3	1165	89	1147	88	1243	88
4	1214	78	1195	78	1291	78
5	1252	71	1233	71	1329	71
6	1285	66	1265	66	1362	66
7	1314	62	1293	62	1390	62
8	1339	59	1319	59	1416	59
9	1363	57	1342	56	1438	56
10	1384	54	1363	54	1460	54
11	1404	53	1383	53	1479	52
12	1423	51	1402	51	1497	50
13	1441	50	1419	50	1515	49
14	1458	49	1436	49	1531	48
15	1474	48	1452	48	1546	46
16	1489	47	1468	47	1560	45
17	1504	46	1483	46	1574	44
18	1519	45	1497	45	1587	43
19	1533	45	1511	44	*1592	42
20	*1535	44	1524	43	1611	41
21	1560	43	1537	42	1622	40
22	1572	42	1549	42	1633	39
23	1584	41	1561	41	1644	38
24	1595	40	*1564	40	1654	38
25	1606	39	1584	40	1663	37
26	1616	38	1594	39	1673	37
27	1626	37	1605	39	1682	37
28	**1634	36	1616	39	1692	37
29	1644	35	1626	39	**1698	37
30	1652	35	1637	39	1710	37
31	1661	35	1647	39	1720	37
32	1669	35	1658	39	1729	37

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1678	35	**1669	39	1739	38
34	1687	36	1680	40	1749	38
35	1696	37	1691	40	1759	39
36	1706	38	1702	41	1770	39
37	1716	39	1714	41	1781	40
38	1726	40	1726	42	1792	41
39	1738	41	1738	42	***1803	41
40	***1749	42	1750	43	1816	42
41	1762	43	1763	44	1828	43
42	1775	44	***1771	45	1842	44
43	1789	45	1791	46	1856	46
44	1804	47	1806	47	1871	47
45	1819	48	1822	49	1887	49
46	1836	50	1839	50	1904	51
47	1854	52	1857	53	1923	53
48	1874	55	1877	55	1944	56
49	1896	58	1900	59	1966	59
50	1922	63	1926	63	1992	63
51	1951	68	1955	68	2022	68
52	1987	75	1991	76	2058	75
53	2032	86	2037	87	2102	86
54	2094	104	2100	105	2164	104
55	2198	146	2205	146	2267	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 2024 STAAR Grade 5 and Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1140		1000	
1	1408	576	1346	641
2	1834	420	1804	461
3	2099	353	2082	382
4	2298	314	2285	337
5	2460	287	2448	306
6	2598	268	2585	284
7	2719	252	2705	267
8	2828	240	2812	253
9	2928	231	2910	243
10	3020	222	3000	234
11	3106	215	3084	226
12	3187	210	3162	220
13	3264	205	3237	215
14	3338	201	3309	210
15	3409	197	3377	206
16	3478	194	3444	203
17	3545	192	3508	200
18	*3550	190	*3550	198
19	3675	188	3633	196
20	3739	187	3693	195
21	3802	187	3753	194
22	3865	187	3812	193
23	3927	187	3871	193
24	3991	188	3930	193
25	**4000	190	3989	193
26	4120	192	**4000	194
27	4187	195	4109	196
28	4257	199	4170	197
29	4331	205	4233	199
30	***4380	212	4297	202
31	4493	221	4362	205
32	4585	232	4430	209

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	4688	247	4501	213
34	4807	267	4575	219
35	4948	295	***4619	225
36	5125	336	4736	232
37	5369	406	4825	241
38	5774	566	4921	252
39	6200		5028	266
40			5147	283
41			5283	305
42			5445	336
43			5648	382
44			5925	461
45			6384	641
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 2024 STAAR Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

		Grade 8
Raw	**SS**	**CSEM**
0	1050	
1	1325	586
2	1750	423
3	2009	353
4	2201	311
5	2355	283
6	2485	263
7	2598	247
8	2700	235
9	2792	225
10	2877	216
11	2956	209
12	3030	203
13	3100	198
14	3167	193
15	3230	189
16	3292	186
17	3351	183
18	3409	180
19	3465	178
20	3520	176
21	*3550	174
22	3626	173
23	3678	172
24	3729	171
25	3780	170
26	3831	170
27	3882	170
28	3933	171
29	3984	171
30	**4000	172
31	4088	174
32	4141	176

(Continued)

	Grade 8	
Raw	**SS**	**CSEM**
33	4196	178
34	4252	181
35	4310	184
36	***4352	188
37	4435	193
38	4502	199
39	4575	207
40	4652	215
41	4737	226
42	4831	239
43	4938	255
44	5060	276
45	5207	304
46	5390	347
47	5642	418
48	6059	582
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 2024 STAAR Spanish Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	860		910		1000	
1	934	133	1025	133	1087	133
2	1031	96	1121	96	1182	96
3	1089	80	1179	79	1240	79
4	1132	70	1221	70	1283	70
5	1166	64	1255	63	1317	63
6	1195	59	1284	59	1345	59
7	1221	56	1309	55	1370	55
8	1244	53	1331	53	1392	52
9	1265	51	1351	50	1412	50
10	1284	50	1370	49	1431	48
11	1303	48	1388	47	1448	47
12	1321	48	1405	46	1464	46
13	1338	47	1421	45	1480	45
14	1354	46	1437	45	1495	44
15	*1360	46	1452	44	1509	43
16	1387	46	*1462	44	*1515	43
17	1403	46	1481	43	1538	43
18	1419	46	1496	43	1552	43
19	1435	46	1510	43	1566	42
20	1451	46	1524	43	1580	43
21	1467	46	1539	43	1594	43
22	**1471	47	1553	43	1608	43
23	1501	48	**1557	44	1622	43
24	1519	48	1583	44	**1634	43
25	1537	50	1598	45	1651	44
26	1557	51	1613	45	1666	44
27	1577	52	1629	46	1681	45
28	1599	54	1646	47	1696	45
29	***1600	56	1664	49	1713	46
30	1647	59	1683	50	1729	47
31	1675	62	***1690	52	1747	48
32	1707	66	1724	54	1765	50

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1744	73	1748	57	***1776	51
34	1789	82	1775	61	1806	53
35	1850	97	1805	65	1829	56
36	1948	134	1841	71	1854	60
37	2070		1885	81	1884	64
38			1944	97	1919	71
39			2042	133	1962	80
40			2130		2021	96
41					2117	133
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 2024 STAAR Spanish Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	600		680		720	
1	726	156	783	156	826	157
2	838	112	895	112	939	113
3	906	93	963	93	1008	94
4	956	82	1013	82	1059	83
5	996	75	1053	75	1101	76
6	1030	69	1087	70	1135	70
7	1060	65	1117	65	1166	66
8	1086	62	1143	62	1193	63
9	1110	59	1168	60	1218	60
10	1133	57	1190	57	1240	58
11	1153	55	1211	56	1261	56
12	1173	54	1231	54	1281	54
13	1191	52	1249	53	1300	53
14	1209	51	1267	52	1318	51
15	1225	50	1284	51	1334	50
16	1241	49	1300	50	1350	49
17	1256	47	1316	49	1366	48
18	1271	46	1332	48	1381	47
19	1284	45	1346	47	1395	46
20	1297	44	1361	47	1408	45
21	1309	43	1375	46	1421	44
22	*1318	42	1389	46	*1431	43
23	1332	41	1402	45	1446	43
24	1343	40	*1408	45	1458	42
25	1354	40	1428	44	1469	42
26	1364	40	1441	44	1480	41
27	1374	39	1453	44	1491	41
28	1384	40	1466	44	1502	41
29	1395	40	1478	44	1513	41
30	1405	40	**1488	44	1524	41
31	1416	41	1504	44	1536	42
32	1427	42	1517	45	1547	42

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1438	42	1530	45	**1556	43
34	**1447	43	1544	46	1572	44
35	1463	44	1558	47	1585	45
36	1476	45	1572	48	1598	46
37	1490	46	***1581	49	1613	47
38	1504	47	1604	50	1628	49
39	***1515	49	1620	51	1644	50
40	1535	50	1638	53	1661	52
41	1552	52	1656	54	***1662	54
42	1570	54	1676	56	1699	56
43	1590	56	1697	58	1721	59
44	1611	59	1721	61	1744	61
45	1634	62	1746	64	1770	65
46	1661	66	1774	68	1799	69
47	1692	72	1807	73	1833	74
48	1729	79	1845	81	1872	81
49	1775	90	1894	92	1921	92
50	1840	110	1959	110	1987	111
51	1948	154	2067	153	2096	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 2024 STAAR Spanish Grade 5 Science Conditional Standard Error of Measurement for Scale Scores

		Grade 5
Raw	**SS**	**CSEM**
0	1140	
1	1408	576
2	1834	420
3	2099	353
4	2298	314
5	2460	287
6	2598	268
7	2719	252
8	2828	240
9	2928	231
10	3020	222
11	3106	215
12	3187	210
13	3264	205
14	3338	201
15	3409	197
16	3478	194
17	3545	192
18	*3550	190
19	3675	188
20	3739	187
21	3802	187
22	3865	187
23	3927	187
24	3991	188
25	**4000	190
26	4120	192
27	4187	195
28	4257	199
29	4331	205
30	***4380	212
31	4493	221
32	4585	232

(Continued)

		Grade 5
Raw	**SS**	**CSEM**
33	4688	247
34	4807	267
35	4948	295
36	5125	336
37	5369	406
38	5774	566
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 2024 STAAR EOC Assessments Mathematics Conditional Standard Error of Measurement for Scale Scores

	Algebra I	
Raw	**SS**	**CSEM**
0	1500	
1	1897	468
2	2231	336
3	2433	279
4	2581	245
5	2698	222
6	2797	205
7	2883	193
8	2959	183
9	3028	175
10	3092	168
11	3151	162
12	3206	157
13	3258	153
14	3308	149
15	3355	146
16	3400	143
17	3444	141
18	3486	139
19	3489	137
20	*3550	135
21	3607	134
22	3645	132
23	3683	131
24	3720	131
25	3757	130
26	3794	129
27	3830	129
28	3865	128
29	3901	128
30	3937	128
31	3972	128
32	**4000	128

(Continued)

	Algebra I	
Raw	**SS**	**CSEM**
33	4044	129
34	4080	129
35	4116	130
36	4153	131
37	4190	132
38	4228	133
39	4267	134
40	4306	136
41	***4345	137
42	4388	140
43	4431	142
44	4476	145
45	4522	148
46	4571	151
47	4622	155
48	4676	160
49	4734	166
50	4796	173
51	4863	181
52	4938	191
53	5022	203
54	5119	220
55	5235	243
56	5380	277
57	5579	334
58	5910	467
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 2024 STAAR EOC Assessments RLA Conditional Standard Error of Measurement for Scale Scores

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1750		1650	
1	1993	436	1915	452
2	2304	313	2238	325
3	2491	259	2434	270
4	2628	228	2577	237
5	2737	207	2692	215
6	2829	191	2789	200
7	2909	180	2873	188
8	2980	170	2948	178
9	3045	163	3016	170
10	3104	157	3079	164
11	3160	151	3137	158
12	3211	147	3192	154
13	3260	143	3244	150
14	3307	140	3293	146
15	3352	137	3340	143
16	3395	134	3385	140
17	3436	132	3429	138
18	3476	130	3471	135
19	3515	128	3511	133
20	3552	127	3551	131
21	3589	125	3589	129
22	3625	123	3626	127
23	3660	122	3662	125
24	3695	121	3696	123
25	3728	119	3730	121
26	3761	118	3762	118
27	*3775	116	*3775	116
28	3824	114	3822	113
29	3854	113	3850	111
30	3883	111	3877	109
31	3911	109	3903	107
32	3938	108	3929	105

(Continued)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	3965	106	3953	104
34	3991	105	3978	104
35	**4000	104	**4000	104
36	4042	104	4026	104
37	4067	104	4051	105
38	4092	104	4076	106
39	4117	105	4102	108
40	4143	106	4129	110
41	4170	108	4156	112
42	4197	110	4185	114
43	4226	112	4215	117
44	4256	114	4246	119
45	4287	117	4279	121
46	4320	120	4312	124
47	4354	123	4348	127
48	4390	126	4385	130
49	4428	129	4423	133
50	4468	132	4464	136
51	4510	136	4507	140
52	4554	140	4553	145
53	4601	144	4601	150
54	***4606	149	4654	156
55	4705	155	4711	163
56	4763	162	***4734	171
57	4828	171	4843	181
58	4900	183	4922	194
59	4984	198	5014	210
60	5086	220	5124	233
61	5215	252	5262	266
62	5394	308	5454	322
63	5698	432	5773	450
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 2024 STAAR EOC Assessments Science Conditional Standard Error of Measurement for Scale Scores

		Biology
Raw	**SS**	**CSEM**
0	1900	
1	2185	444
2	2505	320
3	2699	266
4	2842	235
5	2956	214
6	3054	199
7	3139	187
8	3215	178
9	3285	171
10	3349	165
11	3410	160
12	3466	155
13	3516	151
14	*3550	148
15	3621	145
16	3669	143
17	3715	141
18	3759	139
19	3803	137
20	3845	135
21	3887	134
22	3928	133
23	3968	132
24	**4000	131
25	4046	130
26	4085	130
27	4123	129
28	4162	129
29	4200	129
30	4238	130
31	4277	130
32	4316	131

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	4356	132
34	4396	133
35	4437	134
36	4479	136
37	4522	138
38	***4531	140
39	4612	143
40	4660	146
41	4711	150
42	4764	155
43	4821	160
44	4882	166
45	4948	174
46	5021	183
47	5102	195
48	5196	210
49	5308	232
50	5447	263
51	5638	318
52	5953	443
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 2024 STAAR EOC Assessments Social Studies Conditional Standard Error of Measurement for Scale Scores

	U.S. History	
Raw	**SS**	**CSEM**
0	1420	
1	1762	495
2	2114	355
3	2326	293
4	2480	257
5	2602	233
6	2705	215
7	2793	201
8	2872	190
9	2943	181
10	3007	174
11	3067	168
12	3122	162
13	3175	157
14	3224	153
15	3271	149
16	3316	146
17	3358	143
18	3400	140
19	3439	138
20	3478	136
21	3486	134
22	*3550	132
23	3587	131
24	3621	129
25	3655	128
26	3688	127
27	3721	125
28	3753	124
29	3784	124
30	3815	123
31	3846	122
32	3876	121

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
33	3906	121
34	3936	120
35	3966	120
36	3995	120
37	**4000	119
38	4054	119
39	4083	119
40	4112	119
41	4141	119
42	4170	119
43	4199	119
44	4228	120
45	4258	120
46	4287	120
47	4317	121
48	4347	121
49	4377	122
50	4408	123
51	***4424	123
52	4471	124
53	4503	125
54	4535	127
55	4569	128
56	4603	130
57	4637	131
58	4673	133
59	4710	135
60	4748	137
61	4787	140
62	4828	143
63	4871	146
64	4916	150

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
65	4963	154
66	5013	159
67	5066	164
68	5124	170
69	5186	178
70	5254	187
71	5329	198
72	5415	211
73	5514	229
74	5633	254
75	5783	290
76	5991	352
77	6339	493
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 2024 STAAR
Mean P-Values, Raw Score Summary, and Reliability

Table B.4.1. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	373,236	19.48	7.77	0.88	2.72	50.69
	Numerical Representations and Relationships	10	373,236	5.51	2.51	0.67	1.44	51.73
	Computations and Algebraic Relationships	13	373,236	6.87	3.21	0.77	1.55	52.19
	Geometry and Measurement	8	373,236	4.04	1.84	0.54	1.25	48.79
	Data Analysis and Personal Financial Literacy	6	373,236	3.06	1.55	0.43	1.17	48.58
RLA	OVERALL TEST	52	359,806	26.88	11.16	0.93	3.03	56.58
	Reading	26	359,806	14.39	5.54	0.84	2.22	55.15
	Writing	26	359,806	12.50	6.24	0.89	2.04	58.59

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.4.2. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	183,874	18.87	7.53	0.87	2.74	48.96
	Numerical Representations and Relationships	10	183,874	5.34	2.45	0.65	1.46	50.12
	Computations and Algebraic Relationships	13	183,874	6.56	3.15	0.75	1.57	49.73
	Geometry and Measurement	8	183,874	3.99	1.80	0.52	1.25	48.02
	Data Analysis and Personal Financial Literacy	6	183,874	2.99	1.53	0.41	1.17	46.97
RLA	OVERALL TEST	52	176,934	27.79	11.09	0.93	3.02	58.00
	Reading	26	176,934	14.72	5.46	0.84	2.21	56.35
	Writing	26	176,934	13.08	6.25	0.89	2.03	60.34

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.4.3. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	189,297	20.06	7.95	0.89	2.69	52.37
	Numerical Representations and Relationships	10	189,297	5.67	2.55	0.69	1.42	53.29
	Computations and Algebraic Relationships	13	189,297	7.17	3.25	0.78	1.53	54.59
	Geometry and Measurement	8	189,297	4.09	1.88	0.56	1.24	49.53
	Data Analysis and Personal Financial Literacy	6	189,297	3.13	1.57	0.45	1.16	50.16
RLA	OVERALL TEST	52	182,816	26.00	11.15	0.93	3.03	55.20
	Reading	26	182,816	14.06	5.59	0.84	2.23	53.99
	Writing	26	182,816	11.94	6.18	0.89	2.04	56.90

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 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	48,897	16.69	7.19	0.85	2.77	43.05
	Numerical Representations and Relationships	10	48,897	4.79	2.43	0.63	1.49	45.07
	Computations and Algebraic Relationships	13	48,897	5.78	2.96	0.71	1.58	43.39
	Geometry and Measurement	8	48,897	3.51	1.75	0.49	1.26	41.87
	Data Analysis and Personal Financial Literacy	6	48,897	2.61	1.51	0.36	1.20	41.11
RLA	OVERALL TEST	52	48,928	23.30	10.85	0.92	3.06	49.90
	Reading	26	48,928	12.75	5.45	0.83	2.25	48.95
	Writing	26	48,928	10.55	6.03	0.88	2.05	51.25

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.

Table B.4.5. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	188,007	18.30	7.34	0.86	2.76	47.50
	Numerical Representations and Relationships	10	188,007	5.16	2.40	0.63	1.47	48.32
	Computations and Algebraic Relationships	13	188,007	6.42	3.08	0.74	1.57	48.70
	Geometry and Measurement	8	188,007	3.84	1.81	0.51	1.26	46.08
	Data Analysis and Personal Financial Literacy	6	188,007	2.88	1.49	0.37	1.19	45.72
RLA	OVERALL TEST	52	174,511	25.35	10.75	0.92	3.06	53.46
	Reading	26	174,511	13.62	5.34	0.82	2.25	52.14
	Writing	26	174,511	11.73	6.06	0.88	2.06	55.33

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.4.6. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	97,629	21.72	7.64	0.88	2.65	56.79
	Numerical Representations and Relationships	10	97,629	6.14	2.47	0.68	1.39	57.80
	Computations and Algebraic Relationships	13	97,629	7.74	3.16	0.77	1.51	58.99
	Geometry and Measurement	8	97,629	4.42	1.79	0.53	1.23	53.89
	Data Analysis and Personal Financial Literacy	6	97,629	3.43	1.54	0.46	1.14	54.60
RLA	OVERALL TEST	52	97,644	29.69	10.75	0.92	2.97	62.31
	Reading	26	97,644	15.77	5.38	0.84	2.17	60.51
	Writing	26	97,644	13.92	6.01	0.89	2.01	64.85

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.4.7. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	303,553	19.03	7.78	0.88	2.72	49.49
	Numerical Representations and Relationships	10	303,553	5.38	2.52	0.67	1.45	50.53
	Computations and Algebraic Relationships	13	303,553	6.70	3.21	0.77	1.55	50.86
	Geometry and Measurement	8	303,553	3.95	1.85	0.54	1.25	47.65
	Data Analysis and Personal Financial Literacy	6	303,553	3.00	1.55	0.43	1.17	47.59
RLA	OVERALL TEST	52	119,431	19.23	8.89	0.88	3.10	41.92
	Reading	26	119,431	10.89	4.56	0.75	2.30	41.92
	Writing	26	119,431	8.33	5.07	0.84	2.06	41.93

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.4.8. Spring 2024 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	69,683	21.42	7.42	0.87	2.68	55.91
	Numerical Representations and Relationships	10	69,683	6.05	2.40	0.66	1.41	56.95
	Computations and Algebraic Relationships	13	69,683	7.60	3.11	0.76	1.53	58.01
	Geometry and Measurement	8	69,683	4.43	1.77	0.52	1.23	53.71
	Data Analysis and Personal Financial Literacy	6	69,683	3.33	1.51	0.43	1.15	52.91
RLA	OVERALL TEST	52	240,375	30.69	10.18	0.92	2.97	63.86
	Reading	26	240,375	16.12	5.15	0.82	2.17	61.72
	Writing	26	240,375	14.57	5.71	0.88	2.00	66.87

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.4.9. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	377,744	21.35	9.24	0.91	2.83	51.88
	Numerical Representations and Relationships	11	377,744	6.66	2.88	0.75	1.43	58.21
	Computations and Algebraic Relationships	12	377,744	5.59	3.38	0.78	1.57	45.87
	Geometry and Measurement	12	377,744	5.82	2.91	0.72	1.55	46.06
	Data Analysis and Personal Financial Literacy	5	377,744	3.28	1.38	0.52	0.96	65.77
RLA	OVERALL TEST	52	368,529	27.48	11.27	0.92	3.12	56.13
	Reading	26	368,529	14.74	5.26	0.83	2.19	57.29
	Writing	26	368,529	12.74	6.66	0.89	2.21	54.49

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.4.10. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	185,297	20.76	9.02	0.90	2.85	50.39
	Numerical Representations and Relationships	11	185,297	6.44	2.86	0.74	1.46	56.32
	Computations and Algebraic Relationships	12	185,297	5.46	3.32	0.78	1.57	44.85
	Geometry and Measurement	12	185,297	5.59	2.82	0.70	1.56	44.00
	Data Analysis and Personal Financial Literacy	5	185,297	3.28	1.37	0.51	0.96	65.27
RLA	OVERALL TEST	52	180,353	28.31	11.18	0.92	3.13	57.24
	Reading	26	180,353	14.89	5.16	0.82	2.20	57.77
	Writing	26	180,353	13.42	6.66	0.89	2.21	56.50

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.4.11. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	192,395	21.91	9.41	0.91	2.81	53.32
	Numerical Representations and Relationships	11	192,395	6.87	2.89	0.76	1.41	60.03
	Computations and Algebraic Relationships	12	192,395	5.71	3.43	0.79	1.56	46.87
	Geometry and Measurement	12	192,395	6.05	2.98	0.73	1.54	48.04
	Data Analysis and Personal Financial Literacy	5	192,395	3.28	1.40	0.54	0.95	66.25
RLA	OVERALL TEST	52	188,130	26.69	11.29	0.92	3.12	55.07
	Reading	26	188,130	14.60	5.35	0.83	2.19	56.83
	Writing	26	188,130	12.09	6.59	0.89	2.20	52.57

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.4.12. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	48,389	17.79	8.46	0.89	2.85	42.74
	Numerical Representations and Relationships	11	48,389	5.60	2.85	0.72	1.50	48.92
	Computations and Algebraic Relationships	12	48,389	4.45	3.12	0.75	1.55	36.15
	Geometry and Measurement	12	48,389	4.89	2.53	0.63	1.53	37.61
	Data Analysis and Personal Financial Literacy	5	48,389	2.84	1.40	0.50	0.99	56.89
RLA	OVERALL TEST	52	48,429	24.02	10.82	0.92	3.15	50.06
	Reading	26	48,429	13.26	5.13	0.81	2.23	51.60
	Writing	26	48,429	10.75	6.36	0.88	2.20	47.87

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.4.13. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	191,814	20.16	8.80	0.89	2.86	48.93
	Numerical Representations and Relationships	11	191,814	6.41	2.85	0.74	1.46	56.12
	Computations and Algebraic Relationships	12	191,814	5.14	3.22	0.76	1.59	42.23
	Geometry and Measurement	12	191,814	5.43	2.75	0.68	1.56	42.72
	Data Analysis and Personal Financial Literacy	5	191,814	3.18	1.37	0.50	0.97	63.53
RLA	OVERALL TEST	52	181,773	25.68	10.81	0.91	3.16	52.59
	Reading	26	181,773	13.88	5.07	0.81	2.23	53.90
	Writing	26	181,773	11.80	6.43	0.88	2.22	50.75

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 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	98,818	23.76	9.09	0.91	2.79	57.95
	Numerical Representations and Relationships	11	98,818	7.28	2.73	0.75	1.37	63.41
	Computations and Algebraic Relationships	12	98,818	6.40	3.34	0.78	1.57	52.68
	Geometry and Measurement	12	98,818	6.56	2.98	0.73	1.54	52.50
	Data Analysis and Personal Financial Literacy	5	98,818	3.52	1.33	0.50	0.94	71.12
RLA	OVERALL TEST	52	99,287	30.62	10.82	0.92	3.07	62.33
	Reading	26	99,287	16.27	5.04	0.82	2.14	63.32
	Writing	26	99,287	14.35	6.44	0.88	2.19	60.94

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.4.15. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	129,991	15.89	7.51	0.85	2.88	38.22
	Numerical Representations and Relationships	11	129,991	5.13	2.74	0.68	1.54	45.03
	Computations and Algebraic Relationships	12	129,991	3.71	2.68	0.67	1.54	30.47
	Geometry and Measurement	12	129,991	4.43	2.31	0.55	1.55	33.99
	Data Analysis and Personal Financial Literacy	5	129,991	2.61	1.34	0.44	1.00	51.76
RLA	OVERALL TEST	52	123,113	19.13	8.71	0.87	3.16	41.06
	Reading	26	123,113	11.13	4.43	0.74	2.28	43.27
	Writing	26	123,113	8.00	5.05	0.81	2.18	37.94

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.4.16. Spring 2024 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	247,753	24.21	8.76	0.90	2.79	59.05
	Numerical Representations and Relationships	11	247,753	7.46	2.62	0.73	1.37	65.13
	Computations and Algebraic Relationships	12	247,753	6.57	3.29	0.77	1.58	53.95
	Geometry and Measurement	12	247,753	6.55	2.93	0.72	1.55	52.39
	Data Analysis and Personal Financial Literacy	5	247,753	3.63	1.27	0.47	0.93	73.11
RLA	OVERALL TEST	52	245,416	31.67	10.00	0.90	3.09	63.69
	Reading	26	245,416	16.56	4.67	0.79	2.15	64.33
	Writing	26	245,416	15.12	6.07	0.87	2.20	62.79

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.4.17. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	380,589	23.25	9.25	0.90	2.88	55.31
	Numerical Representations and Relationships	7	380,589	3.71	1.87	0.62	1.15	51.63
	Computations and Algebraic Relationships	20	380,589	10.86	4.91	0.83	2.02	54.13
	Geometry and Measurement	10	380,589	5.59	2.50	0.68	1.41	56.06
	Data Analysis and Personal Financial Literacy	5	380,589	3.09	1.26	0.39	0.99	64.07
RLA	OVERALL TEST	52	375,523	28.87	11.30	0.93	3.09	62.02
	Reading	26	375,523	15.24	5.63	0.85	2.18	59.42
	Writing	26	375,523	13.64	6.31	0.88	2.18	65.67
Science	OVERALL TEST	39	380,984	19.46	7.44	0.86	2.80	50.78
	Matter and Energy	5	380,984	2.75	1.27	0.39	0.99	51.95
	Force, Motion, and Energy	9	380,984	4.04	1.91	0.51	1.34	43.31
	Earth and Space	13	380,984	6.33	2.99	0.72	1.60	51.42
	Organisms and Environments	12	380,984	6.34	2.79	0.68	1.58	55.26

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.4.18. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	187,165	22.97	8.99	0.90	2.90	54.57
	Numerical Representations and Relationships	7	187,165	3.62	1.89	0.62	1.16	50.59
	Computations and Algebraic Relationships	20	187,165	10.70	4.75	0.82	2.04	53.16
	Geometry and Measurement	10	187,165	5.57	2.43	0.66	1.41	55.73
	Data Analysis and Personal Financial Literacy	5	187,165	3.08	1.25	0.38	0.99	63.84
RLA	OVERALL TEST	52	184,339	29.64	11.22	0.92	3.09	63.16
	Reading	26	184,339	15.43	5.57	0.85	2.18	60.11
	Writing	26	184,339	14.22	6.30	0.88	2.18	67.46
Science	OVERALL TEST	39	187,177	19.00	7.16	0.85	2.80	49.48
	Matter and Energy	5	187,177	2.76	1.24	0.36	0.99	51.84
	Force, Motion, and Energy	9	187,177	3.85	1.82	0.47	1.33	41.04
	Earth and Space	13	187,177	6.17	2.93	0.70	1.60	50.16
	Organisms and Environments	12	187,177	6.23	2.75	0.67	1.58	54.20

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 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	193,375	23.52	9.49	0.91	2.86	56.04
	Numerical Representations and Relationships	7	193,375	3.79	1.86	0.62	1.14	52.64
	Computations and Algebraic Relationships	20	193,375	11.02	5.05	0.84	1.99	55.07
	Geometry and Measurement	10	193,375	5.61	2.56	0.70	1.41	56.38
	Data Analysis and Personal Financial Literacy	5	193,375	3.10	1.28	0.41	0.98	64.30
RLA	OVERALL TEST	52	191,141	28.13	11.32	0.93	3.09	60.91
	Reading	26	191,141	15.05	5.68	0.85	2.18	58.76
	Writing	26	191,141	13.08	6.27	0.88	2.17	63.95
Science	OVERALL TEST	39	193,761	19.90	7.67	0.87	2.79	52.04
	Matter and Energy	5	193,761	2.75	1.29	0.41	0.99	52.05
	Force, Motion, and Energy	9	193,761	4.23	1.97	0.53	1.35	45.51
	Earth and Space	13	193,761	6.47	3.05	0.73	1.59	52.65
	Organisms and Environments	12	193,761	6.45	2.83	0.69	1.57	56.28

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.4.20. Spring 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	48,498	19.65	8.68	0.88	2.96	46.61
	Numerical Representations and Relationships	7	48,498	3.05	1.79	0.57	1.17	42.30
	Computations and Algebraic Relationships	20	48,498	9.09	4.68	0.81	2.06	45.15
	Geometry and Measurement	10	48,498	4.78	2.40	0.63	1.45	48.00
	Data Analysis and Personal Financial Literacy	5	48,498	2.73	1.25	0.35	1.01	56.17
RLA	OVERALL TEST	52	48,671	25.42	10.87	0.92	3.15	55.63
	Reading	26	48,671	13.69	5.46	0.83	2.23	53.58
	Writing	26	48,671	11.73	6.06	0.87	2.20	58.52
Science	OVERALL TEST	39	48,700	16.72	6.70	0.82	2.81	43.58
	Matter and Energy	5	48,700	2.39	1.24	0.34	1.01	44.58
	Force, Motion, and Energy	9	48,700	3.43	1.70	0.39	1.33	36.26
	Earth and Space	13	48,700	5.44	2.78	0.67	1.60	44.68
	Organisms and Environments	12	48,700	5.45	2.66	0.65	1.58	47.65

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 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	195,442	21.98	8.69	0.89	2.93	52.36
	Numerical Representations and Relationships	7	195,442	3.53	1.82	0.59	1.17	49.21
	Computations and Algebraic Relationships	20	195,442	10.24	4.66	0.81	2.05	51.01
	Geometry and Measurement	10	195,442	5.25	2.37	0.64	1.42	52.79
	Data Analysis and Personal Financial Literacy	5	195,442	2.96	1.23	0.36	0.99	61.61
RLA	OVERALL TEST	52	188,919	27.15	10.84	0.92	3.14	58.70
	Reading	26	188,919	14.29	5.41	0.83	2.22	55.84
	Writing	26	188,919	12.86	6.11	0.87	2.20	62.73
Science	OVERALL TEST	39	194,316	18.10	6.89	0.83	2.81	47.36
	Matter and Energy	5	194,316	2.59	1.23	0.33	1.00	48.72
	Force, Motion, and Energy	9	194,316	3.77	1.79	0.44	1.34	40.41
	Earth and Space	13	194,316	5.84	2.85	0.69	1.60	47.80
	Organisms and Environments	12	194,316	5.90	2.66	0.65	1.59	51.56

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 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	98,344	25.75	9.06	0.90	2.82	61.22
	Numerical Representations and Relationships	7	98,344	4.08	1.83	0.62	1.12	56.68
	Computations and Algebraic Relationships	20	98,344	12.06	4.79	0.83	1.97	60.18
	Geometry and Measurement	10	98,344	6.25	2.46	0.69	1.38	62.35
	Data Analysis and Personal Financial Literacy	5	98,344	3.37	1.24	0.38	0.98	69.88
RLA	OVERALL TEST	52	99,092	31.77	10.83	0.92	3.02	67.90
	Reading	26	99,092	16.86	5.43	0.85	2.11	65.62
	Writing	26	99,092	14.91	6.06	0.87	2.15	71.12
Science	OVERALL TEST	39	99,066	22.23	7.36	0.86	2.78	57.85
	Matter and Energy	5	99,066	3.12	1.23	0.39	0.96	59.00
	Force, Motion, and Energy	9	99,066	4.60	1.97	0.53	1.34	49.54
	Earth and Space	13	99,066	7.27	2.95	0.71	1.59	58.69
	Organisms and Environments	12	99,066	7.24	2.73	0.67	1.57	62.80

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 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	126,691	17.62	7.60	0.85	2.98	42.33
	Numerical Representations and Relationships	7	126,691	2.83	1.67	0.49	1.19	39.59
	Computations and Algebraic Relationships	20	126,691	8.08	4.17	0.75	2.08	40.92
	Geometry and Measurement	10	126,691	4.20	2.16	0.54	1.46	42.48
	Data Analysis and Personal Financial Literacy	5	126,691	2.51	1.20	0.32	0.99	51.77
RLA	OVERALL TEST	52	120,377	20.25	8.73	0.87	3.19	45.59
	Reading	26	120,377	11.17	4.56	0.75	2.29	43.82
	Writing	26	120,377	9.08	4.96	0.80	2.20	48.08
Science	OVERALL TEST	39	123,770	14.97	6.07	0.79	2.79	39.38
	Matter and Energy	5	123,770	2.24	1.25	0.32	1.03	41.82
	Force, Motion, and Energy	9	123,770	3.23	1.61	0.32	1.33	34.32
	Earth and Space	13	123,770	4.64	2.45	0.60	1.56	39.23
	Organisms and Environments	12	123,770	4.86	2.51	0.60	1.59	42.29

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 2024 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	253,898	26.06	8.70	0.89	2.82	61.79
	Numerical Representations and Relationships	7	253,898	4.14	1.82	0.63	1.11	57.63
	Computations and Algebraic Relationships	20	253,898	12.25	4.65	0.82	1.98	60.73
	Geometry and Measurement	10	253,898	6.29	2.36	0.66	1.38	62.83
	Data Analysis and Personal Financial Literacy	5	253,898	3.37	1.20	0.33	0.98	70.21
RLA	OVERALL TEST	52	255,146	32.94	10.01	0.91	3.02	69.76
	Reading	26	255,146	17.16	5.04	0.82	2.11	66.78
	Writing	26	255,146	15.79	5.71	0.86	2.14	73.97
Science	OVERALL TEST	39	257,214	21.62	7.06	0.84	2.79	56.26
	Matter and Energy	5	257,214	3.00	1.20	0.35	0.96	56.82
	Force, Motion, and Energy	9	257,214	4.43	1.92	0.51	1.34	47.64
	Earth and Space	13	257,214	7.14	2.89	0.69	1.60	57.29
	Organisms and Environments	12	257,214	7.06	2.64	0.65	1.57	61.49

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.4.25. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	387,455	21.03	9.30	0.90	2.96	49.18
	Numerical Representations and Relationships	11	387,455	4.94	2.82	0.70	1.55	42.40
	Computations and Algebraic Relationships	15	387,455	7.95	3.86	0.80	1.72	55.34
	Geometry and Measurement	7	387,455	2.97	1.73	0.51	1.22	39.57
	Data Analysis and Personal Financial Literacy	10	387,455	5.17	2.29	0.62	1.42	54.01
RLA	OVERALL TEST	56	394,029	29.41	12.01	0.93	3.26	55.26
	Reading	28	394,029	15.37	5.64	0.83	2.30	54.45
	Writing	28	394,029	14.04	7.03	0.89	2.30	56.36

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.4.26. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	190,236	20.70	9.03	0.89	2.96	48.50
	Numerical Representations and Relationships	11	190,236	4.76	2.74	0.68	1.55	40.82
	Computations and Algebraic Relationships	15	190,236	7.90	3.79	0.79	1.73	55.13
	Geometry and Measurement	7	190,236	2.88	1.69	0.49	1.21	38.09
	Data Analysis and Personal Financial Literacy	10	190,236	5.16	2.24	0.61	1.41	54.17
RLA	OVERALL TEST	56	192,961	30.78	11.87	0.92	3.26	57.07
	Reading	28	192,961	15.78	5.55	0.83	2.30	55.80
	Writing	28	192,961	15.00	6.99	0.89	2.30	58.82

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.4.27. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	197,143	21.34	9.55	0.90	2.96	49.84
	Numerical Representations and Relationships	11	197,143	5.10	2.87	0.71	1.54	43.93
	Computations and Algebraic Relationships	15	197,143	8.00	3.92	0.81	1.72	55.55
	Geometry and Measurement	7	197,143	3.05	1.77	0.52	1.22	40.99
	Data Analysis and Personal Financial Literacy	10	197,143	5.18	2.34	0.63	1.43	53.85
RLA	OVERALL TEST	56	201,003	28.09	11.99	0.93	3.25	53.52
	Reading	28	201,003	14.97	5.70	0.84	2.30	53.17
	Writing	28	201,003	13.13	6.96	0.89	2.29	54.01

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.4.28. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	48,682	18.02	8.21	0.87	2.96	42.47
	Numerical Representations and Relationships	11	48,682	4.22	2.54	0.63	1.55	35.89
	Computations and Algebraic Relationships	15	48,682	6.83	3.61	0.77	1.74	48.05
	Geometry and Measurement	7	48,682	2.38	1.55	0.37	1.22	31.71
	Data Analysis and Personal Financial Literacy	10	48,682	4.59	2.06	0.55	1.39	48.88
RLA	OVERALL TEST	56	48,921	26.44	11.28	0.92	3.28	50.55
	Reading	28	48,921	14.14	5.40	0.81	2.32	50.12
	Writing	28	48,921	12.30	6.61	0.88	2.31	51.13

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 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	204,603	19.23	8.47	0.87	3.00	45.09
	Numerical Representations and Relationships	11	204,603	4.49	2.62	0.64	1.57	38.32
	Computations and Algebraic Relationships	15	204,603	7.27	3.64	0.77	1.75	50.89
	Geometry and Measurement	7	204,603	2.71	1.61	0.42	1.23	35.88
	Data Analysis and Personal Financial Literacy	10	204,603	4.76	2.13	0.55	1.43	50.19
RLA	OVERALL TEST	56	206,822	27.02	11.45	0.92	3.29	51.16
	Reading	28	206,822	14.26	5.44	0.82	2.33	50.49
	Writing	28	206,822	12.76	6.72	0.88	2.32	52.07

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 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	97,415	24.22	9.27	0.90	2.94	56.41
	Numerical Representations and Relationships	11	97,415	5.67	2.86	0.71	1.55	49.16
	Computations and Algebraic Relationships	15	97,415	9.15	3.75	0.80	1.70	63.30
	Geometry and Measurement	7	97,415	3.50	1.75	0.54	1.19	46.71
	Data Analysis and Personal Financial Literacy	10	97,415	5.90	2.31	0.62	1.42	60.64
RLA	OVERALL TEST	56	99,453	33.44	11.39	0.92	3.21	62.26
	Reading	28	99,453	17.24	5.34	0.82	2.26	61.16
	Writing	28	99,453	16.21	6.76	0.89	2.27	63.77

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.4.31. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	116,085	15.27	6.76	0.81	2.95	36.09
	Numerical Representations and Relationships	11	116,085	3.64	2.19	0.51	1.54	30.55
	Computations and Algebraic Relationships	15	116,085	5.59	3.09	0.68	1.74	39.68
	Geometry and Measurement	7	116,085	2.17	1.42	0.26	1.22	29.32
	Data Analysis and Personal Financial Literacy	10	116,085	3.87	1.85	0.44	1.38	41.56
RLA	OVERALL TEST	56	112,757	19.75	8.90	0.86	3.30	39.29
	Reading	28	112,757	11.14	4.63	0.74	2.36	39.56
	Writing	28	112,757	8.60	5.08	0.80	2.29	38.93

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.4.32. Spring 2024 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	43	271,370	23.49	9.15	0.90	2.95	54.78
	Numerical Representations and Relationships	11	271,370	5.49	2.87	0.71	1.55	47.47
	Computations and Algebraic Relationships	15	271,370	8.96	3.71	0.79	1.71	62.04
	Geometry and Measurement	7	271,370	3.31	1.74	0.53	1.19	43.95
	Data Analysis and Personal Financial Literacy	10	271,370	5.73	2.23	0.59	1.43	59.33
RLA	OVERALL TEST	56	281,272	33.28	10.85	0.91	3.23	61.66
	Reading	28	281,272	17.06	5.10	0.80	2.27	60.42
	Writing	28	281,272	16.22	6.51	0.88	2.28	63.35

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.4.33. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	324,109	22.62	9.63	0.90	2.99	47.72
	Probablility and Numerical Representations	6	324,109	2.72	1.47	0.48	1.06	47.06
	Computations and Algebraic Relationships	20	324,109	9.52	4.67	0.81	2.03	45.69
	Geometry and Measurement	11	324,109	4.93	2.77	0.74	1.40	44.82
	Data Analysis and Personal Financial Literacy	9	324,109	5.46	2.02	0.59	1.30	57.09
RLA	OVERALL TEST	56	397,569	32.23	12.31	0.94	3.13	60.28
	Reading	28	397,569	15.90	5.67	0.84	2.26	56.30
	Writing	28	397,569	16.33	7.26	0.91	2.14	65.73

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 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	159,221	22.30	9.29	0.90	2.99	46.94
	Probablility and Numerical Representations	6	159,221	2.70	1.44	0.45	1.07	46.80
	Computations and Algebraic Relationships	20	159,221	9.30	4.50	0.79	2.04	44.39
	Geometry and Measurement	11	159,221	4.85	2.74	0.74	1.40	44.13
	Data Analysis and Personal Financial Literacy	9	159,221	5.45	1.95	0.56	1.29	56.92
RLA	OVERALL TEST	56	194,536	34.03	12.02	0.93	3.10	62.84
	Reading	28	194,536	16.43	5.61	0.84	2.25	58.07
	Writing	28	194,536	17.61	7.02	0.91	2.11	69.36

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.4.35. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	164,822	22.94	9.93	0.91	2.97	48.47
	Probablility and Numerical Representations	6	164,822	2.75	1.50	0.51	1.05	47.31
	Computations and Algebraic Relationships	20	164,822	9.72	4.81	0.82	2.02	46.94
	Geometry and Measurement	11	164,822	5.00	2.79	0.75	1.39	45.50
	Data Analysis and Personal Financial Literacy	9	164,822	5.46	2.08	0.61	1.30	57.26
RLA	OVERALL TEST	56	202,979	30.50	12.35	0.93	3.15	57.83
	Reading	28	202,979	15.39	5.67	0.84	2.26	54.60
	Writing	28	202,979	15.11	7.28	0.91	2.17	62.26

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.4.36. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	42,193	19.47	8.22	0.86	3.02	40.70
	Probablility and Numerical Representations	6	42,193	2.38	1.38	0.39	1.08	41.07
	Computations and Algebraic Relationships	20	42,193	8.12	4.06	0.75	2.05	38.71
	Geometry and Measurement	11	42,193	4.03	2.40	0.66	1.40	36.64
	Data Analysis and Personal Financial Literacy	9	42,193	4.94	1.88	0.49	1.34	51.08
RLA	OVERALL TEST	56	49,786	29.44	11.79	0.93	3.19	55.93
	Reading	28	49,786	14.74	5.48	0.82	2.29	52.23
	Writing	28	49,786	14.70	6.98	0.90	2.19	61.01

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.4.37. Spring 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	176,633	21.04	8.81	0.88	3.02	44.22
	Probablility and Numerical Representations	6	176,633	2.53	1.39	0.41	1.07	43.79
	Computations and Algebraic Relationships	20	176,633	8.80	4.32	0.77	2.06	42.03
	Geometry and Measurement	11	176,633	4.55	2.59	0.70	1.41	41.37
	Data Analysis and Personal Financial Literacy	9	176,633	5.16	1.94	0.53	1.33	53.70
RLA	OVERALL TEST	56	210,497	29.81	11.92	0.93	3.19	56.23
	Reading	28	210,497	14.83	5.49	0.83	2.29	52.53
	Writing	28	210,497	14.97	7.10	0.90	2.20	61.30

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 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	77,313	25.93	9.72	0.91	2.94	54.98
	Probablility and Numerical Representations	6	77,313	3.12	1.49	0.51	1.04	53.76
	Computations and Algebraic Relationships	20	77,313	10.97	4.73	0.82	2.02	53.05
	Geometry and Measurement	11	77,313	5.75	2.78	0.75	1.39	52.26
	Data Analysis and Personal Financial Literacy	9	77,313	6.09	1.98	0.61	1.23	64.27
RLA	OVERALL TEST	56	98,986	36.36	11.38	0.93	3.02	67.22
	Reading	28	98,986	17.69	5.31	0.83	2.20	62.64
	Writing	28	98,986	18.67	6.71	0.91	2.05	73.48

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 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	102,231	17.39	6.94	0.81	3.05	36.24
	Probablility and Numerical Representations	6	102,231	2.07	1.27	0.27	1.09	35.75
	Computations and Algebraic Relationships	20	102,231	7.20	3.45	0.65	2.05	33.92
	Geometry and Measurement	11	102,231	3.71	2.18	0.58	1.40	33.69
	Data Analysis and Personal Financial Literacy	9	102,231	4.41	1.80	0.39	1.41	45.55
RLA	OVERALL TEST	56	106,153	21.69	9.71	0.89	3.28	42.65
	Reading	28	106,153	11.52	4.71	0.75	2.33	40.87
	Writing	28	106,153	10.17	5.79	0.84	2.28	45.08

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 2024 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	46	221,878	25.03	9.74	0.91	2.95	53.01
	Probablility and Numerical Representations	6	221,878	3.02	1.46	0.49	1.04	52.26
	Computations and Algebraic Relationships	20	221,878	10.58	4.77	0.82	2.02	51.11
	Geometry and Measurement	11	221,878	5.49	2.83	0.76	1.40	49.95
	Data Analysis and Personal Financial Literacy	9	221,878	5.94	1.93	0.58	1.24	62.41
RLA	OVERALL TEST	56	291,416	36.07	10.84	0.92	3.06	66.70
	Reading	28	291,416	17.49	5.12	0.81	2.22	61.92
	Writing	28	291,416	18.57	6.40	0.89	2.08	73.25

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.

Table B.4.41. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	356,728	24.27	9.97	0.90	3.17	49.20
	Numerical Representations and Relationships	4	356,728	2.94	1.07	0.43	0.81	70.40
	Computations and Algebraic Relationships	19	356,728	9.86	4.43	0.81	1.93	51.07
	Geometry and Measurement	19	356,728	8.80	4.38	0.78	2.07	44.90
	Data Analysis and Personal Financial Literacy	6	356,728	2.66	1.57	0.54	1.07	44.36
RLA	OVERALL TEST	56	402,883	30.00	12.02	0.93	3.23	55.36
	Reading	28	402,883	15.75	5.91	0.85	2.29	56.00
	Writing	28	402,883	14.25	6.72	0.88	2.28	54.49
Science	OVERALL TEST	46	398,004	23.58	9.55	0.90	3.02	51.23
	Matter and Energy	12	398,004	7.10	2.77	0.71	1.49	56.81
	Force, Motion, and Energy	9	398,004	4.32	2.40	0.66	1.40	48.82
	Earth and Space	13	398,004	6.10	3.02	0.73	1.57	47.19
	Organisms and Environments	12	398,004	6.06	2.76	0.72	1.45	51.63
Social Studies	OVERALL TEST	49	405,802	24.04	9.87	0.90	3.16	49.13
	History	22	405,802	11.20	4.30	0.75	2.13	49.66
	Geography and Culture	10	405,802	4.96	2.51	0.66	1.46	49.16
	Government and Citizenship	11	405,802	4.60	2.76	0.72	1.45	43.41
	Economics, Science, Technology, and Society	6	405,802	3.28	1.66	0.71	0.89	57.62

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 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	173,372	24.31	9.73	0.89	3.17	49.11
	Numerical Representations and Relationships	4	173,372	2.98	1.05	0.43	0.79	71.40
	Computations and Algebraic Relationships	19	173,372	9.92	4.35	0.80	1.93	51.06
	Geometry and Measurement	19	173,372	8.82	4.32	0.77	2.07	44.94
	Data Analysis and Personal Financial Literacy	6	173,372	2.59	1.53	0.51	1.07	43.18
RLA	OVERALL TEST	56	196,266	31.65	11.70	0.92	3.22	57.50
	Reading	28	196,266	16.35	5.76	0.84	2.28	58.00
	Writing	28	196,266	15.31	6.54	0.88	2.27	56.82
Science	OVERALL TEST	46	194,152	23.11	9.39	0.90	3.04	50.17
	Matter and Energy	12	194,152	7.06	2.72	0.70	1.49	56.44
	Force, Motion, and Energy	9	194,152	4.14	2.36	0.65	1.40	46.90
	Earth and Space	13	194,152	5.95	2.98	0.72	1.59	46.05
	Organisms and Environments	12	194,152	5.96	2.75	0.72	1.46	50.63
Social Studies	OVERALL TEST	49	198,133	23.70	9.63	0.89	3.17	48.39
	History	22	198,133	11.03	4.16	0.73	2.14	48.83
	Geography and Culture	10	198,133	4.88	2.48	0.65	1.47	48.15
	Government and Citizenship	11	198,133	4.60	2.73	0.72	1.45	43.54
	Economics, Science, Technology, and Society	6	198,133	3.19	1.64	0.69	0.90	56.10

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.4.43. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	183,301	24.23	10.19	0.90	3.16	49.29
	Numerical Representations and Relationships	4	183,301	2.90	1.09	0.43	0.82	69.46
	Computations and Algebraic Relationships	19	183,301	9.81	4.50	0.81	1.93	51.09
	Geometry and Measurement	19	183,301	8.79	4.44	0.78	2.06	44.87
	Data Analysis and Personal Financial Literacy	6	183,301	2.73	1.61	0.56	1.07	45.47
RLA	OVERALL TEST	56	206,566	28.43	12.11	0.93	3.24	53.33
	Reading	28	206,566	15.18	5.99	0.85	2.29	54.10
	Writing	28	206,566	13.25	6.72	0.88	2.28	52.28
Science	OVERALL TEST	46	203,800	24.03	9.67	0.90	3.01	52.25
	Matter and Energy	12	203,800	7.14	2.82	0.72	1.48	57.16
	Force, Motion, and Energy	9	203,800	4.49	2.42	0.67	1.39	50.65
	Earth and Space	13	203,800	6.25	3.06	0.74	1.56	48.27
	Organisms and Environments	12	203,800	6.16	2.76	0.73	1.44	52.59
Social Studies	OVERALL TEST	49	207,615	24.37	10.09	0.90	3.15	49.84
	History	22	207,615	11.36	4.42	0.77	2.13	50.45
	Geography and Culture	10	207,615	5.03	2.53	0.67	1.45	50.12
	Government and Citizenship	11	207,615	4.61	2.79	0.73	1.45	43.29
	Economics, Science, Technology, and Society	6	207,615	3.37	1.68	0.73	0.87	59.08

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 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	47,116	21.11	8.51	0.86	3.21	42.51
	Numerical Representations and Relationships	4	47,116	2.75	1.11	0.41	0.85	65.36
	Computations and Algebraic Relationships	19	47,116	8.57	3.96	0.75	1.97	44.27
	Geometry and Measurement	19	47,116	7.58	3.73	0.69	2.08	38.35
	Data Analysis and Personal Financial Literacy	6	47,116	2.21	1.41	0.42	1.08	36.79
RLA	OVERALL TEST	56	50,267	27.10	11.23	0.91	3.28	50.78
	Reading	28	50,267	14.48	5.60	0.83	2.33	51.47
	Writing	28	50,267	12.62	6.29	0.87	2.30	49.82
Science	OVERALL TEST	46	50,164	20.52	8.49	0.87	3.06	44.64
	Matter and Energy	12	50,164	6.44	2.57	0.64	1.53	51.18
	Force, Motion, and Energy	9	50,164	3.62	2.23	0.61	1.40	41.12
	Earth and Space	13	50,164	5.17	2.71	0.65	1.60	40.17
	Organisms and Environments	12	50,164	5.29	2.56	0.66	1.48	45.38
Social Studies	OVERALL TEST	49	50,532	21.39	8.97	0.87	3.17	43.62
	History	22	50,532	10.15	3.87	0.69	2.14	44.65
	Geography and Culture	10	50,532	4.32	2.35	0.60	1.48	42.77
	Government and Citizenship	11	50,532	4.08	2.60	0.70	1.43	38.86
	Economics, Science, Technology, and Society	6	50,532	2.83	1.61	0.68	0.91	50.17

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.4.45. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	189,127	22.69	9.11	0.88	3.21	45.92
	Numerical Representations and Relationships	4	189,127	2.85	1.10	0.42	0.84	68.06
	Computations and Algebraic Relationships	19	189,127	9.32	4.18	0.78	1.97	48.16
	Geometry and Measurement	19	189,127	8.08	3.97	0.72	2.09	41.19
	Data Analysis and Personal Financial Literacy	6	189,127	2.44	1.48	0.47	1.08	40.68
RLA	OVERALL TEST	56	212,652	27.55	11.37	0.92	3.28	51.16
	Reading	28	212,652	14.57	5.64	0.83	2.33	51.75
	Writing	28	212,652	12.98	6.38	0.87	2.31	50.36
Science	OVERALL TEST	46	209,347	21.52	8.85	0.88	3.06	46.80
	Matter and Energy	12	209,347	6.57	2.64	0.66	1.53	52.45
	Force, Motion, and Energy	9	209,347	3.88	2.29	0.63	1.40	43.93
	Earth and Space	13	209,347	5.53	2.85	0.69	1.59	43.04
	Organisms and Environments	12	209,347	5.54	2.59	0.68	1.48	47.21
Social Studies	OVERALL TEST	49	214,653	21.94	9.04	0.88	3.18	44.83
	History	22	214,653	10.37	3.92	0.70	2.15	45.73
	Geography and Culture	10	214,653	4.49	2.38	0.61	1.48	44.47
	Government and Citizenship	11	214,653	4.08	2.58	0.69	1.44	38.84
	Economics, Science, Technology, and Society	6	214,653	3.00	1.62	0.69	0.91	53.17

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.4.46. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	88,300	27.20	10.14	0.91	3.12	55.34
	Numerical Representations and Relationships	4	88,300	3.15	0.97	0.40	0.75	75.67
	Computations and Algebraic Relationships	19	88,300	10.89	4.44	0.82	1.90	56.64
	Geometry and Measurement	19	88,300	10.05	4.54	0.80	2.05	51.32
	Data Analysis and Personal Financial Literacy	6	88,300	3.11	1.59	0.55	1.07	51.76
RLA	OVERALL TEST	56	101,637	34.12	11.45	0.92	3.16	62.41
	Reading	28	101,637	17.74	5.64	0.84	2.22	63.19
	Writing	28	101,637	16.38	6.45	0.88	2.24	61.33
Science	OVERALL TEST	46	100,785	27.32	9.09	0.89	2.98	59.28
	Matter and Energy	12	100,785	8.02	2.67	0.71	1.43	64.36
	Force, Motion, and Energy	9	100,785	5.12	2.28	0.63	1.39	57.85
	Earth and Space	13	100,785	7.18	2.92	0.72	1.54	55.07
	Organisms and Environments	12	100,785	7.00	2.66	0.72	1.41	59.56
Social Studies	OVERALL TEST	49	102,126	27.70	9.76	0.90	3.14	56.71
	History	22	102,126	12.61	4.39	0.77	2.12	56.38
	Geography and Culture	10	102,126	5.84	2.47	0.67	1.43	58.03
	Government and Citizenship	11	102,126	5.42	2.72	0.70	1.48	50.61
	Economics, Science, Technology, and Society	6	102,126	3.83	1.54	0.69	0.86	66.41

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.4.47. Spring 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	99,283	18.78	7.59	0.82	3.22	37.95
	Numerical Representations and Relationships	4	99,283	2.47	1.18	0.40	0.91	59.28
	Computations and Algebraic Relationships	19	99,283	7.70	3.60	0.70	1.98	39.76
	Geometry and Measurement	19	99,283	6.64	3.31	0.61	2.08	33.83
	Data Analysis and Personal Financial Literacy	6	99,283	1.97	1.33	0.36	1.07	32.76
RLA	OVERALL TEST	56	100,202	19.90	8.92	0.86	3.31	38.75
	Reading	28	100,202	11.20	4.74	0.75	2.35	39.78
	Writing	28	100,202	8.69	4.94	0.78	2.32	37.33
Science	OVERALL TEST	46	98,494	17.04	7.20	0.82	3.04	37.39
	Matter and Energy	12	98,494	5.43	2.31	0.53	1.58	43.30
	Force, Motion, and Energy	9	98,494	2.87	1.97	0.53	1.35	32.77
	Earth and Space	13	98,494	4.35	2.46	0.60	1.56	34.74
	Organisms and Environments	12	98,494	4.39	2.22	0.55	1.48	37.82
Social Studies	OVERALL TEST	49	98,952	17.51	7.29	0.82	3.12	35.65
	History	22	98,952	8.78	3.24	0.57	2.13	37.99
	Geography and Culture	10	98,952	3.48	2.09	0.50	1.47	34.78
	Government and Citizenship	11	98,952	2.97	2.13	0.60	1.35	29.19
	Economics, Science, Technology, and Society	6	98,952	2.27	1.50	0.61	0.93	40.90

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 2024 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	48	257,445	26.39	9.97	0.90	3.14	53.54
	Numerical Representations and Relationships	4	257,445	3.12	0.97	0.39	0.76	74.69
	Computations and Algebraic Relationships	19	257,445	10.69	4.43	0.81	1.91	55.43
	Geometry and Measurement	19	257,445	9.64	4.45	0.79	2.06	49.17
	Data Analysis and Personal Financial Literacy	6	257,445	2.93	1.58	0.54	1.07	48.83
RLA	OVERALL TEST	56	302,681	33.34	11.00	0.92	3.20	60.86
	Reading	28	302,681	17.25	5.47	0.83	2.26	61.37
	Writing	28	302,681	16.09	6.19	0.87	2.25	60.16
Science	OVERALL TEST	46	299,510	25.73	9.24	0.89	3.01	55.78
	Matter and Energy	12	299,510	7.65	2.69	0.71	1.46	61.25
	Force, Motion, and Energy	9	299,510	4.79	2.34	0.64	1.41	54.10
	Earth and Space	13	299,510	6.68	2.97	0.72	1.57	51.28
	Organisms and Environments	12	299,510	6.61	2.69	0.72	1.44	56.17
Social Studies	OVERALL TEST	49	306,850	26.15	9.67	0.89	3.16	53.48
	History	22	306,850	11.98	4.31	0.76	2.13	53.42
	Geography and Culture	10	306,850	5.43	2.45	0.65	1.45	53.79
	Government and Citizenship	11	306,850	5.13	2.73	0.71	1.48	47.99
	Economics, Science, Technology, and Society	6	306,850	3.61	1.57	0.69	0.88	63.01

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.4.49. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	20,605	15.38	6.74	0.83	2.77	40.28
	Numerical Representations and Relationships	10	20,605	4.01	2.34	0.61	1.46	38.29
	Computations and Algebraic Relationships	13	20,605	5.53	2.86	0.69	1.59	42.04
	Geometry and Measurement	8	20,605	3.43	1.70	0.44	1.27	40.80
	Data Analysis and Personal Financial Literacy	6	20,605	2.40	1.41	0.30	1.18	38.44
RLA	OVERALL TEST	52	34,257	22.06	10.21	0.91	3.12	46.56
	Reading	26	34,257	12.22	5.12	0.80	2.28	47.53
	Writing	26	34,257	9.84	5.84	0.87	2.11	45.19

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 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	10,492	14.93	6.42	0.81	2.78	39.02
	Numerical Representations and Relationships	10	10,492	3.90	2.25	0.58	1.46	37.25
	Computations and Algebraic Relationships	13	10,492	5.25	2.74	0.66	1.60	39.84
	Geometry and Measurement	8	10,492	3.43	1.67	0.42	1.27	40.70
	Data Analysis and Personal Financial Literacy	6	10,492	2.36	1.38	0.26	1.19	37.34
RLA	OVERALL TEST	52	17,505	22.86	10.22	0.91	3.12	47.97
	Reading	26	17,505	12.53	5.02	0.79	2.27	48.82
	Writing	26	17,505	10.32	5.96	0.87	2.13	46.78

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.4.51. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	10,083	15.86	7.04	0.85	2.76	41.61
	Numerical Representations and Relationships	10	10,083	4.14	2.43	0.64	1.46	39.41
	Computations and Algebraic Relationships	13	10,083	5.83	2.95	0.71	1.59	44.36
	Geometry and Measurement	8	10,083	3.43	1.73	0.47	1.26	40.92
	Data Analysis and Personal Financial Literacy	6	10,083	2.45	1.43	0.33	1.18	39.63
RLA	OVERALL TEST	52	16,727	21.24	10.14	0.91	3.10	45.10
	Reading	26	16,727	11.91	5.21	0.81	2.29	46.20
	Writing	26	16,727	9.34	5.68	0.87	2.08	43.54

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 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	37	34	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	10	34	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	13	34	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	8	34	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	6	34	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	29	N/A	N/A	N/A	N/A	N/A
	Reading	26	29	N/A	N/A	N/A	N/A	N/A
	Writing	26	29	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 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	19,847	15.41	6.72	0.83	2.77	40.35
	Numerical Representations and Relationships	10	19,847	4.02	2.34	0.61	1.46	38.39
	Computations and Algebraic Relationships	13	19,847	5.54	2.85	0.69	1.59	42.11
	Geometry and Measurement	8	19,847	3.43	1.70	0.44	1.27	40.84
	Data Analysis and Personal Financial Literacy	6	19,847	2.41	1.41	0.29	1.18	38.51
RLA	OVERALL TEST	52	33,353	22.12	10.21	0.91	3.12	46.66
	Reading	26	33,353	12.25	5.12	0.80	2.28	47.64
	Writing	26	33,353	9.87	5.84	0.87	2.11	45.29

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 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	286	17.35	7.92	0.88	2.71	45.46
	Numerical Representations and Relationships	10	286	4.51	2.69	0.73	1.40	42.43
	Computations and Algebraic Relationships	13	286	6.27	3.20	0.76	1.57	48.12
	Geometry and Measurement	8	286	3.93	1.86	0.55	1.25	47.20
	Data Analysis and Personal Financial Literacy	6	286	2.64	1.56	0.45	1.15	41.40
RLA	OVERALL TEST	52	408	22.00	10.60	0.91	3.10	45.85
	Reading	26	408	11.93	5.23	0.81	2.26	46.42
	Writing	26	408	10.06	6.07	0.88	2.12	45.05

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.4.55. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	16,480	15.12	6.69	0.83	2.77	39.60
	Numerical Representations and Relationships	10	16,480	3.94	2.32	0.60	1.46	37.67
	Computations and Algebraic Relationships	13	16,480	5.43	2.84	0.69	1.59	41.24
	Geometry and Measurement	8	16,480	3.37	1.69	0.44	1.27	40.10
	Data Analysis and Personal Financial Literacy	6	16,480	2.37	1.41	0.30	1.18	37.98
RLA	OVERALL TEST	52	8,833	16.64	7.71	0.84	3.05	36.39
	Reading	26	8,833	9.68	4.09	0.69	2.29	37.58
	Writing	26	8,833	6.97	4.46	0.80	2.01	34.72

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.4.56. Spring 2024 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	37	4,125	16.42	6.85	0.84	2.78	42.98
	Numerical Representations and Relationships	10	4,125	4.30	2.43	0.64	1.47	40.77
	Computations and Algebraic Relationships	13	4,125	5.94	2.90	0.70	1.59	45.24
	Geometry and Measurement	8	4,125	3.66	1.72	0.46	1.26	43.59
	Data Analysis and Personal Financial Literacy	6	4,125	2.53	1.38	0.26	1.19	40.28
RLA	OVERALL TEST	52	25,424	23.94	10.31	0.91	3.13	50.09
	Reading	26	25,424	13.11	5.15	0.80	2.28	50.99
	Writing	26	25,424	10.83	5.93	0.87	2.14	48.83

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 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	14,990	15.43	7.73	0.86	2.86	37.12
	Numerical Representations and Relationships	11	14,990	4.73	2.85	0.72	1.51	42.28
	Computations and Algebraic Relationships	12	14,990	3.71	2.75	0.70	1.52	30.19
	Geometry and Measurement	12	14,990	4.37	2.27	0.52	1.57	33.28
	Data Analysis and Personal Financial Literacy	5	14,990	2.61	1.39	0.50	0.98	51.47
RLA	OVERALL TEST	52	25,322	23.84	9.65	0.90	3.08	51.08
	Reading	26	25,322	14.11	5.20	0.81	2.26	53.88
	Writing	26	25,322	9.72	5.16	0.84	2.08	47.14

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.4.58. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	7,590	15.00	7.46	0.85	2.85	36.04
	Numerical Representations and Relationships	11	7,590	4.48	2.76	0.70	1.51	40.13
	Computations and Algebraic Relationships	12	7,590	3.66	2.69	0.68	1.51	29.81
	Geometry and Measurement	12	7,590	4.26	2.19	0.50	1.55	32.33
	Data Analysis and Personal Financial Literacy	5	7,590	2.60	1.37	0.48	0.98	50.72
RLA	OVERALL TEST	52	12,904	24.67	9.64	0.90	3.09	52.53
	Reading	26	12,904	14.41	5.11	0.81	2.25	55.02
	Writing	26	12,904	10.26	5.23	0.84	2.10	49.02

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.4.59. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	7,372	15.89	7.98	0.87	2.87	38.28
	Numerical Representations and Relationships	11	7,372	5.01	2.92	0.73	1.51	44.54
	Computations and Algebraic Relationships	12	7,372	3.77	2.81	0.71	1.52	30.63
	Geometry and Measurement	12	7,372	4.49	2.35	0.55	1.58	34.28
	Data Analysis and Personal Financial Literacy	5	7,372	2.63	1.41	0.51	0.98	52.29
RLA	OVERALL TEST	52	12,397	22.98	9.60	0.90	3.07	49.60
	Reading	26	12,397	13.81	5.26	0.82	2.26	52.72
	Writing	26	12,397	9.17	5.03	0.83	2.06	45.20

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 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	40	15	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	11	15	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	12	15	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	12	15	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	5	15	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	18	N/A	N/A	N/A	N/A	N/A
	Reading	26	18	N/A	N/A	N/A	N/A	N/A
	Writing	26	18	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 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	14,423	15.48	7.73	0.86	2.86	37.23
	Numerical Representations and Relationships	11	14,423	4.75	2.85	0.72	1.51	42.42
	Computations and Algebraic Relationships	12	14,423	3.72	2.75	0.69	1.52	30.29
	Geometry and Measurement	12	14,423	4.38	2.27	0.52	1.57	33.34
	Data Analysis and Personal Financial Literacy	5	14,423	2.62	1.39	0.50	0.98	51.66
RLA	OVERALL TEST	52	24,669	23.91	9.65	0.90	3.08	51.22
	Reading	26	24,669	14.15	5.19	0.81	2.26	54.02
	Writing	26	24,669	9.76	5.16	0.84	2.08	47.27

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.4.62. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	176	16.57	8.82	0.90	2.84	39.91
	Numerical Representations and Relationships	11	176	4.85	3.01	0.75	1.51	43.56
	Computations and Algebraic Relationships	12	176	4.11	3.22	0.79	1.49	33.30
	Geometry and Measurement	12	176	4.94	2.35	0.53	1.61	37.69
	Data Analysis and Personal Financial Literacy	5	176	2.67	1.38	0.44	1.03	53.27
RLA	OVERALL TEST	52	237	23.09	9.95	0.90	3.08	49.63
	Reading	26	237	13.68	5.43	0.82	2.27	52.36
	Writing	26	237	9.41	5.19	0.84	2.06	45.77

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.4.63. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	3,665	12.55	5.83	0.77	2.79	30.06
	Numerical Representations and Relationships	11	3,665	3.77	2.39	0.60	1.52	34.11
	Computations and Algebraic Relationships	12	3,665	2.83	2.13	0.54	1.44	23.17
	Geometry and Measurement	12	3,665	3.74	1.88	0.34	1.52	27.90
	Data Analysis and Personal Financial Literacy	5	3,665	2.20	1.25	0.36	1.00	43.02
RLA	OVERALL TEST	52	6,705	18.75	7.50	0.83	3.08	41.04
	Reading	26	6,705	11.55	4.35	0.71	2.33	44.00
	Writing	26	6,705	7.20	3.98	0.75	2.00	36.85

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.4.64. Spring 2024 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	40	11,325	16.36	8.03	0.87	2.88	39.41
	Numerical Representations and Relationships	11	11,325	5.05	2.92	0.74	1.50	44.93
	Computations and Algebraic Relationships	12	11,325	3.99	2.86	0.71	1.54	32.47
	Geometry and Measurement	12	11,325	4.57	2.34	0.55	1.58	35.02
	Data Analysis and Personal Financial Literacy	5	11,325	2.75	1.40	0.51	0.98	54.21
RLA	OVERALL TEST	52	18,617	25.67	9.69	0.90	3.08	54.70
	Reading	26	18,617	15.04	5.17	0.81	2.23	57.44
	Writing	26	18,617	10.64	5.23	0.84	2.11	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.4.65. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	11,552	16.51	7.57	0.85	2.96	39.59
	Numerical Representations and Relationships	7	11,552	2.68	1.67	0.50	1.19	37.70
	Computations and Algebraic Relationships	20	11,552	7.59	4.21	0.76	2.07	38.15
	Geometry and Measurement	10	11,552	3.89	2.12	0.53	1.45	39.58
	Data Analysis and Personal Financial Literacy	5	11,552	2.36	1.16	0.29	0.97	48.22
RLA	OVERALL TEST	52	18,510	25.44	10.09	0.91	3.08	53.08
	Reading	26	18,510	14.38	4.98	0.80	2.21	56.46
	Writing	26	18,510	11.06	5.82	0.86	2.14	48.31
Science	OVERALL TEST	39	13,036	13.62	5.41	0.75	2.72	35.99
	Matter and Energy	5	13,036	2.21	1.25	0.32	1.03	41.39
	Force, Motion, and Energy	9	13,036	3.00	1.49	0.24	1.30	31.75
	Earth and Space	13	13,036	3.92	2.22	0.56	1.47	32.98
	Organisms and Environments	12	13,036	4.49	2.29	0.54	1.55	39.58

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.4.66. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	5,711	16.35	7.35	0.84	2.96	39.10
	Numerical Representations and Relationships	7	5,711	2.56	1.66	0.49	1.18	36.34
	Computations and Algebraic Relationships	20	5,711	7.55	4.11	0.74	2.08	37.72
	Geometry and Measurement	10	5,711	3.88	2.07	0.50	1.45	39.42
	Data Analysis and Personal Financial Literacy	5	5,711	2.36	1.14	0.28	0.97	48.15
RLA	OVERALL TEST	52	9,354	26.86	10.05	0.91	3.09	55.34
	Reading	26	9,354	14.93	4.89	0.80	2.20	58.51
	Writing	26	9,354	11.93	5.88	0.86	2.16	50.86
Science	OVERALL TEST	39	6,478	13.33	5.17	0.73	2.70	35.15
	Matter and Energy	5	6,478	2.20	1.22	0.29	1.03	41.20
	Force, Motion, and Energy	9	6,478	2.89	1.43	0.20	1.28	30.38
	Earth and Space	13	6,478	3.84	2.19	0.56	1.46	32.28
	Organisms and Environments	12	6,478	4.40	2.22	0.51	1.55	38.75

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 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	5,820	16.69	7.78	0.86	2.95	40.10
	Numerical Representations and Relationships	7	5,820	2.79	1.68	0.50	1.19	39.06
	Computations and Algebraic Relationships	20	5,820	7.63	4.32	0.77	2.06	38.59
	Geometry and Measurement	10	5,820	3.90	2.18	0.56	1.45	39.77
	Data Analysis and Personal Financial Literacy	5	5,820	2.37	1.17	0.31	0.97	48.33
RLA	OVERALL TEST	52	9,139	23.99	9.92	0.90	3.08	50.78
	Reading	26	9,139	13.82	5.02	0.80	2.23	54.37
	Writing	26	9,139	10.18	5.62	0.86	2.11	45.71
Science	OVERALL TEST	39	6,538	13.90	5.62	0.76	2.73	36.84
	Matter and Energy	5	6,538	2.21	1.27	0.35	1.03	41.59
	Force, Motion, and Energy	9	6,538	3.12	1.54	0.27	1.32	33.11
	Earth and Space	13	6,538	3.99	2.25	0.57	1.48	33.67
	Organisms and Environments	12	6,538	4.58	2.36	0.56	1.56	40.41

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.4.68. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N[2]	Mean	SD	Alpha	SEM	Mean P-Value
Mathematics	OVERALL TEST	42	14	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	7	14	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	20	14	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	10	14	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	5	14	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	17	N/A	N/A	N/A	N/A	N/A
	Reading	26	17	N/A	N/A	N/A	N/A	N/A
	Writing	26	17	N/A	N/A	N/A	N/A	N/A
Science	OVERALL TEST	39	12	N/A	N/A	N/A	N/A	N/A
	Matter and Energy	5	12	N/A	N/A	N/A	N/A	N/A
	Force, Motion, and Energy	9	12	N/A	N/A	N/A	N/A	N/A
	Earth and Space	13	12	N/A	N/A	N/A	N/A	N/A
	Organisms and Environments	12	12	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 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	11,075	16.63	7.57	0.85	2.96	39.87
	Numerical Representations and Relationships	7	11,075	2.70	1.67	0.50	1.19	37.98
	Computations and Algebraic Relationships	20	11,075	7.65	4.21	0.76	2.07	38.44
	Geometry and Measurement	10	11,075	3.91	2.12	0.53	1.45	39.84
	Data Analysis and Personal Financial Literacy	5	11,075	2.37	1.16	0.29	0.97	48.45
RLA	OVERALL TEST	52	17,969	25.56	10.07	0.91	3.08	53.30
	Reading	26	17,969	14.43	4.98	0.80	2.21	56.67
	Writing	26	17,969	11.13	5.82	0.86	2.14	48.54
Science	OVERALL TEST	39	12,573	13.68	5.41	0.75	2.72	36.15
	Matter and Energy	5	12,573	2.22	1.25	0.32	1.03	41.55
	Force, Motion, and Energy	9	12,573	3.01	1.49	0.24	1.30	31.83
	Earth and Space	13	12,573	3.94	2.23	0.56	1.47	33.18
	Organisms and Environments	12	12,573	4.51	2.29	0.54	1.56	39.75

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 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	145	15.66	9.00	0.90	2.83	37.62
	Numerical Representations and Relationships	7	145	2.57	1.89	0.65	1.13	35.69
	Computations and Algebraic Relationships	20	145	7.14	4.84	0.84	1.94	35.73
	Geometry and Measurement	10	145	3.73	2.25	0.59	1.44	38.75
	Data Analysis and Personal Financial Literacy	5	145	2.22	1.31	0.51	0.92	45.78
RLA	OVERALL TEST	52	197	24.36	10.60	0.91	3.09	50.69
	Reading	26	197	13.77	5.21	0.82	2.21	54.01
	Writing	26	197	10.59	6.03	0.87	2.16	46.00
Science	OVERALL TEST	39	157	12.99	5.94	0.80	2.67	34.40
	Matter and Energy	5	157	2.08	1.23	0.27	1.06	39.25
	Force, Motion, and Energy	9	157	2.87	1.56	0.33	1.28	30.29
	Earth and Space	13	157	3.54	2.37	0.67	1.37	29.69
	Organisms and Environments	12	157	4.51	2.51	0.63	1.54	39.49

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.4.71. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	2,576	14.48	6.39	0.79	2.95	34.99
	Numerical Representations and Relationships	7	2,576	2.38	1.51	0.39	1.18	33.67
	Computations and Algebraic Relationships	20	2,576	6.57	3.61	0.68	2.04	33.58
	Geometry and Measurement	10	2,576	3.40	1.97	0.44	1.47	34.79
	Data Analysis and Personal Financial Literacy	5	2,576	2.13	1.12	0.25	0.98	43.02
RLA	OVERALL TEST	52	4,511	20.36	8.54	0.87	3.09	43.61
	Reading	26	4,511	11.83	4.49	0.74	2.27	46.63
	Writing	26	4,511	8.53	4.89	0.82	2.09	39.35
Science	OVERALL TEST	39	3,065	12.27	4.71	0.67	2.69	32.60
	Matter and Energy	5	3,065	1.97	1.22	0.26	1.05	36.77
	Force, Motion, and Energy	9	3,065	2.82	1.41	0.16	1.29	29.94
	Earth and Space	13	3,065	3.45	1.94	0.47	1.41	29.64
	Organisms and Environments	12	3,065	4.04	2.14	0.48	1.55	35.44

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.4.72. Spring 2024 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Mathematics	OVERALL TEST	42	8,976	17.10	7.78	0.86	2.96	40.91
	Numerical Representations and Relationships	7	8,976	2.76	1.71	0.52	1.18	38.85
	Computations and Algebraic Relationships	20	8,976	7.88	4.32	0.77	2.07	39.46
	Geometry and Measurement	10	8,976	4.03	2.14	0.55	1.44	40.96
	Data Analysis and Personal Financial Literacy	5	8,976	2.43	1.16	0.30	0.97	49.72
RLA	OVERALL TEST	52	13,999	27.08	10.00	0.91	3.07	56.13
	Reading	26	13,999	15.20	4.86	0.80	2.19	59.63
	Writing	26	13,999	11.87	5.87	0.87	2.15	51.20
Science	OVERALL TEST	39	9,971	14.03	5.54	0.76	2.72	37.03
	Matter and Energy	5	9,971	2.28	1.25	0.33	1.02	42.81
	Force, Motion, and Energy	9	9,971	3.06	1.51	0.25	1.31	32.30
	Earth and Space	13	9,971	4.06	2.28	0.58	1.48	34.01
	Organisms and Environments	12	9,971	4.63	2.32	0.55	1.55	40.85

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.4.73. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	467,451	31.06	12.38	0.92	3.40	52.76
	Number and Algebraic Methods	12	467,451	6.12	2.79	0.69	1.54	52.91
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	467,451	7.59	3.42	0.76	1.66	53.48
	Writing and Solving Linear Functions, Equations, and Inequalities	13	467,451	6.73	3.16	0.74	1.62	52.82
	Quadratic Functions and Equations	13	467,451	6.71	2.84	0.67	1.62	50.83
	Exponential Functions and Equations	7	467,451	3.91	1.94	0.65	1.15	54.66

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.4.74. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	224,561	31.60	12.13	0.92	3.40	53.80
	Number and Algebraic Methods	12	224,561	6.34	2.74	0.68	1.54	54.92
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	224,561	7.67	3.38	0.76	1.66	54.11
	Writing and Solving Linear Functions, Equations, and Inequalities	13	224,561	6.86	3.11	0.73	1.61	53.96
	Quadratic Functions and Equations	13	224,561	6.77	2.81	0.67	1.62	51.36
	Exponential Functions and Equations	7	224,561	3.96	1.91	0.63	1.15	55.49

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.4.75. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	242,821	30.57	12.58	0.93	3.40	51.82
	Number and Algebraic Methods	12	242,821	5.92	2.82	0.70	1.54	51.06
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	242,821	7.51	3.45	0.77	1.65	52.90
	Writing and Solving Linear Functions, Equations, and Inequalities	13	242,821	6.61	3.21	0.75	1.62	51.77
	Quadratic Functions and Equations	13	242,821	6.67	2.87	0.68	1.63	50.34
	Exponential Functions and Equations	7	242,821	3.86	1.96	0.66	1.15	53.89

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.4.76. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability 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	63,445	27.10	11.14	0.90	3.45	45.95
	Number and Algebraic Methods	12	63,445	5.38	2.60	0.64	1.56	46.53
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	63,445	6.51	3.14	0.71	1.68	45.84
	Writing and Solving Linear Functions, Equations, and Inequalities	13	63,445	5.85	2.86	0.67	1.65	45.98
	Quadratic Functions and Equations	13	63,445	5.99	2.67	0.62	1.65	45.02
	Exponential Functions and Equations	7	63,445	3.37	1.85	0.60	1.18	46.84

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.4.77. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	252,160	29.22	11.50	0.91	3.44	49.70
	Number and Algebraic Methods	12	252,160	5.81	2.65	0.66	1.55	50.32
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	252,160	7.10	3.25	0.73	1.69	50.03
	Writing and Solving Linear Functions, Equations, and Inequalities	13	252,160	6.28	2.97	0.70	1.63	49.50
	Quadratic Functions and Equations	13	252,160	6.38	2.70	0.63	1.64	48.25
	Exponential Functions and Equations	7	252,160	3.66	1.86	0.61	1.17	51.08

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 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Algebra I	OVERALL TEST	59	109,516	34.82	12.17	0.92	3.35	59.07
	Number and Algebraic Methods	12	109,516	6.73	2.75	0.69	1.54	58.07
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	109,516	8.67	3.35	0.77	1.62	61.12
	Writing and Solving Linear Functions, Equations, and Inequalities	13	109,516	7.61	3.17	0.75	1.59	59.31
	Quadratic Functions and Equations	13	109,516	7.35	2.78	0.68	1.57	55.79
	Exponential Functions and Equations	7	109,516	4.46	1.89	0.64	1.13	62.48

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 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability 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	117,577	23.64	9.22	0.86	3.48	40.10
	Number and Algebraic Methods	12	117,577	4.79	2.36	0.57	1.55	41.47
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	117,577	5.69	2.75	0.62	1.70	40.01
	Writing and Solving Linear Functions, Equations, and Inequalities	13	117,577	4.98	2.44	0.55	1.64	39.65
	Quadratic Functions and Equations	13	117,577	5.31	2.35	0.49	1.67	39.51
	Exponential Functions and Equations	7	117,577	2.88	1.63	0.47	1.19	39.92

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.4.80. Spring 2024 STAAR EOC Assessments Algebra I Mean P-Values, Raw Score Summary, and Reliability 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	349,874	33.56	12.30	0.93	3.37	57.02
	Number and Algebraic Methods	12	349,874	6.57	2.78	0.69	1.53	56.76
	Describing and Graphing Linear Functions, Equations, and Inequalities	14	349,874	8.22	3.39	0.77	1.64	58.00
	Writing and Solving Linear Functions, Equations, and Inequalities	13	349,874	7.32	3.16	0.74	1.60	57.24
	Quadratic Functions and Equations	13	349,874	7.19	2.83	0.68	1.59	54.64
	Exponential Functions and Equations	7	349,874	4.26	1.91	0.65	1.14	59.61

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.4.81. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	487,979	36.18	15.38	0.95	3.32	59.38
	Reading	32	487,979	18.45	7.42	0.90	2.34	57.09
	Writing	32	487,979	17.73	8.47	0.92	2.34	62.27

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.4.82. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	230,425	38.90	15.06	0.95	3.28	62.97
	Reading	32	230,425	19.59	7.24	0.90	2.31	60.53
	Writing	32	230,425	19.31	8.32	0.92	2.32	66.03

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.4.83. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	257,483	33.75	15.26	0.95	3.36	56.17
	Reading	32	257,483	17.43	7.43	0.90	2.37	54.00
	Writing	32	257,483	16.32	8.36	0.92	2.36	58.90

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 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability 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	65,144	32.54	14.25	0.94	3.41	54.18
	Reading	32	65,144	16.81	6.99	0.88	2.41	51.96
	Writing	32	65,144	15.72	7.85	0.91	2.40	56.97

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 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	271,115	33.17	14.88	0.95	3.39	54.92
	Reading	32	271,115	17.02	7.21	0.89	2.39	52.59
	Writing	32	271,115	16.15	8.21	0.92	2.39	57.85

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.4.86. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English I	OVERALL TEST	64	110,111	42.94	14.00	0.95	3.16	69.43
	Reading	32	110,111	21.64	6.76	0.89	2.23	67.16
	Writing	32	110,111	21.30	7.77	0.92	2.23	72.28

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.4.87. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability 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	134,873	23.71	11.13	0.90	3.48	40.78
	Reading	32	134,873	12.77	5.79	0.82	2.46	39.24
	Writing	32	134,873	10.93	6.04	0.84	2.45	42.71

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.4.88. Spring 2024 STAAR EOC Assessments English I Mean P-Values, Raw Score Summary, and Reliability 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	353,106	40.94	14.05	0.95	3.24	66.48
	Reading	32	353,106	20.61	6.80	0.89	2.28	63.90
	Writing	32	353,106	20.33	7.80	0.91	2.28	69.74

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 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	464,024	37.52	13.93	0.94	3.42	60.58
	Reading	32	464,024	18.87	6.19	0.85	2.39	58.78
	Writing	32	464,024	18.65	8.34	0.92	2.42	62.86

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 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	222,193	39.64	13.38	0.94	3.39	63.36
	Reading	32	222,193	19.71	5.93	0.84	2.38	61.27
	Writing	32	222,193	19.93	8.05	0.91	2.41	66.00

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 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	241,776	35.57	14.15	0.94	3.44	58.03
	Reading	32	241,776	18.11	6.33	0.86	2.41	56.49
	Writing	32	241,776	17.47	8.44	0.92	2.43	59.97

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 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability 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	60,445	34.01	13.06	0.93	3.49	55.87
	Reading	32	60,445	17.42	5.95	0.83	2.44	54.21
	Writing	32	60,445	16.59	7.79	0.90	2.48	57.96

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.4.93. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	255,543	34.96	13.71	0.94	3.47	56.88
	Reading	32	255,543	17.83	6.17	0.85	2.43	55.47
	Writing	32	255,543	17.14	8.16	0.91	2.47	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.4.94. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
English II	OVERALL TEST	64	107,904	43.04	12.44	0.93	3.30	68.46
	Reading	32	107,904	21.16	5.47	0.82	2.33	65.97
	Writing	32	107,904	21.88	7.61	0.91	2.33	71.59

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.4.95. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability 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	114,091	25.10	10.89	0.89	3.54	42.54
	Reading	32	114,091	13.76	5.35	0.78	2.48	42.66
	Writing	32	114,091	11.34	6.32	0.84	2.51	42.38

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.4.96. Spring 2024 STAAR EOC Assessments English II Mean P-Values, Raw Score Summary, and Reliability 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	349,933	41.57	12.34	0.93	3.36	66.46
	Reading	32	349,933	20.54	5.50	0.82	2.36	64.03
	Writing	32	349,933	21.03	7.50	0.90	2.38	69.53

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.4.97. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	439,213	26.84	10.51	0.91	3.18	52.32
	Cell Structure and Function	10	439,213	4.26	2.36	0.72	1.24	44.20
	Mechanisms of Genetics	11	439,213	5.36	2.33	0.63	1.41	51.18
	Biological Evolution and Classification	10	439,213	5.67	2.23	0.64	1.34	57.54
	Biological Processes and Systems	10	439,213	5.45	2.43	0.63	1.47	54.98
	Interdependence within Environmental Systems	12	439,213	6.11	3.01	0.77	1.43	53.81

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.4.98. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	213,636	26.79	10.32	0.90	3.19	52.17
	Cell Structure and Function	10	213,636	4.25	2.35	0.72	1.24	44.06
	Mechanisms of Genetics	11	213,636	5.44	2.33	0.63	1.41	51.90
	Biological Evolution and Classification	10	213,636	5.66	2.15	0.62	1.33	57.62
	Biological Processes and Systems	10	213,636	5.40	2.40	0.62	1.47	54.46
	Interdependence within Environmental Systems	12	213,636	6.04	2.97	0.76	1.44	52.98

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.4.99. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	225,517	26.89	10.69	0.91	3.17	52.46
	Cell Structure and Function	10	225,517	4.27	2.36	0.73	1.24	44.34
	Mechanisms of Genetics	11	225,517	5.28	2.32	0.63	1.41	50.49
	Biological Evolution and Classification	10	225,517	5.68	2.30	0.66	1.35	57.48
	Biological Processes and Systems	10	225,517	5.48	2.45	0.65	1.46	55.48
	Interdependence within Environmental Systems	12	225,517	6.17	3.06	0.78	1.43	54.60

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 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	56,133	23.65	9.59	0.89	3.21	46.31
	Cell Structure and Function	10	56,133	3.73	2.16	0.66	1.26	38.86
	Mechanisms of Genetics	11	56,133	4.84	2.19	0.58	1.41	46.08
	Biological Evolution and Classification	10	56,133	5.05	2.16	0.59	1.39	51.18
	Biological Processes and Systems	10	56,133	4.82	2.32	0.60	1.47	49.02
	Interdependence within Environmental Systems	12	56,133	5.21	2.77	0.73	1.44	46.67

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.4.101. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	236,883	24.60	9.62	0.89	3.21	48.15
	Cell Structure and Function	10	236,883	3.85	2.17	0.67	1.25	40.17
	Mechanisms of Genetics	11	236,883	4.92	2.16	0.57	1.42	47.17
	Biological Evolution and Classification	10	236,883	5.26	2.15	0.59	1.37	53.38
	Biological Processes and Systems	10	236,883	5.03	2.30	0.58	1.49	50.86
	Interdependence within Environmental Systems	12	236,883	5.54	2.84	0.74	1.45	49.34

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.4.102. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	106,711	31.19	10.10	0.90	3.14	60.50
	Cell Structure and Function	10	106,711	5.00	2.38	0.74	1.21	51.43
	Mechanisms of Genetics	11	106,711	6.15	2.28	0.63	1.40	58.60
	Biological Evolution and Classification	10	106,711	6.52	2.05	0.61	1.27	66.24
	Biological Processes and Systems	10	106,711	6.24	2.36	0.62	1.45	62.89
	Interdependence within Environmental Systems	12	106,711	7.28	2.91	0.77	1.40	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.4.103. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	107,319	19.68	7.83	0.83	3.19	38.81
	Cell Structure and Function	10	107,319	3.03	1.80	0.50	1.27	32.32
	Mechanisms of Genetics	11	107,319	4.04	1.86	0.42	1.42	38.73
	Biological Evolution and Classification	10	107,319	4.32	2.07	0.52	1.43	43.46
	Biological Processes and Systems	10	107,319	4.01	1.97	0.46	1.45	40.45
	Interdependence within Environmental Systems	12	107,319	4.29	2.46	0.65	1.45	39.21

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.4.104. Spring 2024 STAAR EOC Assessments Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Biology	OVERALL TEST	53	331,894	29.16	10.22	0.90	3.17	56.69
	Cell Structure and Function	10	331,894	4.66	2.38	0.73	1.22	48.04
	Mechanisms of Genetics	11	331,894	5.79	2.30	0.63	1.40	55.20
	Biological Evolution and Classification	10	331,894	6.10	2.10	0.61	1.30	62.10
	Biological Processes and Systems	10	331,894	5.91	2.38	0.62	1.46	59.68
	Interdependence within Environmental Systems	12	331,894	6.69	2.94	0.77	1.42	58.54

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 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	388,334	44.57	14.00	0.92	3.90	56.90
	History	36	388,334	21.91	6.72	0.85	2.60	60.43
	Geography and Culture	14	388,334	7.80	2.84	0.73	1.48	56.06
	Government and Citizenship	11	388,334	5.72	2.48	0.58	1.60	52.20
	Economics, Science, Technology, and Society	17	388,334	9.13	3.51	0.75	1.76	52.99

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.4.106. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	191,230	43.80	13.50	0.92	3.92	55.87
	History	36	191,230	21.37	6.47	0.84	2.63	58.88
	Geography and Culture	14	191,230	7.70	2.74	0.71	1.49	55.40
	Government and Citizenship	11	191,230	5.68	2.44	0.56	1.62	51.71
	Economics, Science, Technology, and Society	17	191,230	9.04	3.43	0.74	1.76	52.47

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 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	197,057	45.32	14.44	0.93	3.88	57.89
	History	36	197,057	22.43	6.91	0.86	2.58	61.95
	Geography and Culture	14	197,057	7.91	2.93	0.74	1.48	56.69
	Government and Citizenship	11	197,057	5.76	2.51	0.60	1.59	52.68
	Economics, Science, Technology, and Society	17	197,057	9.22	3.59	0.76	1.75	53.49

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.4.108. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	48,442	40.33	13.20	0.91	3.97	51.55
	History	36	48,442	19.96	6.44	0.83	2.67	55.08
	Geography and Culture	14	48,442	7.06	2.75	0.70	1.50	50.82
	Government and Citizenship	11	48,442	5.20	2.33	0.52	1.61	47.45
	Economics, Science, Technology, and Society	17	48,442	8.11	3.33	0.71	1.80	47.18

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.4.109. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	207,159	41.96	13.43	0.91	3.96	53.56
	History	36	207,159	20.73	6.55	0.84	2.65	57.16
	Geography and Culture	14	207,159	7.39	2.80	0.71	1.51	53.02
	Government and Citizenship	11	207,159	5.30	2.36	0.53	1.61	48.45
	Economics, Science, Technology, and Society	17	207,159	8.54	3.38	0.72	1.78	49.56

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.4.110. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
U.S. History	OVERALL TEST	78	98,925	49.98	13.11	0.92	3.79	63.76
	History	36	98,925	24.43	6.21	0.84	2.51	67.38
	Geography and Culture	14	98,925	8.69	2.66	0.71	1.44	62.36
	Government and Citizenship	11	98,925	6.52	2.45	0.58	1.58	59.29
	Economics, Science, Technology, and Society	17	98,925	10.34	3.34	0.74	1.71	59.96

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.4.111. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	73,723	33.71	12.09	0.89	4.01	43.02
	History	36	73,723	16.88	6.26	0.81	2.73	46.40
	Geography and Culture	14	73,723	5.79	2.61	0.64	1.57	41.36
	Government and Citizenship	11	73,723	4.31	2.09	0.44	1.57	39.62
	Economics, Science, Technology, and Society	17	73,723	6.73	3.01	0.65	1.79	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.4.112. Spring 2024 STAAR EOC Assessments U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	314,611	47.11	13.18	0.91	3.86	60.15
	History	36	314,611	23.09	6.26	0.83	2.57	63.72
	Geography and Culture	14	314,611	8.28	2.68	0.70	1.46	59.50
	Government and Citizenship	11	314,611	6.05	2.44	0.57	1.60	55.15
	Economics, Science, Technology, and Society	17	314,611	9.70	3.38	0.73	1.74	56.21

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.

Spring 2024 STAAR
Progress Measure

Table B.5.1. Spring 2024 STAAR Grade 4 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	98428	37262	35028	1392	-	15363	4571	-	391	4080	-	-	341	37262	50782	10384
		38	36	1	-	16	5	-	-	4	-	-	-	38	52	11
Approaches	104559	32715	1094	-	13446	27409	-	-	18027	7828	-	-	4040	46161	46530	11868
		31	1	-	13	26	-	-	17	7	-	-	4	44	45	11
Meets	90554	5064	-	-	20395	-	-	14955	27523	-	-	11624	10993	40414	39147	10993
		6	-	-	23	-	-	17	30	-	-	13	12	45	43	12
Masters	65671	330	-	-	2218	-	-	15266	-	-	-	34147	13710	17814	34147	13710
		1	-	-	3	-	-	23	-	-	-	52	21	27	52	21
2023 Total	359212	75371	36122	1392	36059	42772	4571	30221	45941	11908	-	45771	29084	141651	170606	46955
		21	10	-	10	12	1	8	13	3	-	13	8	39	47	13

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.5.2. Spring 2024 STAAR Grade 4 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	77296	25619	18216	2754	540	15547	11619	-	-	2855	-	-	146	26159	33763	17374
		33	24	4	1	20	15	-	-	4	-	-	-	34	44	22
Approaches	90061	12149	-	-	23160	27542	222	-	16940	7288	-	-	2760	35309	44482	10270
		13	-	-	26	31	-	-	19	8	-	-	3	39	49	11
Meets	104780	1229	-	-	23308	-	-	25043	28433	-	716	16468	9583	50296	44901	9583
		1	-	-	22	-	-	24	27	-	1	16	9	48	43	9
Masters	68303	63	-	-	1537	-	-	15864	-	-	-	44723	6116	17464	44723	6116
		-	-	-	2	-	-	23	-	-	-	65	9	26	65	9
2023 Total	340440	39060	18216	2754	48545	43089	11841	40907	45373	10143	716	61191	18605	129228	167869	43343
		11	5	1	14	13	3	12	13	3	-	18	5	38	49	13

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.5.3. Spring 2024 STAAR Grade 5 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	109585	34545	30278	929	798	28626	5369	-	2477	6248	-	-	315	35343	61381	12861
		32	28	1	1	26	5	-	2	6	-	-	-	32	56	12
Approaches	84006	15497	-	-	18603	19962	-	-	23566	4127	-	-	2251	34100	43528	6378
		18	-	-	22	24	-	-	28	5	-	-	3	41	52	8
Meets	91093	3266	-	-	21807	-	-	18811	31275	-	-	9939	5995	43884	41214	5995
		4	-	-	24	-	-	21	34	-	-	11	7	48	45	7
Masters	78302	225	-	-	2559	-	-	23618	-	-	-	39350	12550	26402	39350	12550
		-	-	-	3	-	-	30	-	-	-	50	16	34	50	16
2023 Total	362986	53533	30278	929	43767	48588	5369	42429	57318	10375	-	49289	21111	139729	185473	37784
		15	8	-	12	13	1	12	16	3	-	14	6	38	51	10

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.5.4. Spring 2024 STAAR Grade 5 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	76675	24917	22958	2915	-	8791	13030	-	-	3750	-	-	314	24917	31749	20009
		32	30	4	-	11	17	-	-	5	-	-	-	32	41	26
Approaches	107092	18816	663	-	15230	31577	1664	-	17803	14602	-	-	6737	34046	50043	23003
		18	1	-	14	29	2	-	17	14	-	-	6	32	47	21
Meets	90796	1210	-	-	14160	-	-	13998	27181	-	-	18320	15927	29368	45501	15927
		1	-	-	16	-	-	15	30	-	-	20	18	32	50	18
Masters	75659	68	-	-	1137	-	-	12554	-	-	-	48982	12918	13759	48982	12918
		-	-	-	2	-	-	17	-	-	-	65	17	18	65	17
2023 Total	350222	45011	23621	2915	30527	40368	14694	26552	44984	18352	-	67302	35896	102090	176275	71857
		13	7	1	9	12	4	8	13	5	-	19	10	29	50	21

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.5.5. Spring 2024 STAAR Grade 6 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	75846	27066	28604	655	1	16289	1861	-	106	1221	-	-	43	27067	44999	3780
		36	38	1	-	21	2	-	-	2	-	-	-	36	59	5
Approaches	109968	40943	-	-	30726	23753	-	-	12280	1711	-	-	555	71669	36033	2266
		37	-	-	28	22	-	-	11	2	-	-	1	65	33	2
Meets	106043	8491	-	-	41919	-	-	24747	22298	-	-	6093	2495	75157	28391	2495
		8	-	-	40	-	-	23	21	-	-	6	2	71	27	2
Masters	72801	472	-	-	5347	-	-	28274	-	-	-	31146	7562	34093	31146	7562
		1	-	-	7	-	-	39	-	-	-	43	10	47	43	10
2023 Total	364658	76972	28604	655	77993	40042	1861	53021	34684	2932	-	37239	10655	207986	140569	16103
		21	8	-	21	11	1	15	10	1	-	10	3	57	39	4

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.5.6. Spring 2024 STAAR Grade 6 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	68284	22343	24386	5285	-	7214	6282	-	-	2703	-	-	71	22343	31600	14341
		33	36	8	-	11	9	-	-	4	-	-	-	33	46	21
Approaches	87868	21914	1441	-	13776	22716	-	-	20545	5520	-	-	1956	35690	44702	7476
		25	2	-	16	26	-	-	23	6	-	-	2	41	51	9
Meets	98896	4277	-	-	19947	-	-	25595	28766	-	-	13794	6517	49819	42560	6517
		4	-	-	20	-	-	26	29	-	-	14	7	50	43	7
Masters	101128	357	-	-	2840	-	-	27740	-	-	-	62992	7199	30937	62992	7199
		-	-	-	3	-	-	27	-	-	-	62	7	31	62	7
2023 Total	356176	48891	25827	5285	36563	29930	6282	53335	49311	8223	-	76786	15743	138789	181854	35533
		14	7	1	10	8	2	15	14	2	-	22	4	39	51	10

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.5.7. Spring 2024 STAAR Grade 7 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	89208	28591	47698	515	-	9921	932	-	337	1189	-	-	25	28591	57956	2661
		32	53	1	-	11	1	-	-	1	-	-	-	32	65	3
Approaches	117472	49115	7106	-	12118	28750	-	-	17506	2151	-	-	726	61233	53362	2877
		42	6	-	10	24	-	-	15	2	-	-	1	52	45	2
Meets	62239	4038	-	-	13786	-	-	11331	24670	-	-	5401	3013	29155	30071	3013
		6	-	-	22	-	-	18	40	-	-	9	5	47	48	5
Masters	33390	125	-	-	596	-	-	10243	-	-	-	16534	5892	10964	16534	5892
		-	-	-	2	-	-	31	-	-	-	50	18	33	50	18
2023 Total	302309	81869	54804	515	26500	38671	932	21574	42513	3340	-	21935	9656	129943	157923	14443
		27	18	-	9	13	-	7	14	1	-	7	3	43	52	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.5.8. Spring 2024 STAAR Grade 7 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	90906	45032	23101	2341	-	11749	5224	-	-	3270	-	-	189	45032	34850	11024
		50	25	3	-	13	6	-	-	4	-	-	-	50	38	12
Approaches	92762	24405	-	-	19858	19688	-	-	19534	5832	-	-	3445	44263	39222	9277
		26	-	-	21	21	-	-	21	6	-	-	4	48	42	10
Meets	109298	3089	-	-	18960	-	-	23513	28744	-	674	21177	13141	46236	49921	13141
		3	-	-	17	-	-	22	26	-	1	19	12	42	46	12
Masters	81118	93	-	-	886	-	-	11657	-	-	-	57270	11212	12636	57270	11212
		-	-	-	1	-	-	14	-	-	-	71	14	16	71	14
2023 Total	374084	72619	23101	2341	39704	31437	5224	35170	48278	9102	674	78447	27987	148167	181263	44654
		19	6	1	11	8	1	9	13	2	-	21	7	40	48	12

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.5.9. Spring 2024 STAAR Grade 8 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	120783	31185	37191	610	888	34892	3846	-	3472	8080	-	-	619	32073	75555	13155
		26	31	1	1	29	3	-	3	7	-	-	1	27	63	11
Approaches	88865	21163	356	-	17950	20814	-	-	22334	3431	-	-	2817	39113	43504	6248
		24	-	-	20	23	-	-	25	4	-	-	3	44	49	7
Meets	77222	3440	-	-	18304	-	-	17483	21671	-	139	10615	5570	39366	32286	5570
		4	-	-	24	-	-	23	28	-	-	14	7	51	42	7
Masters	33325	109	-	-	1017	-	-	8151	-	-	-	19021	5027	9277	19021	5027
		-	-	-	3	-	-	24	-	-	-	57	15	28	57	15
2023 Total	320195	55897	37547	610	38159	55706	3846	25634	47477	11511	139	29636	14033	119829	170366	30000
		17	12	-	12	17	1	8	15	4	-	9	4	37	53	9

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.5.10. Spring 2024 STAAR Grade 8 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	87294	25190	24808	7870	-	14569	11103	-	-	3606	-	-	148	25190	39377	22727
		29	28	9	-	17	13	-	-	4	-	-	-	29	45	26
Approaches	92666	15633	-	-	19990	26645	-	-	21591	5940	-	-	2867	35623	48236	8807
		17	-	-	22	29	-	-	23	6	-	-	3	38	52	10
Meets	100558	1989	-	-	19985	-	-	22283	29450	-	-	17821	9030	44257	47271	9030
		2	-	-	20	-	-	22	29	-	-	18	9	44	47	9
Masters	99486	107	-	-	1572	-	-	17578	-	-	-	68444	11785	19257	68444	11785
		-	-	-	2	-	-	18	-	-	-	69	12	19	69	12
2023 Total	380004	42919	24808	7870	41547	41214	11103	39861	51041	9546	-	86265	23830	124327	203328	52349
		11	7	2	11	11	3	10	13	3	-	23	6	33	54	14

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.5.11. Spring 2024 STAAR Spanish Grade 4 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	4350	1795	1522	57	-	590	197	-	15	159	-	-	15	1795	2127	428
		41	35	1	-	14	5	-	-	4	-	-	-	41	49	10
Approaches	3022	1157	46	-	323	726	-	-	433	228	-	-	109	1480	1205	337
		38	2	-	11	24	-	-	14	8	-	-	4	49	40	11
Meets	1424	115	-	-	358	-	-	205	432	-	-	140	174	678	572	174
		8	-	-	25	-	-	14	30	-	-	10	12	48	40	12
Masters	487	6	-	-	24	-	-	152	-	-	-	245	60	182	245	60
		1	-	-	5	-	-	31	-	-	-	50	12	37	50	12
2023 Total	9283	3073	1568	57	705	1316	197	357	880	387	-	385	358	4135	4149	999
		33	17	1	8	14	2	4	9	4	-	4	4	45	45	11

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.5.12. Spring 2024 STAAR Spanish Grade 4 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	8187	1758	2883	2042	-	123	952	-	-	375	-	-	54	1758	3006	3423
		21	35	25	-	2	12	-	-	5	-	-	1	21	37	42
Approaches	5397	1440	345	-	393	1265	97	-	711	716	-	-	430	1833	2321	1243
		27	6	-	7	23	2	-	13	13	-	-	8	34	43	23
Meets	2303	180	-	-	494	-	-	257	723	-	-	228	421	931	951	421
		8	-	-	21	-	-	11	31	-	-	10	18	40	41	18
Masters	2569	33	-	-	187	-	-	754	-	-	-	1223	372	974	1223	372
		1	-	-	7	-	-	29	-	-	-	48	14	38	48	14
2023 Total	18456	3411	3228	2042	1074	1388	1049	1011	1434	1091	-	1451	1277	5496	7501	5459
		18	17	11	6	8	6	5	8	6	-	8	7	30	41	30

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.5.13. Spring 2024 STAAR Spanish Grade 5 Mathematics Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	3738	1104	1134	45	25	919	220	-	47	227	-	-	17	1129	2100	509
		30	30	1	1	25	6	-	1	6	-	-	-	30	56	14
Approaches	1483	275	-	-	323	360	-	-	385	106	-	-	34	598	745	140
		19	-	-	22	24	-	-	26	7	-	-	2	40	50	9
Meets	852	36	-	-	232	-	-	176	285	-	-	76	47	444	361	47
		4	-	-	27	-	-	21	33	-	-	9	6	52	42	6
Masters	297	5	-	-	24	-	-	127	-	-	-	127	14	156	127	14
		2	-	-	8	-	-	43	-	-	-	43	5	53	43	5
2023 Total	6370	1420	1134	45	604	1279	220	303	717	333	-	203	112	2327	3333	710
		22	18	1	9	20	3	5	11	5	-	3	2	37	52	11

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.5.14. Spring 2024 STAAR Spanish Grade 5 RLA Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	5951	1744	1420	340	107	1272	726	-	32	297	-	-	13	1851	2724	1376
		29	24	6	2	21	12	-	1	5	-	-	-	31	46	23
Approaches	2535	270	-	-	809	621	-	-	575	171	-	-	89	1079	1196	260
		11	-	-	32	24	-	-	23	7	-	-	4	43	47	10
Meets	2205	56	-	-	661	-	-	425	744	-	-	174	145	1142	918	145
		3	-	-	30	-	-	19	34	-	-	8	7	52	42	7
Masters	1634	1	-	-	123	-	-	697	-	-	-	713	100	821	713	100
		-	-	-	8	-	-	43	-	-	-	44	6	50	44	6
2023 Total	12325	2071	1420	340	1700	1893	726	1122	1351	468	-	887	347	4893	5551	1881
		17	12	3	14	15	6	9	11	4	-	7	3	40	45	15

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.5.15. Spring 2024 STAAR EOC Assessments Algebra I Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	87337	39928	57	-	32235	7623	75	-	4969	1372	-	-	1078	72163	12649	2525
		46	-	-	37	9	-	-	6	2	-	-	1	83	14	3
Approaches	109987	17133	-	-	53624	5534	-	-	24928	1210	-	-	7558	70757	30462	8768
		16	-	-	49	5	-	-	23	1	-	-	7	64	28	8
Meets	118638	2854	-	-	31037	-	-	8318	35579	-	-	10685	30165	42209	46264	30165
		2	-	-	26	-	-	7	30	-	-	9	25	36	39	25
Masters	75211	94	-	-	2299	-	-	9917	-	-	-	14913	47988	12310	14913	47988
		-	-	-	3	-	-	13	-	-	-	20	64	16	20	64
2023 Total	391173	60009	57	-	119195	13157	75	18235	65476	2582	-	25598	86789	197439	104288	89446
		15	-	-	30	3	-	5	17	1	-	7	22	50	27	23

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.5.16. Spring 2024 STAAR EOC Assessments English I Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	65711	52622	331	-	3603	5742	23	-	3223	132	-	-	35	56225	9296	190
		80	1	-	5	9	-	-	5	-	-	-	-	86	14	-
Approaches	100104	29204	-	-	21930	8825	-	-	39184	422	-	-	539	51134	48009	961
		29	-	-	22	9	-	-	39	-	-	-	1	51	48	1
Meets	112051	4045	-	-	11868	-	-	3669	81164	-	-	3303	8002	19582	84467	8002
		4	-	-	11	-	-	3	72	-	-	3	7	17	75	7
Masters	108332	173	-	-	691	-	-	40330	-	-	-	25631	41507	41194	25631	41507
		-	-	-	1	-	-	37	-	-	-	24	38	38	24	38
2023 Total	386198	86044	331	-	38092	14567	23	43999	123571	554	-	28934	50083	168135	167403	50660
		22	-	-	10	4	-	11	32	-	-	7	13	44	43	13

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.5.17. Spring 2024 STAAR EOC Assessments English II Progress Measure

		2024														
		Did Not Meet			Approaches			Meets			Masters			2024 Total		
2023	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	71050	25367	24759	89	-	14935	124	-	5324	448	-	-	4	25367	45018	665
		36	35	-	-	21	-	-	7	1	-	-	-	36	63	1
Approaches	56296	12467	-	-	8564	12973	-	-	22248	30	-	-	14	21031	35221	44
		22	-	-	15	23	-	-	40	-	-	-	-	37	63	-
Meets	185968	3960	-	-	17245	-	-	72373	84991	-	-	7214	185	93578	92205	185
		2	-	-	9	-	-	39	46	-	-	4	-	50	50	-
Masters	69844	23	-	-	78	-	-	38469	-	-	-	30355	919	38570	30355	919
		-	-	-	-	-	-	55	-	-	-	43	1	55	43	1
2023 Total	383158	41817	24759	89	25887	27908	124	110842	112563	478	-	37569	1122	178546	202799	1813
		11	6	-	7	7	-	29	29	-	-	10	-	47	53	-

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.

Spring 2024 STAAR
On-Track Measure

Table B.6.1. Spring 2024 STAAR Grades 3–8 and English I On-Track Measure

				On Track to Be at or Above the Meets Grade Level Standard	
Current Grade	**Subject**	**Target Grade**	**Total N**	**Count**	**Percentage**
4	Mathematics	5	359212	182000	51
5	Mathematics	8	362986	215133	59
6	Mathematics	8	364658	136808	38
7	Mathematics	8	302309	129241	43
8	Mathematics	Algebra I	320195	158782	50
4	Mathematics Spanish	5	9283	3231	35
4	RLA	5	340440	188191	55
5	RLA	8	350222	235761	67
6	RLA	8	356176	216801	61
7	RLA	8	374084	210084	56
8	RLA	English I	380004	236472	62
English I	RLA	English II	386198	268819	70
4	RLA Spanish	5	18456	9681	52

Spring 2024 STAAR
Scale Score Descriptive Statistics

Table B.7.1. Spring 2024 STAAR Grades 3–8 Scale Score Descriptive Statistics

	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	373,236	1450.57	1435	1354	1210	219	152.01	23108.24	0.51	0.59
Grade 4	Mathematics	377,744	1556.47	1539	1388	1220	243	171.27	29333.51	0.63	0.52
Grade 5	Mathematics	380,589	1639.49	1622	1634	1200	232	165.30	27325.62	0.58	0.70
Grade 6	Mathematics	387,455	1715.58	1684	1600	1280	204	153.65	23608.21	0.84	1.43
Grade 7	Mathematics	324,109	1753.60	1703	1643	1250	202	157.44	24786.07	1.02	1.54
Grade 8	Mathematics	356,728	1848.78	1813	1739	1230	183	152.26	23184.54	1.02	1.92
Grade 3	RLA	359,806	1460.92	1464	1547	1400	235	158.24	25039.58	-0.01	0.17
Grade 4	RLA	368,529	1553.98	1552	1622	1390	213	153.17	23461.11	0.07	0.50
Grade 5	RLA	375,523	1605.99	1615	1592	1390	223	168.08	28251.73	0.05	0.44
Grade 6	RLA	394,029	1642.27	1644	1606	1400	205	150.35	22604.68	-0.01	0.55
Grade 7	RLA	397,569	1668.52	1669	1771	1400	242	167.93	28199.82	0.00	0.28
Grade 8	RLA	402,883	1715.68	1710	1720	1380	205	151.53	22961.29	0.19	0.74
Grade 5	Science	380,984	3689.37	3675	3545	5060	662	554.92	307940.40	0.26	0.99
Grade 8	Science	398,004	3926.69	3871	3377	5800	985	684.47	468494.29	0.50	0.66
Grade 8	Social Studies	405,802	3728.72	3678	3230	5500	849	595.83	355018.41	0.46	0.60

Table B.7.2. Spring 2024 STAAR Spanish Grades 3–5 Scale Score Descriptive Statistics

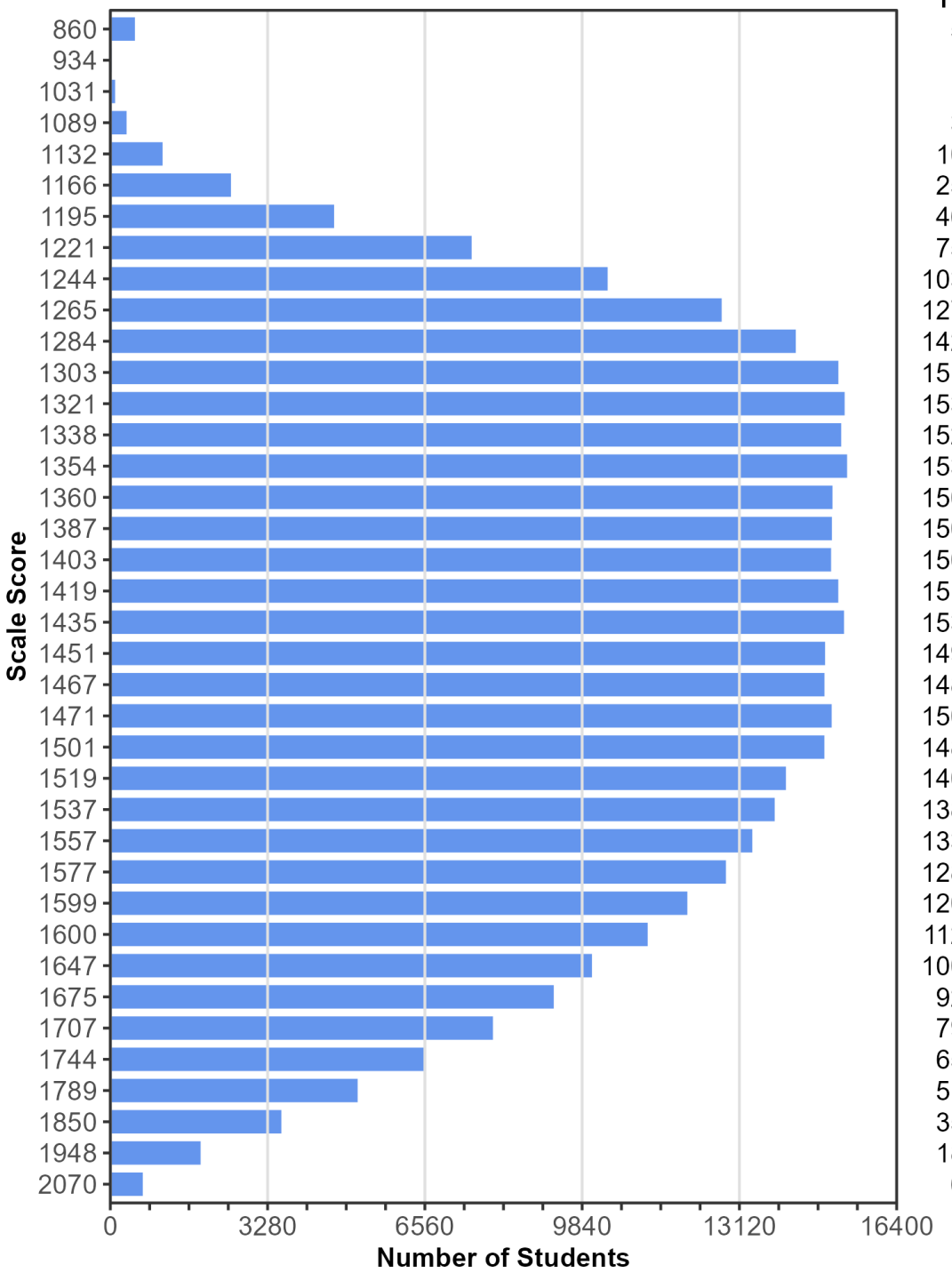
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	20,605	1373.84	1354	1284	1210	167	122.74	15065.02	0.72	0.85
Grade 4	Mathematics	14,990	1453.10	1421	1370	1220	154	131.32	17245.46	0.94	1.61
Grade 5	Mathematics	11,552	1526.63	1509	1448	1200	160	120.48	14514.64	0.65	0.96
Grade 3	RLA	34,257	1307.46	1297	1191	1470	196	138.97	19312.17	0.39	0.08
Grade 4	RLA	25,322	1406.81	1402	1284	1387	204	143.46	20579.36	0.30	0.17
Grade 5	RLA	18,510	1468.31	1469	1334	1460	190	141.25	19952.24	0.07	0.32
Grade 5	Science	13,036	3266.81	3264	3106	3985	525	411.19	169078.64	0.26	0.78

Table B.7.3. Spring 2024 STAAR EOC Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	467,451	4009.10	3901	3489	4930	738	566.93	321411.0	0.91	1.92
English I	487,979	4068.57	4067	3395	4250	843	546.90	299104.7	0.16	-0.25
English II	464,024	4093.52	4102	4553	4400	661	481.61	231943.8	-0.17	0.08
Biology	439,213	4121.82	4085	3803	4360	678	478.39	228858.2	0.44	0.70
U.S. History	388,334	4266.38	4258	4377	5330	697	471.46	222273.6	0.11	0.68

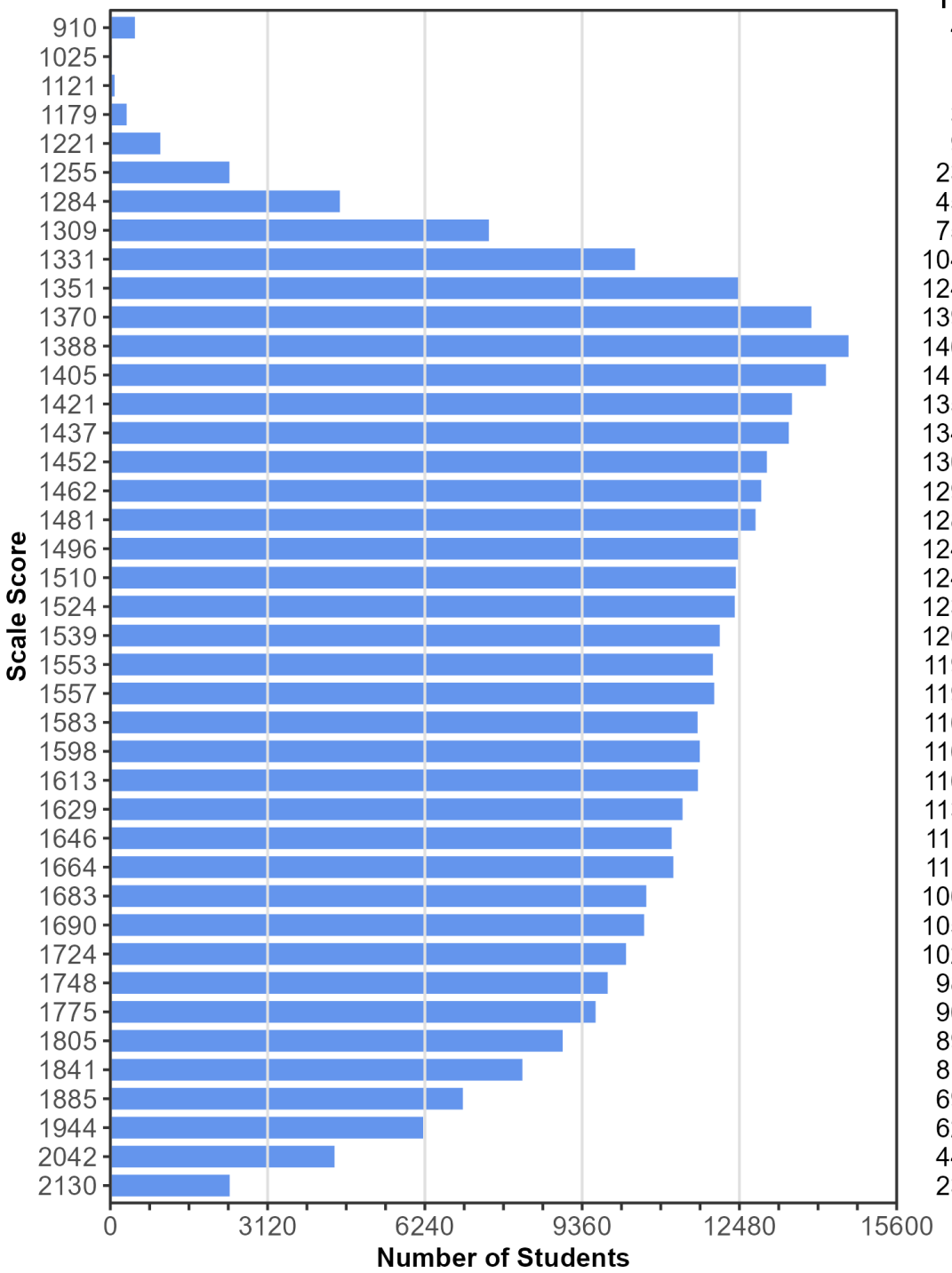
Spring 2024 STAAR
Frequency Distribution of Scale Scores

Figure B.7.1. Spring 2024 STAAR Grade 3 Mathematics
Frequency Distribution of Scale Scores



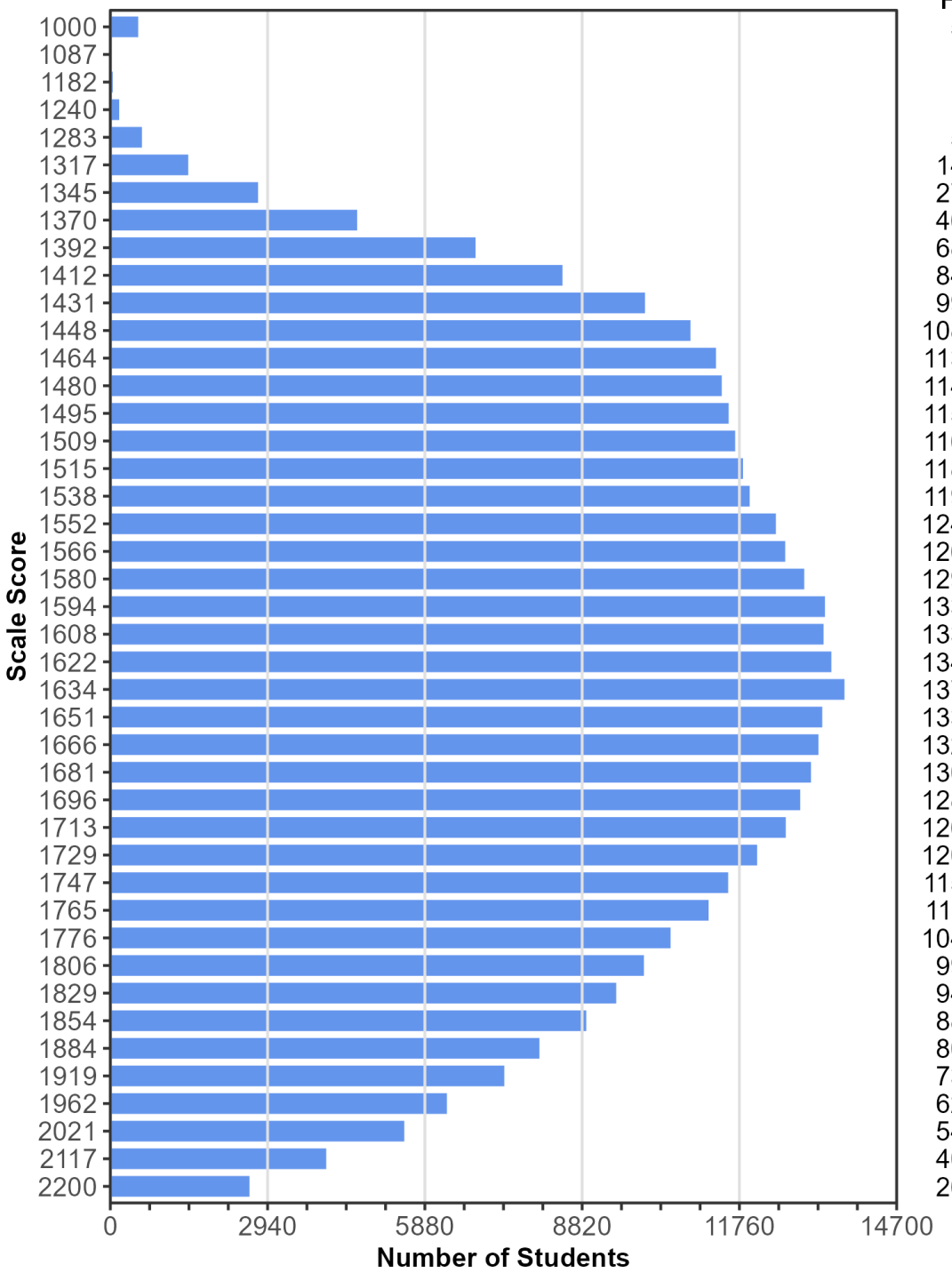
Scale Score	Freq	Cum Freq	Pct	Cum Pct
860	513	513	0.14	0.14
934	27	540	0.01	0.14
1031	100	640	0.03	0.17
1089	339	979	0.09	0.26
1132	1089	2068	0.29	0.55
1166	2516	4584	0.67	1.23
1195	4667	9251	1.25	2.48
1221	7536	16787	2.02	4.50
1244	10372	27159	2.78	7.28
1265	12750	39909	3.42	10.69
1284	14296	54205	3.83	14.52
1303	15184	69389	4.07	18.59
1321	15315	84704	4.10	22.69
1338	15243	99947	4.08	26.78
1354	15366	115313	4.12	30.90
1360	15061	130374	4.04	34.93
1387	15049	145423	4.03	38.96
1403	15032	160455	4.03	42.99
1419	15182	175637	4.07	47.06
1435	15302	190939	4.10	51.16
1451	14907	205846	3.99	55.15
1467	14894	220740	3.99	59.14
1471	15044	235784	4.03	63.17
1501	14892	250676	3.99	67.16
1519	14090	264766	3.78	70.94
1537	13856	278622	3.71	74.65
1557	13389	292011	3.59	78.24
1577	12839	304850	3.44	81.68
1599	12034	316884	3.22	84.90
1600	11207	328091	3.00	87.90
1647	10045	338136	2.69	90.60
1675	9248	347384	2.48	93.07
1707	7980	355364	2.14	95.21
1744	6587	361951	1.76	96.98
1789	5158	367109	1.38	98.36
1850	3567	370676	0.96	99.31
1948	1883	372559	0.50	99.82
2070	677	373236	0.18	100.00

Figure B.7.2. Spring 2024 STAAR Grade 4 Mathematics
Frequency Distribution of Scale Scores



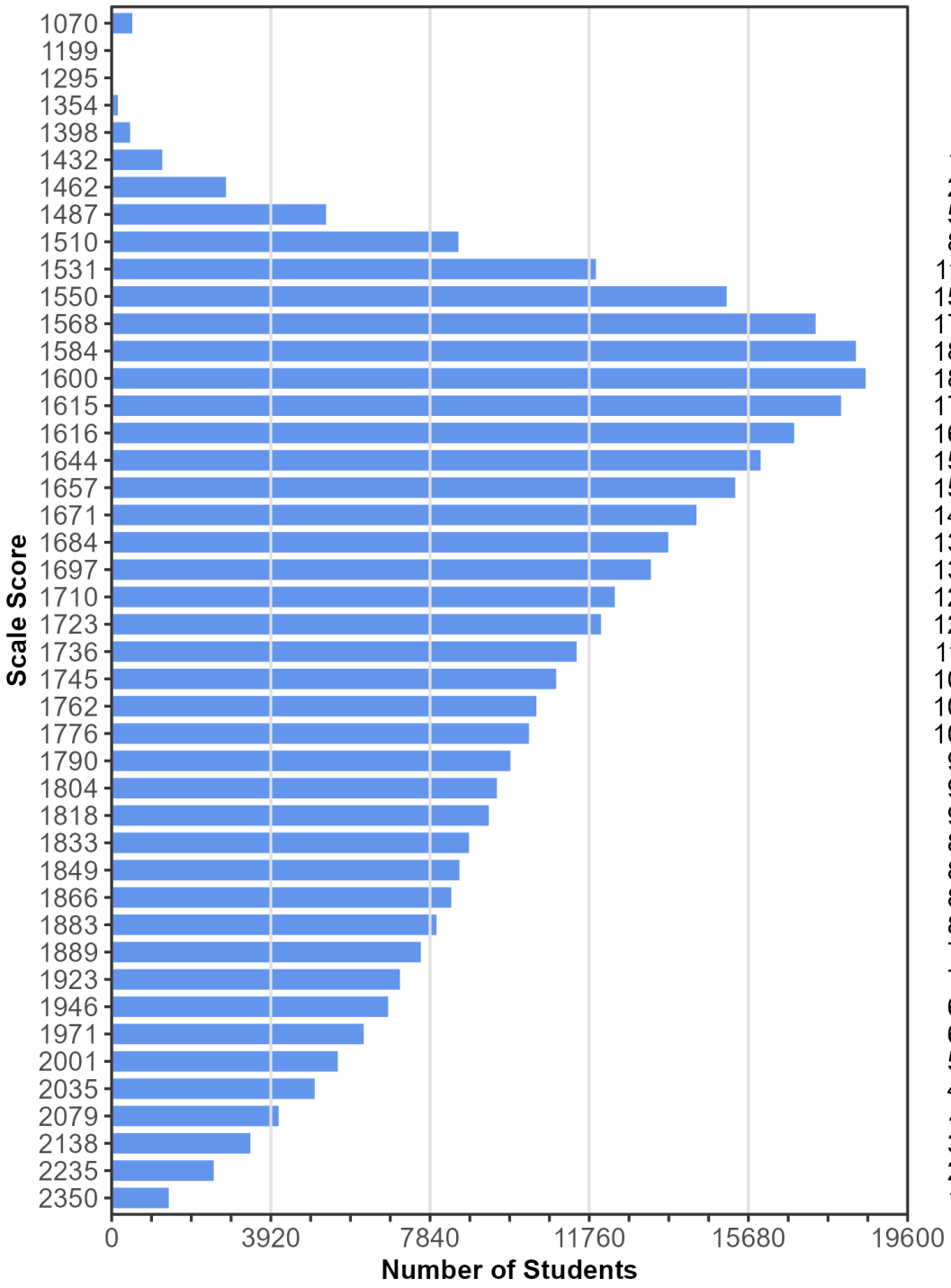
	Freq	Cum Freq	Pct	Cum Pct
910	488	488	0.13	0.13
1025	27	515	0.01	0.14
1121	84	599	0.02	0.16
1179	324	923	0.09	0.24
1221	992	1915	0.26	0.51
1255	2363	4278	0.63	1.13
1284	4554	8832	1.21	2.34
1309	7510	16342	1.99	4.33
1331	10410	26752	2.76	7.08
1351	12465	39217	3.30	10.38
1370	13908	53125	3.68	14.06
1388	14645	67770	3.88	17.94
1405	14197	81967	3.76	21.70
1421	13523	95490	3.58	25.28
1437	13460	108950	3.56	28.84
1452	13025	121975	3.45	32.29
1462	12913	134888	3.42	35.71
1481	12800	147688	3.39	39.10
1496	12464	160152	3.30	42.40
1510	12409	172561	3.29	45.68
1524	12388	184949	3.28	48.96
1539	12090	197039	3.20	52.16
1553	11954	208993	3.16	55.33
1557	11980	220973	3.17	58.50
1583	11650	232623	3.08	61.58
1598	11694	244317	3.10	64.68
1613	11657	255974	3.09	67.76
1629	11352	267326	3.01	70.77
1646	11136	278462	2.95	73.72
1664	11169	289631	2.96	76.67
1683	10633	300264	2.81	79.49
1690	10590	310854	2.80	82.29
1724	10231	321085	2.71	85.00
1748	9867	330952	2.61	87.61
1775	9627	340579	2.55	90.16
1805	8974	349553	2.38	92.54
1841	8175	357728	2.16	94.70
1885	6994	364722	1.85	96.55
1944	6206	370928	1.64	98.20
2042	4448	375376	1.18	99.37
2130	2368	377744	0.63	100.00

Figure B.7.3. Spring 2024 STAAR Grade 5 Mathematics
Frequency Distribution of Scale Scores



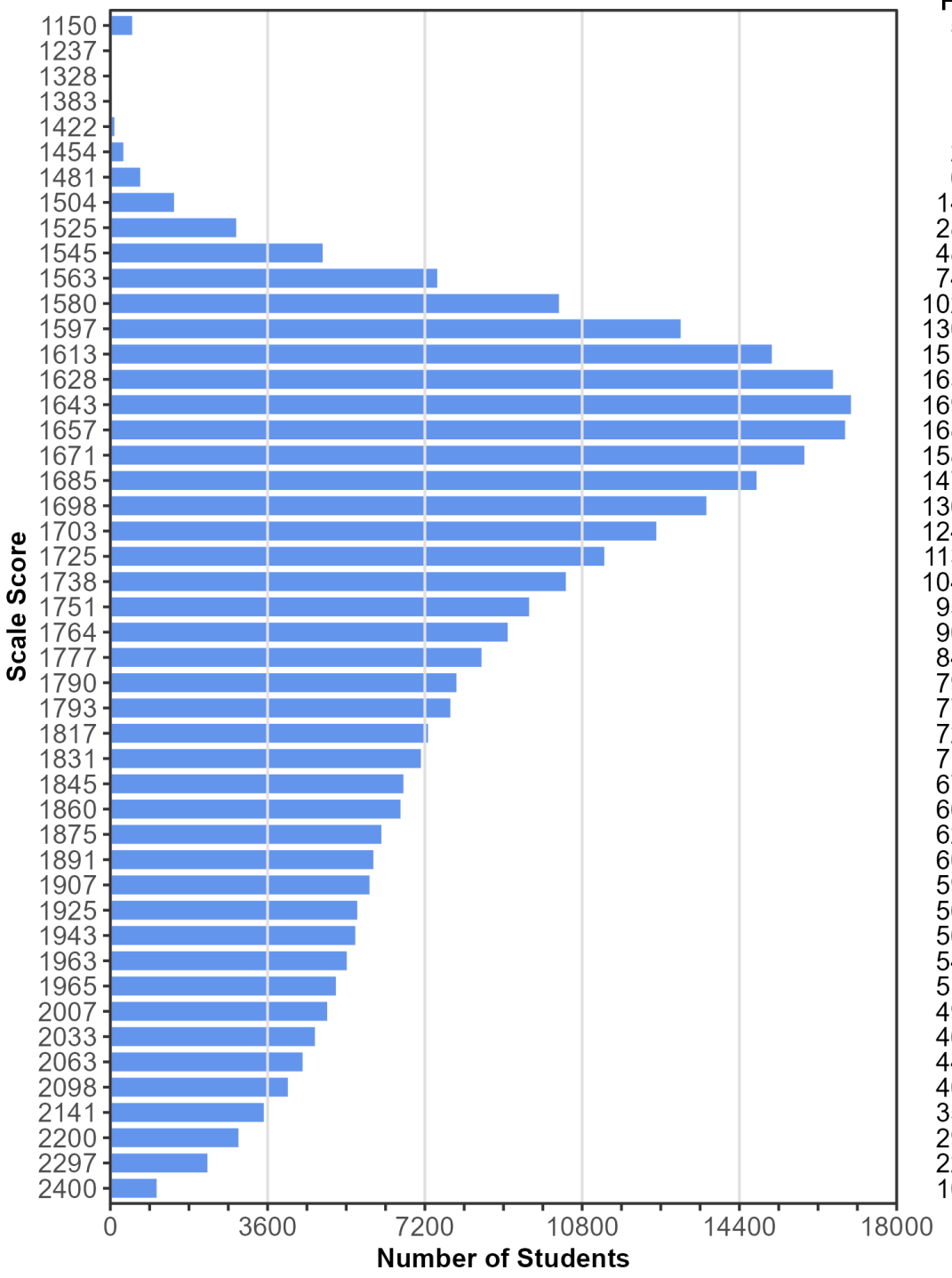
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1000	521	521	0.14	0.14
1087	14	535	0.00	0.14
1182	45	580	0.01	0.15
1240	166	746	0.04	0.20
1283	591	1337	0.16	0.35
1317	1456	2793	0.38	0.73
1345	2762	5555	0.73	1.46
1370	4614	10169	1.21	2.67
1392	6827	16996	1.79	4.47
1412	8452	25448	2.22	6.69
1431	9995	35443	2.63	9.31
1448	10845	46288	2.85	12.16
1464	11322	57610	2.97	15.14
1480	11430	69040	3.00	18.14
1495	11559	80599	3.04	21.18
1509	11681	92280	3.07	24.25
1515	11827	104107	3.11	27.35
1538	11952	116059	3.14	30.49
1552	12442	128501	3.27	33.76
1566	12616	141117	3.31	37.08
1580	12972	154089	3.41	40.49
1594	13359	167448	3.51	44.00
1608	13333	180781	3.50	47.50
1622	13479	194260	3.54	51.04
1634	13724	207984	3.61	54.65
1651	13309	221293	3.50	58.14
1666	13237	234530	3.48	61.62
1681	13098	247628	3.44	65.06
1696	12897	260525	3.39	68.45
1713	12627	273152	3.32	71.77
1729	12091	285243	3.18	74.95
1747	11550	296793	3.03	77.98
1765	11184	307977	2.94	80.92
1776	10473	318450	2.75	83.67
1806	9975	328425	2.62	86.29
1829	9457	337882	2.48	88.78
1854	8897	346779	2.34	91.12
1884	8021	354800	2.11	93.22
1919	7365	362165	1.94	95.16
1962	6292	368457	1.65	96.81
2021	5495	373952	1.44	98.26
2117	4036	377988	1.06	99.32
2200	2601	380589	0.68	100.00

Figure B.7.4. Spring 2024 STAAR Grade 6 Mathematics
Frequency Distribution of Scale Scores



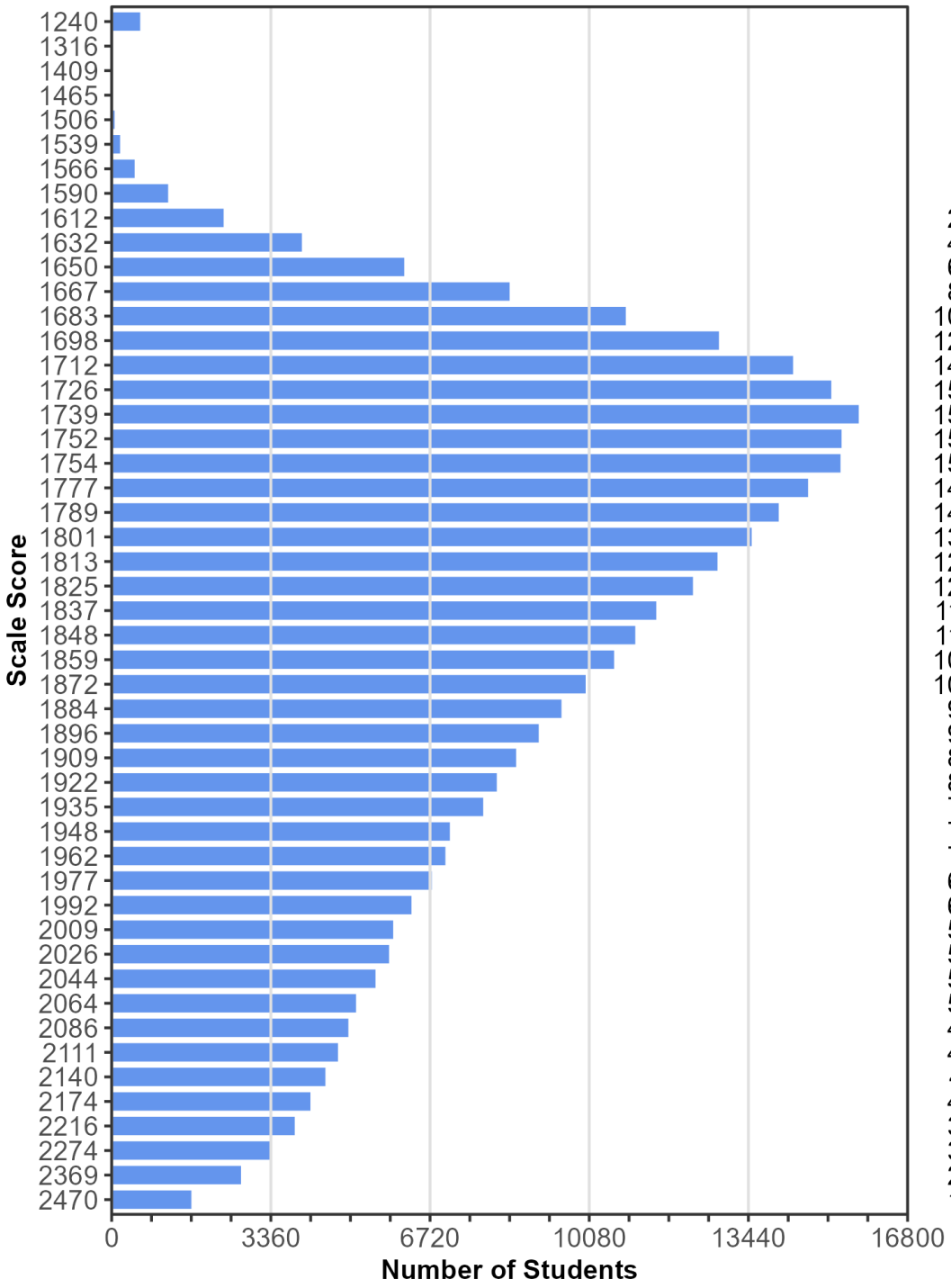
	Freq	Cum Freq	Pct	Cum Pct
1070	507	507	0.13	0.13
1199	14	521	0.00	0.13
1295	25	546	0.01	0.14
1354	154	700	0.04	0.18
1398	454	1154	0.12	0.30
1432	1246	2400	0.32	0.62
1462	2816	5216	0.73	1.35
1487	5280	10496	1.36	2.71
1510	8539	19035	2.20	4.91
1531	11924	30959	3.08	7.99
1550	15145	46104	3.91	11.90
1568	17336	63440	4.47	16.37
1584	18327	81767	4.73	21.10
1600	18566	100333	4.79	25.90
1615	17961	118294	4.64	30.53
1616	16805	135099	4.34	34.87
1644	15981	151080	4.12	38.99
1657	15355	166435	3.96	42.96
1671	14401	180836	3.72	46.67
1684	13706	194542	3.54	50.21
1697	13280	207822	3.43	53.64
1710	12390	220212	3.20	56.84
1723	12050	232262	3.11	59.95
1736	11449	243711	2.95	62.90
1745	10946	254657	2.83	65.73
1762	10460	265117	2.70	68.43
1776	10277	275394	2.65	71.08
1790	9819	285213	2.53	73.61
1804	9487	294700	2.45	76.06
1818	9288	303988	2.40	78.46
1833	8799	312787	2.27	80.73
1849	8564	321351	2.21	82.94
1866	8364	329715	2.16	85.10
1883	8000	337715	2.06	87.16
1889	7614	345329	1.97	89.13
1923	7099	352428	1.83	90.96
1946	6807	359235	1.76	92.72
1971	6207	365442	1.60	94.32
2001	5567	371009	1.44	95.76
2035	4999	376008	1.29	97.05
2079	4112	380120	1.06	98.11
2138	3415	383535	0.88	98.99
2235	2514	386049	0.65	99.64
2350	1406	387455	0.36	100.00

Figure B.7.5. Spring 2024 STAAR Grade 7 Mathematics
Frequency Distribution of Scale Scores



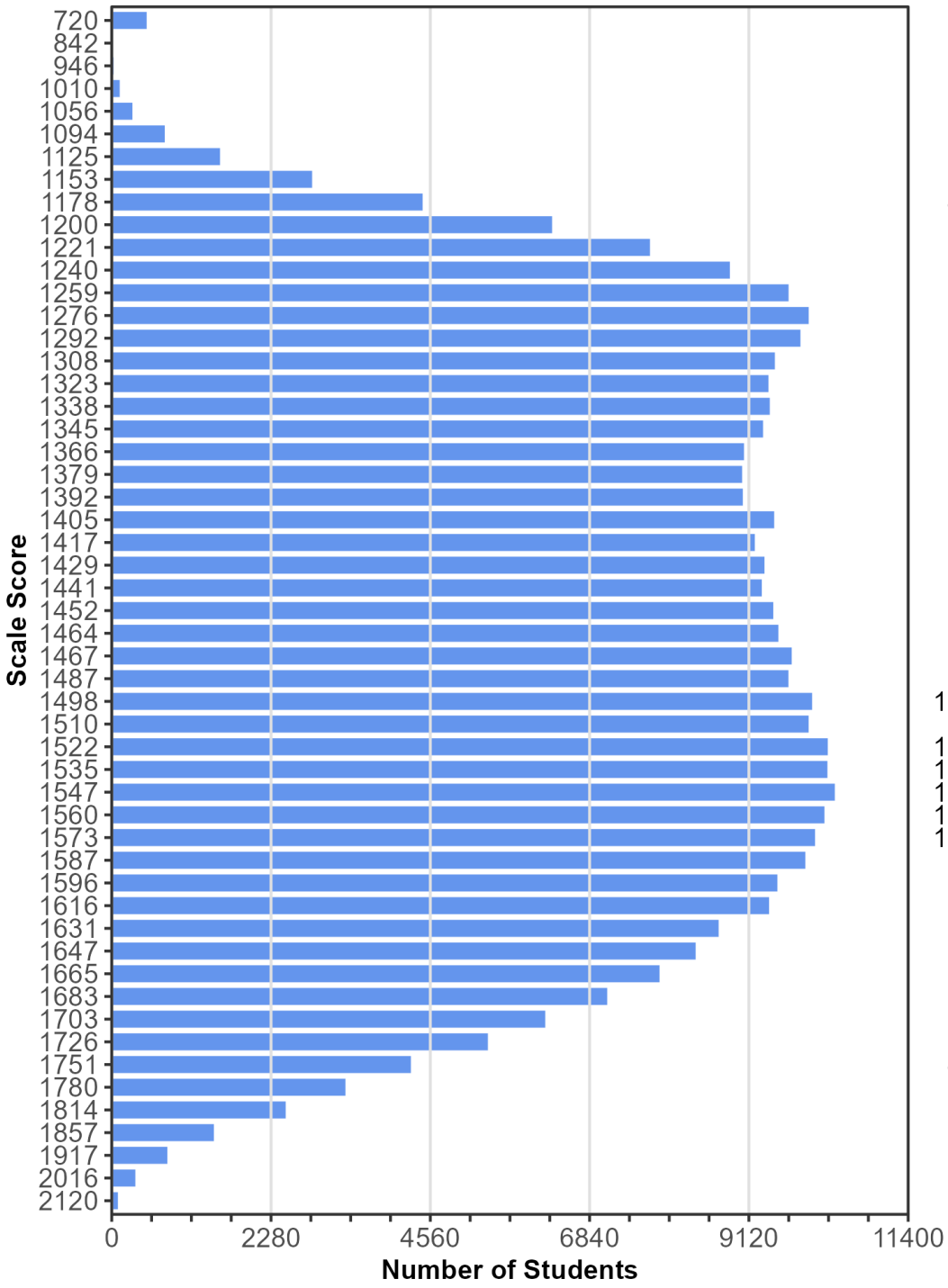
Freq	Cum Freq	Pct	Cum Pct
500	500	0.15	0.15
11	511	0.00	0.16
16	527	0.00	0.16
31	558	0.01	0.17
93	651	0.03	0.20
298	949	0.09	0.29
684	1633	0.21	0.50
1459	3092	0.45	0.95
2881	5973	0.89	1.84
4860	10833	1.50	3.34
7483	18316	2.31	5.65
10270	28586	3.17	8.82
13055	41641	4.03	12.85
15142	56783	4.67	17.52
16542	73325	5.10	22.62
16953	90278	5.23	27.85
16819	107097	5.19	33.04
15887	122984	4.90	37.95
14793	137777	4.56	42.51
13644	151421	4.21	46.72
12498	163919	3.86	50.58
11307	175226	3.49	54.06
10427	185653	3.22	57.28
9586	195239	2.96	60.24
9096	204335	2.81	63.05
8497	212832	2.62	65.67
7923	220755	2.44	68.11
7783	228538	2.40	70.51
7272	235810	2.24	72.76
7109	242919	2.19	74.95
6710	249629	2.07	77.02
6642	256271	2.05	79.07
6205	262476	1.91	80.98
6022	268498	1.86	82.84
5933	274431	1.83	84.67
5652	280083	1.74	86.42
5607	285690	1.73	88.15
5412	291102	1.67	89.82
5164	296266	1.59	91.41
4964	301230	1.53	92.94
4683	305913	1.44	94.39
4402	310315	1.36	95.74
4064	314379	1.25	97.00
3514	317893	1.08	98.08
2933	320826	0.90	98.99
2223	323049	0.69	99.67
1060	324109	0.33	100.00

Figure B.7.6. Spring 2024 STAAR Grade 8 Mathematics
Frequency Distribution of Scale Scores



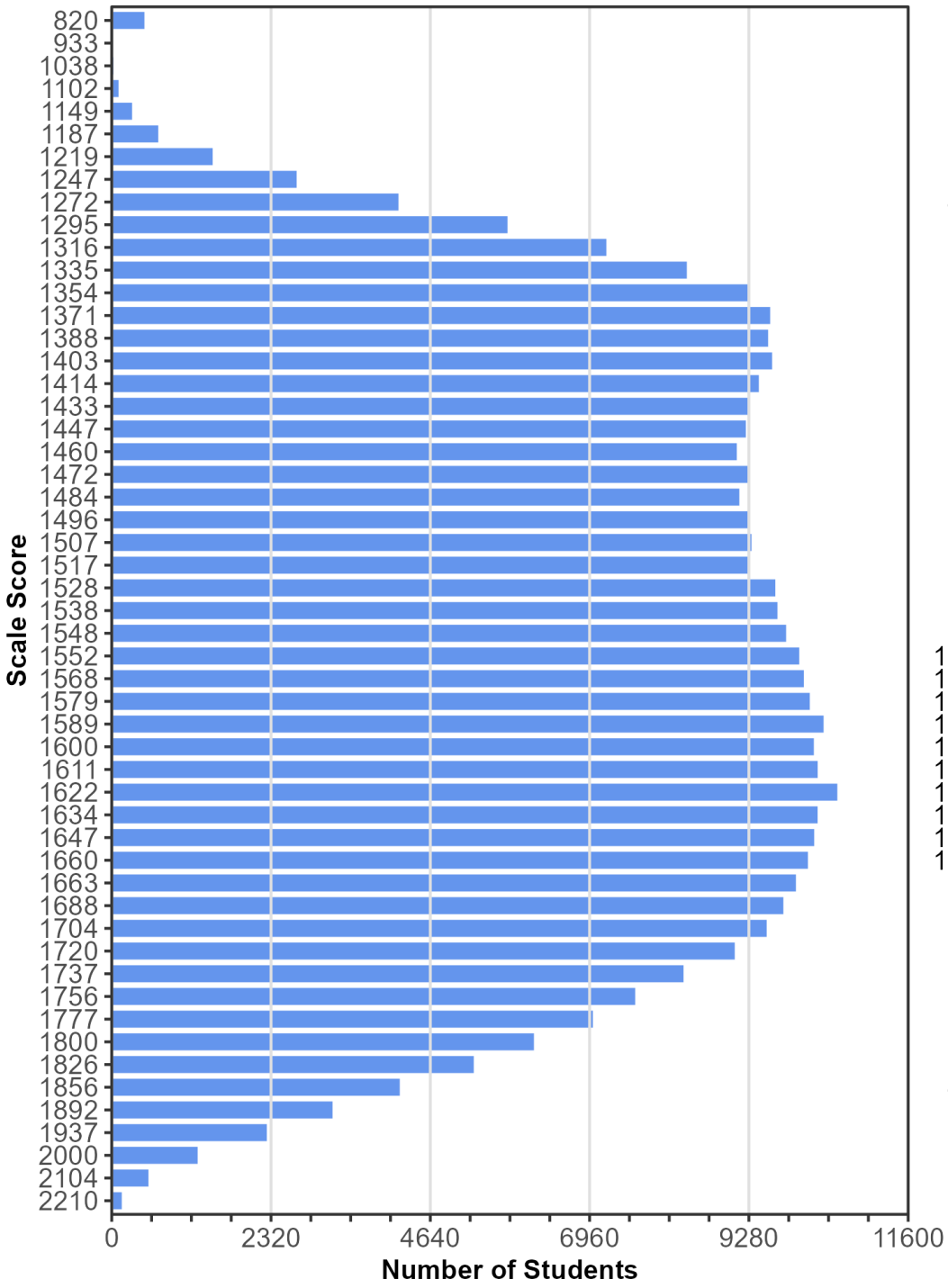
	Freq	Cum Freq	Pct	Cum Pct
1240	602	602	0.17	0.17
1316	16	618	0.00	0.17
1409	12	630	0.00	0.18
1465	18	648	0.01	0.18
1506	61	709	0.02	0.20
1539	178	887	0.05	0.25
1566	485	1372	0.14	0.38
1590	1192	2564	0.33	0.72
1612	2363	4927	0.66	1.38
1632	4015	8942	1.13	2.51
1650	6175	15117	1.73	4.24
1667	8400	23517	2.35	6.59
1683	10852	34369	3.04	9.63
1698	12818	47187	3.59	13.23
1712	14383	61570	4.03	17.26
1726	15189	76759	4.26	21.52
1739	15769	92528	4.42	25.94
1752	15406	107934	4.32	30.26
1754	15384	123318	4.31	34.57
1777	14699	138017	4.12	38.69
1789	14080	152097	3.95	42.64
1801	13509	165606	3.79	46.42
1813	12787	178393	3.58	50.01
1825	12269	190662	3.44	53.45
1837	11495	202157	3.22	56.67
1848	11052	213209	3.10	59.77
1859	10605	223814	2.97	62.74
1872	10005	233819	2.80	65.55
1884	9494	243313	2.66	68.21
1896	9014	252327	2.53	70.73
1909	8537	260864	2.39	73.13
1922	8129	268993	2.28	75.41
1935	7842	276835	2.20	77.60
1948	7138	283973	2.00	79.60
1962	7044	291017	1.97	81.58
1977	6757	297774	1.89	83.47
1992	6328	304102	1.77	85.25
2009	5942	310044	1.67	86.91
2026	5855	315899	1.64	88.55
2044	5569	321468	1.56	90.12
2064	5158	326626	1.45	91.56
2086	4996	331622	1.40	92.96
2111	4777	336399	1.34	94.30
2140	4511	340910	1.26	95.57
2174	4195	345105	1.18	96.74
2216	3864	348969	1.08	97.82
2274	3347	352316	0.94	98.76
2369	2730	355046	0.77	99.53
2470	1682	356728	0.47	100.00

Figure B.7.7. Spring 2024 STAAR Grade 3 RLA
Frequency Distribution of Scale Scores



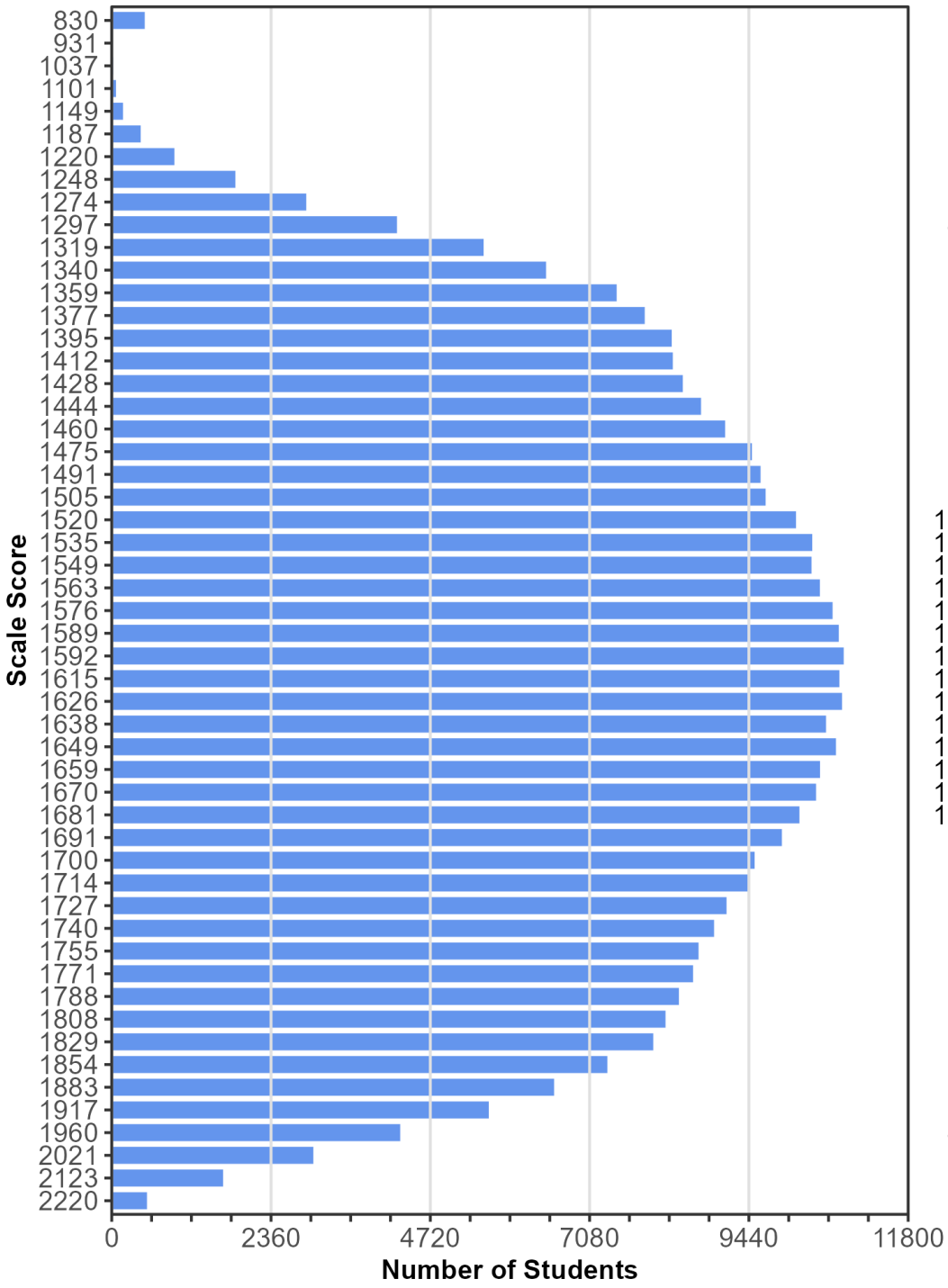
Freq	Cum Freq	Pct	Cum Pct
500	500	0.14	0.14
16	516	0.00	0.14
27	543	0.01	0.15
114	657	0.03	0.18
296	953	0.08	0.26
759	1712	0.21	0.48
1550	3262	0.43	0.91
2868	6130	0.80	1.70
4450	10580	1.24	2.94
6303	16883	1.75	4.69
7704	24587	2.14	6.83
8848	33435	2.46	9.29
9688	43123	2.69	11.99
9976	53099	2.77	14.76
9858	62957	2.74	17.50
9492	72449	2.64	20.14
9400	81849	2.61	22.75
9419	91268	2.62	25.37
9323	100591	2.59	27.96
9050	109641	2.52	30.47
9023	118664	2.51	32.98
9033	127697	2.51	35.49
9482	137179	2.64	38.13
9205	146384	2.56	40.68
9343	155727	2.60	43.28
9305	165032	2.59	45.87
9469	174501	2.63	48.50
9542	184043	2.65	51.15
9733	193776	2.71	53.86
9686	203462	2.69	56.55
10025	213487	2.79	59.33
9975	223462	2.77	62.11
10249	233711	2.85	64.95
10244	243955	2.85	67.80
10350	254305	2.88	70.68
10201	264506	2.84	73.51
10067	274573	2.80	76.31
9928	284501	2.76	79.07
9528	294029	2.65	81.72
9409	303438	2.62	84.33
8687	312125	2.41	86.75
8358	320483	2.32	89.07
7841	328324	2.18	91.25
7092	335416	1.97	93.22
6205	341621	1.72	94.95
5384	347005	1.50	96.44
4282	351287	1.19	97.63
3345	354632	0.93	98.56
2490	357122	0.69	99.25
1462	358584	0.41	99.66
796	359380	0.22	99.88
338	359718	0.09	99.98
88	359806	0.02	100.00

Figure B.7.8. Spring 2024 STAAR Grade 4 RLA
Frequency Distribution of Scale Scores



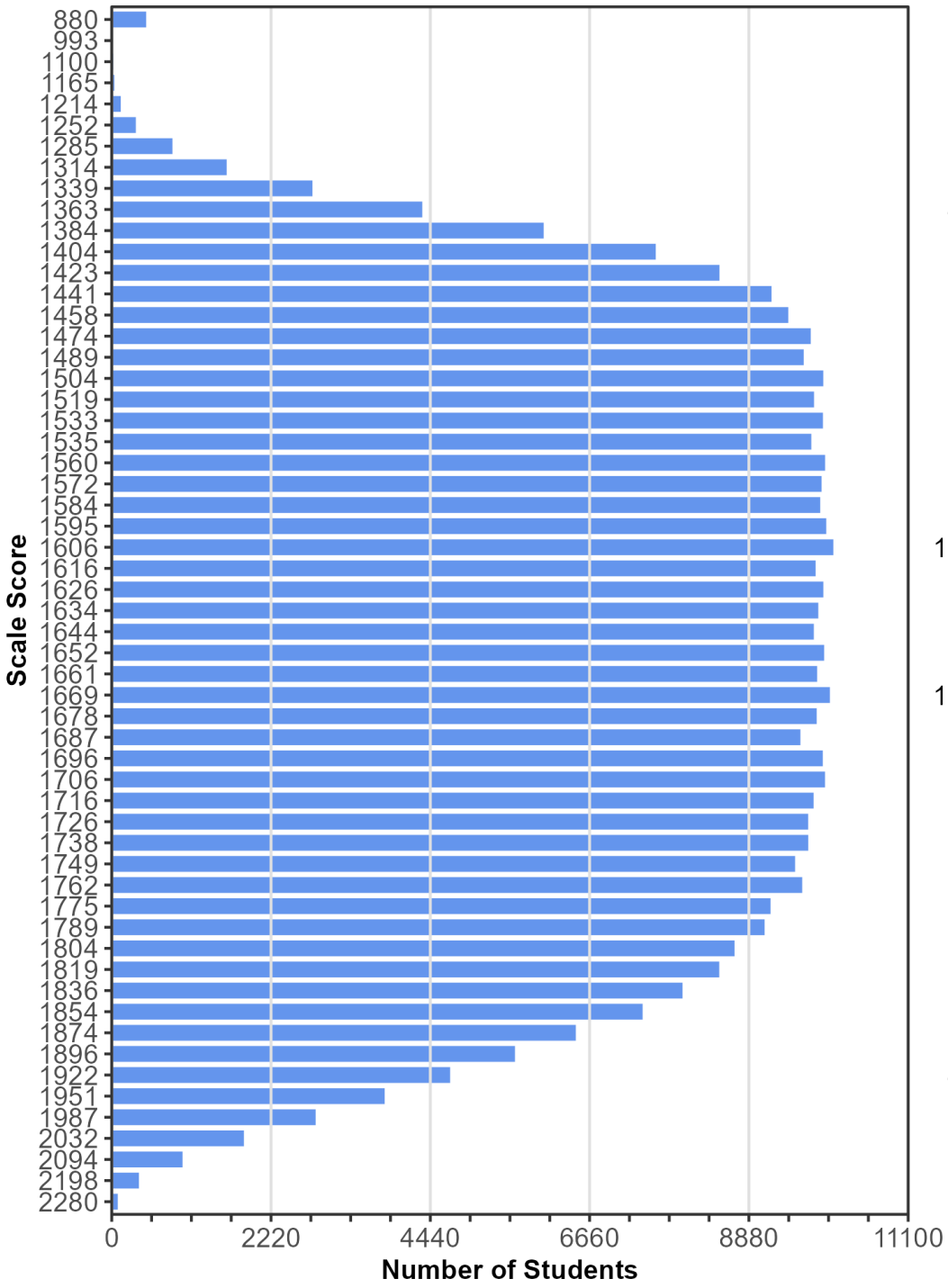
Freq	Cum Freq	Pct	Cum Pct
476	476	0.13	0.13
14	490	0.00	0.13
26	516	0.01	0.14
99	615	0.03	0.17
296	911	0.08	0.25
677	1588	0.18	0.43
1470	3058	0.40	0.83
2693	5751	0.73	1.56
4176	9927	1.13	2.69
5764	15691	1.56	4.26
7204	22895	1.95	6.21
8376	31271	2.27	8.49
9276	40547	2.52	11.00
9590	50137	2.60	13.60
9560	59697	2.59	16.20
9618	69315	2.61	18.81
9425	78740	2.56	21.37
9286	88026	2.52	23.89
9236	97262	2.51	26.39
9104	106366	2.47	28.86
9270	115636	2.52	31.38
9141	124777	2.48	33.86
9289	134066	2.52	36.38
9318	143384	2.53	38.91
9263	152647	2.51	41.42
9665	162312	2.62	44.04
9696	172008	2.63	46.67
9822	181830	2.67	49.34
10013	191843	2.72	52.06
10080	201923	2.74	54.79
10167	212090	2.76	57.55
10368	222458	2.81	60.36
10226	232684	2.77	63.14
10282	242966	2.79	65.93
10568	253534	2.87	68.80
10281	263815	2.79	71.59
10231	274046	2.78	74.36
10140	284186	2.75	77.11
9965	294151	2.70	79.82
9782	303933	2.65	82.47
9539	313472	2.59	85.06
9074	322546	2.46	87.52
8328	330874	2.26	89.78
7624	338498	2.07	91.85
7008	345506	1.90	93.75
6149	351655	1.67	95.42
5273	356928	1.43	96.85
4195	361123	1.14	97.99
3216	364339	0.87	98.86
2259	366598	0.61	99.48
1251	367849	0.34	99.82
534	368383	0.14	99.96
146	368529	0.04	100.00

Figure B.7.9. Spring 2024 STAAR Grade 5 RLA
Frequency Distribution of Scale Scores



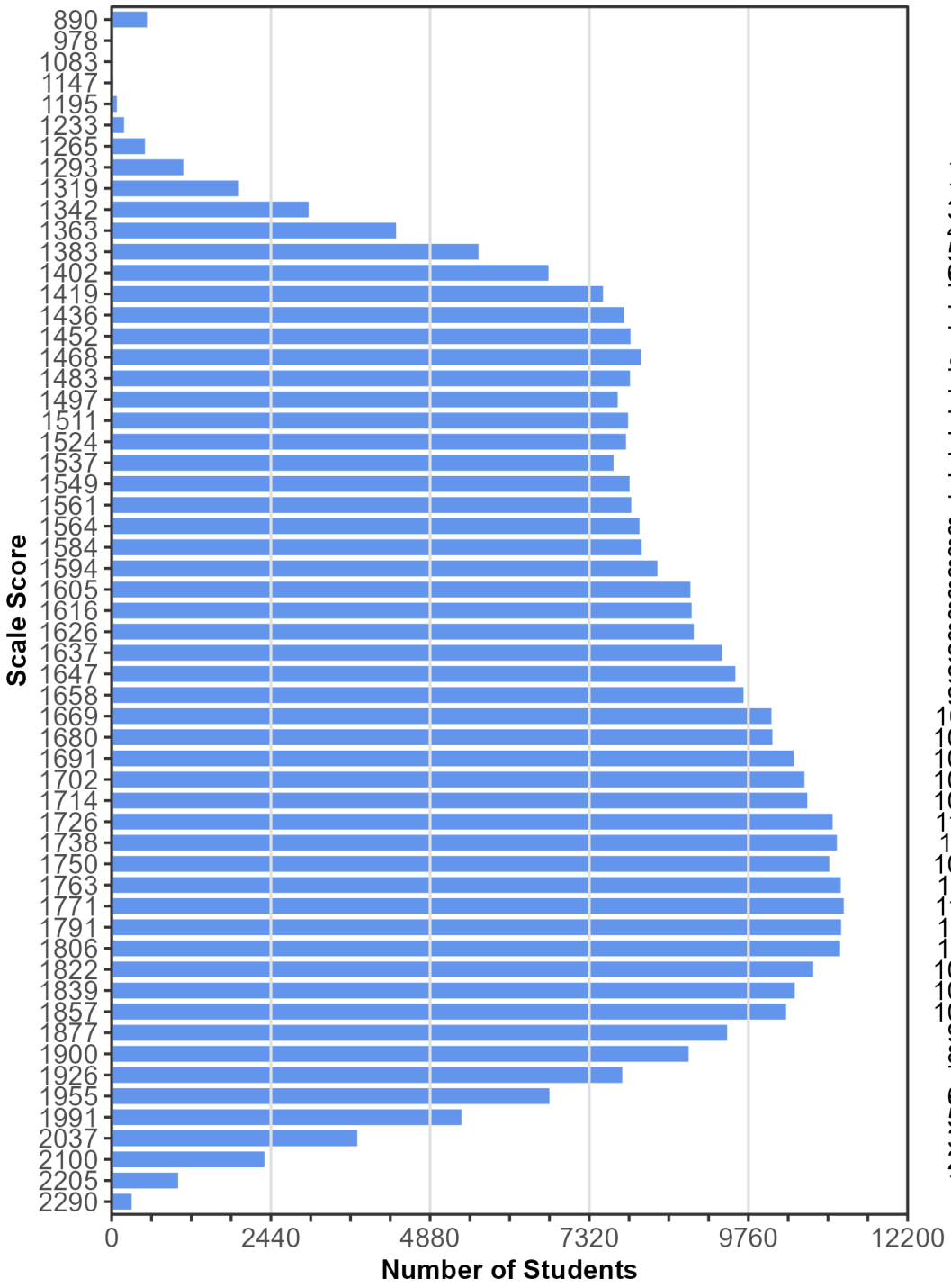
Freq	Cum Freq	Pct	Cum Pct
489	489	0.13	0.13
8	497	0.00	0.13
20	517	0.01	0.14
63	580	0.02	0.15
167	747	0.04	0.20
429	1176	0.11	0.31
928	2104	0.25	0.56
1832	3936	0.49	1.05
2882	6818	0.77	1.82
4226	11044	1.13	2.94
5510	16554	1.47	4.41
6435	22989	1.71	6.12
7481	30470	1.99	8.11
7897	38367	2.10	10.22
8297	46664	2.21	12.43
8313	54977	2.21	14.64
8461	63438	2.25	16.89
8731	72169	2.33	19.22
9088	81257	2.42	21.64
9485	90742	2.53	24.16
9613	100355	2.56	26.72
9688	110043	2.58	29.30
10138	120181	2.70	32.00
10379	130560	2.76	34.77
10368	140928	2.76	37.53
10492	151420	2.79	40.32
10680	162100	2.84	43.17
10772	172872	2.87	46.03
10845	183717	2.89	48.92
10781	194498	2.87	51.79
10820	205318	2.88	54.68
10583	215901	2.82	57.49
10730	226631	2.86	60.35
10494	237125	2.79	63.15
10435	247560	2.78	65.92
10189	257749	2.71	68.64
9929	267678	2.64	71.28
9525	277203	2.54	73.82
9433	286636	2.51	76.33
9109	295745	2.43	78.76
8924	304669	2.38	81.13
8694	313363	2.32	83.45
8613	321976	2.29	85.74
8403	330379	2.24	87.98
8205	338584	2.18	90.16
8023	346607	2.14	92.30
7343	353950	1.96	94.26
6554	360504	1.75	96.00
5586	366090	1.49	97.49
4274	370364	1.14	98.63
2987	373351	0.80	99.42
1650	375001	0.44	99.86
522	375523	0.14	100.00

Figure B.7.10. Spring 2024 STAAR Grade 6 RLA
Frequency Distribution of Scale Scores



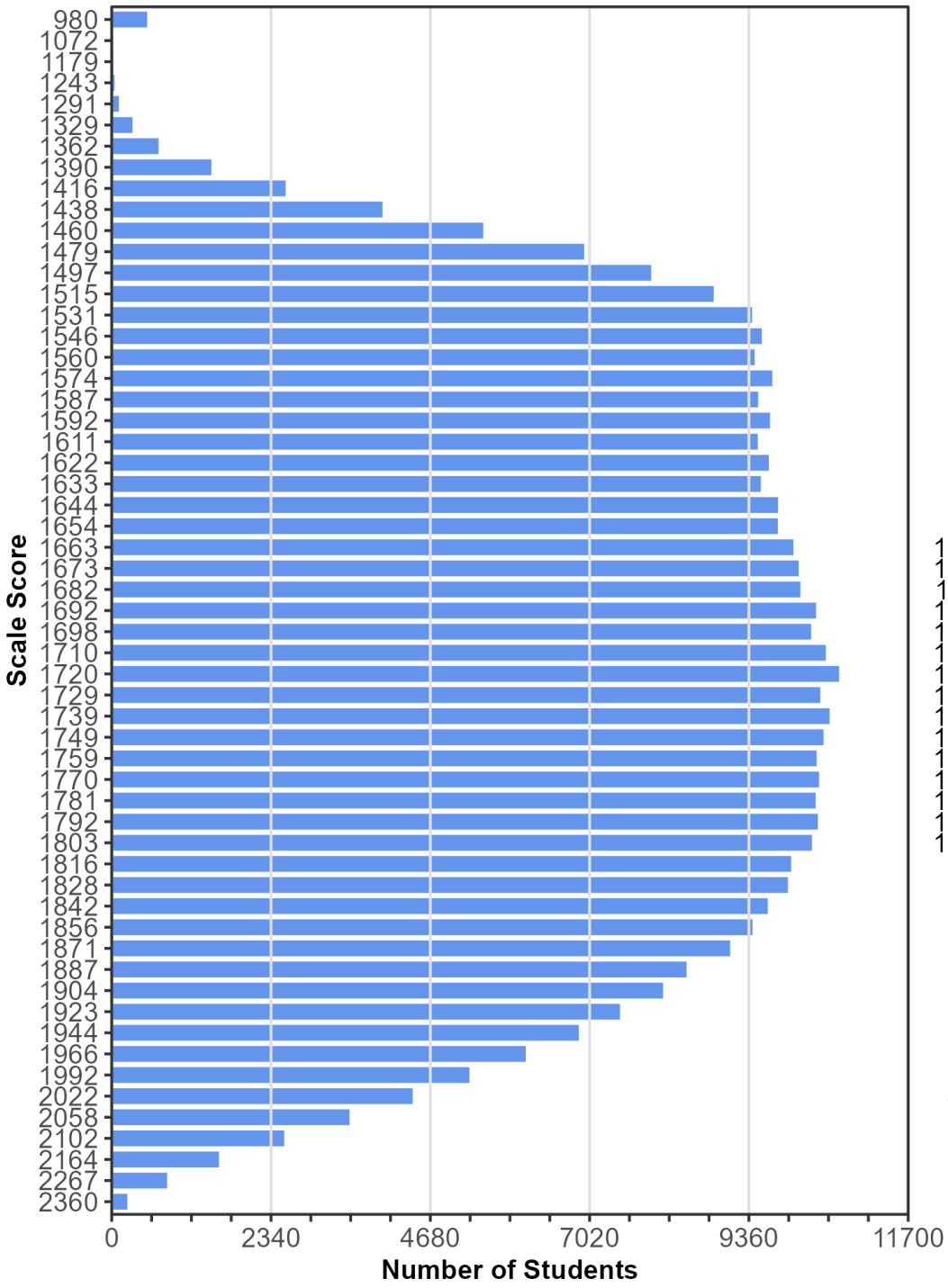
Freq	Cum Freq	Pct	Cum Pct
480	480	0.12	0.12
15	495	0.00	0.13
21	516	0.01	0.13
36	552	0.01	0.14
126	678	0.03	0.17
336	1014	0.09	0.26
846	1860	0.21	0.47
1602	3462	0.41	0.88
2797	6259	0.71	1.59
4328	10587	1.10	2.69
6020	16607	1.53	4.21
7582	24189	1.92	6.14
8470	32659	2.15	8.29
9195	41854	2.33	10.62
9429	51283	2.39	13.02
9742	61025	2.47	15.49
9644	70669	2.45	17.93
9917	80586	2.52	20.45
9788	90374	2.48	22.94
9912	100286	2.52	25.45
9750	110036	2.47	27.93
9942	119978	2.52	30.45
9893	129871	2.51	32.96
9874	139745	2.51	35.47
9958	149703	2.53	37.99
10057	159760	2.55	40.55
9810	169570	2.49	43.03
9918	179488	2.52	45.55
9847	189335	2.50	48.05
9785	199120	2.48	50.53
9929	209049	2.52	53.05
9831	218880	2.49	55.55
10009	228889	2.54	58.09
9825	238714	2.49	60.58
9598	248312	2.44	63.02
9910	258222	2.52	65.53
9941	268163	2.52	68.06
9782	277945	2.48	70.54
9706	287651	2.46	73.00
9707	297358	2.46	75.47
9524	306882	2.42	77.88
9623	316505	2.44	80.33
9182	325687	2.33	82.66
9099	334786	2.31	84.96
8680	343466	2.20	87.17
8467	351933	2.15	89.32
7955	359888	2.02	91.34
7399	367287	1.88	93.21
6468	373755	1.64	94.85
5620	379375	1.43	96.28
4716	384091	1.20	97.48
3804	387895	0.97	98.44
2842	390737	0.72	99.16
1842	392579	0.47	99.63
987	393566	0.25	99.88
379	393945	0.10	99.98
84	394029	0.02	100.00

Figure B.7.11. Spring 2024 STAAR Grade 7 RLA
Frequency Distribution of Scale Scores



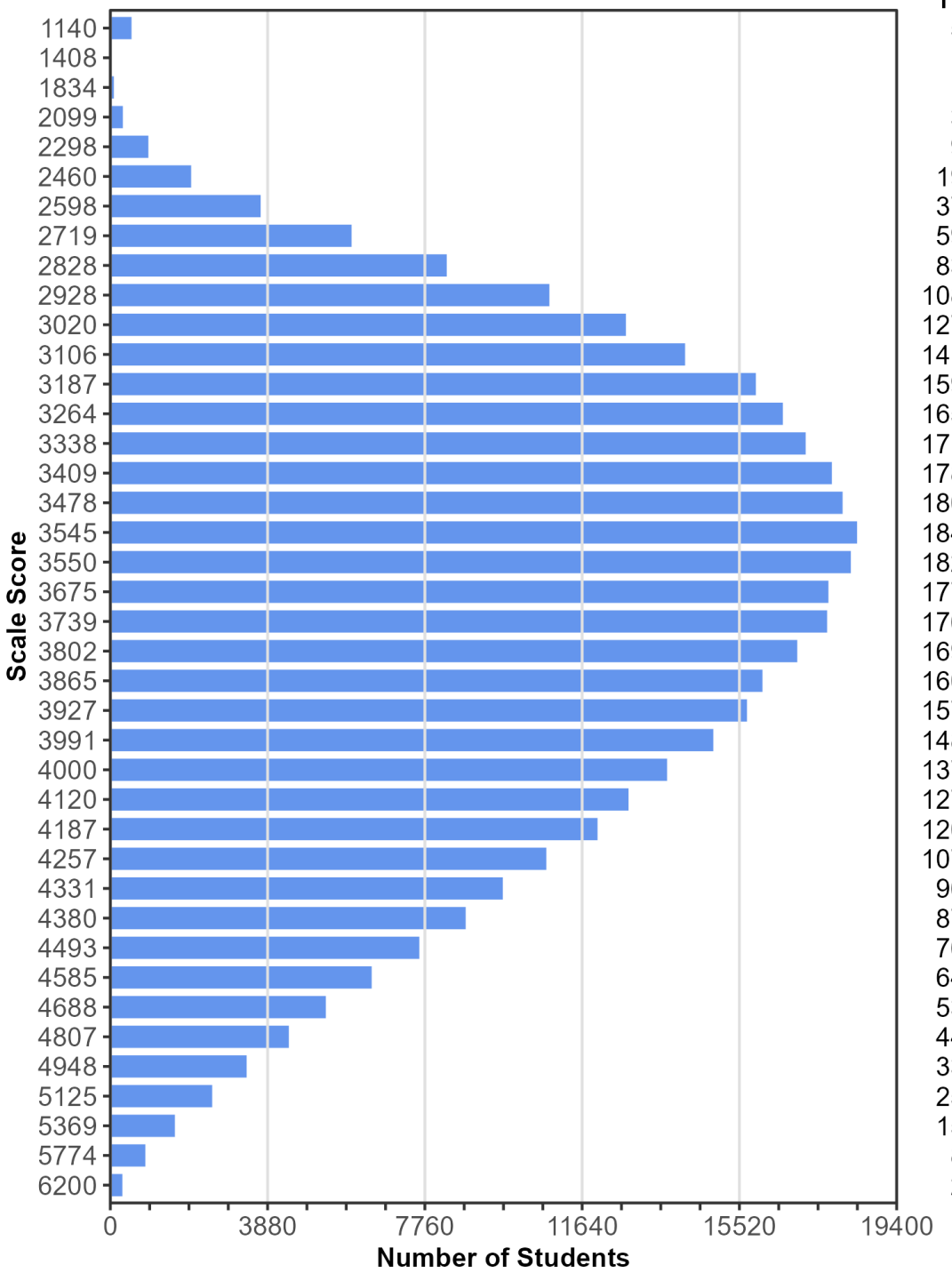
Freq	Cum Freq	Pct	Cum Pct
541	541	0.14	0.14
19	560	0.00	0.14
11	571	0.00	0.14
24	595	0.01	0.15
80	675	0.02	0.17
190	865	0.05	0.22
510	1375	0.13	0.35
1097	2472	0.28	0.62
1950	4422	0.49	1.11
3017	7439	0.76	1.87
4359	11798	1.10	2.97
5623	17421	1.41	4.38
6695	24116	1.68	6.07
7531	31647	1.89	7.96
7853	39500	1.98	9.94
7952	47452	2.00	11.94
8113	55565	2.04	13.98
7945	63510	2.00	15.97
7756	71266	1.95	17.93
7915	79181	1.99	19.92
7884	87065	1.98	21.90
7692	94757	1.93	23.83
7938	102695	2.00	25.83
7964	110659	2.00	27.83
8090	118749	2.03	29.87
8123	126872	2.04	31.91
8365	135237	2.10	34.02
8870	144107	2.23	36.25
8889	152996	2.24	38.48
8922	161918	2.24	40.73
9356	171274	2.35	43.08
9560	180834	2.40	45.48
9685	190519	2.44	47.92
10113	200632	2.54	50.46
10128	210760	2.55	53.01
10455	221215	2.63	55.64
10619	231834	2.67	58.31
10662	242496	2.68	60.99
11051	253547	2.78	63.77
11116	264663	2.80	66.57
10999	275662	2.77	69.34
11175	286837	2.81	72.15
11221	298058	2.82	74.97
11181	309239	2.81	77.78
11165	320404	2.81	80.59
10754	331158	2.70	83.30
10470	341628	2.63	85.93
10338	351966	2.60	88.53
9434	361400	2.37	90.90
8843	370243	2.22	93.13
7826	378069	1.97	95.10
6710	384779	1.69	96.78
5362	390141	1.35	98.13
3764	393905	0.95	99.08
2340	396245	0.59	99.67
1018	397263	0.26	99.92
306	397569	0.08	100.00

Figure B.7.12. Spring 2024 STAAR Grade 8 RLA
Frequency Distribution of Scale Scores



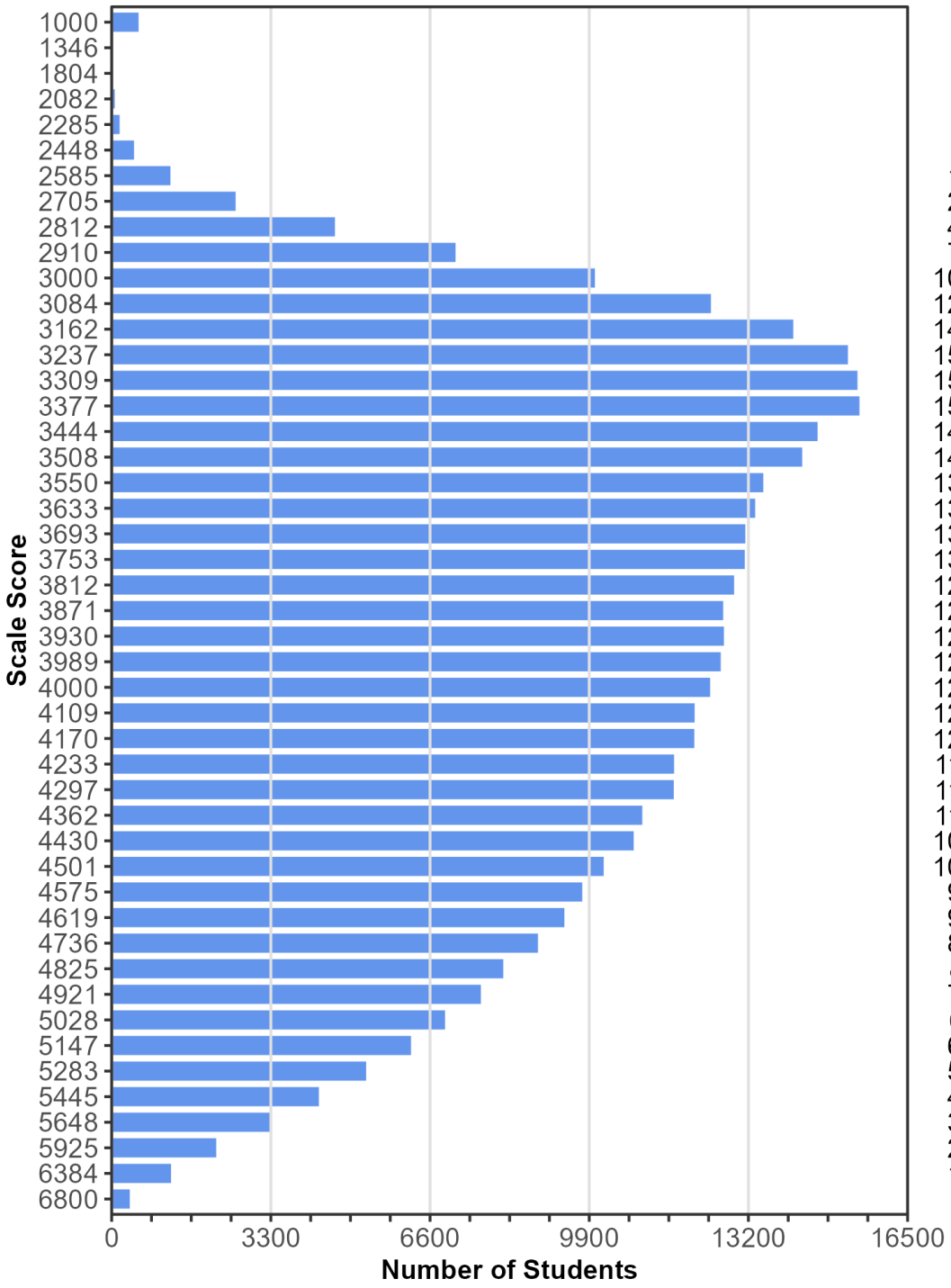
Freq	Cum Freq	Pct	Cum Pct
520	520	0.13	0.13
11	531	0.00	0.13
21	552	0.01	0.14
41	593	0.01	0.15
105	698	0.03	0.17
305	1003	0.08	0.25
688	1691	0.17	0.42
1463	3154	0.36	0.78
2555	5709	0.63	1.42
3977	9686	0.99	2.40
5458	15144	1.35	3.76
6940	22084	1.72	5.48
7925	30009	1.97	7.45
8843	38852	2.19	9.64
9406	48258	2.33	11.98
9550	57808	2.37	14.35
9445	67253	2.34	16.69
9704	76957	2.41	19.10
9497	86454	2.36	21.46
9670	96124	2.40	23.86
9491	105615	2.36	26.21
9653	115268	2.40	28.61
9535	124803	2.37	30.98
9789	134592	2.43	33.41
9786	144378	2.43	35.84
10014	154392	2.49	38.32
10092	164484	2.50	40.83
10116	174600	2.51	43.34
10347	184947	2.57	45.91
10274	195221	2.55	48.46
10490	205711	2.60	51.06
10686	216397	2.65	53.71
10409	226806	2.58	56.30
10546	237352	2.62	58.91
10457	247809	2.60	61.51
10355	258164	2.57	64.08
10391	268555	2.58	66.66
10341	278896	2.57	69.23
10374	289270	2.57	71.80
10287	299557	2.55	74.35
9981	309538	2.48	76.83
9933	319471	2.47	79.30
9638	329109	2.39	81.69
9413	338522	2.34	84.02
9085	347607	2.25	86.28
8444	356051	2.10	88.38
8100	364151	2.01	90.39
7467	371618	1.85	92.24
6861	378479	1.70	93.94
6082	384561	1.51	95.45
5255	389816	1.30	96.76
4422	394238	1.10	97.85
3494	397732	0.87	98.72
2534	400266	0.63	99.35
1575	401841	0.39	99.74
813	402654	0.20	99.94
229	402883	0.06	100.00

Figure B.7.13. Spring 2024 STAAR Grade 5 Science
Frequency Distribution of Scale Scores



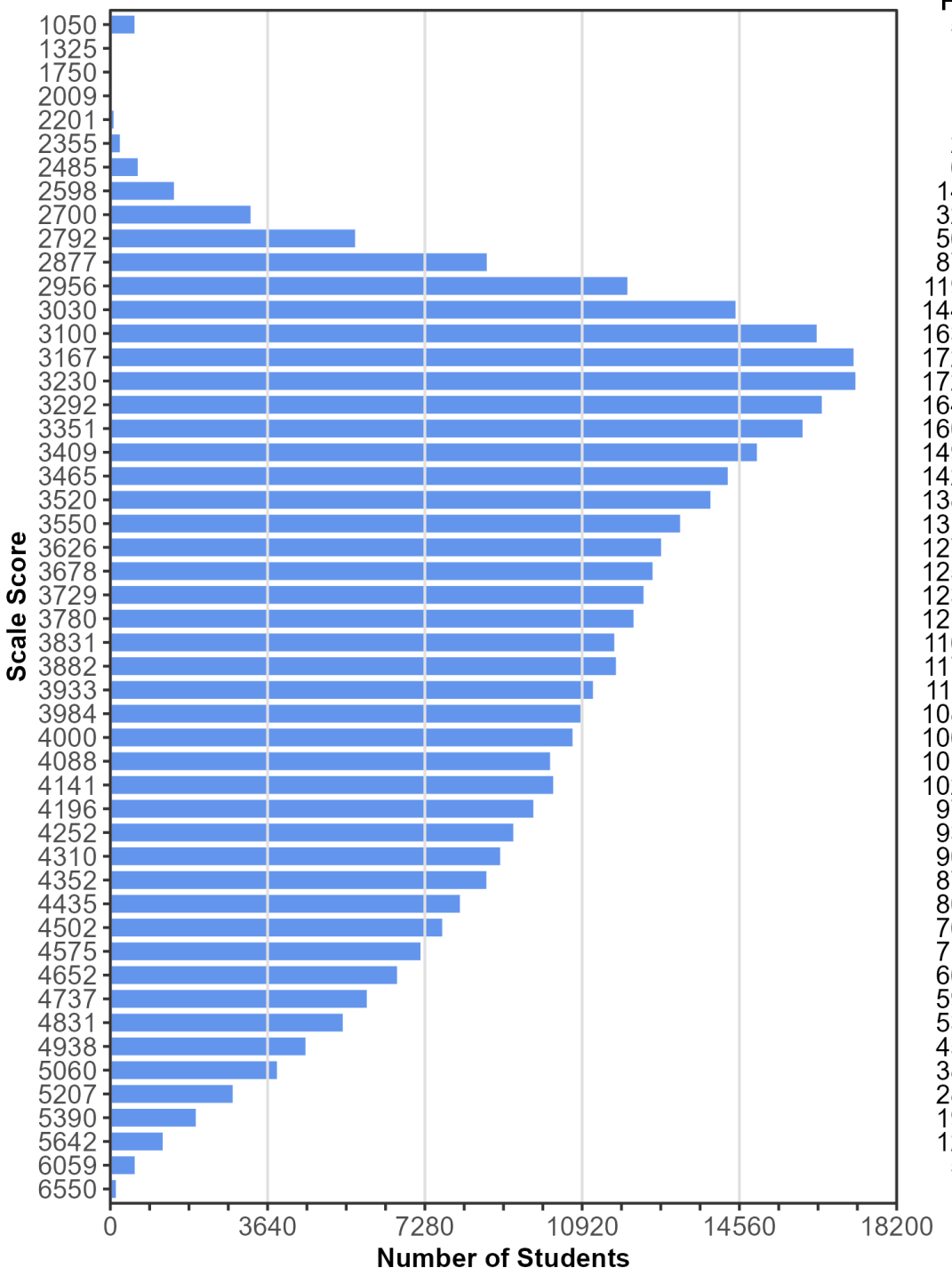
	Cum		Cum
Freq	Freq	Pct	Pct
523	523	0.14	0.14
21	544	0.01	0.14
87	631	0.02	0.17
309	940	0.08	0.25
938	1878	0.25	0.49
1994	3872	0.52	1.02
3709	7581	0.97	1.99
5952	13533	1.56	3.55
8300	21833	2.18	5.73
10832	32665	2.84	8.57
12720	45385	3.34	11.91
14180	59565	3.72	15.63
15926	75491	4.18	19.81
16592	92083	4.36	24.17
17155	109238	4.50	28.67
17801	127039	4.67	33.34
18066	145105	4.74	38.09
18423	163528	4.84	42.92
18269	181797	4.80	47.72
17716	199513	4.65	52.37
17684	217197	4.64	57.01
16947	234144	4.45	61.46
16089	250233	4.22	65.68
15707	265940	4.12	69.80
14876	280816	3.90	73.71
13736	294552	3.61	77.31
12784	307336	3.36	80.67
12020	319356	3.15	83.82
10756	330112	2.82	86.65
9683	339795	2.54	89.19
8767	348562	2.30	91.49
7623	356185	2.00	93.49
6448	362633	1.69	95.18
5317	367950	1.40	96.58
4404	372354	1.16	97.73
3363	375717	0.88	98.62
2513	378230	0.66	99.28
1593	379823	0.42	99.70
864	380687	0.23	99.92
297	380984	0.08	100.00

Figure B.7.14. Spring 2024 STAAR Grade 8 Science
Frequency Distribution of Scale Scores



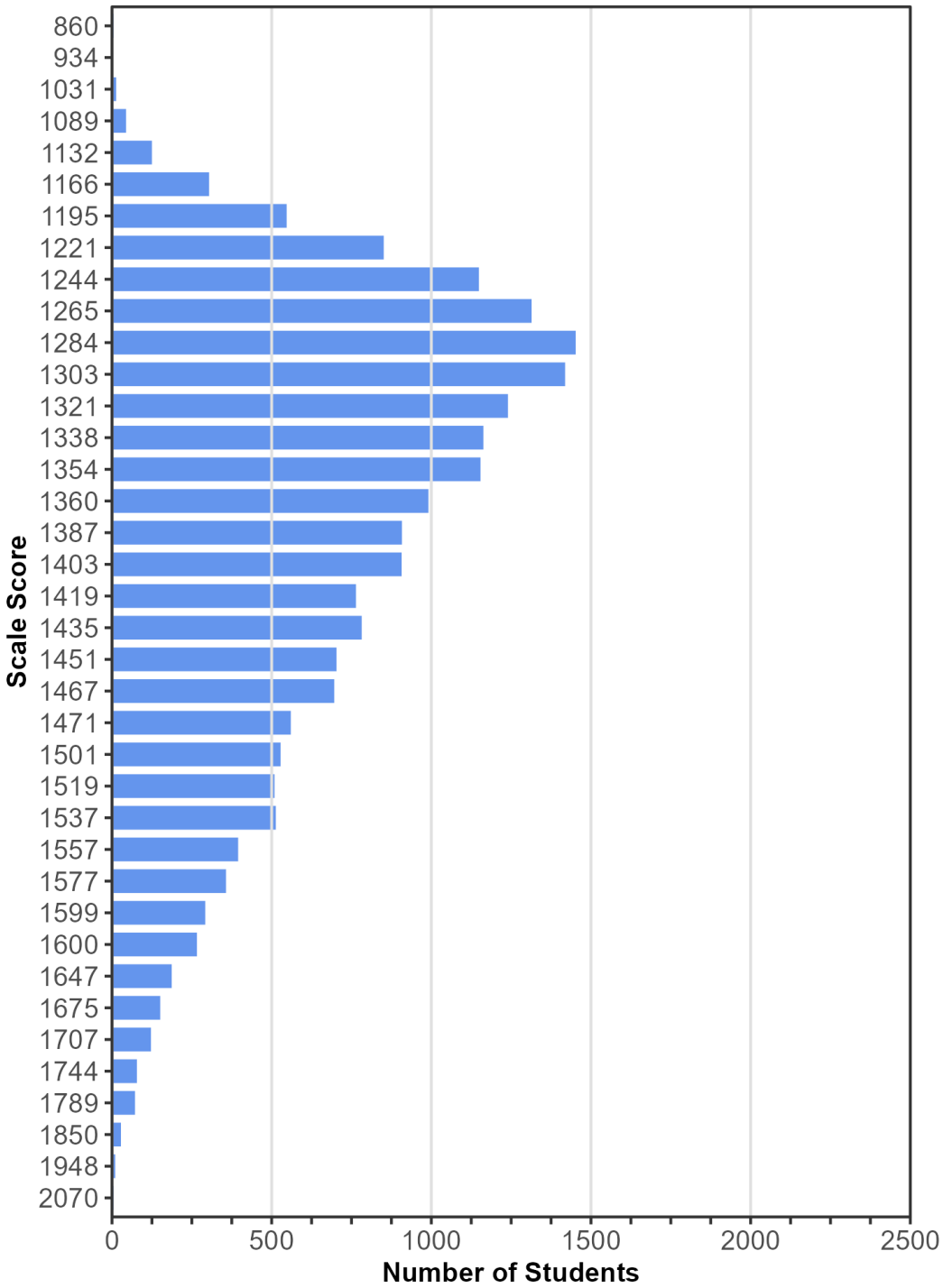
Freq	Cum Freq	Pct	Cum Pct
558	558	0.14	0.14
14	572	0.00	0.14
20	592	0.01	0.15
64	656	0.02	0.16
165	821	0.04	0.21
463	1284	0.12	0.32
1218	2502	0.31	0.63
2570	5072	0.65	1.27
4629	9701	1.16	2.44
7129	16830	1.79	4.23
10018	26848	2.52	6.75
12423	39271	3.12	9.87
14130	53401	3.55	13.42
15263	68664	3.83	17.25
15461	84125	3.88	21.14
15502	99627	3.89	25.03
14634	114261	3.68	28.71
14315	128576	3.60	32.31
13509	142085	3.39	35.70
13340	155425	3.35	39.05
13136	168561	3.30	42.35
13125	181686	3.30	45.65
12903	194589	3.24	48.89
12677	207266	3.19	52.08
12692	219958	3.19	55.27
12626	232584	3.17	58.44
12407	244991	3.12	61.55
12085	257076	3.04	64.59
12078	269154	3.03	67.63
11659	280813	2.93	70.56
11653	292466	2.93	73.48
11000	303466	2.76	76.25
10822	314288	2.72	78.97
10200	324488	2.56	81.53
9755	334243	2.45	83.98
9385	343628	2.36	86.34
8837	352465	2.22	88.56
8118	360583	2.04	90.60
7654	368237	1.92	92.52
6911	375148	1.74	94.26
6205	381353	1.56	95.82
5275	386628	1.33	97.14
4296	390924	1.08	98.22
3305	394229	0.83	99.05
2168	396397	0.54	99.60
1231	397628	0.31	99.91
376	398004	0.09	100.00

Figure B.7.15. Spring 2024 STAAR Grade 8 Social Studies
Frequency Distribution of Scale Scores



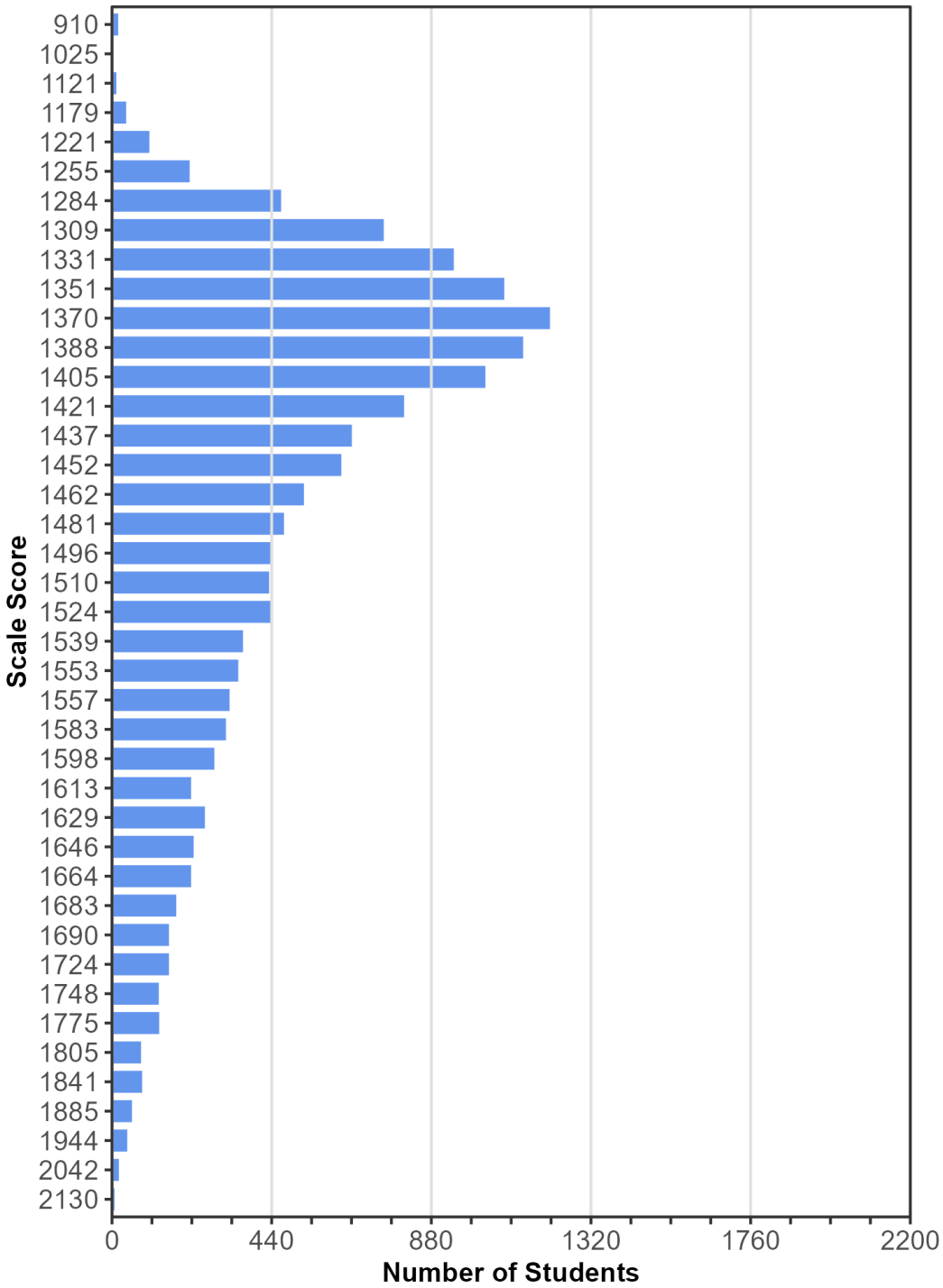
Freq	Cum Freq	Pct	Cum Pct
560	560	0.14	0.14
12	572	0.00	0.14
10	582	0.00	0.14
18	600	0.00	0.15
75	675	0.02	0.17
220	895	0.05	0.22
634	1529	0.16	0.38
1473	3002	0.36	0.74
3246	6248	0.80	1.54
5665	11913	1.40	2.94
8713	20626	2.15	5.08
11968	32594	2.95	8.03
14471	47065	3.57	11.60
16348	63413	4.03	15.63
17202	80615	4.24	19.87
17245	97860	4.25	24.12
16467	114327	4.06	28.17
16021	130348	3.95	32.12
14965	145313	3.69	35.81
14290	159603	3.52	39.33
13889	173492	3.42	42.75
13187	186679	3.25	46.00
12747	199426	3.14	49.14
12550	211976	3.09	52.24
12341	224317	3.04	55.28
12109	236426	2.98	58.26
11664	248090	2.87	61.14
11702	259792	2.88	64.02
11170	270962	2.75	66.77
10881	281843	2.68	69.45
10698	292541	2.64	72.09
10178	302719	2.51	74.60
10250	312969	2.53	77.12
9790	322759	2.41	79.54
9326	332085	2.30	81.83
9023	341108	2.22	84.06
8704	349812	2.14	86.20
8091	357903	1.99	88.20
7681	365584	1.89	90.09
7179	372763	1.77	91.86
6636	379399	1.64	93.49
5936	385335	1.46	94.96
5380	390715	1.33	96.28
4517	395232	1.11	97.40
3856	399088	0.95	98.35
2832	401920	0.70	99.04
1978	403898	0.49	99.53
1212	405110	0.30	99.83
564	405674	0.14	99.97
128	405802	0.03	100.00

Figure B.7.16. Spring 2024 STAAR Spanish Grade 3 Mathematics
Frequency Distribution of Scale Scores



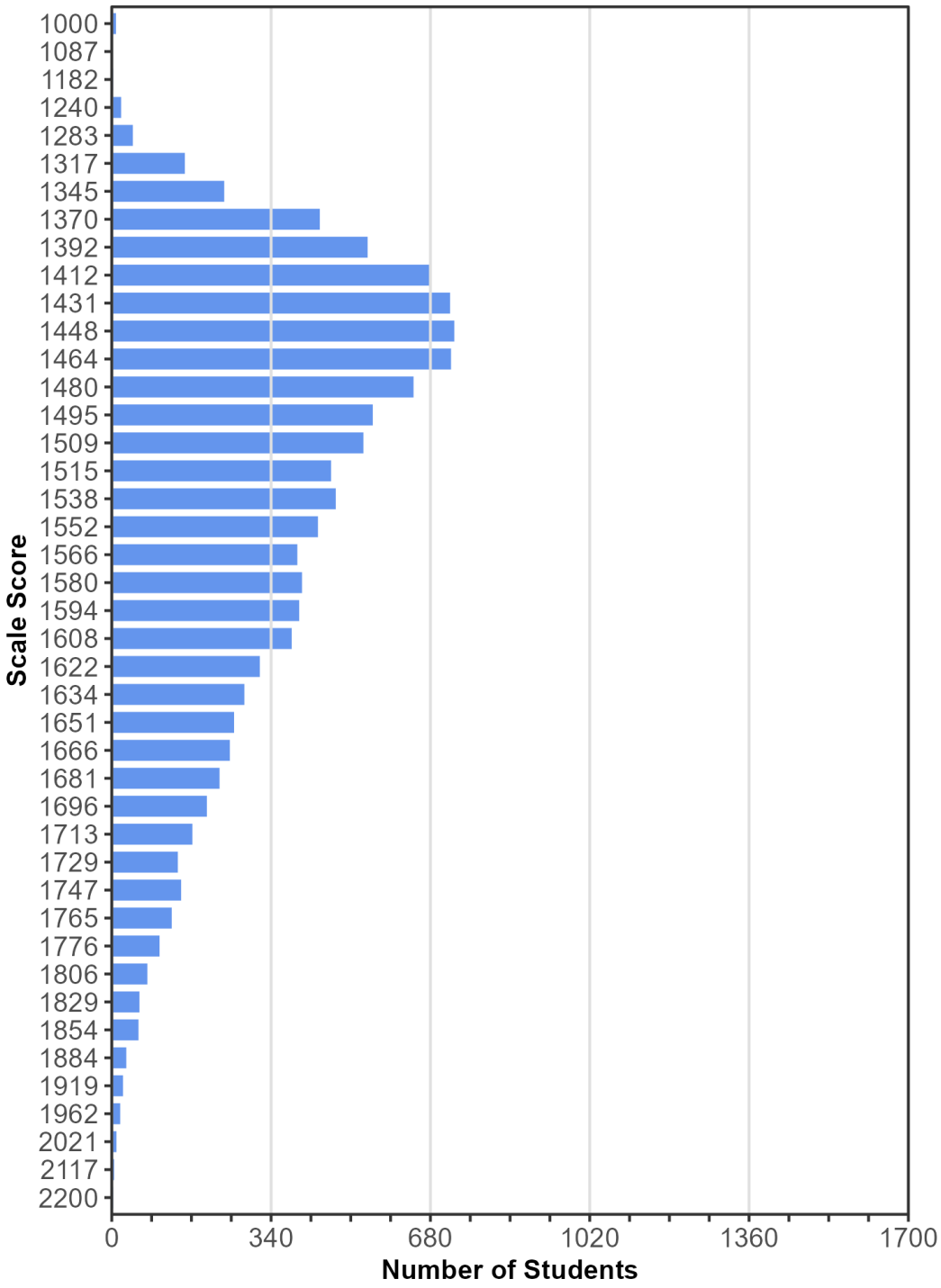
	Cum		Cum
Freq	Freq	Pct	Pct
5	5	0.02	0.02
1	6	0.00	0.03
13	19	0.06	0.09
44	63	0.21	0.31
125	188	0.61	0.91
304	492	1.48	2.39
547	1039	2.65	5.04
851	1890	4.13	9.17
1149	3039	5.58	14.75
1314	4353	6.38	21.13
1452	5805	7.05	28.17
1419	7224	6.89	35.06
1240	8464	6.02	41.08
1163	9627	5.64	46.72
1154	10781	5.60	52.32
991	11772	4.81	57.13
908	12680	4.41	61.54
907	13587	4.40	65.94
764	14351	3.71	69.65
782	15133	3.80	73.44
703	15836	3.41	76.86
696	16532	3.38	80.23
560	17092	2.72	82.95
528	17620	2.56	85.51
509	18129	2.47	87.98
513	18642	2.49	90.47
395	19037	1.92	92.39
357	19394	1.73	94.12
292	19686	1.42	95.54
266	19952	1.29	96.83
187	20139	0.91	97.74
151	20290	0.73	98.47
122	20412	0.59	99.06
78	20490	0.38	99.44
72	20562	0.35	99.79
28	20590	0.14	99.93
10	20600	0.05	99.98
5	20605	0.02	100.00

Figure B.7.17. Spring 2024 STAAR Spanish Grade 4 Mathematics
Frequency Distribution of Scale Scores



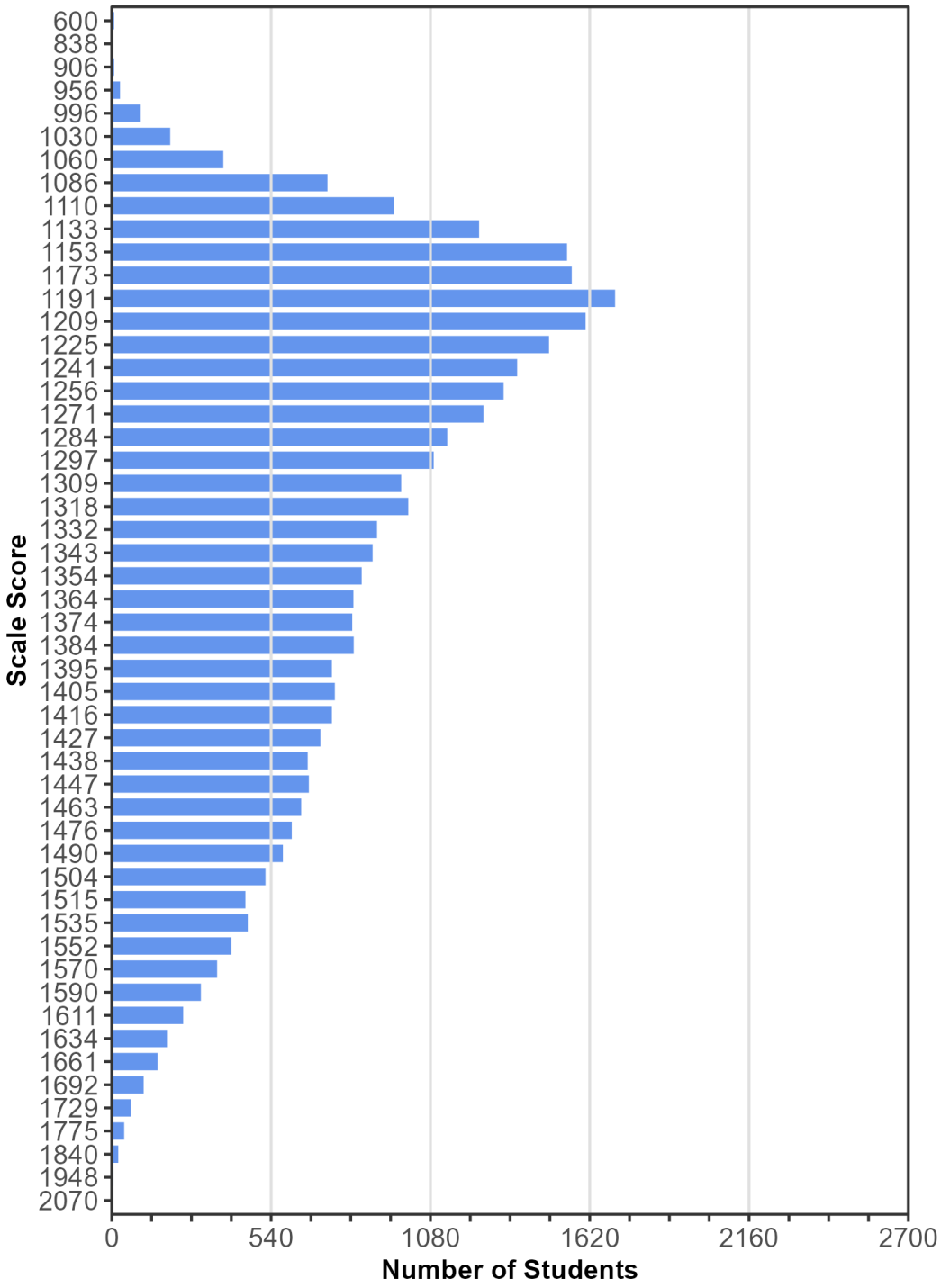
	Freq	Cum Freq	Pct	Cum Pct
910	17	17	0.11	0.11
1025	2	19	0.01	0.13
1121	12	31	0.08	0.21
1179	39	70	0.26	0.47
1221	103	173	0.69	1.15
1255	214	387	1.43	2.58
1284	466	853	3.11	5.69
1309	749	1602	5.00	10.69
1331	942	2544	6.28	16.97
1351	1081	3625	7.21	24.18
1370	1207	4832	8.05	32.23
1388	1133	5965	7.56	39.79
1405	1029	6994	6.86	46.66
1421	805	7799	5.37	52.03
1437	661	8460	4.41	56.44
1452	632	9092	4.22	60.65
1462	529	9621	3.53	64.18
1481	474	10095	3.16	67.34
1496	440	10535	2.94	70.28
1510	433	10968	2.89	73.17
1524	440	11408	2.94	76.10
1539	361	11769	2.41	78.51
1553	348	12117	2.32	80.83
1557	324	12441	2.16	83.00
1583	314	12755	2.09	85.09
1598	282	13037	1.88	86.97
1613	218	13255	1.45	88.43
1629	256	13511	1.71	90.13
1646	225	13736	1.50	91.63
1664	218	13954	1.45	93.09
1683	177	14131	1.18	94.27
1690	157	14288	1.05	95.32
1724	157	14445	1.05	96.36
1748	129	14574	0.86	97.22
1775	130	14704	0.87	98.09
1805	80	14784	0.53	98.63
1841	83	14867	0.55	99.18
1885	55	14922	0.37	99.55
1944	42	14964	0.28	99.83
2042	19	14983	0.13	99.95
2130	7	14990	0.05	100.00

Figure B.7.18. Spring 2024 STAAR Spanish Grade 5 Mathematics
Frequency Distribution of Scale Scores



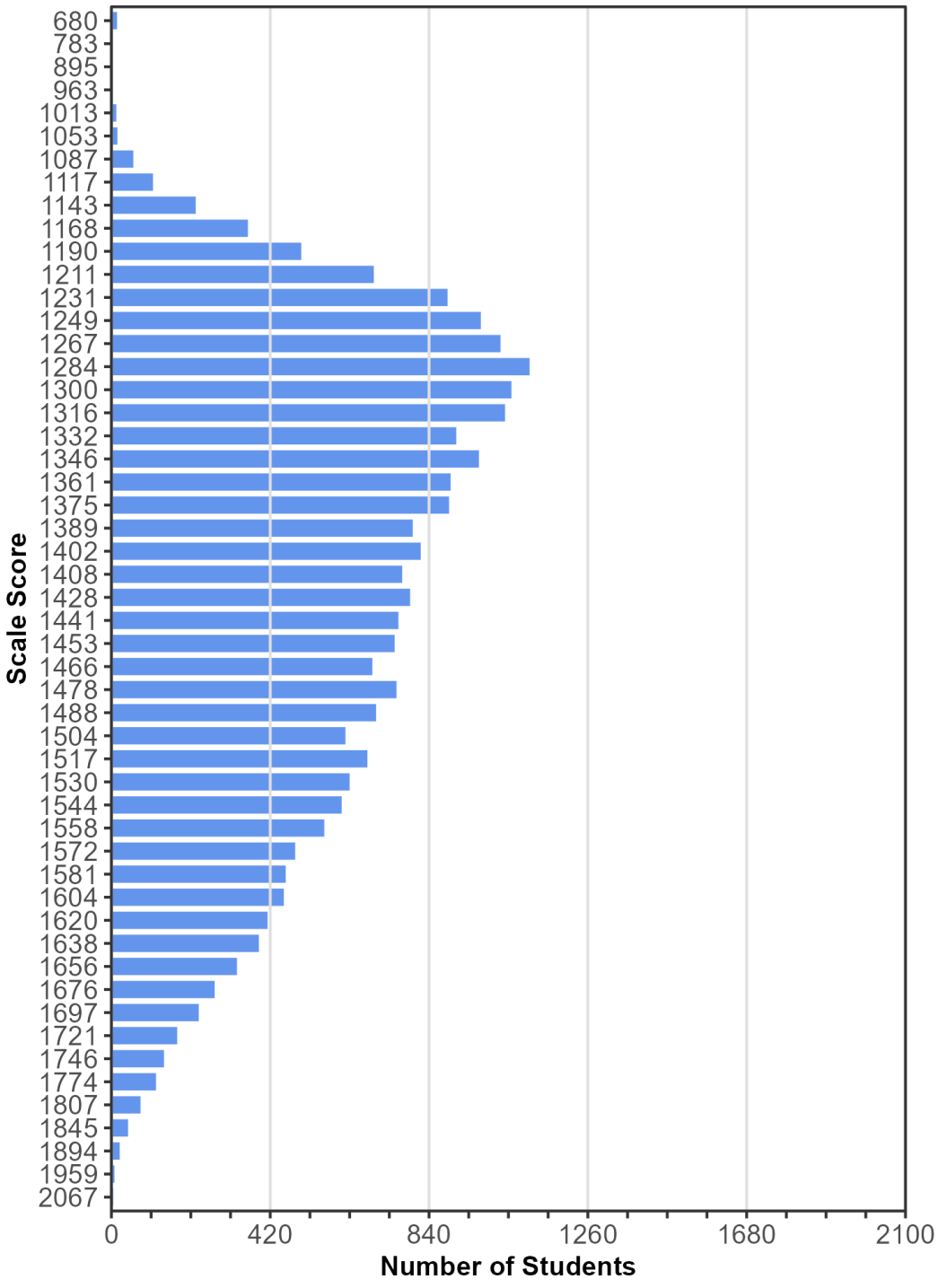
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1000	9	9	0.08	0.08
1087	2	11	0.02	0.10
1182	3	14	0.03	0.12
1240	20	34	0.17	0.29
1283	45	79	0.39	0.68
1317	156	235	1.35	2.03
1345	240	475	2.08	4.11
1370	444	919	3.84	7.96
1392	546	1465	4.73	12.68
1412	681	2146	5.90	18.58
1431	722	2868	6.25	24.83
1448	731	3599	6.33	31.15
1464	724	4323	6.27	37.42
1480	644	4967	5.57	43.00
1495	557	5524	4.82	47.82
1509	537	6061	4.65	52.47
1515	468	6529	4.05	56.52
1538	478	7007	4.14	60.66
1552	440	7447	3.81	64.47
1566	396	7843	3.43	67.89
1580	406	8249	3.51	71.41
1594	400	8649	3.46	74.87
1608	384	9033	3.32	78.19
1622	316	9349	2.74	80.93
1634	283	9632	2.45	83.38
1651	261	9893	2.26	85.64
1666	252	10145	2.18	87.82
1681	230	10375	1.99	89.81
1696	203	10578	1.76	91.57
1713	172	10750	1.49	93.06
1729	141	10891	1.22	94.28
1747	148	11039	1.28	95.56
1765	128	11167	1.11	96.67
1776	102	11269	0.88	97.55
1806	76	11345	0.66	98.21
1829	59	11404	0.51	98.72
1854	57	11461	0.49	99.21
1884	31	11492	0.27	99.48
1919	24	11516	0.21	99.69
1962	18	11534	0.16	99.84
2021	10	11544	0.09	99.93
2117	5	11549	0.04	99.97
2200	3	11552	0.03	100.00

Figure B.7.19. Spring 2024 STAAR Spanish Grade 3 RLA
Frequency Distribution of Scale Scores



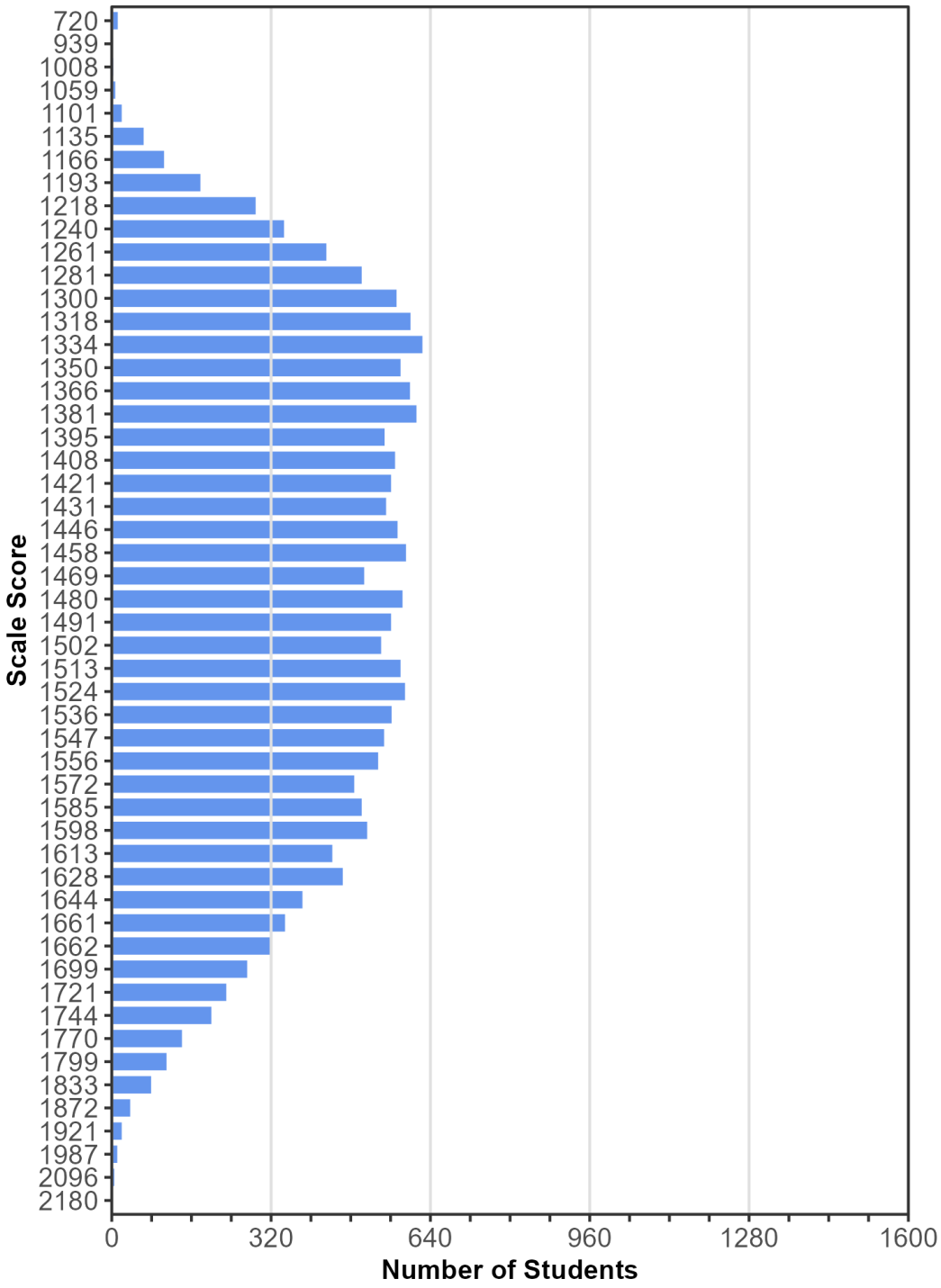
Scale Score	Freq	Cum Freq	Pct	Cum Pct
600	8	8	0.02	0.02
838	1	9	0.00	0.03
906	8	17	0.02	0.05
956	28	45	0.08	0.13
996	98	143	0.29	0.42
1030	198	341	0.58	1.00
1060	378	719	1.10	2.10
1086	731	1450	2.13	4.23
1110	956	2406	2.79	7.02
1133	1245	3651	3.63	10.66
1153	1543	5194	4.50	15.16
1173	1559	6753	4.55	19.71
1191	1706	8459	4.98	24.69
1209	1606	10065	4.69	29.38
1225	1482	11547	4.33	33.71
1241	1374	12921	4.01	37.72
1256	1328	14249	3.88	41.59
1271	1260	15509	3.68	45.27
1284	1137	16646	3.32	48.59
1297	1091	17737	3.18	51.78
1309	981	18718	2.86	54.64
1318	1005	19723	2.93	57.57
1332	899	20622	2.62	60.20
1343	884	21506	2.58	62.78
1354	847	22353	2.47	65.25
1364	819	23172	2.39	67.64
1374	815	23987	2.38	70.02
1384	820	24807	2.39	72.41
1395	746	25553	2.18	74.59
1405	756	26309	2.21	76.80
1416	746	27055	2.18	78.98
1427	707	27762	2.06	81.04
1438	664	28426	1.94	82.98
1447	668	29094	1.95	84.93
1463	642	29736	1.87	86.80
1476	610	30346	1.78	88.58
1490	580	30926	1.69	90.28
1504	521	31447	1.52	91.80
1515	453	31900	1.32	93.12
1535	461	32361	1.35	94.47
1552	405	32766	1.18	95.65
1570	357	33123	1.04	96.69
1590	302	33425	0.88	97.57
1611	242	33667	0.71	98.28
1634	190	33857	0.55	98.83
1661	155	34012	0.45	99.28
1692	108	34120	0.32	99.60
1729	65	34185	0.19	99.79
1775	42	34227	0.12	99.91
1840	22	34249	0.06	99.98
1948	6	34255	0.02	99.99
2070	2	34257	0.01	100.00

Figure B.7.20. Spring 2024 STAAR Spanish Grade 4 RLA
Frequency Distribution of Scale Scores



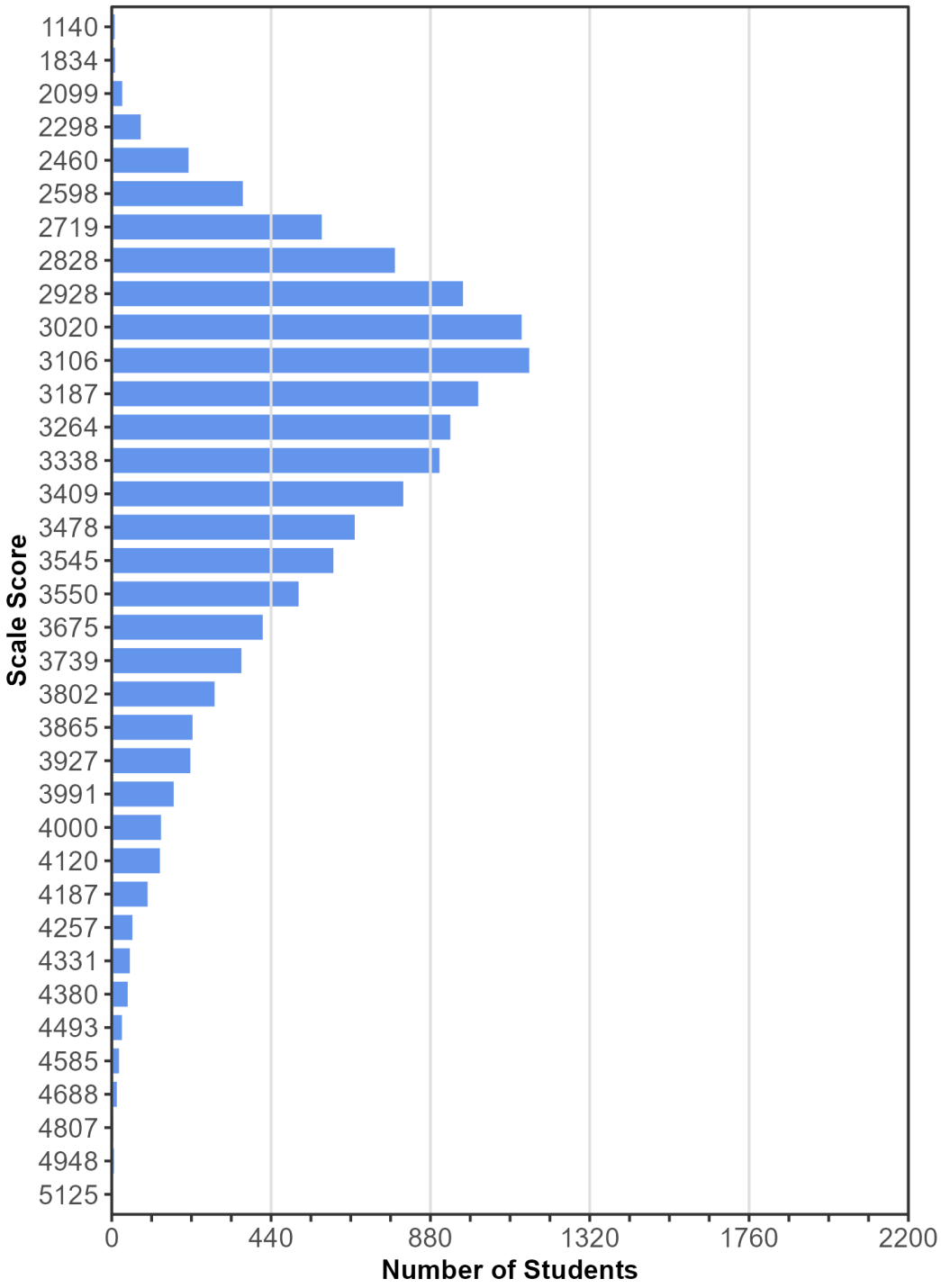
Scale Score	Freq	Cum Freq	Pct	Cum Pct
680	15	15	0.06	0.06
783	1	16	0.00	0.06
895	1	17	0.00	0.07
963	1	18	0.00	0.07
1013	13	31	0.05	0.12
1053	16	47	0.06	0.19
1087	58	105	0.23	0.41
1117	110	215	0.43	0.85
1143	223	438	0.88	1.73
1168	361	799	1.43	3.16
1190	502	1301	1.98	5.14
1211	694	1995	2.74	7.88
1231	889	2884	3.51	11.39
1249	977	3861	3.86	15.25
1267	1029	4890	4.06	19.31
1284	1106	5996	4.37	23.68
1300	1058	7054	4.18	27.86
1316	1041	8095	4.11	31.97
1332	912	9007	3.60	35.57
1346	972	9979	3.84	39.41
1361	897	10876	3.54	42.95
1375	893	11769	3.53	46.48
1389	797	12566	3.15	49.62
1402	818	13384	3.23	52.86
1408	769	14153	3.04	55.89
1428	790	14943	3.12	59.01
1441	759	15702	3.00	62.01
1453	749	16451	2.96	64.97
1466	690	17141	2.72	67.69
1478	754	17895	2.98	70.67
1488	700	18595	2.76	73.43
1504	619	19214	2.44	75.88
1517	677	19891	2.67	78.55
1530	630	20521	2.49	81.04
1544	609	21130	2.41	83.45
1558	563	21693	2.22	85.67
1572	486	22179	1.92	87.59
1581	461	22640	1.82	89.41
1604	456	23096	1.80	91.21
1620	413	23509	1.63	92.84
1638	390	23899	1.54	94.38
1656	332	24231	1.31	95.69
1676	273	24504	1.08	96.77
1697	231	24735	0.91	97.68
1721	174	24909	0.69	98.37
1746	139	25048	0.55	98.92
1774	118	25166	0.47	99.38
1807	77	25243	0.30	99.69
1845	44	25287	0.17	99.86
1894	22	25309	0.09	99.95
1959	8	25317	0.03	99.98
2067	5	25322	0.02	100.00

Figure B.7.21. Spring 2024 STAAR Spanish Grade 5 RLA
Frequency Distribution of Scale Scores



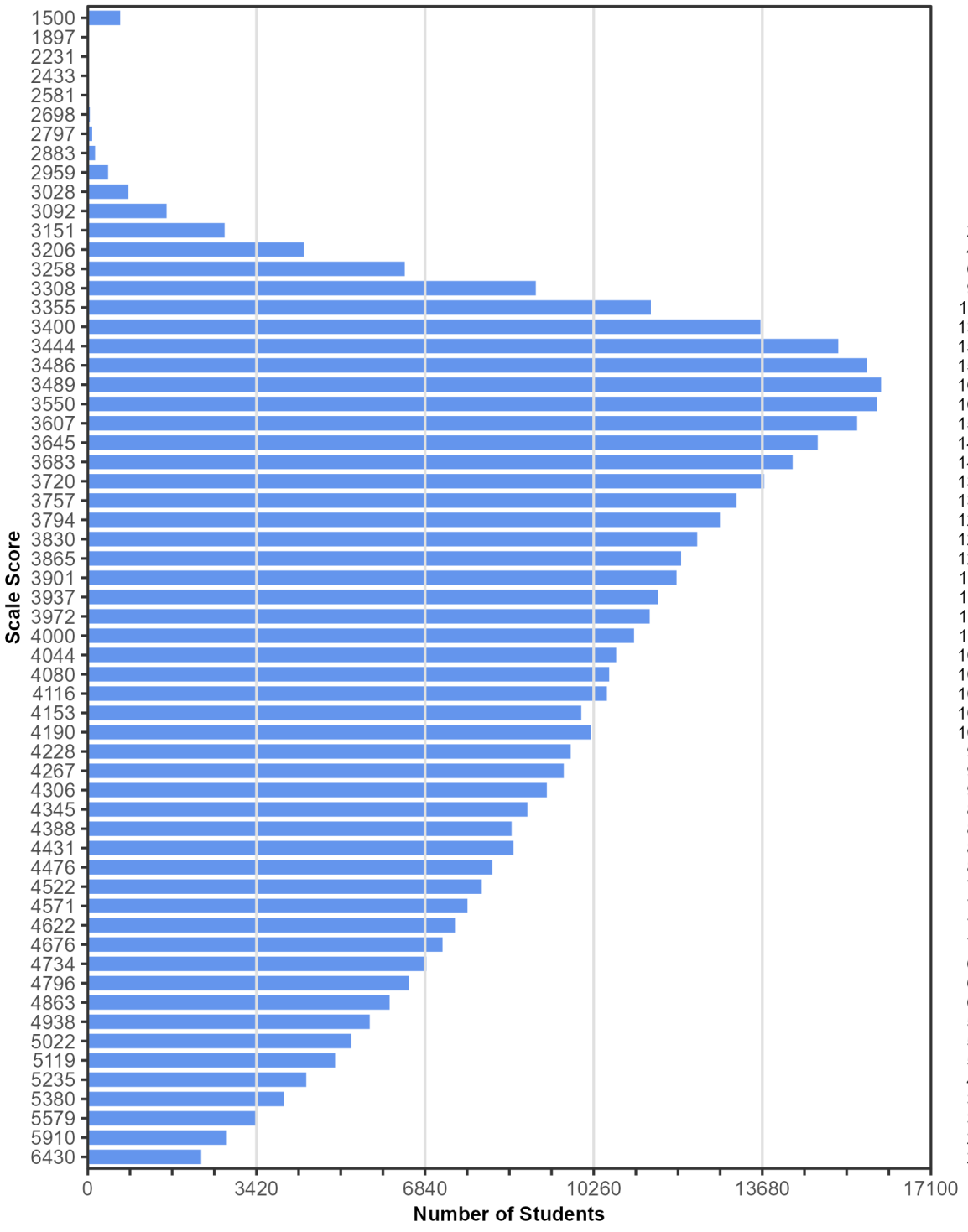
Scale Score	Freq	Cum Freq	Pct	Cum Pct
720	12	12	0.06	0.06
939	1	13	0.01	0.07
1008	3	16	0.02	0.09
1059	7	23	0.04	0.12
1101	20	43	0.11	0.23
1135	64	107	0.35	0.58
1166	105	212	0.57	1.15
1193	178	390	0.96	2.11
1218	289	679	1.56	3.67
1240	346	1025	1.87	5.54
1261	431	1456	2.33	7.87
1281	502	1958	2.71	10.58
1300	572	2530	3.09	13.67
1318	600	3130	3.24	16.91
1334	624	3754	3.37	20.28
1350	580	4334	3.13	23.41
1366	599	4933	3.24	26.65
1381	612	5545	3.31	29.96
1395	548	6093	2.96	32.92
1408	569	6662	3.07	35.99
1421	561	7223	3.03	39.02
1431	551	7774	2.98	42.00
1446	574	8348	3.10	45.10
1458	591	8939	3.19	48.29
1469	507	9446	2.74	51.03
1480	584	10030	3.16	54.19
1491	561	10591	3.03	57.22
1502	541	11132	2.92	60.14
1513	580	11712	3.13	63.27
1524	589	12301	3.18	66.46
1536	562	12863	3.04	69.49
1547	547	13410	2.96	72.45
1556	535	13945	2.89	75.34
1572	487	14432	2.63	77.97
1585	502	14934	2.71	80.68
1598	513	15447	2.77	83.45
1613	443	15890	2.39	85.85
1628	464	16354	2.51	88.35
1644	383	16737	2.07	90.42
1661	348	17085	1.88	92.30
1662	319	17404	1.72	94.02
1699	272	17676	1.47	95.49
1721	230	17906	1.24	96.74
1744	200	18106	1.08	97.82
1770	141	18247	0.76	98.58
1799	110	18357	0.59	99.17
1833	79	18436	0.43	99.60
1872	37	18473	0.20	99.80
1921	20	18493	0.11	99.91
1987	11	18504	0.06	99.97
2096	5	18509	0.03	99.99
2180	1	18510	0.01	100.00

Figure B.7.22. Spring 2024 STAAR Spanish Grade 5 Science
Frequency Distribution of Scale Scores



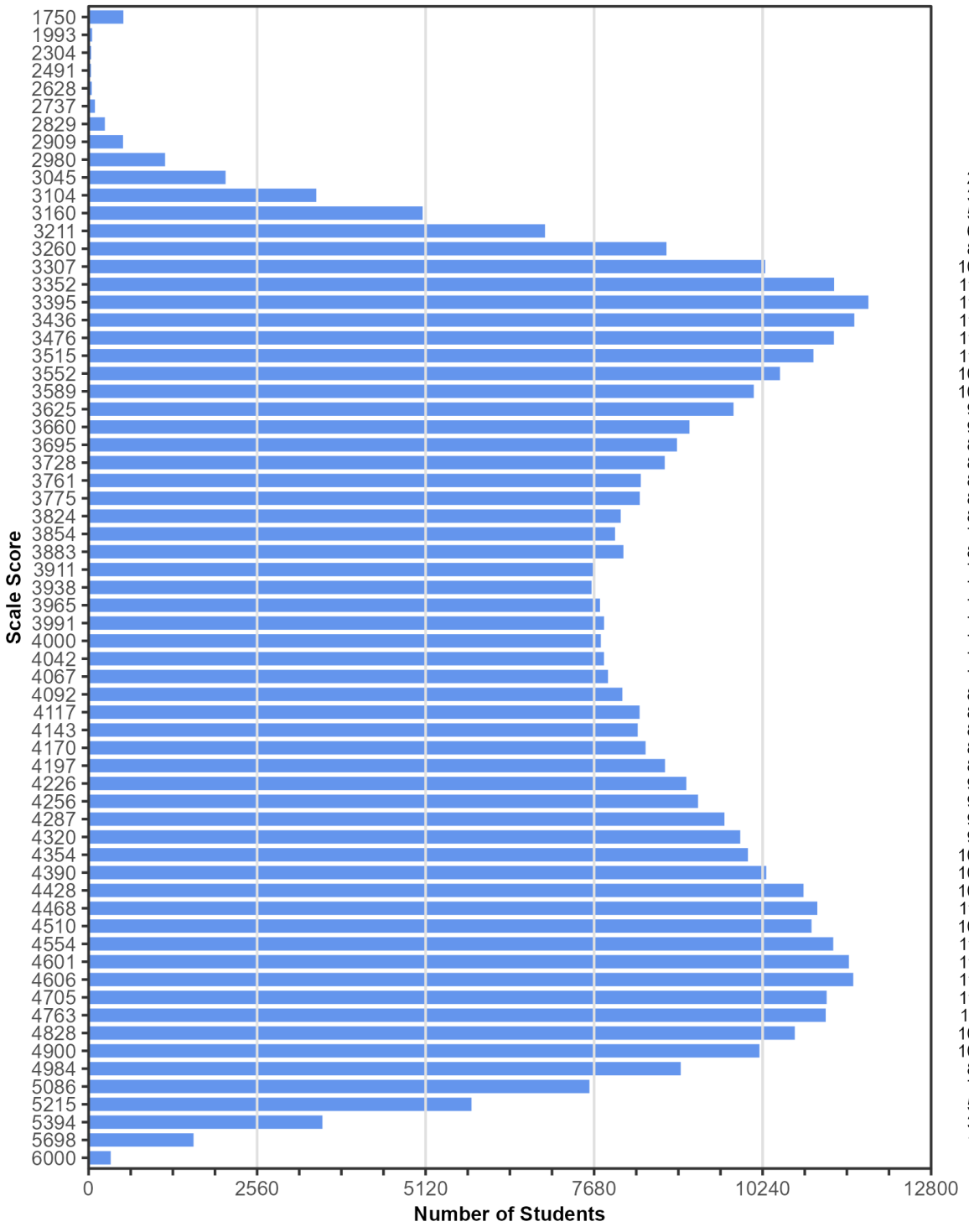
	Freq	Cum Freq	Pct	Cum Pct
1140	8	8	0.06	0.06
1834	9	17	0.07	0.13
2099	29	46	0.22	0.35
2298	80	126	0.61	0.97
2460	212	338	1.63	2.59
2598	362	700	2.78	5.37
2719	580	1280	4.45	9.82
2828	782	2062	6.00	15.82
2928	970	3032	7.44	23.26
3020	1132	4164	8.68	31.94
3106	1153	5317	8.84	40.79
3187	1012	6329	7.76	48.55
3264	935	7264	7.17	55.72
3338	905	8169	6.94	62.66
3409	805	8974	6.18	68.84
3478	671	9645	5.15	73.99
3545	612	10257	4.69	78.68
3550	516	10773	3.96	82.64
3675	417	11190	3.20	85.84
3739	358	11548	2.75	88.59
3802	284	11832	2.18	90.76
3865	223	12055	1.71	92.47
3927	217	12272	1.66	94.14
3991	171	12443	1.31	95.45
4000	136	12579	1.04	96.49
4120	133	12712	1.02	97.51
4187	99	12811	0.76	98.27
4257	57	12868	0.44	98.71
4331	50	12918	0.38	99.09
4380	44	12962	0.34	99.43
4493	28	12990	0.21	99.65
4585	20	13010	0.15	99.80
4688	14	13024	0.11	99.91
4807	4	13028	0.03	99.94
4948	6	13034	0.05	99.98
5125	2	13036	0.02	100.00

Figure B.7.23. Spring 2024 STAAR EOC Assessments Algebra I
Frequency Distribution of Scale Scores



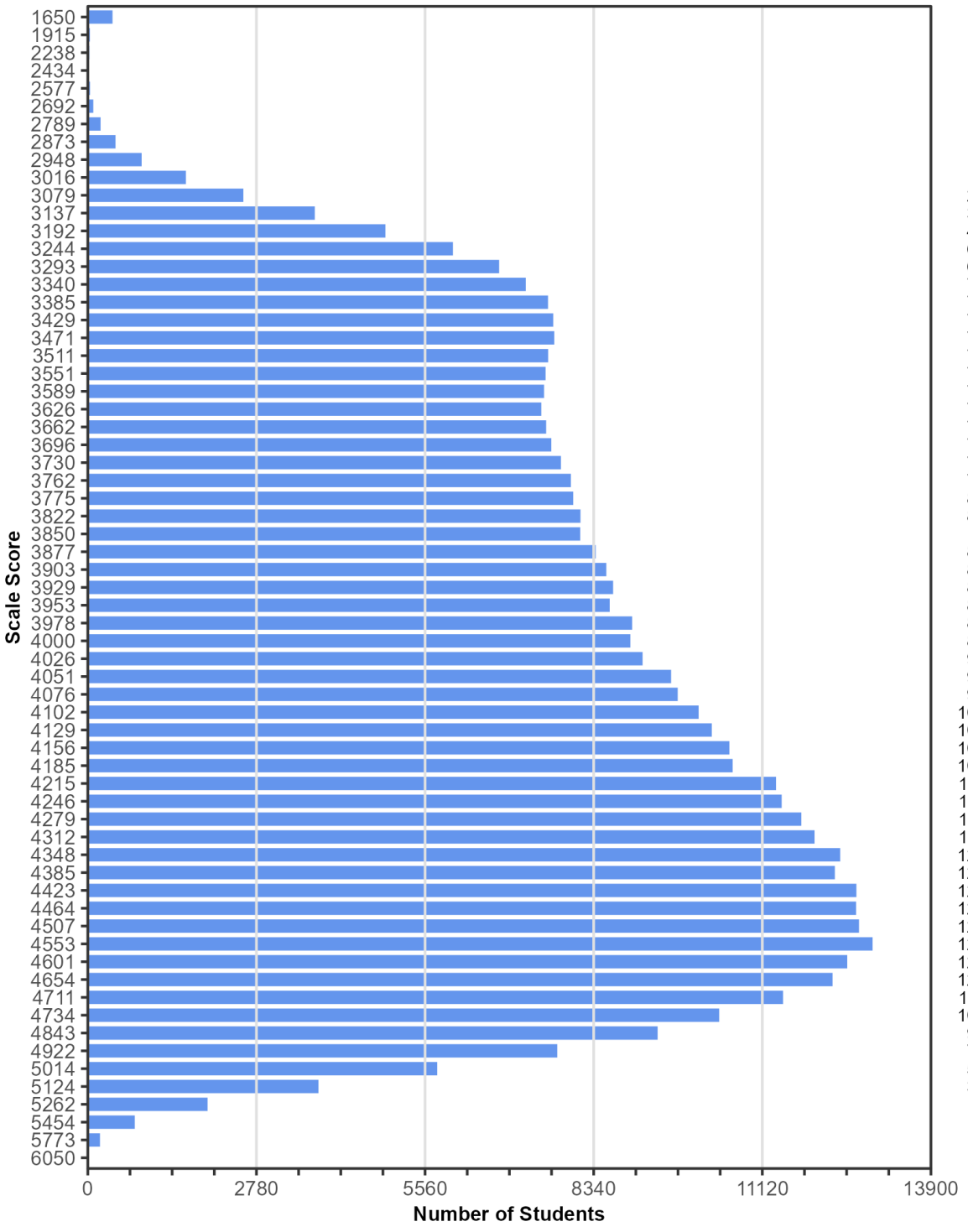
Freq	Cum Freq	Pct	Cum Pct
656	656	0.14	0.14
24	680	0.01	0.15
17	697	0.00	0.15
17	714	0.00	0.15
15	729	0.00	0.16
42	771	0.01	0.16
88	859	0.02	0.18
147	1006	0.03	0.22
411	1417	0.09	0.30
822	2239	0.18	0.48
1597	3836	0.34	0.82
2774	6610	0.59	1.41
4378	10988	0.94	2.35
6428	17416	1.38	3.73
9086	26502	1.94	5.67
11420	37922	2.44	8.11
13643	51565	2.92	11.03
15221	66786	3.26	14.29
15802	82588	3.38	17.67
16089	98677	3.44	21.11
16010	114687	3.42	24.53
15603	130290	3.34	27.87
14804	145094	3.17	31.04
14295	159389	3.06	34.10
13718	173107	2.93	37.03
13157	186264	2.81	39.85
12821	199085	2.74	42.59
12360	211445	2.64	45.23
12033	223478	2.57	47.81
11941	235419	2.55	50.36
11567	246986	2.47	52.84
11396	258382	2.44	55.27
11078	269460	2.37	57.64
10716	280176	2.29	59.94
10574	290750	2.26	62.20
10528	301278	2.25	64.45
10008	311286	2.14	66.59
10202	321488	2.18	68.77
9794	331282	2.10	70.87
9653	340935	2.07	72.93
9310	350245	1.99	74.93
8916	359161	1.91	76.83
8595	367756	1.84	78.67
8632	376388	1.85	80.52
8202	384590	1.75	82.27
7990	392580	1.71	83.98
7700	400280	1.65	85.63
7463	407743	1.60	87.23
7194	414937	1.54	88.77
6868	421805	1.47	90.24
6520	428325	1.39	91.63
6121	434446	1.31	92.94
5717	440163	1.22	94.16
5345	445508	1.14	95.31
5015	450523	1.07	96.38
4430	454953	0.95	97.33
3977	458930	0.85	98.18
3403	462333	0.73	98.91
2819	465152	0.60	99.51
2299	467451	0.49	100.00

Figure B.7.24. Spring 2024 STAAR EOC Assessments English I
Frequency Distribution of Scale Scores



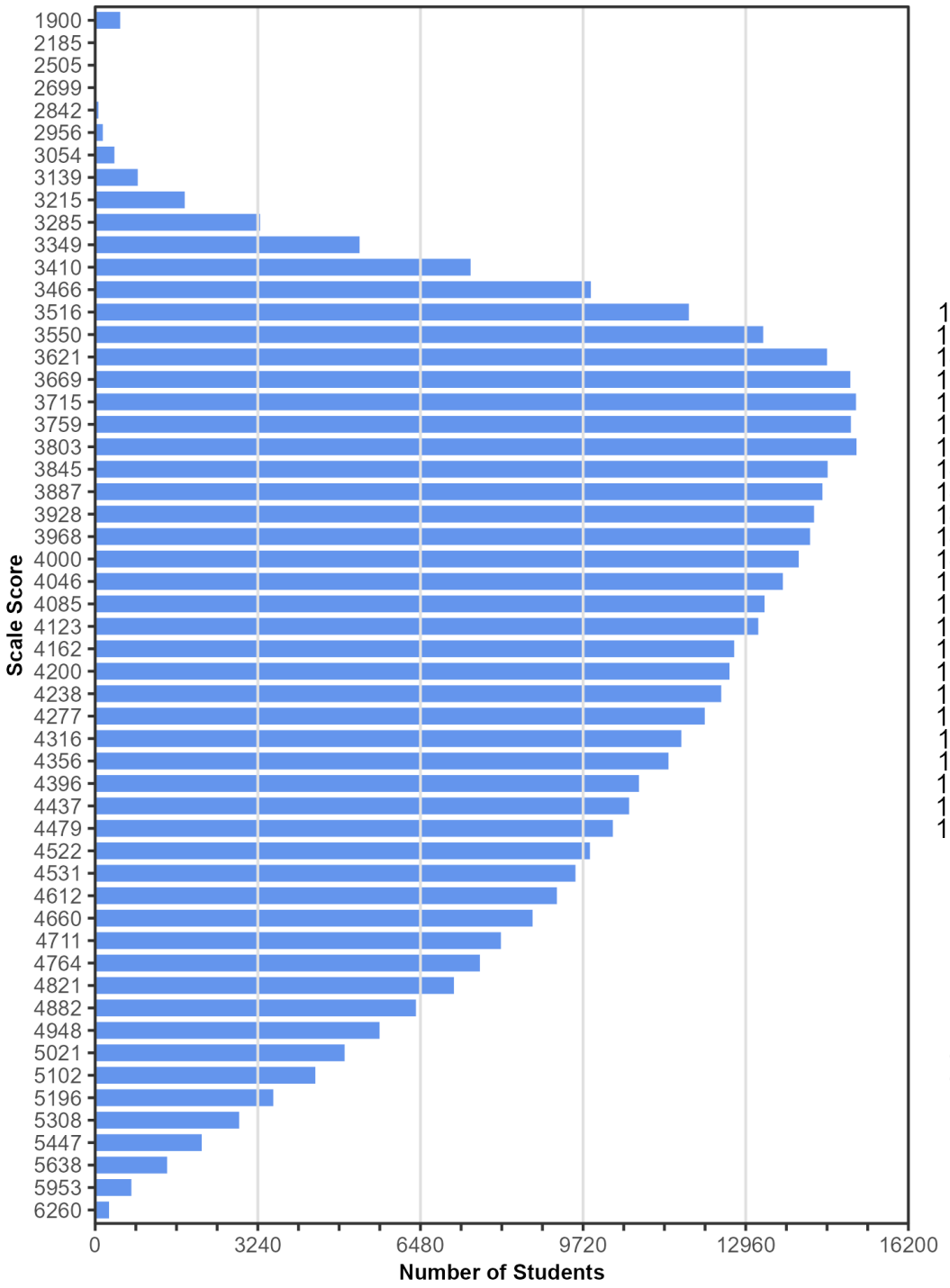
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1750	528	528	0.11	0.11
1993	53	581	0.01	0.12
2304	38	619	0.01	0.13
2491	36	655	0.01	0.13
2628	48	703	0.01	0.14
2737	96	799	0.02	0.16
2829	247	1046	0.05	0.21
2909	523	1569	0.11	0.32
2980	1162	2731	0.24	0.56
3045	2081	4812	0.43	0.99
3104	3459	8271	0.71	1.69
3160	5075	13346	1.04	2.73
3211	6934	20280	1.42	4.16
3260	8779	29059	1.80	5.95
3307	10280	39339	2.11	8.06
3352	11326	50665	2.32	10.38
3395	11847	62512	2.43	12.81
3436	11633	74145	2.38	15.19
3476	11323	85468	2.32	17.51
3515	11012	96480	2.26	19.77
3552	10504	106984	2.15	21.92
3589	10106	117090	2.07	23.99
3625	9798	126888	2.01	26.00
3660	9128	136016	1.87	27.87
3695	8939	144955	1.83	29.71
3728	8753	153708	1.79	31.50
3761	8389	162097	1.72	33.22
3775	8374	170471	1.72	34.93
3824	8084	178555	1.66	36.59
3854	7998	186553	1.64	38.23
3883	8125	194678	1.67	39.89
3911	7671	202349	1.57	41.47
3938	7641	209990	1.57	43.03
3965	7769	217759	1.59	44.62
3991	7831	225590	1.60	46.23
4000	7782	233372	1.59	47.82
4042	7830	241202	1.60	49.43
4067	7891	249093	1.62	51.05
4092	8109	257202	1.66	52.71
4117	8371	265573	1.72	54.42
4143	8342	273915	1.71	56.13
4170	8462	282377	1.73	57.87
4197	8757	291134	1.79	59.66
4226	9081	300215	1.86	61.52
4256	9259	309474	1.90	63.42
4287	9660	319134	1.98	65.40
4320	9900	329034	2.03	67.43
4354	10017	339051	2.05	69.48
4390	10295	349346	2.11	71.59
4428	10861	360207	2.23	73.82
4468	11071	371278	2.27	76.08
4510	10983	382261	2.25	78.34
4554	11313	393574	2.32	80.65
4601	11551	405125	2.37	83.02
4606	11617	416742	2.38	85.40
4705	11213	427955	2.30	87.70
4763	11199	439154	2.29	89.99
4828	10729	449883	2.20	92.19
4900	10192	460075	2.09	94.28
4984	8996	469071	1.84	96.13
5086	7609	476680	1.56	97.68
5215	5817	482497	1.19	98.88
5394	3552	486049	0.73	99.60
5698	1594	487643	0.33	99.93
6000	336	487979	0.07	100.00

Figure B.7.25. Spring 2024 STAAR EOC Assessments English II
Frequency Distribution of Scale Scores



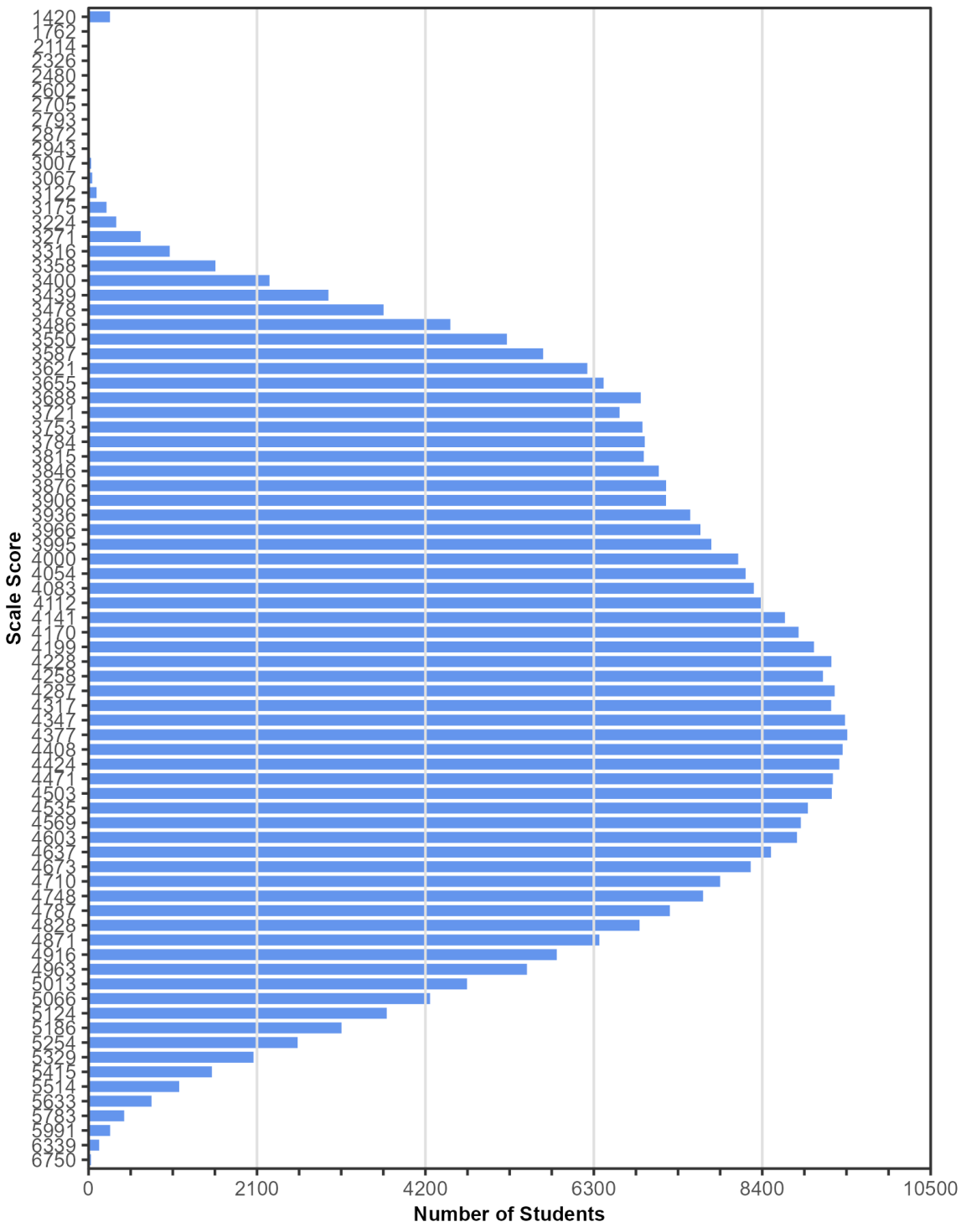
Freq	Cum Freq	Pct	Cum Pct
407	407	0.09	0.09
34	441	0.01	0.10
26	467	0.01	0.10
15	482	0.00	0.10
36	518	0.01	0.11
90	608	0.02	0.13
210	818	0.05	0.18
456	1274	0.10	0.27
887	2161	0.19	0.47
1617	3778	0.35	0.81
2564	6342	0.55	1.37
3742	10084	0.81	2.17
4906	14990	1.06	3.23
6019	21009	1.30	4.53
6781	27790	1.46	5.99
7220	35010	1.56	7.54
7587	42597	1.64	9.18
7673	50270	1.65	10.83
7690	57960	1.66	12.49
7589	65549	1.64	14.13
7547	73096	1.63	15.75
7522	80618	1.62	17.37
7476	88094	1.61	18.98
7556	95650	1.63	20.61
7641	103291	1.65	22.26
7800	111091	1.68	23.94
7964	119055	1.72	25.66
8002	127057	1.72	27.38
8122	135179	1.75	29.13
8118	143297	1.75	30.88
8373	151670	1.80	32.69
8548	160218	1.84	34.53
8660	168878	1.87	36.39
8604	177482	1.85	38.25
8974	186456	1.93	40.18
8944	195400	1.93	42.11
9146	204546	1.97	44.08
9616	214162	2.07	46.15
9725	223887	2.10	48.25
10070	233957	2.17	50.42
10287	244244	2.22	52.64
10577	254821	2.28	54.92
10631	265452	2.29	57.21
11346	276798	2.45	59.65
11439	288237	2.47	62.12
11763	300000	2.53	64.65
11981	311981	2.58	67.23
12404	324385	2.67	69.91
12316	336701	2.65	72.56
12671	349372	2.73	75.29
12666	362038	2.73	78.02
12714	374752	2.74	80.76
12938	387690	2.79	83.55
12520	400210	2.70	86.25
12278	412488	2.65	88.89
11462	423950	2.47	91.36
10409	434359	2.24	93.61
9394	443753	2.02	95.63
7739	451492	1.67	97.30
5760	457252	1.24	98.54
3802	461054	0.82	99.36
1974	463028	0.43	99.79
774	463802	0.17	99.95
200	464002	0.04	100.00
22	464024	0.00	100.00

Figure B.7.26. Spring 2024 STAAR EOC Assessments Biology
Frequency Distribution of Scale Scores



Freq	Cum Freq	Pct	Cum Pct
498	498	0.11	0.11
19	517	0.00	0.12
17	534	0.00	0.12
28	562	0.01	0.13
64	626	0.01	0.14
152	778	0.03	0.18
383	1161	0.09	0.26
848	2009	0.19	0.46
1784	3793	0.41	0.86
3285	7078	0.75	1.61
5267	12345	1.20	2.81
7479	19824	1.70	4.51
9875	29699	2.25	6.76
11827	41526	2.69	9.45
13308	54834	3.03	12.48
14580	69414	3.32	15.80
15044	84458	3.43	19.23
15157	99615	3.45	22.68
15055	114670	3.43	26.11
15167	129837	3.45	29.56
14592	144429	3.32	32.88
14488	158917	3.30	36.18
14321	173238	3.26	39.44
14242	187480	3.24	42.69
14016	201496	3.19	45.88
13700	215196	3.12	49.00
13335	228531	3.04	52.03
13210	241741	3.01	55.04
12730	254471	2.90	57.94
12637	267108	2.88	60.82
12472	279580	2.84	63.65
12144	291724	2.76	66.42
11677	303401	2.66	69.08
11421	314822	2.60	71.68
10833	325655	2.47	74.15
10637	336292	2.42	76.57
10313	346605	2.35	78.92
9856	356461	2.24	81.16
9568	366029	2.18	83.34
9198	375227	2.09	85.43
8714	383941	1.98	87.42
8083	392024	1.84	89.26
7664	399688	1.74	91.00
7147	406835	1.63	92.63
6390	413225	1.45	94.08
5665	418890	1.29	95.37
4969	423859	1.13	96.50
4386	428245	1.00	97.50
3548	431793	0.81	98.31
2868	434661	0.65	98.96
2122	436783	0.48	99.45
1433	438216	0.33	99.77
721	438937	0.16	99.94
276	439213	0.06	100.00

Figure B.7.27. Spring 2024 STAAR EOC Assessments U.S. History
Frequency Distribution of Scale Scores



Freq	Cum Freq	Pct	Cum Pct
267	267	0.07	0.07
10	277	0.00	0.07
9	286	0.00	0.07
9	295	0.00	0.08
12	307	0.00	0.08
4	311	0.00	0.08
14	325	0.00	0.08
9	334	0.00	0.09
8	342	0.00	0.09
14	356	0.00	0.09
32	388	0.01	0.10
46	434	0.01	0.11
100	534	0.03	0.14
224	758	0.06	0.20
345	1103	0.09	0.28
649	1752	0.17	0.45
1012	2764	0.26	0.71
1582	4346	0.41	1.12
2256	6602	0.58	1.70
2991	9593	0.77	2.47
3678	13271	0.95	3.42
4511	17782	1.16	4.58
5216	22998	1.34	5.92
5668	28666	1.46	7.38
6219	34885	1.60	8.98
6422	41307	1.65	10.64
6885	48192	1.77	12.41
6621	54813	1.70	14.11
6906	61719	1.78	15.89
6935	68654	1.79	17.68
6923	75577	1.78	19.46
7110	82687	1.83	21.29
7201	89888	1.85	23.15
7200	97088	1.85	25.00
7502	104590	1.93	26.93
7629	112219	1.96	28.90
7765	119984	2.00	30.90
8100	128084	2.09	32.98
8192	136276	2.11	35.09
8295	144571	2.14	37.23
8381	152952	2.16	39.39
8683	161635	2.24	41.62
8853	170488	2.28	43.90
9045	179533	2.33	46.23
9262	188795	2.39	48.62
9157	197952	2.36	50.97
9303	207255	2.40	53.37
9259	216514	2.38	55.75
9432	225946	2.43	58.18
9460	235406	2.44	60.62
9402	244808	2.42	63.04
9361	254169	2.41	65.45
9281	263450	2.39	67.84
9267	272717	2.39	70.23
8969	281686	2.31	72.54
8881	290567	2.29	74.82
8833	299400	2.27	77.10
8509	307909	2.19	79.29
8256	316165	2.13	81.42
7875	324040	2.03	83.44
7662	331702	1.97	85.42
7248	338950	1.87	87.28
6870	345820	1.77	89.05
6369	352189	1.64	90.69
5838	358027	1.50	92.20
5466	363493	1.41	93.60
4719	368212	1.22	94.82
4258	372470	1.10	95.91
3718	376188	0.96	96.87
3154	379342	0.81	97.68
2607	381949	0.67	98.36
2057	384006	0.53	98.89
1538	385544	0.40	99.28
1130	386674	0.29	99.57
785	387459	0.20	99.77
444	387903	0.11	99.89
269	388172	0.07	99.96
132	388304	0.03	99.99
30	388334	0.01	100.00

Spring 2024 STAAR
English Learner Performance Measure

Table B.8.1. Spring 2024 STAAR Grades 3–8 Mathematics
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 3	7,209	1040 14%	2772 38%	2031 28%	519 7%	847 12%
Grade 4	3,272	668 20%	1056 32%	733 22%	229 7%	586 18%
Grade 5	2,822	451 16%	971 34%	808 29%	168 6%	424 15%
Grade 6	4,385	801 18%	2246 51%	884 20%	162 4%	292 7%
Grade 7	4,132	1451 35%	1775 43%	601 15%	116 3%	189 5%
Grade 8	4,583	897 20%	2153 47%	968 21%	167 4%	398 9%

Table B.8.2. Spring 2024 STAAR Grades 3–8 RLA
English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 3	6,215	1374 22%	1724 28%	1898 31%	343 6%	876 14%
Grade 4	2,745	719 26%	812 30%	760 28%	83 3%	371 14%
Grade 5	2,454	766 31%	736 30%	465 19%	79 3%	408 17%
Grade 6	4,413	2107 48%	1276 29%	713 16%	84 2%	233 5%
Grade 7	4,583	2529 55%	1226 27%	504 11%	73 2%	251 5%
Grade 8	4,785	2412 50%	1394 29%	665 14%	92 2%	222 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 2024 STAAR Grade 5 and Grade 8 Science English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 5	2,766	808 29%	1374 50%	407 15%	39 1%	138 5%
Grade 8	4,656	1750 38%	1949 42%	672 14%	109 2%	176 4%

Table B.8.4. Spring 2024 STAAR Grade 8 Social Studies English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 8	4,765	2102 44%	2127 45%	325 7%	44 1%	167 4%

Table B.8.5. Spring 2024 STAAR EOC Assessments English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Algebra I	6,568	448 7%	3672 56%	1411 21%	237 4%	800 12%
English I	7,848	4274 54%	2566 33%	741 9%	175 2%	92 1%
English II	9,258	4240 46%	3184 34%	1629 18%	179 2%	26 0%
Biology	6,945	958 14%	3382 49%	2069 30%	281 4%	255 4%
U.S. History	4,397	192 4%	2149 49%	1390 32%	246 6%	420 10%

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 2023–2024

Appendix C

STAAR Alternate 2 Statistical Tables and Figures

List of Tables

Table C.1.1. 2024 STAAR Alternate 2 Grades 3–8 Classification Consistency and Accuracy

Table C.1.2. 2024 STAAR Alternate 2 EOC Assessments Classification Consistency and Accuracy

Table C.2.1. 2024 STAAR Alternate 2 Grades 3–8 Scale Score Correlations

Table C.2.2. 2024 STAAR Alternate 2 EOC Assessments Scale Score Correlations

Table C.3.1. 2024 STAAR Alternate 2 Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table C.3.2. 2024 STAAR Alternate 2 Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table C.3.3. 2024 STAAR Alternate 2 Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table C.3.4. 2024 STAAR Alternate 2 Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

Table C.3.5. 2024 STAAR Alternate 2 Grade 5 and Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

Table C.3.6. 2024 STAAR Alternate 2 Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

Table C.3.7. 2024 STAAR Alternate 2 EOC Assessments Mathematics Conditional Standard Error of Measurement for Scale Scores

Table C.3.8. 2024 STAAR Alternate 2 EOC Assessments RLA Conditional Standard Error of Measurement for Scale Scores

Table C.3.9. 2024 STAAR Alternate 2 EOC Assessments Science Conditional Standard Error of Measurement for Scale Scores

Table C.3.10. 2024 STAAR Alternate 2 EOC Assessments Social Studies Conditional Standard Error of Measurement for Scale Scores

Table C.4.1. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.2. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.3. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.4. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.5. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.6. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.7. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.8. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.9. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.10. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.11. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.12. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.13. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.14. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.15. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.16. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.17. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.18. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.19. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.20. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.21. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.22. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.23. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.24. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.25. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.26. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.27. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.28. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.29. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.30. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.31. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.32. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.33. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.34. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.35. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.36. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.37. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.38. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.39. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.40. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.41. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.42. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.43. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.44. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.45. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.46. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.47. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.48. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.49. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.50. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.51. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.52. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.53. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.54. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.55. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.56. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.57. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.58. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.59. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.60. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Table C.5.1. 2024 STAAR Alternate 2 All Mathematics Progress Measure

Table C.5.2. 2024 STAAR Alternate 2 All RLA Progress Measure

Table C.5.3. 2024 STAAR Alternate 2 Grade 4 Mathematics Progress Measure

Table C.5.4. 2024 STAAR Alternate 2 Grade 5 Mathematics Progress Measure

Table C.5.5. 2024 STAAR Alternate 2 Grade 6 Mathematics Progress Measure

Table C.5.6. 2024 STAAR Alternate 2 Grade 7 Mathematics Progress Measure

Table C.5.7. 2024 STAAR Alternate 2 Grade 8 Mathematics Progress Measure

Table C.5.8. 2024 STAAR Alternate 2 EOC Assessments Algebra I Progress Measure

Table C.5.9. 2024 STAAR Alternate 2 Grade 4 RLA Progress Measure

Table C.5.10. 2024 STAAR Alternate 2 Grade 5 RLA Progress Measure

Table C.5.11. 2024 STAAR Alternate 2 Grade 6 RLA Progress Measure

Table C.5.12. 2024 STAAR Alternate 2 Grade 7 RLA Progress Measure

Table C.5.13. 2024 STAAR Alternate 2 Grade 8 RLA Progress Measure

Table C.5.14. 2024 STAAR Alternate 2 EOC Assessments English I Progress Measure

Table C.5.15. 2024 STAAR Alternate 2 EOC Assessments English II Progress Measure

Table C.6.1. 2024 STAAR Alternate 2 Grades 3–8 Scale Score Descriptive Statistics

Table C.6.2. 2024 STAAR Alternate 2 EOC Assessments Scale Score Descriptive Statistics

List of Figures

Figure C.6.1. 2024 STAAR Alternate 2 Grade 3 Mathematics Frequency Distribution of Scale Scores

Figure C.6.2. 2024 STAAR Alternate 2 Grade 4 Mathematics Frequency Distribution of Scale Scores

Figure C.6.3. 2024 STAAR Alternate 2 Grade 5 Mathematics Frequency Distribution of Scale Scores

Figure C.6.4. 2024 STAAR Alternate 2 Grade 6 Mathematics Frequency Distribution of Scale Scores

Figure C.6.5. 2024 STAAR Alternate 2 Grade 7 Mathematics Frequency Distribution of Scale Scores

Figure C.6.6. 2024 STAAR Alternate 2 Grade 8 Mathematics Frequency Distribution of Scale Scores

Figure C.6.7. 2024 STAAR Alternate 2 Grade 3 RLA Frequency Distribution of Scale Scores

Figure C.6.8. 2024 STAAR Alternate 2 Grade 4 RLA Frequency Distribution of Scale Scores

Figure C.6.9. 2024 STAAR Alternate 2 Grade 5 RLA Frequency Distribution of Scale Scores

Figure C.6.10. 2024 STAAR Alternate 2 Grade 6 RLA Frequency Distribution of Scale Scores

Figure C.6.11. 2024 STAAR Alternate 2 Grade 7 RLA Frequency Distribution of Scale Scores

Figure C.6.12. 2024 STAAR Alternate 2 Grade 8 RLA Frequency Distribution of Scale Scores

Figure C.6.13. 2024 STAAR Alternate 2 Grade 5 Science Frequency Distribution of Scale Scores

Figure C.6.14. 2024 STAAR Alternate 2 Grade 8 Science Frequency Distribution of Scale Scores

Figure C.6.15. 2024 STAAR Alternate 2 Grade 8 Social Studies Frequency Distribution of Scale Scores

Figure C.6.16. 2024 STAAR Alternate 2 Algebra I Frequency Distribution of Scale Scores

Figure C.6.17. 2024 STAAR Alternate 2 English I Frequency Distribution of Scale Scores

Figure C.6.18. 2024 STAAR Alternate 2 English II Frequency Distribution of Scale Scores

Figure C.6.19. 2024 STAAR Alternate 2 Biology Frequency Distribution of Scale Scores

Figure C.6.20. 2024 STAAR Alternate 2 U.S. History Frequency Distribution of Scale Scores

2024 STAAR Alternate 2 Classification Consistency and Accuracy

Table C.1.1. 2024 STAAR Alternate 2 Grades 3–8
Classification Consistency and Accuracy

	Decision Consistency				Decision Accuracy			
Grade	**Mathematics**	**RLA**	**Science**	**Social Studies**	**Mathematics**	**RLA**	**Science**	**Social Studies**
3	79.0	79.8			85.0	85.7		
4	83.1	78.1			88.0	84.4		
5	79.9	74.1	80.5		85.7	81.5	86.1	
6	76.2	73.8			83.0	81.3		
7	77.8	75.0			84.2	82.2		
8	75.2	72.7	80.4	75.0	82.5	80.5	86.0	82.1

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. 2024 STAAR Alternate 2 EOC Assessments
Classification Consistency and Accuracy

Test	Decision Consistency	Decision Accuracy
Algebra I	73.4	81.1
English I	72.6	80.2
English II	74.3	81.8
Biology	78.7	84.7
U.S. History	75.5	82.5

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.

2024 STAAR Alternate 2
Scale Score Correlations

Table C.2.1. 2024 STAAR Alternate 2 Grades 3–8
Scale Score Correlations

Grade	STAAR Alternate 2	STAAR Alternate 2	N	Correlation
3	Mathematics	RLA	6,617	0.78
4	Mathematics	RLA	6,484	0.78
5	Mathematics	RLA	6,032	0.73
	Mathematics	Science	6,030	0.75
	RLA	Science	6,030	0.76
6	Mathematics	RLA	5,578	0.74
7	Mathematics	RLA	5,220	0.73
8	Mathematics	RLA	4,981	0.73
	Mathematics	Science	4,976	0.74
	Mathematics	Social Studies	4,980	0.72
	RLA	Science	4,976	0.74
	RLA	Social Studies	4,982	0.75
	Science	Social Studies	4,982	0.76

Table C.2.2. 2024 STAAR Alternate 2 EOC Assessments
Scale Score Correlations

STAAR Alternate 2	STAAR Alternate 2	N	Correlation
Algebra I	English I	5,020	0.67
Algebra I	English II	109	0.77
Algebra I	Biology	3,721	0.69
Algebra I	US History	78	0.66
English I	English II	58	0.90
English I	Biology	3,721	0.69
English I	US History	73	0.78
English II	Biology	1,203	0.74
English II	US History	485	0.78
Biology	US History	235	0.80

2024 STAAR Alternate 2 Conditional Standard Error of Measurement for Scale Scores

Table C.3.1. 2024 STAAR Alternate 2 Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	108		123		117	
1	153	40	168	39	162	39
2	178	28	192	27	187	28
3	193	24	207	22	201	23
4	205	21	217	20	212	20
5	214	20	225	18	220	18
6	223	19	232	17	227	17
7	231	18	239	16	234	16
8	238	17	245	15	240	16
9	244	17	250	15	245	15
10	250	16	255	15	251	15
11	256	16	260	14	256	14
12	262	15	265	14	260	14
13	267	15	270	14	265	14
14	272	15	274	14	270	14
15	277	14	279	14	274	14
16	281	14	283	14	279	14
17	286	14	287	14	283	14
18	290	14	292	14	287	14
19	294	14	296	14	291	13
20	**300	14	**300	14	296	13
21	303	14	305	14	**300	13
22	307	14	309	14	304	13
23	311	14	313	14	308	13
24	315	14	318	14	312	13
25	320	14	322	14	317	14
26	324	14	327	14	321	14
27	328	14	332	14	325	14
28	333	14	337	14	330	14
29	337	15	342	15	335	14
30	342	15	347	15	340	15
31	348	15	352	16	345	15
32	353	16	358	16	350	16

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	359	17	365	17	357	17
34	366	18	372	18	363	18
35	***375	19	380	19	371	19
36	383	21	***387	21	***379	21
37	395	24	401	24	392	24
38	411	30	418	29	408	29
39	439	42	446	41	436	42
40	489		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. 2024 STAAR Alternate 2 Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		103		100	
1	146	44	151	41	139	42
2	173	30	176	29	166	30
3	189	25	192	24	181	24
4	201	22	203	21	193	21
5	210	20	212	19	202	20
6	218	19	219	18	209	18
7	225	18	226	17	217	18
8	232	17	233	17	223	17
9	238	17	239	16	229	17
10	243	16	244	16	235	16
11	249	16	250	16	241	16
12	254	16	255	15	246	16
13	259	16	260	15	252	16
14	265	16	265	15	257	16
15	270	15	271	15	263	16
16	275	15	276	15	268	16
17	280	15	280	15	274	16
18	284	15	285	15	279	15
19	289	15	290	15	284	15
20	294	15	295	15	289	15
21	**300	15	**300	15	294	15
22	304	15	304	15	**300	15
23	309	15	309	15	304	15
24	313	15	314	15	309	15
25	318	15	319	15	314	15
26	323	15	324	15	319	15
27	328	16	328	15	324	15
28	333	16	334	15	329	16
29	339	16	339	15	335	16
30	344	16	344	16	340	16
31	350	17	350	16	346	17
32	356	18	356	17	352	17

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	363	18	363	18	359	18
34	***373	20	370	19	***365	19
35	380	21	***375	20	376	21
36	390	23	389	22	386	23
37	403	27	402	26	399	26
38	421	33	419	31	417	32
39	452	46	450	44	448	45
40	508		503		502	

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. 2024 STAAR Alternate 2 Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100		100	
1	141	48	133	50	125	50
2	172	34	165	35	157	35
3	190	28	184	29	176	28
4	203	25	198	25	189	25
5	214	23	208	23	200	22
6	223	21	217	21	209	21
7	231	20	225	20	217	20
8	239	19	232	19	224	19
9	246	19	239	18	230	18
10	252	18	245	18	237	18
11	258	18	250	17	242	17
12	264	17	256	17	248	17
13	270	17	261	16	253	17
14	276	17	266	16	258	16
15	281	17	271	16	263	16
16	286	16	276	16	268	16
17	291	16	280	16	273	16
18	296	16	285	16	278	16
19	**300	16	290	16	283	16
20	306	16	294	15	287	16
21	311	16	**300	16	292	16
22	316	16	303	16	297	16
23	320	16	308	16	**300	16
24	325	16	313	16	306	16
25	330	16	317	16	311	16
26	335	16	322	16	316	16
27	340	16	327	16	321	16
28	345	17	333	17	326	17
29	351	17	338	17	331	17
30	356	17	344	18	337	17
31	363	18	350	18	343	18
32	369	19	357	19	349	19

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	376	20	364	20	356	20
34	***388	21	372	21	364	21
35	394	23	***380	23	***374	23
36	405	25	393	26	385	25
37	419	29	407	29	398	29
38	438	35	427	36	418	35
39	472	50	461	51	452	51
40	532		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. 2024 STAAR Alternate 2 Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100		100	
1	121	50	116	51	122	49
2	153	35	149	35	154	34
3	173	29	168	29	173	28
4	187	26	182	25	186	24
5	198	23	193	23	196	22
6	208	22	202	22	205	20
7	216	21	210	20	213	19
8	224	20	218	20	220	19
9	231	19	225	19	227	18
10	237	18	231	18	233	17
11	243	18	237	18	239	17
12	249	17	243	18	245	17
13	255	17	249	17	250	16
14	260	17	254	17	256	16
15	265	16	260	17	261	16
16	270	16	265	17	266	16
17	275	16	271	17	271	16
18	280	16	276	17	276	16
19	285	16	281	17	281	16
20	290	16	286	17	286	16
21	295	16	292	17	291	16
22	**300	16	297	17	296	16
23	304	16	**300	17	**300	16
24	309	16	307	17	306	16
25	314	16	313	17	311	16
26	319	16	318	17	316	16
27	324	17	324	17	321	16
28	329	17	329	18	326	17
29	335	17	335	18	332	17
30	340	18	342	19	338	17
31	347	18	348	19	344	18
32	353	19	355	20	351	19

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	360	20	363	21	358	20
34	***370	21	372	22	***371	21
35	378	23	***378	24	376	23
36	389	26	394	27	387	25
37	403	30	409	30	401	29
38	424	36	430	37	421	35
39	458	52	465	53	455	50
40	520		528		514	

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. 2024 STAAR Alternate 2 Grade 5 and Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	110		131	
1	154	39	173	36
2	178	28	196	26
3	193	23	210	21
4	203	20	220	19
5	212	19	229	17
6	220	18	236	16
7	227	17	243	15
8	233	17	249	15
9	239	16	254	14
10	245	16	259	14
11	251	16	264	14
12	257	16	269	13
13	262	16	273	13
14	268	15	277	13
15	273	15	282	13
16	278	15	286	13
17	284	15	290	12
18	289	15	294	12
19	293	15	298	12
20	**300	14	**300	12
21	303	14	305	12
22	307	14	309	12
23	312	14	313	12
24	317	14	317	12
25	321	14	321	13
26	326	14	325	13
27	330	14	330	13
28	335	15	334	13
29	340	15	338	13
30	345	15	343	14
31	351	16	348	14
32	357	16	353	15

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	363	17	359	15
34	370	18	366	16
35	378	19	373	17
36	***387	21	***382	19
37	399	25	393	22
38	416	30	408	27
39	444	43	434	38
40	495		479	

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. 2024 STAAR Alternate 2 Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

		Grade 8
Raw	**SS**	**CSEM**
0	110	
1	152	37
2	175	26
3	188	22
4	198	19
5	206	18
6	213	17
7	219	16
8	225	15
9	231	15
10	236	15
11	241	15
12	247	15
13	252	15
14	257	14
15	262	14
16	267	14
17	272	14
18	277	14
19	282	14
20	287	14
21	292	14
22	297	14
23	**300	14
24	306	14
25	311	14
26	316	14
27	321	14
28	326	15
29	331	15
30	337	15
31	343	16
32	349	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	436	41
40	485	

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. 2024 STAAR Alternate 2 EOC Assessments Mathematics Conditional Standard Error of Measurement for Scale Scores

	Algebra I	
Raw	**SS**	**CSEM**
0	100	
1	139	44
2	168	31
3	185	25
4	197	22
5	206	20
6	214	19
7	222	18
8	228	17
9	234	16
10	240	16
11	246	16
12	251	15
13	256	15
14	261	15
15	266	15
16	271	15
17	275	15
18	280	15
19	285	15
20	290	15
21	294	15
22	**300	15
23	303	15
24	308	15
25	313	15
26	318	15
27	323	15
28	328	15
29	333	16
30	339	16
31	345	17
32	351	17

(Continued)

	Algebra I	
Raw	**SS**	**CSEM**
33	***361	18
34	365	19
35	374	21
36	384	23
37	398	26
38	416	32
39	447	45
40	501	

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. 2024 STAAR Alternate 2 EOC Assessments RLA Conditional Standard Error of Measurement for Scale Scores

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	100		100	
1	134	47	134	48
2	165	33	166	34
3	183	27	185	28
4	197	24	198	24
5	207	22	209	22
6	216	20	218	20
7	224	19	226	19
8	231	18	233	18
9	238	18	240	18
10	244	17	246	17
11	250	17	252	17
12	256	17	257	16
13	261	16	263	16
14	267	16	268	16
15	272	16	273	16
16	277	16	278	15
17	282	16	283	15
18	287	16	287	15
19	292	16	292	15
20	297	15	297	15
21	**300	16	**300	15
22	306	16	306	15
23	311	16	310	15
24	316	16	315	15
25	321	16	320	15
26	326	16	325	16
27	332	16	330	16
28	337	17	335	16
29	343	17	340	17
30	349	17	346	17
31	355	18	352	18
32	***365	19	359	19

(Continued)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	369	20	***370	20
34	378	21	374	21
35	387	23	384	23
36	399	25	395	25
37	413	29	409	29
38	433	35	429	35
39	467	49	463	49
40	526		523	

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. 2024 STAAR Alternate 2 EOC Assessments Science Conditional Standard Error of Measurement for Scale Scores

		Biology
Raw	**SS**	**CSEM**
0	129	
1	170	35
2	192	25
3	205	20
4	215	18
5	223	17
6	230	16
7	236	15
8	241	14
9	247	14
10	252	14
11	256	13
12	261	13
13	266	13
14	270	13
15	275	13
16	279	13
17	283	13
18	287	13
19	292	13
20	296	13
21	**300	13
22	304	13
23	308	13
24	313	13
25	317	13
26	321	13
27	326	13
28	330	13
29	335	14
30	340	14
31	345	14
32	351	15

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	357	16
34	364	17
35	371	18
36	***383	20
37	392	22
38	407	27
39	433	38
40	479	

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. 2024 STAAR Alternate 2 EOC Assessments Social Studies Conditional Standard Error of Measurement for Scale Scores

	U.S. History	
Raw	**SS**	**CSEM**
0	105	
1	149	38
2	172	26
3	186	22
4	197	19
5	205	18
6	212	17
7	219	16
8	225	16
9	230	15
10	236	15
11	241	15
12	246	14
13	251	14
14	256	14
15	261	14
16	266	14
17	271	14
18	276	14
19	280	14
20	285	14
21	290	14
22	295	14
23	**300	14
24	304	14
25	309	14
26	314	14
27	319	15
28	324	15
29	330	15
30	335	15
31	341	16
32	348	16

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
33	354	17
34	362	18
35	***368	19
36	380	21
37	393	24
38	409	29
39	438	41
40	487	

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.

2024 STAAR Alternate 2
Mean P-Values, Raw Score Summary, and Reliability

Table C.4.1. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,620	28.38	7.58	0.85	2.91	1.42
	Numerical Representations and Relationships	16	6,620	11.06	3.32	0.68	1.88	1.38
	Computations and Algebraic Relationships	8	6,620	5.50	2.09	0.56	1.39	1.37
	Geometry and Measurement	8	6,620	6.28	1.82	0.59	1.16	1.57
	Data Analysis and Personal Financial Literacy	8	6,620	5.54	1.81	0.47	1.31	1.39
RLA	OVERALL TEST	40	6,623	26.31	7.39	0.82	3.06	1.32
	Reading	24	6,623	16.50	4.53	0.75	2.26	1.37
	Writing	16	6,623	9.81	3.48	0.64	2.06	1.23

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. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,960	28.10	7.34	0.84	2.94	1.40
	Numerical Representations and Relationships	16	1,960	10.93	3.25	0.66	1.90	1.37
	Computations and Algebraic Relationships	8	1,960	5.41	2.08	0.54	1.40	1.35
	Geometry and Measurement	8	1,960	6.26	1.79	0.56	1.17	1.56
	Data Analysis and Personal Financial Literacy	8	1,960	5.50	1.78	0.45	1.31	1.37
RLA	OVERALL TEST	40	1,962	26.35	7.22	0.82	3.07	1.32
	Reading	24	1,962	16.51	4.39	0.73	2.27	1.38
	Writing	16	1,962	9.84	3.44	0.63	2.07	1.23

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. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,657	28.50	7.67	0.85	2.90	1.43
	Numerical Representations and Relationships	16	4,657	11.11	3.35	0.68	1.87	1.39
	Computations and Algebraic Relationships	8	4,657	5.54	2.10	0.56	1.38	1.38
	Geometry and Measurement	8	4,657	6.29	1.83	0.60	1.15	1.57
	Data Analysis and Personal Financial Literacy	8	4,657	5.56	1.82	0.48	1.31	1.39
RLA	OVERALL TEST	40	4,658	26.29	7.45	0.83	3.06	1.31
	Reading	24	4,658	16.49	4.59	0.75	2.25	1.37
	Writing	16	4,658	9.80	3.49	0.65	2.06	1.22

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. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,142	27.77	7.60	0.85	2.97	1.39
	Numerical Representations and Relationships	16	1,142	10.80	3.37	0.69	1.90	1.35
	Computations and Algebraic Relationships	8	1,142	5.32	2.12	0.56	1.41	1.33
	Geometry and Measurement	8	1,142	6.17	1.81	0.57	1.19	1.54
	Data Analysis and Personal Financial Literacy	8	1,142	5.48	1.82	0.46	1.34	1.37
RLA	OVERALL TEST	40	1,144	25.83	7.51	0.83	3.09	1.29
	Reading	24	1,144	16.22	4.65	0.76	2.28	1.35
	Writing	16	1,144	9.61	3.50	0.65	2.08	1.20

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. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,627	28.98	7.51	0.85	2.85	1.45
	Numerical Representations and Relationships	16	3,627	11.30	3.28	0.68	1.84	1.41
	Computations and Algebraic Relationships	8	3,627	5.65	2.05	0.55	1.36	1.41
	Geometry and Measurement	8	3,627	6.39	1.80	0.60	1.13	1.60
	Data Analysis and Personal Financial Literacy	8	3,627	5.65	1.79	0.48	1.28	1.41
RLA	OVERALL TEST	40	3,626	26.92	7.39	0.83	3.00	1.35
	Reading	24	3,626	16.82	4.51	0.75	2.21	1.40
	Writing	16	3,626	10.10	3.47	0.65	2.02	1.26

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. 2024 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,228	27.99	7.50	0.84	2.97	1.40
	Numerical Representations and Relationships	16	1,228	10.86	3.35	0.67	1.92	1.36
	Computations and Algebraic Relationships	8	1,228	5.43	2.09	0.54	1.40	1.36
	Geometry and Measurement	8	1,228	6.25	1.84	0.58	1.17	1.56
	Data Analysis and Personal Financial Literacy	8	1,228	5.45	1.79	0.44	1.33	1.36
RLA	OVERALL TEST	40	1,228	25.88	7.03	0.80	3.13	1.29
	Reading	24	1,228	16.37	4.35	0.71	2.30	1.36
	Writing	16	1,228	9.51	3.38	0.60	2.12	1.19

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. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,490	29.24	6.88	0.83	2.78	1.46
	Numerical Representations and Relationships	16	6,490	11.56	3.21	0.69	1.78	1.45
	Computations and Algebraic Relationships	8	6,490	5.69	1.74	0.43	1.30	1.42
	Geometry and Measurement	8	6,490	6.18	1.55	0.48	1.11	1.55
	Data Analysis and Personal Financial Literacy	8	6,490	5.80	1.83	0.51	1.28	1.45
RLA	OVERALL TEST	40	6,488	28.00	7.18	0.82	2.99	1.40
	Reading	24	6,488	17.72	4.52	0.76	2.21	1.48
	Writing	16	6,488	10.28	3.34	0.64	2.00	1.29

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. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,980	29.36	6.81	0.83	2.76	1.47
	Numerical Representations and Relationships	16	1,980	11.58	3.14	0.68	1.78	1.45
	Computations and Algebraic Relationships	8	1,980	5.71	1.71	0.43	1.29	1.43
	Geometry and Measurement	8	1,980	6.21	1.54	0.48	1.11	1.55
	Data Analysis and Personal Financial Literacy	8	1,980	5.85	1.84	0.52	1.27	1.46
RLA	OVERALL TEST	40	1,979	28.21	7.25	0.83	2.96	1.41
	Reading	24	1,979	17.85	4.56	0.76	2.19	1.49
	Writing	16	1,979	10.35	3.36	0.65	1.98	1.29

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. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,509	29.18	6.91	0.83	2.79	1.46
	Numerical Representations and Relationships	16	4,509	11.55	3.24	0.69	1.79	1.44
	Computations and Algebraic Relationships	8	4,509	5.68	1.75	0.43	1.31	1.42
	Geometry and Measurement	8	4,509	6.17	1.56	0.48	1.12	1.54
	Data Analysis and Personal Financial Literacy	8	4,509	5.77	1.83	0.50	1.29	1.44
RLA	OVERALL TEST	40	4,508	27.91	7.15	0.82	3.01	1.40
	Reading	24	4,508	17.66	4.51	0.76	2.22	1.47
	Writing	16	4,508	10.25	3.33	0.63	2.01	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.4.10. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,110	28.93	6.83	0.83	2.80	1.45
	Numerical Representations and Relationships	16	1,110	11.54	3.20	0.69	1.79	1.44
	Computations and Algebraic Relationships	8	1,110	5.56	1.73	0.42	1.31	1.39
	Geometry and Measurement	8	1,110	6.09	1.56	0.47	1.13	1.52
	Data Analysis and Personal Financial Literacy	8	1,110	5.75	1.79	0.48	1.30	1.44
RLA	OVERALL TEST	40	1,112	27.76	7.26	0.83	3.01	1.39
	Reading	24	1,112	17.59	4.65	0.77	2.22	1.47
	Writing	16	1,112	10.17	3.30	0.62	2.02	1.27

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. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,581	29.87	6.56	0.83	2.73	1.49
	Numerical Representations and Relationships	16	3,581	11.80	3.08	0.68	1.75	1.48
	Computations and Algebraic Relationships	8	3,581	5.81	1.68	0.42	1.28	1.45
	Geometry and Measurement	8	3,581	6.32	1.50	0.48	1.09	1.58
	Data Analysis and Personal Financial Literacy	8	3,581	5.93	1.78	0.50	1.26	1.48
RLA	OVERALL TEST	40	3,579	28.59	6.86	0.82	2.94	1.43
	Reading	24	3,579	18.04	4.31	0.75	2.17	1.50
	Writing	16	3,579	10.55	3.25	0.63	1.97	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.4.12. 2024 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,174	28.25	7.52	0.84	2.86	1.41
	Numerical Representations and Relationships	16	1,174	11.11	3.44	0.69	1.86	1.39
	Computations and Algebraic Relationships	8	1,174	5.53	1.85	0.45	1.34	1.38
	Geometry and Measurement	8	1,174	5.99	1.66	0.49	1.14	1.50
	Data Analysis and Personal Financial Literacy	8	1,174	5.62	1.96	0.53	1.30	1.40
RLA	OVERALL TEST	40	1,172	27.17	7.62	0.83	3.06	1.36
	Reading	24	1,172	17.31	4.80	0.76	2.26	1.44
	Writing	16	1,172	9.85	3.50	0.64	2.05	1.23

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	6,033	30.49	6.88	0.84	2.74	1.52
	Numerical Representations and Relationships	8	6,033	6.06	1.73	0.46	1.27	1.52
	Computations and Algebraic Relationships	16	6,033	12.41	3.08	0.70	1.69	1.55
	Geometry and Measurement	8	6,033	5.78	1.74	0.45	1.29	1.45
	Data Analysis and Personal Financial Literacy	8	6,033	6.24	1.70	0.47	1.23	1.56
RLA	OVERALL TEST	40	6,032	28.71	6.80	0.81	2.97	1.44
	Reading	24	6,032	17.73	4.21	0.72	2.22	1.48
	Writing	16	6,032	10.97	3.24	0.63	1.97	1.37
Science	OVERALL TEST	40	6,033	31.06	6.65	0.83	2.68	1.55
	Matter and Energy	8	6,033	6.35	1.68	0.47	1.22	1.59
	Force, Motion, and Energy	8	6,033	6.17	1.71	0.49	1.21	1.54
	Earth and Space	16	6,033	12.54	2.88	0.66	1.68	1.57
	Organisms and Environments	8	6,033	6.00	1.69	0.44	1.26	1.50

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,914	30.06	6.52	0.81	2.80	1.50
	Numerical Representations and Relationships	8	1,914	5.92	1.69	0.41	1.29	1.48
	Computations and Algebraic Relationships	16	1,914	12.24	2.94	0.65	1.74	1.53
	Geometry and Measurement	8	1,914	5.72	1.70	0.40	1.31	1.43
	Data Analysis and Personal Financial Literacy	8	1,914	6.18	1.65	0.42	1.26	1.54
RLA	OVERALL TEST	40	1,914	28.78	6.54	0.79	2.97	1.44
	Reading	24	1,914	17.86	4.06	0.70	2.21	1.49
	Writing	16	1,914	10.92	3.15	0.60	1.98	1.36
Science	OVERALL TEST	40	1,914	31.06	6.36	0.82	2.69	1.55
	Matter and Energy	8	1,914	6.37	1.65	0.45	1.22	1.59
	Force, Motion, and Energy	8	1,914	6.15	1.68	0.48	1.21	1.54
	Earth and Space	16	1,914	12.55	2.77	0.63	1.67	1.57
	Organisms and Environments	8	1,914	5.99	1.64	0.41	1.26	1.50

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,119	30.69	7.03	0.85	2.72	1.53
	Numerical Representations and Relationships	8	4,119	6.13	1.75	0.48	1.25	1.53
	Computations and Algebraic Relationships	16	4,119	12.49	3.15	0.71	1.67	1.56
	Geometry and Measurement	8	4,119	5.81	1.76	0.47	1.28	1.45
	Data Analysis and Personal Financial Literacy	8	4,119	6.27	1.72	0.49	1.22	1.57
RLA	OVERALL TEST	40	4,118	28.67	6.92	0.81	2.97	1.43
	Reading	24	4,118	17.67	4.27	0.73	2.22	1.47
	Writing	16	4,118	11.00	3.28	0.64	1.96	1.38
Science	OVERALL TEST	40	4,119	31.07	6.78	0.84	2.68	1.55
	Matter and Energy	8	4,119	6.35	1.70	0.48	1.21	1.59
	Force, Motion, and Energy	8	4,119	6.17	1.72	0.49	1.22	1.54
	Earth and Space	16	4,119	12.54	2.93	0.67	1.68	1.57
	Organisms and Environments	8	4,119	6.01	1.71	0.46	1.25	1.50

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,027	30.04	7.15	0.84	2.76	1.50
	Numerical Representations and Relationships	8	1,027	6.01	1.78	0.48	1.26	1.50
	Computations and Algebraic Relationships	16	1,027	12.16	3.18	0.70	1.71	1.52
	Geometry and Measurement	8	1,027	5.68	1.81	0.47	1.30	1.42
	Data Analysis and Personal Financial Literacy	8	1,027	6.20	1.75	0.49	1.23	1.55
RLA	OVERALL TEST	40	1,027	28.54	7.17	0.82	2.95	1.43
	Reading	24	1,027	17.75	4.45	0.75	2.19	1.48
	Writing	16	1,027	10.79	3.33	0.64	1.97	1.35
Science	OVERALL TEST	40	1,027	30.76	7.06	0.85	2.69	1.54
	Matter and Energy	8	1,027	6.30	1.73	0.49	1.21	1.57
	Force, Motion, and Energy	8	1,027	6.09	1.80	0.53	1.21	1.52
	Earth and Space	16	1,027	12.37	3.01	0.67	1.69	1.55
	Organisms and Environments	8	1,027	5.99	1.79	0.50	1.25	1.50

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,309	31.20	6.57	0.84	2.67	1.56
	Numerical Representations and Relationships	8	3,309	6.19	1.68	0.46	1.23	1.55
	Computations and Algebraic Relationships	16	3,309	12.69	2.94	0.69	1.65	1.59
	Geometry and Measurement	8	3,309	5.93	1.68	0.45	1.25	1.48
	Data Analysis and Personal Financial Literacy	8	3,309	6.38	1.63	0.46	1.21	1.60
RLA	OVERALL TEST	40	3,309	29.29	6.47	0.80	2.92	1.46
	Reading	24	3,309	18.06	3.99	0.71	2.18	1.51
	Writing	16	3,309	11.23	3.14	0.62	1.93	1.40
Science	OVERALL TEST	40	3,309	31.58	6.33	0.83	2.62	1.58
	Matter and Energy	8	3,309	6.47	1.61	0.47	1.18	1.62
	Force, Motion, and Energy	8	3,309	6.26	1.64	0.48	1.19	1.56
	Earth and Space	16	3,309	12.77	2.77	0.66	1.63	1.60
	Organisms and Environments	8	3,309	6.09	1.62	0.43	1.23	1.52

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. 2024 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,090	29.46	7.34	0.84	2.84	1.47
	Numerical Representations and Relationships	8	1,090	5.85	1.81	0.45	1.32	1.46
	Computations and Algebraic Relationships	16	1,090	12.01	3.34	0.72	1.74	1.50
	Geometry and Measurement	8	1,090	5.55	1.81	0.43	1.35	1.39
	Data Analysis and Personal Financial Literacy	8	1,090	6.04	1.78	0.47	1.28	1.51
RLA	OVERALL TEST	40	1,089	27.90	7.14	0.81	3.06	1.40
	Reading	24	1,089	17.24	4.45	0.72	2.29	1.44
	Writing	16	1,089	10.66	3.35	0.63	2.03	1.33
Science	OVERALL TEST	40	1,090	30.50	7.12	0.84	2.75	1.53
	Matter and Energy	8	1,090	6.26	1.77	0.48	1.25	1.57
	Force, Motion, and Energy	8	1,090	6.05	1.79	0.50	1.25	1.51
	Earth and Space	16	1,090	12.28	3.04	0.67	1.71	1.54
	Organisms and Environments	8	1,090	5.90	1.76	0.44	1.30	1.48

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. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	5,585	30.09	6.65	0.82	2.77	1.50
	Numerical Representations and Relationships	8	5,585	6.62	1.65	0.55	1.10	1.65
	Computations and Algebraic Relationships	16	5,585	11.89	3.01	0.64	1.80	1.49
	Geometry and Measurement	8	5,585	5.97	1.75	0.53	1.19	1.49
	Data Analysis and Personal Financial Literacy	8	5,585	5.62	1.77	0.43	1.33	1.41
RLA	OVERALL TEST	40	5,585	28.93	6.83	0.81	2.93	1.45
	Reading	24	5,585	17.74	4.25	0.73	2.21	1.48
	Writing	16	5,585	11.19	3.21	0.63	1.93	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.4.20. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,804	29.99	6.65	0.82	2.77	1.50
	Numerical Representations and Relationships	8	1,804	6.52	1.67	0.54	1.13	1.63
	Computations and Algebraic Relationships	16	1,804	11.89	3.04	0.65	1.78	1.49
	Geometry and Measurement	8	1,804	5.95	1.73	0.51	1.21	1.49
	Data Analysis and Personal Financial Literacy	8	1,804	5.62	1.74	0.43	1.31	1.41
RLA	OVERALL TEST	40	1,805	29.31	6.80	0.82	2.89	1.47
	Reading	24	1,805	17.99	4.19	0.73	2.18	1.50
	Writing	16	1,805	11.32	3.20	0.64	1.90	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.4.21. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,779	30.14	6.65	0.82	2.76	1.51
	Numerical Representations and Relationships	8	3,779	6.66	1.64	0.56	1.08	1.66
	Computations and Algebraic Relationships	16	3,779	11.88	3.00	0.63	1.81	1.49
	Geometry and Measurement	8	3,779	5.97	1.75	0.54	1.19	1.49
	Data Analysis and Personal Financial Literacy	8	3,779	5.62	1.78	0.43	1.33	1.41
RLA	OVERALL TEST	40	3,778	28.75	6.84	0.81	2.95	1.44
	Reading	24	3,778	17.63	4.28	0.72	2.22	1.47
	Writing	16	3,778	11.12	3.21	0.63	1.94	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.4.22. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	991	30.02	6.50	0.82	2.79	1.50
	Numerical Representations and Relationships	8	991	6.58	1.63	0.52	1.12	1.64
	Computations and Algebraic Relationships	16	991	11.94	3.00	0.64	1.79	1.49
	Geometry and Measurement	8	991	5.92	1.76	0.54	1.20	1.48
	Data Analysis and Personal Financial Literacy	8	991	5.58	1.74	0.41	1.33	1.40
RLA	OVERALL TEST	40	988	29.10	6.68	0.81	2.95	1.46
	Reading	24	988	17.80	4.22	0.72	2.23	1.48
	Writing	16	988	11.30	3.12	0.62	1.93	1.41

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. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,045	30.51	6.60	0.82	2.70	1.53
	Numerical Representations and Relationships	8	3,045	6.70	1.63	0.57	1.05	1.68
	Computations and Algebraic Relationships	16	3,045	12.03	2.95	0.63	1.77	1.50
	Geometry and Measurement	8	3,045	6.07	1.73	0.54	1.17	1.52
	Data Analysis and Personal Financial Literacy	8	3,045	5.71	1.77	0.45	1.30	1.43
RLA	OVERALL TEST	40	3,045	29.30	6.81	0.81	2.88	1.46
	Reading	24	3,045	17.96	4.25	0.73	2.16	1.50
	Writing	16	3,045	11.34	3.17	0.63	1.90	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.4.24. 2024 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,068	29.56	6.62	0.82	2.84	1.48
	Numerical Representations and Relationships	8	1,068	6.49	1.66	0.52	1.15	1.62
	Computations and Algebraic Relationships	16	1,068	11.67	3.08	0.64	1.84	1.46
	Geometry and Measurement	8	1,068	5.84	1.73	0.49	1.24	1.46
	Data Analysis and Personal Financial Literacy	8	1,068	5.56	1.71	0.36	1.37	1.39
RLA	OVERALL TEST	40	1,071	28.29	6.83	0.81	3.00	1.41
	Reading	24	1,071	17.43	4.23	0.72	2.25	1.45
	Writing	16	1,071	10.86	3.23	0.63	1.98	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.4.25. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	5,225	31.35	6.17	0.82	2.63	1.57
	Numerical Representations and Relationships	16	5,225	12.82	2.82	0.65	1.66	1.60
	Computations and Algebraic Relationships	8	5,225	6.05	1.68	0.47	1.22	1.51
	Geometry and Measurement	8	5,225	6.71	1.49	0.52	1.02	1.68
	Data Analysis and Personal Financial Literacy	8	5,225	5.77	1.68	0.41	1.28	1.44
RLA	OVERALL TEST	40	5,227	29.79	6.42	0.81	2.80	1.49
	Reading	24	5,227	18.44	4.07	0.74	2.09	1.54
	Writing	16	5,227	11.34	2.98	0.61	1.86	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.4.26. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,717	31.16	6.03	0.80	2.64	1.56
	Numerical Representations and Relationships	16	1,717	12.71	2.74	0.62	1.69	1.59
	Computations and Algebraic Relationships	8	1,717	6.04	1.67	0.47	1.21	1.51
	Geometry and Measurement	8	1,717	6.68	1.48	0.52	1.02	1.67
	Data Analysis and Personal Financial Literacy	8	1,717	5.73	1.67	0.42	1.27	1.43
RLA	OVERALL TEST	40	1,718	29.88	6.40	0.81	2.78	1.49
	Reading	24	1,718	18.54	4.04	0.74	2.06	1.55
	Writing	16	1,718	11.34	2.98	0.61	1.85	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. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,508	31.44	6.24	0.82	2.63	1.57
	Numerical Representations and Relationships	16	3,508	12.87	2.86	0.67	1.64	1.61
	Computations and Algebraic Relationships	8	3,508	6.05	1.69	0.48	1.22	1.51
	Geometry and Measurement	8	3,508	6.73	1.49	0.53	1.02	1.68
	Data Analysis and Personal Financial Literacy	8	3,508	5.79	1.68	0.41	1.29	1.45
RLA	OVERALL TEST	40	3,508	29.75	6.41	0.81	2.81	1.49
	Reading	24	3,508	18.40	4.07	0.73	2.10	1.53
	Writing	16	3,508	11.35	2.97	0.60	1.87	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.4.28. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	915	31.28	6.12	0.82	2.65	1.56
	Numerical Representations and Relationships	16	915	12.75	2.79	0.64	1.69	1.59
	Computations and Algebraic Relationships	8	915	6.02	1.70	0.48	1.24	1.51
	Geometry and Measurement	8	915	6.79	1.45	0.54	1.00	1.70
	Data Analysis and Personal Financial Literacy	8	915	5.72	1.66	0.39	1.30	1.43
RLA	OVERALL TEST	40	920	30.20	6.33	0.81	2.77	1.51
	Reading	24	920	18.73	4.01	0.74	2.06	1.56
	Writing	16	920	11.47	2.98	0.62	1.84	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.4.29. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,835	31.72	5.96	0.81	2.59	1.59
	Numerical Representations and Relationships	16	2,835	12.96	2.75	0.65	1.62	1.62
	Computations and Algebraic Relationships	8	2,835	6.13	1.65	0.47	1.20	1.53
	Geometry and Measurement	8	2,835	6.76	1.44	0.50	1.01	1.69
	Data Analysis and Personal Financial Literacy	8	2,835	5.87	1.63	0.39	1.26	1.47
RLA	OVERALL TEST	40	2,832	30.12	6.22	0.80	2.76	1.51
	Reading	24	2,832	18.65	3.93	0.73	2.05	1.55
	Writing	16	2,832	11.47	2.92	0.60	1.84	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. 2024 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	994	31.01	6.32	0.82	2.66	1.55
	Numerical Representations and Relationships	16	994	12.70	2.90	0.67	1.67	1.59
	Computations and Algebraic Relationships	8	994	5.97	1.68	0.45	1.24	1.49
	Geometry and Measurement	8	994	6.67	1.51	0.53	1.03	1.67
	Data Analysis and Personal Financial Literacy	8	994	5.68	1.72	0.43	1.30	1.42
RLA	OVERALL TEST	40	993	29.30	6.49	0.81	2.86	1.47
	Reading	24	993	18.17	4.16	0.74	2.13	1.51
	Writing	16	993	11.13	2.96	0.59	1.90	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.4.31. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	4,984	31.09	5.94	0.80	2.66	1.55
	Numerical Representations and Relationships	16	4,984	12.73	2.78	0.64	1.65	1.59
	Computations and Algebraic Relationships	8	4,984	6.11	1.58	0.41	1.21	1.53
	Geometry and Measurement	8	4,984	6.03	1.53	0.32	1.26	1.51
	Data Analysis and Personal Financial Literacy	8	4,984	6.21	1.61	0.48	1.16	1.55
RLA	OVERALL TEST	40	4,985	29.79	6.36	0.80	2.80	1.49
	Reading	24	4,985	17.82	4.08	0.72	2.14	1.49
	Writing	16	4,985	11.97	2.94	0.63	1.79	1.50
Science	OVERALL TEST	40	4,982	31.53	6.26	0.82	2.60	1.58
	Matter and Energy	8	4,982	6.38	1.65	0.52	1.15	1.59
	Force, Motion, and Energy	8	4,982	6.54	1.55	0.49	1.11	1.63
	Earth and Space	8	4,982	6.08	1.67	0.43	1.26	1.52
	Organisms and Environments	16	4,982	12.54	2.73	0.63	1.65	1.57
Social Studies	OVERALL TEST	40	4,988	31.52	5.87	0.80	2.57	1.58
	History	16	4,988	12.36	2.64	0.59	1.69	1.55
	Geography and Culture	8	4,988	6.48	1.51	0.44	1.13	1.62
	Government and Citizenship	8	4,988	6.21	1.51	0.36	1.21	1.55
	Economics, Science, Technology, and Society	8	4,988	6.47	1.46	0.41	1.12	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.4.32. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,679	30.61	5.88	0.79	2.70	1.53
	Numerical Representations and Relationships	16	1,679	12.48	2.74	0.62	1.68	1.56
	Computations and Algebraic Relationships	8	1,679	6.05	1.58	0.41	1.21	1.51
	Geometry and Measurement	8	1,679	5.93	1.57	0.34	1.27	1.48
	Data Analysis and Personal Financial Literacy	8	1,679	6.15	1.59	0.45	1.17	1.54
RLA	OVERALL TEST	40	1,680	29.76	6.42	0.81	2.78	1.49
	Reading	24	1,680	17.81	4.11	0.73	2.12	1.48
	Writing	16	1,680	11.94	2.93	0.62	1.79	1.49
Science	OVERALL TEST	40	1,677	31.43	6.28	0.83	2.60	1.57
	Matter and Energy	8	1,677	6.32	1.70	0.54	1.14	1.58
	Force, Motion, and Energy	8	1,677	6.54	1.55	0.50	1.09	1.64
	Earth and Space	8	1,677	6.03	1.66	0.41	1.27	1.51
	Organisms and Environments	16	1,677	12.54	2.78	0.65	1.64	1.57
Social Studies	OVERALL TEST	40	1,680	31.32	5.90	0.81	2.57	1.57
	History	16	1,680	12.32	2.64	0.59	1.67	1.54
	Geography and Culture	8	1,680	6.42	1.53	0.45	1.13	1.60
	Government and Citizenship	8	1,680	6.19	1.51	0.36	1.20	1.55
	Economics, Science, Technology, and Society	8	1,680	6.39	1.47	0.41	1.12	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. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	3,305	31.33	5.96	0.80	2.64	1.57
	Numerical Representations and Relationships	16	3,305	12.86	2.79	0.65	1.63	1.61
	Computations and Algebraic Relationships	8	3,305	6.15	1.58	0.42	1.20	1.54
	Geometry and Measurement	8	3,305	6.08	1.51	0.30	1.25	1.52
	Data Analysis and Personal Financial Literacy	8	3,305	6.24	1.62	0.49	1.15	1.56
RLA	OVERALL TEST	40	3,305	29.80	6.33	0.80	2.81	1.49
	Reading	24	3,305	17.82	4.07	0.72	2.15	1.49
	Writing	16	3,305	11.98	2.94	0.63	1.79	1.50
Science	OVERALL TEST	40	3,305	31.59	6.25	0.82	2.60	1.58
	Matter and Energy	8	3,305	6.41	1.63	0.50	1.15	1.60
	Force, Motion, and Energy	8	3,305	6.54	1.56	0.48	1.12	1.63
	Earth and Space	8	3,305	6.10	1.68	0.44	1.26	1.53
	Organisms and Environments	16	3,305	12.54	2.71	0.62	1.65	1.57
Social Studies	OVERALL TEST	40	3,308	31.62	5.85	0.80	2.58	1.58
	History	16	3,308	12.39	2.64	0.58	1.70	1.55
	Geography and Culture	8	3,308	6.50	1.50	0.43	1.13	1.63
	Government and Citizenship	8	3,308	6.21	1.51	0.36	1.21	1.55
	Economics, Science, Technology, and Society	8	3,308	6.51	1.45	0.40	1.11	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.4.34. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	878	31.09	5.93	0.80	2.66	1.55
	Numerical Representations and Relationships	16	878	12.79	2.79	0.65	1.64	1.60
	Computations and Algebraic Relationships	8	878	6.14	1.59	0.43	1.20	1.53
	Geometry and Measurement	8	878	5.98	1.54	0.32	1.27	1.50
	Data Analysis and Personal Financial Literacy	8	878	6.18	1.59	0.46	1.16	1.54
RLA	OVERALL TEST	40	878	29.92	6.62	0.82	2.81	1.50
	Reading	24	878	17.91	4.29	0.75	2.14	1.49
	Writing	16	878	12.01	2.99	0.64	1.80	1.50
Science	OVERALL TEST	40	878	31.50	6.42	0.84	2.61	1.58
	Matter and Energy	8	878	6.32	1.66	0.50	1.17	1.58
	Force, Motion, and Energy	8	878	6.58	1.57	0.51	1.10	1.64
	Earth and Space	8	878	6.07	1.71	0.46	1.25	1.52
	Organisms and Environments	16	878	12.54	2.79	0.65	1.65	1.57
Social Studies	OVERALL TEST	40	878	31.57	6.11	0.82	2.57	1.58
	History	16	878	12.41	2.74	0.63	1.68	1.55
	Geography and Culture	8	878	6.50	1.49	0.42	1.14	1.63
	Government and Citizenship	8	878	6.22	1.55	0.40	1.20	1.56
	Economics, Science, Technology, and Society	8	878	6.44	1.51	0.45	1.12	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.4.35. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	2,687	31.54	5.91	0.81	2.59	1.58
	Numerical Representations and Relationships	16	2,687	12.91	2.75	0.66	1.60	1.61
	Computations and Algebraic Relationships	8	2,687	6.21	1.55	0.42	1.18	1.55
	Geometry and Measurement	8	2,687	6.11	1.50	0.33	1.23	1.53
	Data Analysis and Personal Financial Literacy	8	2,687	6.31	1.58	0.47	1.14	1.58
RLA	OVERALL TEST	40	2,690	30.30	6.15	0.80	2.73	1.52
	Reading	24	2,690	18.14	3.94	0.72	2.09	1.51
	Writing	16	2,690	12.17	2.83	0.62	1.73	1.52
Science	OVERALL TEST	40	2,685	31.98	6.10	0.83	2.53	1.60
	Matter and Energy	8	2,685	6.47	1.62	0.53	1.10	1.62
	Force, Motion, and Energy	8	2,685	6.61	1.51	0.49	1.08	1.65
	Earth and Space	8	2,685	6.14	1.66	0.44	1.23	1.54
	Organisms and Environments	16	2,685	12.75	2.63	0.63	1.59	1.59
Social Studies	OVERALL TEST	40	2,690	31.95	5.68	0.80	2.51	1.60
	History	16	2,690	12.54	2.56	0.58	1.65	1.57
	Geography and Culture	8	2,690	6.57	1.47	0.45	1.09	1.64
	Government and Citizenship	8	2,690	6.28	1.49	0.36	1.18	1.57
	Economics, Science, Technology, and Society	8	2,690	6.56	1.39	0.38	1.09	1.64

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. 2024 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Mathematics	OVERALL TEST	40	1,001	30.67	5.55	0.74	2.75	1.53
	Numerical Representations and Relationships	16	1,001	12.57	2.64	0.56	1.71	1.57
	Computations and Algebraic Relationships	8	1,001	6.01	1.57	0.36	1.24	1.50
	Geometry and Measurement	8	1,001	5.97	1.52	0.25	1.30	1.49
	Data Analysis and Personal Financial Literacy	8	1,001	6.11	1.60	0.45	1.17	1.53
RLA	OVERALL TEST	40	1,000	29.31	6.18	0.77	2.88	1.47
	Reading	24	1,000	17.49	3.98	0.69	2.20	1.46
	Writing	16	1,000	11.82	2.91	0.59	1.84	1.48
Science	OVERALL TEST	40	1,001	31.36	6.15	0.80	2.66	1.57
	Matter and Energy	8	1,001	6.39	1.63	0.46	1.17	1.60
	Force, Motion, and Energy	8	1,001	6.49	1.58	0.46	1.14	1.62
	Earth and Space	8	1,001	6.12	1.60	0.34	1.28	1.53
	Organisms and Environments	16	1,001	12.37	2.76	0.61	1.69	1.55
Social Studies	OVERALL TEST	40	1,002	31.13	5.75	0.78	2.63	1.56
	History	16	1,002	12.16	2.61	0.55	1.72	1.52
	Geography and Culture	8	1,002	6.42	1.49	0.37	1.17	1.60
	Government and Citizenship	8	1,002	6.15	1.47	0.27	1.24	1.54
	Economics, Science, Technology, and Society	8	1,002	6.41	1.50	0.40	1.14	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.4.37. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	5,112	29.86	6.42	0.81	2.77	1.49
	Number and Algebraic Methods	8	5,112	5.78	1.63	0.42	1.24	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	5,112	5.72	1.69	0.39	1.32	1.43
	Writing and Solving Linear Functions, Equations, and Inequalities	8	5,112	6.37	1.63	0.47	1.17	1.59
	Quadratic Functions and Equations	8	5,112	6.05	1.82	0.58	1.17	1.51
	Exponential Functions and Equations	8	5,112	5.93	1.69	0.42	1.27	1.48

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. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	1,755	29.46	6.74	0.83	2.76	1.47
	Number and Algebraic Methods	8	1,755	5.77	1.64	0.44	1.23	1.44
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	1,755	5.68	1.70	0.39	1.32	1.42
	Writing and Solving Linear Functions, Equations, and Inequalities	8	1,755	6.27	1.72	0.53	1.18	1.57
	Quadratic Functions and Equations	8	1,755	5.94	1.89	0.61	1.18	1.49
	Exponential Functions and Equations	8	1,755	5.80	1.76	0.48	1.27	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.4.39. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	3,357	30.06	6.23	0.80	2.77	1.50
	Number and Algebraic Methods	8	3,357	5.79	1.63	0.41	1.25	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	3,357	5.75	1.68	0.38	1.31	1.44
	Writing and Solving Linear Functions, Equations, and Inequalities	8	3,357	6.42	1.58	0.44	1.17	1.60
	Quadratic Functions and Equations	8	3,357	6.11	1.78	0.56	1.17	1.53
	Exponential Functions and Equations	8	3,357	5.99	1.65	0.39	1.28	1.50

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. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability 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	855	29.71	6.23	0.81	2.75	1.49
	Number and Algebraic Methods	8	855	5.76	1.65	0.45	1.22	1.44
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	855	5.70	1.68	0.40	1.30	1.42
	Writing and Solving Linear Functions, Equations, and Inequalities	8	855	6.36	1.55	0.44	1.16	1.59
	Quadratic Functions and Equations	8	855	5.99	1.77	0.56	1.18	1.50
	Exponential Functions and Equations	8	855	5.90	1.65	0.39	1.29	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.4.41. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	2,744	30.30	6.27	0.81	2.72	1.51
	Number and Algebraic Methods	8	2,744	5.88	1.61	0.41	1.23	1.47
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	2,744	5.81	1.65	0.38	1.30	1.45
	Writing and Solving Linear Functions, Equations, and Inequalities	8	2,744	6.46	1.58	0.47	1.14	1.62
	Quadratic Functions and Equations	8	2,744	6.11	1.78	0.57	1.16	1.53
	Exponential Functions and Equations	8	2,744	6.03	1.64	0.41	1.25	1.51

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. 2024 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Algebra I	OVERALL TEST	40	1,113	29.32	6.71	0.81	2.82	1.47
	Number and Algebraic Methods	8	1,113	5.62	1.65	0.40	1.26	1.40
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	1,113	5.61	1.73	0.38	1.35	1.40
	Writing and Solving Linear Functions, Equations, and Inequalities	8	1,113	6.24	1.75	0.50	1.22	1.56
	Quadratic Functions and Equations	8	1,113	6.04	1.90	0.61	1.17	1.51
	Exponential Functions and Equations	8	1,113	5.82	1.76	0.45	1.29	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.4.43. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	5,116	29.44	6.55	0.82	2.78	1.47
	Reading	24	5,116	18.15	4.12	0.74	2.07	1.51
	Writing	16	5,116	11.29	3.09	0.64	1.84	1.41
English II	OVERALL TEST	40	4,683	29.76	6.81	0.83	2.78	1.49
	Reading	24	4,683	18.35	4.29	0.76	2.08	1.53
	Writing	16	4,683	11.41	3.15	0.65	1.84	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.4.44. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	1,754	29.32	6.84	0.84	2.77	1.47
	Reading	24	1,754	18.10	4.28	0.77	2.06	1.51
	Writing	16	1,754	11.21	3.18	0.67	1.82	1.40
English II	OVERALL TEST	40	1,530	29.76	6.76	0.83	2.78	1.49
	Reading	24	1,530	18.40	4.25	0.76	2.07	1.53
	Writing	16	1,530	11.37	3.13	0.65	1.85	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.4.45. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	3,362	29.50	6.40	0.81	2.79	1.47
	Reading	24	3,362	18.17	4.03	0.73	2.08	1.51
	Writing	16	3,362	11.33	3.03	0.63	1.84	1.42
English II	OVERALL TEST	40	3,153	29.76	6.84	0.83	2.79	1.49
	Reading	24	3,153	18.32	4.31	0.76	2.08	1.53
	Writing	16	3,153	11.43	3.15	0.66	1.84	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.4.46. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability 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	862	29.58	6.73	0.83	2.74	1.48
	Reading	24	862	18.24	4.17	0.76	2.05	1.52
	Writing	16	862	11.34	3.20	0.68	1.80	1.42
English II	OVERALL TEST	40	815	29.92	6.76	0.82	2.79	1.50
	Reading	24	815	18.42	4.34	0.76	2.06	1.54
	Writing	16	815	11.50	3.12	0.64	1.84	1.44

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. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	2,753	29.77	6.31	0.81	2.75	1.49
	Reading	24	2,753	18.37	3.95	0.73	2.05	1.53
	Writing	16	2,753	11.40	2.99	0.63	1.82	1.43
English II	OVERALL TEST	40	2,485	30.03	6.75	0.83	2.74	1.50
	Reading	24	2,485	18.51	4.22	0.76	2.04	1.54
	Writing	16	2,485	11.52	3.12	0.66	1.81	1.44

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. 2024 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
English I	OVERALL TEST	40	1,105	28.91	6.71	0.82	2.83	1.45
	Reading	24	1,105	17.80	4.27	0.75	2.11	1.48
	Writing	16	1,105	11.12	3.10	0.63	1.87	1.39
English II	OVERALL TEST	40	1,028	29.51	6.59	0.81	2.85	1.48
	Reading	24	1,028	18.21	4.17	0.73	2.13	1.52
	Writing	16	1,028	11.30	3.06	0.62	1.89	1.41

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. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	5,027	32.20	6.09	0.83	2.48	1.61
	Cell Structure and Function	8	5,027	6.66	1.52	0.51	1.05	1.66
	Mechanisms of Genetics	8	5,027	6.22	1.57	0.38	1.22	1.55
	Biological Evolution and Classification	8	5,027	6.19	1.58	0.45	1.16	1.55
	Biological Processes and Systems	8	5,027	6.85	1.46	0.54	0.98	1.71
	Interdependence within Environmental Systems	8	5,027	6.29	1.62	0.49	1.15	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.4.50. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	1,698	32.07	6.28	0.84	2.47	1.60
	Cell Structure and Function	8	1,698	6.63	1.56	0.55	1.04	1.66
	Mechanisms of Genetics	8	1,698	6.17	1.59	0.40	1.23	1.54
	Biological Evolution and Classification	8	1,698	6.15	1.60	0.46	1.17	1.54
	Biological Processes and Systems	8	1,698	6.86	1.49	0.57	0.97	1.72
	Interdependence within Environmental Systems	8	1,698	6.26	1.64	0.52	1.14	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.4.51. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	3,329	32.27	5.99	0.82	2.48	1.61
	Cell Structure and Function	8	3,329	6.67	1.50	0.49	1.06	1.67
	Mechanisms of Genetics	8	3,329	6.24	1.55	0.37	1.22	1.56
	Biological Evolution and Classification	8	3,329	6.21	1.57	0.44	1.16	1.55
	Biological Processes and Systems	8	3,329	6.84	1.45	0.52	0.98	1.71
	Interdependence within Environmental Systems	8	3,329	6.31	1.61	0.48	1.15	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.4.52. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Black or African American

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	866	32.27	5.76	0.82	2.45	1.61
	Cell Structure and Function	8	866	6.67	1.52	0.55	1.02	1.67
	Mechanisms of Genetics	8	866	6.20	1.52	0.36	1.22	1.55
	Biological Evolution and Classification	8	866	6.22	1.53	0.43	1.15	1.55
	Biological Processes and Systems	8	866	6.88	1.42	0.55	0.96	1.72
	Interdependence within Environmental Systems	8	866	6.30	1.52	0.43	1.15	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.4.53. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	2,697	32.51	5.92	0.83	2.43	1.63
	Cell Structure and Function	8	2,697	6.72	1.47	0.49	1.04	1.68
	Mechanisms of Genetics	8	2,697	6.30	1.54	0.39	1.20	1.58
	Biological Evolution and Classification	8	2,697	6.25	1.54	0.43	1.15	1.56
	Biological Processes and Systems	8	2,697	6.88	1.43	0.54	0.96	1.72
	Interdependence within Environmental Systems	8	2,697	6.36	1.59	0.50	1.12	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.4.54. 2024 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
Biology	OVERALL TEST	40	1,082	31.92	6.51	0.84	2.52	1.60
	Cell Structure and Function	8	1,082	6.60	1.61	0.55	1.06	1.65
	Mechanisms of Genetics	8	1,082	6.10	1.63	0.39	1.25	1.52
	Biological Evolution and Classification	8	1,082	6.16	1.66	0.47	1.18	1.54
	Biological Processes and Systems	8	1,082	6.83	1.51	0.53	0.99	1.71
	Interdependence within Environmental Systems	8	1,082	6.23	1.71	0.50	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.4.55. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	4,475	32.28	5.72	0.81	2.46	1.61
	History	8	4,475	5.99	1.55	0.37	1.23	1.50
	Geography and Culture	8	4,475	6.61	1.38	0.42	1.05	1.65
	Government and Citizenship	8	4,475	6.40	1.51	0.40	1.16	1.60
	Economics, Science, Technology, and Society	16	4,475	13.28	2.56	0.65	1.50	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.4.56. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	1,540	32.19	5.64	0.81	2.46	1.61
	History	8	1,540	5.96	1.54	0.35	1.24	1.49
	Geography and Culture	8	1,540	6.57	1.39	0.44	1.04	1.64
	Government and Citizenship	8	1,540	6.44	1.50	0.42	1.15	1.61
	Economics, Science, Technology, and Society	16	1,540	13.22	2.51	0.64	1.50	1.65

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. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	2,935	32.33	5.77	0.82	2.46	1.62
	History	8	2,935	6.01	1.55	0.37	1.22	1.50
	Geography and Culture	8	2,935	6.64	1.37	0.41	1.05	1.66
	Government and Citizenship	8	2,935	6.37	1.51	0.40	1.17	1.59
	Economics, Science, Technology, and Society	16	2,935	13.31	2.58	0.66	1.50	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.4.58. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	808	32.26	5.60	0.81	2.46	1.61
	History	8	808	6.03	1.49	0.33	1.22	1.51
	Geography and Culture	8	808	6.59	1.30	0.33	1.06	1.65
	Government and Citizenship	8	808	6.36	1.52	0.41	1.17	1.59
	Economics, Science, Technology, and Society	16	808	13.28	2.58	0.67	1.49	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.4.59. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability 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	2,335	32.65	5.79	0.83	2.39	1.63
	History	8	2,335	6.09	1.56	0.40	1.20	1.52
	Geography and Culture	8	2,335	6.66	1.41	0.48	1.01	1.67
	Government and Citizenship	8	2,335	6.49	1.49	0.42	1.13	1.62
	Economics, Science, Technology, and Society	16	2,335	13.41	2.54	0.67	1.45	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.4.60. 2024 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean Item Score[3]
U.S. History	OVERALL TEST	40	1,004	31.84	5.62	0.80	2.53	1.59
	History	8	1,004	5.82	1.53	0.32	1.26	1.46
	Geography and Culture	8	1,004	6.61	1.34	0.36	1.07	1.65
	Government and Citizenship	8	1,004	6.26	1.52	0.36	1.21	1.57
	Economics, Science, Technology, and Society	16	1,004	13.15	2.57	0.64	1.54	1.64

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.

2024 STAAR Alternate 2
Progress Measure

Table C.5.1. 2024 STAAR Alternate 2 All Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	373	87	32	79	96	65	14	0	119	254
		23%	9%	21%	26%	17%	4%	0%	32%	68%
Stage 2	207	20	21	50	67	41	8	20	71	116
		10%	10%	24%	32%	20%	4%	10%	34%	56%
Stage 3	1,456	63	61	293	545	410	84	124	838	494
		4%	4%	20%	37%	28%	6%	9%	58%	34%
Stage 4	5,694	91	59	479	2,010	2,451	604	629	4,461	604
		2%	1%	8%	35%	43%	11%	11%	78%	11%
Stage 5	13,355	63	38	376	2,662	7,095	3,121	3,139	10,216	0
		0%	0%	3%	20%	53%	23%	24%	76%	0%
Stage 6	9,050	12	7	96	769	3,139	5,027	4,023	0	5,027
		0%	0%	1%	8%	35%	56%	44%	0%	56%
2023 Total*	30,135	336	218	1,373	6,149	13,201	8,858	7,935	15,705	6,495
		1%	1%	5%	20%	44%	29%	26%	52%	22%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.2. 2024 STAAR Alternate 2 All RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	537	168	48	120	120	58	23	0	216	321
		31%	9%	22%	22%	11%	4%	0%	40%	60%
Stage 2	440	35	52	133	147	56	17	35	185	220
		8%	12%	30%	33%	13%	4%	8%	42%	50%
Stage 3	2,623	114	96	636	986	606	185	210	1,622	791
		4%	4%	24%	38%	23%	7%	8%	62%	30%
Stage 4	8,031	107	98	1,028	3,003	2,817	978	1,233	5,820	978
		1%	1%	13%	37%	35%	12%	15%	72%	12%
Stage 5	14,119	68	65	740	3,335	6,415	3,496	4,208	9,911	0
		0%	0%	5%	24%	45%	25%	30%	70%	0%
Stage 6	8,834	20	22	175	946	2,908	4,763	4,071	0	4,763
		0%	0%	2%	11%	33%	54%	46%	0%	54%
2023 Total*	34,584	512	381	2,832	8,537	12,860	9,462	9,757	17,754	7,073
		1%	1%	8%	25%	37%	27%	28%	51%	20%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.3. 2024 STAAR Alternate 2 Grade 4 Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	123	31	16	26	29	20	1	0	47	76
		25%	13%	21%	24%	16%	1%	0%	38%	62%
Stage 2	61	8	7	17	15	13	1	8	24	29
		13%	11%	28%	25%	21%	2%	13%	39%	48%
Stage 3	394	13	13	97	146	116	9	26	243	125
		3%	3%	25%	37%	29%	2%	7%	62%	32%
Stage 4	1,136	10	12	99	393	576	46	121	969	46
		1%	1%	9%	35%	51%	4%	11%	85%	4%
Stage 5	2,789	9	10	63	462	1,913	332	544	2,245	0
		0%	0%	2%	17%	69%	12%	20%	80%	0%
Stage 6	1,253	1	1	11	60	718	462	791	0	462
		0%	0%	1%	5%	57%	37%	63%	0%	37%
2023 Total*	5,756	72	59	313	1,105	3,356	851	1,490	3,528	738
		1%	1%	5%	19%	58%	15%	26%	61%	13%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.4. 2024 STAAR Alternate 2 Grade 5 Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	62	19	3	18	12	7	3	0	22	40
		31%	5%	29%	19%	11%	5%	0%	35%	65%
Stage 2	29	3	1	9	11	4	1	3	10	16
		10%	3%	31%	38%	14%	3%	10%	34%	55%
Stage 3	268	17	14	65	88	70	14	31	153	84
		6%	5%	24%	33%	26%	5%	12%	57%	31%
Stage 4	950	10	11	100	337	426	66	121	763	66
		1%	1%	11%	35%	45%	7%	13%	80%	7%
Stage 5	3,199	14	3	102	563	1,796	721	682	2,517	0
		0%	0%	3%	18%	56%	23%	21%	79%	0%
Stage 6	942	1	1	9	40	355	536	406	0	536
		0%	0%	1%	4%	38%	57%	43%	0%	57%
2023 Total*	5,450	64	33	303	1,051	2,658	1,341	1,243	3,465	742
		1%	1%	6%	19%	49%	25%	23%	64%	14%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.5. 2024 STAAR Alternate 2 Grade 6 Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	72	13	4	10	22	15	8	0	17	55
		18%	6%	14%	31%	21%	11%	0%	24%	76%
Stage 2	25	1	3	3	10	7	1	1	6	18
		4%	12%	12%	40%	28%	4%	4%	24%	72%
Stage 3	308	13	11	43	109	112	20	24	152	132
		4%	4%	14%	35%	36%	6%	8%	49%	43%
Stage 4	1,113	26	17	88	340	482	160	131	822	160
		2%	2%	8%	31%	43%	14%	12%	74%	14%
Stage 5	2,179	16	10	53	395	1,032	673	474	1,705	0
		1%	0%	2%	18%	47%	31%	22%	78%	0%
Stage 6	1,312	1	0	9	100	430	772	540	0	772
		0%	0%	1%	8%	33%	59%	41%	0%	59%
2023 Total*	5,009	70	45	206	976	2,078	1,634	1,170	2,702	1,137
		1%	1%	4%	19%	41%	33%	23%	54%	23%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.6. 2024 STAAR Alternate 2 Grade 7 Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	46	8	3	10	14	10	1	0	11	35
		17%	7%	22%	30%	22%	2%	0%	24%	76%
Stage 2	33	2	4	10	7	10	0	2	14	17
		6%	12%	30%	21%	30%	0%	6%	42%	52%
Stage 3	180	4	8	27	73	56	12	12	100	68
		2%	4%	15%	41%	31%	7%	7%	56%	38%
Stage 4	768	12	7	58	238	391	62	77	629	62
		2%	1%	8%	31%	51%	8%	10%	82%	8%
Stage 5	1,851	6	3	43	304	1,063	432	356	1,495	0
		0%	0%	2%	16%	57%	23%	19%	81%	0%
Stage 6	1,934	1	2	21	115	673	1,122	812	0	1,122
		0%	0%	1%	6%	35%	58%	42%	0%	58%
2023 Total*	4,812	33	27	169	751	2,203	1,629	1,259	2,249	1,304
		1%	1%	4%	16%	46%	34%	26%	47%	27%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.7. 2024 STAAR Alternate 2 Grade 8 Mathematics Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	37	7	5	10	6	9	0	0	12	25
		19%	14%	27%	16%	24%	0%	0%	32%	68%
Stage 2	28	5	4	6	11	1	1	5	10	13
		18%	14%	21%	39%	4%	4%	18%	36%	46%
Stage 3	166	6	8	39	77	26	10	14	116	36
		4%	5%	23%	46%	16%	6%	8%	70%	22%
Stage 4	814	7	5	62	339	311	90	74	650	90
		1%	1%	8%	42%	38%	11%	9%	80%	11%
Stage 5	1,830	7	2	45	446	822	508	500	1,330	0
		0%	0%	2%	24%	45%	28%	27%	73%	0%
Stage 6	1,758	3	1	10	134	470	1,140	618	0	1,140
		0%	0%	1%	8%	27%	65%	35%	0%	65%
2023 Total*	4,633	35	25	172	1,013	1,639	1,749	1,211	2,118	1,304
		1%	1%	4%	22%	35%	38%	26%	46%	28%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.8. 2024 STAAR Alternate 2 EOC Assessments Algebra I Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	33	9	1	5	13	4	1	0	10	23
		27%	3%	15%	39%	12%	3%	0%	30%	70%
Stage 2	31	1	2	5	13	6	4	1	7	23
		3%	6%	16%	42%	19%	13%	3%	23%	74%
Stage 3	140	10	7	22	52	30	19	17	74	49
		7%	5%	16%	37%	21%	14%	12%	53%	35%
Stage 4	913	26	7	72	363	265	180	105	628	180
		3%	1%	8%	40%	29%	20%	12%	69%	20%
Stage 5	1,507	11	10	70	492	469	455	583	924	0
		1%	1%	5%	33%	31%	30%	39%	61%	0%
Stage 6	1,851	5	2	36	320	493	995	856	0	995
		0%	0%	2%	17%	27%	54%	46%	0%	54%
2023 Total*	4,475	62	29	210	1,253	1,267	1,654	1,562	1,643	1,270
		1%	1%	5%	28%	28%	37%	35%	37%	28%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.9. 2024 STAAR Alternate 2 Grade 4 RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	150	50	16	36	29	12	7	0	66	84
		33%	11%	24%	19%	8%	5%	0%	44%	56%
Stage 2	102	9	12	33	30	16	2	9	45	48
		9%	12%	32%	29%	16%	2%	9%	44%	47%
Stage 3	460	26	23	110	176	110	15	49	286	125
		6%	5%	24%	38%	24%	3%	11%	62%	27%
Stage 4	1,394	22	18	186	577	516	75	226	1,093	75
		2%	1%	13%	41%	37%	5%	16%	78%	5%
Stage 5	2,842	13	12	120	655	1,576	466	800	2,042	0
		0%	0%	4%	23%	55%	16%	28%	72%	0%
Stage 6	808	0	2	9	84	373	340	468	0	340
		0%	0%	1%	10%	46%	42%	58%	0%	42%
2023 Total*	5,756	120	83	494	1,551	2,603	905	1,552	3,532	672
		2%	1%	9%	27%	45%	16%	27%	61%	12%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.10. 2024 STAAR Alternate 2 Grade 5 RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	89	33	10	24	13	8	1	0	43	46
		37%	11%	27%	15%	9%	1%	0%	48%	52%
Stage 2	74	5	14	29	20	3	3	5	43	26
		7%	19%	39%	27%	4%	4%	7%	58%	35%
Stage 3	499	16	27	176	176	86	18	43	352	104
		3%	5%	35%	35%	17%	4%	9%	71%	21%
Stage 4	1,217	10	17	254	445	416	75	281	861	75
		1%	1%	21%	37%	34%	6%	23%	71%	6%
Stage 5	2,545	14	15	190	629	1,220	477	848	1,697	0
		1%	1%	7%	25%	48%	19%	33%	67%	0%
Stage 6	1,023	2	5	21	119	411	465	558	0	465
		0%	0%	2%	12%	40%	45%	55%	0%	45%
2023 Total*	5,447	80	88	694	1,402	2,144	1,039	1,735	2,996	716
		1%	2%	13%	26%	39%	19%	32%	55%	13%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.11. 2024 STAAR Alternate 2 Grade 6 RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	93	15	9	14	38	12	5	0	24	69
		16%	10%	15%	41%	13%	5%	0%	26%	74%
Stage 2	81	6	5	21	30	14	5	6	26	49
		7%	6%	26%	37%	17%	6%	7%	32%	60%
Stage 3	469	17	12	77	200	119	44	29	277	163
		4%	3%	16%	43%	25%	9%	6%	59%	35%
Stage 4	1,303	18	22	136	555	406	166	176	961	166
		1%	2%	10%	43%	31%	13%	14%	74%	13%
Stage 5	1,877	13	12	112	559	665	516	696	1,181	0
		1%	1%	6%	30%	35%	27%	37%	63%	0%
Stage 6	1,189	8	2	34	192	363	590	599	0	590
		1%	0%	3%	16%	31%	50%	50%	0%	50%
2023 Total*	5,012	77	62	394	1,574	1,579	1,326	1,506	2,469	1,037
		2%	1%	8%	31%	32%	26%	30%	49%	21%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.12. 2024 STAAR Alternate 2 Grade 7 RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	52	13	3	18	12	6	0	0	16	36
		25%	6%	35%	23%	12%	0%	0%	31%	69%
Stage 2	47	5	4	12	22	4	0	5	16	26
		11%	9%	26%	47%	9%	0%	11%	34%	55%
Stage 3	429	14	13	111	177	86	28	27	288	114
		3%	3%	26%	41%	20%	7%	6%	67%	27%
Stage 4	1,310	21	19	148	512	498	112	188	1,010	112
		2%	1%	11%	39%	38%	9%	14%	77%	9%
Stage 5	1,650	5	7	76	387	803	372	475	1,175	0
		0%	0%	5%	23%	49%	23%	29%	71%	0%
Stage 6	1,325	0	3	33	155	496	638	687	0	638
		0%	0%	2%	12%	37%	48%	52%	0%	48%
2023 Total*	4,813	58	49	398	1,265	1,893	1,150	1,382	2,505	926
		1%	1%	8%	26%	39%	24%	29%	52%	19%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.13. 2024 STAAR Alternate 2 Grade 8 RLA Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	49	14	6	13	6	9	1	0	20	29
		29%	12%	27%	12%	18%	2%	0%	41%	59%
Stage 2	64	4	11	19	20	7	3	4	30	30
		6%	17%	30%	31%	11%	5%	6%	47%	47%
Stage 3	255	9	10	65	105	47	19	19	170	66
		4%	4%	25%	41%	18%	7%	7%	67%	26%
Stage 4	1,034	9	7	155	439	318	106	171	757	106
		1%	1%	15%	42%	31%	10%	17%	73%	10%
Stage 5	2,043	8	8	99	545	813	570	660	1,383	0
		0%	0%	5%	27%	40%	28%	32%	68%	0%
Stage 6	1,193	2	1	19	107	328	736	457	0	736
		0%	0%	2%	9%	27%	62%	38%	0%	62%
2023 Total*	4,638	46	43	370	1,222	1,522	1,435	1,311	2,360	967
		1%	1%	8%	26%	33%	31%	28%	51%	21%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.14. 2024 STAAR Alternate 2 EOC Assessments English I Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	40	12	2	5	11	6	4	0	14	26
		30%	5%	13%	28%	15%	10%	0%	35%	65%
Stage 2	38	4	1	8	15	7	3	4	9	25
		11%	3%	21%	39%	18%	8%	11%	24%	66%
Stage 3	296	17	4	39	80	105	51	21	119	156
		6%	1%	13%	27%	35%	17%	7%	40%	53%
Stage 4	1,070	16	8	70	248	394	334	94	642	334
		1%	1%	7%	23%	37%	31%	9%	60%	31%
Stage 5	1,680	13	7	75	264	622	699	359	1,321	0
		1%	0%	4%	16%	37%	42%	21%	79%	0%
Stage 6	1,392	4	6	30	155	390	807	585	0	807
		0%	0%	2%	11%	28%	58%	42%	0%	58%
2023 Total*	4,516	66	28	227	773	1,524	1,898	1,063	2,105	1,348
		1%	1%	5%	17%	34%	42%	24%	47%	30%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.15. 2024 STAAR Alternate 2 EOC Assessments English II Progress Measure

		2024						2024 Total*		
2023	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	64	31	2	10	11	5	5	0	33	31
		48%	3%	16%	17%	8%	8%	0%	52%	48%
Stage 2	34	2	5	11	10	5	1	2	16	16
		6%	15%	32%	29%	15%	3%	6%	47%	47%
Stage 3	215	15	7	58	72	53	10	22	130	63
		7%	3%	27%	33%	25%	5%	10%	60%	29%
Stage 4	703	11	7	79	227	269	110	97	496	110
		2%	1%	11%	32%	38%	16%	14%	71%	16%
Stage 5	1,482	2	4	68	296	716	396	370	1,112	0
		0%	0%	5%	20%	48%	27%	25%	75%	0%
Stage 6	1,904	4	3	29	134	547	1,187	717	0	1,187
		0%	0%	2%	7%	29%	62%	38%	0%	62%
2023 Total*	4,402	65	28	255	750	1,595	1,709	1,208	1,787	1,407
		1%	1%	6%	17%	36%	39%	27%	41%	32%

Notes:

* Due to rounding, percentages may not sum to 100.

2024 STAAR Alternate 2
Scale Score Descriptive Statistics

Table C.6.1. 2024 STAAR Alternate 2 Grades 3–8
Scale Score Descriptive Statistics

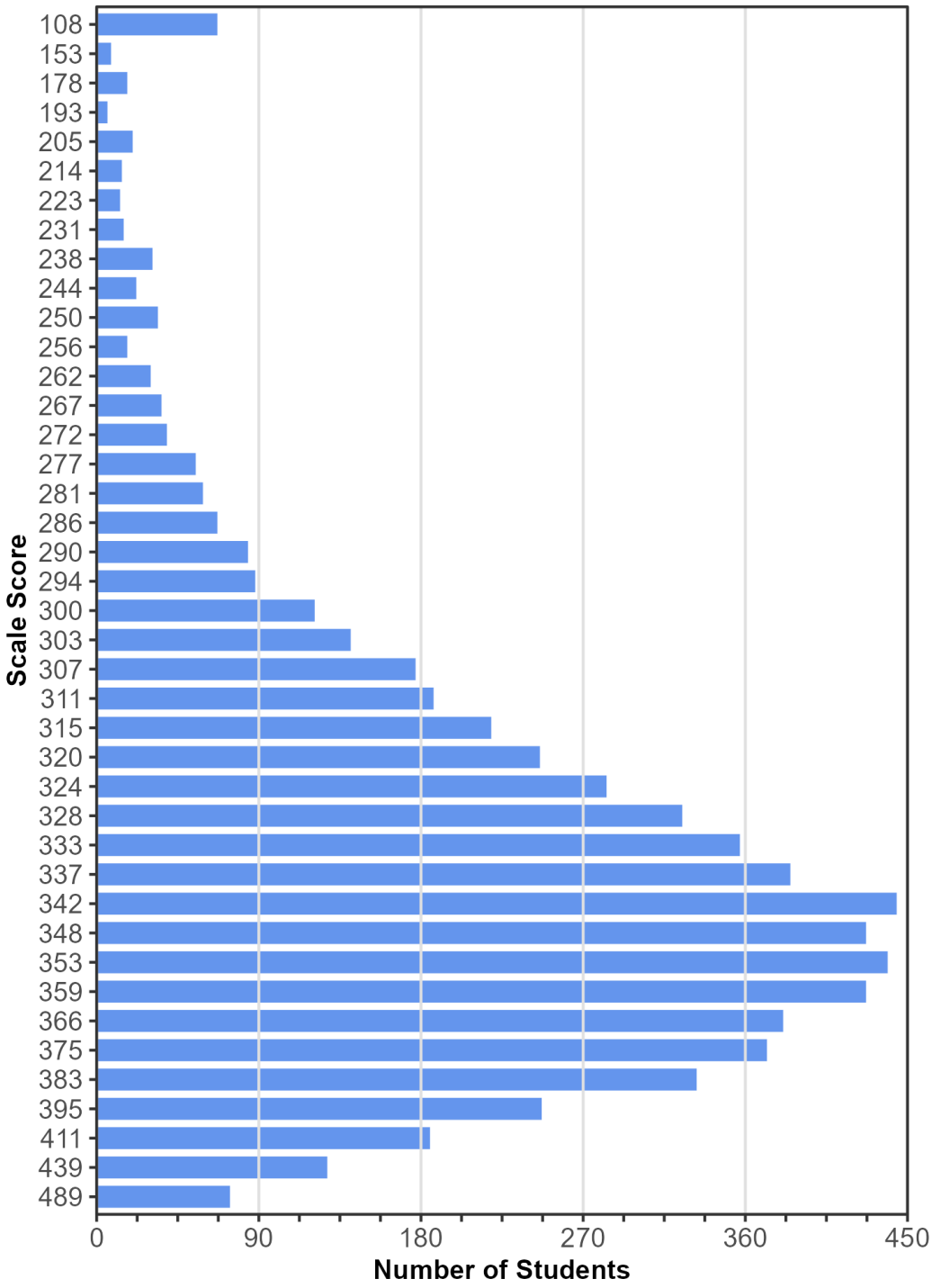
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	6,620	339.62	342	342	381	46	47.87	2291.89	-1.06	5.78
Grade 4	Mathematics	6,490	348.85	347	358	372	45	44.34	1965.64	-0.58	5.06
Grade 5	Mathematics	6,033	352.02	350	350	369	46	48.44	2346.33	-0.07	3.44
Grade 6	Mathematics	5,585	352.77	350	356	408	52	50.64	2564.40	-0.50	5.55
Grade 7	Mathematics	5,225	362.72	356	370	400	55	49.96	2496.13	0.10	3.48
Grade 8	Mathematics	4,984	355.53	352	352	402	47	47.46	2252.26	-0.09	4.49
Grade 3	RLA	6,623	339.38	340	351	432	43	49.66	2466.21	-1.14	6.13
Grade 4	RLA	6,488	338.48	338	344	422	47	49.28	2428.86	-0.68	5.38
Grade 5	RLA	6,032	336.36	337	343	412	53	47.84	2288.93	-0.30	4.84
Grade 6	RLA	5,585	342.17	340	340	420	56	51.16	2617.67	-0.23	4.47
Grade 7	RLA	5,227	349.67	348	355	428	48	53.34	2845.30	0.27	3.38
Grade 8	RLA	4,985	345.75	338	332	414	55	49.75	2475.56	0.13	3.70
Grade 5	Science	6,033	361.69	357	363	385	52	50.71	2571.24	-0.19	3.57
Grade 8	Science	4,982	360.81	359	373	348	44	44.12	1946.47	-0.06	3.65
Grade 8	Social Studies	4,988	354.11	352	372	375	49	44.51	1980.82	-0.25	4.26

Table C.6.2. 2024 STAAR Alternate 2 EOC Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	5,112	345.62	345	339	401	42	47.58	2263.64	-0.41	5.48
English I	5,116	353.21	349	365	426	52	50.44	2544.29	-0.34	5.17
English II	4,683	354.34	352	346	423	59	52.05	2708.79	-0.24	4.23
Biology	5,027	361.83	357	371	350	43	44.17	1950.75	-0.52	4.70
US History	4,475	359.73	354	368	382	50	47.48	2254.10	-0.16	3.81

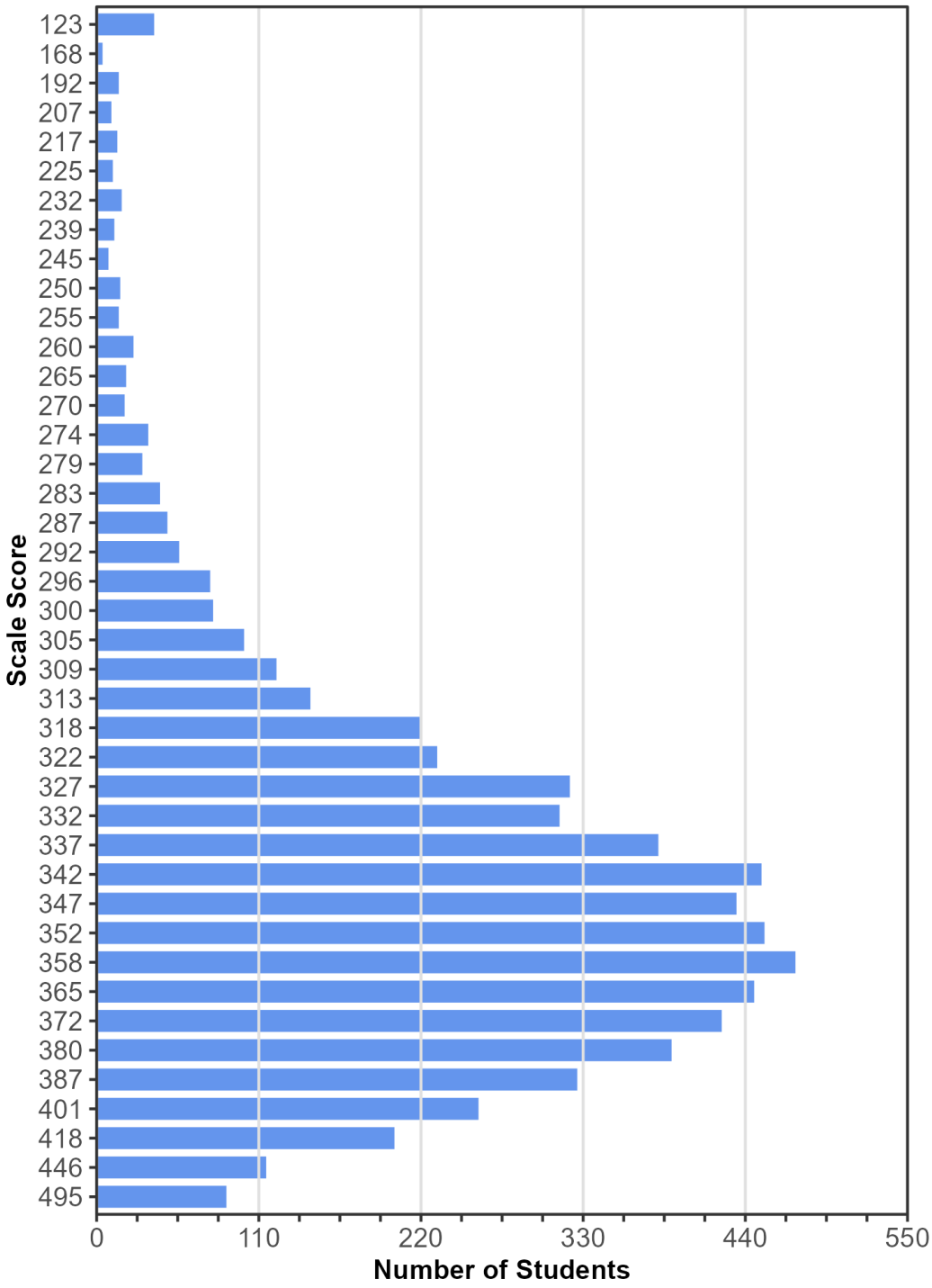
2024 STAAR Alternate 2
Frequency Distribution of Scale Scores

Figure C.6.1. 2024 STAAR Alternate 2 Grade 3 Mathematics
Frequency Distribution of Scale Scores



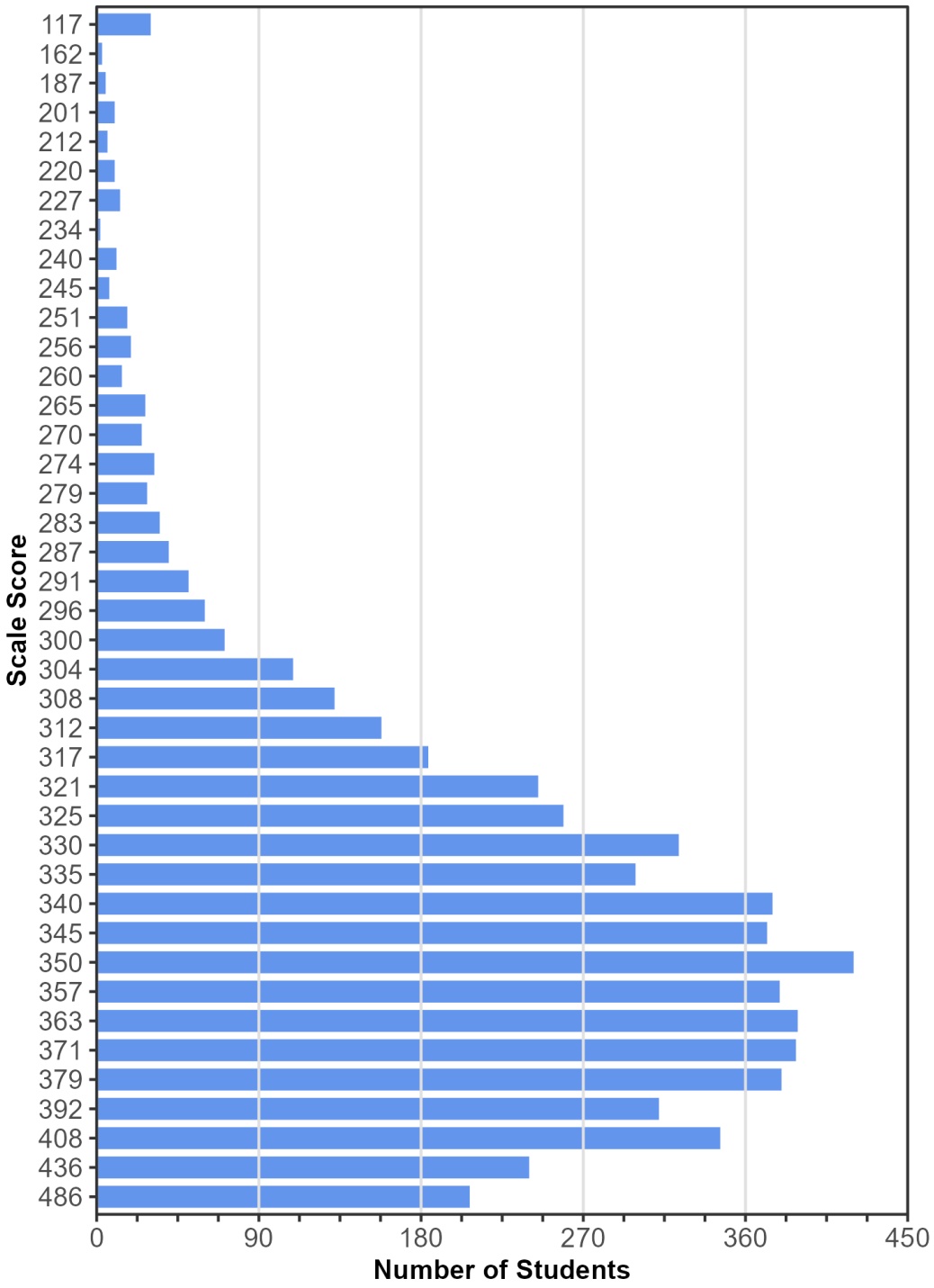
Scale Score	Freq	Cum Freq	Pct	Cum Pct
108	67	67	1.01	1.01
153	8	75	0.12	1.13
178	17	92	0.26	1.39
193	6	98	0.09	1.48
205	20	118	0.30	1.78
214	14	132	0.21	1.99
223	13	145	0.20	2.19
231	15	160	0.23	2.42
238	31	191	0.47	2.89
244	22	213	0.33	3.22
250	34	247	0.51	3.73
256	17	264	0.26	3.99
262	30	294	0.45	4.44
267	36	330	0.54	4.98
272	39	369	0.59	5.57
277	55	424	0.83	6.40
281	59	483	0.89	7.30
286	67	550	1.01	8.31
290	84	634	1.27	9.58
294	88	722	1.33	10.91
300	121	843	1.83	12.73
303	141	984	2.13	14.86
307	177	1161	2.67	17.54
311	187	1348	2.82	20.36
315	219	1567	3.31	23.67
320	246	1813	3.72	27.39
324	283	2096	4.27	31.66
328	325	2421	4.91	36.57
333	357	2778	5.39	41.96
337	385	3163	5.82	47.78
342	444	3607	6.71	54.49
348	427	4034	6.45	60.94
353	439	4473	6.63	67.57
359	427	4900	6.45	74.02
366	381	5281	5.76	79.77
375	372	5653	5.62	85.39
383	333	5986	5.03	90.42
395	247	6233	3.73	94.15
411	185	6418	2.79	96.95
439	128	6546	1.93	98.88
489	74	6620	1.12	100.00

Figure C.6.2. 2024 STAAR Alternate 2 Grade 4 Mathematics
Frequency Distribution of Scale Scores



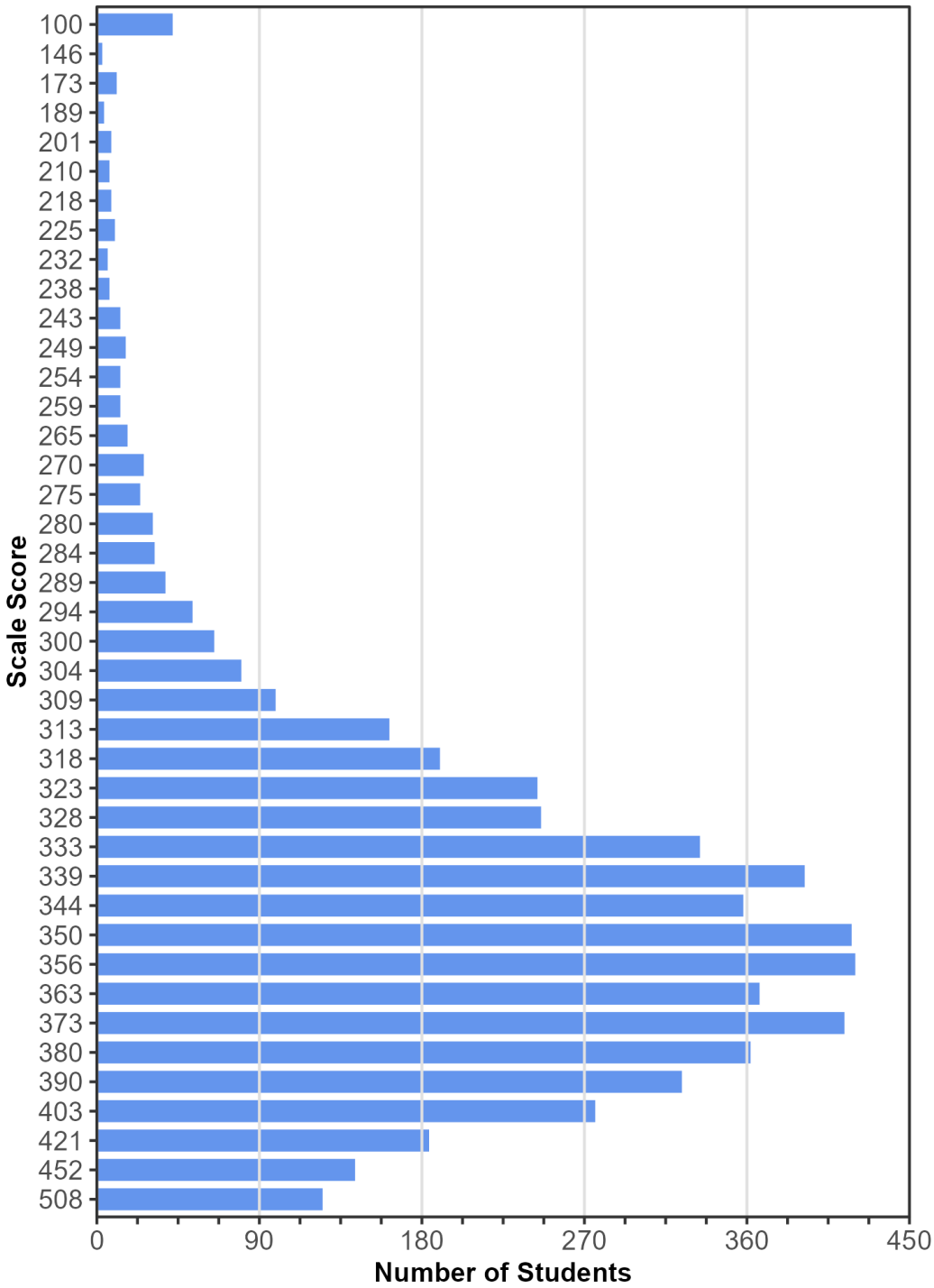
Scale Score	Freq	Cum Freq	Pct	Cum Pct
123	39	39	0.60	0.60
168	4	43	0.06	0.66
192	15	58	0.23	0.89
207	10	68	0.15	1.05
217	14	82	0.22	1.26
225	11	93	0.17	1.43
232	17	110	0.26	1.69
239	12	122	0.18	1.88
245	8	130	0.12	2.00
250	16	146	0.25	2.25
255	15	161	0.23	2.48
260	25	186	0.39	2.87
265	20	206	0.31	3.17
270	19	225	0.29	3.47
274	35	260	0.54	4.01
279	31	291	0.48	4.48
283	43	334	0.66	5.15
287	48	382	0.74	5.89
292	56	438	0.86	6.75
296	77	515	1.19	7.94
300	79	594	1.22	9.15
305	100	694	1.54	10.69
309	122	816	1.88	12.57
313	145	961	2.23	14.81
318	220	1181	3.39	18.20
322	231	1412	3.56	21.76
327	321	1733	4.95	26.70
332	314	2047	4.84	31.54
337	381	2428	5.87	37.41
342	451	2879	6.95	44.36
347	434	3313	6.69	51.05
352	453	3766	6.98	58.03
358	474	4240	7.30	65.33
365	446	4686	6.87	72.20
372	424	5110	6.53	78.74
380	390	5500	6.01	84.75
387	326	5826	5.02	89.77
401	259	6085	3.99	93.76
418	202	6287	3.11	96.87
446	115	6402	1.77	98.64
495	88	6490	1.36	100.00

Figure C.6.3. 2024 STAAR Alternate 2 Grade 5 Mathematics
Frequency Distribution of Scale Scores



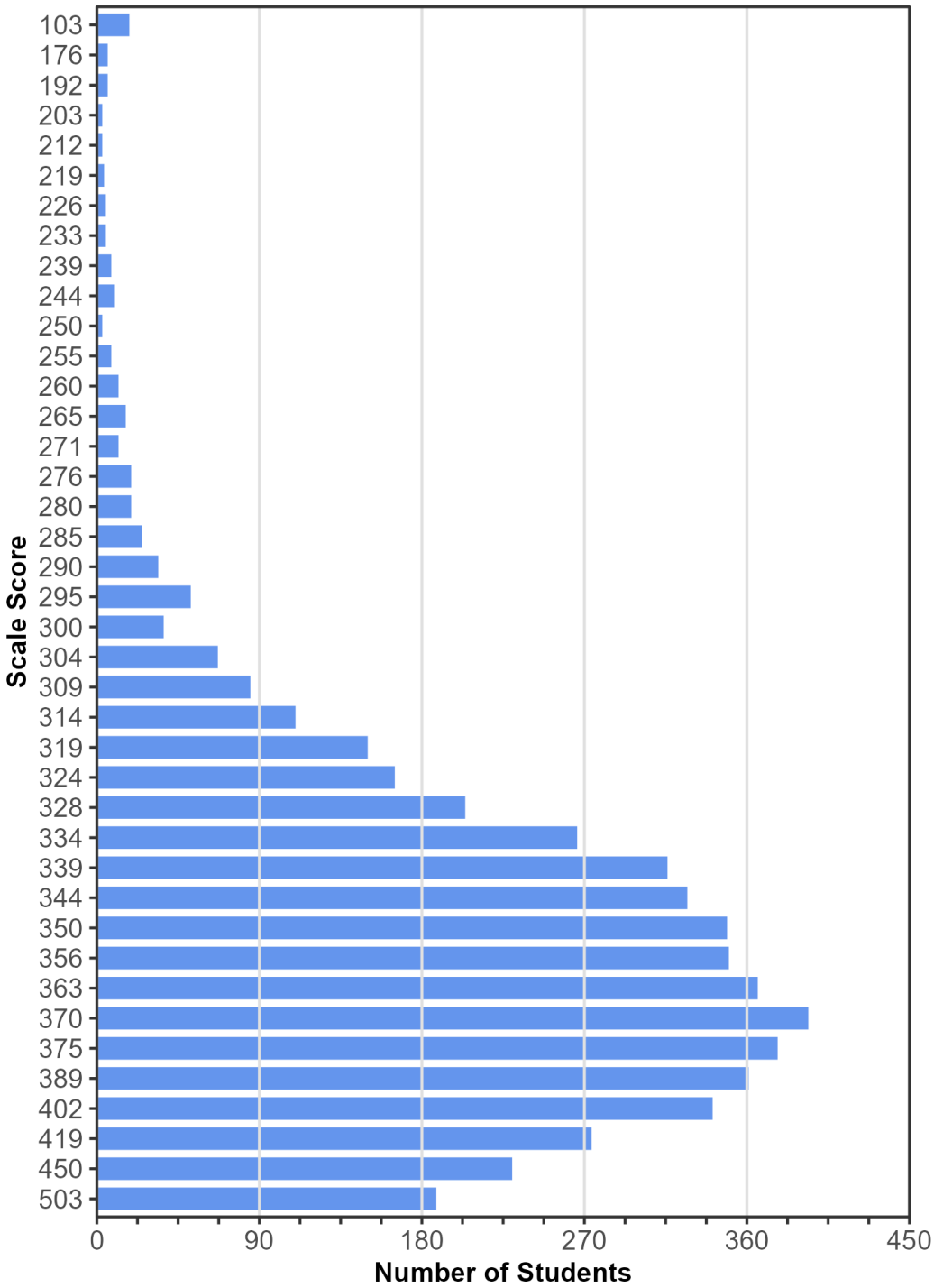
Scale Score	Freq	Cum Freq	Pct	Cum Pct
117	30	30	0.50	0.50
162	3	33	0.05	0.55
187	5	38	0.08	0.63
201	10	48	0.17	0.80
212	6	54	0.10	0.90
220	10	64	0.17	1.06
227	13	77	0.22	1.28
234	2	79	0.03	1.31
240	11	90	0.18	1.49
245	7	97	0.12	1.61
251	17	114	0.28	1.89
256	19	133	0.31	2.20
260	14	147	0.23	2.44
265	27	174	0.45	2.88
270	25	199	0.41	3.30
274	32	231	0.53	3.83
279	28	259	0.46	4.29
283	35	294	0.58	4.87
287	40	334	0.66	5.54
291	51	385	0.85	6.38
296	60	445	0.99	7.38
300	71	516	1.18	8.55
304	109	625	1.81	10.36
308	132	757	2.19	12.55
312	158	915	2.62	15.17
317	184	1099	3.05	18.22
321	245	1344	4.06	22.28
325	259	1603	4.29	26.57
330	323	1926	5.35	31.92
335	299	2225	4.96	36.88
340	375	2600	6.22	43.10
345	372	2972	6.17	49.26
350	420	3392	6.96	56.22
357	379	3771	6.28	62.51
363	389	4160	6.45	68.95
371	388	4548	6.43	75.39
379	380	4928	6.30	81.68
392	312	5240	5.17	86.86
408	346	5586	5.74	92.59
436	240	5826	3.98	96.57
486	207	6033	3.43	100.00

Figure C.6.4. 2024 STAAR Alternate 2 Grade 6 Mathematics
Frequency Distribution of Scale Scores



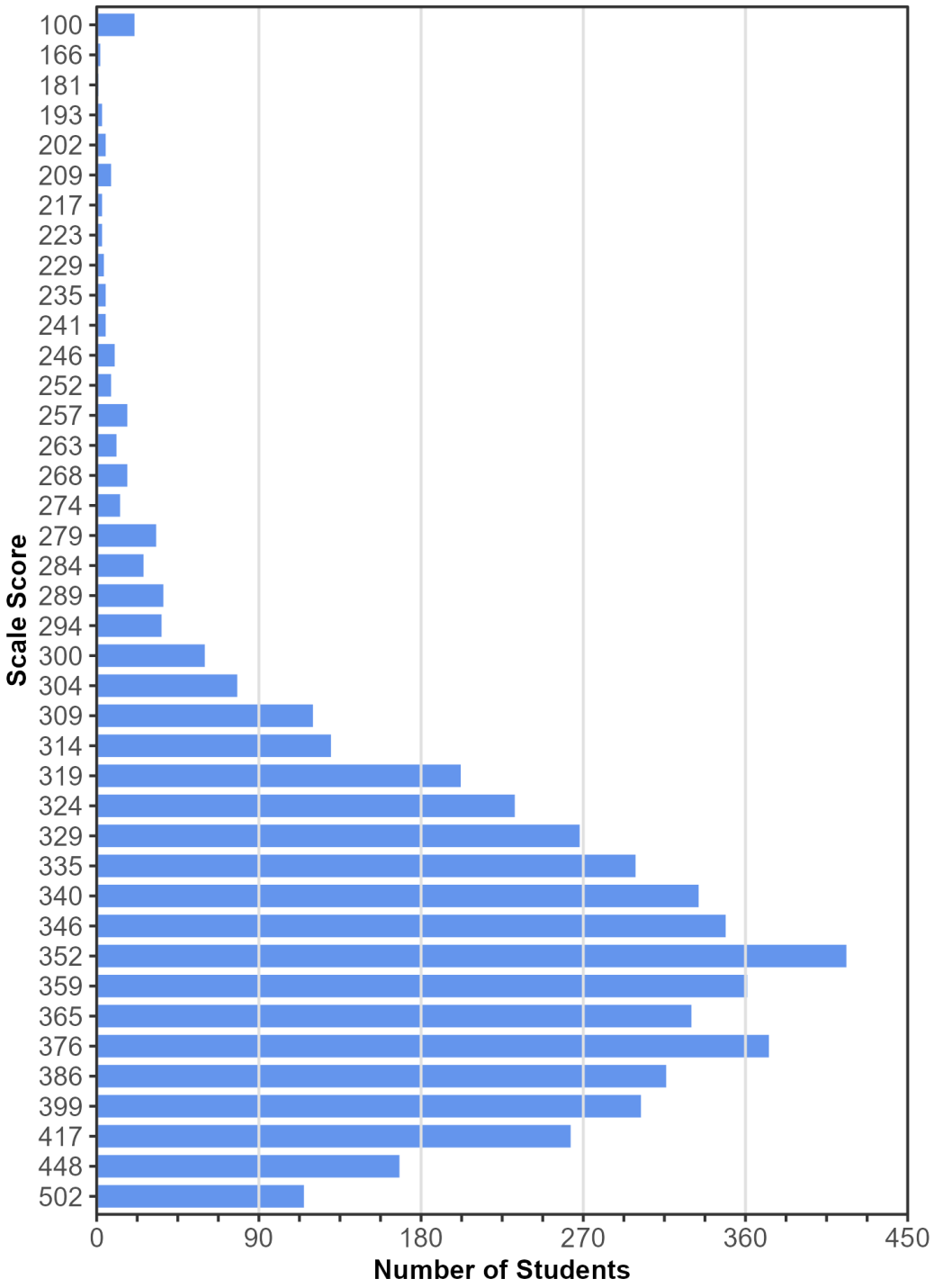
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	42	42	0.75	0.75
146	3	45	0.05	0.81
173	11	56	0.20	1.00
189	4	60	0.07	1.07
201	8	68	0.14	1.22
210	7	75	0.13	1.34
218	8	83	0.14	1.49
225	10	93	0.18	1.67
232	6	99	0.11	1.77
238	7	106	0.13	1.90
243	13	119	0.23	2.13
249	16	135	0.29	2.42
254	13	148	0.23	2.65
259	13	161	0.23	2.88
265	17	178	0.30	3.19
270	26	204	0.47	3.65
275	24	228	0.43	4.08
280	31	259	0.56	4.64
284	32	291	0.57	5.21
289	38	329	0.68	5.89
294	53	382	0.95	6.84
300	65	447	1.16	8.00
304	80	527	1.43	9.44
309	99	626	1.77	11.21
313	162	788	2.90	14.11
318	190	978	3.40	17.51
323	244	1222	4.37	21.88
328	246	1468	4.40	26.28
333	334	1802	5.98	32.26
339	392	2194	7.02	39.28
344	358	2552	6.41	45.69
350	418	2970	7.48	53.18
356	420	3390	7.52	60.70
363	367	3757	6.57	67.27
373	414	4171	7.41	74.68
380	362	4533	6.48	81.16
390	324	4857	5.80	86.97
403	276	5133	4.94	91.91
421	184	5317	3.29	95.20
452	143	5460	2.56	97.76
508	125	5585	2.24	100.00

Figure C.6.5. 2024 STAAR Alternate 2 Grade 7 Mathematics
Frequency Distribution of Scale Scores



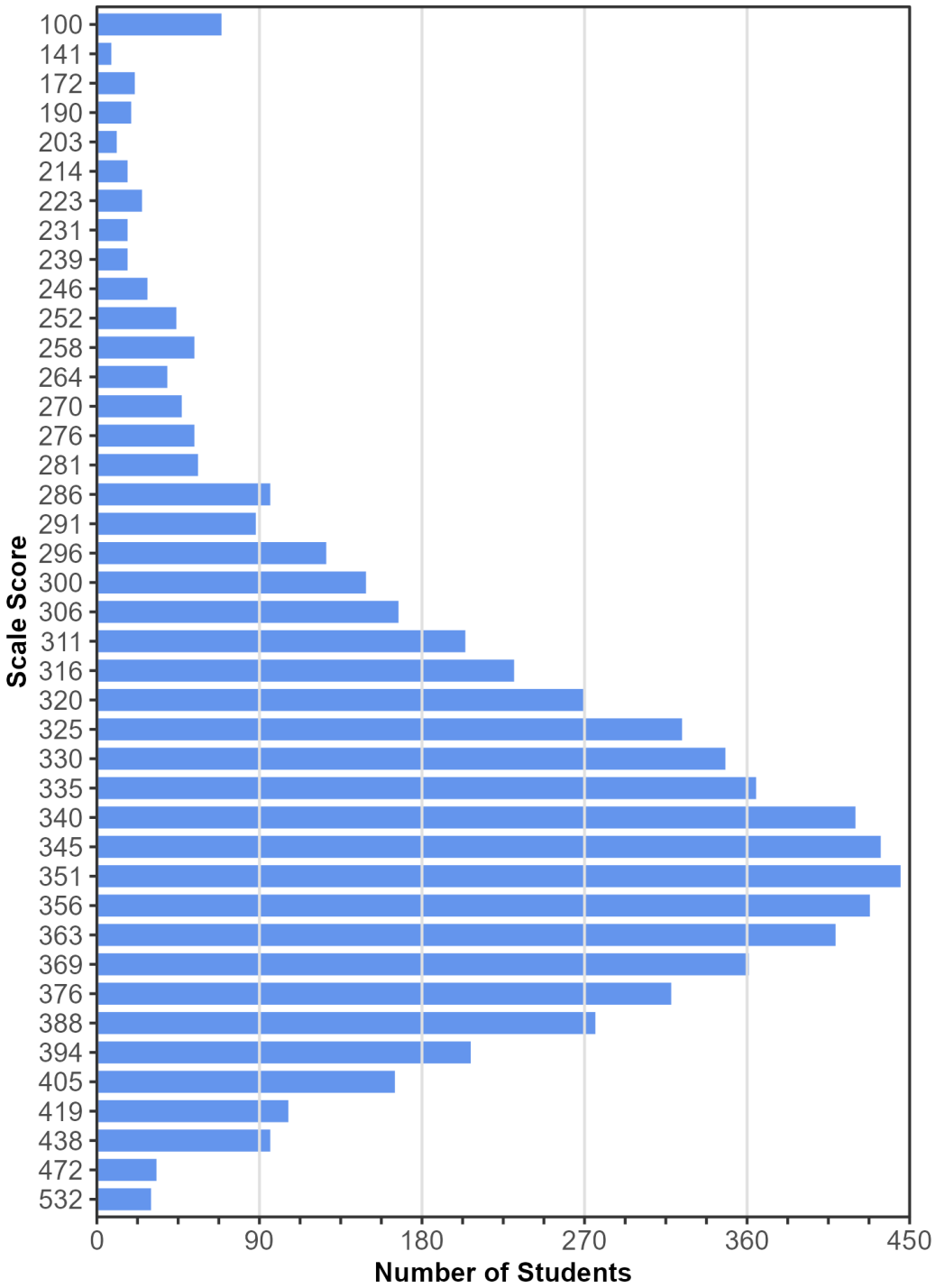
Scale Score	Freq	Cum Freq	Pct	Cum Pct
103	18	18	0.34	0.34
176	6	24	0.11	0.46
192	6	30	0.11	0.57
203	3	33	0.06	0.63
212	3	36	0.06	0.69
219	4	40	0.08	0.77
226	5	45	0.10	0.86
233	5	50	0.10	0.96
239	8	58	0.15	1.11
244	10	68	0.19	1.30
250	3	71	0.06	1.36
255	8	79	0.15	1.51
260	12	91	0.23	1.74
265	16	107	0.31	2.05
271	12	119	0.23	2.28
276	19	138	0.36	2.64
280	19	157	0.36	3.00
285	25	182	0.48	3.48
290	34	216	0.65	4.13
295	52	268	1.00	5.13
300	37	305	0.71	5.84
304	67	372	1.28	7.12
309	85	457	1.63	8.75
314	110	567	2.11	10.85
319	150	717	2.87	13.72
324	165	882	3.16	16.88
328	204	1086	3.90	20.78
334	266	1352	5.09	25.88
339	316	1668	6.05	31.92
344	327	1995	6.26	38.18
350	349	2344	6.68	44.86
356	350	2694	6.70	51.56
363	366	3060	7.00	58.56
370	394	3454	7.54	66.11
375	377	3831	7.22	73.32
389	361	4192	6.91	80.23
402	341	4533	6.53	86.76
419	274	4807	5.24	92.00
450	230	5037	4.40	96.40
503	188	5225	3.60	100.00

Figure C.6.6. 2024 STAAR Alternate 2 Grade 8 Mathematics
Frequency Distribution of Scale Scores



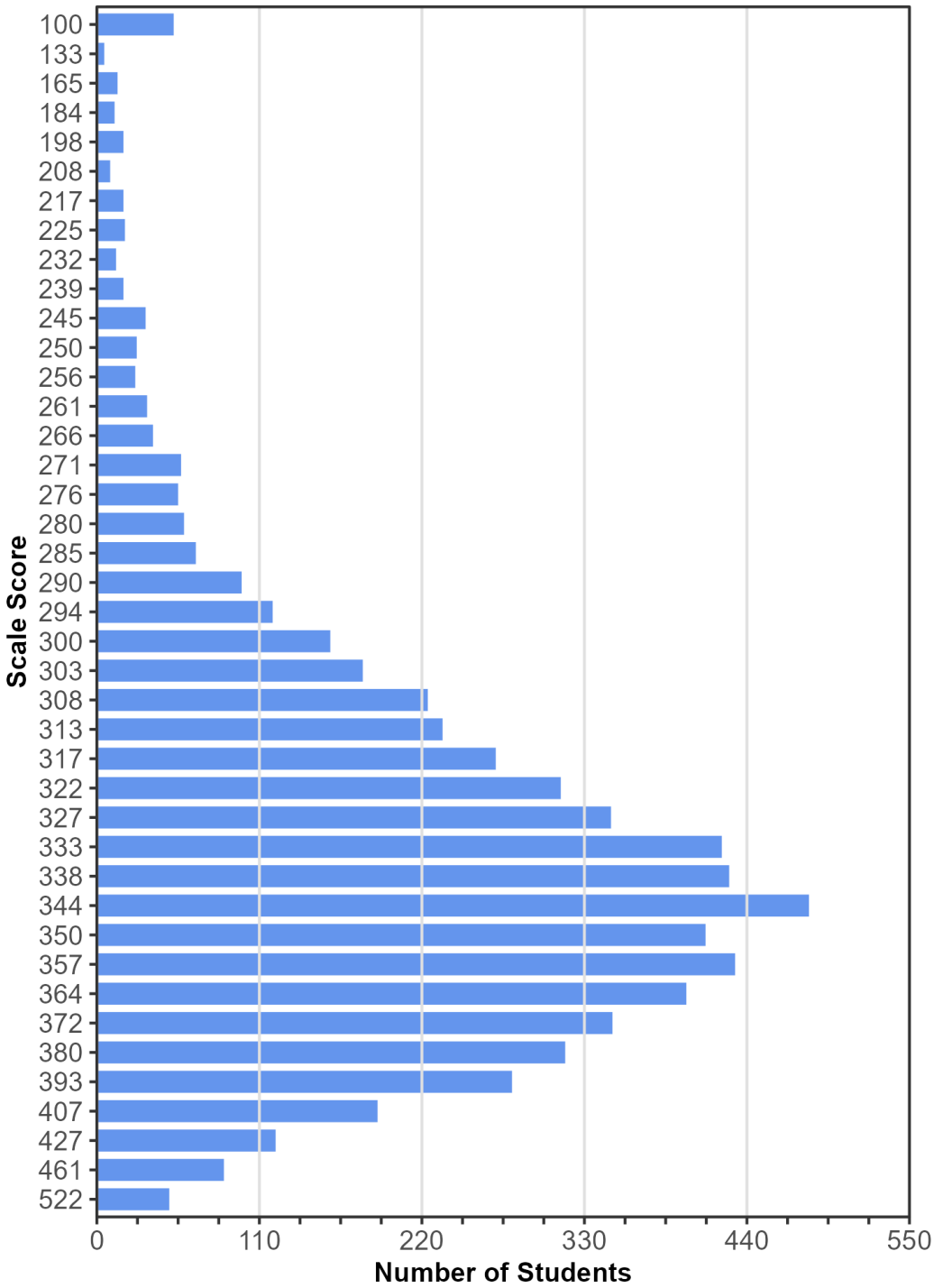
	Freq	Cum Freq	Pct	Cum Pct
100	21	21	0.42	0.42
166	2	23	0.04	0.46
181	1	24	0.02	0.48
193	3	27	0.06	0.54
202	5	32	0.10	0.64
209	8	40	0.16	0.80
217	3	43	0.06	0.86
223	3	46	0.06	0.92
229	4	50	0.08	1.00
235	5	55	0.10	1.10
241	5	60	0.10	1.20
246	10	70	0.20	1.40
252	8	78	0.16	1.57
257	17	95	0.34	1.91
263	11	106	0.22	2.13
268	17	123	0.34	2.47
274	13	136	0.26	2.73
279	33	169	0.66	3.39
284	26	195	0.52	3.91
289	37	232	0.74	4.65
294	36	268	0.72	5.38
300	60	328	1.20	6.58
304	78	406	1.57	8.15
309	120	526	2.41	10.55
314	130	656	2.61	13.16
319	202	858	4.05	17.22
324	232	1090	4.65	21.87
329	268	1358	5.38	27.25
335	299	1657	6.00	33.25
340	334	1991	6.70	39.95
346	349	2340	7.00	46.95
352	416	2756	8.35	55.30
359	361	3117	7.24	62.54
365	330	3447	6.62	69.16
376	373	3820	7.48	76.65
386	316	4136	6.34	82.99
399	302	4438	6.06	89.04
417	263	4701	5.28	94.32
448	168	4869	3.37	97.69
502	115	4984	2.31	100.00

Figure C.6.7. 2024 STAAR Alternate 2 Grade 3 RLA
Frequency Distribution of Scale Scores



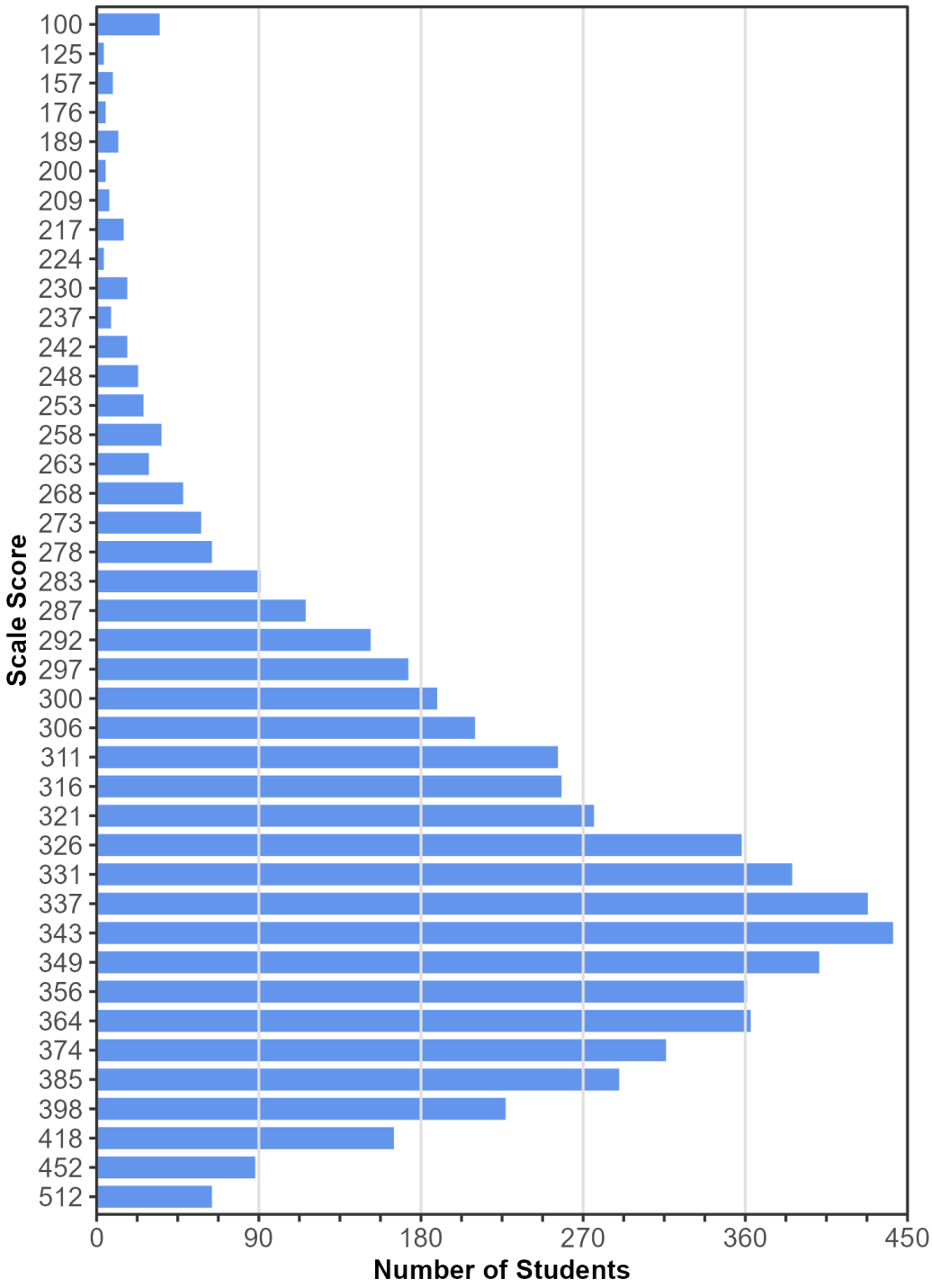
	Freq	Cum Freq	Pct	Cum Pct
100	69	69	1.04	1.04
141	8	77	0.12	1.16
172	21	98	0.32	1.48
190	19	117	0.29	1.77
203	11	128	0.17	1.93
214	17	145	0.26	2.19
223	25	170	0.38	2.57
231	17	187	0.26	2.82
239	17	204	0.26	3.08
246	28	232	0.42	3.50
252	44	276	0.66	4.17
258	54	330	0.82	4.98
264	39	369	0.59	5.57
270	47	416	0.71	6.28
276	54	470	0.82	7.10
281	56	526	0.85	7.94
286	96	622	1.45	9.39
291	88	710	1.33	10.72
296	127	837	1.92	12.64
300	149	986	2.25	14.89
306	167	1153	2.52	17.41
311	204	1357	3.08	20.49
316	231	1588	3.49	23.98
320	270	1858	4.08	28.05
325	324	2182	4.89	32.95
330	348	2530	5.25	38.20
335	365	2895	5.51	43.71
340	420	3315	6.34	50.05
345	434	3749	6.55	56.61
351	445	4194	6.72	63.32
356	428	4622	6.46	69.79
363	409	5031	6.18	75.96
369	361	5392	5.45	81.41
376	318	5710	4.80	86.21
388	276	5986	4.17	90.38
394	207	6193	3.13	93.51
405	165	6358	2.49	96.00
419	106	6464	1.60	97.60
438	96	6560	1.45	99.05
472	33	6593	0.50	99.55
532	30	6623	0.45	100.00

Figure C.6.8. 2024 STAAR Alternate 2 Grade 4 RLA
Frequency Distribution of Scale Scores



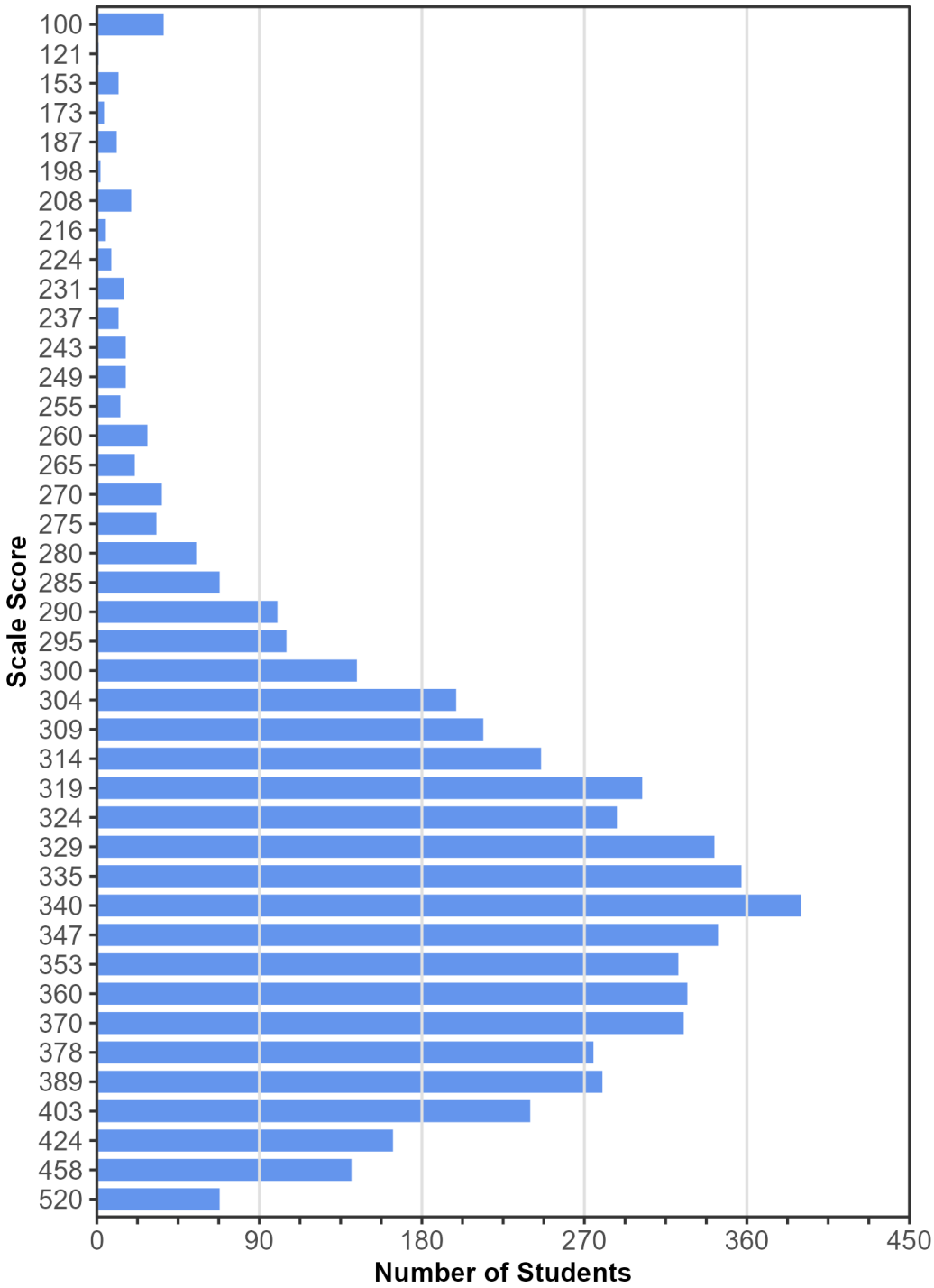
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	52	52	0.80	0.80
133	5	57	0.08	0.88
165	14	71	0.22	1.09
184	12	83	0.18	1.28
198	18	101	0.28	1.56
208	9	110	0.14	1.70
217	18	128	0.28	1.97
225	19	147	0.29	2.27
232	13	160	0.20	2.47
239	18	178	0.28	2.74
245	33	211	0.51	3.25
250	27	238	0.42	3.67
256	26	264	0.40	4.07
261	34	298	0.52	4.59
266	38	336	0.59	5.18
271	57	393	0.88	6.06
276	55	448	0.85	6.91
280	59	507	0.91	7.81
285	67	574	1.03	8.85
290	98	672	1.51	10.36
294	119	791	1.83	12.19
300	158	949	2.44	14.63
303	180	1129	2.77	17.40
308	224	1353	3.45	20.85
313	234	1587	3.61	24.46
317	270	1857	4.16	28.62
322	314	2171	4.84	33.46
327	348	2519	5.36	38.83
333	423	2942	6.52	45.35
338	428	3370	6.60	51.94
344	482	3852	7.43	59.37
350	412	4264	6.35	65.72
357	432	4696	6.66	72.38
364	399	5095	6.15	78.53
372	349	5444	5.38	83.91
380	317	5761	4.89	88.79
393	281	6042	4.33	93.13
407	190	6232	2.93	96.05
427	121	6353	1.86	97.92
461	86	6439	1.33	99.24
522	49	6488	0.76	100.00

Figure C.6.9. 2024 STAAR Alternate 2 Grade 5 RLA
Frequency Distribution of Scale Scores



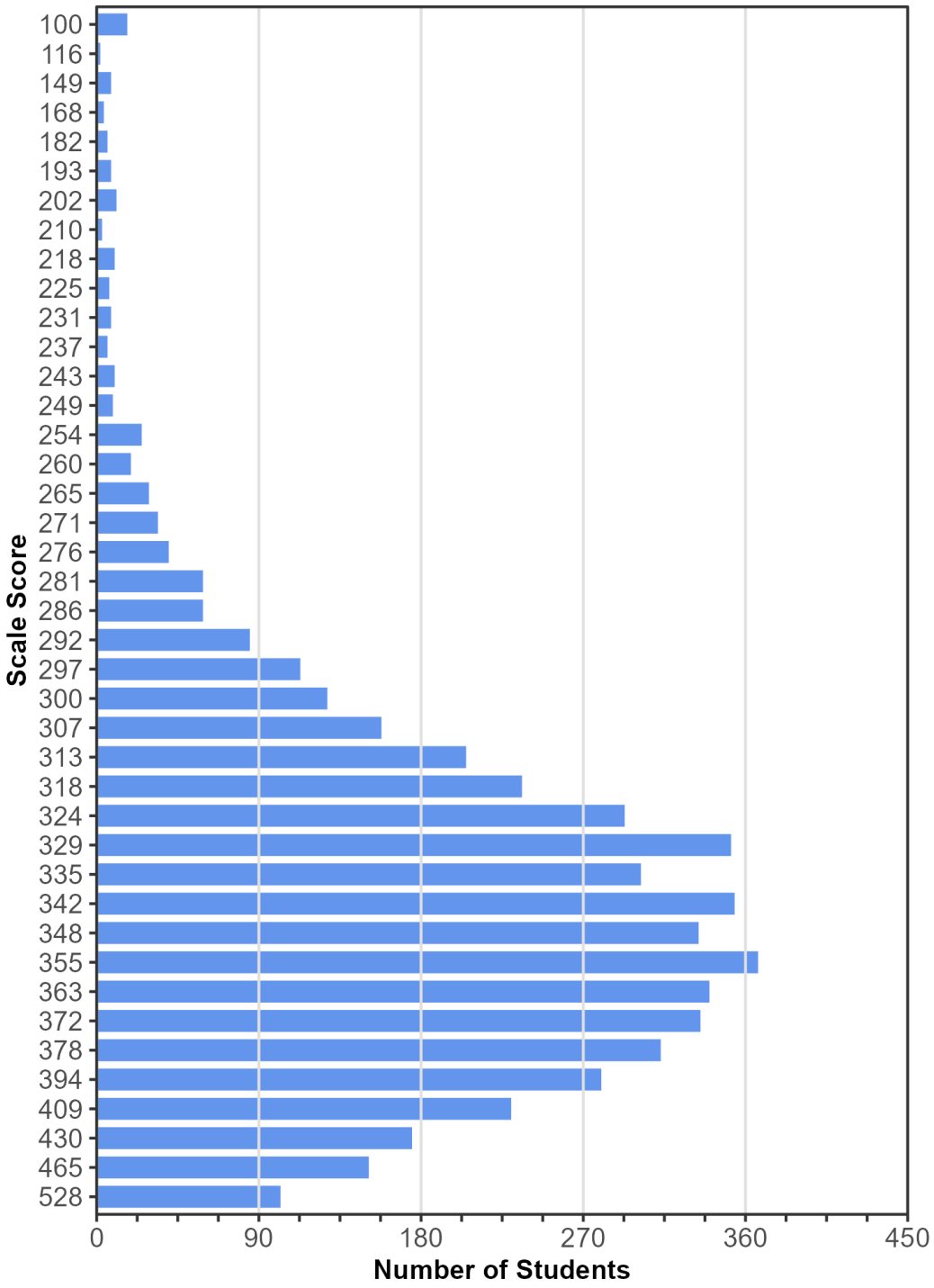
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	35	35	0.58	0.58
125	4	39	0.07	0.65
157	9	48	0.15	0.80
176	5	53	0.08	0.88
189	12	65	0.20	1.08
200	5	70	0.08	1.16
209	7	77	0.12	1.28
217	15	92	0.25	1.53
224	4	96	0.07	1.59
230	17	113	0.28	1.87
237	8	121	0.13	2.01
242	17	138	0.28	2.29
248	23	161	0.38	2.67
253	26	187	0.43	3.10
258	36	223	0.60	3.70
263	29	252	0.48	4.18
268	48	300	0.80	4.97
273	58	358	0.96	5.94
278	64	422	1.06	7.00
283	91	513	1.51	8.50
287	116	629	1.92	10.43
292	152	781	2.52	12.95
297	173	954	2.87	15.82
300	189	1143	3.13	18.95
306	210	1353	3.48	22.43
311	256	1609	4.24	26.67
316	258	1867	4.28	30.95
321	276	2143	4.58	35.53
326	358	2501	5.94	41.46
331	386	2887	6.40	47.86
337	428	3315	7.10	54.96
343	442	3757	7.33	62.28
349	401	4158	6.65	68.93
356	361	4519	5.98	74.92
364	363	4882	6.02	80.94
374	316	5198	5.24	86.17
385	290	5488	4.81	90.98
398	227	5715	3.76	94.74
418	165	5880	2.74	97.48
452	88	5968	1.46	98.94
512	64	6032	1.06	100.00

Figure C.6.10. 2024 STAAR Alternate 2 Grade 6 RLA
Frequency Distribution of Scale Scores



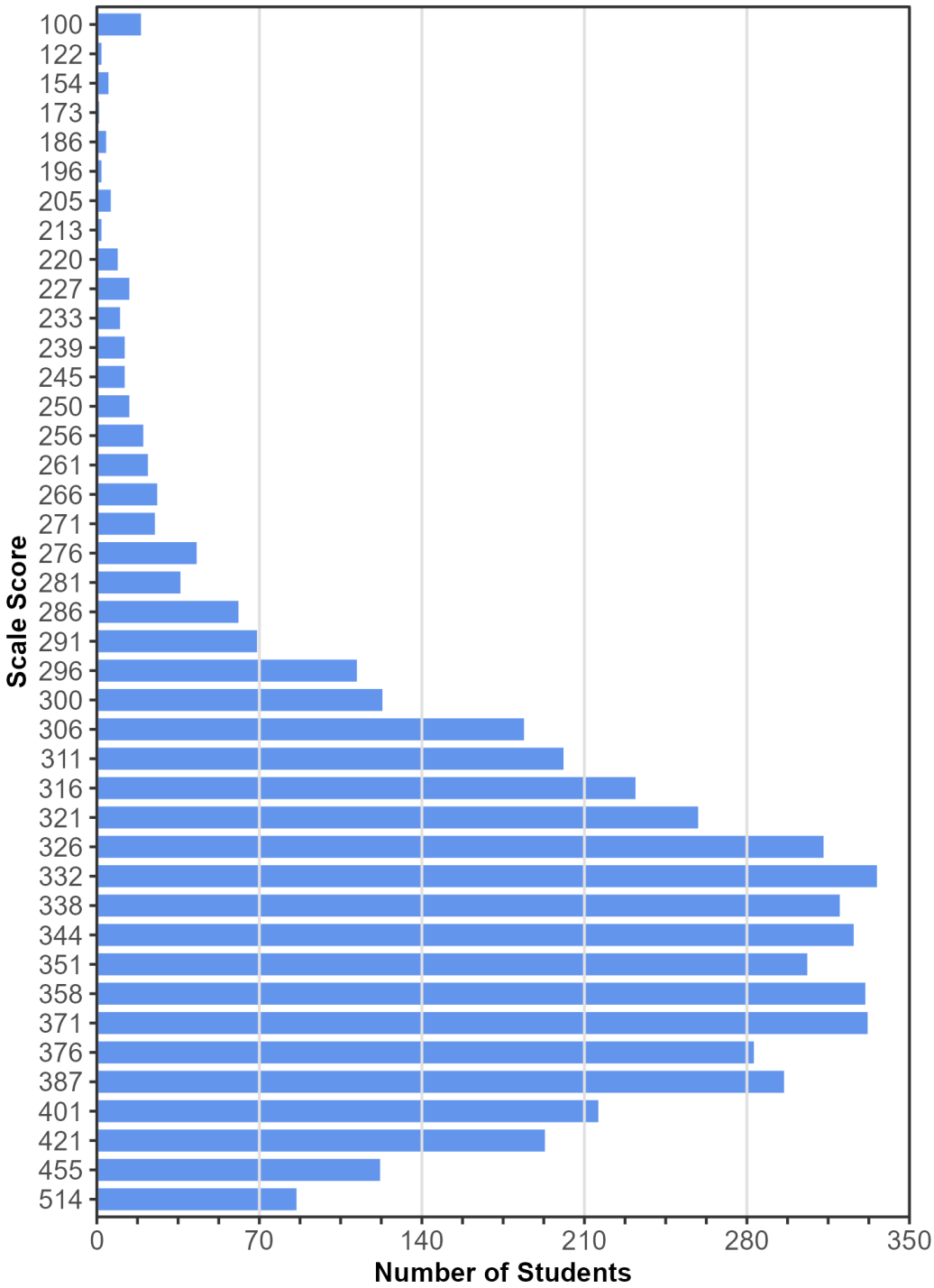
	Freq	Cum Freq	Pct	Cum Pct
100	37	37	0.66	0.66
121	1	38	0.02	0.68
153	12	50	0.21	0.90
173	4	54	0.07	0.97
187	11	65	0.20	1.16
198	2	67	0.04	1.20
208	19	86	0.34	1.54
216	5	91	0.09	1.63
224	8	99	0.14	1.77
231	15	114	0.27	2.04
237	12	126	0.21	2.26
243	16	142	0.29	2.54
249	16	158	0.29	2.83
255	13	171	0.23	3.06
260	28	199	0.50	3.56
265	21	220	0.38	3.94
270	36	256	0.64	4.58
275	33	289	0.59	5.17
280	55	344	0.98	6.16
285	68	412	1.22	7.38
290	100	512	1.79	9.17
295	105	617	1.88	11.05
300	144	761	2.58	13.63
304	199	960	3.56	17.19
309	214	1174	3.83	21.02
314	246	1420	4.40	25.43
319	302	1722	5.41	30.83
324	288	2010	5.16	35.99
329	342	2352	6.12	42.11
335	357	2709	6.39	48.50
340	390	3099	6.98	55.49
347	344	3443	6.16	61.65
353	322	3765	5.77	67.41
360	327	4092	5.85	73.27
370	325	4417	5.82	79.09
378	275	4692	4.92	84.01
389	280	4972	5.01	89.02
403	240	5212	4.30	93.32
424	164	5376	2.94	96.26
458	141	5517	2.52	98.78
520	68	5585	1.22	100.00

Figure C.6.11. 2024 STAAR Alternate 2 Grade 7 RLA
Frequency Distribution of Scale Scores



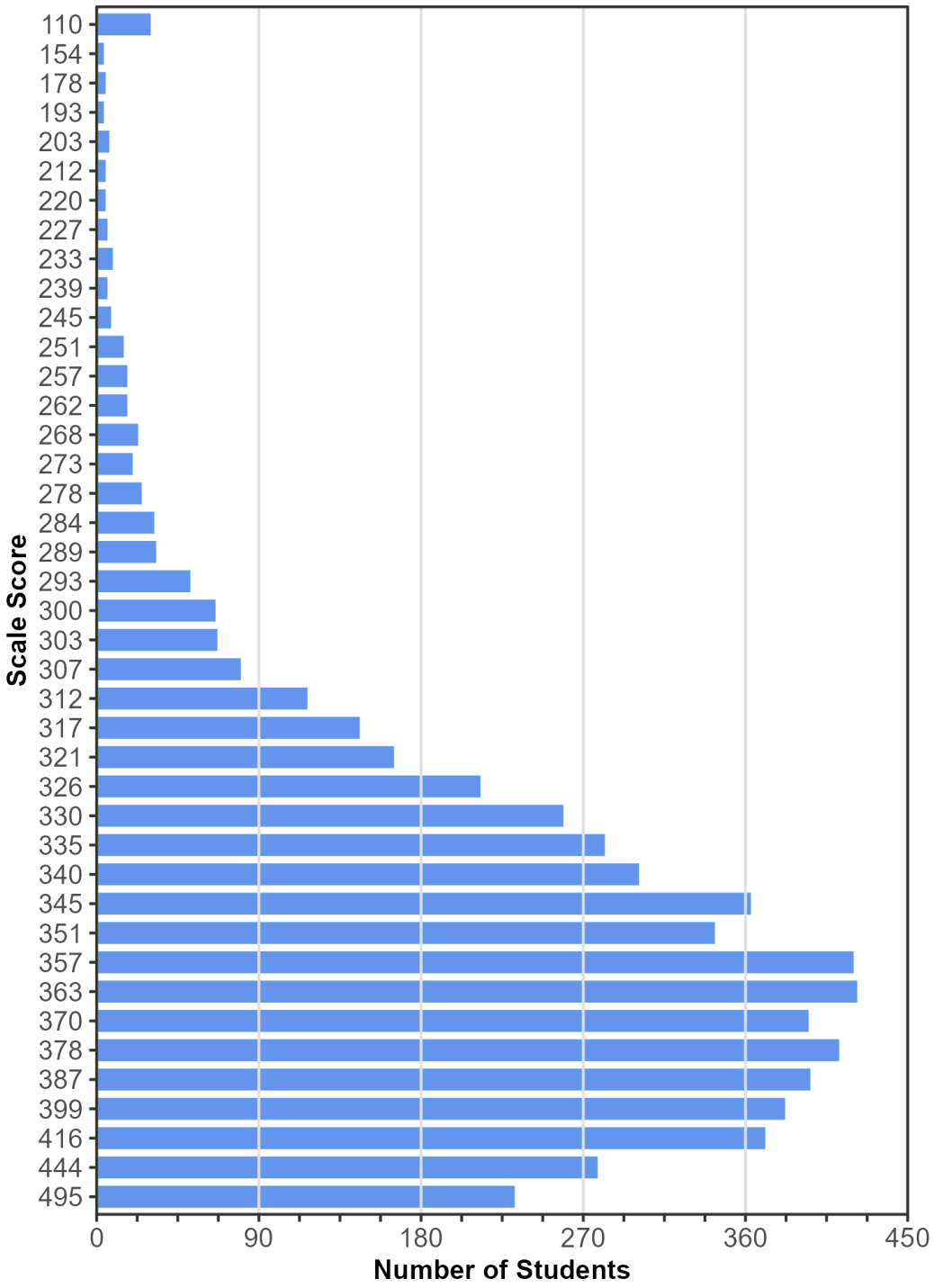
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	17	17	0.33	0.33
116	2	19	0.04	0.36
149	8	27	0.15	0.52
168	4	31	0.08	0.59
182	6	37	0.11	0.71
193	8	45	0.15	0.86
202	11	56	0.21	1.07
210	3	59	0.06	1.13
218	10	69	0.19	1.32
225	7	76	0.13	1.45
231	8	84	0.15	1.61
237	6	90	0.11	1.72
243	10	100	0.19	1.91
249	9	109	0.17	2.09
254	25	134	0.48	2.56
260	19	153	0.36	2.93
265	29	182	0.55	3.48
271	34	216	0.65	4.13
276	40	256	0.77	4.90
281	59	315	1.13	6.03
286	59	374	1.13	7.16
292	85	459	1.63	8.78
297	113	572	2.16	10.94
300	128	700	2.45	13.39
307	158	858	3.02	16.41
313	205	1063	3.92	20.34
318	236	1299	4.52	24.85
324	293	1592	5.61	30.46
329	352	1944	6.73	37.19
335	302	2246	5.78	42.97
342	354	2600	6.77	49.74
348	334	2934	6.39	56.13
355	367	3301	7.02	63.15
363	340	3641	6.50	69.66
372	335	3976	6.41	76.07
378	313	4289	5.99	82.05
394	280	4569	5.36	87.41
409	230	4799	4.40	91.81
430	175	4974	3.35	95.16
465	151	5125	2.89	98.05
528	102	5227	1.95	100.00

Figure C.6.12. 2024 STAAR Alternate 2 Grade 8 RLA
Frequency Distribution of Scale Scores



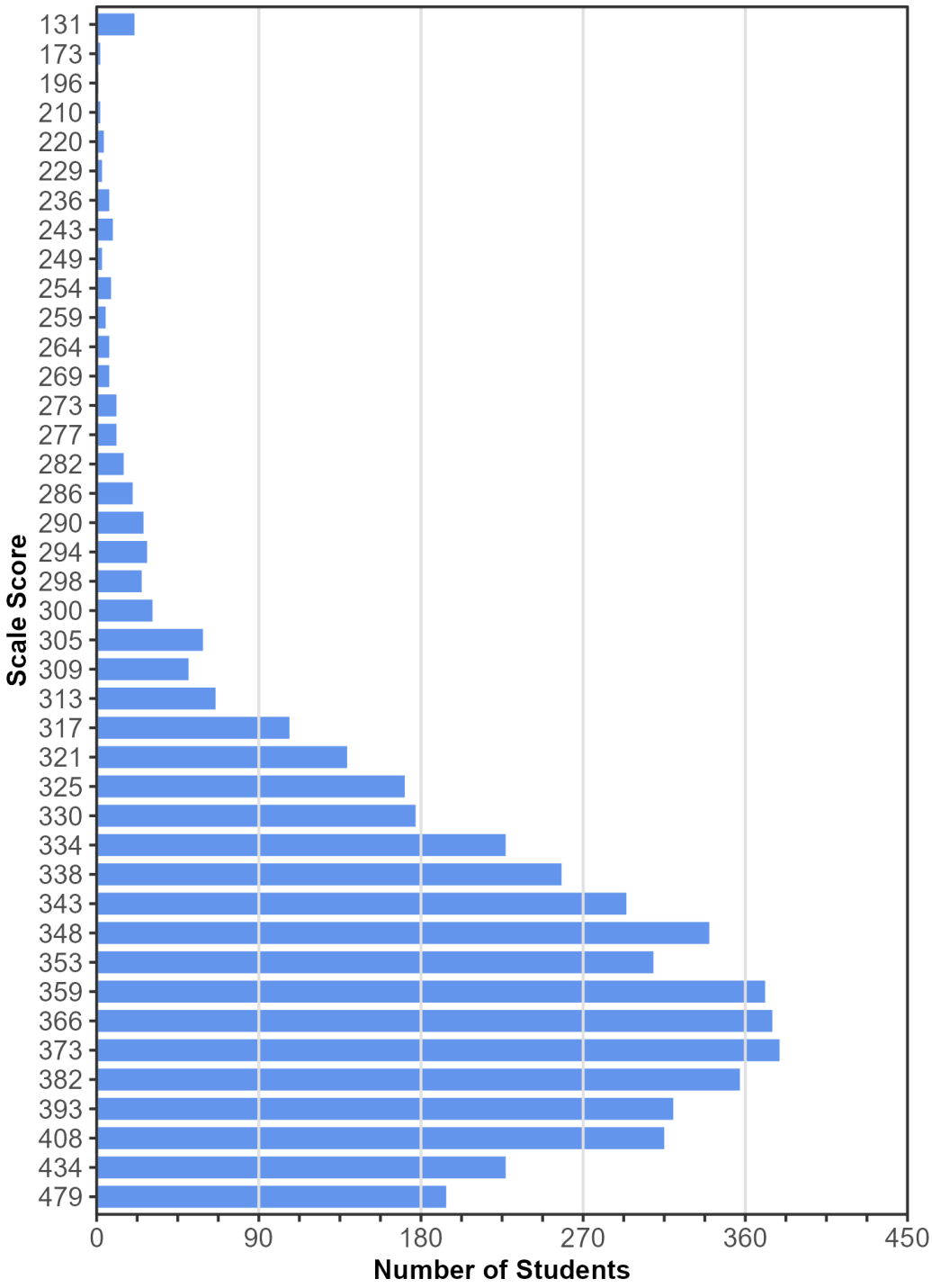
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	19	19	0.38	0.38
122	2	21	0.04	0.42
154	5	26	0.10	0.52
173	1	27	0.02	0.54
186	4	31	0.08	0.62
196	2	33	0.04	0.66
205	6	39	0.12	0.78
213	2	41	0.04	0.82
220	9	50	0.18	1.00
227	14	64	0.28	1.28
233	10	74	0.20	1.48
239	12	86	0.24	1.73
245	12	98	0.24	1.97
250	14	112	0.28	2.25
256	20	132	0.40	2.65
261	22	154	0.44	3.09
266	26	180	0.52	3.61
271	25	205	0.50	4.11
276	43	248	0.86	4.97
281	36	284	0.72	5.70
286	61	345	1.22	6.92
291	69	414	1.38	8.30
296	112	526	2.25	10.55
300	123	649	2.47	13.02
306	184	833	3.69	16.71
311	201	1034	4.03	20.74
316	232	1266	4.65	25.40
321	259	1525	5.20	30.59
326	313	1838	6.28	36.87
332	336	2174	6.74	43.61
338	320	2494	6.42	50.03
344	326	2820	6.54	56.57
351	306	3126	6.14	62.71
358	331	3457	6.64	69.35
371	332	3789	6.66	76.01
376	283	4072	5.68	81.69
387	296	4368	5.94	87.62
401	216	4584	4.33	91.96
421	193	4777	3.87	95.83
455	122	4899	2.45	98.27
514	86	4985	1.73	100.00

Figure C.6.13. 2024 STAAR Alternate 2 Grade 5 Science
Frequency Distribution of Scale Scores



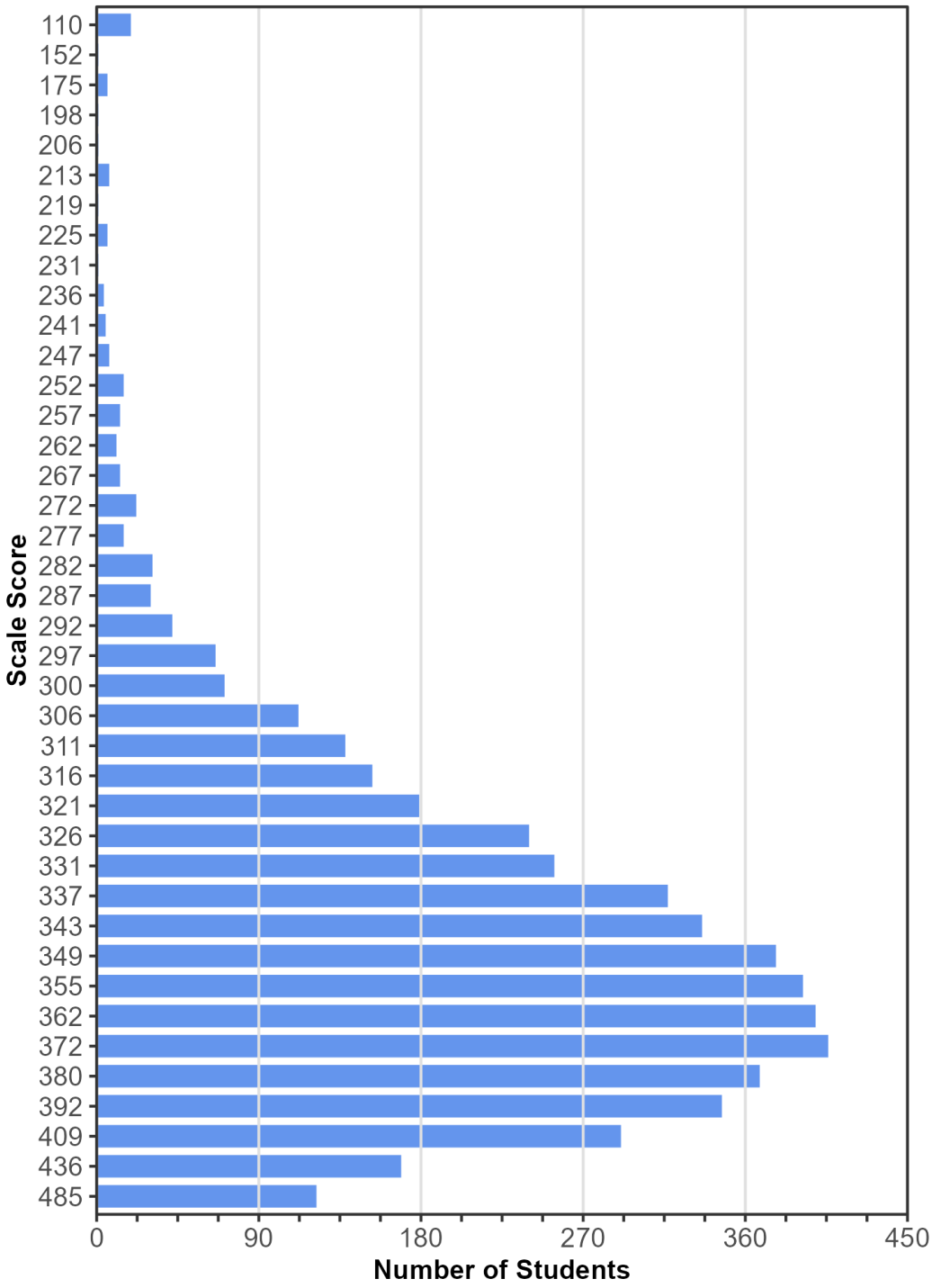
	Freq	Cum Freq	Pct	Cum Pct
110	30	30	0.50	0.50
154	4	34	0.07	0.56
178	5	39	0.08	0.65
193	4	43	0.07	0.71
203	7	50	0.12	0.83
212	5	55	0.08	0.91
220	5	60	0.08	0.99
227	6	66	0.10	1.09
233	9	75	0.15	1.24
239	6	81	0.10	1.34
245	8	89	0.13	1.48
251	15	104	0.25	1.72
257	17	121	0.28	2.01
262	17	138	0.28	2.29
268	23	161	0.38	2.67
273	20	181	0.33	3.00
278	25	206	0.41	3.41
284	32	238	0.53	3.94
289	33	271	0.55	4.49
293	52	323	0.86	5.35
300	66	389	1.09	6.45
303	67	456	1.11	7.56
307	80	536	1.33	8.88
312	117	653	1.94	10.82
317	146	799	2.42	13.24
321	165	964	2.73	15.98
326	213	1177	3.53	19.51
330	259	1436	4.29	23.80
335	282	1718	4.67	28.48
340	301	2019	4.99	33.47
345	363	2382	6.02	39.48
351	343	2725	5.69	45.17
357	420	3145	6.96	52.13
363	422	3567	6.99	59.12
370	395	3962	6.55	65.67
378	412	4374	6.83	72.50
387	396	4770	6.56	79.07
399	382	5152	6.33	85.40
416	371	5523	6.15	91.55
444	278	5801	4.61	96.15
495	232	6033	3.85	100.00

Figure C.6.14. 2024 STAAR Alternate 2 Grade 8 Science
Frequency Distribution of Scale Scores



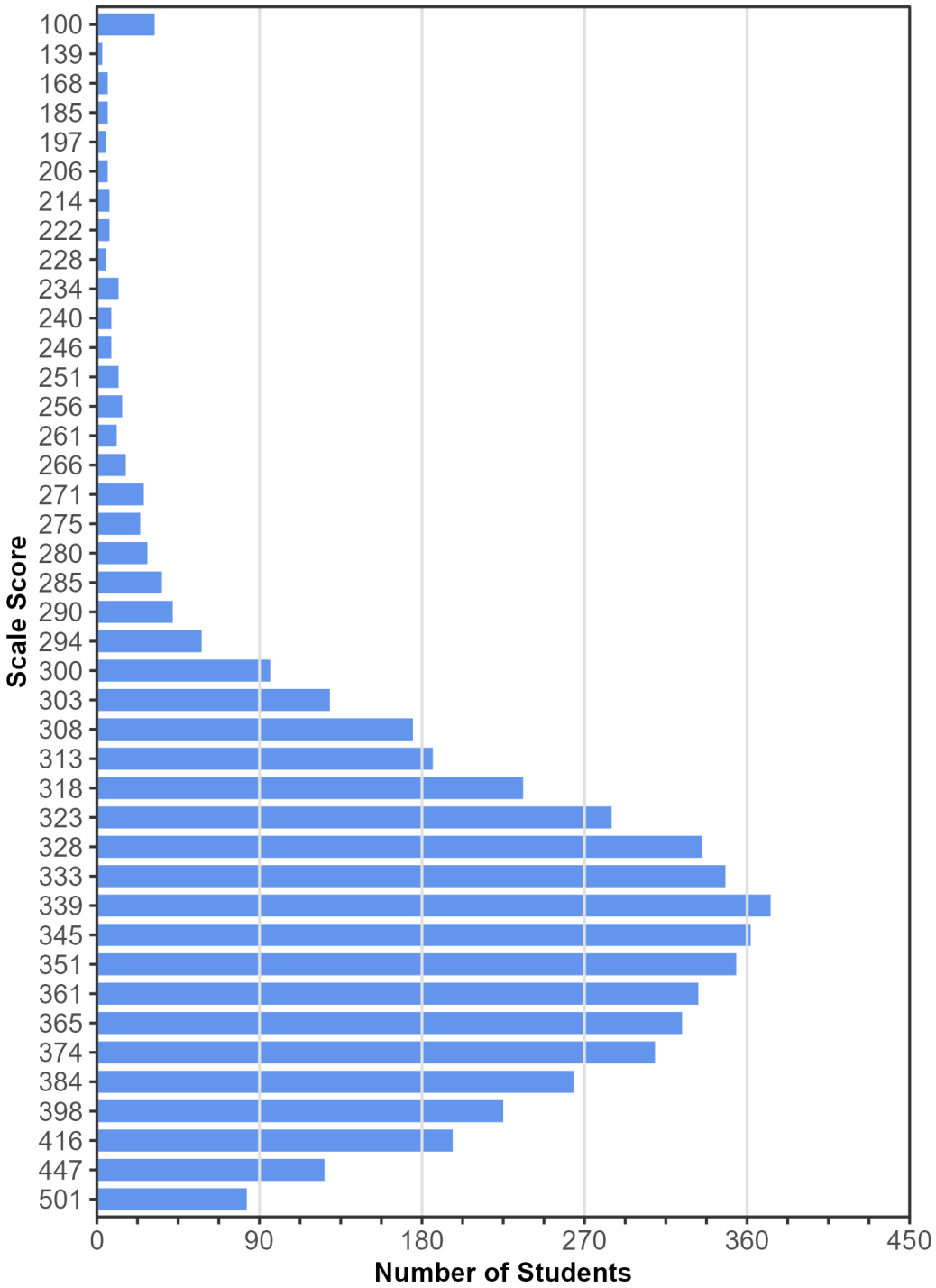
	Freq	Cum Freq	Pct	Cum Pct
131	21	21	0.42	0.42
173	2	23	0.04	0.46
196	1	24	0.02	0.48
210	2	26	0.04	0.52
220	4	30	0.08	0.60
229	3	33	0.06	0.66
236	7	40	0.14	0.80
243	9	49	0.18	0.98
249	3	52	0.06	1.04
254	8	60	0.16	1.20
259	5	65	0.10	1.30
264	7	72	0.14	1.45
269	7	79	0.14	1.59
273	11	90	0.22	1.81
277	11	101	0.22	2.03
282	15	116	0.30	2.33
286	20	136	0.40	2.73
290	26	162	0.52	3.25
294	28	190	0.56	3.81
298	25	215	0.50	4.32
300	31	246	0.62	4.94
305	59	305	1.18	6.12
309	51	356	1.02	7.15
313	66	422	1.32	8.47
317	107	529	2.15	10.62
321	139	668	2.79	13.41
325	171	839	3.43	16.84
330	177	1016	3.55	20.39
334	227	1243	4.56	24.95
338	258	1501	5.18	30.13
343	294	1795	5.90	36.03
348	340	2135	6.82	42.85
353	309	2444	6.20	49.06
359	371	2815	7.45	56.50
366	375	3190	7.53	64.03
373	379	3569	7.61	71.64
382	357	3926	7.17	78.80
393	320	4246	6.42	85.23
408	315	4561	6.32	91.55
434	227	4788	4.56	96.11
479	194	4982	3.89	100.00

Figure C.6.15. 2024 STAAR Alternate 2 Grade 8 Social Studies
Frequency Distribution of Scale Scores



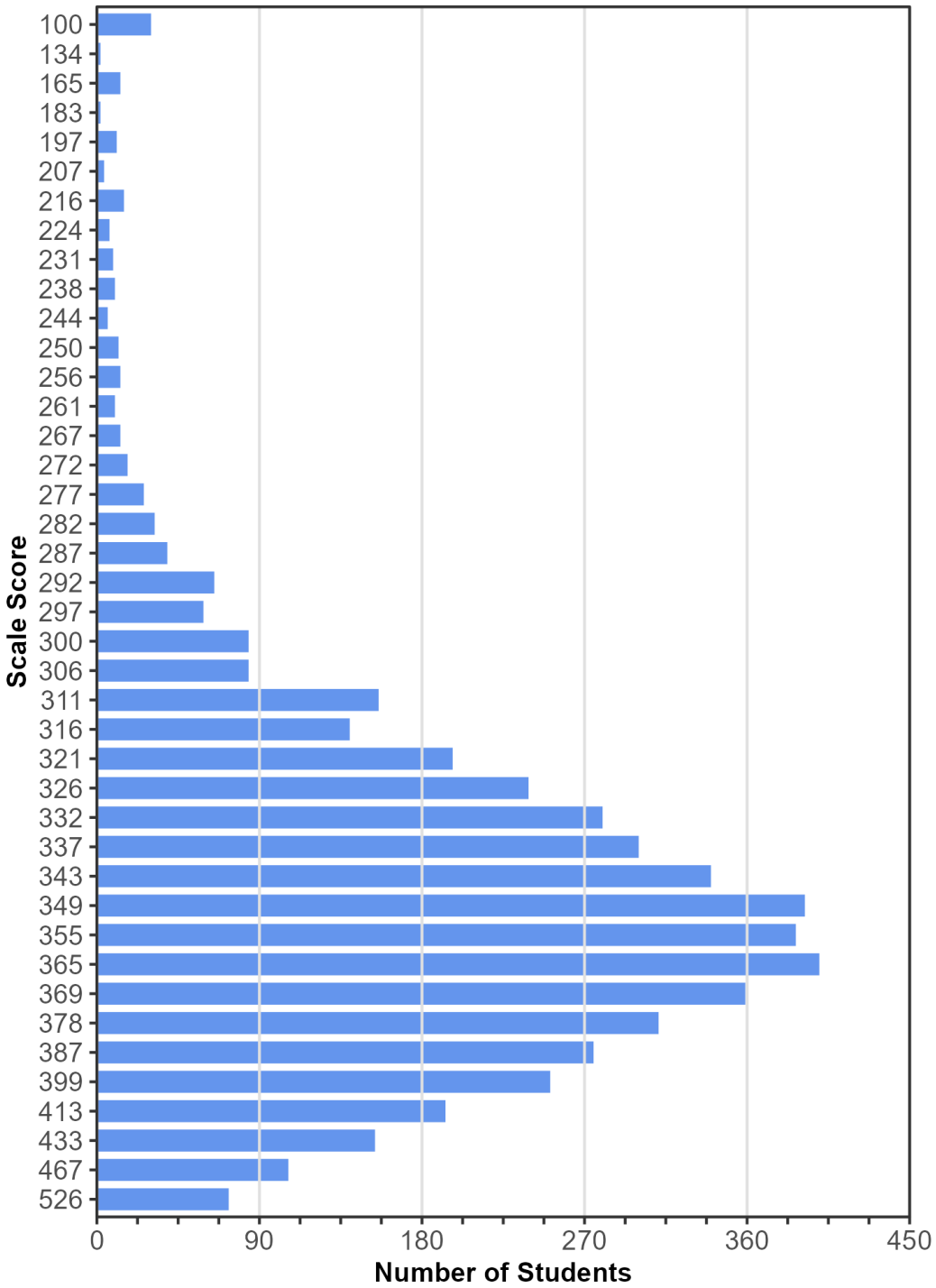
Scale Score	Freq	Cum Freq	Pct	Cum Pct
110	19	19	0.38	0.38
152	1	20	0.02	0.40
175	6	26	0.12	0.52
198	1	27	0.02	0.54
206	1	28	0.02	0.56
213	7	35	0.14	0.70
219	1	36	0.02	0.72
225	6	42	0.12	0.84
231	1	43	0.02	0.86
236	4	47	0.08	0.94
241	5	52	0.10	1.04
247	7	59	0.14	1.18
252	15	74	0.30	1.48
257	13	87	0.26	1.74
262	11	98	0.22	1.96
267	13	111	0.26	2.23
272	22	133	0.44	2.67
277	15	148	0.30	2.97
282	31	179	0.62	3.59
287	30	209	0.60	4.19
292	42	251	0.84	5.03
297	66	317	1.32	6.36
300	71	388	1.42	7.78
306	112	500	2.25	10.02
311	138	638	2.77	12.79
316	153	791	3.07	15.86
321	179	970	3.59	19.45
326	240	1210	4.81	24.26
331	254	1464	5.09	29.35
337	317	1781	6.36	35.71
343	336	2117	6.74	42.44
349	377	2494	7.56	50.00
355	392	2886	7.86	57.86
362	399	3285	8.00	65.86
372	406	3691	8.14	74.00
380	368	4059	7.38	81.38
392	347	4406	6.96	88.33
409	291	4697	5.83	94.17
436	169	4866	3.39	97.55
485	122	4988	2.45	100.00

Figure C.6.16. 2024 STAAR Alternate 2 Algebra I
Frequency Distribution of Scale Scores



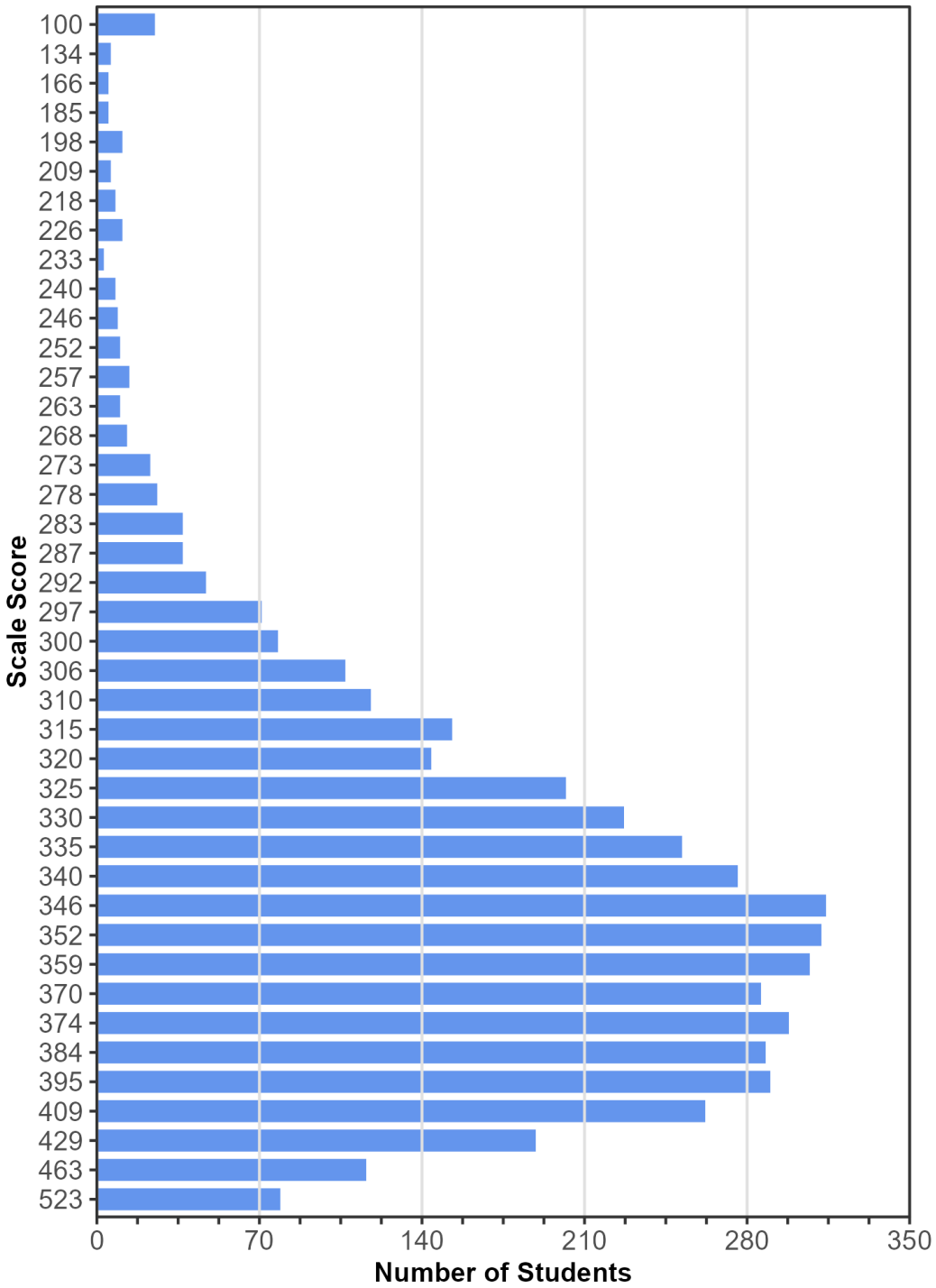
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	32	32	0.63	0.63
139	3	35	0.06	0.68
168	6	41	0.12	0.80
185	6	47	0.12	0.92
197	5	52	0.10	1.02
206	6	58	0.12	1.13
214	7	65	0.14	1.27
222	7	72	0.14	1.41
228	5	77	0.10	1.51
234	12	89	0.23	1.74
240	8	97	0.16	1.90
246	8	105	0.16	2.05
251	12	117	0.23	2.29
256	14	131	0.27	2.56
261	11	142	0.22	2.78
266	16	158	0.31	3.09
271	26	184	0.51	3.60
275	24	208	0.47	4.07
280	28	236	0.55	4.62
285	36	272	0.70	5.32
290	42	314	0.82	6.14
294	58	372	1.13	7.28
300	96	468	1.88	9.15
303	129	597	2.52	11.68
308	175	772	3.42	15.10
313	186	958	3.64	18.74
318	236	1194	4.62	23.36
323	285	1479	5.58	28.93
328	335	1814	6.55	35.49
333	348	2162	6.81	42.29
339	373	2535	7.30	49.59
345	362	2897	7.08	56.67
351	354	3251	6.92	63.60
361	333	3584	6.51	70.11
365	324	3908	6.34	76.45
374	309	4217	6.04	82.49
384	264	4481	5.16	87.66
398	225	4706	4.40	92.06
416	197	4903	3.85	95.91
447	126	5029	2.46	98.38
501	83	5112	1.62	100.00

Figure C.6.17. 2024 STAAR Alternate 2 English I
Frequency Distribution of Scale Scores



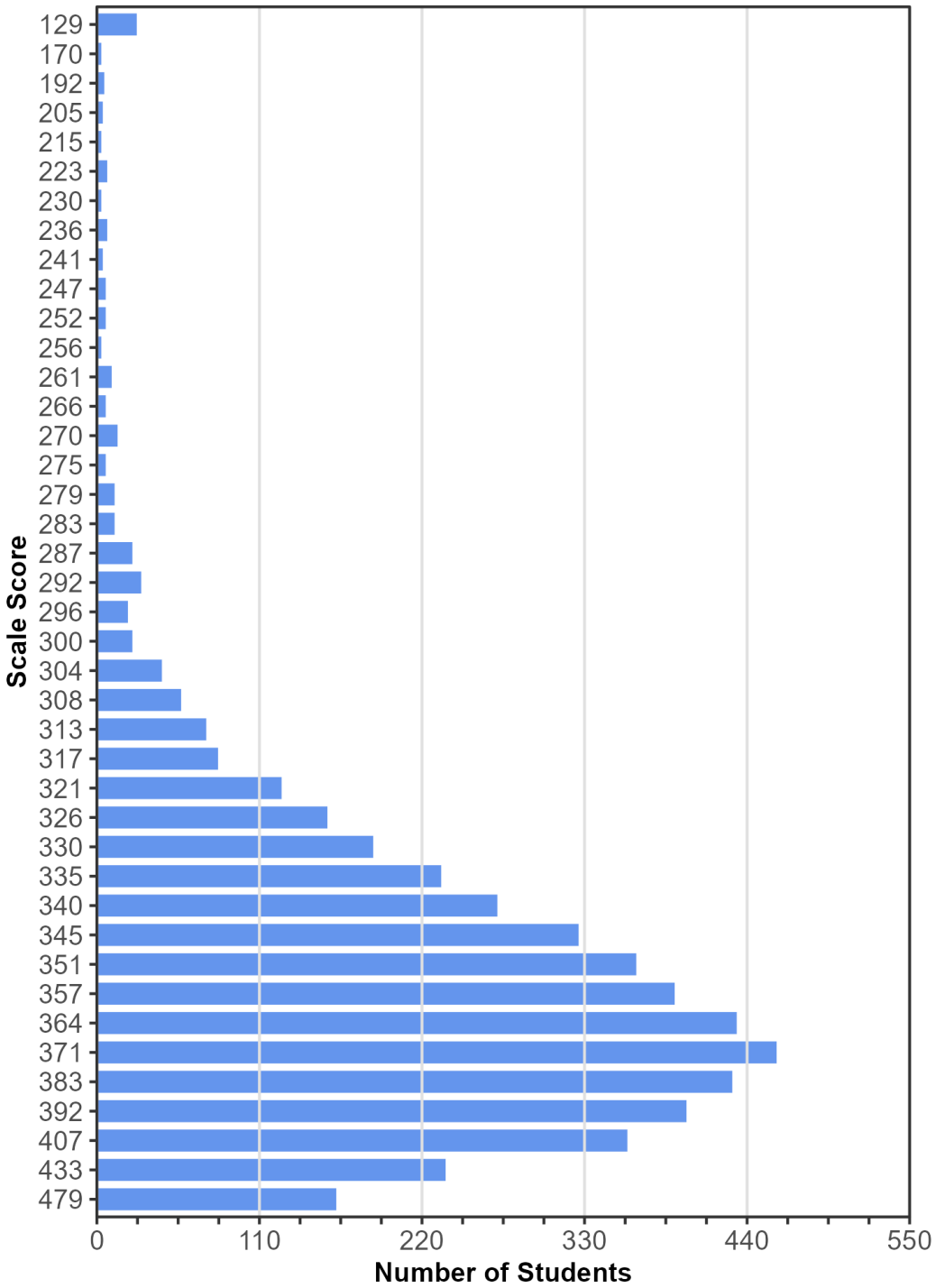
	Freq	Cum Freq	Pct	Cum Pct
100	30	30	0.59	0.59
134	2	32	0.04	0.63
165	13	45	0.25	0.88
183	2	47	0.04	0.92
197	11	58	0.22	1.13
207	4	62	0.08	1.21
216	15	77	0.29	1.51
224	7	84	0.14	1.64
231	9	93	0.18	1.82
238	10	103	0.20	2.01
244	6	109	0.12	2.13
250	12	121	0.23	2.37
256	13	134	0.25	2.62
261	10	144	0.20	2.81
267	13	157	0.25	3.07
272	17	174	0.33	3.40
277	26	200	0.51	3.91
282	32	232	0.63	4.53
287	39	271	0.76	5.30
292	65	336	1.27	6.57
297	59	395	1.15	7.72
300	84	479	1.64	9.36
306	84	563	1.64	11.00
311	156	719	3.05	14.05
316	140	859	2.74	16.79
321	197	1056	3.85	20.64
326	239	1295	4.67	25.31
332	280	1575	5.47	30.79
337	300	1875	5.86	36.65
343	340	2215	6.65	43.30
349	392	2607	7.66	50.96
355	387	2994	7.56	58.52
365	400	3394	7.82	66.34
369	359	3753	7.02	73.36
378	311	4064	6.08	79.44
387	275	4339	5.38	84.81
399	251	4590	4.91	89.72
413	193	4783	3.77	93.49
433	154	4937	3.01	96.50
467	106	5043	2.07	98.57
526	73	5116	1.43	100.00

Figure C.6.18. 2024 STAAR Alternate 2 English II
Frequency Distribution of Scale Scores



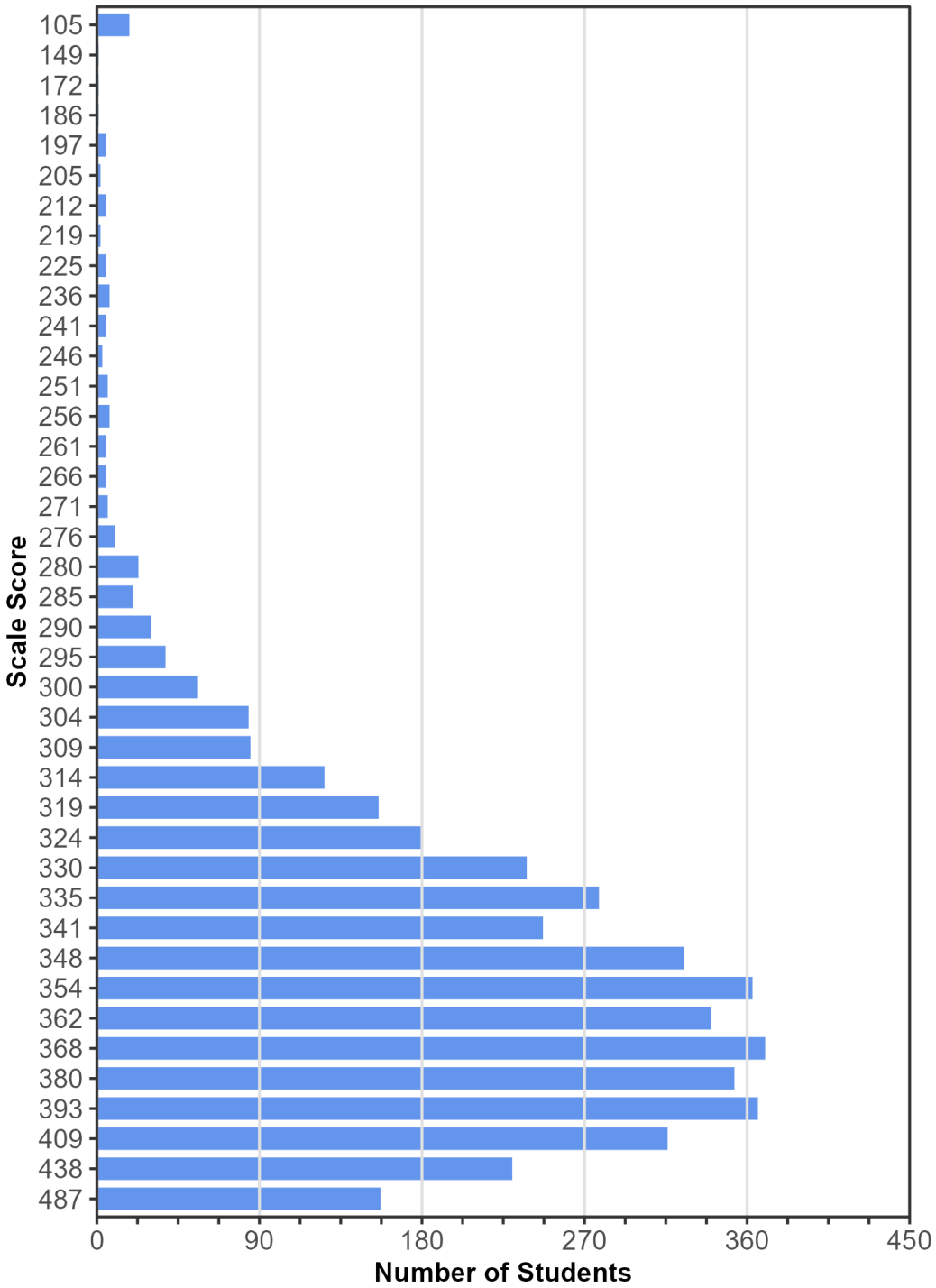
	Freq	Cum Freq	Pct	Cum Pct
100	25	25	0.53	0.53
134	6	31	0.13	0.66
166	5	36	0.11	0.77
185	5	41	0.11	0.88
198	11	52	0.23	1.11
209	6	58	0.13	1.24
218	8	66	0.17	1.41
226	11	77	0.23	1.64
233	3	80	0.06	1.71
240	8	88	0.17	1.88
246	9	97	0.19	2.07
252	10	107	0.21	2.28
257	14	121	0.30	2.58
263	10	131	0.21	2.80
268	13	144	0.28	3.07
273	23	167	0.49	3.57
278	26	193	0.56	4.12
283	37	230	0.79	4.91
287	37	267	0.79	5.70
292	47	314	1.00	6.71
297	71	385	1.52	8.22
300	78	463	1.67	9.89
306	107	570	2.28	12.17
310	118	688	2.52	14.69
315	153	841	3.27	17.96
320	144	985	3.07	21.03
325	202	1187	4.31	25.35
330	227	1414	4.85	30.19
335	252	1666	5.38	35.58
340	276	1942	5.89	41.47
346	314	2256	6.71	48.17
352	312	2568	6.66	54.84
359	307	2875	6.56	61.39
370	286	3161	6.11	67.50
374	298	3459	6.36	73.86
384	288	3747	6.15	80.01
395	290	4037	6.19	86.21
409	262	4299	5.59	91.80
429	189	4488	4.04	95.84
463	116	4604	2.48	98.31
523	79	4683	1.69	100.00

Figure C.6.19. 2024 STAAR Alternate 2 Biology
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
129	27	27	0.54	0.54
170	3	30	0.06	0.60
192	5	35	0.10	0.70
205	4	39	0.08	0.78
215	3	42	0.06	0.84
223	7	49	0.14	0.97
230	3	52	0.06	1.03
236	7	59	0.14	1.17
241	4	63	0.08	1.25
247	6	69	0.12	1.37
252	6	75	0.12	1.49
256	3	78	0.06	1.55
261	10	88	0.20	1.75
266	6	94	0.12	1.87
270	14	108	0.28	2.15
275	6	114	0.12	2.27
279	12	126	0.24	2.51
283	12	138	0.24	2.75
287	24	162	0.48	3.22
292	30	192	0.60	3.82
296	21	213	0.42	4.24
300	24	237	0.48	4.71
304	44	281	0.88	5.59
308	57	338	1.13	6.72
313	74	412	1.47	8.20
317	82	494	1.63	9.83
321	125	619	2.49	12.31
326	156	775	3.10	15.42
330	187	962	3.72	19.14
335	233	1195	4.63	23.77
340	271	1466	5.39	29.16
345	326	1792	6.48	35.65
351	365	2157	7.26	42.91
357	391	2548	7.78	50.69
364	433	2981	8.61	59.30
371	460	3441	9.15	68.45
383	430	3871	8.55	77.00
392	399	4270	7.94	84.94
407	359	4629	7.14	92.08
433	236	4865	4.69	96.78
479	162	5027	3.22	100.00

Figure C.6.20. 2024 STAAR Alternate 2 U.S. History
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
105	18	18	0.40	0.40
149	1	19	0.02	0.42
172	1	20	0.02	0.45
186	1	21	0.02	0.47
197	5	26	0.11	0.58
205	2	28	0.04	0.63
212	5	33	0.11	0.74
219	2	35	0.04	0.78
225	5	40	0.11	0.89
236	7	47	0.16	1.05
241	5	52	0.11	1.16
246	3	55	0.07	1.23
251	6	61	0.13	1.36
256	7	68	0.16	1.52
261	5	73	0.11	1.63
266	5	78	0.11	1.74
271	6	84	0.13	1.88
276	10	94	0.22	2.10
280	23	117	0.51	2.61
285	20	137	0.45	3.06
290	30	167	0.67	3.73
295	38	205	0.85	4.58
300	56	261	1.25	5.83
304	84	345	1.88	7.71
309	85	430	1.90	9.61
314	126	556	2.82	12.42
319	156	712	3.49	15.91
324	180	892	4.02	19.93
330	238	1130	5.32	25.25
335	278	1408	6.21	31.46
341	247	1655	5.52	36.98
348	325	1980	7.26	44.25
354	363	2343	8.11	52.36
362	340	2683	7.60	59.96
368	370	3053	8.27	68.22
380	353	3406	7.89	76.11
393	366	3772	8.18	84.29
409	316	4088	7.06	91.35
438	230	4318	5.14	96.49
487	157	4475	3.51	100.00

TECHNICAL DIGEST 2023–2024

Appendix D

TELPAS Statistical Tables and Figures

List of Tables

Table D.1.1. 2024 TELPAS Classification Consistency and Accuracy

Table D.2.1. 2024 TELPAS Scale Score Correlations

Table D.3.1. 2024 TELPAS Reading Grades 2–5 Conditional Standard Error of Measurement for Scale Scores

Table D.3.2. 2024 TELPAS Reading Grades 6–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.3. 2024 TELPAS Writing Grades 2–5 Conditional Standard Error of Measurement for Scale Scores

Table D.3.4. 2024 TELPAS Writing Grades 6–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.5. 2024 TELPAS Listening Grades 2–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.6. 2024 TELPAS Speaking Grades 2–12 Conditional Standard Error of Measurement for Scale Scores

Table D.4.1. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.2. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.3. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.4. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.5. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.6. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.7. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.8. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.9. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.10. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.11. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.12. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.13. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.14. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.15. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.16. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.17. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.18. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.19. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.20. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.21. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.22. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.23. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.24. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.25. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.26. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.27. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.28. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.29. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.30. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.5.1. 2024 TELPAS Reading Scale Score Descriptive Statistics

Table D.5.2. 2024 TELPAS Writing Scale Score Descriptive Statistics

Table D.5.3. 2024 TELPAS Listening Scale Score Descriptive Statistics

Table D.5.4. 2024 TELPAS Speaking Scale Score Descriptive Statistics

Table D.6.1. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades K–1

Table D.6.2. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 2–3

Table D.6.3. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 4–5

Table D.6.4. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 6–7

Table D.6.5. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 8–9

Table D.6.6. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 10–11

Table D.6.7. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grade 12

List of Figures

Figure D.5.1. 2024 TELPAS Grade 2 Reading Frequency Distribution of Scale Scores
Figure D.5.2. 2024 TELPAS Grade 3 Reading Frequency Distribution of Scale Scores
Figure D.5.3. 2024 TELPAS Grade 4 Reading Frequency Distribution of Scale Scores
Figure D.5.4. 2024 TELPAS Grade 5 Reading Frequency Distribution of Scale Scores
Figure D.5.5. 2024 TELPAS Grade 6 Reading Frequency Distribution of Scale Scores
Figure D.5.6. 2024 TELPAS Grade 7 Reading Frequency Distribution of Scale Scores
Figure D.5.7. 2024 TELPAS Grade 8 Reading Frequency Distribution of Scale Scores
Figure D.5.8. 2024 TELPAS Grade 9 Reading Frequency Distribution of Scale Scores
Figure D.5.9. 2024 TELPAS Grade 10 Reading Frequency Distribution of Scale Scores
Figure D.5.10. 2024 TELPAS Grade 11 Reading Frequency Distribution of Scale Scores
Figure D.5.11. 2024 TELPAS Grade 12 Reading Frequency Distribution of Scale Scores
Figure D.5.12. 2024 TELPAS Grade 2 Writing Frequency Distribution of Scale Scores
Figure D.5.13. 2024 TELPAS Grade 3 Writing Frequency Distribution of Scale Scores
Figure D.5.14. 2024 TELPAS Grade 4 Writing Frequency Distribution of Scale Scores
Figure D.5.15. 2024 TELPAS Grade 5 Writing Frequency Distribution of Scale Scores
Figure D.5.16. 2024 TELPAS Grade 6 Writing Frequency Distribution of Scale Scores
Figure D.5.17. 2024 TELPAS Grade 7 Writing Frequency Distribution of Scale Scores
Figure D.5.18. 2024 TELPAS Grade 8 Writing Frequency Distribution of Scale Scores
Figure D.5.19. 2024 TELPAS Grade 9 Writing Frequency Distribution of Scale Scores
Figure D.5.20. 2024 TELPAS Grade 10 Writing Frequency Distribution of Scale Scores
Figure D.5.21. 2024 TELPAS Grade 11 Writing Frequency Distribution of Scale Scores
Figure D.5.22. 2024 TELPAS Grade 12 Writing Frequency Distribution of Scale Scores
Figure D.5.23. 2024 TELPAS Grade 2 Listening Frequency Distribution of Scale Scores
Figure D.5.24. 2024 TELPAS Grade 3 Listening Frequency Distribution of Scale Scores
Figure D.5.25. 2024 TELPAS Grade 4 Listening Frequency Distribution of Scale Scores
Figure D.5.26. 2024 TELPAS Grade 5 Listening Frequency Distribution of Scale Scores
Figure D.5.27. 2024 TELPAS Grade 6 Listening Frequency Distribution of Scale Scores

Figure D.5.28. 2024 TELPAS Grade 7 Listening Frequency Distribution of Scale Scores
Figure D.5.29. 2024 TELPAS Grade 8 Listening Frequency Distribution of Scale Scores
Figure D.5.30. 2024 TELPAS Grade 9 Listening Frequency Distribution of Scale Scores
Figure D.5.31. 2024 TELPAS Grade 10 Listening Frequency Distribution of Scale Scores
Figure D.5.32. 2024 TELPAS Grade 11 Listening Frequency Distribution of Scale Scores
Figure D.5.33. 2024 TELPAS Grade 12 Listening Frequency Distribution of Scale Scores
Figure D.5.34. 2024 TELPAS Grade 2 Speaking Frequency Distribution of Scale Scores
Figure D.5.35. 2024 TELPAS Grade 3 Speaking Frequency Distribution of Scale Scores
Figure D.5.36. 2024 TELPAS Grade 4 Speaking Frequency Distribution of Scale Scores
Figure D.5.37. 2024 TELPAS Grade 5 Speaking Frequency Distribution of Scale Scores
Figure D.5.38. 2024 TELPAS Grade 6 Speaking Frequency Distribution of Scale Scores
Figure D.5.39. 2024 TELPAS Grade 7 Speaking Frequency Distribution of Scale Scores
Figure D.5.40. 2024 TELPAS Grade 8 Speaking Frequency Distribution of Scale Scores
Figure D.5.41. 2024 TELPAS Grade 9 Speaking Frequency Distribution of Scale Scores
Figure D.5.42. 2024 TELPAS Grade 10 Speaking Frequency Distribution of Scale Scores
Figure D.5.43. 2024 TELPAS Grade 11 Speaking Frequency Distribution of Scale Scores
Figure D.5.44. 2024 TELPAS Grade 12 Speaking Frequency Distribution of Scale Scores

2024 TELPAS
Classification Consistency and Accuracy

Table D.1.1. 2024 TELPAS Classification Consistency and Accuracy

Grade	Domain	Decision Consistency[1]	Decision Accuracy[2]
2	Reading	68.5	76.9
3	Reading	67.8	76.4
4–5	Reading	68.7	77.1
6–7	Reading	66.6	75.6
8–9	Reading	70.6	78.7
10–12	Reading	70.6	78.8
2	Writing	72.5	80.6
3	Writing	77.9	84.1
4–5	Writing	73.6	81.5
6–7	Writing	71.6	80.0
8–9	Writing	73.0	81.0
10–12	Writing	68.7	77.4
2–3	Listening	68.0	76.7
4–5	Listening	64.4	73.8
6–8	Listening	66.5	75.6
9–12	Listening	66.3	75.3
2–3	Speaking	79.9	85.1
4–5	Speaking	75.7	82.3
6–8	Speaking	81.2	86.3
9–12	Speaking	80.7	86.4

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.

2024 TELPAS
Scale Score Correlations

Table D.2.1. 2024 TELPAS Scale Score Correlations

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
2	Reading	Listening	104,902	0.69
	Reading	Speaking	104,902	0.43
	Reading	Writing	102,632	0.77
	Listening	Speaking	105,029	0.52
	Listening	Writing	102,399	0.66
	Speaking	Writing	102,399	0.49
3	Reading	Listening	105,556	0.74
	Reading	Speaking	105,556	0.47
	Reading	Writing	104,842	0.78
	Listening	Speaking	105,680	0.53
	Listening	Writing	104,636	0.70
	Speaking	Writing	104,636	0.54
4	Reading	Listening	105,770	0.77
	Reading	Speaking	105,770	0.47
	Reading	Writing	105,295	0.77
	Listening	Speaking	105,860	0.51
	Listening	Writing	105,087	0.69
	Speaking	Writing	105,087	0.54
5	Reading	Listening	106,381	0.79
	Reading	Speaking	106,381	0.48
	Reading	Writing	105,993	0.77
	Listening	Speaking	106,484	0.51
	Listening	Writing	105,804	0.71
	Speaking	Writing	105,804	0.55
6	Reading	Listening	105,357	0.78
	Reading	Speaking	105,357	0.43
	Reading	Writing	104,860	0.75
	Listening	Speaking	105,521	0.47
	Listening	Writing	104,669	0.69
	Speaking	Writing	104,669	0.50
7	Reading	Listening	104,689	0.79
	Reading	Speaking	104,689	0.44
	Reading	Writing	104,063	0.76
	Listening	Speaking	104,861	0.47
	Listening	Writing	103,792	0.70
	Speaking	Writing	103,792	0.51

(Continued)

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
8	Reading	Listening	101,252	0.78
	Reading	Speaking	101,252	0.44
	Reading	Writing	100,225	0.71
	Listening	Speaking	101,432	0.48
	Listening	Writing	99,958	0.69
	Speaking	Writing	99,958	0.52
9	Reading	Listening	108,754	0.76
	Reading	Speaking	108,754	0.50
	Reading	Writing	105,718	0.72
	Listening	Speaking	109,302	0.52
	Listening	Writing	105,052	0.71
	Speaking	Writing	105,052	0.52
10	Reading	Listening	85,629	0.76
	Reading	Speaking	85,629	0.47
	Reading	Writing	83,503	0.74
	Listening	Speaking	86,046	0.49
	Listening	Writing	82,937	0.68
	Speaking	Writing	82,937	0.51
11	Reading	Listening	67,823	0.74
	Reading	Speaking	67,823	0.45
	Reading	Writing	66,244	0.73
	Listening	Speaking	68,221	0.48
	Listening	Writing	65,790	0.65
	Speaking	Writing	65,790	0.49
12	Reading	Listening	50,746	0.72
	Reading	Speaking	50,746	0.40
	Reading	Writing	49,771	0.70
	Listening	Speaking	51,120	0.44
	Listening	Writing	49,320	0.60
	Speaking	Writing	49,320	0.44

2024 TELPAS
Conditional Standard Error of Measurement for Scale Scores

Table D.3.1. 2024 TELPAS Reading Grades 2–5
Conditional Standard Error of Measurement for Scale Scores

	Grade 2		Grade 3		Grades 4–5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1167		1063		1087	
1	1250	69	1173	90	1195	89
2	1300	50	1238	65	1259	64
3	1331	42	1279	55	1299	54
4	1354	37	1309	49	1329	48
5	1373	34	1333	44	1353	44
6	1389	32	1354	42	1373	41
7	1403	30	1373	39	1391	39
8	1416	29	1389	38	1408	37
9	1428	28	1405	36	1423	35
10	*1440	27	1420	35	*1437	34
11	1450	27	1433	34	1450	33
12	1461	26	*1447	34	1463	33
13	1471	26	1459	33	1475	32
14	1481	26	1472	33	1486	32
15	1491	26	1484	33	1498	31
16	1501	26	1496	32	1509	31
17	1510	26	1508	32	1520	31
18	1520	26	1520	32	**1531	31
19	**1531	26	**1531	32	1542	31
20	1541	27	1543	33	1553	31
21	1552	28	1555	33	1564	31
22	1564	28	1568	33	1575	31
23	1577	30	1581	34	1586	32
24	1591	31	1594	35	1598	32
25	***1606	33	***1608	36	***1610	33
26	1625	37	1623	37	1622	34
27	1647	41	1639	39	1636	35
28	1677	49	1657	41	1650	36
29	1727	68	1678	44	1666	38
30	1809		1701	48	1683	40
31			1731	54	1703	43
32			1770	65	1726	47
33			1835	90	1754	53
34			1944		1793	64
35					1856	88
36					1963	

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. 2024 TELPAS Reading Grades 6–12
Conditional Standard Error of Measurement for Scale Scores

	Grades 6–7		Grades 8–9		Grades 10–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1110		1162		1150	
1	1208	81	1247	71	1230	66
2	1266	58	1298	51	1279	48
3	1301	49	1330	43	1309	41
4	1327	43	1353	38	1332	36
5	1348	39	1372	35	1350	33
6	1366	36	1388	32	1366	31
7	1382	34	1402	31	1380	29
8	1396	33	1415	29	1393	28
9	1409	31	1427	28	1405	27
10	1421	31	*1438	27	1416	26
11	1432	30	1449	26	1426	26
12	1443	29	1459	26	*1436	25
13	*1454	29	1468	25	1446	25
14	1464	28	1477	25	1455	24
15	1474	28	1486	25	1464	24
16	1483	28	1495	24	1473	24
17	1493	27	1503	24	1482	24
18	1502	27	1512	24	1490	23
19	1512	27	1520	24	1499	23
20	1521	28	**1529	24	1507	23
21	**1531	28	1537	24	1516	23
22	1540	28	1545	24	1524	24
23	1550	28	1554	24	**1533	24
24	1560	29	1563	25	1542	24
25	1571	29	1572	25	1551	24
26	1582	30	1581	26	1560	25
27	1593	31	1591	26	1570	25
28	***1605	32	***1602	27	1580	26
29	1618	33	1613	28	1592	27
30	1633	34	1625	30	***1603	28
31	1648	36	1639	31	1617	30
32	1666	39	1654	34	1632	32
33	1687	43	1672	37	1649	35
34	1713	49	1695	42	1671	40
35	1749	58	1725	50	1700	48
36	1807	81	1775	70	1747	66
37	1905		1860		1827	

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. 2024 TELPAS Writing Grades 2–5
Conditional Standard Error of Measurement for Scale Scores

	Grade 2		Grade 3		Grades 4–5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
6	1356		1262			
7	1405	40	1320	46		
8	*1435	29	1357	34		
9	1453	24	1380	29	1260	
10	1467	22	1398	26	1322	52
11	1478	20	*1413	25	1352	35
12	1489	20	1428	24	1369	28
13	1499	20	1442	24	1382	25
14	1509	20	1456	24	1392	24
15	1521	21	1470	25	1401	23
16	**1533	22	1486	26	*1410	23
17	1546	22	1503	27	1420	24
18	1559	22	1520	27	1430	25
19	1571	21	**1536	26	1442	27
20	1582	21	1552	25	1455	29
21	1593	20	1566	24	1470	29
22	***1604	21	1580	24	1485	28
23	1616	21	1594	24	1498	26
24	1628	21	***1608	25	1508	24
25	1640	22	1623	25	1518	23
26	1653	22	1640	26	**1526	22
27	1667	24	1658	28	1535	21
28	1684	27	1679	32	1543	21
29	1710	36	1710	42	1551	22
30	1752		1759		1559	22
31					1567	22
32					1576	22
33					1584	22
34					1593	22
35					***1601	22
36					1610	22
37					1618	22
38					1627	23
39					1638	26
40					1651	30
41					1672	42
42					1714	

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. 2024 TELPAS Writing Grades 6–12
Conditional Standard Error of Measurement for Scale Scores

	Grades 6–7		Grades 8–9		Grades 10–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
9	1291		1234		1249	
10	1344	47	1305	57	1336	69
11	1369	32	1340	37	1379	44
12	1383	27	1358	30	1401	34
13	1394	25	1371	26	1415	30
14	1404	23	1382	24	1427	27
15	1413	23	1391	23	1437	26
16	1422	23	1400	23	*1447	25
17	*1432	24	1409	23	1456	25
18	1442	25	*1418	23	1465	25
19	1454	27	1428	25	1474	25
20	1467	28	1439	26	1484	26
21	1480	27	1452	29	1493	26
22	1492	25	1467	30	1503	26
23	1502	24	1482	28	1513	26
24	1511	22	1494	26	1522	25
25	1519	21	1505	24	**1531	25
26	**1527	21	1514	23	1540	25
27	1534	21	1523	22	1550	25
28	1541	21	**1530	21	1559	26
29	1549	22	1538	21	1569	27
30	1557	22	1546	21	1580	28
31	1566	23	1554	22	1592	29
32	1575	23	1562	22	***1604	29
33	1583	22	1571	22	1616	28
34	1591	22	1579	23	1627	27
35	1599	22	1588	23	1638	27
36	***1607	22	1597	23	1648	27
37	1616	22	***1606	24	1658	27
38	1625	24	1616	25	1670	28
39	1635	26	1627	27	1682	31
40	1649	32	1641	31	1699	37
41	1674	48	1663	43	1728	54
42	1728		1705		1788	

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. 2024 TELPAS Listening Grades 2–12
Conditional Standard Error of Measurement for Scale Scores

	Grades 2–3		Grades 4–5		Grades 6–8		Grades 9–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
0	1192		1166		1177		1219	
1	1276	70	1247	67	1261	70	1286	55
2	1327	51	1295	48	1312	51	1327	40
3	1359	43	1326	41	1344	43	1352	34
4	1383	38	1349	36	1367	38	1371	30
5	1402	35	1367	34	1387	35	1387	28
6	1419	33	1384	32	1404	33	1401	26
7	1435	31	1398	30	1419	31	1414	25
8	*1449	30	1412	29	*1433	30	1425	24
9	1462	30	1425	28	1446	29	1436	24
10	1475	29	1437	28	1458	29	1446	23
11	1487	29	1448	27	1470	28	*1457	23
12	1499	28	*1460	27	1481	28	1466	23
13	1511	28	1471	27	1493	28	1476	23
14	1522	28	1482	27	1504	28	1486	23
15	**1534	28	1493	27	1515	28	1496	23
16	1546	28	1504	27	**1527	28	1506	23
17	1558	29	1516	28	1539	28	1516	24
18	1571	29	**1528	28	1551	29	**1526	24
19	1584	30	1541	29	1564	30	1537	25
20	1598	31	1554	30	1577	31	1549	26
21	***1613	33	1569	31	1592	32	1562	27
22	1630	35	1585	33	***1609	35	1576	28
23	1649	38	***1604	36	1628	38	1592	31
24	1673	42	1626	41	1651	42	***1611	34
25	1705	51	1657	48	1682	50	1637	41
26	1755	69	1705	66	1733	69	1678	56
27	1839		1786		1817		1746	

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. 2024 TELPAS Speaking Grades 2–12
Conditional Standard Error of Measurement for Scale Scores

	Grades 2–3		Grades 4–5		Grades 6–8		Grades 9–12	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
10	1262		1359		1381		1415	
11	1307	37	1391	26	1407	21	1441	21
12	1336	28	1411	19	1423	16	1456	15
13	1354	24	1423	17	1435	14	1465	13
14	1369	22	1433	15	1444	13	1471	11
15	1382	20	1442	14	1452	12	1477	10
16	1393	20	1450	13	*1460	12	1481	9
17	1404	19	1457	13	1467	12	*1485	9
18	*1415	19	1463	13	1474	12	1489	9
19	1425	19	*1470	12	1482	12	1493	9
20	1436	20	1476	13	1489	12	1496	9
21	1448	20	1482	13	1497	12	1500	9
22	1460	21	1489	13	1505	13	1504	9
23	1474	22	1497	14	1514	13	1508	10
24	1488	23	1505	15	1523	14	1512	10
25	1504	24	1514	16	**1533	14	1518	11
26	1522	26	**1525	17	1543	14	**1525	13
27	**1542	27	1538	19	1554	15	1534	15
28	1564	29	1554	21	1565	15	1546	17
29	1588	29	1574	23	1577	15	1561	17
30	***1610	28	1595	22	1588	14	1574	16
31	1633	28	***1613	21	1598	14	1586	15
32	1656	29	1630	20	***1608	14	1597	15
33	1681	30	1646	20	1619	15	***1608	16
34	1710	34	1664	22	1631	16	1621	18
35	1748	41	1688	27	1649	21	1639	23
36	1799		1721		1674		1667	

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.

2024 TELPAS
Mean P-Values, Raw Score Summary, and Reliability

Table D.4.1. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	30	105,149	12.88	6.33	0.86	2.39	42.92
	Understand words and language structures	12	105,149	5.41	2.85	0.73	1.50	45.11
	Basic understanding of variety of texts written in English	7	105,149	2.61	1.74	0.56	1.16	37.35
	Analyze and evaluate information and ideas in a variety of texts written in English	11	105,149	4.85	2.60	0.69	1.45	44.08
Writing	OVERALL TEST	30	102,632	10.44	4.81	0.91	1.40	26.78
	Express ideas in writing and address writing assignments	24	102,632	8.43	3.21	0.88	1.11	13.52
	Use standard grammar, usage, and spelling to edit writing tasks	6	102,632	2.00	1.95	0.81	0.86	33.42

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. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	30	51,027	12.99	6.28	0.86	2.39	43.31
	Understand words and language structures	12	51,027	5.35	2.80	0.71	1.50	44.62
	Basic understanding of variety of texts written in English	7	51,027	2.68	1.75	0.56	1.16	38.27
	Analyze and evaluate information and ideas in a variety of texts written in English	11	51,027	4.96	2.59	0.69	1.45	45.10
Writing	OVERALL TEST	30	50,065	10.73	4.96	0.92	1.44	27.87
	Express ideas in writing and address writing assignments	24	50,065	8.66	3.34	0.88	1.14	14.79
	Use standard grammar, usage, and spelling to edit writing tasks	6	50,065	2.06	1.96	0.80	0.87	34.41

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. 2024 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	30	54,077	12.77	6.38	0.86	2.38	42.56
	Understand words and language structures	12	54,077	5.47	2.90	0.74	1.49	45.58
	Basic understanding of variety of texts written in English	7	54,077	2.55	1.73	0.55	1.16	36.48
	Analyze and evaluate information and ideas in a variety of texts written in English	11	54,077	4.74	2.61	0.69	1.45	43.13
Writing	OVERALL TEST	30	52,526	10.16	4.66	0.91	1.37	25.76
	Express ideas in writing and address writing assignments	24	52,526	8.22	3.06	0.88	1.08	12.31
	Use standard grammar, usage, and spelling to edit writing tasks	6	52,526	1.95	1.94	0.81	0.85	32.48

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. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	34	105,769	16.96	7.72	0.90	2.49	49.89
	Understand words and language structures	12	105,769	7.46	3.20	0.81	1.38	62.13
	Basic understanding of variety of texts written in English	9	105,769	4.01	2.32	0.68	1.31	44.52
	Analyze and evaluate information and ideas in a variety of texts written in English	13	105,769	5.50	3.20	0.76	1.57	42.30
Writing	OVERALL TEST	30	104,842	13.86	5.65	0.92	1.61	46.23
	Express ideas in writing and address writing assignments	24	104,842	10.44	3.87	0.87	1.37	24.66
	Use standard grammar, usage, and spelling to edit writing tasks	6	104,842	3.42	2.10	0.84	0.85	57.02

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. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	34	51,469	17.15	7.74	0.90	2.49	50.44
	Understand words and language structures	12	51,469	7.49	3.17	0.81	1.39	62.43
	Basic understanding of variety of texts written in English	9	51,469	4.06	2.32	0.68	1.31	45.15
	Analyze and evaluate information and ideas in a variety of texts written in English	13	51,469	5.60	3.23	0.77	1.56	43.04
Writing	OVERALL TEST	30	51,103	14.33	5.77	0.92	1.66	47.90
	Express ideas in writing and address writing assignments	24	51,103	10.83	4.00	0.87	1.42	26.82
	Use standard grammar, usage, and spelling to edit writing tasks	6	51,103	3.51	2.09	0.83	0.85	58.44

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. 2024 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	34	54,276	16.79	7.70	0.89	2.50	49.37
	Understand words and language structures	12	54,276	7.42	3.24	0.82	1.38	61.86
	Basic understanding of variety of texts written in English	9	54,276	3.95	2.32	0.68	1.31	43.93
	Analyze and evaluate information and ideas in a variety of texts written in English	13	54,276	5.41	3.16	0.75	1.57	41.61
Writing	OVERALL TEST	30	53,718	13.41	5.50	0.92	1.57	44.65
	Express ideas in writing and address writing assignments	24	53,718	10.07	3.72	0.87	1.31	22.60
	Use standard grammar, usage, and spelling to edit writing tasks	6	53,718	3.34	2.11	0.84	0.85	55.68

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. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	36	212,562	20.40	8.11	0.90	2.56	56.66
	Understand words and language structures	12	212,562	7.36	3.03	0.78	1.43	61.31
	Basic understanding of variety of texts written in English	10	212,562	5.39	2.60	0.73	1.36	53.93
	Analyze and evaluate information and ideas in a variety of texts written in English	14	212,562	7.65	3.29	0.76	1.62	54.63
Writing	OVERALL TEST	42	211,288	23.38	8.99	0.93	2.42	51.08
	Express ideas in writing and address writing assignments	36	211,288	20.01	7.35	0.91	2.25	40.77
	Use standard grammar, usage, and spelling to edit writing tasks	6	211,288	3.37	2.11	0.81	0.91	56.24

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. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	36	102,595	20.60	8.05	0.90	2.55	57.22
	Understand words and language structures	12	102,595	7.29	2.98	0.77	1.43	60.74
	Basic understanding of variety of texts written in English	10	102,595	5.52	2.61	0.73	1.35	55.18
	Analyze and evaluate information and ideas in a variety of texts written in English	14	102,595	7.79	3.27	0.76	1.62	55.67
Writing	OVERALL TEST	42	102,029	24.56	9.12	0.93	2.48	53.30
	Express ideas in writing and address writing assignments	36	102,029	21.10	7.49	0.91	2.31	44.83
	Use standard grammar, usage, and spelling to edit writing tasks	6	102,029	3.45	2.09	0.81	0.91	57.54

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. 2024 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	36	109,925	20.21	8.17	0.90	2.56	56.15
	Understand words and language structures	12	109,925	7.42	3.08	0.79	1.42	61.85
	Basic understanding of variety of texts written in English	10	109,925	5.28	2.59	0.72	1.36	52.79
	Analyze and evaluate information and ideas in a variety of texts written in English	14	109,925	7.51	3.30	0.76	1.63	53.67
Writing	OVERALL TEST	42	109,224	22.29	8.72	0.93	2.37	49.02
	Express ideas in writing and address writing assignments	36	109,224	18.99	7.07	0.90	2.19	36.98
	Use standard grammar, usage, and spelling to edit writing tasks	6	109,224	3.30	2.12	0.82	0.91	55.04

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. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	210,523	21.85	8.02	0.89	2.62	59.04
	Understand words and language structures	12	210,523	7.84	3.07	0.79	1.41	65.37
	Basic understanding of variety of texts written in English	10	210,523	6.20	2.65	0.75	1.32	61.99
	Analyze and evaluate information and ideas in a variety of texts written in English	15	210,523	7.80	3.22	0.71	1.74	52.01
Writing	OVERALL TEST	42	208,923	24.99	8.29	0.90	2.57	50.29
	Express ideas in writing and address writing assignments	36	208,923	21.90	7.07	0.89	2.36	47.78
	Use standard grammar, usage, and spelling to edit writing tasks	6	208,923	3.09	1.72	0.64	1.03	51.55

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. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	101,070	22.66	7.94	0.89	2.60	61.24
	Understand words and language structures	12	101,070	8.00	3.05	0.79	1.40	66.67
	Basic understanding of variety of texts written in English	10	101,070	6.47	2.59	0.75	1.30	64.74
	Analyze and evaluate information and ideas in a variety of texts written in English	15	101,070	8.19	3.20	0.71	1.74	54.58
Writing	OVERALL TEST	42	100,386	26.44	8.32	0.90	2.59	53.42
	Express ideas in writing and address writing assignments	36	100,386	23.22	7.10	0.89	2.38	52.65
	Use standard grammar, usage, and spelling to edit writing tasks	6	100,386	3.23	1.71	0.64	1.02	53.81

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. 2024 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	109,422	21.10	8.02	0.89	2.64	57.02
	Understand words and language structures	12	109,422	7.70	3.09	0.79	1.42	64.19
	Basic understanding of variety of texts written in English	10	109,422	5.95	2.68	0.75	1.34	59.47
	Analyze and evaluate information and ideas in a variety of texts written in English	15	109,422	7.45	3.20	0.70	1.75	49.65
Writing	OVERALL TEST	42	108,515	23.65	8.04	0.90	2.53	47.40
	Express ideas in writing and address writing assignments	36	108,515	20.68	6.82	0.89	2.31	43.27
	Use standard grammar, usage, and spelling to edit writing tasks	6	108,515	2.97	1.73	0.64	1.03	49.47

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. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	211,015	21.58	8.40	0.91	2.58	58.32
	Understand words and language structures	12	211,015	7.46	2.98	0.78	1.40	62.14
	Basic understanding of variety of texts written in English	10	211,015	5.50	2.48	0.68	1.40	55.02
	Analyze and evaluate information and ideas in a variety of texts written in English	15	211,015	8.62	3.84	0.82	1.63	57.47
Writing	OVERALL TEST	42	205,943	24.40	7.96	0.89	2.64	58.76
	Express ideas in writing and address writing assignments	36	205,943	20.38	6.58	0.86	2.49	42.15
	Use standard grammar, usage, and spelling to edit writing tasks	6	205,943	4.02	1.96	0.80	0.87	67.07

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.4.14. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	99,468	22.23	8.29	0.90	2.56	60.09
	Understand words and language structures	12	99,468	7.57	2.88	0.77	1.39	63.05
	Basic understanding of variety of texts written in English	10	99,468	5.73	2.50	0.69	1.39	57.30
	Analyze and evaluate information and ideas in a variety of texts written in English	15	99,468	8.94	3.80	0.82	1.61	59.58
Writing	OVERALL TEST	42	97,318	25.87	7.90	0.88	2.73	62.28
	Express ideas in writing and address writing assignments	36	97,318	21.67	6.55	0.84	2.59	46.94
	Use standard grammar, usage, and spelling to edit writing tasks	6	97,318	4.20	1.92	0.81	0.85	69.95

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. 2024 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	111,493	21.00	8.45	0.91	2.59	56.76
	Understand words and language structures	12	111,493	7.36	3.07	0.79	1.41	61.34
	Basic understanding of variety of texts written in English	10	111,493	5.30	2.44	0.67	1.41	52.99
	Analyze and evaluate information and ideas in a variety of texts written in English	15	111,493	8.34	3.85	0.82	1.63	55.59
Writing	OVERALL TEST	42	108,592	23.09	7.77	0.89	2.55	55.63
	Express ideas in writing and address writing assignments	36	108,592	19.22	6.38	0.86	2.39	37.87
	Use standard grammar, usage, and spelling to edit writing tasks	6	108,592	3.87	1.98	0.80	0.89	64.51

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. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	205,754	21.62	8.03	0.90	2.53	58.43
	Understand words and language structures	12	205,754	8.90	2.95	0.82	1.24	74.16
	Basic understanding of variety of texts written in English	10	205,754	5.39	2.57	0.72	1.35	53.91
	Analyze and evaluate information and ideas in a variety of texts written in English	15	205,754	7.33	3.47	0.76	1.70	48.87
Writing	OVERALL TEST	42	199,518	23.73	8.33	0.90	2.69	50.35
	Express ideas in writing and address writing assignments	36	199,518	20.47	7.21	0.88	2.52	42.49
	Use standard grammar, usage, and spelling to edit writing tasks	6	199,518	3.26	1.72	0.70	0.94	54.27

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. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	96,087	22.03	7.90	0.90	2.52	59.54
	Understand words and language structures	12	96,087	8.96	2.89	0.82	1.24	74.69
	Basic understanding of variety of texts written in English	10	96,087	5.61	2.54	0.72	1.34	56.10
	Analyze and evaluate information and ideas in a variety of texts written in English	15	96,087	7.46	3.43	0.75	1.70	49.72
Writing	OVERALL TEST	42	93,490	25.61	8.19	0.89	2.77	53.97
	Express ideas in writing and address writing assignments	36	93,490	22.22	7.06	0.86	2.60	48.96
	Use standard grammar, usage, and spelling to edit writing tasks	6	93,490	3.39	1.71	0.70	0.93	56.47

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. 2024 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Reading	OVERALL TEST	37	109,644	21.26	8.12	0.90	2.53	57.46
	Understand words and language structures	12	109,644	8.84	3.01	0.83	1.24	73.69
	Basic understanding of variety of texts written in English	10	109,644	5.20	2.58	0.72	1.36	52.00
	Analyze and evaluate information and ideas in a variety of texts written in English	15	109,644	7.22	3.50	0.76	1.70	48.13
Writing	OVERALL TEST	42	106,010	22.08	8.10	0.90	2.61	47.16
	Express ideas in writing and address writing assignments	36	106,010	18.94	6.99	0.88	2.44	36.80
	Use standard grammar, usage, and spelling to edit writing tasks	6	106,010	3.14	1.71	0.70	0.94	52.34

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.4.19. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	210,709	17.75	6.13	0.88	2.09	65.75
	Understand spoken words and language structures	5	210,709	3.33	1.39	0.61	0.87	66.63
	Basic understanding of spoken English	16	210,709	10.36	3.65	0.80	1.65	64.72
	Analyze and evaluate information in spoken English	6	210,709	4.07	1.73	0.70	0.94	67.77
Speaking	OVERALL TEST	36	210,709	20.98	6.32	0.93	1.67	46.91
	Provide and summarize information	16	210,709	9.64	2.63	0.82	1.13	51.53
	Share opinions and analyze information	20	210,709	11.35	3.88	0.91	1.18	42.30

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.20. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	102,404	17.85	5.94	0.88	2.09	66.12
	Understand spoken words and language structures	5	102,404	3.34	1.38	0.61	0.86	66.71
	Basic understanding of spoken English	16	102,404	10.42	3.51	0.78	1.65	65.10
	Analyze and evaluate information in spoken English	6	102,404	4.10	1.68	0.69	0.94	68.32
Speaking	OVERALL TEST	36	102,404	21.50	6.27	0.93	1.69	48.81
	Provide and summarize information	16	102,404	9.82	2.60	0.81	1.14	53.13
	Share opinions and analyze information	20	102,404	11.67	3.86	0.90	1.19	44.48

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.21. 2024 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	108,237	17.67	6.30	0.89	2.09	65.43
	Understand spoken words and language structures	5	108,237	3.33	1.39	0.60	0.88	66.58
	Basic understanding of spoken English	16	108,237	10.30	3.77	0.81	1.64	64.38
	Analyze and evaluate information in spoken English	6	108,237	4.04	1.76	0.71	0.94	67.27
Speaking	OVERALL TEST	36	108,237	20.50	6.33	0.93	1.65	45.14
	Provide and summarize information	16	108,237	9.46	2.64	0.82	1.11	50.03
	Share opinions and analyze information	20	108,237	11.04	3.88	0.91	1.16	40.26

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.22. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	212,344	18.55	5.93	0.88	2.05	68.69
	Understand spoken words and language structures	5	212,344	3.78	1.33	0.63	0.82	75.69
	Basic understanding of spoken English	16	212,344	11.04	3.46	0.79	1.59	69.01
	Analyze and evaluate information in spoken English	6	212,344	3.72	1.78	0.68	1.01	62.01
Speaking	OVERALL TEST	36	212,344	23.84	6.08	0.92	1.73	57.46
	Provide and summarize information	16	212,344	11.59	2.85	0.80	1.28	66.62
	Share opinions and analyze information	20	212,344	12.25	3.45	0.89	1.15	48.31

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.23. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	102,509	18.58	5.81	0.87	2.06	68.81
	Understand spoken words and language structures	5	102,509	3.75	1.32	0.61	0.83	75.05
	Basic understanding of spoken English	16	102,509	11.07	3.36	0.78	1.58	69.16
	Analyze and evaluate information in spoken English	6	102,509	3.76	1.77	0.68	1.01	62.67
Speaking	OVERALL TEST	36	102,509	24.00	6.04	0.92	1.75	58.03
	Provide and summarize information	16	102,509	11.64	2.83	0.79	1.30	66.98
	Share opinions and analyze information	20	102,509	12.36	3.44	0.89	1.16	49.07

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.24. 2024 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	109,793	18.52	6.04	0.89	2.05	68.60
	Understand spoken words and language structures	5	109,793	3.82	1.34	0.64	0.80	76.32
	Basic understanding of spoken English	16	109,793	11.02	3.54	0.80	1.59	68.89
	Analyze and evaluate information in spoken English	6	109,793	3.68	1.79	0.68	1.01	61.41
Speaking	OVERALL TEST	36	109,793	23.68	6.12	0.92	1.70	56.95
	Provide and summarize information	16	109,793	11.54	2.86	0.80	1.27	66.29
	Share opinions and analyze information	20	109,793	12.14	3.46	0.89	1.13	47.61

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.25. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	311,814	18.16	5.77	0.86	2.12	67.25
	Understand spoken words and language structures	5	311,814	3.48	1.41	0.62	0.87	69.65
	Basic understanding of spoken English	16	311,814	10.94	3.50	0.79	1.62	68.37
	Analyze and evaluate information in spoken English	6	311,814	3.74	1.58	0.55	1.06	62.28
Speaking	OVERALL TEST	36	311,814	22.27	6.02	0.92	1.66	52.27
	Provide and summarize information	16	311,814	10.21	2.56	0.82	1.08	57.44
	Share opinions and analyze information	20	311,814	12.06	3.64	0.88	1.23	47.10

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.26. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	149,173	18.48	5.67	0.86	2.09	68.45
	Understand spoken words and language structures	5	149,173	3.49	1.40	0.63	0.86	69.89
	Basic understanding of spoken English	16	149,173	11.17	3.43	0.79	1.59	69.82
	Analyze and evaluate information in spoken English	6	149,173	3.82	1.55	0.54	1.05	63.60
Speaking	OVERALL TEST	36	149,173	22.41	5.98	0.92	1.69	52.80
	Provide and summarize information	16	149,173	10.24	2.54	0.81	1.10	57.78
	Share opinions and analyze information	20	149,173	12.17	3.62	0.88	1.26	47.83

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.27. 2024 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	162,592	17.87	5.84	0.86	2.15	66.17
	Understand spoken words and language structures	5	162,592	3.47	1.41	0.62	0.87	69.45
	Basic understanding of spoken English	16	162,592	10.73	3.55	0.79	1.64	67.05
	Analyze and evaluate information in spoken English	6	162,592	3.67	1.61	0.56	1.07	61.09
Speaking	OVERALL TEST	36	162,592	22.15	6.06	0.93	1.63	51.79
	Provide and summarize information	16	162,592	10.18	2.57	0.83	1.06	57.14
	Share opinions and analyze information	20	162,592	11.97	3.65	0.89	1.21	46.43

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.28. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	314,689	18.67	5.88	0.88	2.01	69.15
	Understand spoken words and language structures	5	314,689	3.71	1.36	0.65	0.81	74.29
	Basic understanding of spoken English	16	314,689	10.86	3.56	0.81	1.57	67.90
	Analyze and evaluate information in spoken English	6	314,689	4.09	1.56	0.61	0.97	68.19
Speaking	OVERALL TEST	36	314,689	21.76	8.53	0.96	1.76	47.95
	Provide and summarize information	16	314,689	10.29	3.81	0.88	1.29	52.81
	Share opinions and analyze information	20	314,689	11.46	4.87	0.94	1.16	43.08

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.29. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	147,129	18.89	5.67	0.88	1.99	69.95
	Understand spoken words and language structures	5	147,129	3.75	1.33	0.63	0.80	74.99
	Basic understanding of spoken English	16	147,129	11.03	3.46	0.80	1.55	68.92
	Analyze and evaluate information in spoken English	6	147,129	4.11	1.49	0.58	0.96	68.51
Speaking	OVERALL TEST	36	147,129	22.17	8.53	0.95	1.81	49.52
	Provide and summarize information	16	147,129	10.45	3.81	0.88	1.33	54.24
	Share opinions and analyze information	20	147,129	11.72	4.88	0.94	1.20	44.80

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.
3. Mean of percent correct (0–100%) for the items

Table D.4.30. 2024 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha[2]	SEM	Mean P-Value[3]
Listening	OVERALL TEST	27	167,503	18.48	6.05	0.89	2.03	68.46
	Understand spoken words and language structures	5	167,503	3.68	1.38	0.65	0.81	73.68
	Basic understanding of spoken English	16	167,503	10.72	3.64	0.81	1.58	67.03
	Analyze and evaluate information in spoken English	6	167,503	4.08	1.61	0.63	0.98	67.92
Speaking	OVERALL TEST	36	167,503	21.39	8.51	0.96	1.72	46.58
	Provide and summarize information	16	167,503	10.15	3.81	0.89	1.26	51.57
	Share opinions and analyze information	20	167,503	11.24	4.84	0.94	1.14	41.58

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.
3. Mean of percent correct (0–100%) for the items

2024 TELPAS
Scale Score Descriptive Statistics

Table D.5.1. 2024 TELPAS Reading Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Reading	105,149	1468.21	1450	1428	642	94	74.97	5619.97	0.76	0.97
Grade 3	Reading	105,769	1512.00	1496	1420	881	148	114.91	13203.52	0.80	0.75
Grade 4	Reading	105,986	1547.48	1531	1463	876	135	104.37	10892.99	0.59	0.52
Grade 5	Reading	106,576	1585.37	1575	1703	876	168	118.09	13945.28	0.46	0.18
Grade 6	Reading	105,551	1541.14	1540	1633	795	131	90.48	8185.81	0.16	-0.23
Grade 7	Reading	104,972	1552.81	1550	1618	795	144	94.97	9019.09	0.15	-0.16
Grade 8	Reading	101,536	1553.44	1545	1639	698	127	89.00	7921.04	0.45	0.18
Grade 9	Reading	109,479	1547.34	1537	1468	698	136	92.53	8562.53	0.37	0.39
Grade 10	Reading	86,230	1524.89	1516	1499	677	116	86.83	7538.75	0.26	0.61
Grade 11	Reading	68,308	1531.14	1524	1499	677	119	86.85	7543.34	0.19	0.73
Grade 12	Reading	51,216	1525.71	1516	1507	677	107	84.45	7131.80	0.14	1.31

Table D.5.2. 2024 TELPAS Writing Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Writing	102,632	1448.43	1435	1356	396	153	77.25	5966.92	0.36	-0.88
Grade 3	Writing	104,842	1444.47	1456	1320	497	163	102.66	10539.26	-0.20	-0.85
Grade 4	Writing	105,295	1465.96	1485	1260	454	151	107.20	11492.60	-0.45	-0.61
Grade 5	Writing	105,993	1498.72	1526	1260	454	134	107.10	11470.73	-0.68	-0.10
Grade 6	Writing	104,860	1501.80	1519	1291	437	90	87.71	7693.43	-0.70	0.36
Grade 7	Writing	104,063	1513.71	1527	1291	437	108	88.90	7903.74	-0.71	0.49
Grade 8	Writing	100,225	1486.78	1505	1494	471	107	92.75	8602.47	-0.91	0.89
Grade 9	Writing	105,718	1479.62	1494	1494	471	118	97.68	9541.44	-0.85	0.50
Grade 10	Writing	83,503	1508.96	1522	1249	539	124	105.10	11046.60	-0.66	0.32
Grade 11	Writing	66,244	1517.84	1531	1604	539	127	100.67	10134.79	-0.69	0.58
Grade 12	Writing	49,771	1513.22	1522	1569	539	115	96.24	9261.79	-0.65	0.62

Table D.5.3. 2024 TELPAS Listening Scale Score Descriptive Statistics

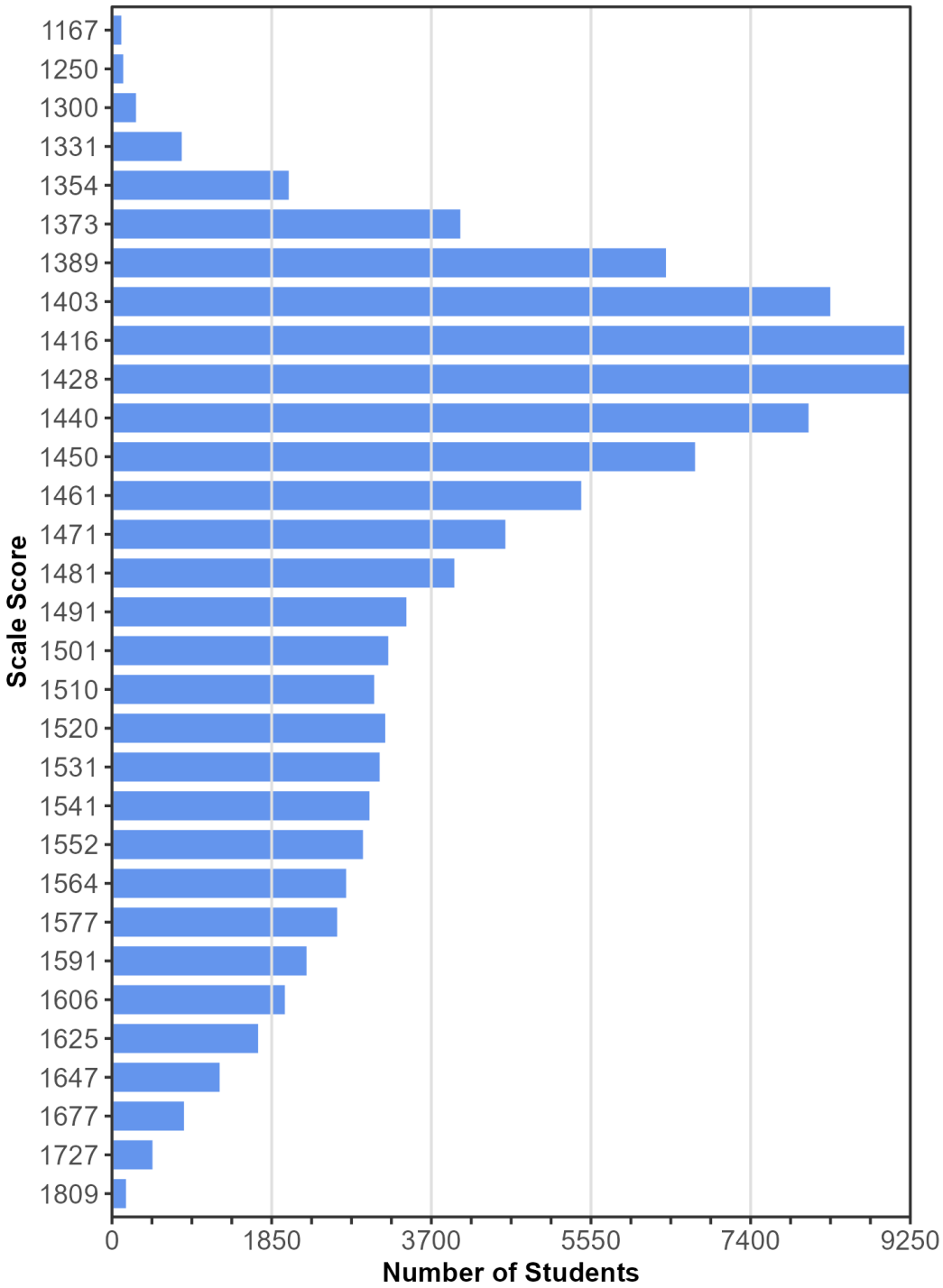
Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Listening	105,029	1555.54	1546	1571	647	126	89.51	8012.56	0.40	0.46
Grade 3	Listening	105,680	1613.40	1613	1705	647	127	103.68	10748.93	0.17	-0.23
Grade 4	Listening	105,860	1538.73	1541	1604	620	122	89.38	7988.78	0.28	0.14
Grade 5	Listening	106,484	1566.63	1569	1657	620	122	97.89	9581.92	0.12	-0.17
Grade 6	Listening	105,521	1558.08	1551	1609	640	105	86.88	7548.88	0.32	0.40
Grade 7	Listening	104,861	1567.83	1564	1628	640	124	91.48	8369.42	0.29	0.22
Grade 8	Listening	101,432	1579.36	1577	1651	640	136	95.42	9105.73	0.20	0.07
Grade 9	Listening	109,302	1540.29	1549	1592	527	106	80.55	6488.79	-0.17	0.21
Grade 10	Listening	86,046	1547.97	1549	1611	527	96	78.18	6111.51	-0.21	0.41
Grade 11	Listening	68,221	1552.26	1562	1611	527	105	76.21	5807.52	-0.23	0.66
Grade 12	Listening	51,120	1547.24	1549	1592	527	96	75.18	5651.74	-0.30	1.08

Table D.5.4. 2024 TELPAS Speaking Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Speaking	105,029	1435.49	1436	1262	537	122	95.20	9063.57	0.01	-0.14
Grade 3	Speaking	105,680	1473.33	1474	1522	537	127	101.67	10337.60	-0.19	-0.11
Grade 4	Speaking	105,860	1513.66	1514	1554	362	78	65.01	4226.06	-0.21	0.31
Grade 5	Speaking	106,484	1519.61	1525	1554	362	98	69.03	4764.99	-0.11	0.16
Grade 6	Speaking	105,521	1515.36	1514	1533	293	72	55.41	3070.01	-0.18	0.34
Grade 7	Speaking	104,861	1509.30	1514	1514	293	69	57.92	3355.07	-0.12	0.11
Grade 8	Speaking	101,432	1509.01	1514	1523	293	69	60.09	3610.25	-0.14	0.00
Grade 9	Speaking	109,302	1507.08	1512	1415	252	90	65.91	4343.66	0.15	-0.74
Grade 10	Speaking	86,046	1510.59	1512	1415	252	105	66.86	4470.59	0.09	-0.77
Grade 11	Speaking	68,221	1513.97	1518	1415	252	105	68.11	4639.43	0.07	-0.78
Grade 12	Speaking	51,120	1507.71	1512	1415	252	120	67.95	4617.81	0.17	-0.79

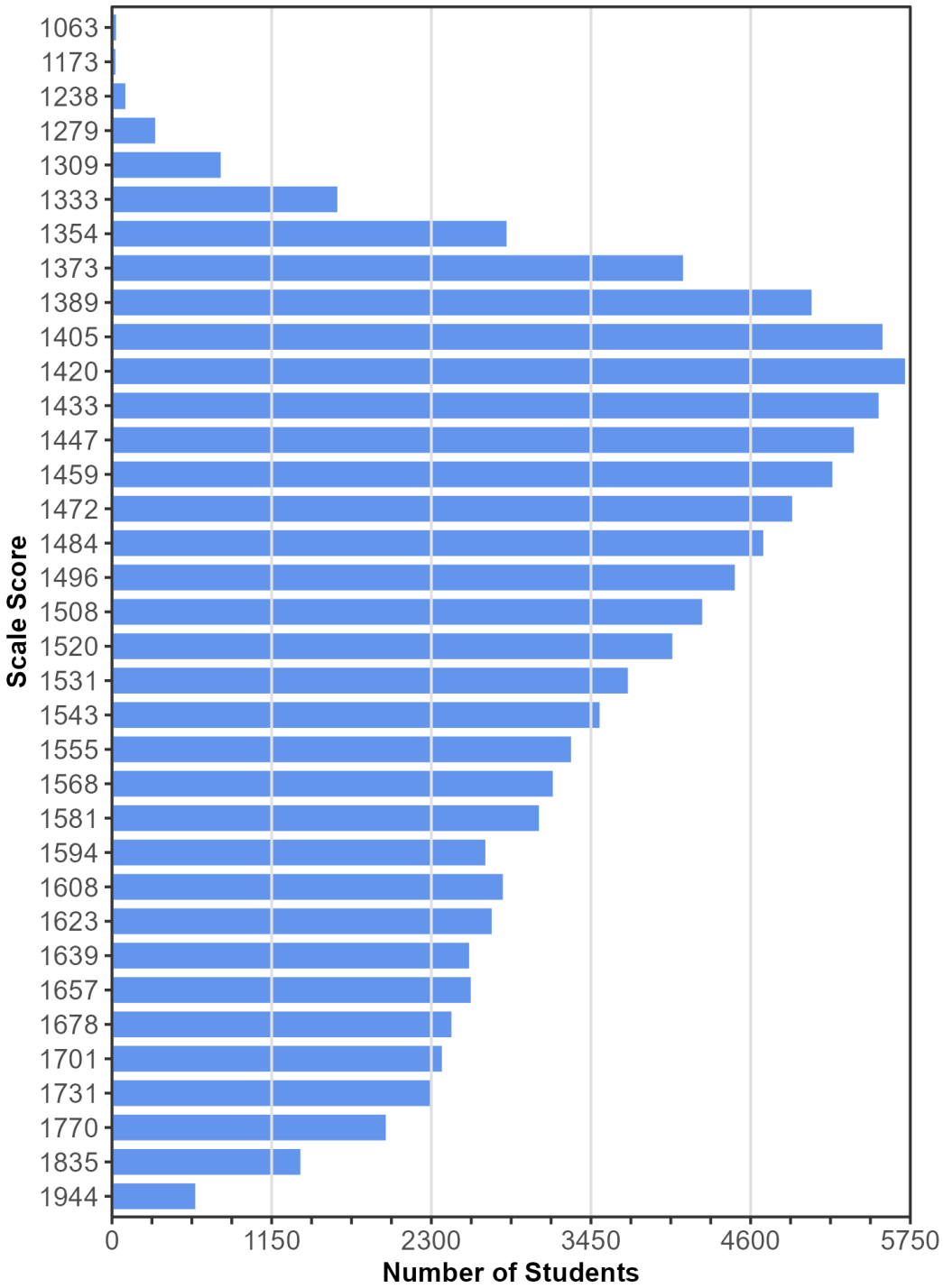
2024 TELPAS
Frequency Distribution of Scale Scores

Figure D.5.1. 2024 TELPAS Grade 2 Reading
Frequency Distribution of Scale Scores



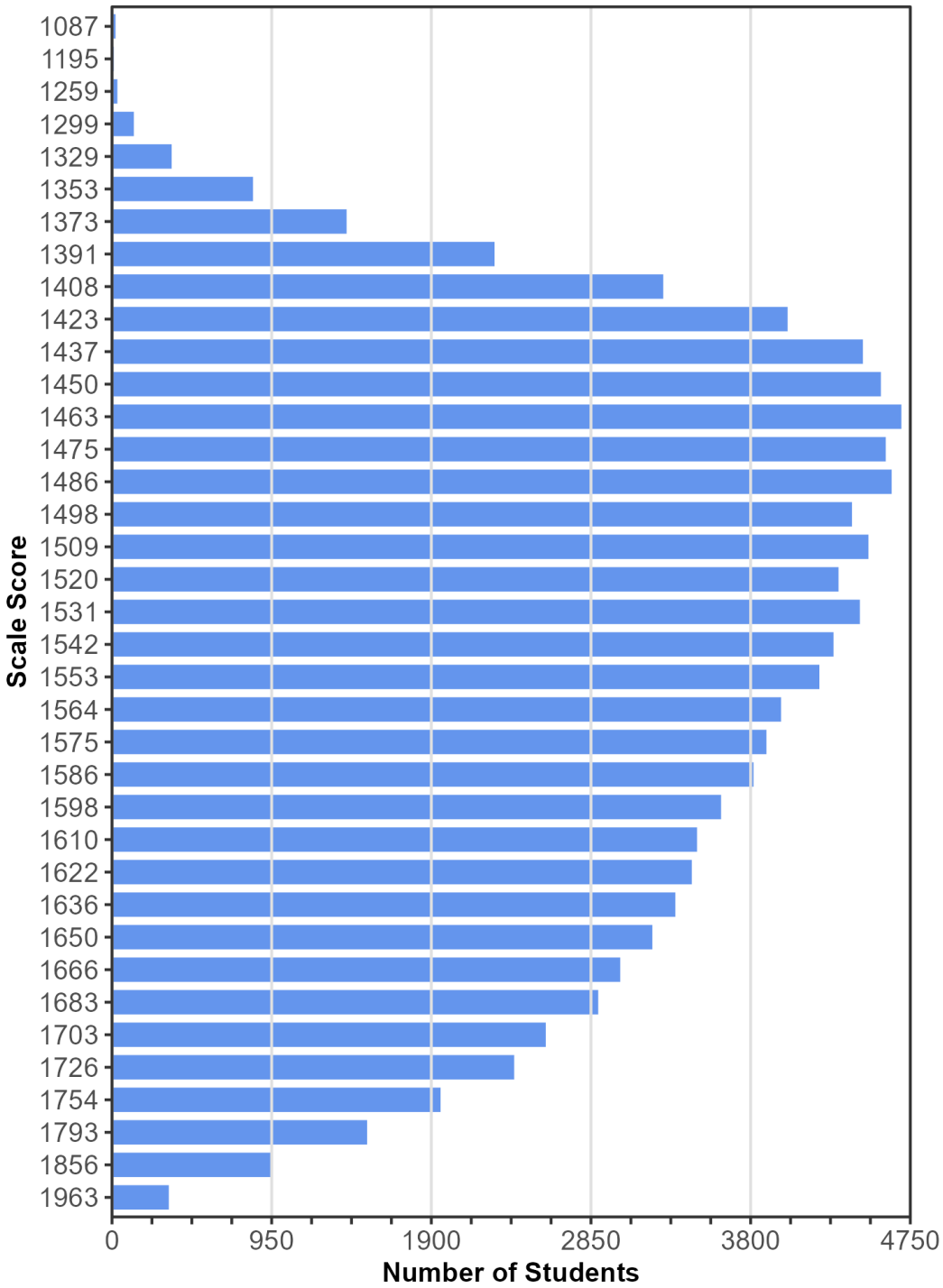
	Freq	Cum Freq	Pct	Cum Pct
1167	108	108	0.10	0.10
1250	130	238	0.12	0.23
1300	278	516	0.26	0.49
1331	808	1324	0.77	1.26
1354	2048	3372	1.95	3.21
1373	4036	7408	3.84	7.05
1389	6419	13827	6.10	13.15
1403	8323	22150	7.92	21.07
1416	9180	31330	8.73	29.80
1428	9238	40568	8.79	38.58
1440	8071	48639	7.68	46.26
1450	6756	55395	6.43	52.68
1461	5437	60832	5.17	57.85
1471	4558	65390	4.33	62.19
1481	3967	69357	3.77	65.96
1491	3411	72768	3.24	69.20
1501	3201	75969	3.04	72.25
1510	3038	79007	2.89	75.14
1520	3166	82173	3.01	78.15
1531	3101	85274	2.95	81.10
1541	2983	88257	2.84	83.94
1552	2908	91165	2.77	86.70
1564	2713	93878	2.58	89.28
1577	2609	96487	2.48	91.76
1591	2254	98741	2.14	93.91
1606	2003	100744	1.90	95.81
1625	1693	102437	1.61	97.42
1647	1247	103684	1.19	98.61
1677	834	104518	0.79	99.40
1727	469	104987	0.45	99.85
1809	162	105149	0.15	100.00

Figure D.5.2. 2024 TELPAS Grade 3 Reading
Frequency Distribution of Scale Scores



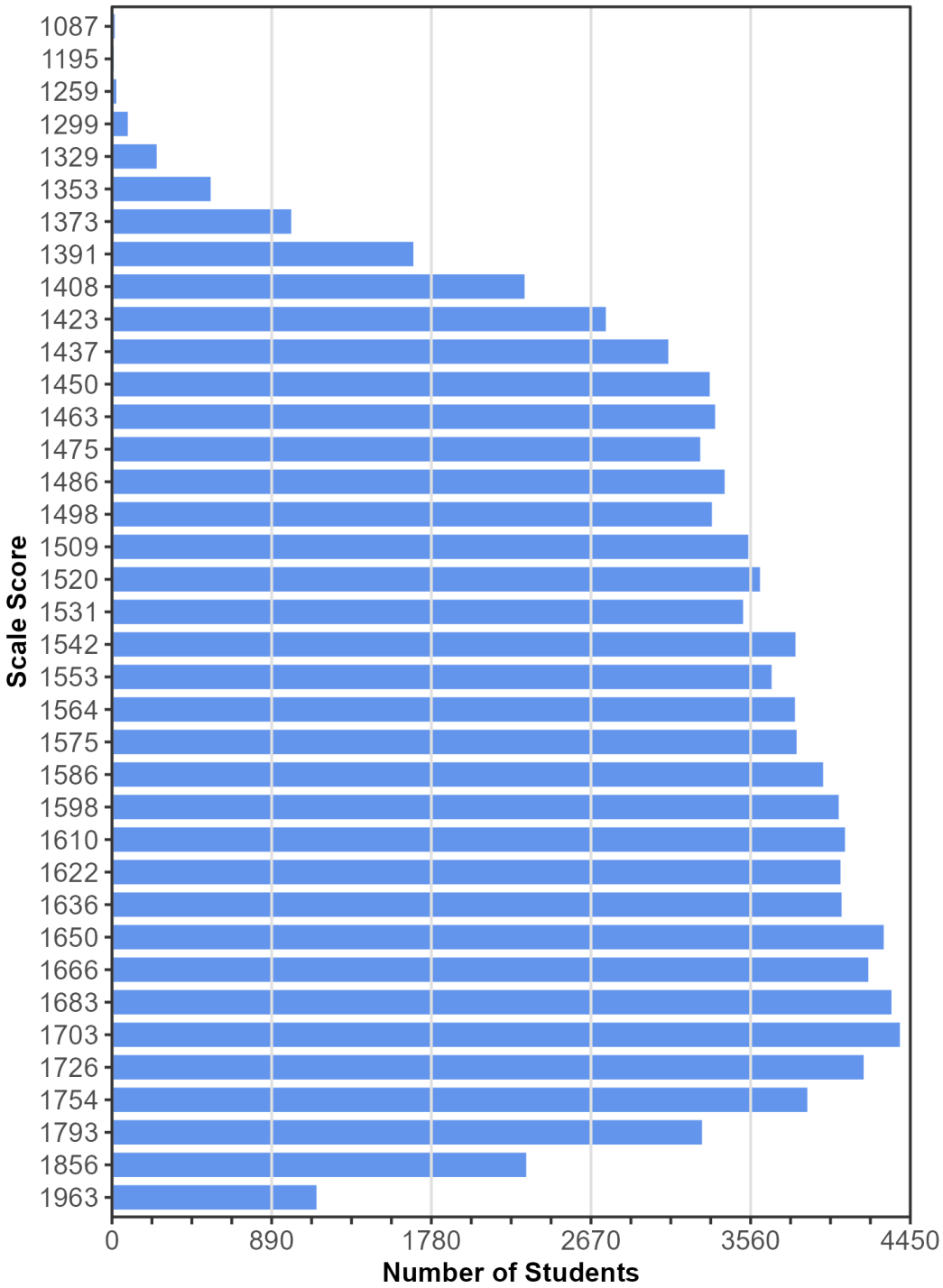
	Freq	Cum Freq	Pct	Cum Pct
1063	29	29	0.03	0.03
1173	25	54	0.02	0.05
1238	96	150	0.09	0.14
1279	311	461	0.29	0.44
1309	783	1244	0.74	1.18
1333	1623	2867	1.53	2.71
1354	2842	5709	2.69	5.40
1373	4113	9822	3.89	9.29
1389	5039	14861	4.76	14.05
1405	5550	20411	5.25	19.30
1420	5712	26123	5.40	24.70
1433	5522	31645	5.22	29.92
1447	5344	36989	5.05	34.97
1459	5189	42178	4.91	39.88
1472	4899	47077	4.63	44.51
1484	4691	51768	4.44	48.94
1496	4486	56254	4.24	53.19
1508	4251	60505	4.02	57.20
1520	4036	64541	3.82	61.02
1531	3716	68257	3.51	64.53
1543	3512	71769	3.32	67.85
1555	3306	75075	3.13	70.98
1568	3175	78250	3.00	73.98
1581	3075	81325	2.91	76.89
1594	2689	84014	2.54	79.43
1608	2815	86829	2.66	82.09
1623	2735	89564	2.59	84.68
1639	2572	92136	2.43	87.11
1657	2584	94720	2.44	89.55
1678	2445	97165	2.31	91.87
1701	2376	99541	2.25	94.11
1731	2299	101840	2.17	96.29
1770	1972	103812	1.86	98.15
1835	1357	105169	1.28	99.43
1944	600	105769	0.57	100.00

Figure D.5.3. 2024 TELPAS Grade 4 Reading
Frequency Distribution of Scale Scores



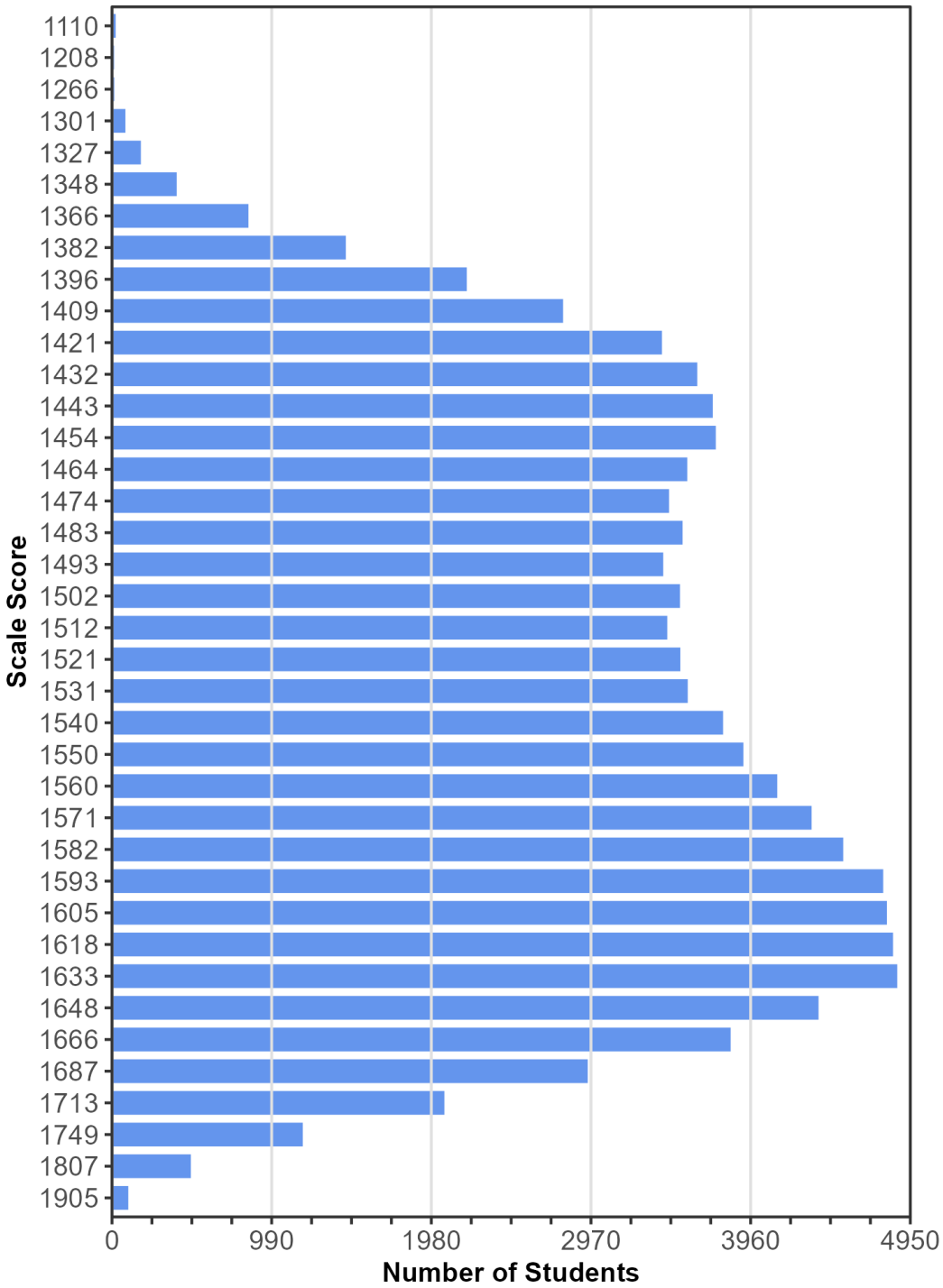
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1087	21	21	0.02	0.02
1195	11	32	0.01	0.03
1259	32	64	0.03	0.06
1299	130	194	0.12	0.18
1329	355	549	0.33	0.52
1353	839	1388	0.79	1.31
1373	1396	2784	1.32	2.63
1391	2276	5060	2.15	4.77
1408	3280	8340	3.09	7.87
1423	4020	12360	3.79	11.66
1437	4468	16828	4.22	15.88
1450	4575	21403	4.32	20.19
1463	4697	26100	4.43	24.63
1475	4604	30704	4.34	28.97
1486	4639	35343	4.38	33.35
1498	4403	39746	4.15	37.50
1509	4501	44247	4.25	41.75
1520	4323	48570	4.08	45.83
1531	4450	53020	4.20	50.03
1542	4293	57313	4.05	54.08
1553	4209	61522	3.97	58.05
1564	3981	65503	3.76	61.80
1575	3894	69397	3.67	65.48
1586	3817	73214	3.60	69.08
1598	3624	76838	3.42	72.50
1610	3481	80319	3.28	75.78
1622	3450	83769	3.26	79.04
1636	3352	87121	3.16	82.20
1650	3215	90336	3.03	85.23
1666	3024	93360	2.85	88.09
1683	2893	96253	2.73	90.82
1703	2581	98834	2.44	93.25
1726	2393	101227	2.26	95.51
1754	1955	103182	1.84	97.35
1793	1518	104700	1.43	98.79
1856	948	105648	0.89	99.68
1963	338	105986	0.32	100.00

Figure D.5.4. 2024 TELPAS Grade 5 Reading
Frequency Distribution of Scale Scores



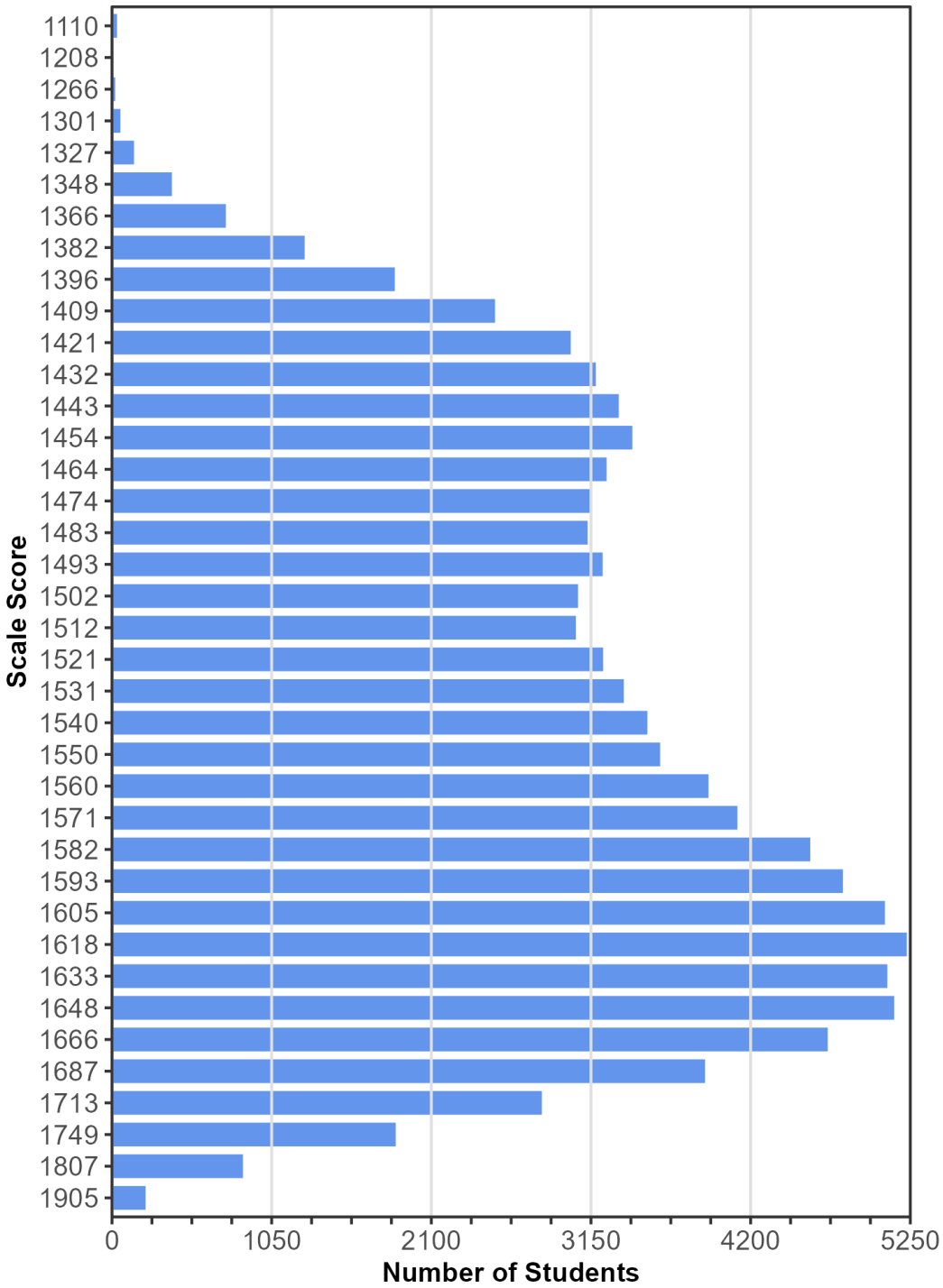
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1087	15	15	0.01	0.01
1195	10	25	0.01	0.02
1259	24	49	0.02	0.05
1299	88	137	0.08	0.13
1329	249	386	0.23	0.36
1353	550	936	0.52	0.88
1373	999	1935	0.94	1.82
1391	1680	3615	1.58	3.39
1408	2300	5915	2.16	5.55
1423	2753	8668	2.58	8.13
1437	3101	11769	2.91	11.04
1450	3332	15101	3.13	14.17
1463	3362	18463	3.15	17.32
1475	3279	21742	3.08	20.40
1486	3415	25157	3.20	23.60
1498	3344	28501	3.14	26.74
1509	3547	32048	3.33	30.07
1520	3612	35660	3.39	33.46
1531	3518	39178	3.30	36.76
1542	3810	42988	3.57	40.34
1553	3677	46665	3.45	43.79
1564	3807	50472	3.57	47.36
1575	3816	54288	3.58	50.94
1586	3964	58252	3.72	54.66
1598	4051	62303	3.80	58.46
1610	4086	66389	3.83	62.29
1622	4061	70450	3.81	66.10
1636	4067	74517	3.82	69.92
1650	4302	78819	4.04	73.96
1666	4216	83035	3.96	77.91
1683	4345	87380	4.08	81.99
1703	4392	91772	4.12	86.11
1726	4190	95962	3.93	90.04
1754	3876	99838	3.64	93.68
1793	3289	103127	3.09	96.76
1856	2309	105436	2.17	98.93
1963	1140	106576	1.07	100.00

Figure D.5.5. 2024 TELPAS Grade 6 Reading
Frequency Distribution of Scale Scores



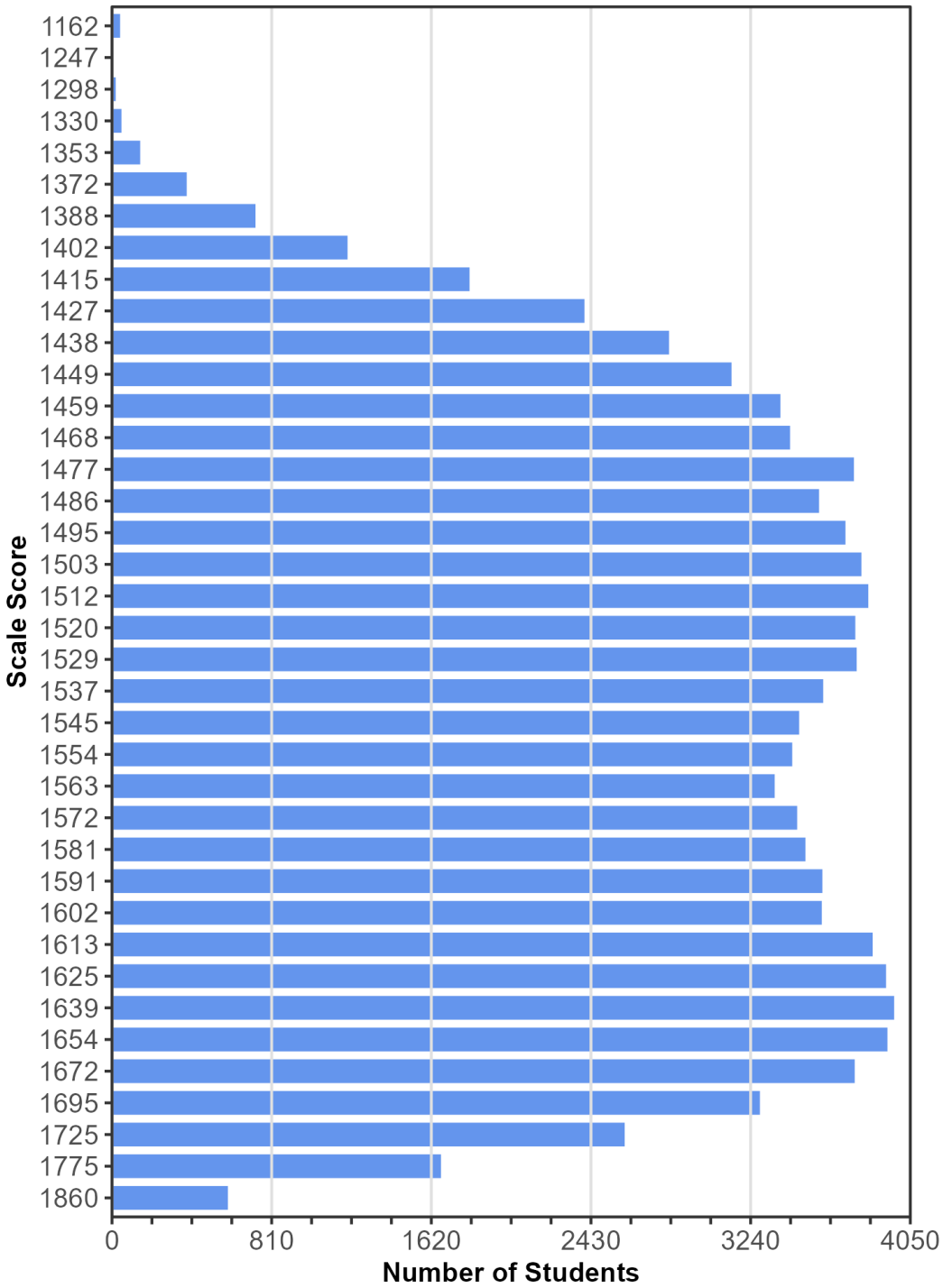
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1110	23	23	0.02	0.02
1208	13	36	0.01	0.03
1266	14	50	0.01	0.05
1301	83	133	0.08	0.13
1327	179	312	0.17	0.30
1348	401	713	0.38	0.68
1366	846	1559	0.80	1.48
1382	1450	3009	1.37	2.85
1396	2200	5209	2.08	4.94
1409	2797	8006	2.65	7.58
1421	3410	11416	3.23	10.82
1432	3629	15045	3.44	14.25
1443	3725	18770	3.53	17.78
1454	3744	22514	3.55	21.33
1464	3567	26081	3.38	24.71
1474	3454	29535	3.27	27.98
1483	3538	33073	3.35	31.33
1493	3418	36491	3.24	34.57
1502	3522	40013	3.34	37.91
1512	3443	43456	3.26	41.17
1521	3524	46980	3.34	44.51
1531	3570	50550	3.38	47.89
1540	3789	54339	3.59	51.48
1550	3915	58254	3.71	55.19
1560	4125	62379	3.91	59.10
1571	4338	66717	4.11	63.21
1582	4534	71251	4.30	67.50
1593	4782	76033	4.53	72.03
1605	4805	80838	4.55	76.59
1618	4843	85681	4.59	81.17
1633	4869	90550	4.61	85.79
1648	4381	94931	4.15	89.94
1666	3836	98767	3.63	93.57
1687	2950	101717	2.79	96.37
1713	2061	103778	1.95	98.32
1749	1183	104961	1.12	99.44
1807	489	105450	0.46	99.90
1905	101	105551	0.10	100.00

Figure D.5.6. 2024 TELPAS Grade 7 Reading
Frequency Distribution of Scale Scores



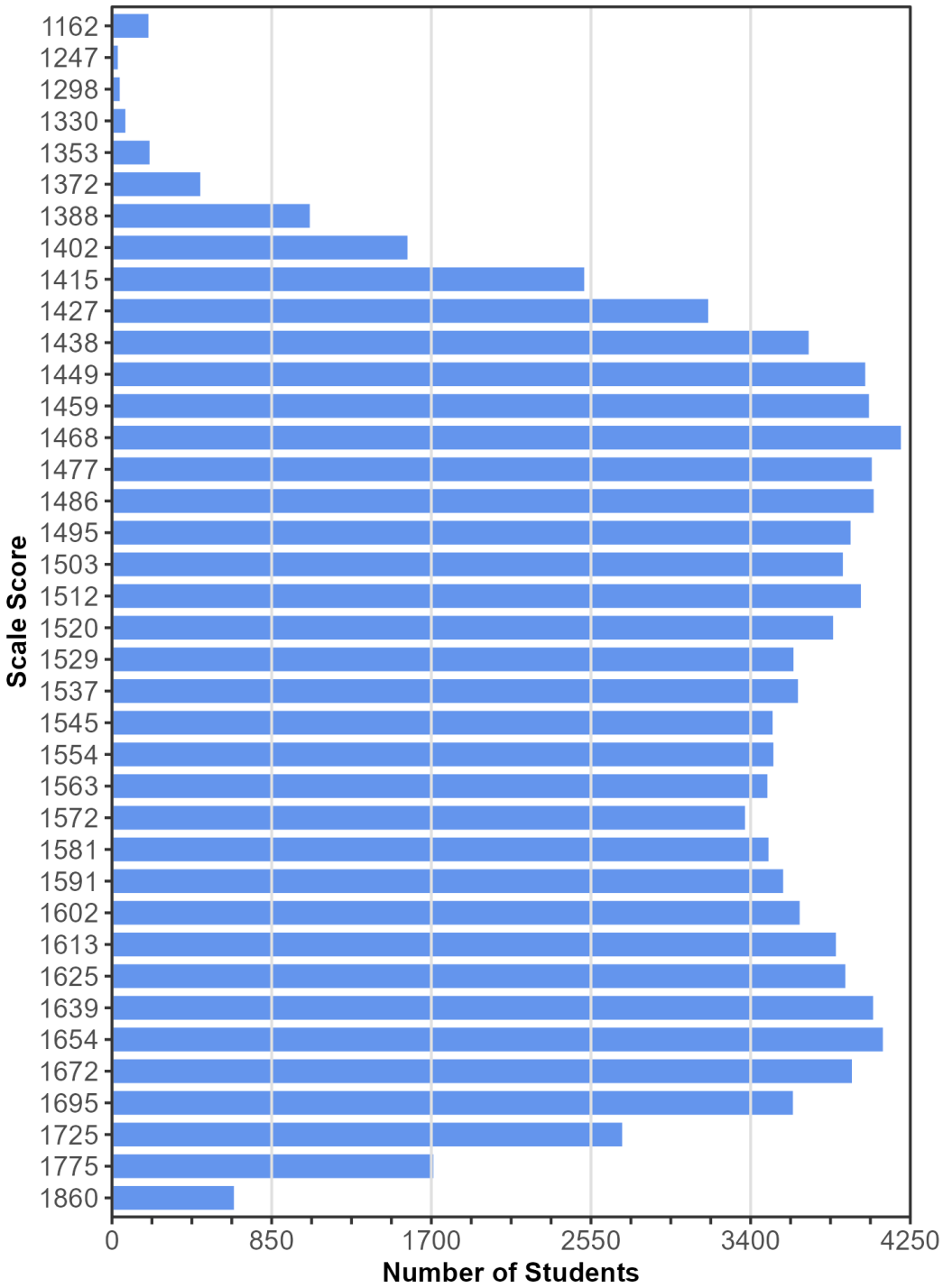
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1110	33	33	0.03	0.03
1208	5	38	0.00	0.04
1266	21	59	0.02	0.06
1301	55	114	0.05	0.11
1327	145	259	0.14	0.25
1348	394	653	0.38	0.62
1366	749	1402	0.71	1.34
1382	1267	2669	1.21	2.54
1396	1860	4529	1.77	4.31
1409	2519	7048	2.40	6.71
1421	3017	10065	2.87	9.59
1432	3182	13247	3.03	12.62
1443	3333	16580	3.18	15.79
1454	3422	20002	3.26	19.05
1464	3253	23255	3.10	22.15
1474	3154	26409	3.00	25.16
1483	3128	29537	2.98	28.14
1493	3227	32764	3.07	31.21
1502	3065	35829	2.92	34.13
1512	3051	38880	2.91	37.04
1521	3230	42110	3.08	40.12
1531	3366	45476	3.21	43.32
1540	3521	48997	3.35	46.68
1550	3605	52602	3.43	50.11
1560	3923	56525	3.74	53.85
1571	4113	60638	3.92	57.77
1582	4592	65230	4.37	62.14
1593	4807	70037	4.58	66.72
1605	5083	75120	4.84	71.56
1618	5227	80347	4.98	76.54
1633	5099	85446	4.86	81.40
1648	5144	90590	4.90	86.30
1666	4707	95297	4.48	90.78
1687	3900	99197	3.72	94.50
1713	2827	102024	2.69	97.19
1749	1866	103890	1.78	98.97
1807	861	104751	0.82	99.79
1905	221	104972	0.21	100.00

Figure D.5.7. 2024 TELPAS Grade 8 Reading
Frequency Distribution of Scale Scores



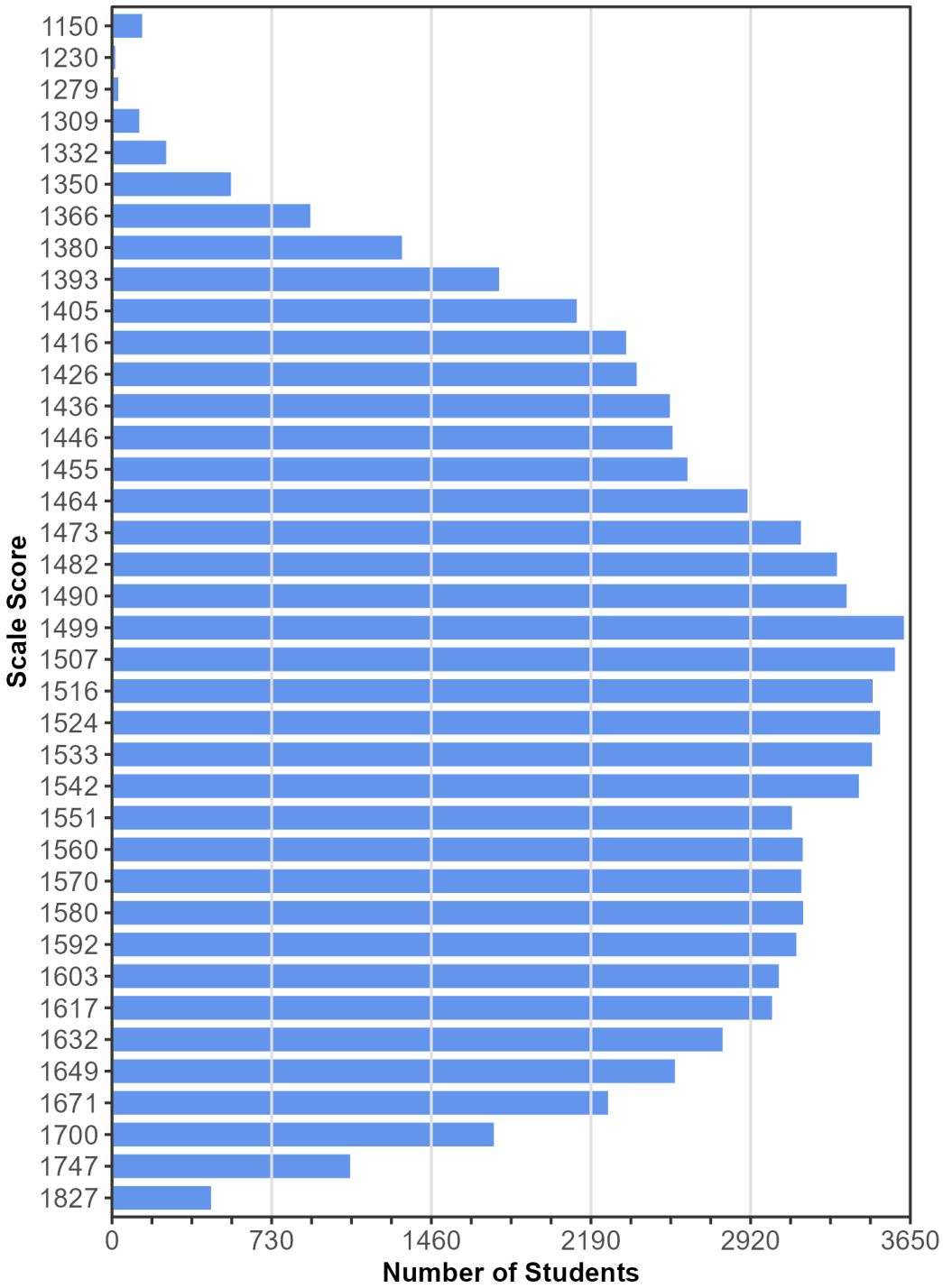
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1162	41	41	0.04	0.04
1247	5	46	0.00	0.05
1298	19	65	0.02	0.06
1330	48	113	0.05	0.11
1353	143	256	0.14	0.25
1372	379	635	0.37	0.63
1388	728	1363	0.72	1.34
1402	1195	2558	1.18	2.52
1415	1814	4372	1.79	4.31
1427	2397	6769	2.36	6.67
1438	2826	9595	2.78	9.45
1449	3143	12738	3.10	12.55
1459	3391	16129	3.34	15.89
1468	3440	19569	3.39	19.27
1477	3764	23333	3.71	22.98
1486	3587	26920	3.53	26.51
1495	3721	30641	3.66	30.18
1503	3802	34443	3.74	33.92
1512	3837	38280	3.78	37.70
1520	3771	42051	3.71	41.41
1529	3778	45829	3.72	45.14
1537	3608	49437	3.55	48.69
1545	3486	52923	3.43	52.12
1554	3451	56374	3.40	55.52
1563	3362	59736	3.31	58.83
1572	3476	63212	3.42	62.26
1581	3518	66730	3.46	65.72
1591	3604	70334	3.55	69.27
1602	3601	73935	3.55	72.82
1613	3859	77794	3.80	76.62
1625	3927	81721	3.87	80.48
1639	3968	85689	3.91	84.39
1654	3934	89623	3.87	88.27
1672	3768	93391	3.71	91.98
1695	3287	96678	3.24	95.22
1725	2601	99279	2.56	97.78
1775	1669	100948	1.64	99.42
1860	588	101536	0.58	100.00

Figure D.5.8. 2024 TELPAS Grade 9 Reading
Frequency Distribution of Scale Scores



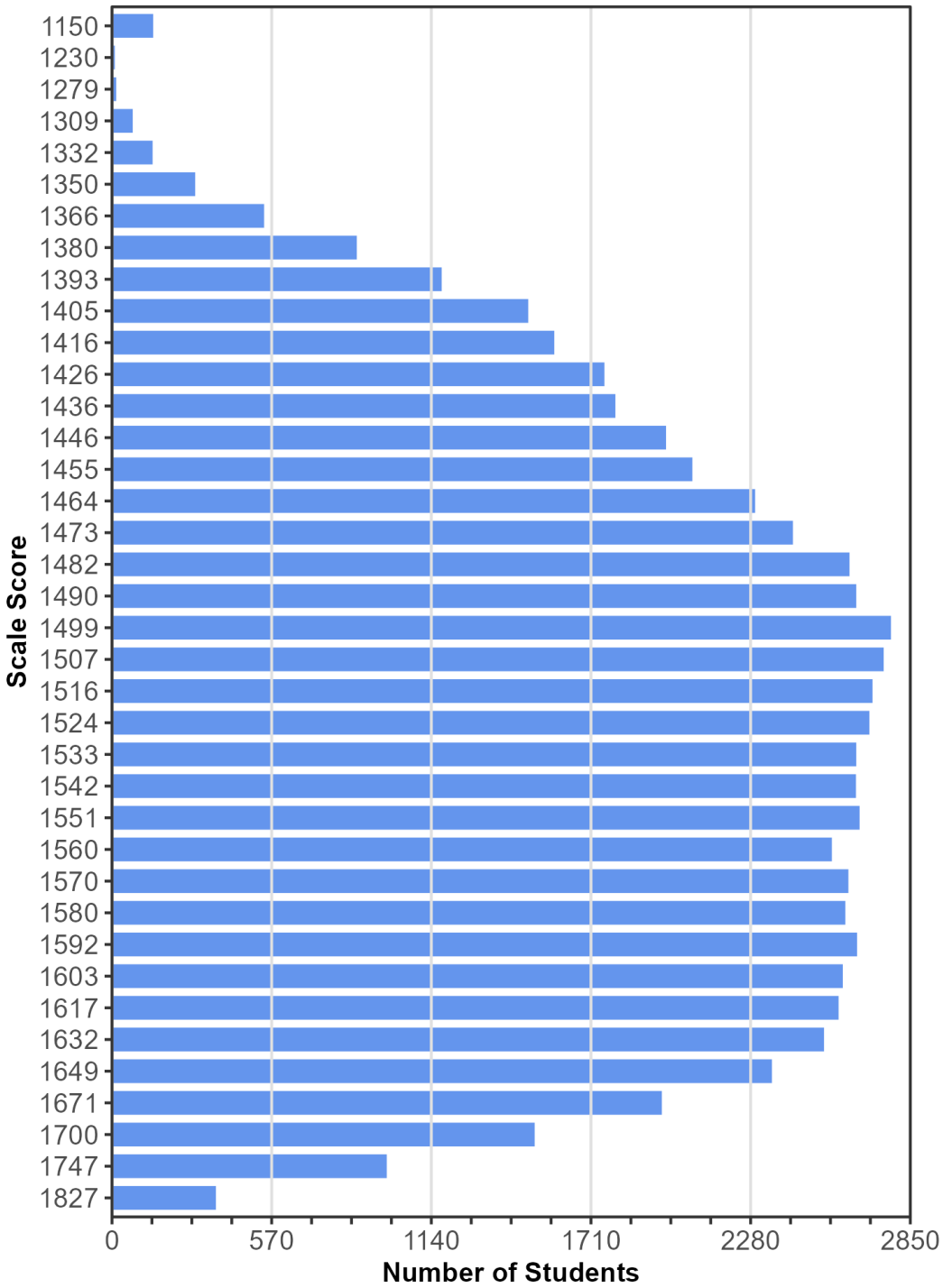
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1162	194	194	0.18	0.18
1247	31	225	0.03	0.21
1298	41	266	0.04	0.24
1330	71	337	0.06	0.31
1353	200	537	0.18	0.49
1372	470	1007	0.43	0.92
1388	1053	2060	0.96	1.88
1402	1573	3633	1.44	3.32
1415	2514	6147	2.30	5.61
1427	3174	9321	2.90	8.51
1438	3709	13030	3.39	11.90
1449	4010	17040	3.66	15.56
1459	4030	21070	3.68	19.25
1468	4200	25270	3.84	23.08
1477	4045	29315	3.69	26.78
1486	4055	33370	3.70	30.48
1495	3932	37302	3.59	34.07
1503	3891	41193	3.55	37.63
1512	3987	45180	3.64	41.27
1520	3839	49019	3.51	44.77
1529	3628	52647	3.31	48.09
1537	3652	56299	3.34	51.42
1545	3517	59816	3.21	54.64
1554	3521	63337	3.22	57.85
1563	3489	66826	3.19	61.04
1572	3370	70196	3.08	64.12
1581	3495	73691	3.19	67.31
1591	3573	77264	3.26	70.57
1602	3661	80925	3.34	73.92
1613	3854	84779	3.52	77.44
1625	3904	88683	3.57	81.00
1639	4052	92735	3.70	84.71
1654	4104	96839	3.75	88.45
1672	3939	100778	3.60	92.05
1695	3625	104403	3.31	95.36
1725	2716	107119	2.48	97.84
1775	1711	108830	1.56	99.41
1860	649	109479	0.59	100.00

Figure D.5.9. 2024 TELPAS Grade 10 Reading
Frequency Distribution of Scale Scores



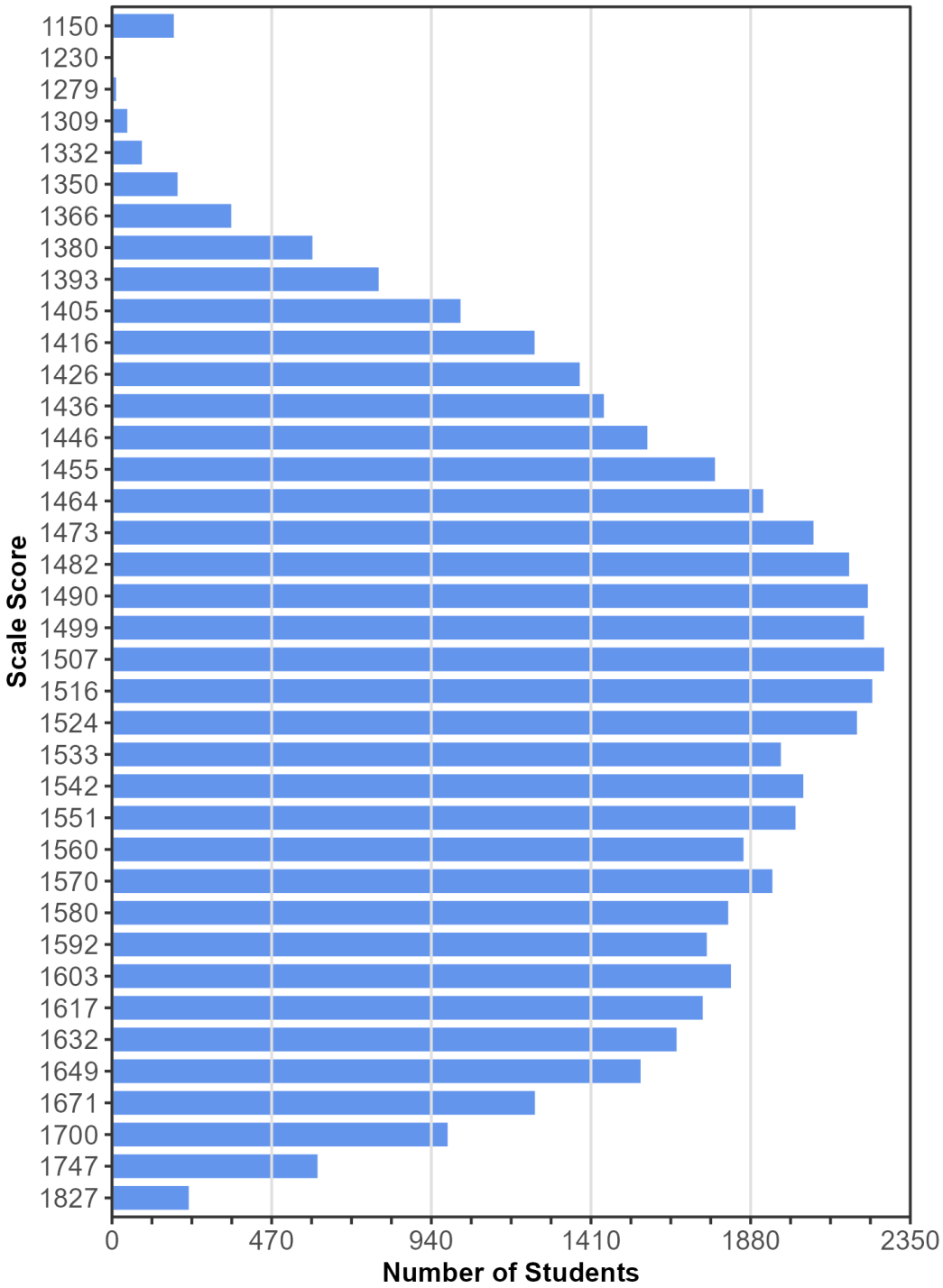
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1150	138	138	0.16	0.16
1230	15	153	0.02	0.18
1279	29	182	0.03	0.21
1309	125	307	0.14	0.36
1332	248	555	0.29	0.64
1350	544	1099	0.63	1.27
1366	907	2006	1.05	2.33
1380	1326	3332	1.54	3.86
1393	1770	5102	2.05	5.92
1405	2125	7227	2.46	8.38
1416	2351	9578	2.73	11.11
1426	2399	11977	2.78	13.89
1436	2551	14528	2.96	16.85
1446	2563	17091	2.97	19.82
1455	2631	19722	3.05	22.87
1464	2906	22628	3.37	26.24
1473	3150	25778	3.65	29.89
1482	3315	29093	3.84	33.74
1490	3359	32452	3.90	37.63
1499	3621	36073	4.20	41.83
1507	3580	39653	4.15	45.99
1516	3478	43131	4.03	50.02
1524	3512	46643	4.07	54.09
1533	3475	50118	4.03	58.12
1542	3415	53533	3.96	62.08
1551	3109	56642	3.61	65.69
1560	3158	59800	3.66	69.35
1570	3152	62952	3.66	73.00
1580	3160	66112	3.66	76.67
1592	3129	69241	3.63	80.30
1603	3049	72290	3.54	83.83
1617	3018	75308	3.50	87.33
1632	2792	78100	3.24	90.57
1649	2574	80674	2.99	93.56
1671	2268	82942	2.63	96.19
1700	1746	84688	2.02	98.21
1747	1089	85777	1.26	99.47
1827	453	86230	0.53	100.00

Figure D.5.10. 2024 TELPAS Grade 11 Reading
Frequency Distribution of Scale Scores



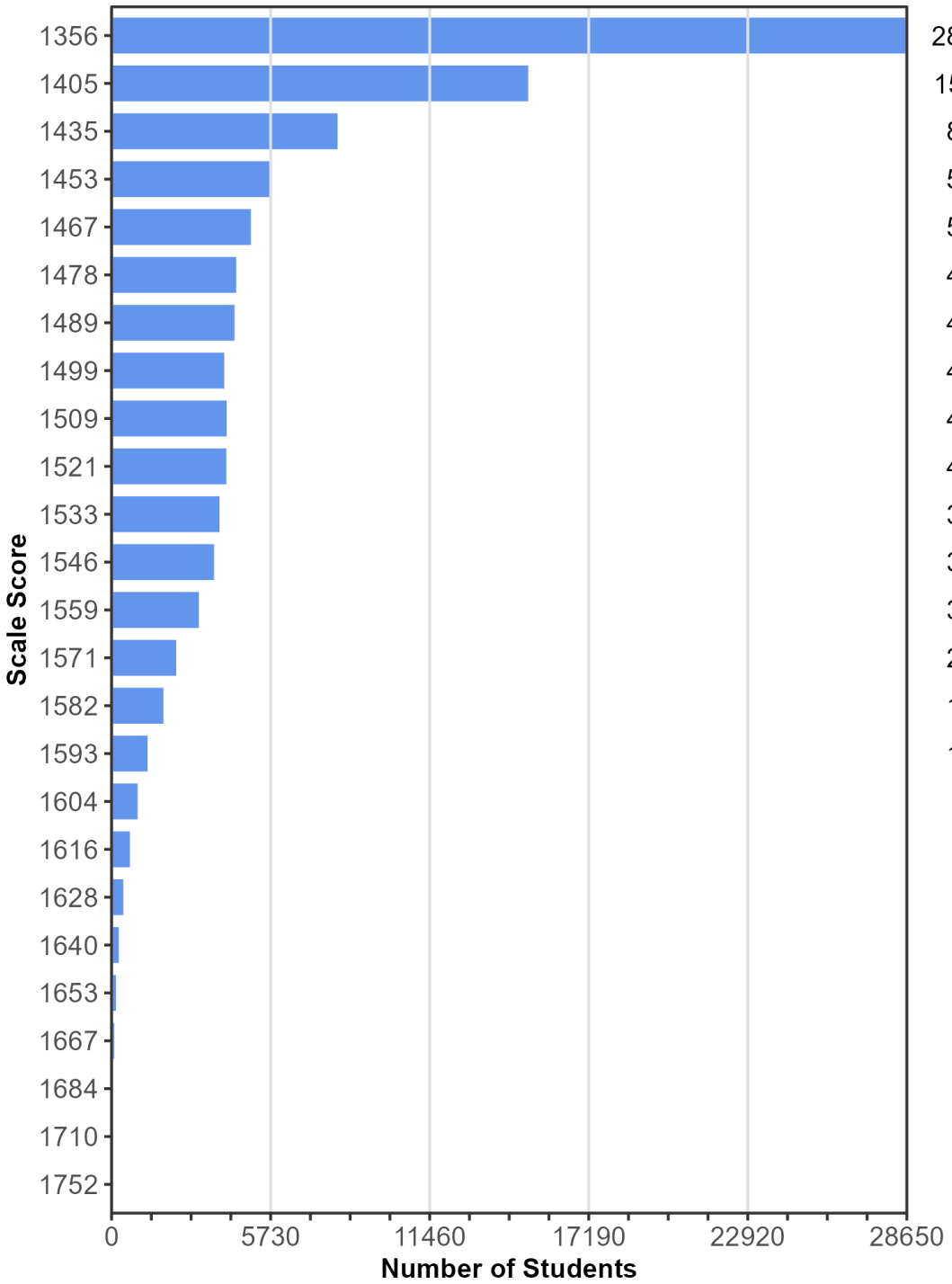
	Freq	Cum Freq	Pct	Cum Pct
1150	147	147	0.22	0.22
1230	10	157	0.01	0.23
1279	15	172	0.02	0.25
1309	74	246	0.11	0.36
1332	145	391	0.21	0.57
1350	297	688	0.43	1.01
1366	543	1231	0.79	1.80
1380	874	2105	1.28	3.08
1393	1177	3282	1.72	4.80
1405	1486	4768	2.18	6.98
1416	1579	6347	2.31	9.29
1426	1758	8105	2.57	11.87
1436	1797	9902	2.63	14.50
1446	1978	11880	2.90	17.39
1455	2072	13952	3.03	20.43
1464	2296	16248	3.36	23.79
1473	2431	18679	3.56	27.35
1482	2633	21312	3.85	31.20
1490	2657	23969	3.89	35.09
1499	2781	26750	4.07	39.16
1507	2755	29505	4.03	43.19
1516	2715	32220	3.97	47.17
1524	2704	34924	3.96	51.13
1533	2657	37581	3.89	55.02
1542	2656	40237	3.89	58.91
1551	2669	42906	3.91	62.81
1560	2570	45476	3.76	66.57
1570	2629	48105	3.85	70.42
1580	2618	50723	3.83	74.26
1592	2660	53383	3.89	78.15
1603	2609	55992	3.82	81.97
1617	2594	58586	3.80	85.77
1632	2542	61128	3.72	89.49
1649	2356	63484	3.45	92.94
1671	1963	65447	2.87	95.81
1700	1509	66956	2.21	98.02
1747	981	67937	1.44	99.46
1827	371	68308	0.54	100.00

Figure D.5.11. 2024 TELPAS Grade 12 Reading
Frequency Distribution of Scale Scores



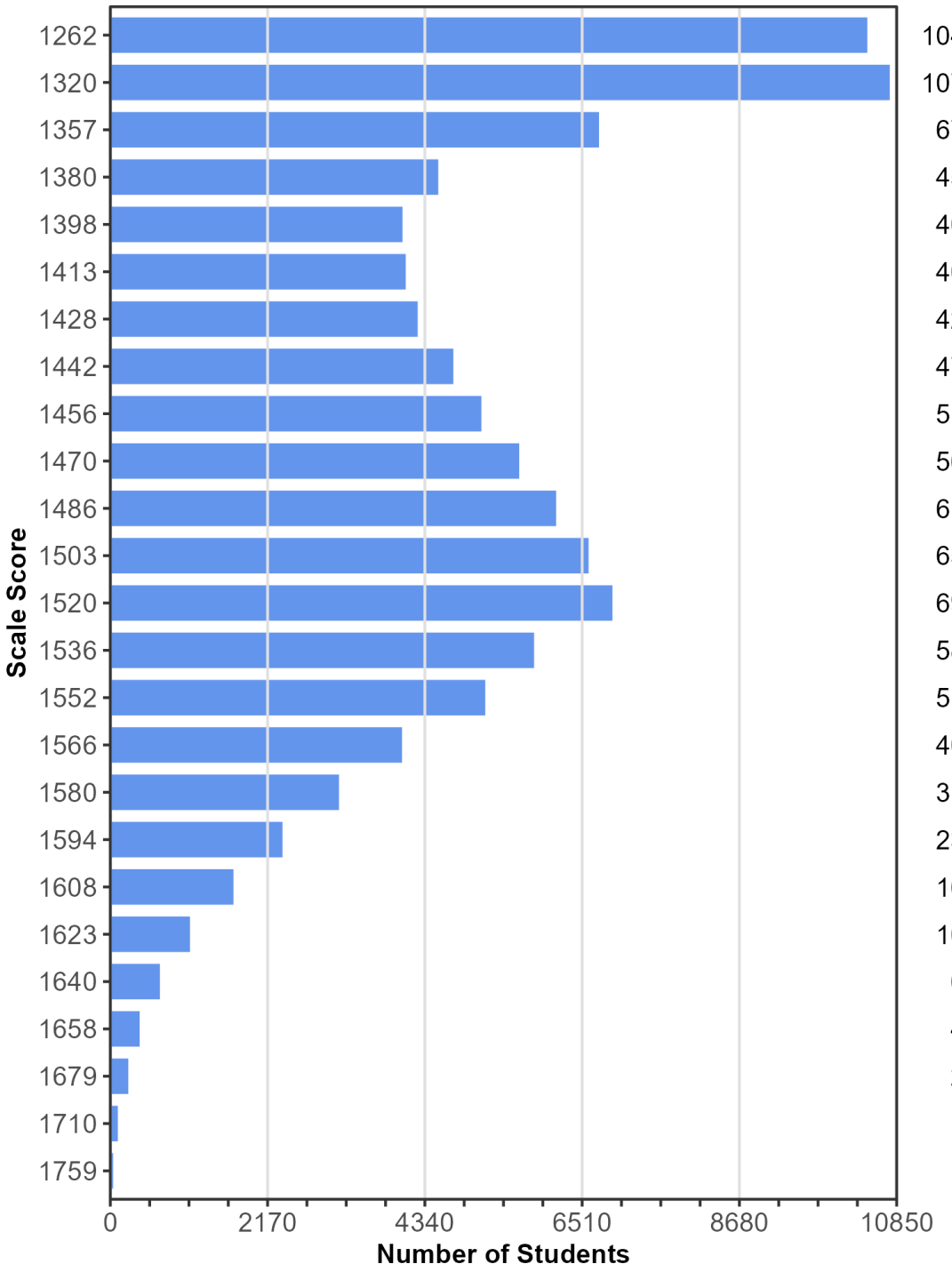
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1150	182	182	0.36	0.36
1230	2	184	0.00	0.36
1279	12	196	0.02	0.38
1309	45	241	0.09	0.47
1332	88	329	0.17	0.64
1350	193	522	0.38	1.02
1366	351	873	0.69	1.70
1380	590	1463	1.15	2.86
1393	785	2248	1.53	4.39
1405	1026	3274	2.00	6.39
1416	1244	4518	2.43	8.82
1426	1377	5895	2.69	11.51
1436	1448	7343	2.83	14.34
1446	1576	8919	3.08	17.41
1455	1775	10694	3.47	20.88
1464	1917	12611	3.74	24.62
1473	2065	14676	4.03	28.66
1482	2170	16846	4.24	32.89
1490	2225	19071	4.34	37.24
1499	2214	21285	4.32	41.56
1507	2273	23558	4.44	46.00
1516	2238	25796	4.37	50.37
1524	2193	27989	4.28	54.65
1533	1969	29958	3.84	58.49
1542	2035	31993	3.97	62.47
1551	2012	34005	3.93	66.40
1560	1859	35864	3.63	70.02
1570	1944	37808	3.80	73.82
1580	1814	39622	3.54	77.36
1592	1751	41373	3.42	80.78
1603	1822	43195	3.56	84.34
1617	1739	44934	3.40	87.73
1632	1662	46596	3.25	90.98
1649	1556	48152	3.04	94.02
1671	1245	49397	2.43	96.45
1700	988	50385	1.93	98.38
1747	605	50990	1.18	99.56
1827	226	51216	0.44	100.00

Figure D.5.12. 2024 TELPAS Grade 2 Writing
Frequency Distribution of Scale Scores



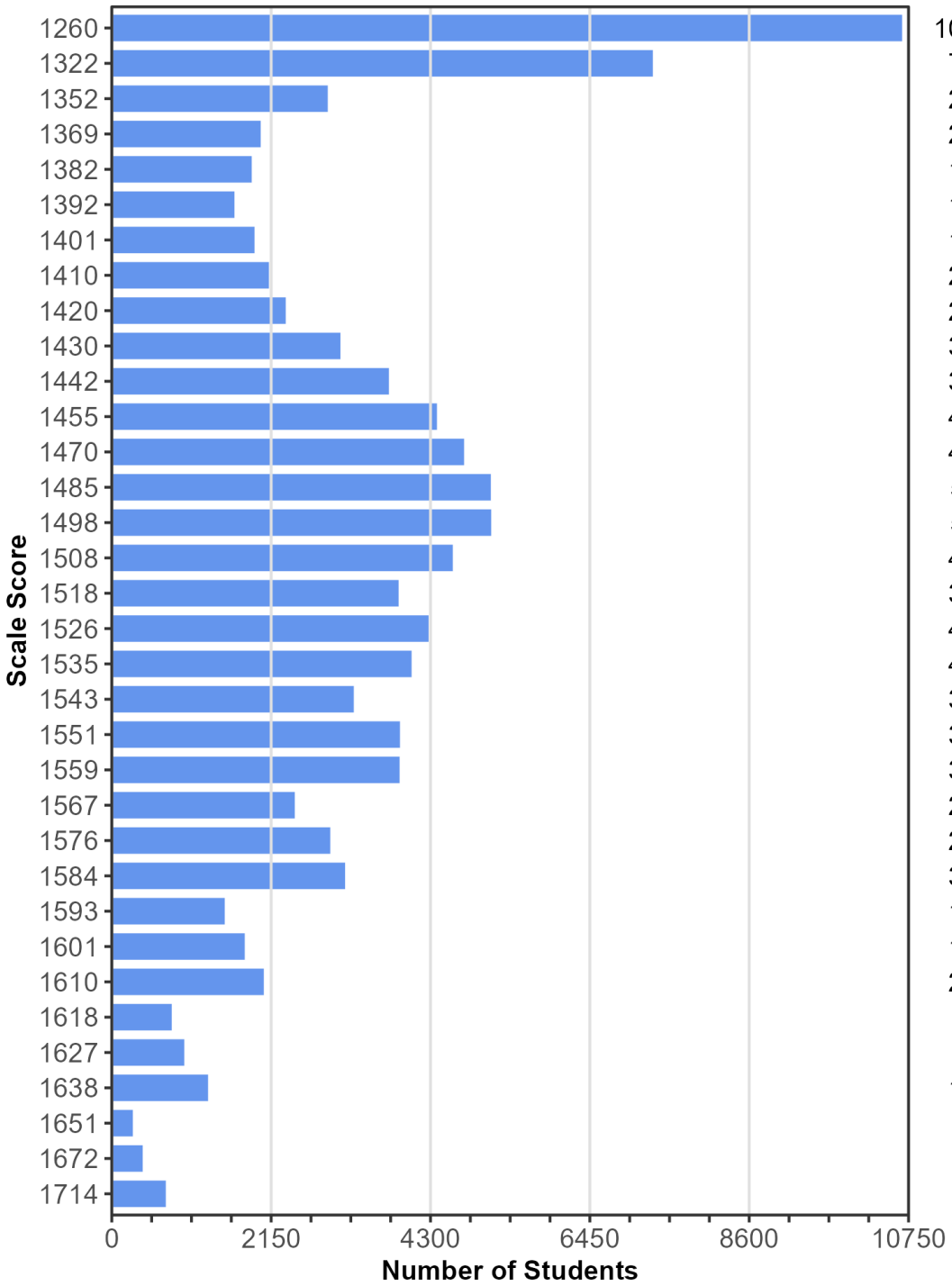
	Freq	Cum Freq	Pct	Cum Pct
1356	28605	28605	27.87	27.87
1405	15011	43616	14.63	42.50
1435	8143	51759	7.93	50.43
1453	5748	57507	5.60	56.03
1467	5020	62527	4.89	60.92
1478	4489	67016	4.37	65.30
1489	4432	71448	4.32	69.62
1499	4059	75507	3.95	73.57
1509	4148	79655	4.04	77.61
1521	4134	83789	4.03	81.64
1533	3889	87678	3.79	85.43
1546	3692	91370	3.60	89.03
1559	3146	94516	3.07	92.09
1571	2329	96845	2.27	94.36
1582	1870	98715	1.82	96.18
1593	1299	100014	1.27	97.45
1604	939	100953	0.91	98.36
1616	662	101615	0.65	99.01
1628	423	102038	0.41	99.42
1640	261	102299	0.25	99.68
1653	160	102459	0.16	99.83
1667	96	102555	0.09	99.92
1684	41	102596	0.04	99.96
1710	23	102619	0.02	99.99
1752	13	102632	0.01	100.00

Figure D.5.13. 2024 TELPAS Grade 3 Writing
Frequency Distribution of Scale Scores



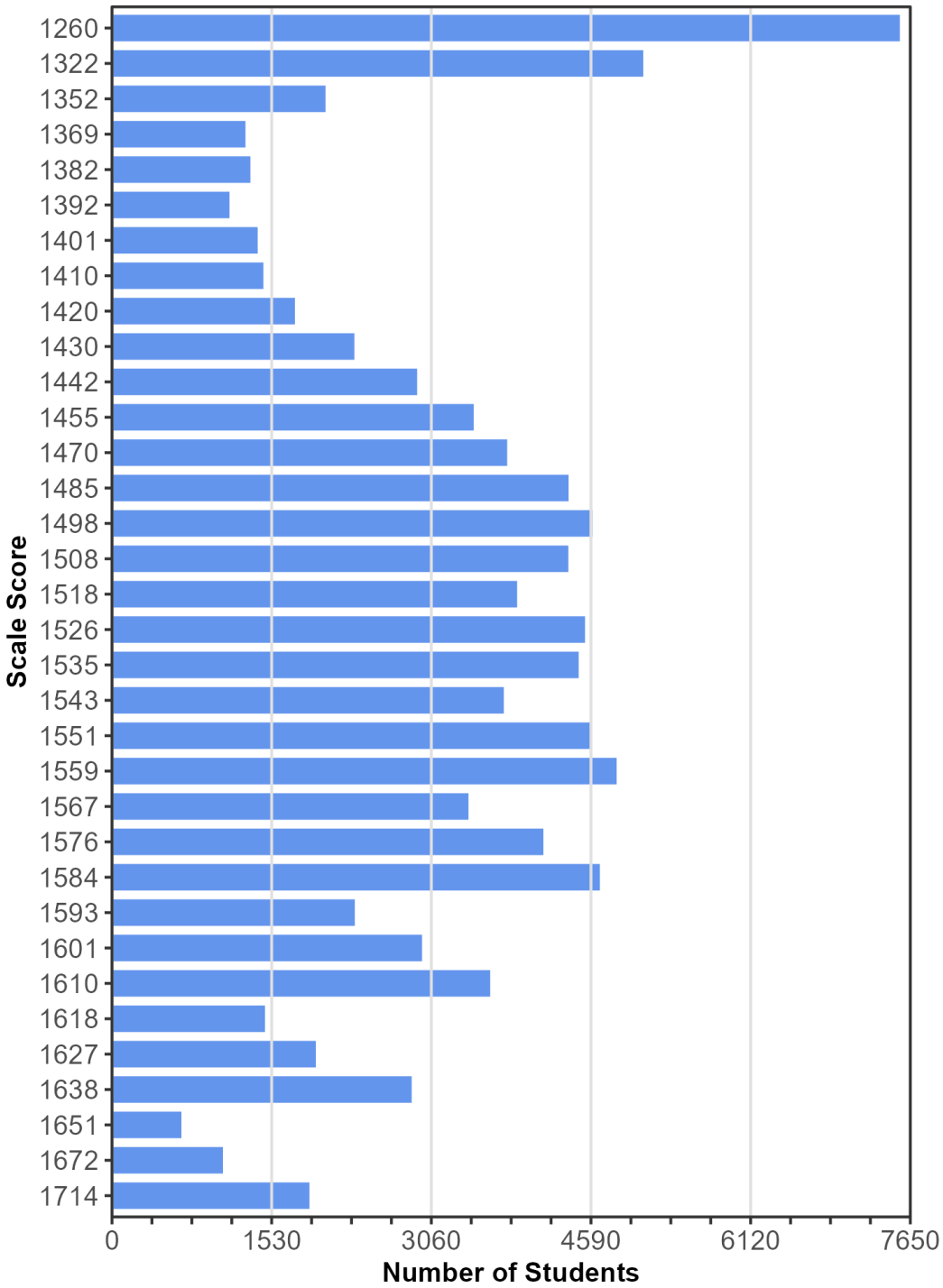
	Freq	Cum Freq	Pct	Cum Pct
1262	10445	10445	9.96	9.96
1320	10755	21200	10.26	20.22
1357	6743	27943	6.43	26.65
1380	4525	32468	4.32	30.97
1398	4031	36499	3.84	34.81
1413	4076	40575	3.89	38.70
1428	4241	44816	4.05	42.75
1442	4733	49549	4.51	47.26
1456	5120	54669	4.88	52.14
1470	5642	60311	5.38	57.53
1486	6151	66462	5.87	63.39
1503	6599	73061	6.29	69.69
1520	6928	79989	6.61	76.29
1536	5847	85836	5.58	81.87
1552	5173	91009	4.93	86.81
1566	4025	95034	3.84	90.64
1580	3155	98189	3.01	93.65
1594	2376	100565	2.27	95.92
1608	1699	102264	1.62	97.54
1623	1099	103363	1.05	98.59
1640	684	104047	0.65	99.24
1658	405	104452	0.39	99.63
1679	248	104700	0.24	99.86
1710	104	104804	0.10	99.96
1759	38	104842	0.04	100.00

Figure D.5.14. 2024 TELPAS Grade 4 Writing
Frequency Distribution of Scale Scores



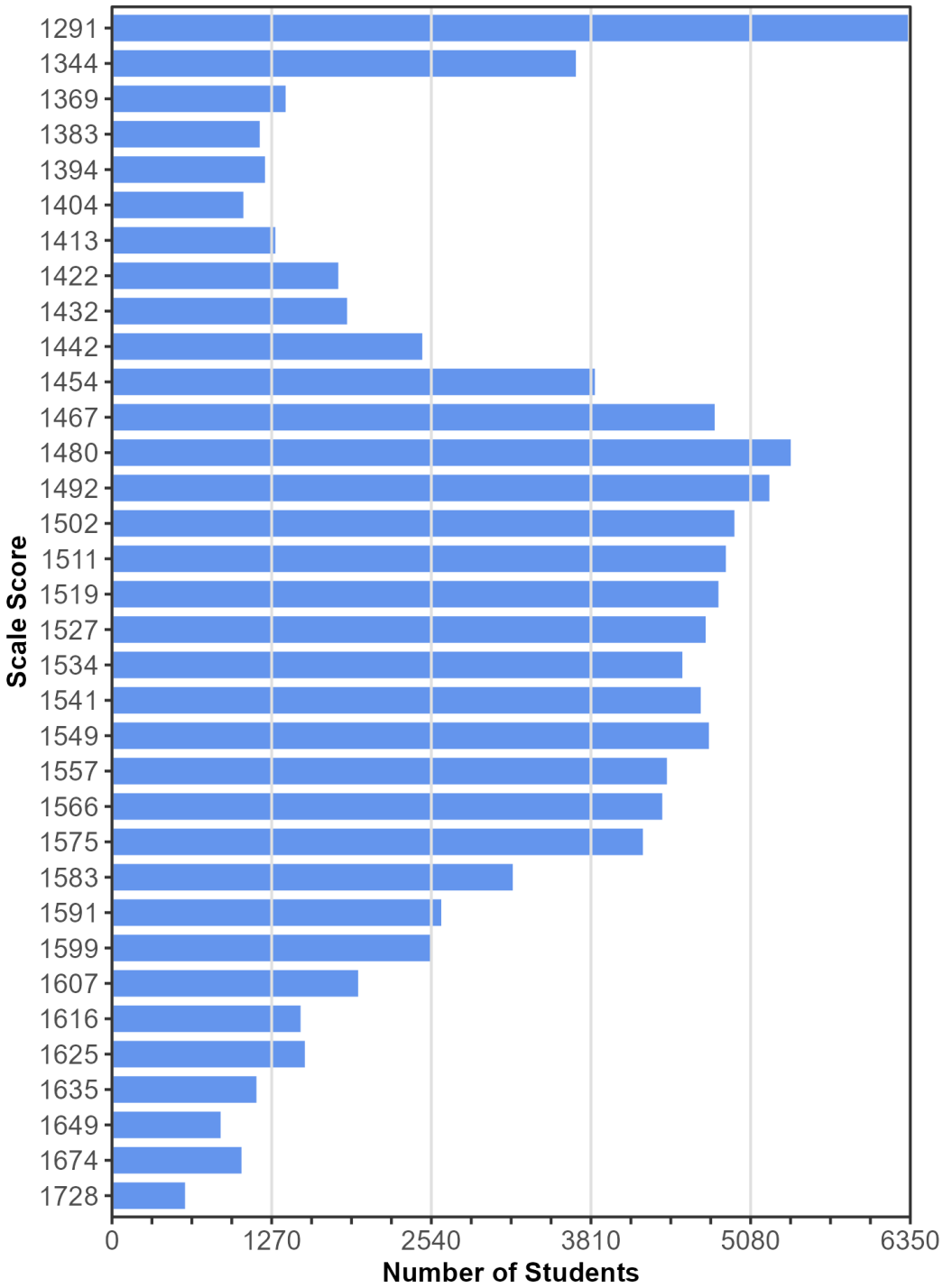
	Freq	Cum Freq	Pct	Cum Pct
1260	10663	10663	10.13	10.13
1322	7300	17963	6.93	17.06
1352	2914	20877	2.77	19.83
1369	2009	22886	1.91	21.74
1382	1888	24774	1.79	23.53
1392	1655	26429	1.57	25.10
1401	1926	28355	1.83	26.93
1410	2120	30475	2.01	28.94
1420	2347	32822	2.23	31.17
1430	3085	35907	2.93	34.10
1442	3739	39646	3.55	37.65
1455	4387	44033	4.17	41.82
1470	4753	48786	4.51	46.33
1485	5114	53900	4.86	51.19
1498	5119	59019	4.86	56.05
1508	4601	63620	4.37	60.42
1518	3870	67490	3.68	64.10
1526	4274	71764	4.06	68.16
1535	4045	75809	3.84	72.00
1543	3265	79074	3.10	75.10
1551	3888	82962	3.69	78.79
1559	3883	86845	3.69	82.48
1567	2469	89314	2.34	84.82
1576	2948	92262	2.80	87.62
1584	3148	95410	2.99	90.61
1593	1524	96934	1.45	92.06
1601	1794	98728	1.70	93.76
1610	2051	100779	1.95	95.71
1618	809	101588	0.77	96.48
1627	978	102566	0.93	97.41
1638	1299	103865	1.23	98.64
1651	284	104149	0.27	98.91
1672	417	104566	0.40	99.31
1714	729	105295	0.69	100.00

Figure D.5.15. 2024 TELPAS Grade 5 Writing
Frequency Distribution of Scale Scores



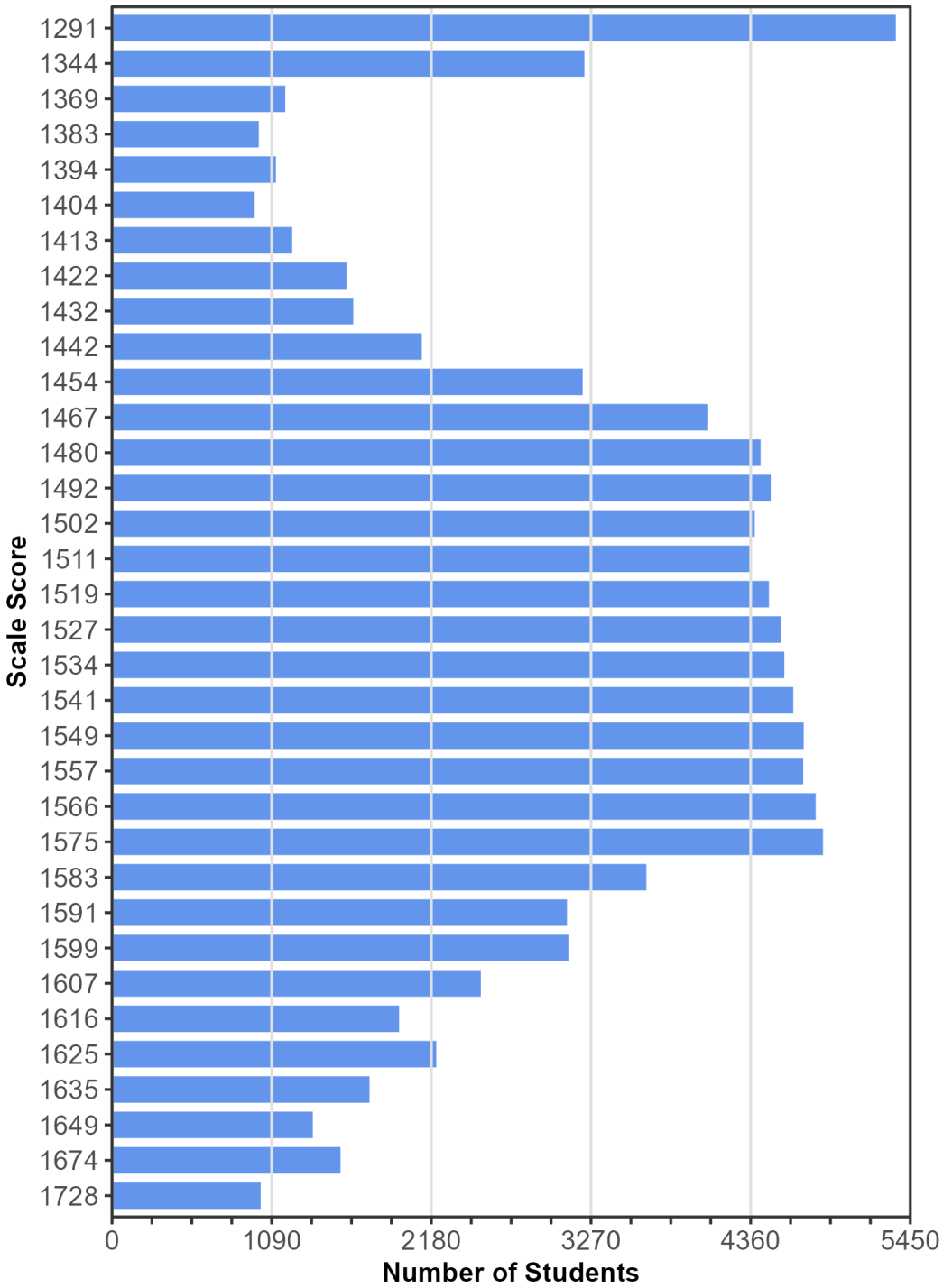
	Freq	Cum Freq	Pct	Cum Pct
1260	7550	7550	7.12	7.12
1322	5091	12641	4.80	11.93
1352	2046	14687	1.93	13.86
1369	1279	15966	1.21	15.06
1382	1326	17292	1.25	16.31
1392	1126	18418	1.06	17.38
1401	1396	19814	1.32	18.69
1410	1451	21265	1.37	20.06
1420	1753	23018	1.65	21.72
1430	2323	25341	2.19	23.91
1442	2924	28265	2.76	26.67
1455	3466	31731	3.27	29.94
1470	3787	35518	3.57	33.51
1485	4375	39893	4.13	37.64
1498	4605	44498	4.34	41.98
1508	4373	48871	4.13	46.11
1518	3882	52753	3.66	49.77
1526	4533	57286	4.28	54.05
1535	4472	61758	4.22	58.27
1543	3755	65513	3.54	61.81
1551	4586	70099	4.33	66.14
1559	4836	74935	4.56	70.70
1567	3416	78351	3.22	73.92
1576	4134	82485	3.90	77.82
1584	4675	87160	4.41	82.23
1593	2327	89487	2.20	84.43
1601	2971	92458	2.80	87.23
1610	3624	96082	3.42	90.65
1618	1466	97548	1.38	92.03
1627	1953	99501	1.84	93.88
1638	2872	102373	2.71	96.58
1651	665	103038	0.63	97.21
1672	1063	104101	1.00	98.21
1714	1892	105993	1.79	100.00

Figure D.5.16. 2024 TELPAS Grade 6 Writing
Frequency Distribution of Scale Scores



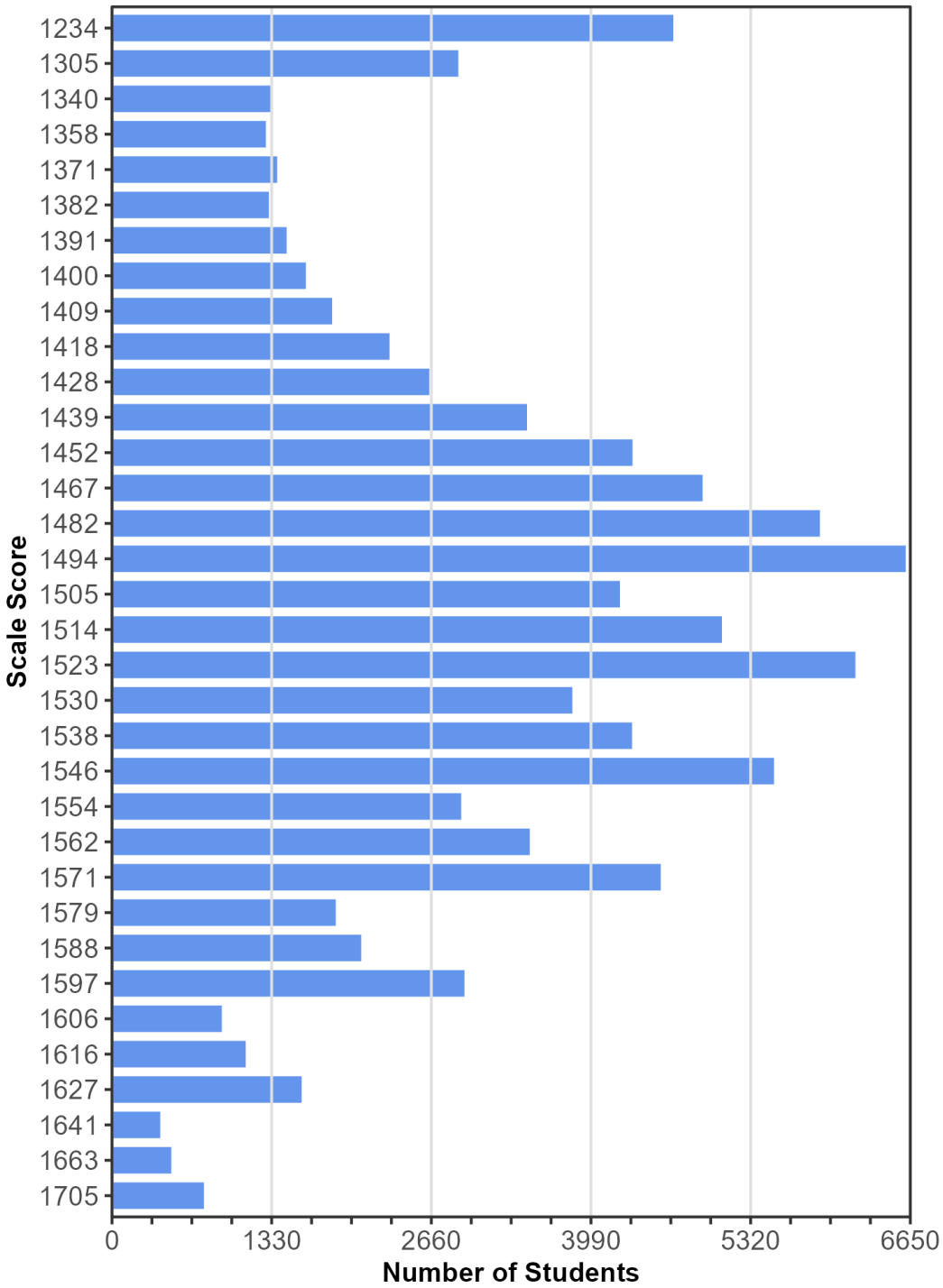
	Freq	Cum Freq	Pct	Cum Pct
1291	6331	6331	6.04	6.04
1344	3690	10021	3.52	9.56
1369	1381	11402	1.32	10.87
1383	1175	12577	1.12	11.99
1394	1217	13794	1.16	13.15
1404	1045	14839	1.00	14.15
1413	1299	16138	1.24	15.39
1422	1800	17938	1.72	17.11
1432	1870	19808	1.78	18.89
1442	2468	22276	2.35	21.24
1454	3841	26117	3.66	24.91
1467	4794	30911	4.57	29.48
1480	5399	36310	5.15	34.63
1492	5229	41539	4.99	39.61
1502	4951	46490	4.72	44.34
1511	4883	51373	4.66	48.99
1519	4824	56197	4.60	53.59
1527	4722	60919	4.50	58.10
1534	4537	65456	4.33	62.42
1541	4683	70139	4.47	66.89
1549	4748	74887	4.53	71.42
1557	4414	79301	4.21	75.63
1566	4377	83678	4.17	79.80
1575	4223	87901	4.03	83.83
1583	3188	91089	3.04	86.87
1591	2619	93708	2.50	89.36
1599	2536	96244	2.42	91.78
1607	1958	98202	1.87	93.65
1616	1500	99702	1.43	95.08
1625	1534	101236	1.46	96.54
1635	1149	102385	1.10	97.64
1649	864	103249	0.82	98.46
1674	1030	104279	0.98	99.45
1728	581	104860	0.55	100.00

Figure D.5.17. 2024 TELPAS Grade 7 Writing
Frequency Distribution of Scale Scores



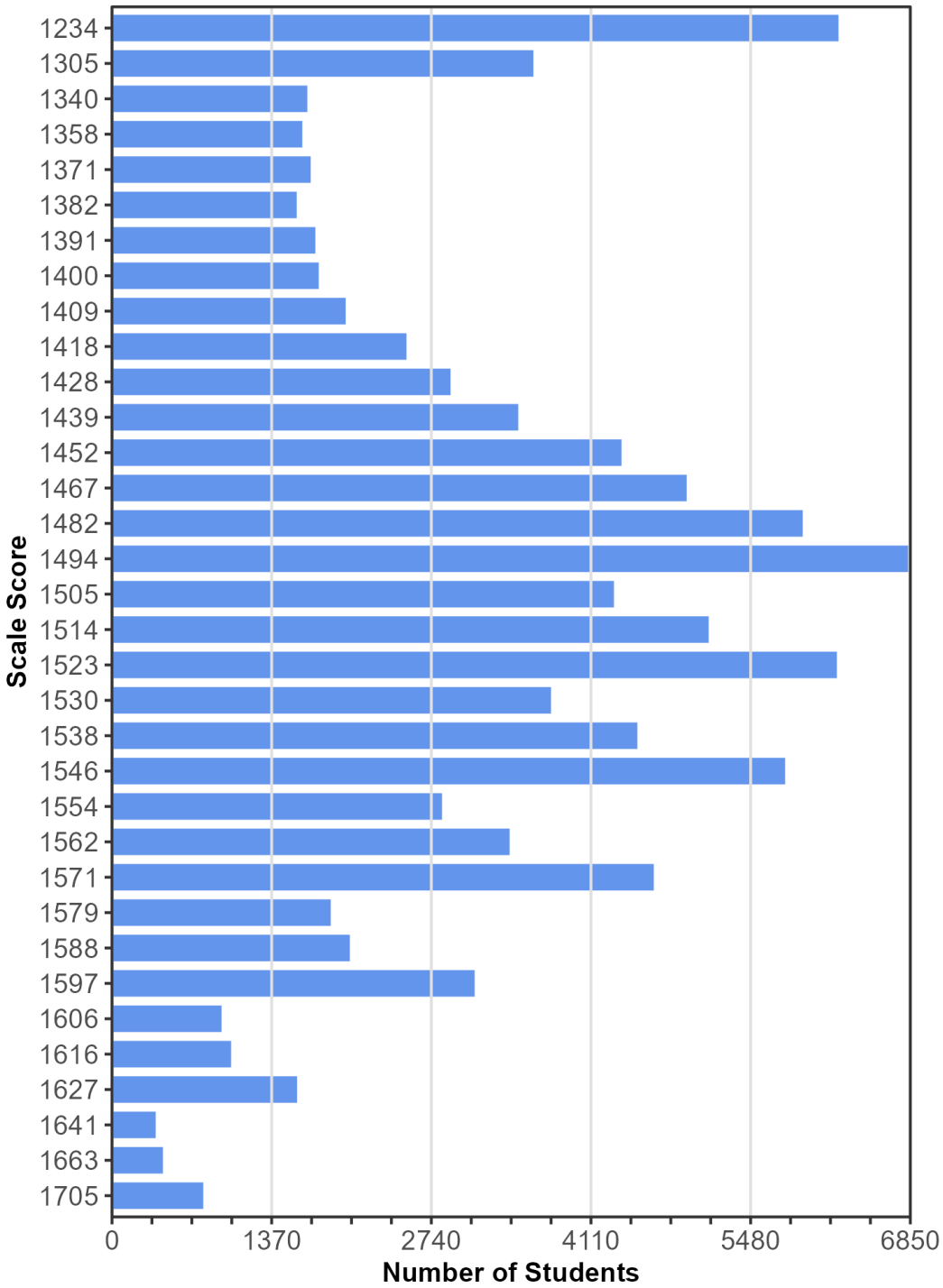
	Freq	Cum Freq	Pct	Cum Pct
1291	5351	5351	5.14	5.14
1344	3225	8576	3.10	8.24
1369	1183	9759	1.14	9.38
1383	1002	10761	0.96	10.34
1394	1119	11880	1.08	11.42
1404	973	12853	0.94	12.35
1413	1230	14083	1.18	13.53
1422	1602	15685	1.54	15.07
1432	1647	17332	1.58	16.66
1442	2115	19447	2.03	18.69
1454	3213	22660	3.09	21.78
1467	4070	26730	3.91	25.69
1480	4428	31158	4.26	29.94
1492	4497	35655	4.32	34.26
1502	4388	40043	4.22	38.48
1511	4362	44405	4.19	42.67
1519	4485	48890	4.31	46.98
1527	4567	53457	4.39	51.37
1534	4590	58047	4.41	55.78
1541	4651	62698	4.47	60.25
1549	4722	67420	4.54	64.79
1557	4719	72139	4.53	69.32
1566	4804	76943	4.62	73.94
1575	4854	81797	4.66	78.60
1583	3648	85445	3.51	82.11
1591	3106	88551	2.98	85.09
1599	3116	91667	2.99	88.09
1607	2518	94185	2.42	90.51
1616	1960	96145	1.88	92.39
1625	2214	98359	2.13	94.52
1635	1758	100117	1.69	96.21
1649	1371	101488	1.32	97.53
1674	1560	103048	1.50	99.02
1728	1015	104063	0.98	100.00

Figure D.5.18. 2024 TELPAS Grade 8 Writing
Frequency Distribution of Scale Scores



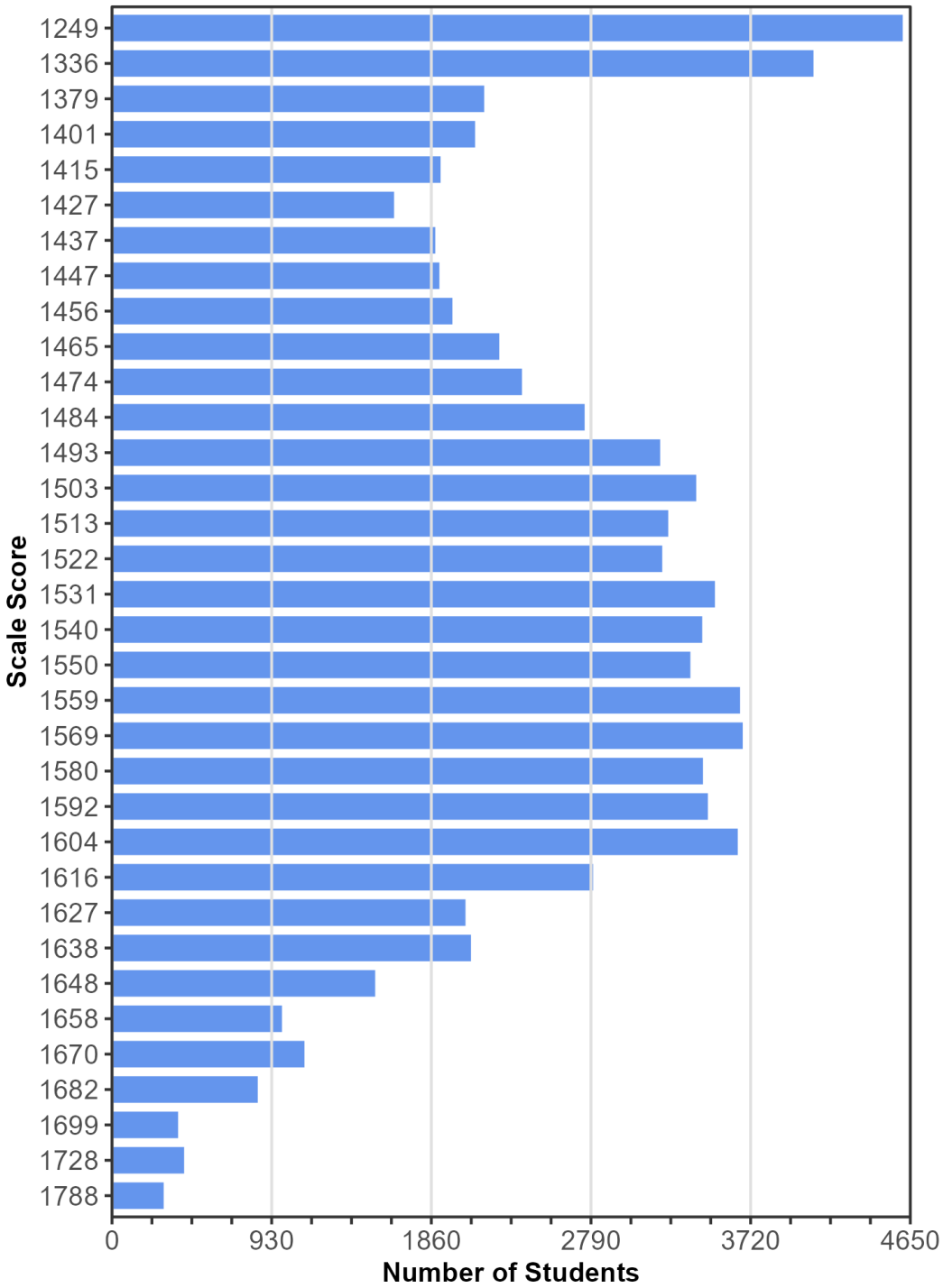
	Freq	Cum Freq	Pct	Cum Pct
1234	4676	4676	4.67	4.67
1305	2885	7561	2.88	7.54
1340	1324	8885	1.32	8.87
1358	1282	10167	1.28	10.14
1371	1376	11543	1.37	11.52
1382	1307	12850	1.30	12.82
1391	1455	14305	1.45	14.27
1400	1615	15920	1.61	15.88
1409	1834	17754	1.83	17.71
1418	2312	20066	2.31	20.02
1428	2643	22709	2.64	22.66
1439	3457	26166	3.45	26.11
1452	4336	30502	4.33	30.43
1467	4920	35422	4.91	35.34
1482	5897	41319	5.88	41.23
1494	6612	47931	6.60	47.82
1505	4231	52162	4.22	52.04
1514	5081	57243	5.07	57.11
1523	6193	63436	6.18	63.29
1530	3835	67271	3.83	67.12
1538	4332	71603	4.32	71.44
1546	5515	77118	5.50	76.94
1554	2909	80027	2.90	79.85
1562	3480	83507	3.47	83.32
1571	4571	88078	4.56	87.88
1579	1864	89942	1.86	89.74
1588	2076	92018	2.07	91.81
1597	2936	94954	2.93	94.74
1606	915	95869	0.91	95.65
1616	1114	96983	1.11	96.77
1627	1580	98563	1.58	98.34
1641	402	98965	0.40	98.74
1663	494	99459	0.49	99.24
1705	766	100225	0.76	100.00

Figure D.5.19. 2024 TELPAS Grade 9 Writing
Frequency Distribution of Scale Scores



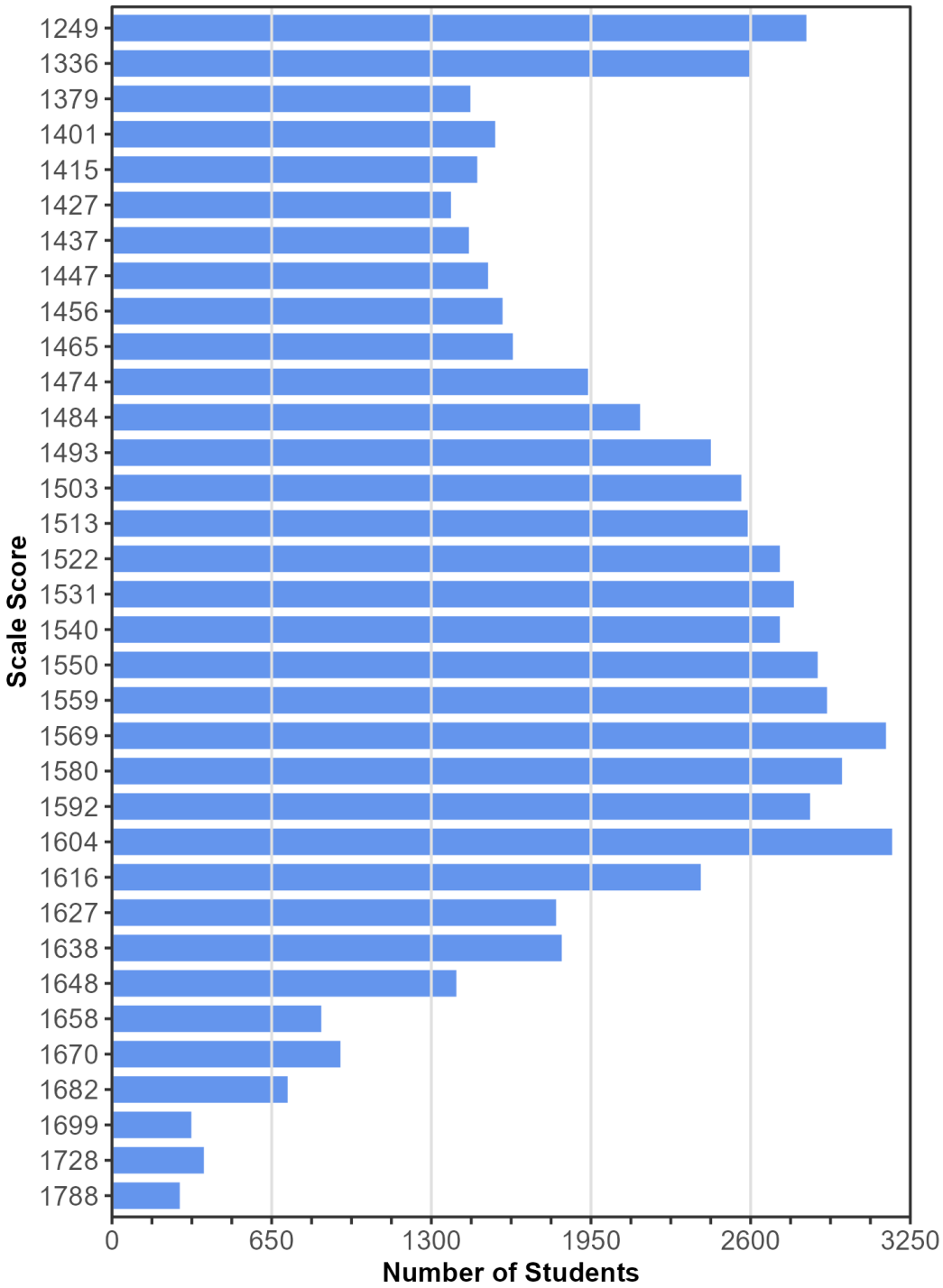
	Freq	Cum Freq	Pct	Cum Pct
1234	6234	6234	5.90	5.90
1305	3616	9850	3.42	9.32
1340	1676	11526	1.59	10.90
1358	1633	13159	1.54	12.45
1371	1705	14864	1.61	14.06
1382	1585	16449	1.50	15.56
1391	1745	18194	1.65	17.21
1400	1774	19968	1.68	18.89
1409	2005	21973	1.90	20.78
1418	2527	24500	2.39	23.17
1428	2905	27405	2.75	25.92
1439	3486	30891	3.30	29.22
1452	4372	35263	4.14	33.36
1467	4931	40194	4.66	38.02
1482	5927	46121	5.61	43.63
1494	6831	52952	6.46	50.09
1505	4307	57259	4.07	54.16
1514	5120	62379	4.84	59.01
1523	6220	68599	5.88	64.89
1530	3767	72366	3.56	68.45
1538	4508	76874	4.26	72.72
1546	5776	82650	5.46	78.18
1554	2832	85482	2.68	80.86
1562	3412	88894	3.23	84.09
1571	4649	93543	4.40	88.48
1579	1877	95420	1.78	90.26
1588	2041	97461	1.93	92.19
1597	3112	100573	2.94	95.13
1606	940	101513	0.89	96.02
1616	1022	102535	0.97	96.99
1627	1588	104123	1.50	98.49
1641	375	104498	0.35	98.85
1663	437	104935	0.41	99.26
1705	783	105718	0.74	100.00

Figure D.5.20. 2024 TELPAS Grade 10 Writing
Frequency Distribution of Scale Scores



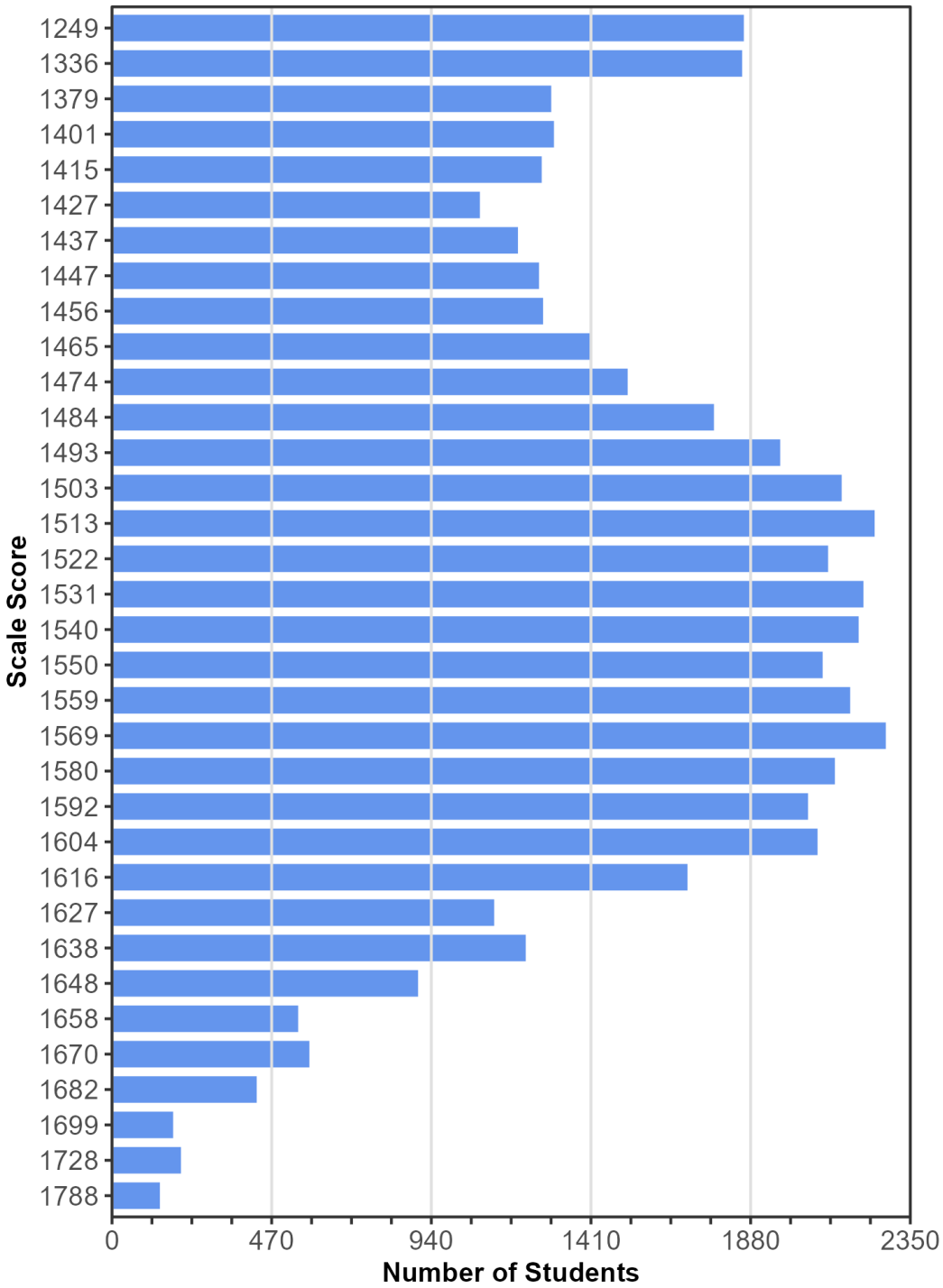
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1249	4606	4606	5.52	5.52
1336	4086	8692	4.89	10.41
1379	2168	10860	2.60	13.01
1401	2115	12975	2.53	15.54
1415	1914	14889	2.29	17.83
1427	1643	16532	1.97	19.80
1437	1883	18415	2.26	22.05
1447	1907	20322	2.28	24.34
1456	1983	22305	2.37	26.71
1465	2256	24561	2.70	29.41
1474	2388	26949	2.86	32.27
1484	2753	29702	3.30	35.57
1493	3193	32895	3.82	39.39
1503	3403	36298	4.08	43.47
1513	3240	39538	3.88	47.35
1522	3205	42743	3.84	51.19
1531	3512	46255	4.21	55.39
1540	3438	49693	4.12	59.51
1550	3369	53062	4.03	63.55
1559	3658	56720	4.38	67.93
1569	3674	60394	4.40	72.33
1580	3442	63836	4.12	76.45
1592	3471	67307	4.16	80.60
1604	3645	70952	4.37	84.97
1616	2802	73754	3.36	88.32
1627	2059	75813	2.47	90.79
1638	2091	77904	2.50	93.29
1648	1533	79437	1.84	95.13
1658	990	80427	1.19	96.32
1670	1121	81548	1.34	97.66
1682	849	82397	1.02	98.68
1699	385	82782	0.46	99.14
1728	420	83202	0.50	99.64
1788	301	83503	0.36	100.00

Figure D.5.21. 2024 TELPAS Grade 11 Writing
Frequency Distribution of Scale Scores



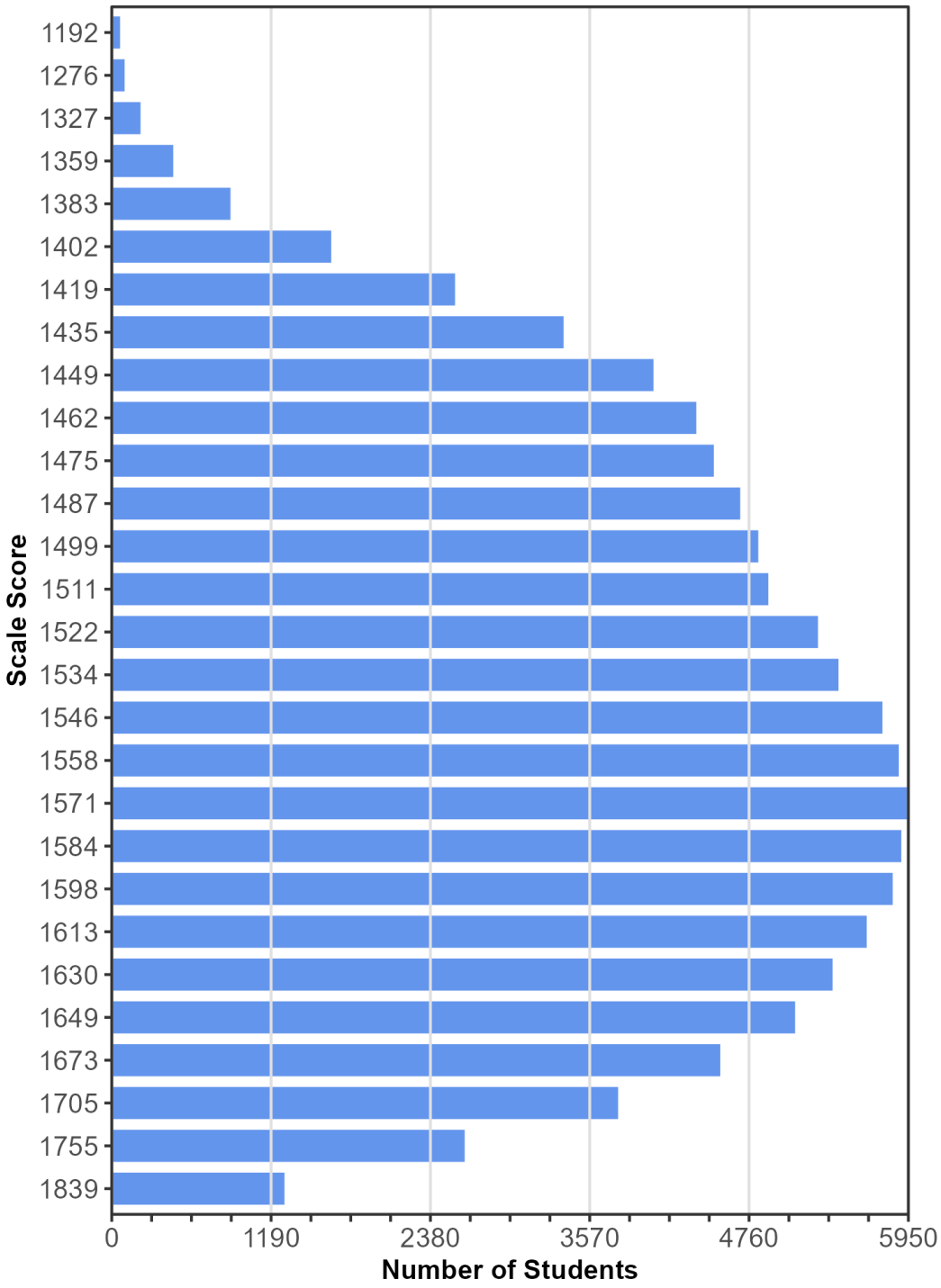
	Freq	Cum Freq	Pct	Cum Pct
1249	2827	2827	4.27	4.27
1336	2602	5429	3.93	8.20
1379	1459	6888	2.20	10.40
1401	1560	8448	2.35	12.75
1415	1487	9935	2.24	15.00
1427	1380	11315	2.08	17.08
1437	1453	12768	2.19	19.27
1447	1531	14299	2.31	21.59
1456	1590	15889	2.40	23.99
1465	1632	17521	2.46	26.45
1474	1938	19459	2.93	29.37
1484	2150	21609	3.25	32.62
1493	2438	24047	3.68	36.30
1503	2562	26609	3.87	40.17
1513	2588	29197	3.91	44.07
1522	2719	31916	4.10	48.18
1531	2776	34692	4.19	52.37
1540	2719	37411	4.10	56.47
1550	2873	40284	4.34	60.81
1559	2911	43195	4.39	65.21
1569	3151	46346	4.76	69.96
1580	2972	49318	4.49	74.45
1592	2842	52160	4.29	78.74
1604	3176	55336	4.79	83.53
1616	2397	57733	3.62	87.15
1627	1808	59541	2.73	89.88
1638	1831	61372	2.76	92.65
1648	1402	62774	2.12	94.76
1658	852	63626	1.29	96.05
1670	930	64556	1.40	97.45
1682	715	65271	1.08	98.53
1699	323	65594	0.49	99.02
1728	374	65968	0.56	99.58
1788	276	66244	0.42	100.00

Figure D.5.22. 2024 TELPAS Grade 12 Writing
Frequency Distribution of Scale Scores



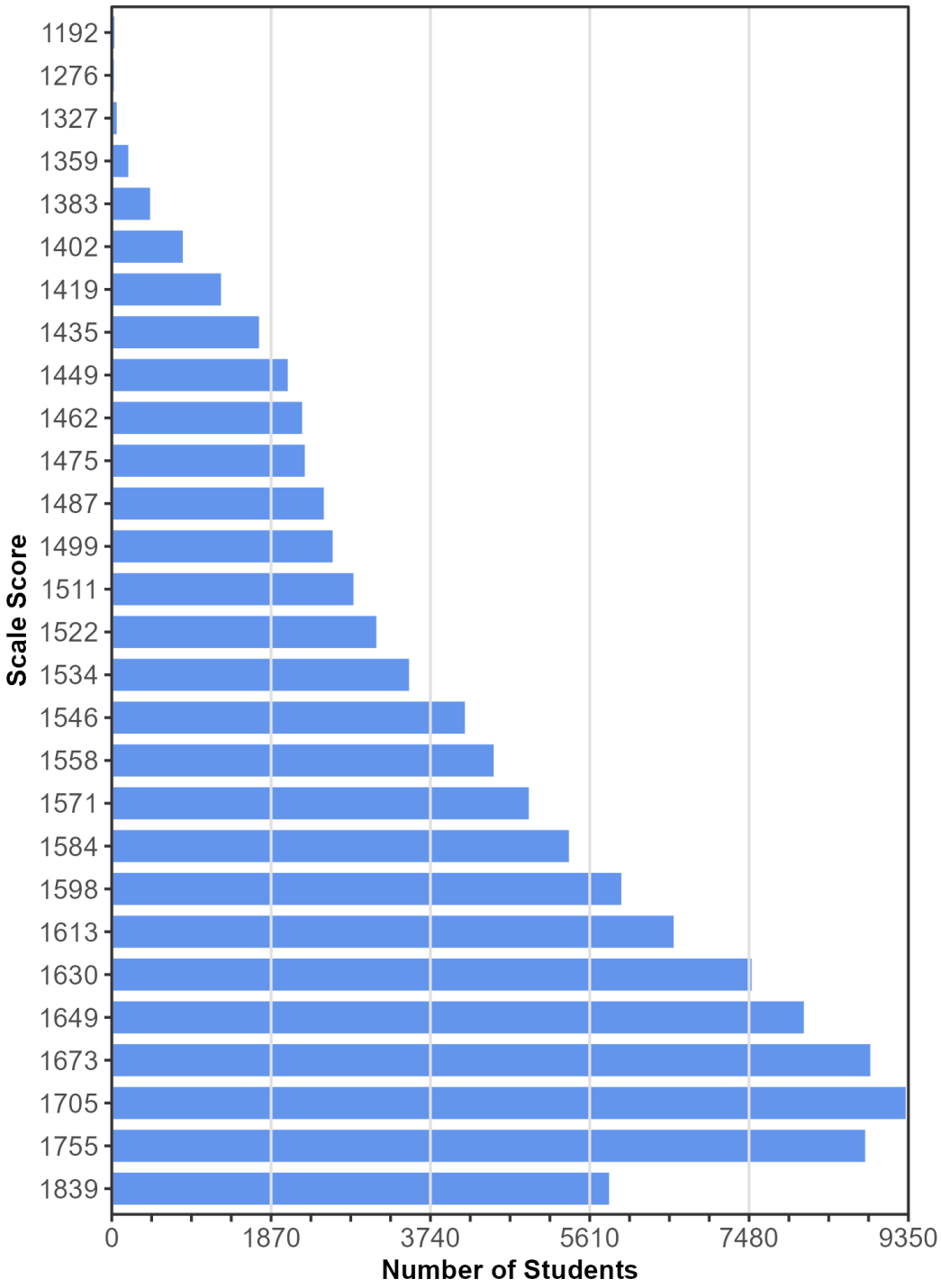
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1249	1860	1860	3.74	3.74
1336	1855	3715	3.73	7.46
1379	1293	5008	2.60	10.06
1401	1301	6309	2.61	12.68
1415	1265	7574	2.54	15.22
1427	1083	8657	2.18	17.39
1437	1195	9852	2.40	19.79
1447	1257	11109	2.53	22.32
1456	1269	12378	2.55	24.87
1465	1411	13789	2.83	27.70
1474	1518	15307	3.05	30.75
1484	1772	17079	3.56	34.32
1493	1967	19046	3.95	38.27
1503	2148	21194	4.32	42.58
1513	2245	23439	4.51	47.09
1522	2108	25547	4.24	51.33
1531	2212	27759	4.44	55.77
1540	2198	29957	4.42	60.19
1550	2092	32049	4.20	64.39
1559	2173	34222	4.37	68.76
1569	2278	36500	4.58	73.34
1580	2128	38628	4.28	77.61
1592	2049	40677	4.12	81.73
1604	2077	42754	4.17	85.90
1616	1694	44448	3.40	89.31
1627	1125	45573	2.26	91.57
1638	1218	46791	2.45	94.01
1648	901	47692	1.81	95.82
1658	548	48240	1.10	96.92
1670	581	48821	1.17	98.09
1682	426	49247	0.86	98.95
1699	180	49427	0.36	99.31
1728	203	49630	0.41	99.72
1788	141	49771	0.28	100.00

Figure D.5.23. 2024 TELPAS Grade 2 Listening
Frequency Distribution of Scale Scores



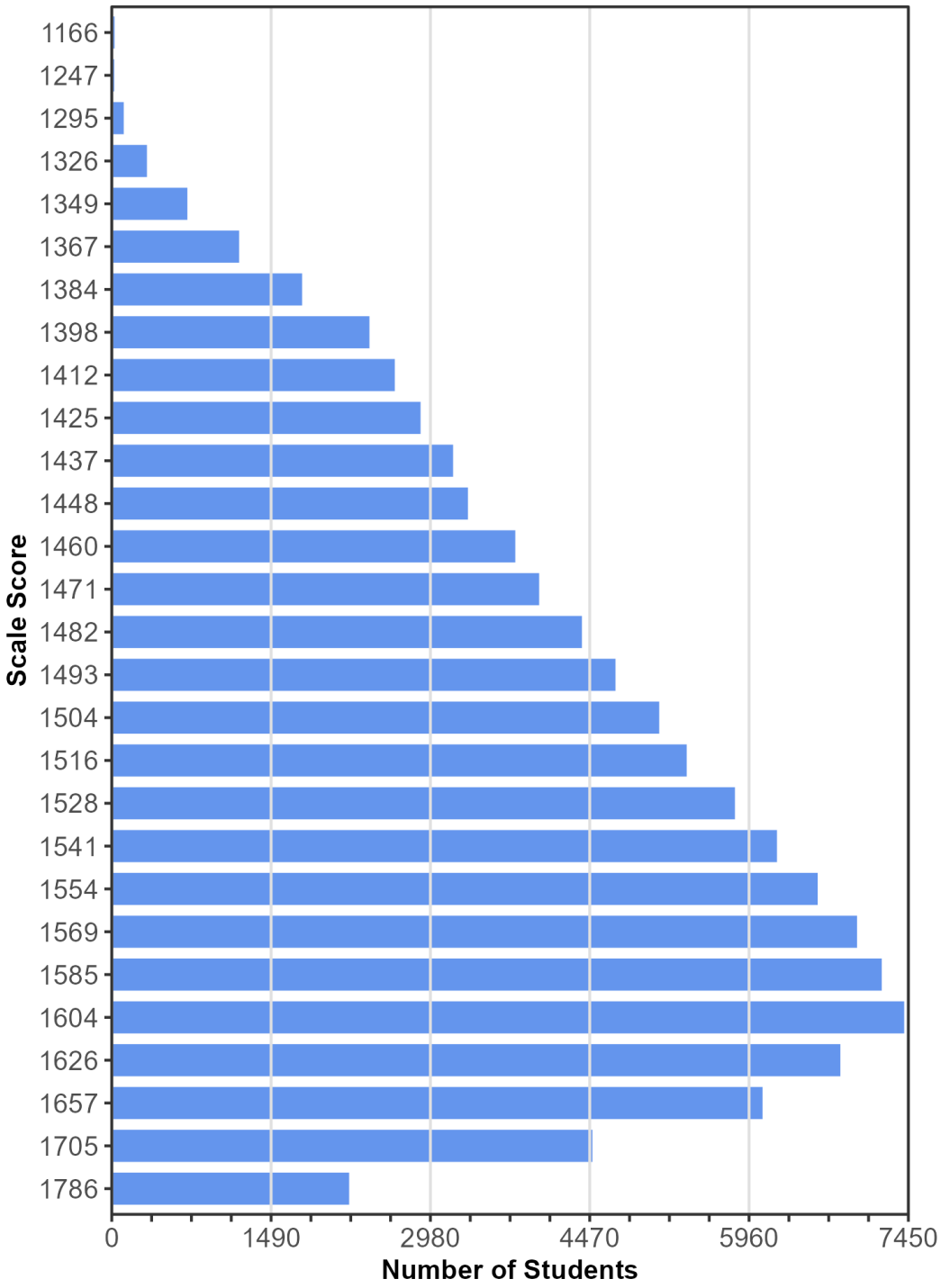
	Freq	Cum Freq	Pct	Cum Pct
1192	62	62	0.06	0.06
1276	96	158	0.09	0.15
1327	215	373	0.20	0.36
1359	459	832	0.44	0.79
1383	887	1719	0.84	1.64
1402	1639	3358	1.56	3.20
1419	2564	5922	2.44	5.64
1435	3375	9297	3.21	8.85
1449	4046	13343	3.85	12.70
1462	4366	17709	4.16	16.86
1475	4497	22206	4.28	21.14
1487	4694	26900	4.47	25.61
1499	4829	31729	4.60	30.21
1511	4904	36633	4.67	34.88
1522	5275	41908	5.02	39.90
1534	5428	47336	5.17	45.07
1546	5756	53092	5.48	50.55
1558	5878	58970	5.60	56.15
1571	5949	64919	5.66	61.81
1584	5897	70816	5.61	67.43
1598	5833	76649	5.55	72.98
1613	5639	82288	5.37	78.35
1630	5384	87672	5.13	83.47
1649	5104	92776	4.86	88.33
1673	4545	97321	4.33	92.66
1705	3782	101103	3.60	96.26
1755	2636	103739	2.51	98.77
1839	1290	105029	1.23	100.00

Figure D.5.24. 2024 TELPAS Grade 3 Listening
Frequency Distribution of Scale Scores



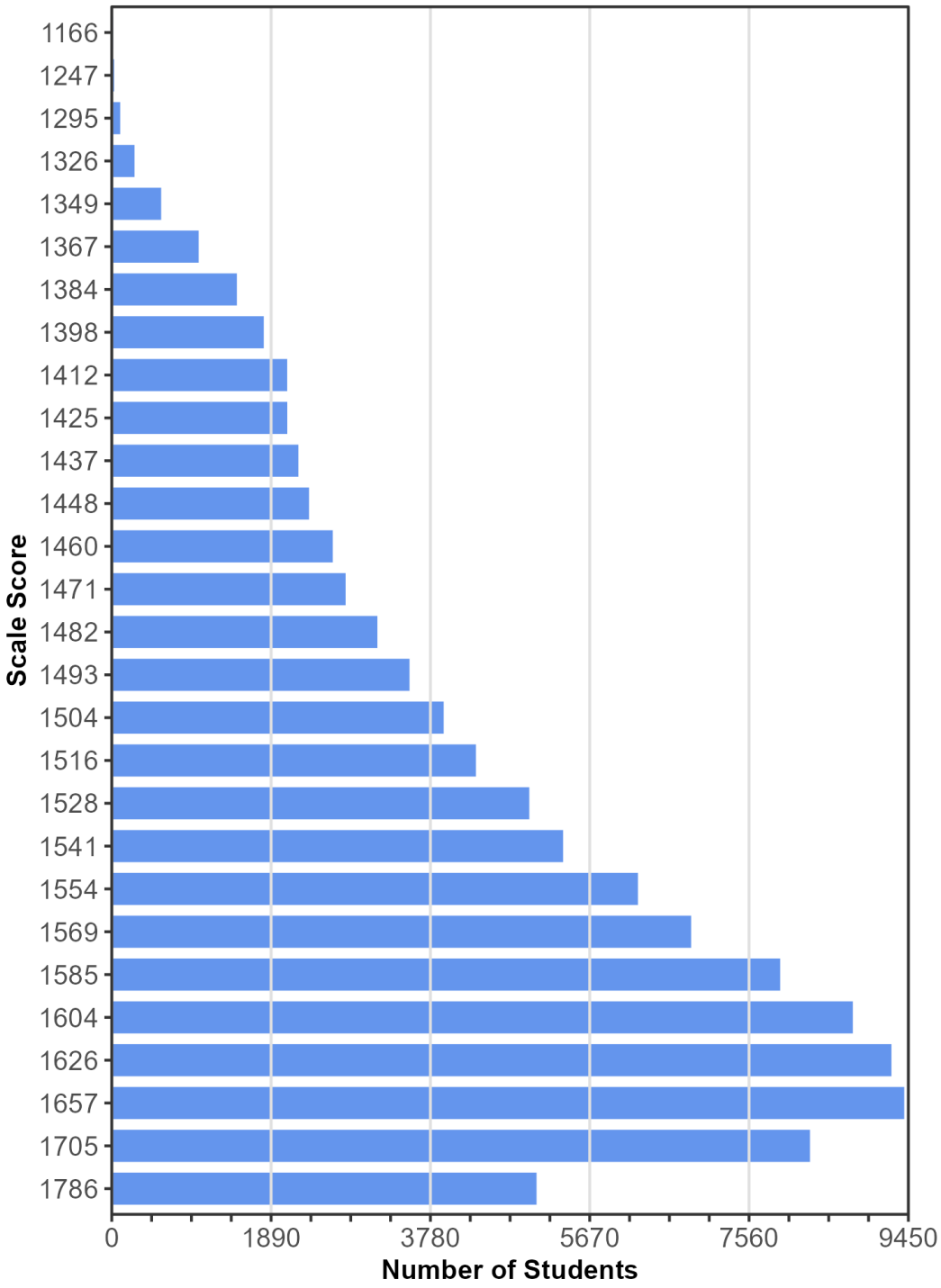
	Freq	Cum Freq	Pct	Cum Pct
1192	30	30	0.03	0.03
1276	25	55	0.02	0.05
1327	58	113	0.05	0.11
1359	194	307	0.18	0.29
1383	450	757	0.43	0.72
1402	834	1591	0.79	1.51
1419	1282	2873	1.21	2.72
1435	1728	4601	1.64	4.35
1449	2066	6667	1.95	6.31
1462	2234	8901	2.11	8.42
1475	2266	11167	2.14	10.57
1487	2489	13656	2.36	12.92
1499	2593	16249	2.45	15.38
1511	2837	19086	2.68	18.06
1522	3106	22192	2.94	21.00
1534	3489	25681	3.30	24.30
1546	4145	29826	3.92	28.22
1558	4483	34309	4.24	32.46
1571	4895	39204	4.63	37.10
1584	5366	44570	5.08	42.17
1598	5981	50551	5.66	47.83
1613	6595	57146	6.24	54.07
1630	7510	64656	7.11	61.18
1649	8123	72779	7.69	68.87
1673	8903	81682	8.42	77.29
1705	9318	91000	8.82	86.11
1755	8843	99843	8.37	94.48
1839	5837	105680	5.52	100.00

Figure D.5.25. 2024 TELPAS Grade 4 Listening
Frequency Distribution of Scale Scores



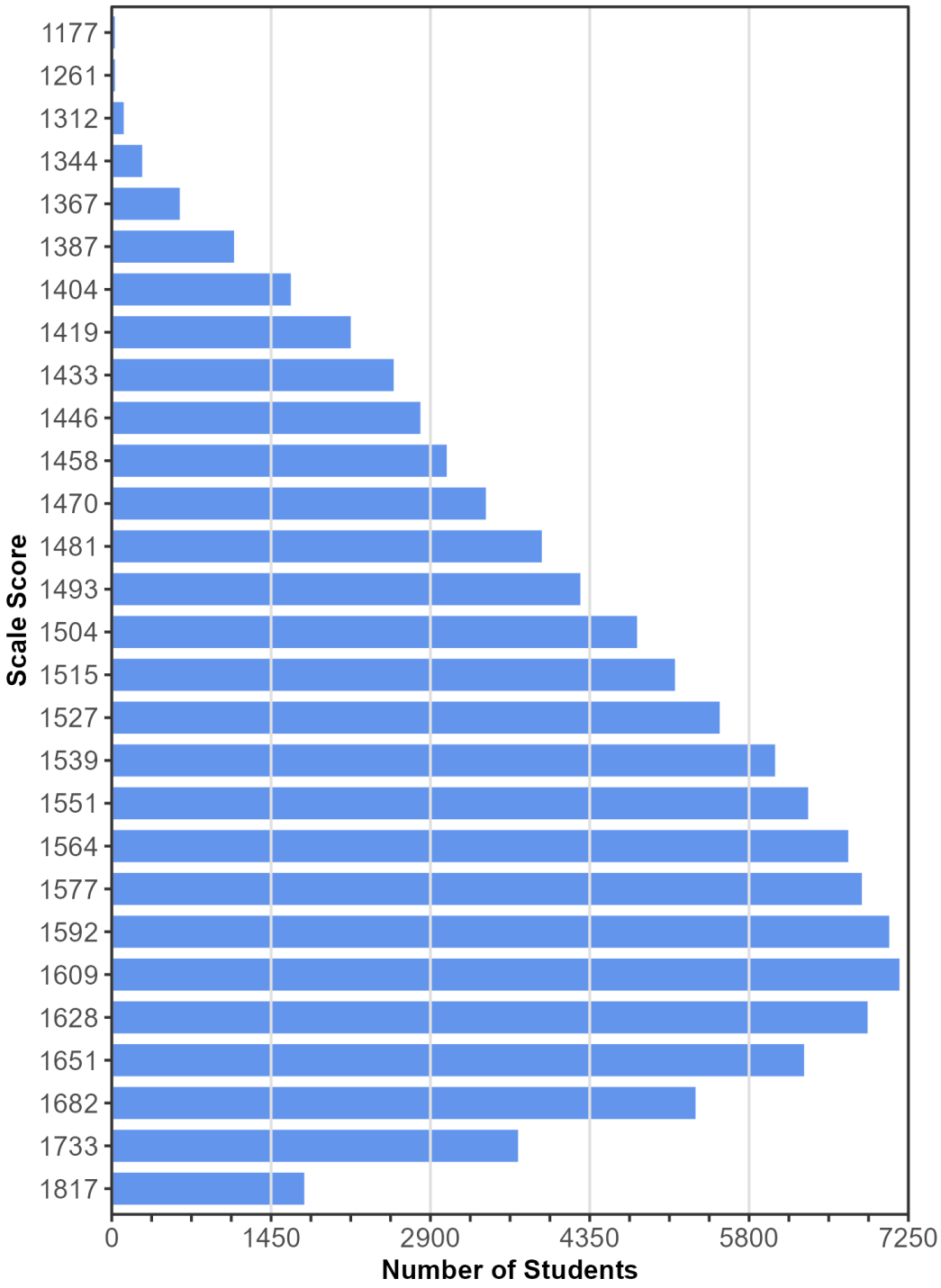
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1166	26	26	0.02	0.02
1247	23	49	0.02	0.05
1295	112	161	0.11	0.15
1326	329	490	0.31	0.46
1349	707	1197	0.67	1.13
1367	1191	2388	1.13	2.26
1384	1780	4168	1.68	3.94
1398	2410	6578	2.28	6.21
1412	2647	9225	2.50	8.71
1425	2888	12113	2.73	11.44
1437	3192	15305	3.02	14.46
1448	3331	18636	3.15	17.60
1460	3774	22410	3.57	21.17
1471	3997	26407	3.78	24.95
1482	4397	30804	4.15	29.10
1493	4711	35515	4.45	33.55
1504	5120	40635	4.84	38.39
1516	5377	46012	5.08	43.46
1528	5828	51840	5.51	48.97
1541	6221	58061	5.88	54.85
1554	6602	64663	6.24	61.08
1569	6970	71633	6.58	67.67
1585	7201	78834	6.80	74.47
1604	7411	86245	7.00	81.47
1626	6814	93059	6.44	87.91
1657	6086	99145	5.75	93.66
1705	4495	103640	4.25	97.90
1786	2220	105860	2.10	100.00

Figure D.5.26. 2024 TELPAS Grade 5 Listening
Frequency Distribution of Scale Scores



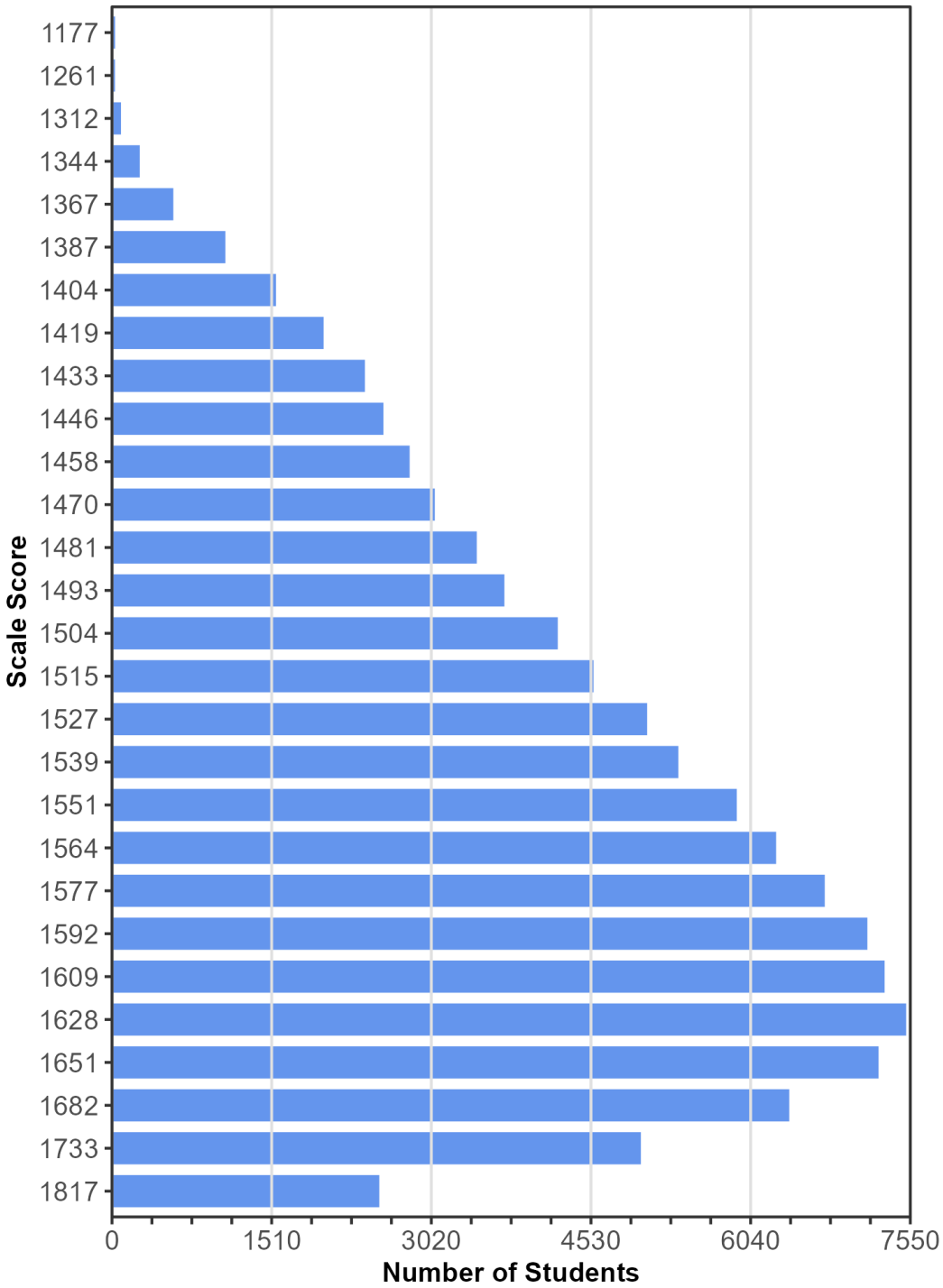
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1166	15	15	0.01	0.01
1247	28	43	0.03	0.04
1295	100	143	0.09	0.13
1326	270	413	0.25	0.39
1349	586	999	0.55	0.94
1367	1031	2030	0.97	1.91
1384	1484	3514	1.39	3.30
1398	1803	5317	1.69	4.99
1412	2081	7398	1.95	6.95
1425	2082	9480	1.96	8.90
1437	2214	11694	2.08	10.98
1448	2340	14034	2.20	13.18
1460	2622	16656	2.46	15.64
1471	2774	19430	2.61	18.25
1482	3151	22581	2.96	21.21
1493	3532	26113	3.32	24.52
1504	3937	30050	3.70	28.22
1516	4320	34370	4.06	32.28
1528	4953	39323	4.65	36.93
1541	5354	44677	5.03	41.96
1554	6242	50919	5.86	47.82
1569	6873	57792	6.45	54.27
1585	7929	65721	7.45	61.72
1604	8791	74512	8.26	69.97
1626	9249	83761	8.69	78.66
1657	9401	93162	8.83	87.49
1705	8283	101445	7.78	95.27
1786	5039	106484	4.73	100.00

Figure D.5.27. 2024 TELPAS Grade 6 Listening
Frequency Distribution of Scale Scores



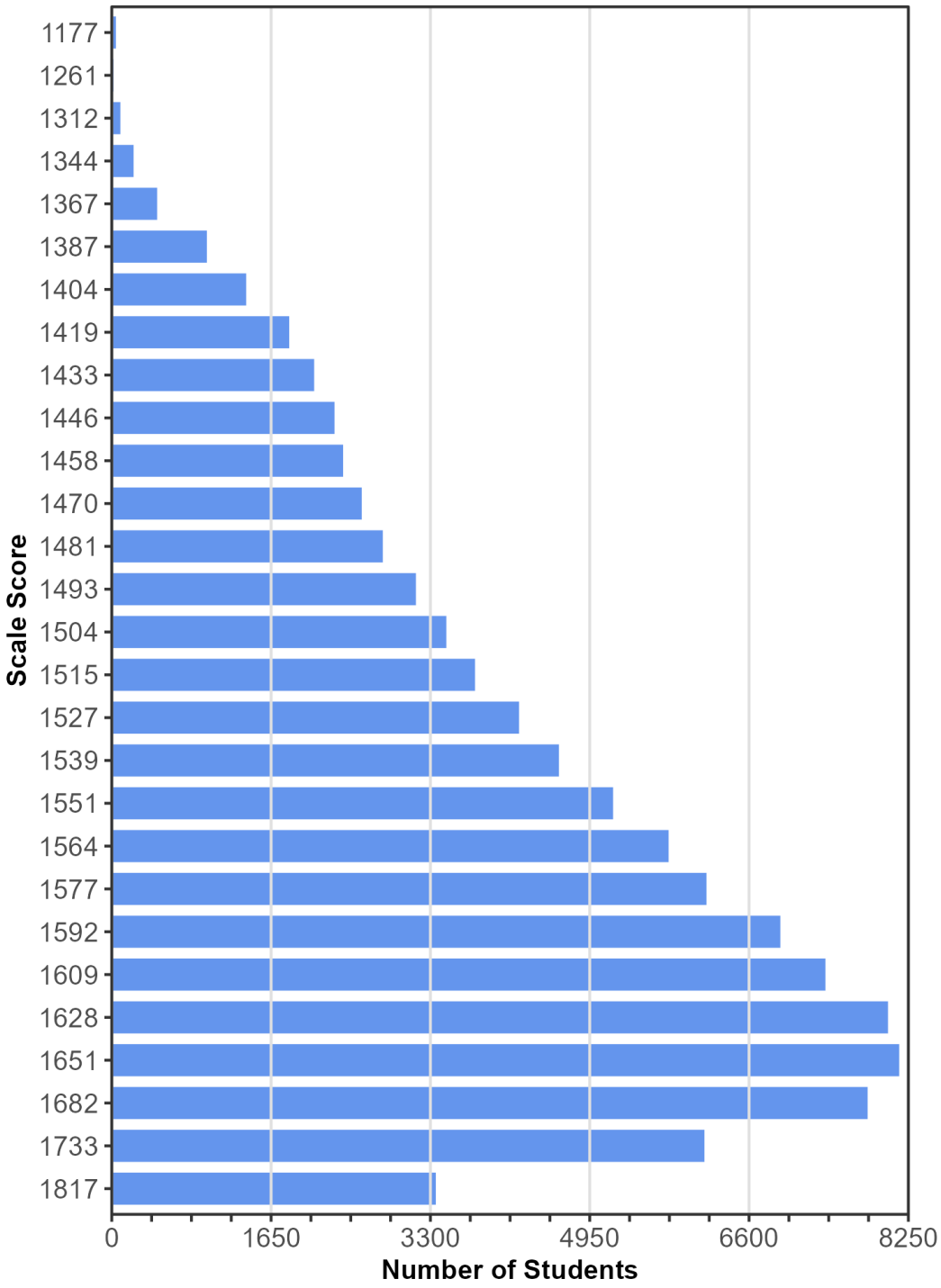
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1177	28	28	0.03	0.03
1261	29	57	0.03	0.05
1312	109	166	0.10	0.16
1344	277	443	0.26	0.42
1367	619	1062	0.59	1.01
1387	1113	2175	1.05	2.06
1404	1630	3805	1.54	3.61
1419	2175	5980	2.06	5.67
1433	2566	8546	2.43	8.10
1446	2809	11355	2.66	10.76
1458	3049	14404	2.89	13.65
1470	3405	17809	3.23	16.88
1481	3914	21723	3.71	20.59
1493	4265	25988	4.04	24.63
1504	4781	30769	4.53	29.16
1515	5126	35895	4.86	34.02
1527	5533	41428	5.24	39.26
1539	6037	47465	5.72	44.98
1551	6338	53803	6.01	50.99
1564	6703	60506	6.35	57.34
1577	6827	67333	6.47	63.81
1592	7076	74409	6.71	70.52
1609	7169	81578	6.79	77.31
1628	6879	88457	6.52	83.83
1651	6301	94758	5.97	89.80
1682	5313	100071	5.04	94.84
1733	3698	103769	3.50	98.34
1817	1752	105521	1.66	100.00

Figure D.5.28. 2024 TELPAS Grade 7 Listening
Frequency Distribution of Scale Scores



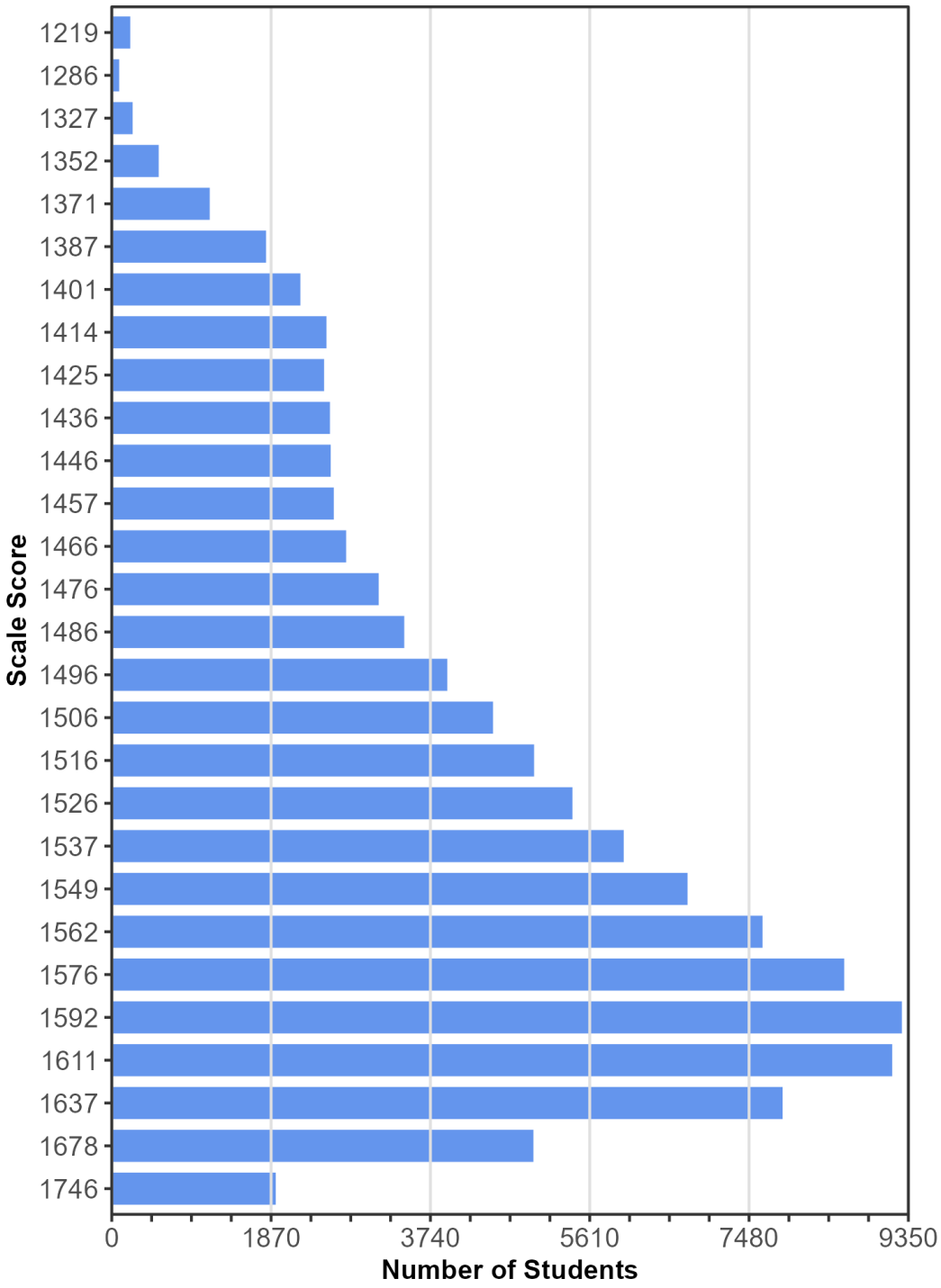
	Freq	Cum Freq	Pct	Cum Pct
1177	29	29	0.03	0.03
1261	28	57	0.03	0.05
1312	85	142	0.08	0.14
1344	262	404	0.25	0.39
1367	580	984	0.55	0.94
1387	1072	2056	1.02	1.96
1404	1551	3607	1.48	3.44
1419	2001	5608	1.91	5.35
1433	2392	8000	2.28	7.63
1446	2567	10567	2.45	10.08
1458	2815	13382	2.68	12.76
1470	3053	16435	2.91	15.67
1481	3450	19885	3.29	18.96
1493	3711	23596	3.54	22.50
1504	4216	27812	4.02	26.52
1515	4554	32366	4.34	30.87
1527	5061	37427	4.83	35.69
1539	5356	42783	5.11	40.80
1551	5909	48692	5.64	46.43
1564	6281	54973	5.99	52.42
1577	6741	61714	6.43	58.85
1592	7144	68858	6.81	65.67
1609	7307	76165	6.97	72.63
1628	7511	83676	7.16	79.80
1651	7250	90926	6.91	86.71
1682	6405	97331	6.11	92.82
1733	5002	102333	4.77	97.59
1817	2528	104861	2.41	100.00

Figure D.5.29. 2024 TELPAS Grade 8 Listening
Frequency Distribution of Scale Scores



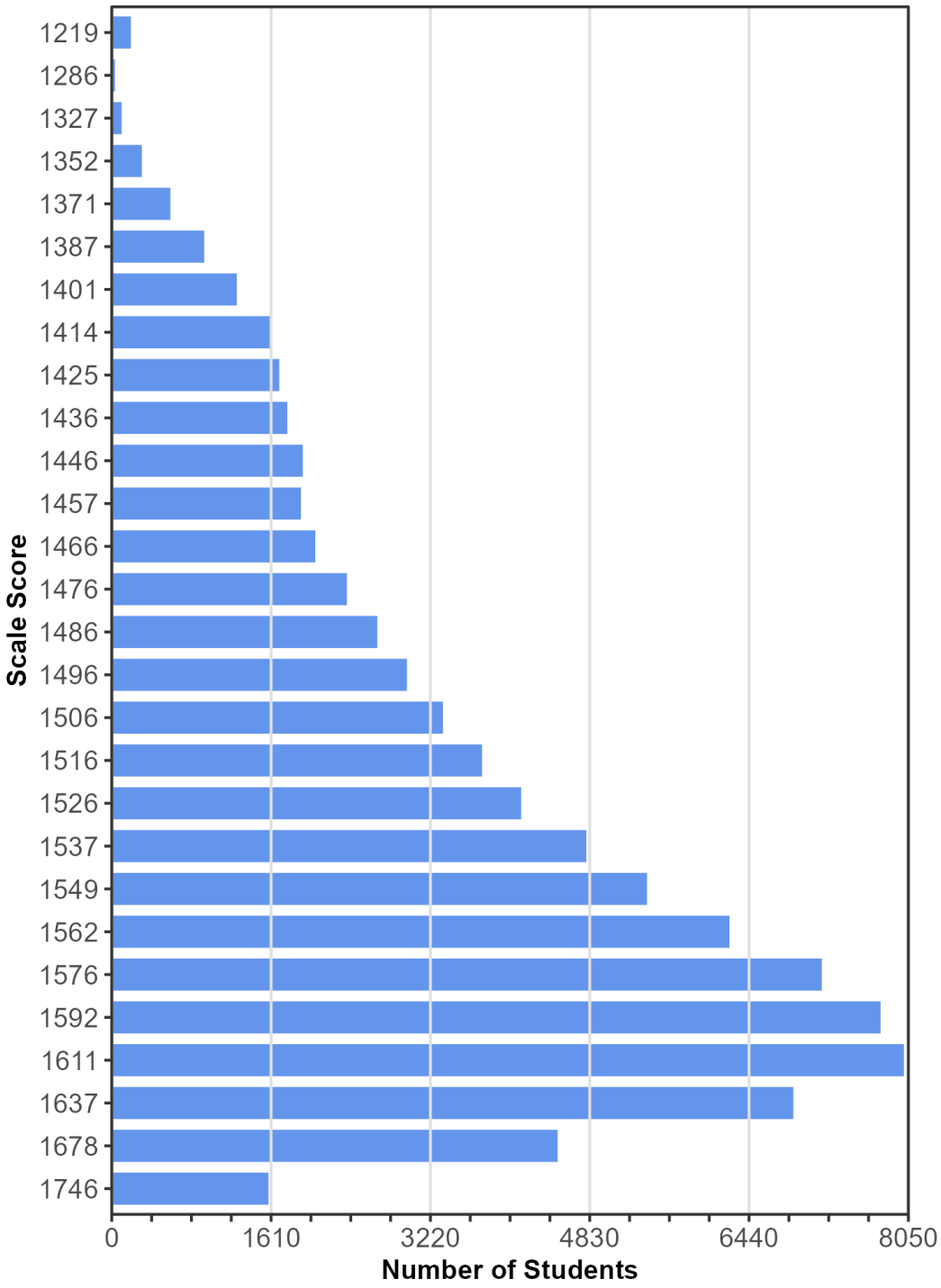
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1177	43	43	0.04	0.04
1261	19	62	0.02	0.06
1312	89	151	0.09	0.15
1344	226	377	0.22	0.37
1367	470	847	0.46	0.84
1387	985	1832	0.97	1.81
1404	1392	3224	1.37	3.18
1419	1838	5062	1.81	4.99
1433	2096	7158	2.07	7.06
1446	2307	9465	2.27	9.33
1458	2396	11861	2.36	11.69
1470	2589	14450	2.55	14.25
1481	2807	17257	2.77	17.01
1493	3150	20407	3.11	20.12
1504	3465	23872	3.42	23.53
1515	3762	27634	3.71	27.24
1527	4218	31852	4.16	31.40
1539	4631	36483	4.57	35.97
1551	5192	41675	5.12	41.09
1564	5767	47442	5.69	46.77
1577	6159	53601	6.07	52.84
1592	6924	60525	6.83	59.67
1609	7391	67916	7.29	66.96
1628	8039	75955	7.93	74.88
1651	8155	84110	8.04	82.92
1682	7828	91938	7.72	90.64
1733	6138	98076	6.05	96.69
1817	3356	101432	3.31	100.00

Figure D.5.30. 2024 TELPAS Grade 9 Listening
Frequency Distribution of Scale Scores



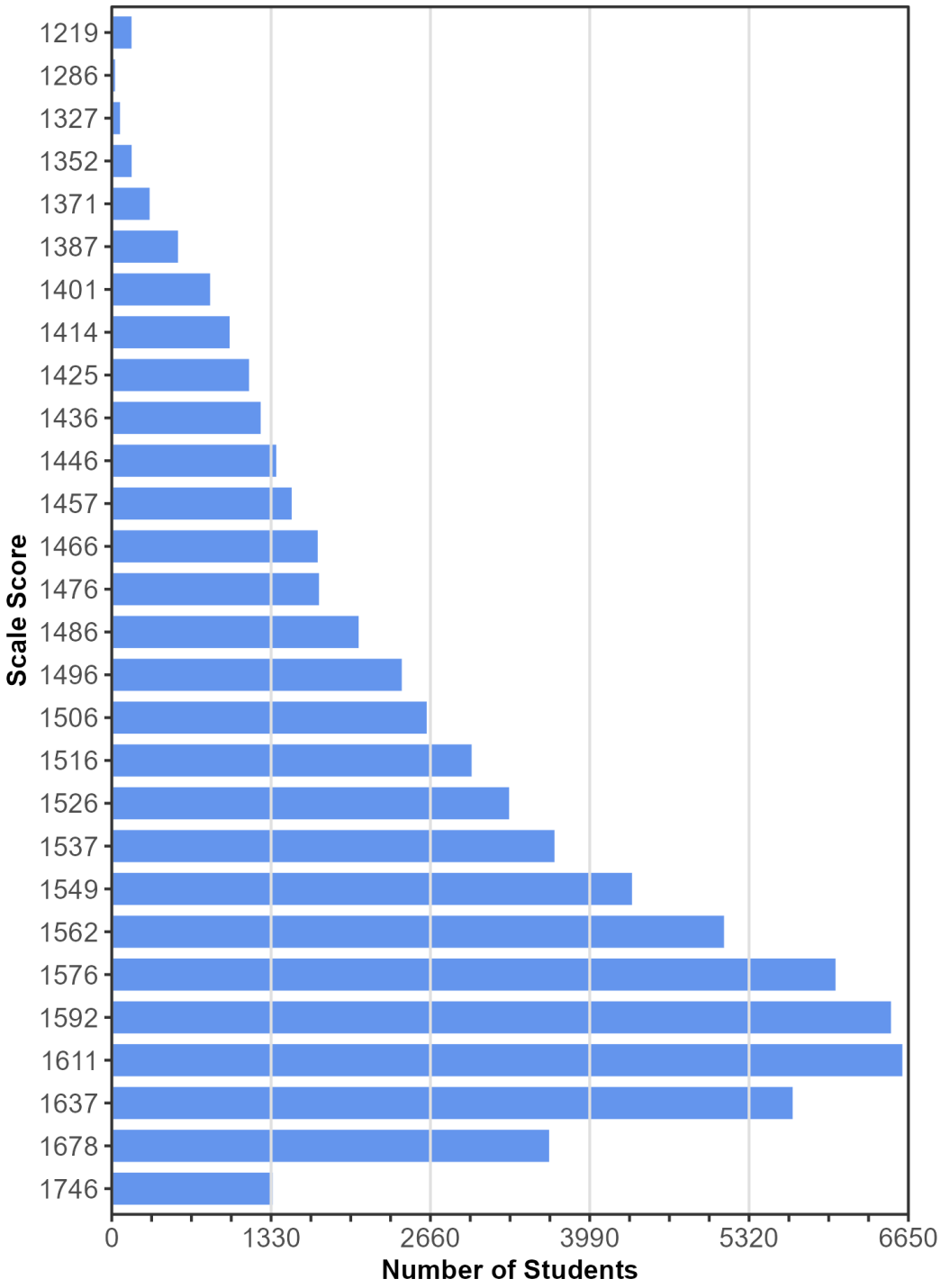
	Freq	Cum Freq	Pct	Cum Pct
1219	217	217	0.20	0.20
1286	87	304	0.08	0.28
1327	245	549	0.22	0.50
1352	551	1100	0.50	1.01
1371	1150	2250	1.05	2.06
1387	1811	4061	1.66	3.72
1401	2214	6275	2.03	5.74
1414	2520	8795	2.31	8.05
1425	2492	11287	2.28	10.33
1436	2561	13848	2.34	12.67
1446	2570	16418	2.35	15.02
1457	2606	19024	2.38	17.40
1466	2751	21775	2.52	19.92
1476	3133	24908	2.87	22.79
1486	3433	28341	3.14	25.93
1496	3938	32279	3.60	29.53
1506	4475	36754	4.09	33.63
1516	4957	41711	4.54	38.16
1526	5407	47118	4.95	43.11
1537	6009	53127	5.50	48.61
1549	6758	59885	6.18	54.79
1562	7638	67523	6.99	61.78
1576	8597	76120	7.87	69.64
1592	9274	85394	8.48	78.13
1611	9160	94554	8.38	86.51
1637	7874	102428	7.20	93.71
1678	4949	107377	4.53	98.24
1746	1925	109302	1.76	100.00

Figure D.5.31. 2024 TELPAS Grade 10 Listening
Frequency Distribution of Scale Scores



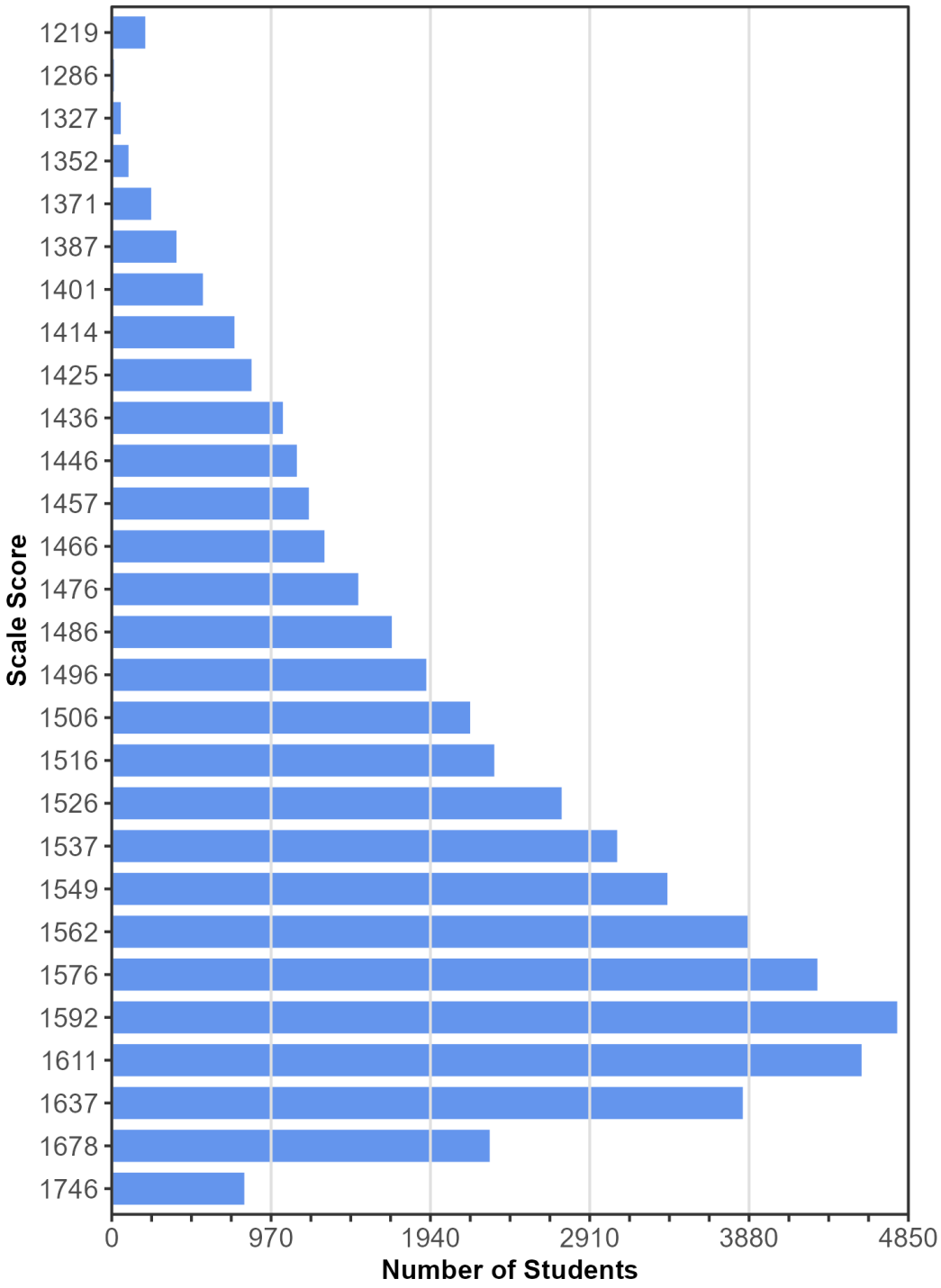
	Freq	Cum Freq	Pct	Cum Pct
1219	193	193	0.22	0.22
1286	32	225	0.04	0.26
1327	100	325	0.12	0.38
1352	304	629	0.35	0.73
1371	592	1221	0.69	1.42
1387	934	2155	1.09	2.50
1401	1264	3419	1.47	3.97
1414	1619	5038	1.88	5.86
1425	1693	6731	1.97	7.82
1436	1774	8505	2.06	9.88
1446	1931	10436	2.24	12.13
1457	1912	12348	2.22	14.35
1466	2057	14405	2.39	16.74
1476	2376	16781	2.76	19.50
1486	2683	19464	3.12	22.62
1496	2983	22447	3.47	26.09
1506	3347	25794	3.89	29.98
1516	3742	29536	4.35	34.33
1526	4137	33673	4.81	39.13
1537	4796	38469	5.57	44.71
1549	5409	43878	6.29	50.99
1562	6242	50120	7.25	58.25
1576	7175	57295	8.34	66.59
1592	7769	65064	9.03	75.62
1611	8005	73069	9.30	84.92
1637	6888	79957	8.01	92.92
1678	4506	84463	5.24	98.16
1746	1583	86046	1.84	100.00

Figure D.5.32. 2024 TELPAS Grade 11 Listening
Frequency Distribution of Scale Scores



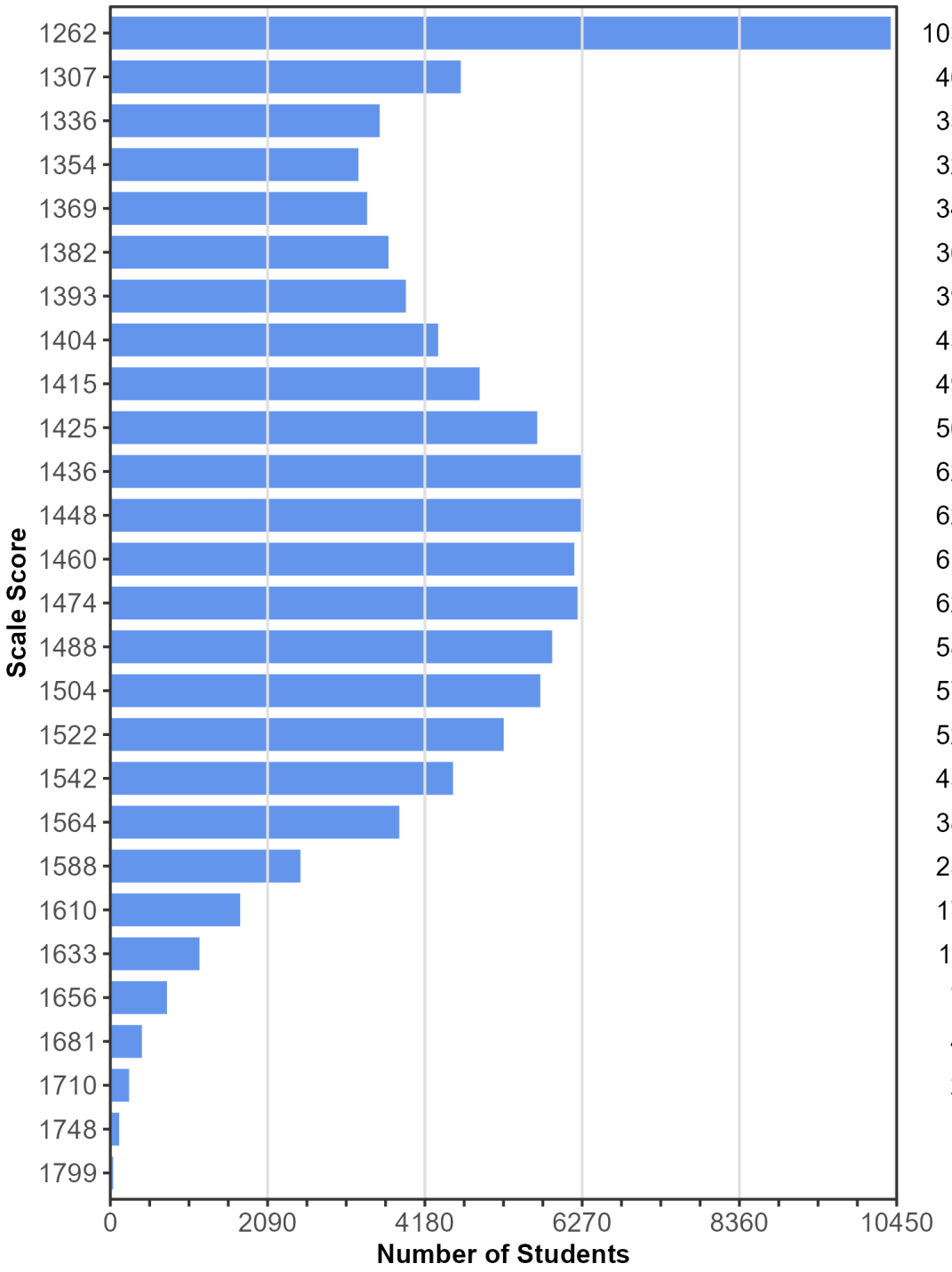
	Freq	Cum Freq	Pct	Cum Pct
1219	165	165	0.24	0.24
1286	27	192	0.04	0.28
1327	69	261	0.10	0.38
1352	166	427	0.24	0.63
1371	316	743	0.46	1.09
1387	553	1296	0.81	1.90
1401	821	2117	1.20	3.10
1414	984	3101	1.44	4.55
1425	1146	4247	1.68	6.23
1436	1243	5490	1.82	8.05
1446	1373	6863	2.01	10.06
1457	1502	8365	2.20	12.26
1466	1719	10084	2.52	14.78
1476	1730	11814	2.54	17.32
1486	2061	13875	3.02	20.34
1496	2421	16296	3.55	23.89
1506	2629	18925	3.85	27.74
1516	3004	21929	4.40	32.14
1526	3317	25246	4.86	37.01
1537	3696	28942	5.42	42.42
1549	4343	33285	6.37	48.79
1562	5111	38396	7.49	56.28
1576	6042	44438	8.86	65.14
1592	6505	50943	9.54	74.67
1611	6600	57543	9.67	84.35
1637	5684	63227	8.33	92.68
1678	3651	66878	5.35	98.03
1746	1343	68221	1.97	100.00

Figure D.5.33. 2024 TELPAS Grade 12 Listening
Frequency Distribution of Scale Scores



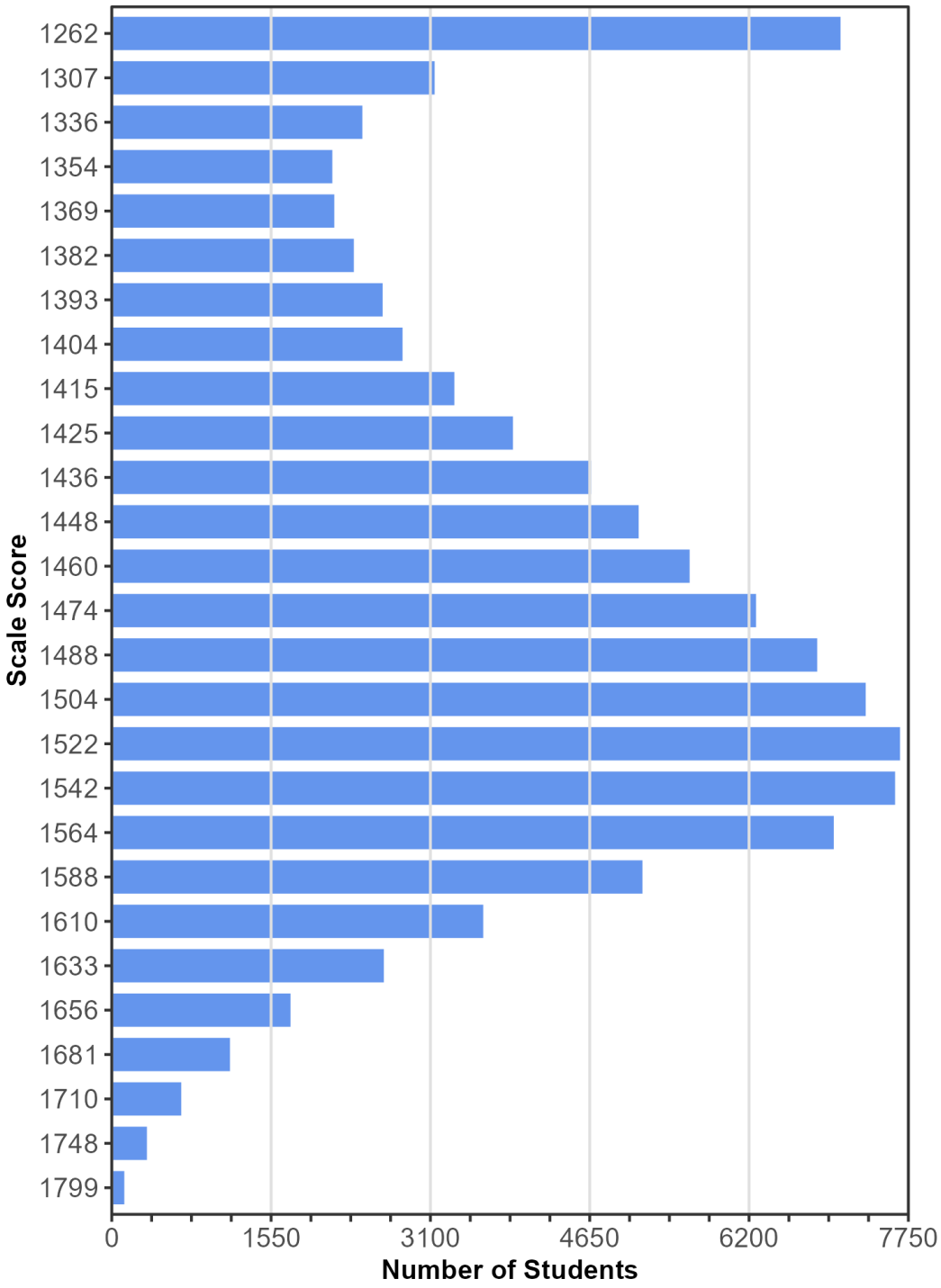
	Freq	Cum Freq	Pct	Cum Pct
1219	204	204	0.40	0.40
1286	13	217	0.03	0.42
1327	55	272	0.11	0.53
1352	102	374	0.20	0.73
1371	240	614	0.47	1.20
1387	394	1008	0.77	1.97
1401	555	1563	1.09	3.06
1414	747	2310	1.46	4.52
1425	851	3161	1.66	6.18
1436	1042	4203	2.04	8.22
1446	1127	5330	2.20	10.43
1457	1200	6530	2.35	12.77
1466	1295	7825	2.53	15.31
1476	1501	9326	2.94	18.24
1486	1705	11031	3.34	21.58
1496	1915	12946	3.75	25.32
1506	2182	15128	4.27	29.59
1516	2329	17457	4.56	34.15
1526	2739	20196	5.36	39.51
1537	3077	23273	6.02	45.53
1549	3383	26656	6.62	52.14
1562	3871	30527	7.57	59.72
1576	4296	34823	8.40	68.12
1592	4782	39605	9.35	77.47
1611	4565	44170	8.93	86.40
1637	3842	48012	7.52	93.92
1678	2301	50313	4.50	98.42
1746	807	51120	1.58	100.00

Figure D.5.34. 2024 TELPAS Grade 2 Speaking
Frequency Distribution of Scale Scores



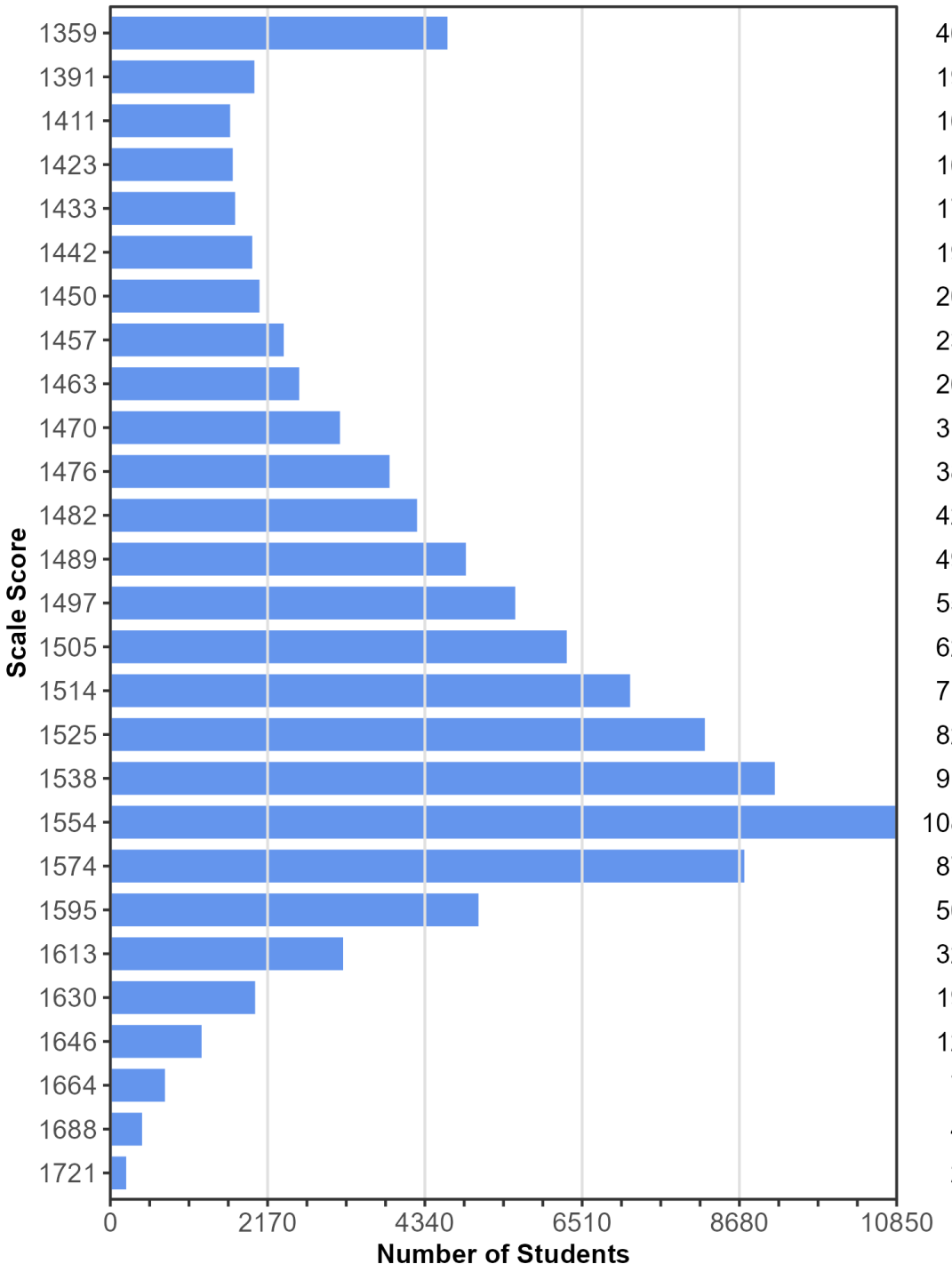
	Freq	Cum Freq	Pct	Cum Pct
1262	10371	10371	9.87	9.87
1307	4657	15028	4.43	14.31
1336	3580	18608	3.41	17.72
1354	3298	21906	3.14	20.86
1369	3413	25319	3.25	24.11
1382	3697	29016	3.52	27.63
1393	3928	32944	3.74	31.37
1404	4357	37301	4.15	35.51
1415	4908	42209	4.67	40.19
1425	5673	47882	5.40	45.59
1436	6257	54139	5.96	51.55
1448	6291	60430	5.99	57.54
1460	6166	66596	5.87	63.41
1474	6210	72806	5.91	69.32
1488	5872	78678	5.59	74.91
1504	5714	84392	5.44	80.35
1522	5228	89620	4.98	85.33
1542	4554	94174	4.34	89.66
1564	3840	98014	3.66	93.32
1588	2527	100541	2.41	95.73
1610	1726	102267	1.64	97.37
1633	1185	103452	1.13	98.50
1656	753	104205	0.72	99.22
1681	420	104625	0.40	99.62
1710	249	104874	0.24	99.85
1748	118	104992	0.11	99.96
1799	37	105029	0.04	100.00

Figure D.5.35. 2024 TELPAS Grade 3 Speaking
Frequency Distribution of Scale Scores



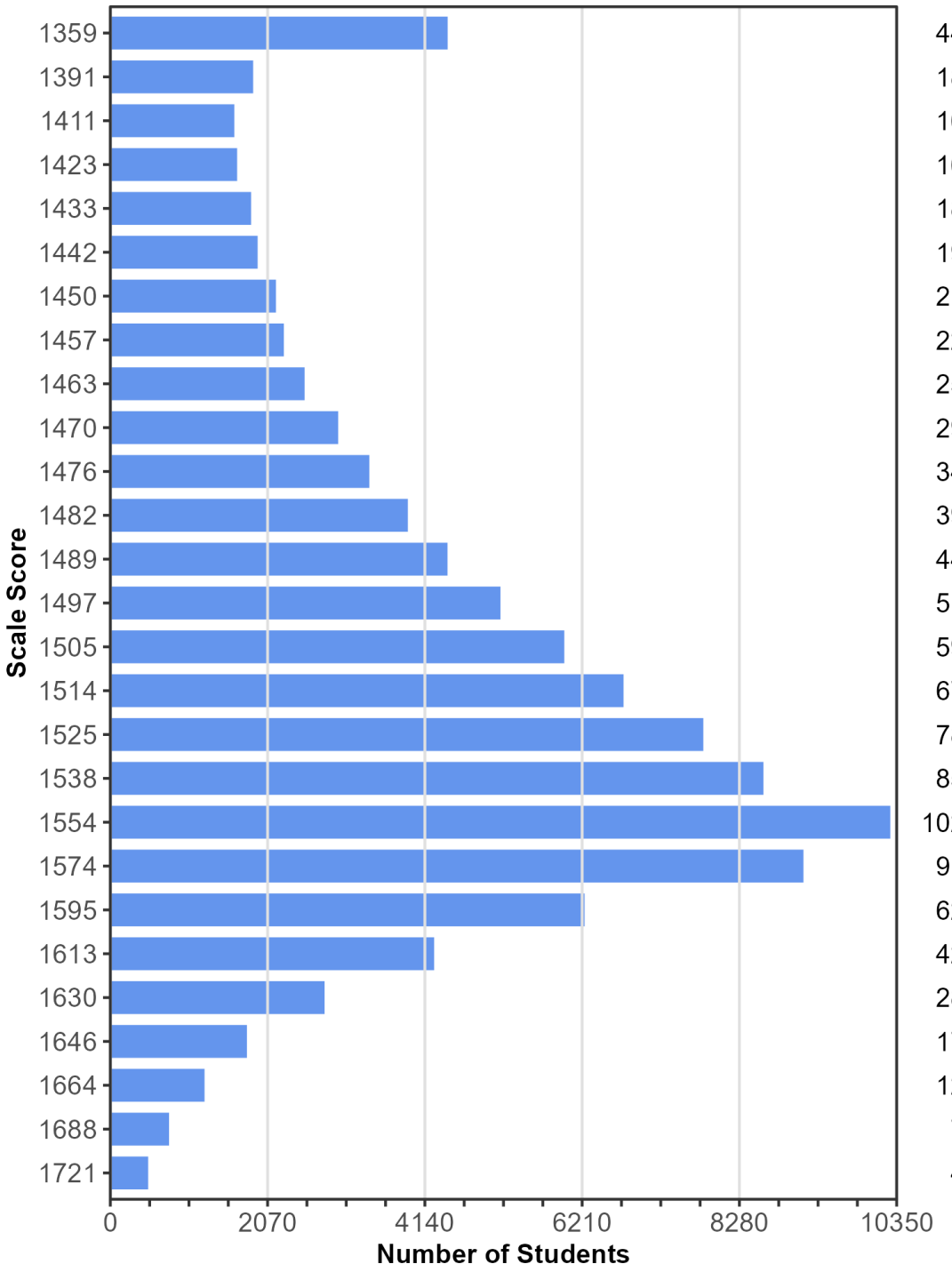
	Freq	Cum Freq	Pct	Cum Pct
1262	7090	7090	6.71	6.71
1307	3141	10231	2.97	9.68
1336	2438	12669	2.31	11.99
1354	2146	14815	2.03	14.02
1369	2165	16980	2.05	16.07
1382	2355	19335	2.23	18.30
1393	2635	21970	2.49	20.79
1404	2829	24799	2.68	23.47
1415	3333	28132	3.15	26.62
1425	3903	32035	3.69	30.31
1436	4664	36699	4.41	34.73
1448	5126	41825	4.85	39.58
1460	5622	47447	5.32	44.90
1474	6268	53715	5.93	50.83
1488	6863	60578	6.49	57.32
1504	7335	67913	6.94	64.26
1522	7669	75582	7.26	71.52
1542	7621	83203	7.21	78.73
1564	7024	90227	6.65	85.38
1588	5163	95390	4.89	90.26
1610	3614	99004	3.42	93.68
1633	2647	101651	2.50	96.19
1656	1739	103390	1.65	97.83
1681	1150	104540	1.09	98.92
1710	676	105216	0.64	99.56
1748	342	105558	0.32	99.88
1799	122	105680	0.12	100.00

Figure D.5.36. 2024 TELPAS Grade 4 Speaking
Frequency Distribution of Scale Scores



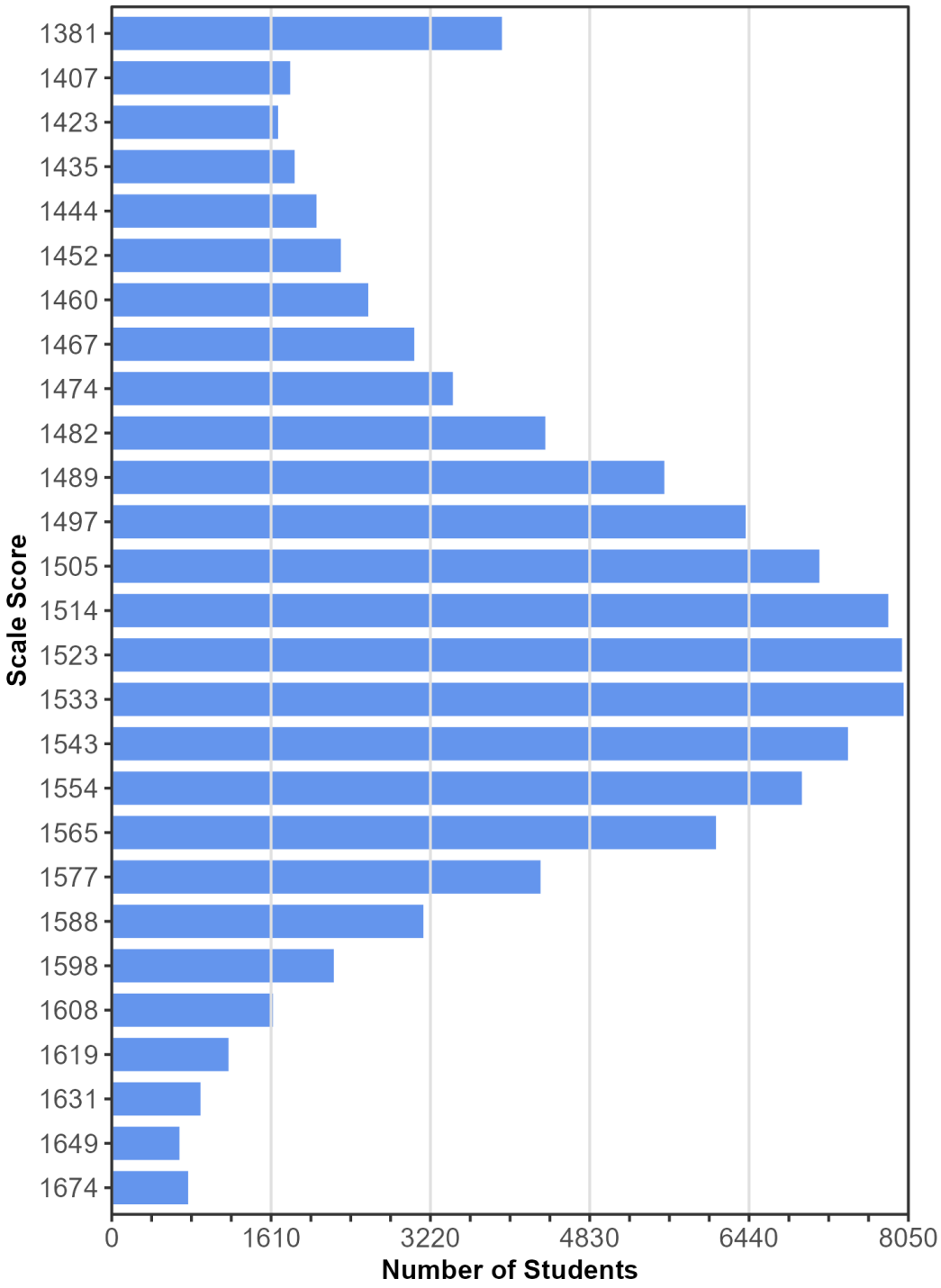
	Freq	Cum Freq	Pct	Cum Pct
1359	4652	4652	4.39	4.39
1391	1988	6640	1.88	6.27
1411	1653	8293	1.56	7.83
1423	1689	9982	1.60	9.43
1433	1722	11704	1.63	11.06
1442	1958	13662	1.85	12.91
1450	2058	15720	1.94	14.85
1457	2393	18113	2.26	17.11
1463	2606	20719	2.46	19.57
1470	3169	23888	2.99	22.57
1476	3853	27741	3.64	26.21
1482	4232	31973	4.00	30.20
1489	4906	36879	4.63	34.84
1497	5587	42466	5.28	40.12
1505	6298	48764	5.95	46.06
1514	7172	55936	6.77	52.84
1525	8203	64139	7.75	60.59
1538	9169	73308	8.66	69.25
1554	10844	84152	10.24	79.49
1574	8748	92900	8.26	87.76
1595	5080	97980	4.80	92.56
1613	3212	101192	3.03	95.59
1630	1998	103190	1.89	97.48
1646	1259	104449	1.19	98.67
1664	754	105203	0.71	99.38
1688	438	105641	0.41	99.79
1721	219	105860	0.21	100.00

Figure D.5.37. 2024 TELPAS Grade 5 Speaking
Frequency Distribution of Scale Scores



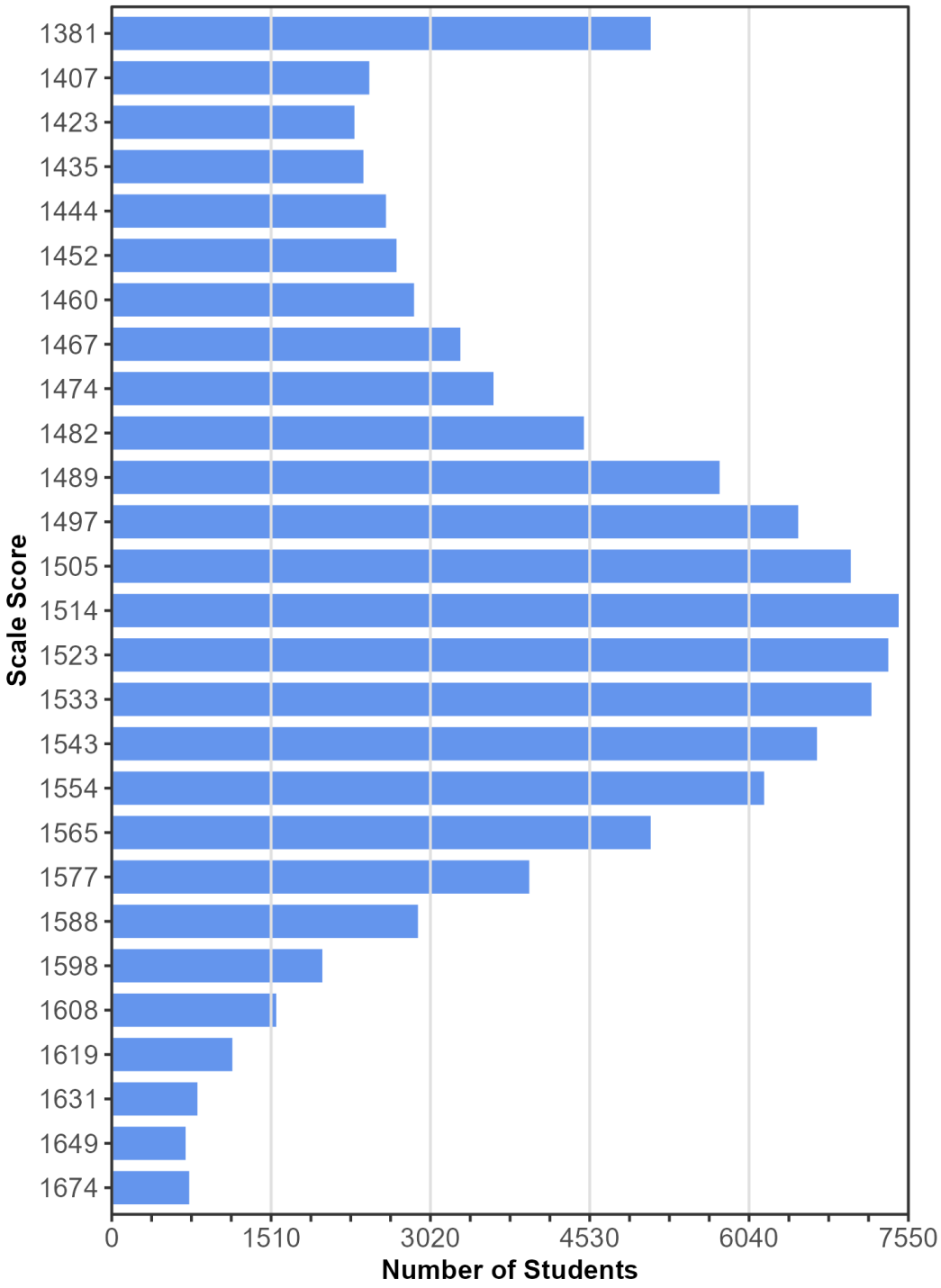
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1359	4440	4440	4.17	4.17
1391	1880	6320	1.77	5.94
1411	1632	7952	1.53	7.47
1423	1669	9621	1.57	9.04
1433	1853	11474	1.74	10.78
1442	1939	13413	1.82	12.60
1450	2180	15593	2.05	14.64
1457	2284	17877	2.14	16.79
1463	2557	20434	2.40	19.19
1470	2999	23433	2.82	22.01
1476	3409	26842	3.20	25.21
1482	3916	30758	3.68	28.89
1489	4438	35196	4.17	33.05
1497	5135	40331	4.82	37.88
1505	5975	46306	5.61	43.49
1514	6755	53061	6.34	49.83
1525	7805	60866	7.33	57.16
1538	8596	69462	8.07	65.23
1554	10266	79728	9.64	74.87
1574	9123	88851	8.57	83.44
1595	6242	95093	5.86	89.30
1613	4262	99355	4.00	93.31
1630	2820	102175	2.65	95.95
1646	1798	103973	1.69	97.64
1664	1240	105213	1.16	98.81
1688	773	105986	0.73	99.53
1721	498	106484	0.47	100.00

Figure D.5.38. 2024 TELPAS Grade 6 Speaking
Frequency Distribution of Scale Scores



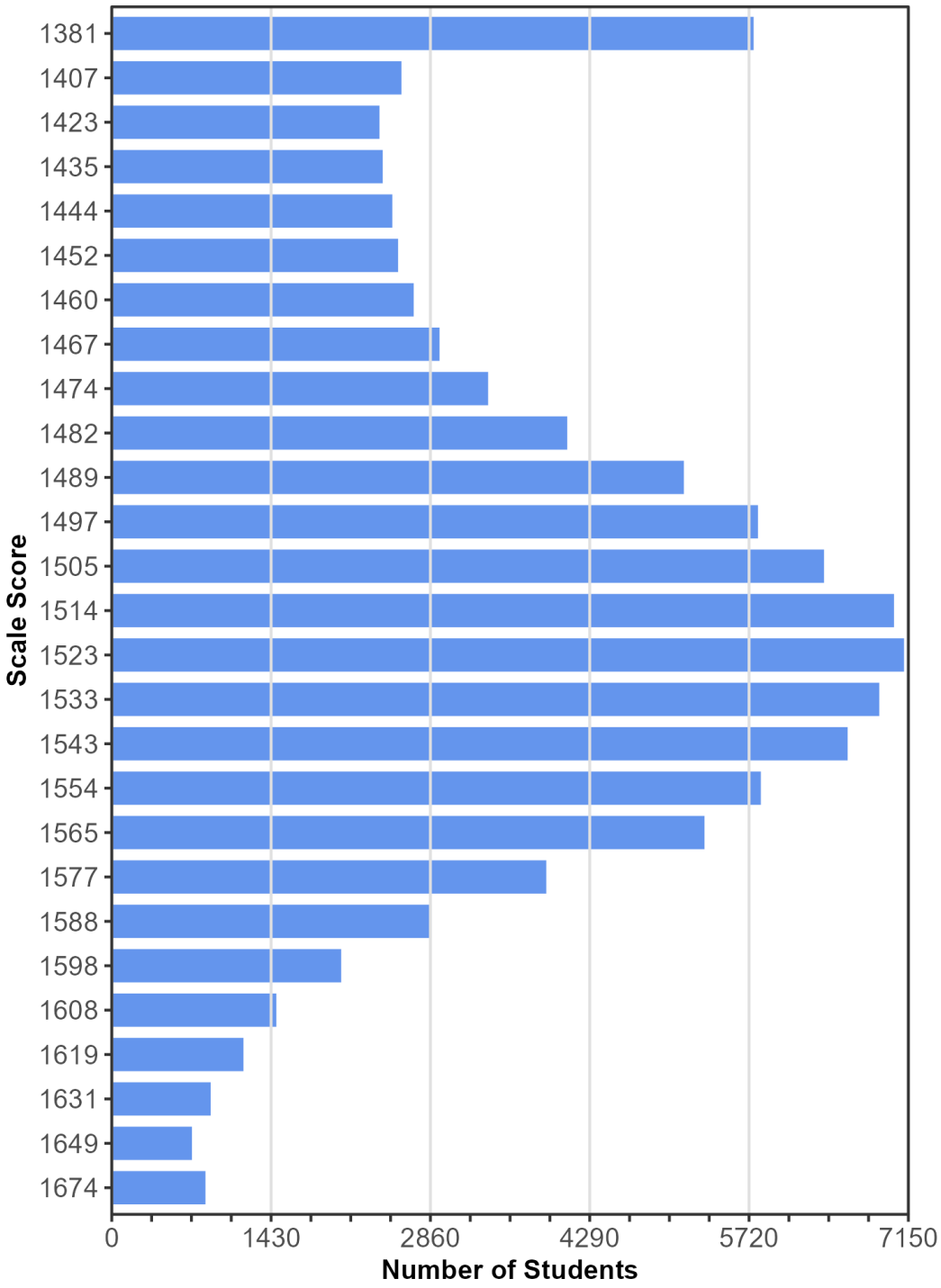
	Freq	Cum Freq	Pct	Cum Pct
1381	3943	3943	3.74	3.74
1407	1803	5746	1.71	5.45
1423	1681	7427	1.59	7.04
1435	1848	9275	1.75	8.79
1444	2069	11344	1.96	10.75
1452	2315	13659	2.19	12.94
1460	2592	16251	2.46	15.40
1467	3057	19308	2.90	18.30
1474	3447	22755	3.27	21.56
1482	4381	27136	4.15	25.72
1489	5584	32720	5.29	31.01
1497	6406	39126	6.07	37.08
1505	7151	46277	6.78	43.86
1514	7847	54124	7.44	51.29
1523	7986	62110	7.57	58.86
1533	8002	70112	7.58	66.44
1543	7440	77552	7.05	73.49
1554	6974	84526	6.61	80.10
1565	6106	90632	5.79	85.89
1577	4333	94965	4.11	90.00
1588	3149	98114	2.98	92.98
1598	2244	100358	2.13	95.11
1608	1630	101988	1.54	96.65
1619	1180	103168	1.12	97.77
1631	897	104065	0.85	98.62
1649	684	104749	0.65	99.27
1674	772	105521	0.73	100.00

Figure D.5.39. 2024 TELPAS Grade 7 Speaking
Frequency Distribution of Scale Scores



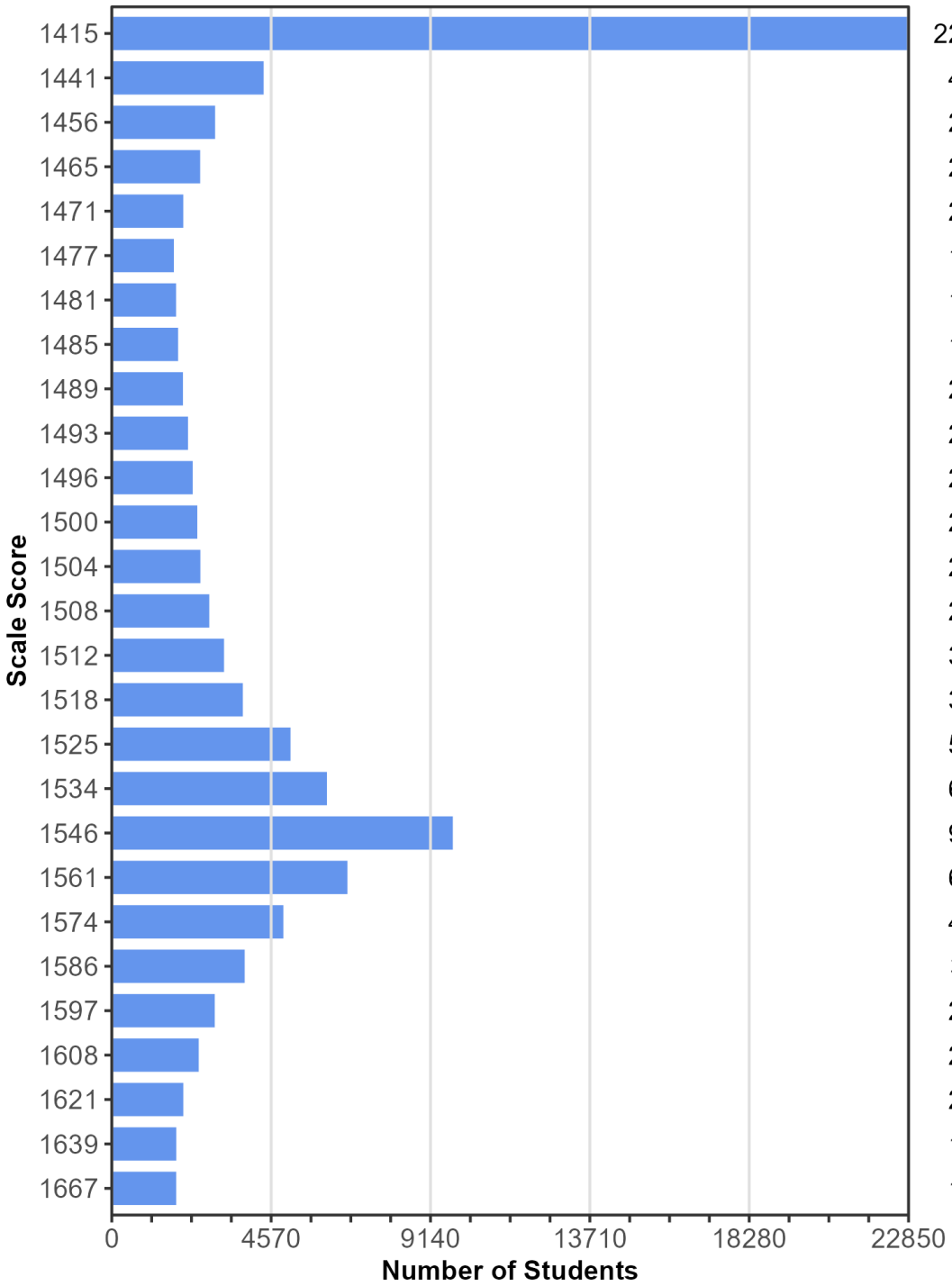
	Freq	Cum Freq	Pct	Cum Pct
1381	5108	5108	4.87	4.87
1407	2441	7549	2.33	7.20
1423	2300	9849	2.19	9.39
1435	2385	12234	2.27	11.67
1444	2599	14833	2.48	14.15
1452	2699	17532	2.57	16.72
1460	2865	20397	2.73	19.45
1467	3304	23701	3.15	22.60
1474	3618	27319	3.45	26.05
1482	4475	31794	4.27	30.32
1489	5761	37555	5.49	35.81
1497	6506	44061	6.20	42.02
1505	7004	51065	6.68	48.70
1514	7458	58523	7.11	55.81
1523	7360	65883	7.02	62.83
1533	7200	73083	6.87	69.70
1543	6684	79767	6.37	76.07
1554	6183	85950	5.90	81.97
1565	5108	91058	4.87	86.84
1577	3957	95015	3.77	90.61
1588	2902	97917	2.77	93.38
1598	1996	99913	1.90	95.28
1608	1559	101472	1.49	96.77
1619	1143	102615	1.09	97.86
1631	812	103427	0.77	98.63
1649	700	104127	0.67	99.30
1674	734	104861	0.70	100.00

Figure D.5.40. 2024 TELPAS Grade 8 Speaking
Frequency Distribution of Scale Scores



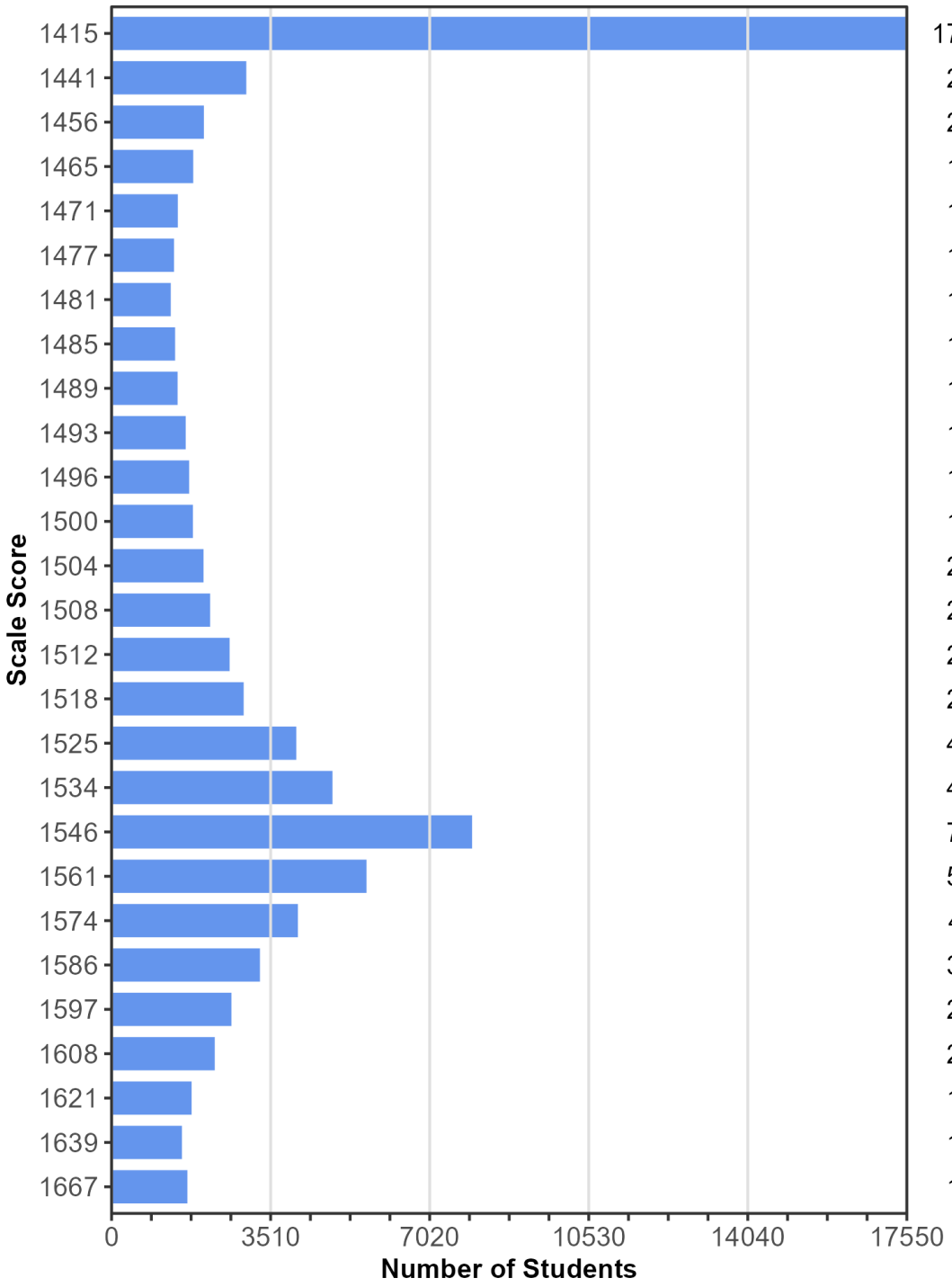
	Freq	Cum Freq	Pct	Cum Pct
1381	5761	5761	5.68	5.68
1407	2600	8361	2.56	8.24
1423	2403	10764	2.37	10.61
1435	2432	13196	2.40	13.01
1444	2518	15714	2.48	15.49
1452	2570	18284	2.53	18.03
1460	2710	20994	2.67	20.70
1467	2942	23936	2.90	23.60
1474	3378	27314	3.33	26.93
1482	4088	31402	4.03	30.96
1489	5135	36537	5.06	36.02
1497	5800	42337	5.72	41.74
1505	6393	48730	6.30	48.04
1514	7021	55751	6.92	54.96
1523	7111	62862	7.01	61.97
1533	6889	69751	6.79	68.77
1543	6605	76356	6.51	75.28
1554	5826	82182	5.74	81.02
1565	5320	87502	5.24	86.27
1577	3900	91402	3.84	90.11
1588	2863	94265	2.82	92.93
1598	2059	96324	2.03	94.96
1608	1477	97801	1.46	96.42
1619	1182	98983	1.17	97.59
1631	888	99871	0.88	98.46
1649	720	100591	0.71	99.17
1674	841	101432	0.83	100.00

Figure D.5.41. 2024 TELPAS Grade 9 Speaking
Frequency Distribution of Scale Scores



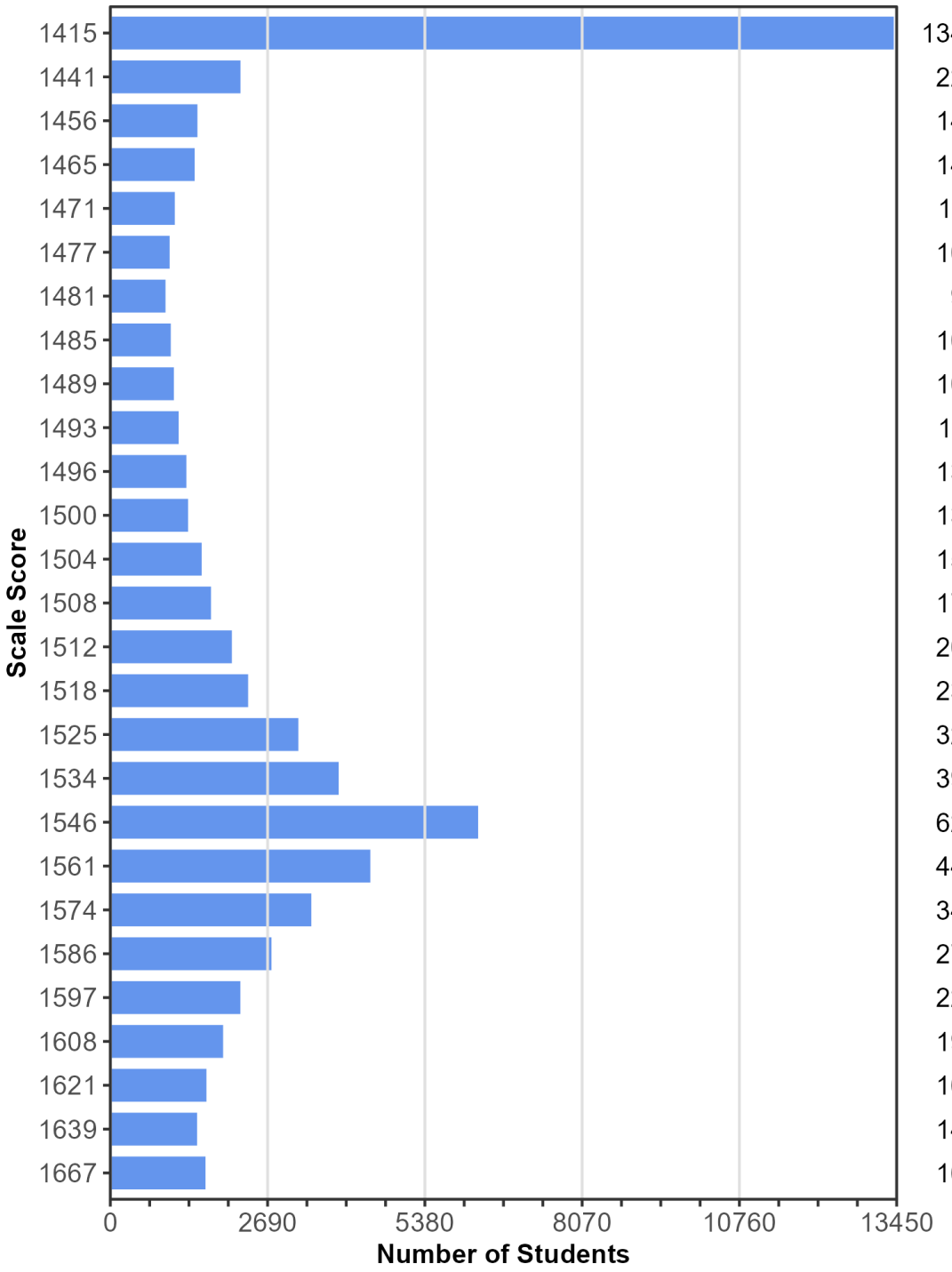
	Freq	Cum Freq	Pct	Cum Pct
1415	22806	22806	20.87	20.87
1441	4355	27161	3.98	24.85
1456	2963	30124	2.71	27.56
1465	2534	32658	2.32	29.88
1471	2050	34708	1.88	31.75
1477	1781	36489	1.63	33.38
1481	1844	38333	1.69	35.07
1485	1902	40235	1.74	36.81
1489	2041	42276	1.87	38.68
1493	2186	44462	2.00	40.68
1496	2320	46782	2.12	42.80
1500	2452	49234	2.24	45.04
1504	2540	51774	2.32	47.37
1508	2796	54570	2.56	49.93
1512	3219	57789	2.95	52.87
1518	3758	61547	3.44	56.31
1525	5125	66672	4.69	61.00
1534	6169	72841	5.64	66.64
1546	9780	82621	8.95	75.59
1561	6757	89378	6.18	81.77
1574	4920	94298	4.50	86.27
1586	3811	98109	3.49	89.76
1597	2953	101062	2.70	92.46
1608	2493	103555	2.28	94.74
1621	2051	105606	1.88	96.62
1639	1849	107455	1.69	98.31
1667	1847	109302	1.69	100.00

Figure D.5.42. 2024 TELPAS Grade 10 Speaking
Frequency Distribution of Scale Scores



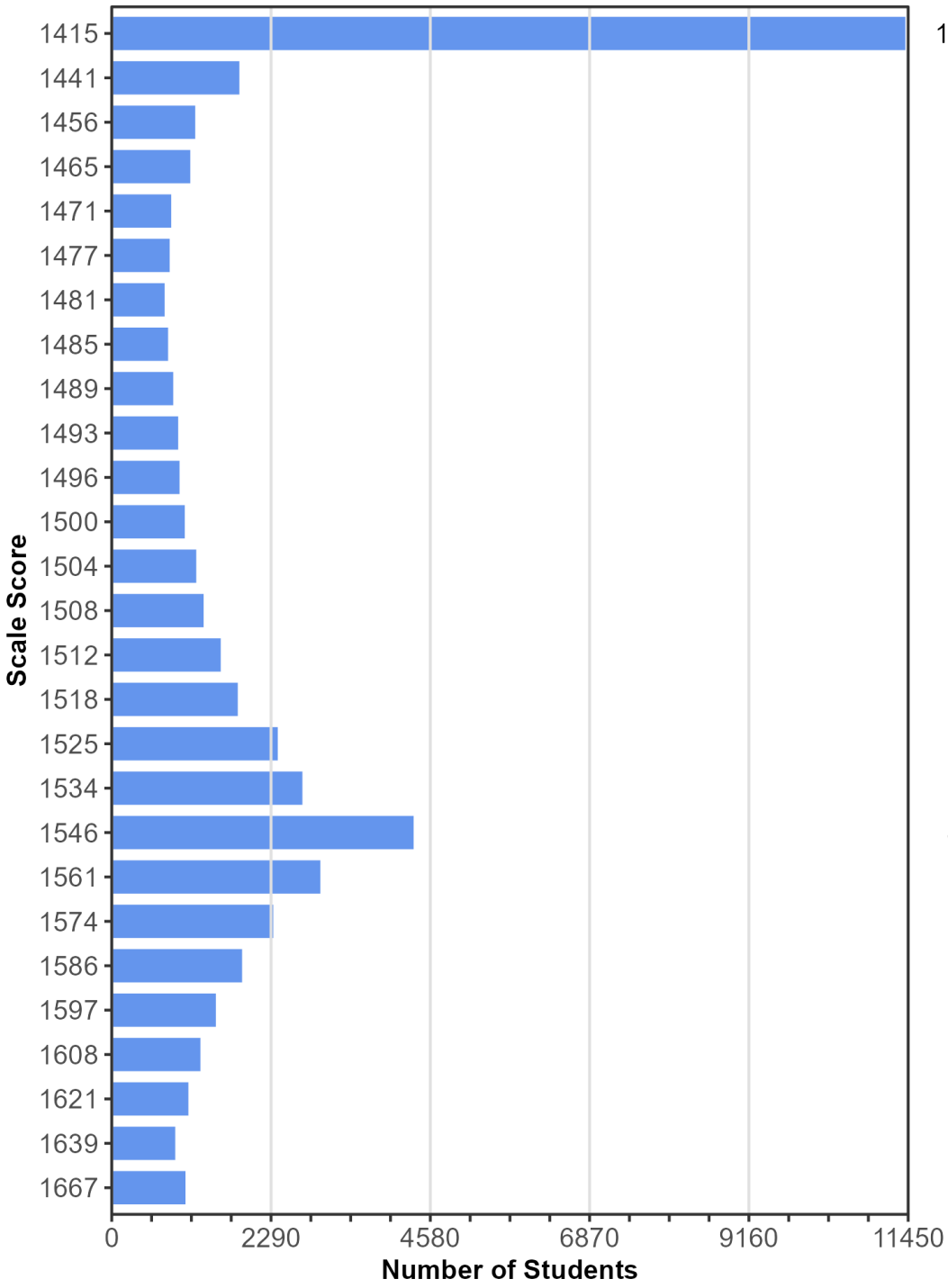
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1415	17529	17529	20.37	20.37
1441	2973	20502	3.46	23.83
1456	2037	22539	2.37	26.19
1465	1802	24341	2.09	28.29
1471	1462	25803	1.70	29.99
1477	1378	27181	1.60	31.59
1481	1306	28487	1.52	33.11
1485	1402	29889	1.63	34.74
1489	1457	31346	1.69	36.43
1493	1636	32982	1.90	38.33
1496	1712	34694	1.99	40.32
1500	1795	36489	2.09	42.41
1504	2029	38518	2.36	44.76
1508	2175	40693	2.53	47.29
1512	2605	43298	3.03	50.32
1518	2915	46213	3.39	53.71
1525	4076	50289	4.74	58.44
1534	4875	55164	5.67	64.11
1546	7956	63120	9.25	73.36
1561	5626	68746	6.54	79.89
1574	4112	72858	4.78	84.67
1586	3274	76132	3.80	88.48
1597	2645	78777	3.07	91.55
1608	2277	81054	2.65	94.20
1621	1766	82820	2.05	96.25
1639	1553	84373	1.80	98.06
1667	1673	86046	1.94	100.00

Figure D.5.43. 2024 TELPAS Grade 11 Speaking
Frequency Distribution of Scale Scores



	Freq	Cum Freq	Pct	Cum Pct
1415	13400	13400	19.64	19.64
1441	2226	15626	3.26	22.90
1456	1490	17116	2.18	25.09
1465	1442	18558	2.11	27.20
1471	1103	19661	1.62	28.82
1477	1014	20675	1.49	30.31
1481	944	21619	1.38	31.69
1485	1034	22653	1.52	33.21
1489	1087	23740	1.59	34.80
1493	1168	24908	1.71	36.51
1496	1301	26209	1.91	38.42
1500	1330	27539	1.95	40.37
1504	1562	29101	2.29	42.66
1508	1722	30823	2.52	45.18
1512	2079	32902	3.05	48.23
1518	2357	35259	3.45	51.68
1525	3216	38475	4.71	56.40
1534	3905	42380	5.72	62.12
1546	6291	48671	9.22	71.34
1561	4447	53118	6.52	77.86
1574	3438	56556	5.04	82.90
1586	2756	59312	4.04	86.94
1597	2224	61536	3.26	90.20
1608	1929	63465	2.83	93.03
1621	1644	65109	2.41	95.44
1639	1484	66593	2.18	97.61
1667	1628	68221	2.39	100.00

Figure D.5.44. 2024 TELPAS Grade 12 Speaking
Frequency Distribution of Scale Scores



	Freq	Cum Freq	Pct	Cum Pct
1415	11409	11409	22.32	22.32
1441	1835	13244	3.59	25.91
1456	1200	14444	2.35	28.26
1465	1129	15573	2.21	30.46
1471	855	16428	1.67	32.14
1477	833	17261	1.63	33.77
1481	760	18021	1.49	35.25
1485	810	18831	1.58	36.84
1489	884	19715	1.73	38.57
1493	955	20670	1.87	40.43
1496	975	21645	1.91	42.34
1500	1050	22695	2.05	44.40
1504	1215	23910	2.38	46.77
1508	1321	25231	2.58	49.36
1512	1567	26798	3.07	52.42
1518	1812	28610	3.54	55.97
1525	2387	30997	4.67	60.64
1534	2741	33738	5.36	66.00
1546	4340	38078	8.49	74.49
1561	2999	41077	5.87	80.35
1574	2322	43399	4.54	84.90
1586	1874	45273	3.67	88.56
1597	1497	46770	2.93	91.49
1608	1275	48045	2.49	93.98
1621	1102	49147	2.16	96.14
1639	913	50060	1.79	97.93
1667	1060	51120	2.07	100.00

2024 TELPAS
Distribution of Proficiency Levels

Table D.6.1. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades K–1

		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**
K	2024	101808	44	29	17	10	101556	52	25	15	8	101597	65	17	11	7	101553	67	18	9	6	101408	50	29	13	7
	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
1	2024	107163	23	33	25	18	106961	32	31	21	16	106941	44	26	16	14	106868	48	27	14	11	106742	30	35	20	15
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.2. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - 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	2024	105120	9	31	33	27	105116	36	50	10	4	105180	39	40	16	6	105175	43	39	15	2	105074	21	54	22	3
	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
3	2024	105763	4	17	27	52	105749	23	48	19	10	105803	30	31	18	21	105796	35	41	19	4	105695	14	44	33	10
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.3. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - 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	2024	105941	18	26	31	26	105929	20	33	40	7	106014	12	34	27	27	106010	27	37	28	8	105904	13	36	37	14
	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
5	2024	106555	13	19	29	38	106546	19	31	39	11	106609	8	25	25	42	106605	19	31	34	15	106500	10	28	38	23
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.4. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 6–7

		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**
6	2024	105561	6	28	36	29	105559	13	46	36	5	105582	18	27	28	28	105578	18	36	38	8	105440	8	34	44	14
	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
7	2024	104907	5	26	35	34	104902	17	46	32	5	104994	16	24	27	33	104990	16	32	41	12	104764	8	32	44	16
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.5. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - 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	2024	101474	5	22	32	40	101469	18	44	33	5	101563	7	35	28	31	101557	19	45	31	5	101321	6	36	44	14
	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
9	2024	109473	15	23	40	22	109471	35	21	36	8	109615	9	36	26	29	109608	23	43	29	5	108903	13	39	36	13
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.6. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grades 10–11

			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**
10	2024	86288	12	22	41	24	86287	33	21	38	8	86475	14	40	26	20	86468	24	29	29	19	85866	13	38	34	16
	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
11	2024	68599	10	22	43	25	68598	32	20	38	10	68714	12	39	27	22	68711	21	29	30	21	68193	10	38	34	17
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

Table D.6.7. Distribution of TELPAS Proficiency Levels by Administration Since 2012 - Grade 12

			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**
12	2024	51161	10	24	43	23	51159	35	21	36	9	51248	12	43	26	19	51249	21	31	30	18	50806	10	43	34	14
	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

Notes:

B = Beginning; I = Intermediate; A = Advanced; H = Advanced High.

Comparisons in performance are only appropriate across certain years due to TELPAS redesigns. For more information and examples of comparison periods for each domain, refer to "Evidence Based on Consequences of Testing" in the Validity section of Chapter 6, "Texas English Language Proficiency Assessment System."

TECHNICAL DIGEST 2023–2024

Appendix E

TELPAS Alternate Statistical Tables and Figures

List of Tables

Table E.1.1. 2024 TELPAS Alternate Classification Consistency and Accuracy

Table E.2.1. 2024 TELPAS Alternate Scale Score Correlations

Table E.3.1. 2024 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

Table E.4.1. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Total Group

Table E.4.2. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Female

Table E.4.3. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Male

Table E.5.1. 2024 TELPAS Alternate Scale Score Descriptive Statistics

List of Figures

Figure E.5.1. 2024 TELPAS Alternate Listening Frequency Distribution of Scale Scores

Figure E.5.2. 2024 TELPAS Alternate Speaking Frequency Distribution of Scale Scores

Figure E.5.3. 2024 TELPAS Alternate Reading Frequency Distribution of Scale Scores

Figure E.5.4. 2024 TELPAS Alternate Writing Frequency Distribution of Scale Scores

2024 TELPAS Alternate Classification Consistency and Accuracy

Table E.1.1. 2024 TELPAS Alternate Classification Consistency and Accuracy

Domain	Decision Consistency[1]	Decision Accuracy[2]
Listening	79.1	85.1
Speaking	83.6	88.3
Reading	80.8	86.2
Writing	82.3	87.3

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.

2024 TELPAS Alternate
Scale Score Correlations

Table E.2.1. 2024 TELPAS Alternate Scale Score Correlations

TELPAS Alternate Domain	TELPAS Alternate Domain	N	Correlation
Listening	Speaking	10,606	0.89
Listening	Reading	10,606	0.89
Listening	Writing	10,606	0.86
Speaking	Reading	10,606	0.84
Speaking	Writing	10,606	0.83
Reading	Writing	10,606	0.92

2024 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

Table E.3.1. 2024 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

	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.

2024 TELPAS Alternate
Mean P-Values, Raw Score Summary, and Reliability

Table E.4.1. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Total Group

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
Listening	OVERALL TEST	50	10,606	29.58	11.87	0.97	1.89	48.95
	Understand spoken words and language structures	20	10,606	11.71	5.05	0.94	1.20	48.17
	Basic understanding of spoken English	30	10,606	17.87	7.03	0.96	1.40	49.47
Speaking	OVERALL TEST	50	10,606	27.99	12.90	0.98	1.61	44.97
	Provide and summarize information	35	10,606	19.49	9.16	0.98	1.36	44.62
	Share opinions and analyze information	15	10,606	8.49	3.84	0.95	0.85	45.77
Reading	OVERALL TEST	50	10,606	25.31	11.75	0.98	1.72	38.27
	Understand words and language structures	25	10,606	13.33	6.24	0.97	1.13	41.64
	Basic understanding of a variety of texts written in English	15	10,606	7.28	3.53	0.94	0.86	35.68
	Analyze and evaluate information and ideas in a variety of texts written in English	10	10,606	4.70	2.38	0.94	0.59	33.70
Writing	OVERALL TEST	50	10,606	22.33	10.97	0.98	1.63	30.83
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	10,606	9.86	4.81	0.96	1.00	36.63
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	10,606	12.47	6.42	0.97	1.10	26.97

Notes:

1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

Table E.4.2. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Female

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
Listening	OVERALL TEST	50	3,457	29.63	11.76	0.97	1.91	49.07
	Understand spoken words and language structures	20	3,457	11.68	5.03	0.94	1.20	47.97
	Basic understanding of spoken English	30	3,457	17.95	6.94	0.96	1.42	49.80
Speaking	OVERALL TEST	50	3,457	28.43	12.78	0.98	1.65	46.07
	Provide and summarize information	35	3,457	19.79	9.07	0.98	1.39	45.68
	Share opinions and analyze information	15	3,457	8.64	3.82	0.95	0.87	46.98
Reading	OVERALL TEST	50	3,457	25.00	11.67	0.98	1.69	37.51
	Understand words and language structures	25	3,457	13.07	6.17	0.97	1.11	40.35
	Basic understanding of a variety of texts written in English	15	3,457	7.24	3.50	0.94	0.86	35.30
	Analyze and evaluate information and ideas in a variety of texts written in English	10	3,457	4.70	2.38	0.94	0.60	33.70
Writing	OVERALL TEST	50	3,457	22.35	10.98	0.98	1.62	30.87
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	3,457	9.81	4.80	0.96	0.99	36.28
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	3,457	12.54	6.43	0.97	1.11	27.26

Notes:

1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

Table E.4.3. 2024 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Male

Subject	Reporting Category	Score Point[1]	N	Mean	SD	Alpha	SEM	Mean P-Value[2]
Listening	OVERALL TEST	50	7,149	29.56	11.93	0.98	1.88	48.89
	Understand spoken words and language structures	20	7,149	11.72	5.07	0.94	1.19	48.26
	Basic understanding of spoken English	30	7,149	17.83	7.07	0.96	1.39	49.31
Speaking	OVERALL TEST	50	7,149	27.77	12.96	0.98	1.60	44.43
	Provide and summarize information	35	7,149	19.35	9.21	0.98	1.34	44.11
	Share opinions and analyze information	15	7,149	8.42	3.85	0.95	0.85	45.19
Reading	OVERALL TEST	50	7,149	25.45	11.79	0.98	1.73	38.64
	Understand words and language structures	25	7,149	13.45	6.27	0.97	1.14	42.27
	Basic understanding of a variety of texts written in English	15	7,149	7.30	3.55	0.94	0.87	35.86
	Analyze and evaluate information and ideas in a variety of texts written in English	10	7,149	4.70	2.38	0.94	0.59	33.71
Writing	OVERALL TEST	50	7,149	22.33	10.96	0.98	1.64	30.81
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	7,149	9.89	4.82	0.96	1.01	36.80
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	7,149	12.44	6.42	0.97	1.09	26.83

Notes:

1. Total number of Score Points
2. Mean of percent correct (0–100%) for the items

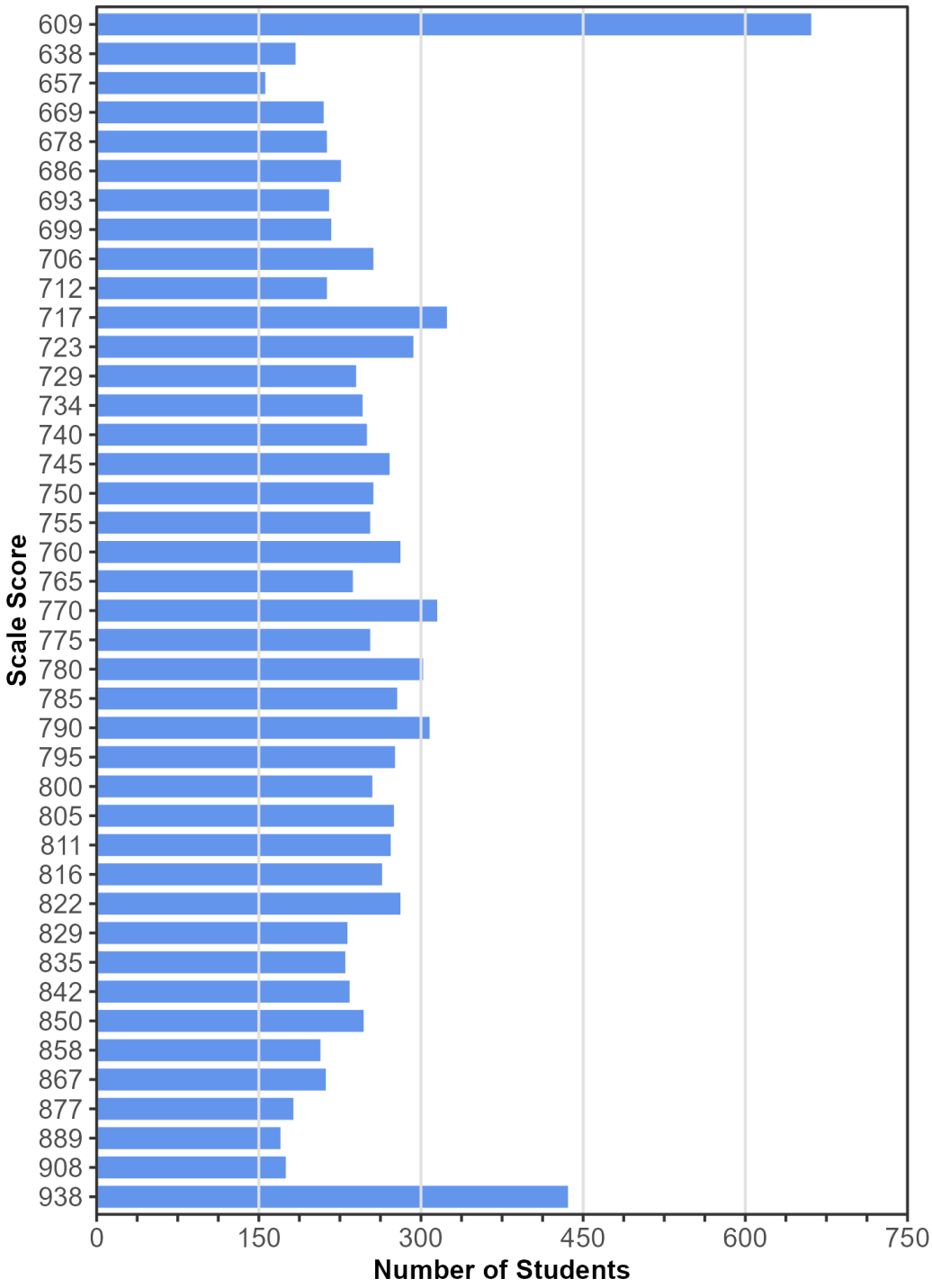
2024 TELPAS Alternate
Scale Score Descriptive Statistics

Table E.5.1. 2024 TELPAS Alternate Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
TELPAS Alternate Listening	10,606	767.38	770	609	329	99	78.79	6208.08	0.0	-0.25
TELPAS Alternate Speaking	10,606	756.93	756	600	352	137	97.63	9531.03	0.1	-0.77
TELPAS Alternate Reading	10,606	744.10	743	601	347	126	87.65	7683.35	0.1	-0.60
TELPAS Alternate Writing	10,606	727.91	728	612	343	122	82.61	6823.73	0.3	-0.48

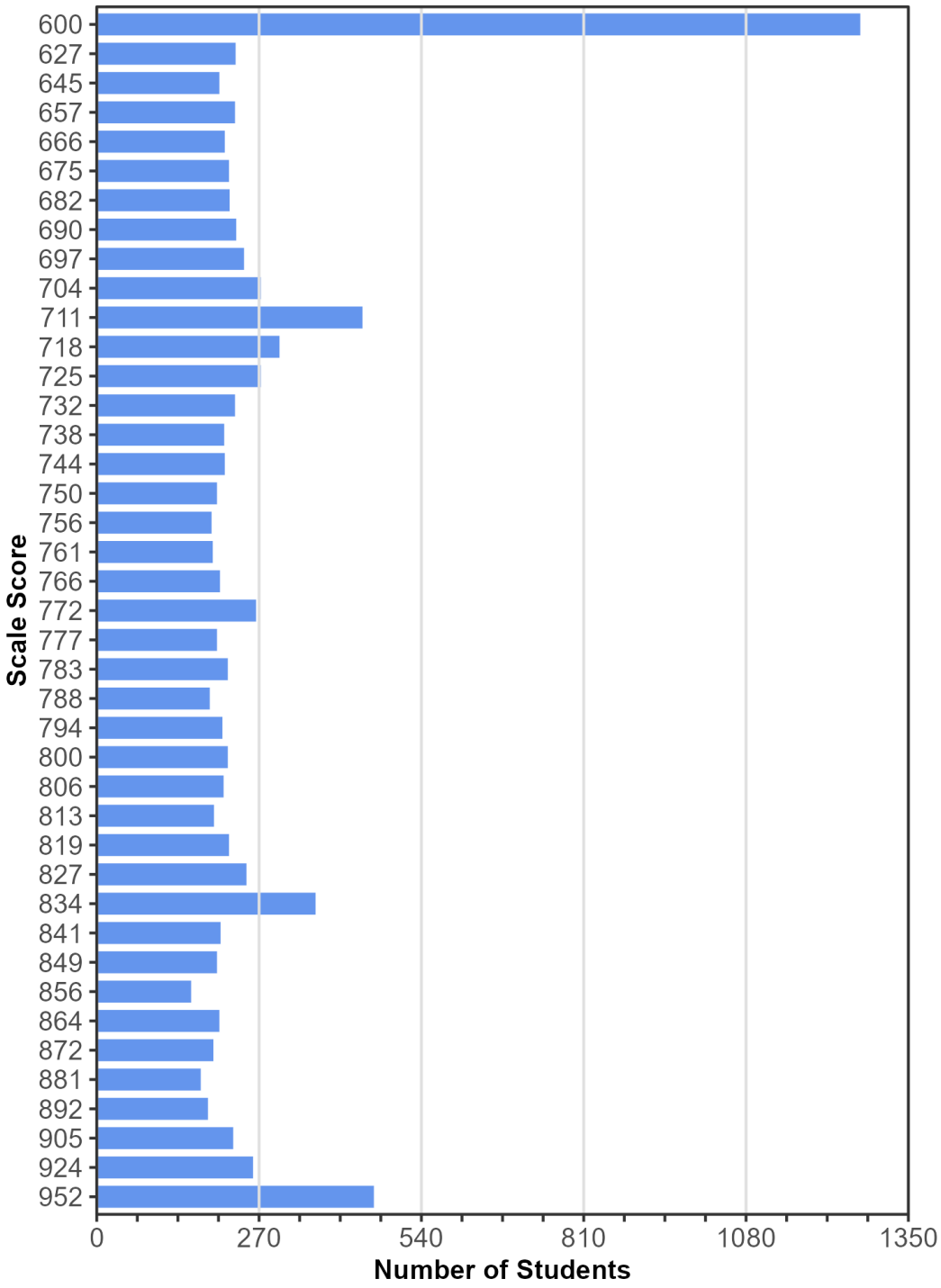
2024 TELPAS Alternate
Frequency Distribution of Scale Scores

Figure E.5.1. 2024 TELPAS Alternate Listening
Frequency Distribution of Scale Scores



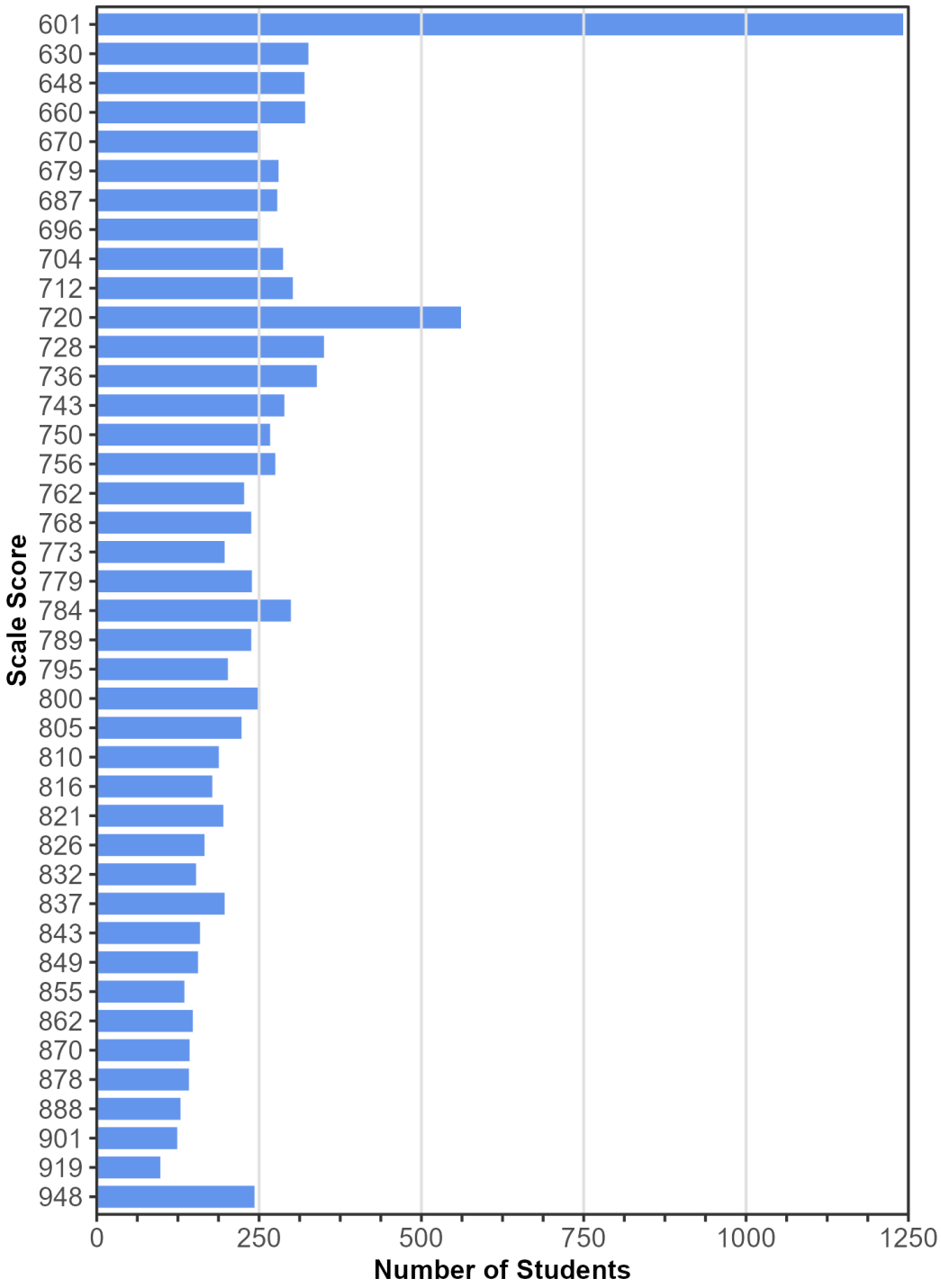
Scale Score	Freq	Cum Freq	Pct	Cum Pct
609	661	661	6.23	6.23
638	184	845	1.73	7.97
657	156	1001	1.47	9.44
669	210	1211	1.98	11.42
678	213	1424	2.01	13.43
686	226	1650	2.13	15.56
693	215	1865	2.03	17.58
699	217	2082	2.05	19.63
706	256	2338	2.41	22.04
712	213	2551	2.01	24.05
717	324	2875	3.05	27.11
723	293	3168	2.76	29.87
729	240	3408	2.26	32.13
734	246	3654	2.32	34.45
740	250	3904	2.36	36.81
745	271	4175	2.56	39.36
750	256	4431	2.41	41.78
755	253	4684	2.39	44.16
760	281	4965	2.65	46.81
765	237	5202	2.23	49.05
770	315	5517	2.97	52.02
775	253	5770	2.39	54.40
780	302	6072	2.85	57.25
785	278	6350	2.62	59.87
790	308	6658	2.90	62.78
795	276	6934	2.60	65.38
800	255	7189	2.40	67.78
805	275	7464	2.59	70.38
811	272	7736	2.56	72.94
816	264	8000	2.49	75.43
822	281	8281	2.65	78.08
829	232	8513	2.19	80.27
835	230	8743	2.17	82.43
842	234	8977	2.21	84.64
850	247	9224	2.33	86.97
858	207	9431	1.95	88.92
867	212	9643	2.00	90.92
877	182	9825	1.72	92.64
889	170	9995	1.60	94.24
908	175	10170	1.65	95.89
938	436	10606	4.11	100.00

Figure E.5.2. 2024 TELPAS Alternate Speaking
Frequency Distribution of Scale Scores



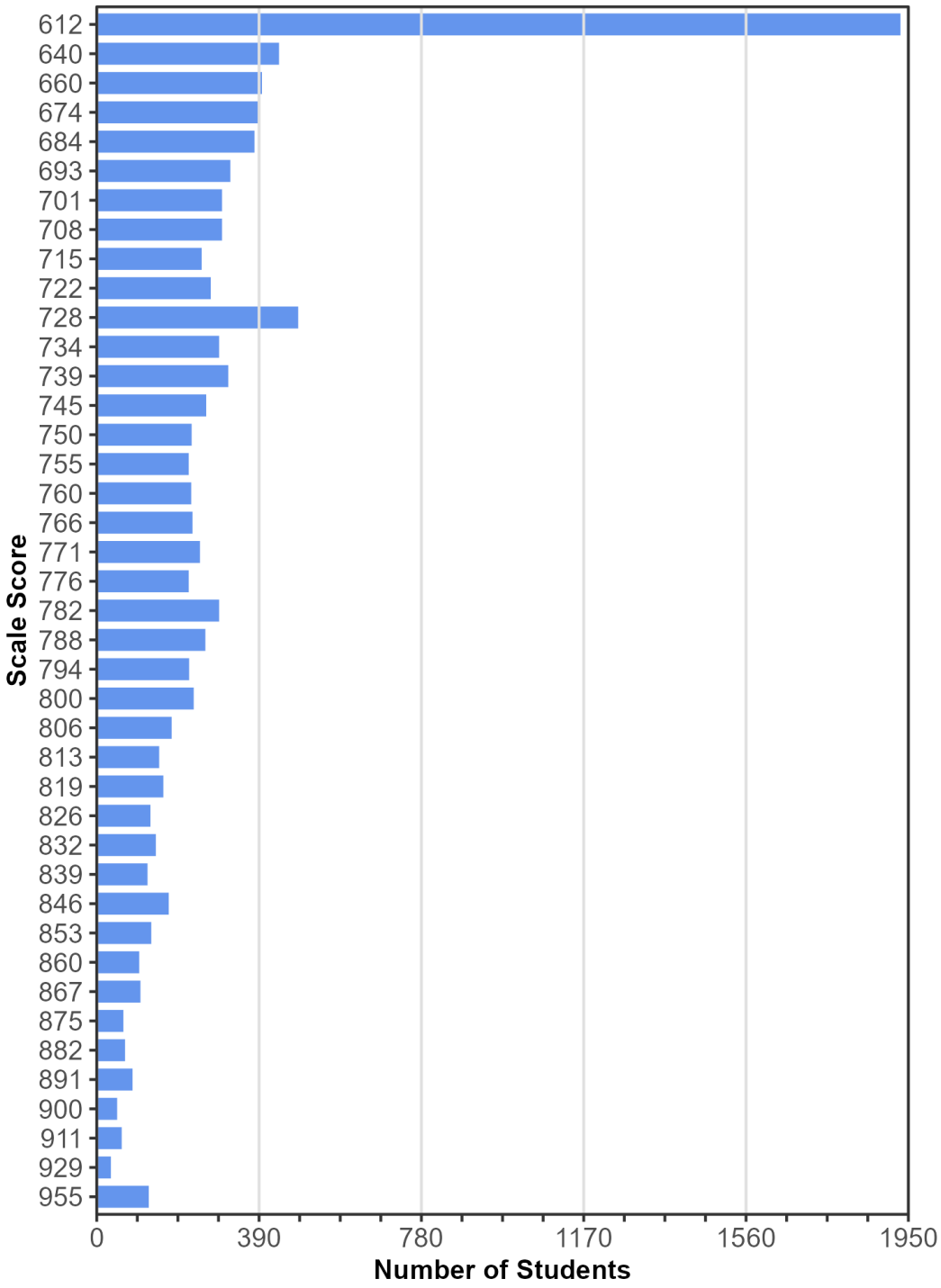
	Freq	Cum Freq	Pct	Cum Pct
600	1270	1270	11.97	11.97
627	231	1501	2.18	14.15
645	204	1705	1.92	16.08
657	230	1935	2.17	18.24
666	213	2148	2.01	20.25
675	220	2368	2.07	22.33
682	221	2589	2.08	24.41
690	232	2821	2.19	26.60
697	245	3066	2.31	28.91
704	273	3339	2.57	31.48
711	442	3781	4.17	35.65
718	304	4085	2.87	38.52
725	273	4358	2.57	41.09
732	230	4588	2.17	43.26
738	212	4800	2.00	45.26
744	213	5013	2.01	47.27
750	200	5213	1.89	49.15
756	191	5404	1.80	50.95
761	193	5597	1.82	52.77
766	205	5802	1.93	54.70
772	265	6067	2.50	57.20
777	200	6267	1.89	59.09
783	218	6485	2.06	61.14
788	188	6673	1.77	62.92
794	209	6882	1.97	64.89
800	218	7100	2.06	66.94
806	211	7311	1.99	68.93
813	195	7506	1.84	70.77
819	220	7726	2.07	72.85
827	249	7975	2.35	75.19
834	364	8339	3.43	78.63
841	206	8545	1.94	80.57
849	200	8745	1.89	82.45
856	157	8902	1.48	83.93
864	204	9106	1.92	85.86
872	194	9300	1.83	87.69
881	173	9473	1.63	89.32
892	185	9658	1.74	91.06
905	227	9885	2.14	93.20
924	260	10145	2.45	95.65
952	461	10606	4.35	100.00

Figure E.5.3. 2024 TELPAS Alternate Reading
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
601	1242	1242	11.71	11.71
630	326	1568	3.07	14.78
648	320	1888	3.02	17.80
660	321	2209	3.03	20.83
670	252	2461	2.38	23.20
679	280	2741	2.64	25.84
687	278	3019	2.62	28.47
696	250	3269	2.36	30.82
704	287	3556	2.71	33.53
712	302	3858	2.85	36.38
720	561	4419	5.29	41.67
728	350	4769	3.30	44.97
736	339	5108	3.20	48.16
743	289	5397	2.72	50.89
750	267	5664	2.52	53.40
756	275	5939	2.59	56.00
762	227	6166	2.14	58.14
768	238	6404	2.24	60.38
773	197	6601	1.86	62.24
779	239	6840	2.25	64.49
784	299	7139	2.82	67.31
789	238	7377	2.24	69.55
795	202	7579	1.90	71.46
800	250	7829	2.36	73.82
805	223	8052	2.10	75.92
810	188	8240	1.77	77.69
816	178	8418	1.68	79.37
821	195	8613	1.84	81.21
826	166	8779	1.57	82.77
832	153	8932	1.44	84.22
837	197	9129	1.86	86.07
843	159	9288	1.50	87.57
849	156	9444	1.47	89.04
855	135	9579	1.27	90.32
862	148	9727	1.40	91.71
870	143	9870	1.35	93.06
878	142	10012	1.34	94.40
888	129	10141	1.22	95.62
901	124	10265	1.17	96.78
919	98	10363	0.92	97.71
948	243	10606	2.29	100.00

Figure E.5.4. 2024 TELPAS Alternate Writing
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
612	1931	1931	18.21	18.21
640	438	2369	4.13	22.34
660	397	2766	3.74	26.08
674	390	3156	3.68	29.76
684	379	3535	3.57	33.33
693	321	3856	3.03	36.36
701	301	4157	2.84	39.19
708	301	4458	2.84	42.03
715	252	4710	2.38	44.41
722	274	4984	2.58	46.99
728	484	5468	4.56	51.56
734	294	5762	2.77	54.33
739	316	6078	2.98	57.31
745	263	6341	2.48	59.79
750	228	6569	2.15	61.94
755	221	6790	2.08	64.02
760	227	7017	2.14	66.16
766	230	7247	2.17	68.33
771	248	7495	2.34	70.67
776	221	7716	2.08	72.75
782	294	8010	2.77	75.52
788	261	8271	2.46	77.98
794	222	8493	2.09	80.08
800	233	8726	2.20	82.27
806	180	8906	1.70	83.97
813	150	9056	1.41	85.39
819	160	9216	1.51	86.89
826	129	9345	1.22	88.11
832	142	9487	1.34	89.45
839	122	9609	1.15	90.60
846	173	9782	1.63	92.23
853	131	9913	1.24	93.47
860	102	10015	0.96	94.43
867	105	10120	0.99	95.42
875	64	10184	0.60	96.02
882	68	10252	0.64	96.66
891	86	10338	0.81	97.47
900	49	10387	0.46	97.94
911	60	10447	0.57	98.50
929	34	10481	0.32	98.82
955	125	10606	1.18	100.00