TECHNICAL DIGEST

2024–2025 School Year

Information about the Texas Assessment Program can be found on the Student Assessment website.

Copyright © 2025, Texas Education Agency. All rights reserved.

Table of Contents

Chapter 1 Historical Overview of Assessment in Texas 1-1
Overview 1-3
Assessment Timeline 1-3
Changes to the Assessment Program Over Time 1-11
Chapter 2 Building a High-Quality Assessment System 2-1
Overview 2-3
Test Development Activities 2-3
Groups Involved 2-4
Item Development and Review 2-6
Pilot Testing 2-8
Field Testing and Data Review 2-8
Test Construction 2-10
Accommodations 2-11
Security 2-12
Performance Tasks 2-13
Quality Control Procedures for Scoring and Reporting 2-18
Chapter 3 Standard Technical Processes 3-1
Overview 3-3
Performance Standards 3-3
Measures of Student Progress 3-6
Item Analysis 3-8
Scaling 3-10
Equating 3-15
Reliability 3-20
Validity 3-24
Sampling 3-28
Chapter 4 State of Texas Assessments of Academic Readiness 4-1
Overview 4-3
Testing Requirements 4-4
Test Development 4-4
Accommodations 4-6
Training 4-7
Test Administration 4-7
Performance Standards 4-9
Scores and Reports 4-11
Measures of Student Progress 4-13
Scaling 4-14

Equating 4-16
Reliability 4-17
Validity 4-17
Sampling 4-27
Test Results 4-28
Chapter 5 STAAR Alternate 2 5-1
Overview 5-3
Participation Requirements 5-3
Test Development 5-4
Accommodations 5-5
Training 5-5
Test Administration 5-6
Performance Standards 5-7
Scores and Reports 5-9
Measures of Student Progress 5-10
Scaling 5-11
Equating 5-13
Reliability 5-13
Validity 5-13
Sampling 5-17
Test Results 5-17
Chapter 6 Texas English Language Proficiency Assessment System 6-1
Overview 6-3
Testing Requirements 6-3
Test Development 6-3
Accommodations 6-4
Training 6-5
Test Administration 6-6
Proficiency Standards 6-8
Scores and Reports 6-10
Scaling 6-14
Equating 6-16
Reliability 6-17
Validity 6-20
Sampling 6-33
Test Results 6-33
Chapter 7 TELPAS Alternate 7-1
Overview 7-3

Participation Requirements 7-3
Test Development 7-4
Training 7-4
Test Administration 7-5
Proficiency Standards 7-6
Scores and Reports 7-7
Scaling 7-10
Reliability 7-12
Validity 7-12
Sampling 7-19
Test Results 7-19
Appendix A Resources, Glossary, and References A-1
Resources A-3
Glossary of Statistical Terminology A-4
References A-6
Appendix B STAAR Statistical Tables and Figures B-1
Appendix C STAAR Alternate 2 Statistical Tables and Figures C-1
Appendix D TELPAS Statistical Tables and Figures D-1
Appendix E TELPAS Alternate Statistical Tables and Figures E-1

TECHNICAL DIGEST 2024–2025

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.

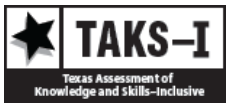
State-Developed Alternative Assessment II (SDAA II)

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

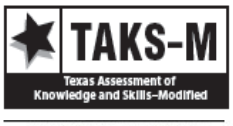
LAT administrations of TAKS in reading for grades 3–8 and in ELA for grade 10 were implemented in spring 2007 for eligible EB students.

—2007–2008

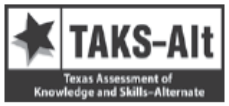
LAT administrations of TAKS in science for grades 5, 8, and 10 were implemented in spring 2008 for eligible EB students.

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

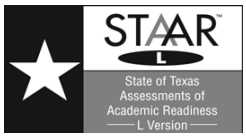
EOC assessments in Algebra II and English I were first administered on a voluntary basis.

—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.

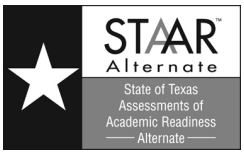
In compliance with SSI, satisfactory performance on STAAR grades 5 and 8 mathematics and reading were requirements for promotion to the next grade level through spring 2021.

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.

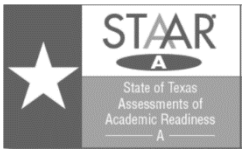
STAAR Modified EOC assessments in Algebra I, Geometry, English I reading, English I writing, English II reading, English II writing, Biology, World Geography, and World History were added to the testing program.

—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 suspension of STAAR at the start of the COVID-19 pandemic, the Texas Education Agency (TEA) launched optional end-of-year (EOY)

assessments that school systems and parents could choose to administer 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 from 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 2024–2025

Chapter 2

Building a High-Quality Assessment System

Chapter 2

Overview

Test Development Activities

Groups Involved

Item Development and Review

Pilot Testing

Field Testing and Data Review

Test Construction

Accommodations

Security

Performance Tasks

Quality Control Procedures for Scoring and Reporting

Overview

TEA implements comprehensive processes and collaborative efforts to build and maintain a high-quality assessment system. These processes incorporate detailed test development procedures and stringent quality control measures designed to ensure that all assessments, including STAAR, STAAR Spanish, STAAR Alternate 2, TELPAS, and TELPAS Alternate, provide equitable and accurate measures of student learning.

A cornerstone of this system is the extensive involvement of Texas educators, including teachers, curriculum specialists, and administrators, who play a vital role in every phase of the test development process. Their input, alongside expertise from the TEA Assessment Development and Student Assessment Divisions and contractors, is critical in shaping assessments that are aligned with the state's curriculum standards and best instructional practices.

This chapter details the various components of assessment development and maintenance. Program-specific development processes are covered in subsequent chapters.

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	Drafts of the reporting categories and TEKS or ELPS student expectations to be assessed are 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.

Step	Process
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.
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	The *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 ensure all students can access the items. 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, applicability to all students, 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 problematic issues or topics.

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 problematic issues or topics. 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 problematic issues or topics. 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 any issues found. 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.

Category	Guidelines
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.
Avoidance of Problematic Issues or Topics	• The item or passage avoids needless reference to topics that are extremely controversial, upsetting, or inappropriate for students. • The item or passage addresses sensitive topics in a careful, fair, and balanced way.

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 stand-alone 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 stand-alone 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 five 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 differential item functioning (DIF), 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 DIF. DIF flags indicate the potential for bias. Content specialists review flagged items, 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 web page. 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 web page.

TELPAS is based on the ELPS. TELPAS assessed curriculum documents and blueprints can be found on the TELPAS Resources web page.

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.

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 web page.

Accommodations

TEA defines accommodations as changes to materials or procedures that enable students to participate meaningfully in learning and testing. 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 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.

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 web page 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 incidents 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 serious testing violation 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 incidents 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 incidents,
- a system to review and respond to each reported testing incident, 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 0 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 web page.

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 web page. 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 web pages.

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 web page.

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 web page.

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 web pages, 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 online scoring and reporting (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. 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 updated student results to the 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 2023–24 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 web page.

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.

Data and Report Processing

TEA undertakes an extensive QC process to confirm score validity and 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 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, refer to chapter 3, "Standard Technical Processes."

TECHNICAL DIGEST 2024–2025

Chapter 3

Standard Technical Processes

Chapter 3

Overview

Performance Standards

Measures of Student Progress

Item Analysis

Scaling

Equating

Reliability

Validity

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
- measures of student progress
- item analysis
- scaling
- equating
- reliability
- validity
- 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

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 2022–23 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 web page 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 web page 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 web page 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 web page for more detailed information.

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 2012–13 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," in chapter 5, “STAAR Alternate 2,” and on the Progress Measures web page.

Item Analysis

Statistical analyses are conducted on student performance data to gauge the level of difficulty of the item, to examine the degree to which the item appropriately distinguishes between students of different performance or proficiency levels, and to 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, which is a common odds ratio, 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 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 $\theta - \delta_{i0} = 0$.

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

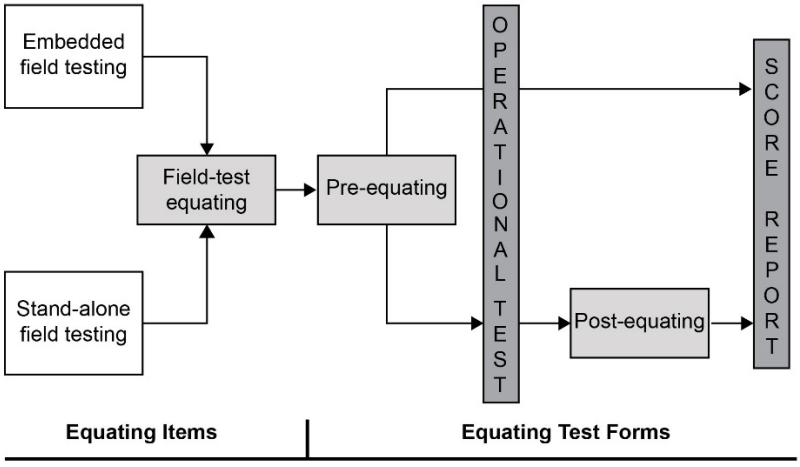
The following are the three types of equating used in the item and test development process:

- 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 have post-equated forms.

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 reflect circumstances, such as current events, that may change how students interact with the item.

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

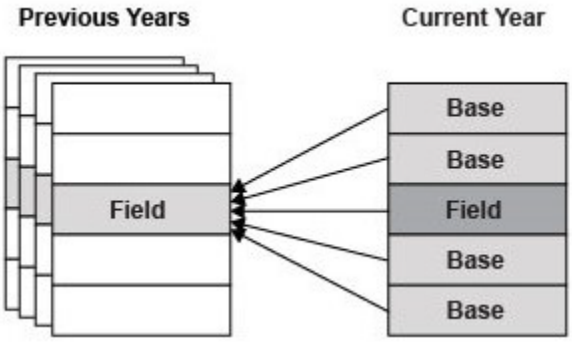
equating constant is calculated and used to transform the Rasch difficulty obtained from the current calibration to the underlying Rasch difficulty scale established by the original test form in the scaling year. This equating constant is defined as:

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

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.

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:

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

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:

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

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 raw test 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:

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

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:

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

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:

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

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 web page.

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.

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 2024–2025

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 2011–12 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 2022–23 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 web page.

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 2011–12 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 fair 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 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 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 the version for the new school year 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 numbers of STAAR assessments that were administered this school year are shown in Table 4.1.

Table 4.1. STAAR Assessments Administered in 2024–25

Assessment	Assessments Administered
STAAR Grade 3 Mathematics	376,511
STAAR Grade 3 RLA	363,696
STAAR Spanish Grade 3 Mathematics	21,323
STAAR Spanish Grade 3 RLA	34,189
STAAR Grade 4 Mathematics	381,087
STAAR Grade 4 RLA	371,514
STAAR Spanish Grade 4 Mathematics	15,390
STAAR Spanish Grade 4 RLA	26,383
STAAR Grade 5 Mathematics	383,302
STAAR Grade 5 RLA	378,299
STAAR Grade 5 Science	383,161
STAAR Spanish Grade 5 Mathematics	12,213
STAAR Spanish Grade 5 RLA	19,421

Assessment	Assessments Administered
STAAR Spanish Grade 5 Science	14,396
STAAR Grade 6 Mathematics	390,074
STAAR Grade 6 RLA	396,501
STAAR Grade 7 Mathematics	299,218
STAAR Grade 7 RLA	398,260
STAAR Grade 8 Mathematics	372,216
STAAR Grade 8 RLA	397,840
STAAR Grade 8 Science	393,057
STAAR Grade 8 Social Studies	401,348
STAAR Algebra I	569,650
STAAR English I	473,821
STAAR English II	658,642
STAAR Biology	593,802
STAAR U.S. History	434,768

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 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 a wide variety of stakeholder groups from across the state 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 2012–15 (Scale Score)	Approaches Grade Level 2016–22 (Scale Score)	Approaches Grade Level (Scale Score)	Meets Grade Level (Scale Score)	Masters Grade Level (Scale Score)
Algebra I	3489	3541	3550	4000	4345
English I	3775	3775	3775	4000	4606
English II	3766	3775	3775	4000	4734
Biology	3516	3550	3550	4000	4531
U.S. History	3486	3536	3550	4000	4424

Refer to the STAAR standard-setting technical reports, which are available on the Assessment Reports and Studies web page, 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 the calibration sample used for the given assessment during the current year's spring calibration.

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 2018–19 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 web page.

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, TEA provides student report cards, student labels, campus rosters, summary reports, and reporting data files.

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. 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 consider 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 consider 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 2012–13 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 web page. STAAR progress measure results are available for grades 4–8

mathematics, grades 4–8 RLA, Spanish grades 4–5 mathematics, Spanish grades 4–5 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 2013–14. 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 web page. 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 and 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 represent an accurate measurement of 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 are followed while developing STAAR assessments for the Texas Assessment Program, supporting 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 appropriateness. Two distinct types of educator committees meet regularly 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. Content validation committees do not edit or revise items prior to field testing, but 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 by item type. Across STAAR, 15 types of response interactions are available to measure student learning. For more information about the item types, refer to the STAAR Resources web page.

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 for each 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 per 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 to determine 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 the ASE and human scores are computed and evaluated on a 10 percent random sample (refer to Tables 4.7 and 4.9). For STAAR Spanish, the rates between the two human scorers are computed on a 100 percent reliability read for ECRs and a 25 percent reliability read for SCRs (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 web page.

Table 4.7. Summary of Score Agreement (Reliability) for Spring 2025 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	30,016	69.8%	27.7%	97.5%	0.74
Conventions	30,016	74.3%	23.7%	98.0%	0.71
STAAR Grade 4 RLA					
Ideas	33,028	70.3%	28.3%	98.5%	0.79
Conventions	33,028	72.6%	25.8%	98.4%	0.76
STAAR Grade 5 RLA					
Ideas	34,236	68.8%	29.6%	98.3%	0.76
Conventions	34,236	68.4%	30.6%	99.0%	0.69
STAAR Grade 6 RLA					
Ideas	36,177	69.2%	29.1%	98.2%	0.82
Conventions	36,177	70.0%	28.6%	98.6%	0.75
STAAR Grade 7 RLA					
Ideas	36,568	67.3%	30.6%	97.9%	0.81
Conventions	36,568	70.0%	29.0%	99.0%	0.73

ECR Trait	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Grade 8 RLA					
Ideas	36,314	63.9%	33.3%	97.2%	0.80
Conventions	36,314	72.4%	25.2%	97.6%	0.75
STAAR English I					
Ideas	41,526	68.5%	29.7%	98.2%	0.83
Conventions	41,526	73.2%	25.6%	98.8%	0.77
STAAR English II					
Ideas	39,940	66.3%	31.4%	97.7%	0.80
Conventions	39,940	72.1%	24.6%	96.7%	0.76

Table 4.8. Summary of Score Agreement (Reliability) for Spring 2025 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	27,258	70.3%	27.8%	98.1%	0.70
Conventions	27,258	74.3%	23.2%	97.5%	0.66
STAAR Spanish Grade 4 RLA					
Ideas	23,094	69.0%	29.3%	98.3%	0.66
Conventions	23,094	71.2%	27.8%	99.0%	0.69
STAAR Spanish Grade 5 RLA					
Ideas	18,270	71.3%	27.3%	98.6%	0.78
Conventions	18,270	72.4%	26.1%	98.5%	0.75

Table 4.9. Summary of Scorer Agreement (Reliability) for Spring 2025 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	34,574	77.1%	22.6%	99.7%	0.77
Writing SCR	33,930	95.9%	4.1%	100%	0.92
STAAR Grade 4 RLA					
Reading SCR	36,429	82.6%	17.0%	99.6%	0.85
Writing SCR	36,343	94.8%	5.2%	100%	0.89
STAAR Grade 5 RLA					
Reading SCR	37,374	78.9%	20.8%	99.7%	0.78
Writing SCR	37,349	91.6%	8.4%	100%	0.83
STAAR Grade 6 RLA					
Reading SCR	39,132	79.5%	20.3%	99.8%	0.81
Writing SCR	38,814	90.7%	9.3%	100%	0.81
STAAR Grade 7 RLA					
Reading SCR	38,954	82.6%	17.1%	99.7%	0.82
Writing SCR	38,568	93.6%	6.4%	100%	0.86
STAAR Grade 8 RLA					
Reading SCR	38,648	81.5%	17.4%	98.9%	0.83
Writing SCR	38,530	95.9%	4.1%	100%	0.92
STAAR English I					
Reading SCR	45,595	79.7%	19.9%	99.6%	0.82
Writing SCR	45,453	90.1%	9.9%	100%	0.80
STAAR English II					
Reading SCR	43,977	91.1%	8.9%	100%	0.81
Writing SCR	44,676	78.9%	21.1%	100%	0.81
STAAR Grade 5 Science					
SCR 1	37,311	94.3%	5.5%	99.8%	0.95
STAAR Grade 8 Science					
SCR 1	37,400	81.4%	18.2%	99.6%	0.80
SCR 2	36,618	90.7%	9.2%	99.9%	0.81

SCR Indicator	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
		Exact	Adjacent	Exact + Adjacent	
STAAR Biology					
SCR 1	39,969	95.1%	4.8%	99.9%	0.96
SCR 2	40,099	93.5%	6.4%	99.9%	0.92
STAAR Grade 8 Social Studies					
SCR 1	37,395	88.3%	10.9%	99.2%	0.91
SCR 2	34,776	80.2%	19.4%	99.6%	0.81
STAAR U.S. History					
SCR 1	35,707	87.0%	12.8%	99.8%	0.90
SCR 2	34,393	89.6%	10.1%	99.7%	0.90

Table 4.10. Summary of Scorer Agreement (Reliability) for Spring 2025 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					
Writing SCR	7,970	91.5%	8.5%	100%	0.83
STAAR Spanish Grade 4 RLA					
Writing SCR	6,394	91.4%	8.6%	100%	0.83
STAAR Spanish Grade 5 RLA					
Writing SCR	4,673	94.8%	5.2%	100%	0.87
STAAR Spanish Grade 5 Science					
Writing SCR	3,485	93.9%	5.4%	99.2%	0.93

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 2025 STAAR RLA ECRs

ECR Trait	Exact Agreement Rate
STAAR Grade 3 RLA	
Ideas	90.8
Conventions	87.8
STAAR Grade 4 RLA	
Ideas	85.9
Conventions	91.5
STAAR Grade 5 RLA	
Ideas	85.9
Conventions	86.3
STAAR Grade 6 RLA	
Ideas	88.6
Conventions	89.5
STAAR Grade 7 RLA	
Ideas	87.9
Conventions	92.1
STAAR Grade 8 RLA	
Ideas	82.0
Conventions	89.8
STAAR English I	
Ideas	85.7
Conventions	85.4
STAAR English II	
Ideas	85.4
Conventions	88.7
STAAR Spanish Grade 3 RLA	
Ideas	86.8
Conventions	78.6
STAAR Spanish Grade 4 RLA	
Ideas	90.9
Conventions	83.5
STAAR Spanish Grade 5 RLA	
Ideas	83.2
Conventions	82.2

Table 4.12. Summary of Validity Results for Spring 2025 STAAR SCRs

SCR Indicator	Exact Agreement Rate
STAAR Grade 3 RLA, Reading SCR	94.6
STAAR Grade 3 RLA, Writing SCR	98.7
STAAR Grade 4 RLA, Reading SCR	93.7
STAAR Grade 4 RLA, Writing SCR	97.6
STAAR Grade 5 RLA, Reading SCR	91.0
STAAR Grade 5 RLA, Writing SCR	95.7
STAAR Grade 6 RLA, Reading SCR	92.9
STAAR Grade 6 RLA, Writing SCR	97.1
STAAR Grade 7 RLA, Reading SCR	94.7
STAAR Grade 7 RLA, Writing SCR	98.4
STAAR Grade 8 RLA, Reading SCR	96.8
STAAR Grade 8 RLA, Writing SCR	98.6
STAAR English I, Reading SCR	92.9
STAAR English I, Writing SCR	96.1
STAAR English II, Reading SCR	92.3
STAAR English II, Writing SCR	96.4
STAAR Grade 5 Science, SCR	97.3
STAAR Grade 8 Science, SCR 1	88.0
STAAR Grade 8 Science, SCR 2	93.1
STAAR Biology, SCR 1	96.4
STAAR Biology, SCR 2	98.2
STAAR Grade 8 Social Studies, SCR 1	91.3
STAAR Grade 8 Social Studies, SCR 2	90.8
STAAR U.S. History, SCR 1	92.2
STAAR U.S. History, SCR 2	92.3
STAAR Spanish Grade 3 RLA, Writing SCR	97.2
STAAR Spanish Grade 4 RLA, Writing SCR	99.3
STAAR Spanish Grade 5 RLA, Writing SCR	98.8
STAAR Spanish Grade 5 Science, SCR 1	97.0

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 dichotomously and polytomously scored 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 Consequences of Testing

Documenting the intended and unintended consequences of administering an assessment is another method of providing validity evidence. 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 testing, 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 STAAR performance levels, as discussed 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 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 that measure similar knowledge and skills, such as the ACT or SAT.

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 in the Research Portal.

TECHNICAL DIGEST 2024–2025

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 2014–15 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.

STAAR Alternate 2 reading and writing assessments were combined into an RLA assessment for each grade to mirror the redesigned STAAR, and 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 2023–24 to clarify the scope of students who are assessed with STAAR Alternate 2. The STAAR Alternate 2 participation requirements, available in English and Spanish on the STAAR Alternate 2 Resources web page, 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—in the school year (August through July) during which 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 web page.

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 fair 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 curriculum framework tools and vertical alignments for each content area. 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, students with disabilities must 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 web page.

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 web page, 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 numbers of STAAR Alternate 2 assessments that were administered this school year are shown in Table 5.1.

Table 5.1. STAAR Alternate 2 Assessments Administered in 2024–25

Assessment	Assessments Administered
Grade 3 Mathematics	7,510
Grade 3 RLA	7,516
Grade 4 Mathematics	6,666
Grade 4 RLA	6,668
Grade 5 Mathematics	6,324
Grade 5 RLA	6,321
Grade 5 Science	6,315
Grade 6 Mathematics	5,877
Grade 6 RLA	5,874
Grade 7 Mathematics	5,385
Grade 7 RLA	5,385
Grade 8 Mathematics	4,900
Grade 8 RLA	4,906
Grade 8 Science	4,901
Grade 8 Social Studies	4,904
Algebra I	4,940
English I	4,924
English II	4,731
Biology	4,995
U.S. History	4,344

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

The following are the performance level categories and policy definitions for STAAR Alternate 2.

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 the content measured for the grade or course assessed. Students exhibit the ability to use higher-level thinking and more complex skills, which includes making inferences and comparisons and solving multistep 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 the content measured for the grade or course assessed. 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 can recognize 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 a wide variety of stakeholder groups from across the state 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
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

Assessment	Level II: Satisfactory	Level III: Accomplished
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 web page, 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 web page.

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. 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 consider 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 consider 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 2015–16 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 web page. 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 subjects 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 web page.

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
Science	Grade 5	43.8943	291.6601
	Grade 8	38.5892	298.4950
	Biology	38.2614	293.1129

Content Area	Grade/Course	*A*	*B*
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 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 represent an accurate measurement of 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 2015–16, 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 appropriateness. 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

Documenting the intended and unintended consequences of administering an assessment is another method of providing validity evidence. 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 testing, 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 STAAR Alternate 2 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 for a given test title 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 STAAR Alternate 2 administration. Pass rates for STAAR Alternate 2 are available in the Research Portal.

TECHNICAL DIGEST 2024–2025

Chapter 6

Texas English Language Proficiency Assessment System

Overview

Testing Requirements

Test Development

Accommodations

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.

Testing 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, including 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, are assessed annually in English language proficiency until they are determined to be proficient by meeting the EB reclassification criteria.

In rare circumstances it might be necessary for the ARD committee, in collaboration 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 fair 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 web page. 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.

Accommodations

The goal of TELPAS 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, TELPAS accommodations are designed to allow all students to demonstrate their English language proficiency 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 TELPAS are also designed to be the same as or similar to those accommodations commonly used during classroom instruction.

Accommodations policies for TELPAS 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 TELPAS. 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 at the campus level (e.g., 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 TELPAS 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 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 TELPAS.

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 web page.

Test Administration

TELPAS is administered once a year, in the spring, during a six-week testing window. The numbers of TELPAS assessments that were administered this school year are shown in Table 6.2.

Table 6.2. TELPAS Assessments Administered in 2024–25

Grade	Listening	Speaking	Reading	Writing
Kindergarten	107,156	107,084	106,954	107,004
Grade 1	110,053	109,906	109,810	109,939
Grade 2	103,734	103,702	103,783	103,785
Grade 3	107,643	107,627	107,726	107,724
Grade 4	103,861	103,857	103,936	103,931
Grade 5	99,748	99,741	99,843	99,837
Grade 6	92,128	92,125	92,194	92,191
Grade 7	97,088	97,085	97,111	97,109
Grade 8	93,795	93,790	93,839	93,837
Grade 9	103,255	103,249	103,442	103,439
Grade 10	88,385	88,383	88,592	88,589
Grade 11	67,640	67,639	67,725	67,726
Grade 12	57,743	57,743	57,910	57,910
Total	1,232,229	1,231,931	1,232,865	1,233,021

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, or 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 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

Domain	Grade or Grade Band	Beginning	Intermediate	Advanced	Advanced High
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 web page, 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 6.4 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 collaboration 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. District-level summary reports provide the number and percentage of students who progressed one, two, or three proficiency levels. The 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 also be used when evaluating instruction or programs that require average-score or year-to-year comparisons. TELPAS is 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

Domain	Grade or Grade Band	*A*	*B*
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 TELPAS administration is provided in Appendix D. Refer to chapter 3, “Standard Technical Processes,” for detailed information about reliability. 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. A median of 0.00 seconds means that at least half of the students did not speak at all.

Table 6.7. TELPAS Speaking Median Response Time per 45-Second Item in 2024–25

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	6.9	6.5	5.1	5.8
Advanced	14.4	12.1	14.6	12.0
Advanced High	26.5	22.6	33.7	25.3

Table 6.8. TELPAS Speaking Median Response Time per 90-Second Item in 2024–25

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	19.5	18.7	9.4	12.6
Advanced	37.4	38.1	25.8	28.5
Advanced High	58.2	61.0	57.1	53.5

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 (which is 0.25 for all domains, as shown in Table 6.4), X_i represents the domain score of each domain, $\sigma_{X_i}^2$ is the variance of the domain score, $\rho_{X_i X_i'}$ is the internal consistency of each domain, Z is the composite score, and σ_Z^2 is the variance of 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* = 103,493)	Listening	2.767	0.938	0.772	0.921
	Speaking	1.797	0.784	0.827	
	Reading	1.879	0.877	0.762	
	Writing	1.676	0.807	0.837	
Grade 3 (*n* = 107,411)	Listening	3.270	0.892	0.796	0.932
	Speaking	2.100	0.898	0.846	
	Reading	2.276	1.094	0.822	
	Writing	1.887	0.827	0.810	

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 4 (*n* = 103,663)	Listening	2.548	1.043	0.801	0.931
	Speaking	2.300	0.867	0.828	
	Reading	2.623	0.985	0.797	
	Writing	2.069	0.875	0.833	
Grade 5 (*n* = 99,593)	Listening	2.889	1.065	0.820	0.938
	Speaking	2.378	0.910	0.834	
	Reading	2.927	1.008	0.820	
	Writing	2.358	0.942	0.840	
Grade 6 (*n* = 91,958)	Listening	2.706	0.935	0.752	0.924
	Speaking	2.289	0.772	0.819	
	Reading	2.490	1.049	0.812	
	Writing	2.232	0.866	0.812	
Grade 7 (*n* = 96,861)	Listening	2.923	0.965	0.773	0.932
	Speaking	2.263	0.792	0.824	
	Reading	2.748	1.082	0.837	
	Writing	2.457	0.920	0.824	
Grade 8 (*n* = 93,539)	Listening	3.041	0.963	0.787	0.923
	Speaking	2.249	0.818	0.837	
	Reading	2.739	0.929	0.802	
	Writing	2.163	0.786	0.764	
Grade 9 (*n* = 102,601)	Listening	2.628	0.970	0.809	0.939
	Speaking	2.117	0.999	0.904	
	Reading	2.747	0.970	0.830	
	Writing	2.131	0.819	0.776	
Grade 10 (*n* = 87,748)	Listening	2.809	0.963	0.805	0.937
	Speaking	2.180	1.017	0.905	
	Reading	2.493	0.941	0.812	
	Writing	2.304	0.917	0.785	
Grade 11 (*n* = 66,899)	Listening	2.822	0.943	0.795	0.933
	Speaking	2.151	1.018	0.908	
	Reading	2.547	0.944	0.810	
	Writing	2.325	0.910	0.775	

Grade	Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Grade 12 (*n* = 57,318)	Listening	2.754	0.943	0.794	0.930
	Speaking	2.040	1.012	0.903	
	Reading	2.495	0.927	0.807	
	Writing	2.294	0.907	0.774	

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

Validity

Validity refers to the extent to which test scores represent an accurate measurement of 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.

Relationship to the Statewide Curriculum

TELPAS assesses English language proficiency 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.

The ELPS are second language acquisition curriculum standards that support the ability of EB students to learn the academic English they need for meaningful engagement in grade-appropriate subject-area instruction. The second language acquisition knowledge and skills in the ELPS are the means for helping EB students learn English simultaneously with academic subject matter. TELPAS assesses the ELPS, which district personnel are required to implement as an integral part of each foundation and enrichment subject of the TEKS state-required curriculum. The ELPS are implemented as part of ongoing subject-area instruction and, as such, are integrally linked with the subject-area TEKS. The ELPS help teachers meet the language and subject-matter needs of EB students simultaneously. The ELPS and TELPAS

work together to provide formative and summative assessment opportunities that support teaching and learning.

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.

Educator Input

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 appropriateness. 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.

Test Developer Input

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 developed 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
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 ratings be based on the ability of students to use English in a variety of content areas.

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 TELPAS reflects accurate measurement of the construct.

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 as 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, then approximately 25 percent are routed for human scoring: a random sample of approximately

10 percent to monitor performance and another approximately 15 percent to verify scoring of unusual or difficult-to-score responses. For any response routed for human scoring, the human score is used to determine a student's score. TELPAS grades 2 and 3 writing constructed responses and sentence-rewrite items are scored completely by humans.

Score reliability for TELPAS 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 2025 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	102,230	87.3%	12.5%	99.8%	0.8068
Item 2	102,156	85.3%	14.5%	99.8%	0.7890
Item 3	102,254	84.9%	14.9%	99.8%	0.8050
Grade 3					
Item 1	107,164	80.0%	19.6%	99.6%	0.7892
Item 2	107,255	76.3%	23.2%	99.5%	0.8110
Item 3	107,021	78.6%	21.0%	99.6%	0.7852

Table 6.12. Summary of Scorer Agreement (Reliability) for 2025 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	19,405	73.6%	26.0%	99.6%	0.8559
	Usage	19,405	72.9%	26.7%	99.7%	0.8204
	Vocabulary	19,405	75.0%	24.7%	99.7%	0.8400
Item 2	Completeness	19,210	76.4%	23.4%	99.8%	0.8269
	Usage	19,210	75.9%	23.8%	99.7%	0.8011
	Vocabulary	19,210	76.0%	23.8%	99.8%	0.8151

Item	Dimension	Number of Responses	Agreement Rate after Two Scores			Quadratic Weighted Kappa
			Exact	Adjacent	Exact + Adjacent	
Item 3	Completeness	19,854	73.0%	26.7%	99.7%	0.8474
	Usage	19,854	71.5%	28.1%	99.6%	0.8086
	Vocabulary	19,854	72.2%	27.5%	99.7%	0.8237
Grades 6–7						
Item 1	Completeness	18,584	83.2%	16.7%	99.9%	0.8790
	Usage	18,584	79.3%	20.6%	99.9%	0.8304
	Vocabulary	18,584	78.1%	21.8%	99.9%	0.8395
Item 2	Completeness	18,370	76.2%	23.7%	99.9%	0.8380
	Usage	18,370	75.0%	24.8%	99.7%	0.8086
	Vocabulary	18,370	76.1%	23.7%	99.8%	0.8313
Item 3	Completeness	18,524	74.3%	25.5%	99.8%	0.8540
	Usage	18,524	73.8%	26.0%	99.8%	0.8357
	Vocabulary	18,524	74.2%	25.7%	99.8%	0.8508
Grades 8–9						
Item 1	Completeness	19,178	76.2%	23.5%	99.7%	0.8072
	Usage	19,178	73.4%	26.4%	99.7%	0.7620
	Vocabulary	19,178	73.2%	26.6%	99.7%	0.7693
Item 2	Completeness	18,466	74.8%	24.8%	99.6%	0.8441
	Usage	18,466	71.4%	28.1%	99.5%	0.8007
	Vocabulary	18,466	73.4%	26.3%	99.7%	0.8255
Item 3	Completeness	18,446	79.1%	20.7%	99.8%	0.8616
	Usage	18,446	75.3%	24.4%	99.7%	0.8192
	Vocabulary	18,446	77.6%	22.2%	99.8%	0.8443
Grades 10–12						
Item 1	Completeness	19,974	78.7%	21.0%	99.7%	0.8842
	Usage	19,974	77.5%	22.2%	99.7%	0.8630
	Vocabulary	19,974	76.9%	22.8%	99.7%	0.8665
Item 2	Completeness	20,299	80.3%	19.6%	99.9%	0.8980
	Usage	20,299	77.5%	22.4%	99.9%	0.8726
	Vocabulary	20,299	79.4%	20.5%	99.9%	0.8888
Item 3	Completeness	18,844	75.6%	24.2%	99.8%	0.8658
	Usage	18,844	72.5%	27.3%	99.8%	0.8340
	Vocabulary	18,844	73.4%	26.3%	99.8%	0.8485

Table 6.13. Summary of Scorer Agreement (Reliability) for 2025 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	5,130	98.7%	1.3%	100%	0.9671
Grade 3	5,365	95.0%	5.0%	100%	0.8972
Grades 4–5	19,077	97.7%	2.3%	100%	0.9519
Grades 6–7	18,449	95.7%	4.3%	100%	0.8953
Grades 8–9	19,023	97.3%	2.7%	100%	0.9421
Grades 10–12	20,574	96.3%	3.7%	100%	0.8964

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	31,981	0.85
Grades 4–5	31,298	0.83
Grades 6–8	42,386	0.86
Grades 9–12	47,826	0.92

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 dichotomously and polytomously scored 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 the 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 Relationships to Other Variables

Another method TEA uses to provide validity evidence for TELPAS is analyzing 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)

Table 6.15 shows the TELPAS reading data for students who participated in both STAAR English RLA and TELPAS. Table 6.16 shows the TELPAS reading data for students who participated in both STAAR Spanish RLA and TELPAS. Table 6.17 shows the TELPAS writing data for students who participated in both STAAR English RLA and TELPAS. Table 6.18 shows the TELPAS writing data for students who participated in both STAAR Spanish RLA and TELPAS. In Tables 6.15 and 6.17, the number of students in the STAAR English I and English II counts includes both students taking the test for the first time and those retaking it.

Table 6.15. TELPAS Reading Proficiency Compared to STAAR English RLA Performance

Grade or Course	TELPAS Reading Proficiency Level	Number of Students	STAAR RLA Average Scale Score	STAAR RLA Passing Rate
Grade 3	Beginning	15,621	1271	19%
	Intermediate	22,512	1372	62%
	Advanced	16,404	1481	94%
	Advanced High	18,636	1599	100%
Grade 4	Beginning	6,170	1329	12%
	Intermediate	24,870	1406	47%
	Advanced	23,419	1528	91%
	Advanced High	22,612	1642	100%
Grade 5	Beginning	4,317	1333	3%
	Intermediate	18,218	1404	20%
	Advanced	22,023	1522	67%
	Advanced High	35,338	1666	97%
Grade 6	Beginning	19,239	1422	6%
	Intermediate	27,132	1498	35%
	Advanced	25,473	1608	82%
	Advanced High	19,430	1718	98%

Grade or Course	TELPAS Reading Proficiency Level	Number of Students	STAAR RLA Average Scale Score	STAAR RLA Passing Rate
Grade 7	Beginning	15,913	1413	4%
	Intermediate	23,001	1491	23%
	Advanced	25,849	1601	68%
	Advanced High	31,161	1713	95%
Grade 8	Beginning	6,893	1471	5%
	Intermediate	34,139	1543	32%
	Advanced	27,298	1660	82%
	Advanced High	23,572	1769	98%
English I	Beginning	17,648	3376	5%
	Intermediate	55,563	3527	17%
	Advanced	28,072	3840	61%
	Advanced High	27,518	4217	95%
English II	Beginning	19,146	3399	8%
	Intermediate	51,892	3603	29%
	Advanced	26,291	3930	75%
	Advanced High	16,135	4245	96%

Table 6.16. TELPAS Reading Proficiency Compared to STAAR Spanish RLA Performance

Grade	TELPAS Reading Proficiency Level	Number of Students	STAAR Spanish RLA Average Scale Score	STAAR Spanish RLA Passing Rate
Grade 3	Beginning	16,433	1276	35%
	Intermediate	11,195	1349	59%
	Advanced	3,490	1422	82%
	Advanced High	2,403	1508	96%
Grade 4	Beginning	6,696	1341	30%
	Intermediate	12,695	1393	46%
	Advanced	4,171	1497	76%
	Advanced High	2,317	1599	96%
Grade 5	Beginning	4,654	1405	43%
	Intermediate	8,587	1458	59%
	Advanced	3,192	1554	85%
	Advanced High	2,642	1645	97%

Table 6.17. TELPAS Writing Proficiency Compared to STAAR English RLA Performance

Grade or Course	TELPAS Writing Proficiency Level	Number of Students	STAAR RLA Average Scale Score	STAAR Passing Rate
Grade 3	Beginning	18,437	1273	19%
	Intermediate	33,863	1432	79%
	Advanced	17,447	1559	98%
	Advanced High	3,426	1653	100%
Grade 4	Beginning	15,105	1344	19%
	Intermediate	32,809	1485	76%
	Advanced	23,888	1598	97%
	Advanced High	5,269	1688	100%
Grade 5	Beginning	10,195	1349	7%
	Intermediate	27,508	1485	50%
	Advanced	31,353	1612	88%
	Advanced High	10,840	1713	99%
Grade 6	Beginning	19,796	1419	5%
	Intermediate	36,230	1535	50%
	Advanced	29,133	1646	88%
	Advanced High	6,115	1748	98%
Grade 7	Beginning	16,546	1411	4%
	Intermediate	30,517	1531	39%
	Advanced	36,880	1644	79%
	Advanced High	11,981	1739	96%
Grade 8	Beginning	18,496	1483	10%
	Intermediate	43,688	1623	63%
	Advanced	26,264	1725	92%
	Advanced High	3,454	1797	98%
English I	Beginning	35,818	3397	6%
	Intermediate	57,472	3712	39%
	Advanced	31,372	4036	78%
	Advanced High	4,139	4303	94%

Grade or Course	TELPAS Writing Proficiency Level	Number of Students	STAAR RLA Average Scale Score	STAAR Passing Rate
English II	Beginning	31,011	3434	11%
	Intermediate	42,038	3678	38%
	Advanced	32,489	3977	76%
	Advanced High	7,926	4231	95%

Table 6.18. TELPAS Writing Proficiency Compared to STAAR Spanish RLA Performance

Grade	TELPAS Writing Proficiency Level	Number of Students	STAAR Spanish RLA Average Scale Score	STAAR Spanish RLA Passing Rate
Grade 3	Beginning	20,840	1283	38%
	Intermediate	9,845	1394	72%
	Advanced	2,495	1474	91%
	Advanced High	341	1545	96%
Grade 4	Beginning	15,138	1359	36%
	Intermediate	7,815	1467	68%
	Advanced	2,522	1549	87%
	Advanced High	404	1632	96%
Grade 5	Beginning	10,652	1427	49%
	Intermediate	5,379	1532	79%
	Advanced	2,400	1604	92%
	Advanced High	644	1674	98%

As shown in Tables 6.15–6.18, the average STAAR scale score increases as students' TELPAS proficiency levels increase within a grade. In addition, passing rates on STAAR tend to increase as TELPAS proficiency levels increase within each grade. These patterns demonstrate a clear positive relationship between TELPAS proficiency and STAAR performance, supporting the validity of TELPAS as a measure of English language proficiency.

Evidence Based on Consequences of Testing

Documenting the intended and unintended consequences of administering an assessment is another method of providing validity evidence. The collection of consequential validity evidence typically occurs after a program has been in place for some time.

Given the important stakes associated with TELPAS testing, 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 TELPAS proficiency levels, as discussed in

the Proficiency Standards section of this chapter. 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 2012–22 and 2023–present. For grades 2–12 listening and speaking, results can be compared for the periods of 2012–17 and 2018–present. For reading, results can be compared for the periods of 2012–13, 2014–17, and 2018–present. Direct comparisons across these distinct periods are not appropriate; however, if historical trends hold, 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 TELPAS administration. The percentages of students in each proficiency level for all four domains as well as for the composite rating are available in the Research Portal.

TECHNICAL DIGEST 2024–2025

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, including students classified as emergent bilingual (EB)/English learner (EL) in PEIMS whose parents have declined bilingual or ESL program services, are assessed annually in English language proficiency until they are determined to be proficient by meeting the EB reclassification criteria.

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 collaboration with the LPAC, must review carefully and consider annually. The TELPAS Alternate participation requirements, available in English and Spanish on the TELPAS Alternate Resources web page, 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 collaboration 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 web page.

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 fair 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 web page. 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 numbers of TELPAS Alternate assessments that were administered this school year are shown in Table 7.1.

Table 7.1. TELPAS Alternate Assessments Administered in 2024–25

Grade	Assessments Administered
Grade 2	2,079
Grade 3	1,954
Grade 4	1,536
Grade 5	1,421
Grade 6	1,110
Grade 7	888
Grade 8	743
Grade 9	748
Grade 10	854
Grade 11	451

Grade	Assessments Administered
Grade 12	388

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 web page, 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 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. District-level summary reports provide the number and percentage of students who progressed one, two, three, or four proficiency levels. The 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 also be used when evaluating instruction or programs that require average-score or year-to-year comparisons. TELPAS Alternate is 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 7.5.

Table 7.5. 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 TELPAS Alternate administration is provided in Appendix E. Refer to chapter 3, "Standard Technical Processes," for detailed information about reliability. 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 (which is 0.25 for all domains, as shown in Table 7.3), X_i represents the domain score of each domain, $\sigma_{X_i}^2$ is the variance of the domain score, $\rho_{X_i X_i'}$ is the internal consistency of each domain, Z is the composite score, and σ_Z^2 is the variance of 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 7.6.

Table 7.6. TELPAS Alternate Composite Score Reliability Estimates

Domain	Mean	Standard Deviation	Internal Consistency	Composite Reliability
Listening	2.878	1.290	0.942	0.985
Speaking	2.743	1.392	0.950	
Reading	2.569	1.390	0.949	
Writing	2.295	1.292	0.948	

Refer to chapter 3, "Standard Technical Processes," for detailed information about reliability.

Validity

Validity refers to the extent to which test scores represent an accurate measurement of 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 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.

Relationship to the Statewide Curriculum

TELPAS Alternate measures student performance in direct alignment with the English language acquisition skills defined by the ELPS that are part of the TEKS state-required curriculum. The ELPS outline the instruction that EB students must receive to support their ability to develop academic English language proficiency.

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. Test 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 design and purpose of these two assessments, one would not expect EB students to perform at the same level on both. 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 passing rate for each STAAR Alternate 2 RLA assessment was calculated based on the TELPAS Alternate proficiency level achieved.

Table 7.7 shows the TELPAS Alternate reading data for students who participated in both STAAR Alternate 2 and TELPAS Alternate. Table 7.8 shows the TELPAS Alternate writing data for students who participated in both assessments.

Table 7.7. TELPAS Alternate 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	697	309	68%
	Imitation	434	335	85%
	Early Independence	442	348	95%
	Developing Independence	238	353	97%
	Basic Fluency	105	370	98%
Grade 4	Awareness	443	307	69%
	Imitation	332	329	83%
	Early Independence	355	347	94%
	Developing Independence	246	358	97%
	Basic Fluency	130	369	98%
Grade 5	Awareness	339	304	62%
	Imitation	286	325	78%
	Early Independence	304	339	91%
	Developing Independence	278	350	93%
	Basic Fluency	197	366	97%
Grade 6	Awareness	281	314	71%
	Imitation	189	336	89%
	Early Independence	202	343	89%
	Developing Independence	213	356	98%
	Basic Fluency	206	370	96%
Grade 7	Awareness	226	318	74%
	Imitation	164	342	92%
	Early Independence	161	354	94%
	Developing Independence	155	360	97%
	Basic Fluency	171	378	96%
Grade 8	Awareness	173	320	78%
	Imitation	120	335	97%
	Early Independence	145	345	90%
	Developing Independence	148	353	97%
	Basic Fluency	143	379	99%

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	326	78%
	Imitation	131	348	95%
	Early Independence	110	355	99%
	Developing Independence	132	364	95%
	Basic Fluency	138	386	99%
English II	Awareness	144	323	75%
	Imitation	108	351	95%
	Early Independence	109	354	95%
	Developing Independence	102	370	99%
	Basic Fluency	116	386	99%

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

Table 7.8. TELPAS Alternate 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	832	312	70%
	Imitation	509	341	91%
	Early Independence	403	352	95%
	Developing Independence	134	365	99%
	Basic Fluency	38	376	95%
Grade 4	Awareness	544	309	70%
	Imitation	376	336	88%
	Early Independence	349	352	96%
	Developing Independence	165	364	99%
	Basic Fluency	72	376	96%
Grade 5	Awareness	442	305	64%
	Imitation	310	334	85%
	Early Independence	329	343	92%
	Developing Independence	228	358	96%
	Basic Fluency	95	374	97%

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	333	317	74%
	Imitation	213	338	89%
	Early Independence	231	349	94%
	Developing Independence	191	363	97%
	Basic Fluency	123	372	97%
Grade 7	Awareness	250	318	74%
	Imitation	196	345	94%
	Early Independence	183	361	97%
	Developing Independence	129	359	95%
	Basic Fluency	119	384	98%
Grade 8	Awareness	204	321	76%
	Imitation	132	336	87%
	Early Independence	173	354	94%
	Developing Independence	122	361	98%
	Basic Fluency	98	379	98%
English I	Awareness	199	326	79%
	Imitation	168	350	95%
	Early Independence	124	360	98%
	Developing Independence	108	368	94%
	Basic Fluency	104	389	99%
English II	Awareness	174	326	78%
	Imitation	101	351	93%
	Early Independence	124	358	99%
	Developing Independence	81	378	99%
	Basic Fluency	99	386	99%

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

As shown in Tables 7.7 and 7.8, students with higher TELPAS Alternate proficiency levels generally achieve higher STAAR Alternate 2 scale scores and passing rates within each grade, with only minor exceptions to this overall pattern. These results demonstrate a clear positive relationship between TELPAS Alternate proficiency and STAAR Alternate 2 performance, supporting the validity of TELPAS Alternate as a measure of English language proficiency for EB students.

Evidence Based on Consequences of Testing

Documenting the intended and unintended consequences of administering an assessment is another method of providing validity evidence. 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 testing, 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 TELPAS Alternate proficiency levels, as discussed in the Proficiency Standards section of this chapter. 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 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 TELPAS Alternate administration. The percentages of students in each proficiency level for all four domains as well as for the composite rating are available in the Research Portal.

TECHNICAL DIGEST 2024–2025

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 2024–2025

Appendix B

STAAR Statistical Tables and Figures

List of Tables

Table B.1.1. Spring 2025 STAAR Grades 3–8 Classification Consistency and Accuracy

Table B.1.2. Spring 2025 STAAR Spanish Grades 3–5 Classification Consistency and Accuracy

Table B.1.3. Spring 2025 STAAR EOC Assessments Classification Consistency and Accuracy

Table B.2.1. Spring 2025 STAAR Grades 3–8 Scale Score Correlations

Table B.2.2. Spring 2025 STAAR Spanish Grades 3–5 Scale Score Correlations

Table B.2.3. Spring 2025 STAAR EOC Assessments Scale Score Correlations

Table B.3.1. Spring 2025 STAAR and STAAR Spanish Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.2. Spring 2025 STAAR Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table B.3.3. Spring 2025 STAAR Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.4. Spring 2025 STAAR Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.5. Spring 2025 STAAR and STAAR Spanish Grade 5 and STAAR Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

Table B.3.6. Spring 2025 STAAR Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

Table B.3.7. Spring 2025 STAAR Spanish Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table B.3.8. Spring 2025 STAAR Algebra I Conditional Standard Error of Measurement for Scale Scores

Table B.3.9. Spring 2025 STAAR English I and English II Conditional Standard Error of Measurement for Scale Scores

Table B.3.10. Spring 2025 STAAR Biology Conditional Standard Error of Measurement for Scale Scores

Table B.3.11. Spring 2025 STAAR U.S. History Conditional Standard Error of Measurement for Scale Scores

Table B.4.1. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.2. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.3. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.4. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.5. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.6. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.7. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.8. Spring 2025 STAAR Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.9. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.10. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.11. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.12. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.13. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.14. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.15. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.16. Spring 2025 STAAR Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.17. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.18. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.19. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.20. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.21. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.22. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.23. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.24. Spring 2025 STAAR Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.25. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.26. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.27. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.28. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.29. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.30. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.31. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.32. Spring 2025 STAAR Grade 6 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.33. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.34. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.35. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.36. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.37. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.38. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.39. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.40. Spring 2025 STAAR Grade 7 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.41. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.42. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.43. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.44. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.45. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.46. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.47. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.48. Spring 2025 STAAR Grade 8 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.49. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.50. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.51. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.52. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.53. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.54. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.55. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.56. Spring 2025 STAAR Spanish Grade 3 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.57. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.58. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.59. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.60. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.61. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.62. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.63. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.64. Spring 2025 STAAR Spanish Grade 4 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.65. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.66. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.67. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.68. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.69. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.70. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.71. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.72. Spring 2025 STAAR Spanish Grade 5 Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.73. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.74. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.75. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.76. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.77. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.78. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.79. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.80. Spring 2025 STAAR Algebra I Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.81. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.82. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.83. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.84. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.85. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.86. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.87. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.88. Spring 2025 STAAR English I Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.89. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.90. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.91. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.92. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.93. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.94. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.95. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.96. Spring 2025 STAAR English II Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.97. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.98. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.99. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.100. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.101. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.102. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.103. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.104. Spring 2025 STAAR Biology Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.4.105. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Table B.4.106. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Table B.4.107. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Table B.4.108. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Black or African-American

Table B.4.109. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table B.4.110. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Table B.4.111. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Students Tested with Accommodations

Table B.4.112. Spring 2025 STAAR U.S. History Mean P-Values, Raw Score Summary, and Reliability Students Tested without Accommodations

Table B.5.1. Spring 2025 STAAR Grade 4 Mathematics Progress Measure

Table B.5.2. Spring 2025 STAAR Grade 4 RLA Progress Measure

Table B.5.3. Spring 2025 STAAR Grade 5 Mathematics Progress Measure

Table B.5.4. Spring 2025 STAAR Grade 5 RLA Progress Measure

Table B.5.5. Spring 2025 STAAR Grade 6 Mathematics Progress Measure

Table B.5.6. Spring 2025 STAAR Grade 6 RLA Progress Measure

Table B.5.7. Spring 2025 STAAR Grade 7 Mathematics Progress Measure

Table B.5.8. Spring 2025 STAAR Grade 7 RLA Progress Measure

Table B.5.9. Spring 2025 STAAR Grade 8 Mathematics Progress Measure

Table B.5.10. Spring 2025 STAAR Grade 8 RLA Progress Measure

Table B.5.11. Spring 2025 STAAR Spanish Grade 4 Mathematics Progress Measure

Table B.5.12. Spring 2025 STAAR Spanish Grade 4 RLA Progress Measure

Table B.5.13. Spring 2025 STAAR Spanish Grade 5 Mathematics Progress Measure

Table B.5.14. Spring 2025 STAAR Spanish Grade 5 RLA Progress Measure

Table B.5.15. Spring 2025 STAAR Algebra I Progress Measure

Table B.5.16. Spring 2025 STAAR English I Progress Measure

Table B.5.17. Spring 2025 STAAR English II Progress Measure

Table B.6.1. Spring 2025 STAAR Grades 3–8 and English I On-Track Measure

Table B.7.1. Spring 2025 STAAR Grades 3–8 Scale Score Descriptive Statistics

Table B.7.2. Spring 2025 STAAR Spanish Grades 3–5 Scale Score Descriptive Statistics

Table B.7.3. Spring 2025 STAAR EOC Assessments Scale Score Descriptive Statistics

Table B.8.1. Spring 2025 STAAR Grades 3–8 Mathematics English Learner Performance Measure

Table B.8.2. Spring 2025 STAAR Grades 3–8 RLA English Learner Performance Measure

Table B.8.3. Spring 2025 STAAR Grade 5 and Grade 8 Science English Learner Performance Measure

Table B.8.4. Spring 2025 STAAR Grade 8 Social Studies English Learner Performance Measure

Table B.8.5. Spring 2025 STAAR EOC Assessments English Learner Performance Measure

List of Figures

Figure B.7.1. Spring 2025 STAAR Grade 3 Mathematics Frequency Distribution of Scale Scores
Figure B.7.2. Spring 2025 STAAR Grade 4 Mathematics Frequency Distribution of Scale Scores
Figure B.7.3. Spring 2025 STAAR Grade 5 Mathematics Frequency Distribution of Scale Scores
Figure B.7.4. Spring 2025 STAAR Grade 6 Mathematics Frequency Distribution of Scale Scores
Figure B.7.5. Spring 2025 STAAR Grade 7 Mathematics Frequency Distribution of Scale Scores
Figure B.7.6. Spring 2025 STAAR Grade 8 Mathematics Frequency Distribution of Scale Scores
Figure B.7.7. Spring 2025 STAAR Grade 3 RLA Frequency Distribution of Scale Scores
Figure B.7.8. Spring 2025 STAAR Grade 4 RLA Frequency Distribution of Scale Scores
Figure B.7.9. Spring 2025 STAAR Grade 5 RLA Frequency Distribution of Scale Scores
Figure B.7.10. Spring 2025 STAAR Grade 6 RLA Frequency Distribution of Scale Scores
Figure B.7.11. Spring 2025 STAAR Grade 7 RLA Frequency Distribution of Scale Scores
Figure B.7.12. Spring 2025 STAAR Grade 8 RLA Frequency Distribution of Scale Scores
Figure B.7.13. Spring 2025 STAAR Grade 5 Science Frequency Distribution of Scale Scores
Figure B.7.14. Spring 2025 STAAR Grade 8 Science Frequency Distribution of Scale Scores
Figure B.7.15. Spring 2025 STAAR Grade 8 Social Studies Frequency Distribution of Scale Scores
Figure B.7.16. Spring 2025 STAAR Spanish Grade 3 Mathematics Frequency Distribution of Scale Scores
Figure B.7.17. Spring 2025 STAAR Spanish Grade 4 Mathematics Frequency Distribution of Scale Scores
Figure B.7.18. Spring 2025 STAAR Spanish Grade 5 Mathematics Frequency Distribution of Scale Scores
Figure B.7.19. Spring 2025 STAAR Spanish Grade 3 RLA Frequency Distribution of Scale Scores
Figure B.7.20. Spring 2025 STAAR Spanish Grade 4 RLA Frequency Distribution of Scale Scores
Figure B.7.21. Spring 2025 STAAR Spanish Grade 5 RLA Frequency Distribution of Scale Scores
Figure B.7.22. Spring 2025 STAAR Spanish Grade 5 Science Frequency Distribution of Scale Scores
Figure B.7.23. Spring 2025 STAAR Algebra I Frequency Distribution of Scale Scores
Figure B.7.24. Spring 2025 STAAR English I Frequency Distribution of Scale Scores
Figure B.7.25. Spring 2025 STAAR English II Frequency Distribution of Scale Scores
Figure B.7.26. Spring 2025 STAAR Biology Frequency Distribution of Scale Scores
Figure B.7.27. Spring 2025 STAAR U.S. History Frequency Distribution of Scale Scores

Spring 2025 STAAR Classification Consistency and Accuracy

Table B.1.1. Spring 2025 STAAR Grades 3–8
Classification Consistency and Accuracy

	Decision Consistency				Decision Accuracy			
Grade	**Mathematics**	**RLA**	**Science**	**Social Studies**	**Mathematics**	**RLA**	**Science**	**Social Studies**
3	69.0	70.3			77.4	78.6		
4	70.6	70.0			78.6	78.4		
5	71.5	70.9	66.3		79.5	78.8	75.0	
6	69.4	71.8			77.8	79.6		
7	74.3	71.5			81.4	79.3		
8	72.3	71.3	69.0	72.8	80.0	79.2	77.5	80.1

Table B.1.2. Spring 2025 STAAR Spanish Grades 3–5
Classification Consistency and Accuracy

	Decision Consistency			Decision Accuracy		
Grade	**Mathematics**	**RLA**	**Science**	**Mathematics**	**RLA**	**Science**
3	70.2	73.7		78.2	80.5	
4	73.0	71.5		80.4	78.5	
5	75.3	70.8	76.2	82.3	78.6	82.8

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 2025 STAAR EOC Assessments
Classification Consistency and Accuracy

Test	Design Consistency	Design Accuracy
Algebra I	73.1	80.6
English I	77.0	83.4
English II	77.6	83.7
Biology	71.5	79.5
U.S. History	75.1	82.2

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

Spring 2025 STAAR
Scale Score Correlations

Table B.2.1. Spring 2025 STAAR Grades 3–8 Scale Score Correlations

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	360,839	0.76
4	Mathematics	RLA	368,011	0.73
5	Mathematics	RLA	372,724	0.75
	Mathematics	Science	376,612	0.77
	RLA	Science	376,117	0.74
6	Mathematics	RLA	384,347	0.73
7	Mathematics	RLA	291,058	0.71
8	Mathematics	RLA	268,114	0.68
	Mathematics	Science	274,992	0.76
	Mathematics	Social Studies	268,912	0.69
	RLA	Science	378,059	0.78
	RLA	Social Studies	393,433	0.76
	Science	Social Studies	380,890	0.80

Table B.2.2. Spring 2025 STAAR Spanish Grades 3–5 Scale Score Correlations

Grade	STAAR	STAAR	N	Correlation
3	Mathematics	RLA	19,836	0.68
4	Mathematics	RLA	14,767	0.65
5	Mathematics	RLA	11,655	0.66
	Mathematics	Science	10,466	0.64
	RLA	Science	13,424	0.64

Table B.2.3. Spring 2025 STAAR EOC Assessments Scale Score Correlations

STAAR	STAAR	N	Correlation
Algebra I	English I	290,848	0.65
Algebra I	English II	36,782	0.46
Algebra I	Biology	250,662	0.69
Algebra I	U.S. History	11,941	0.52
English I	English II	67,181	0.63
English I	Biology	357,792	0.81
English I	U.S. History	24,561	0.58
English II	Biology	59,372	0.78
English II	U.S. History	70,900	0.75
Biology	U.S. History	12,682	0.76

Spring 2025 STAAR
Conditional Standard Error of Measurement for Scale Scores

Table B.3.1. Spring 2025 STAAR and 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	922	134	1018	132	1061	132
2	1019	96	1113	95	1156	95
3	1078	80	1170	79	1214	79
4	1120	70	1212	69	1256	70
5	1155	64	1245	62	1290	63
6	1184	60	1272	58	1319	59
7	1210	56	1296	54	1344	55
8	1233	54	1318	52	1366	53
9	1255	52	1337	50	1387	51
10	1275	50	1356	48	1406	49
11	1294	49	1373	47	1424	47
12	1312	48	1389	46	1441	46
13	1330	48	1405	45	1457	45
14	1347	47	1421	44	1473	45
15	*1360	47	1436	44	1488	44
16	1380	46	1450	44	1502	44
17	1397	46	*1462	44	*1515	43
18	1413	46	1480	44	1531	43
19	1430	46	1494	44	1545	43
20	1446	47	1509	44	1559	43
21	1463	47	1524	44	1573	43
22	**1471	48	1538	44	1588	43
23	1498	48	1554	45	1602	43
24	1516	49	**1557	45	1616	43
25	1535	50	1585	46	1631	44
26	1555	51	1601	46	**1634	44
27	1576	53	1618	47	1661	45
28	1598	54	1635	48	1677	46
29	***1600	56	1654	49	1693	47
30	1646	59	1673	51	1711	48
31	1674	62	***1690	53	1728	49
32	1706	66	1716	55	1747	50

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1743	72	1740	58	1768	52
34	1788	81	1768	61	***1776	54
35	1848	97	1798	66	1813	57
36	1946	134	1835	72	1840	60
37	2070		1879	81	1870	65
38			1940	97	1905	71
39			2038	134	1949	81
40			2130		2008	97
41					2105	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 2025 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	1172	133	1251	132	1324	132
2	1267	96	1345	95	1418	95
3	1325	79	1402	78	1474	78
4	1368	70	1443	69	1516	69
5	1402	64	1476	62	1549	62
6	1431	59	1504	58	1576	58
7	1457	56	1528	54	1600	54
8	1480	53	1549	51	1622	52
9	1501	51	1569	49	1642	49
10	1520	49	1587	47	1660	48
11	1538	48	1603	46	1676	46
12	1555	47	1619	45	1692	45
13	1572	46	1634	44	1707	44
14	1588	45	1649	43	1722	43
15	1603	44	1663	42	1736	42
16	*1616	44	1676	42	1749	41
17	1633	43	1689	41	*1754	41
18	1647	43	1702	41	1775	40
19	1661	43	*1703	41	1787	40
20	1675	43	1727	40	1800	40
21	1689	43	1740	40	1812	40
22	1703	43	1752	40	1824	39
23	1717	43	1765	40	1836	39
24	1732	43	1777	40	1847	39
25	**1745	43	1790	41	**1859	39
26	1760	44	**1793	41	1871	39
27	1775	44	1816	41	1883	40
28	1790	45	1829	42	1895	40
29	1806	45	1842	42	1908	40
30	1822	46	1856	43	1920	41
31	1839	47	1870	43	1933	41
32	1857	49	1885	44	1946	42

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1876	50	1900	45	1960	42
34	***1889	52	1916	46	1974	43
35	1917	54	1933	47	1988	44
36	1941	57	1951	49	2003	45
37	1967	60	***1965	51	***2009	46
38	1997	65	1990	53	2037	48
39	2033	71	2013	56	2055	50
40	2077	81	2038	59	2075	52
41	2137	97	2067	64	2097	55
42	2234	134	2101	70	2122	59
43	2350		2143	79	2150	63
44			2201	95	2184	70
45			2296	133	2226	79
46			2400		2284	95
47					2379	133
48					2470	

Notes:
* Approaches, ** Meets, *** Masters
Conditional Standard Error of Measurement at minimum and maximum scores left intentionally blank, because sufficient information was not available to accurately estimate these values.

Table B.3.3. Spring 2025 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	806	148	905	147	913	146
2	914	107	1012	106	1018	105
3	980	90	1077	89	1082	88
4	1030	79	1125	78	1128	77
5	1069	72	1164	71	1166	70
6	1103	67	1196	66	1198	65
7	1133	63	1225	62	1225	61
8	1159	60	1251	59	1250	58
9	1183	58	1274	57	1273	56
10	1206	56	1296	55	1294	54
11	1226	54	1316	53	1314	52
12	1246	53	1335	52	1333	51
13	1265	51	1353	50	1350	50
14	1283	50	1370	49	1367	49
15	1300	49	1387	48	1384	48
16	1317	49	1402	47	1400	48
17	1333	48	*1414	46	1415	47
18	*1345	47	1432	46	1430	46
19	1365	47	1446	45	1445	46
20	1380	47	1460	44	1460	46
21	1395	46	1473	43	1474	45
22	1409	46	1486	42	*1475	45
23	1424	45	1498	42	1502	44
24	1438	45	1510	41	1515	44
25	1452	45	1522	40	1529	44
26	1466	44	1533	40	1542	43
27	**1467	44	1543	39	1555	43
28	1493	43	**1552	39	1567	42
29	1506	43	1564	39	1579	42
30	1519	43	1575	39	**1592	42
31	1531	42	1585	39	1604	42
32	1544	42	1596	39	1616	42

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1556	42	1606	40	1628	42
34	1568	42	1618	40	1641	43
35	1580	42	1629	41	1653	43
36	1593	43	1641	42	1667	44
37	***1596	43	1654	43	1681	45
38	1619	44	***1663	44	1696	47
39	1633	45	1682	46	***1700	48
40	1648	47	1697	47	1728	49
41	1664	49	1713	49	1745	51
42	1681	51	1731	51	1764	53
43	1700	54	1750	54	1784	56
44	1721	57	1771	57	1807	59
45	1745	61	1795	60	1833	62
46	1774	67	1822	65	1862	67
47	1808	74	1854	70	1896	74
48	1851	84	1892	77	1939	82
49	1910	100	1938	88	1992	94
50	1996	126	2002	105	2065	112
51	2149	176	2106	145	2181	152
52	2200		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 2025 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	994	147	989	147	1064	146
2	1100	106	1094	106	1169	106
3	1164	88	1158	88	1233	88
4	1212	78	1205	77	1280	77
5	1250	71	1243	70	1317	70
6	1282	65	1275	65	1349	65
7	1310	62	1303	61	1377	61
8	1335	58	1328	58	1402	58
9	1358	56	1351	56	1424	56
10	1378	53	1372	54	1445	54
11	1398	52	1391	52	1465	52
12	1416	50	1409	51	1483	50
13	1432	48	1427	49	1500	49
14	1448	47	1443	48	1516	47
15	1463	46	1459	47	1531	46
16	1477	45	1474	46	1545	45
17	1491	44	1489	45	1559	43
18	1504	43	1503	45	1572	42
19	1516	42	1517	44	1584	41
20	1528	41	1530	43	*1592	40
21	*1535	40	1543	43	1606	39
22	1551	40	1555	42	1616	38
23	1562	39	*1564	41	1625	37
24	1572	39	1579	41	1635	36
25	1583	38	1591	40	1644	36
26	1593	38	1602	40	1652	35
27	1603	38	1613	39	1661	35
28	1613	38	1623	39	1670	35
29	1623	38	1634	39	1678	36
30	1633	38	1644	38	1687	36
31	**1634	38	1654	38	1696	36
32	1654	38	1664	38	**1698	37

(Continued)

	Grade 6		Grade 7		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	1664	39	**1669	39	1716	38
34	1674	39	1685	39	1726	39
35	1685	39	1696	39	1736	39
36	1696	39	1706	40	1747	40
37	1707	40	1718	40	1759	41
38	1718	40	1729	41	1771	42
39	1729	41	1741	42	1783	43
40	1741	41	1753	42	1796	44
41	***1749	42	1766	43	***1803	45
42	1766	43	***1771	44	1824	46
43	1779	44	1793	46	1839	47
44	1793	45	1808	47	1856	49
45	1808	47	1824	48	1873	50
46	1824	49	1841	50	1891	52
47	1841	51	1859	52	1911	55
48	1860	54	1878	54	1932	57
49	1881	57	1899	57	1956	60
50	1906	61	1924	61	1983	64
51	1934	67	1951	66	2015	70
52	1968	74	1985	73	2052	77
53	2012	85	2028	84	2098	87
54	2073	103	2087	103	2162	105
55	2175	145	2189	145	2267	146
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 2025 STAAR and STAAR Spanish Grade 5 and STAAR Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1140		1000	
1	1479	579	1413	639
2	1907	421	1868	458
3	2172	351	2141	378
4	2367	309	2339	331
5	2522	280	2495	299
6	2653	259	2626	276
7	2766	243	2740	259
8	2866	230	2841	245
9	2957	220	2932	234
10	3040	211	3016	226
11	3118	204	3094	219
12	3191	199	3167	213
13	3260	194	3237	208
14	3327	190	3305	204
15	3391	187	3369	200
16	3453	185	3432	198
17	3514	183	3494	196
18	*3550	182	*3550	194
19	3633	181	3613	193
20	3692	181	3672	192
21	3751	181	3730	192
22	3810	182	3788	192
23	3870	183	3847	192
24	3931	186	3905	192
25	3994	188	3964	193
26	**4000	192	**4000	195
27	4127	197	4085	197
28	4198	202	4147	199
29	4274	209	4211	202
30	4355	217	4276	205
31	***4380	227	4344	209
32	4542	239	4415	214

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	4651	255	4490	219
34	4777	275	4568	226
35	4926	302	***4619	233
36	5112	343	4741	243
37	5365	412	4839	254
38	5778	570	4947	267
39	6200		5067	284
40			5205	307
41			5369	337
42			5571	380
43			5842	452
44			6260	589
45			7099	884
46			7980	

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 2025 STAAR Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

		Grade 8
Raw	**SS**	**CSEM**
0	1050	
1	1321	588
2	1750	426
3	2013	355
4	2207	314
5	2363	285
6	2495	265
7	2610	249
8	2713	236
9	2807	226
10	2893	218
11	2973	210
12	3048	204
13	3119	198
14	3186	194
15	3250	189
16	3312	186
17	3371	182
18	3428	179
19	3483	177
20	3537	175
21	*3550	173
22	3642	171
23	3693	170
24	3743	169
25	3794	169
26	3843	169
27	3893	169
28	3944	170
29	3994	171
30	**4000	172
31	4098	174
32	4151	176

(Continued)

	Grade 8	
Raw	**SS**	**CSEM**
33	4206	178
34	4263	181
35	4322	185
36	***4352	189
37	4447	194
38	4515	200
39	4588	207
40	4665	215
41	4750	225
42	4844	237
43	4949	253
44	5069	273
45	5213	302
46	5394	344
47	5642	415
48	6054	579
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 2025 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	721	156	777	157	830	156
2	834	113	890	113	942	112
3	903	94	959	94	1009	93
4	953	82	1010	83	1059	82
5	993	75	1051	76	1099	75
6	1027	69	1085	70	1132	69
7	1056	65	1115	66	1161	65
8	1083	62	1142	62	1188	62
9	1107	59	1167	60	1211	59
10	1129	57	1189	57	1233	57
11	1149	55	1210	55	1254	55
12	1168	53	1229	54	1273	53
13	1186	52	1247	52	1291	52
14	1204	51	1264	51	1308	51
15	1220	50	1281	49	1324	49
16	1236	49	1296	48	1340	48
17	1252	48	1311	47	1355	47
18	1267	48	1325	46	1369	46
19	1282	47	1339	45	1383	46
20	1296	46	1352	45	1396	45
21	1310	46	1365	44	1409	44
22	*1318	45	1377	44	1422	43
23	1337	45	1390	43	*1431	43
24	1350	44	1402	43	1446	42
25	1362	44	*1408	43	1457	42
26	1375	43	1426	43	1469	42
27	1387	43	1439	43	1480	42
28	1399	42	1451	44	1492	42
29	1410	42	1464	44	1503	42
30	1422	42	1476	45	1515	43
31	1434	42	**1488	45	1527	43
32	1445	43	1503	46	1539	44

(Continued)

	Grade 3		Grade 4		Grade 5	
Raw	**SS**	**CSEM**	**SS**	**CSEM**	**SS**	**CSEM**
33	**1447	43	1517	47	1552	44
34	1470	44	1532	47	**1556	45
35	1482	44	1547	48	1579	46
36	1496	45	1562	49	1593	47
37	1509	47	1578	50	1608	48
38	***1515	48	***1581	51	1624	50
39	1540	50	1612	52	1640	51
40	1556	51	1630	53	1658	53
41	1574	53	1649	55	***1662	55
42	1593	55	1669	56	1697	57
43	1613	57	1691	58	1719	59
44	1636	60	1714	61	1743	62
45	1661	63	1739	64	1770	66
46	1689	68	1768	68	1800	70
47	1721	73	1800	74	1835	76
48	1759	80	1839	81	1876	83
49	1807	91	1888	92	1926	93
50	1871	109	1954	111	1994	111
51	1978	152	2063	154	2103	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.3.8. Spring 2025 STAAR Algebra I
Conditional Standard Error of Measurement for Scale Scores

	Algebra I	
Raw	**SS**	**CSEM**
0	1500	
1	1921	468
2	2256	336
3	2458	279
4	2606	245
5	2723	222
6	2822	205
7	2908	193
8	2984	183
9	3053	174
10	3116	167
11	3175	162
12	3230	157
13	3282	152
14	3331	149
15	3378	146
16	3423	143
17	3467	140
18	3489	138
19	3541	136
20	*3550	134
21	3628	133
22	3666	132
23	3703	130
24	3740	129
25	3776	129
26	3811	128
27	3847	127
28	3882	127
29	3917	127
30	3952	127
31	3987	127
32	**4000	127

(Continued)

		Algebra I
Raw	**SS**	**CSEM**
33	4057	127
34	4092	128
35	4127	128
36	4163	129
37	4199	130
38	4236	131
39	4274	132
40	4312	133
41	***4345	135
42	4391	137
43	4433	139
44	4476	142
45	4521	145
46	4567	149
47	4617	153
48	4669	158
49	4725	164
50	4785	170
51	4851	179
52	4925	189
53	5007	202
54	5103	219
55	5217	242
56	5361	276
57	5560	334
58	5891	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.3.9. Spring 2025 STAAR English I and English II Conditional Standard Error of Measurement for Scale Scores

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
0	1750		1650	
1	1976	438	1930	452
2	2290	315	2253	325
3	2481	262	2448	269
4	2621	230	2591	236
5	2733	209	2705	214
6	2827	194	2800	198
7	2909	182	2883	186
8	2982	172	2957	177
9	3048	165	3024	169
10	3109	158	3086	162
11	3165	153	3143	157
12	3217	148	3197	152
13	3267	144	3247	148
14	3314	140	3295	145
15	3359	137	3341	141
16	3401	134	3386	139
17	3442	131	3428	136
18	3481	128	3469	134
19	3518	125	3509	132
20	3554	123	3547	130
21	3589	120	3584	128
22	3622	117	3621	126
23	3653	115	3656	124
24	3683	112	3690	122
25	3712	110	3723	120
26	3739	108	3755	118
27	3766	106	*3775	116
28	*3775	104	3815	114
29	3816	103	3844	112
30	3841	102	3872	110
31	3865	102	3898	108
32	3889	102	3925	107

(Continued)

	English I		English II	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	3913	102	3950	106
34	3938	103	3975	105
35	3963	104	**4000	105
36	3988	105	4025	105
37	**4000	106	4049	105
38	4040	108	4075	106
39	4068	109	4100	107
40	4096	111	4126	109
41	4125	113	4153	111
42	4155	114	4182	113
43	4186	116	4211	115
44	4218	118	4241	118
45	4251	120	4274	121
46	4285	122	4307	124
47	4320	124	4343	128
48	4357	127	4381	131
49	4396	130	4420	135
50	4436	133	4463	139
51	4478	137	4508	144
52	4523	141	4556	148
53	4570	145	4607	154
54	***4606	151	4662	160
55	4676	157	4722	167
56	4737	165	***4734	175
57	4803	174	4860	185
58	4878	186	4942	197
59	4966	202	5036	213
60	5070	223	5149	235
61	5202	255	5291	269
62	5384	309	5486	325
63	5690	433	5810	453
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.3.10. Spring 2025 STAAR Biology
Conditional Standard Error of Measurement for Scale Scores

		Biology
Raw	**SS**	**CSEM**
0	1900	
1	2038	447
2	2362	323
3	2561	270
4	2708	238
5	2826	217
6	2926	202
7	3014	190
8	3093	181
9	3164	173
10	3230	167
11	3292	161
12	3350	157
13	3405	153
14	3458	150
15	3509	147
16	*3550	145
17	3604	142
18	3650	141
19	3695	139
20	3739	138
21	3782	136
22	3824	135
23	3866	134
24	3907	134
25	3948	133
26	3988	133
27	**4000	132
28	4069	132
29	4109	132
30	4149	133
31	4190	133
32	4230	134

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	4272	134
34	4313	136
35	4356	137
36	4400	139
37	4444	141
38	4490	143
39	***4531	146
40	4588	149
41	4641	153
42	4696	158
43	4755	163
44	4819	170
45	4888	178
46	4965	187
47	5050	199
48	5148	215
49	5265	236
50	5409	268
51	5606	322
52	5928	446
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.3.11. Spring 2025 STAAR U.S. History Conditional Standard Error of Measurement for Scale Scores

	U.S. History	
Raw	**SS**	**CSEM**
0	1420	
1	1765	494
2	2116	354
3	2327	293
4	2480	256
5	2602	232
6	2703	214
7	2791	200
8	2869	189
9	2939	180
10	3003	173
11	3062	167
12	3117	161
13	3169	156
14	3217	152
15	3264	149
16	3308	145
17	3350	142
18	3391	140
19	3431	137
20	3469	135
21	3486	133
22	3536	132
23	*3550	130
24	3611	129
25	3645	128
26	3678	126
27	3711	125
28	3743	125
29	3774	124
30	3805	123
31	3836	122
32	3867	122

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
33	3897	122
34	3928	121
35	3958	121
36	3987	121
37	**4000	121
38	4047	120
39	4077	120
40	4107	121
41	4136	121
42	4166	121
43	4196	121
44	4226	121
45	4257	122
46	4287	122
47	4318	123
48	4349	123
49	4381	124
50	4412	125
51	***4424	126
52	4477	127
53	4511	128
54	4544	129
55	4579	130
56	4614	132
57	4650	133
58	4687	135
59	4725	137
60	4764	139
61	4804	142
62	4846	144
63	4890	148
64	4936	151

(Continued)

	U.S. History	
Raw	**SS**	**CSEM**
65	4984	155
66	5034	160
67	5088	165
68	5146	171
69	5209	178
70	5277	187
71	5353	198
72	5439	212
73	5539	230
74	5659	254
75	5810	291
76	6018	353
77	6367	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 2025 STAAR
Mean P-Values, Raw Score Summary, and Reliability

Table B.4.1. Spring 2025 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	376,507	20.11	8.22	0.90	2.59	51.67
	Numerical Representations and Relationships	12	376,507	7.08	2.82	0.74	1.43	57.64
	Computations and Algebraic Relationships	14	376,507	6.75	3.81	0.82	1.64	46.07
	Geometry and Measurement	7	376,507	4.38	1.60	0.55	1.07	57.64
	Data Analysis and Personal Financial Literacy	4	376,507	1.90	1.14	0.46	0.84	47.59
RLA	OVERALL TEST	52	363,664	27.30	10.69	0.92	3.05	59.33
	Reading	26	363,664	15.64	5.27	0.84	2.12	60.67
	Writing	26	363,664	11.66	6.04	0.87	2.17	57.44

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.4.2. Spring 2025 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	185,356	19.35	7.97	0.89	2.61	49.65
	Numerical Representations and Relationships	12	185,356	6.75	2.77	0.72	1.45	55.01
	Computations and Algebraic Relationships	14	185,356	6.47	3.71	0.80	1.65	44.08
	Geometry and Measurement	7	185,356	4.26	1.56	0.52	1.08	55.70
	Data Analysis and Personal Financial Literacy	4	185,356	1.87	1.11	0.44	0.84	46.73
RLA	OVERALL TEST	52	178,726	28.09	10.71	0.92	3.04	60.57
	Reading	26	178,726	15.89	5.22	0.84	2.12	61.55
	Writing	26	178,726	12.20	6.10	0.87	2.17	59.19

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.4.3. Spring 2025 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	191,081	20.85	8.39	0.91	2.56	53.63
	Numerical Representations and Relationships	12	191,081	7.40	2.83	0.75	1.41	60.19
	Computations and Algebraic Relationships	14	191,081	7.02	3.89	0.83	1.62	48.00
	Geometry and Measurement	7	191,081	4.49	1.62	0.58	1.06	59.51
	Data Analysis and Personal Financial Literacy	4	191,081	1.94	1.16	0.47	0.84	48.43
RLA	OVERALL TEST	52	184,872	26.54	10.62	0.92	3.05	58.14
	Reading	26	184,872	15.41	5.30	0.84	2.13	59.83
	Writing	26	184,872	11.13	5.94	0.87	2.17	55.76

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.4. Spring 2025 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	49,183	17.11	7.74	0.89	2.61	43.20
	Numerical Representations and Relationships	12	49,183	6.12	2.77	0.71	1.48	48.95
	Computations and Algebraic Relationships	14	49,183	5.50	3.54	0.79	1.62	37.13
	Geometry and Measurement	7	49,183	3.90	1.58	0.51	1.10	50.07
	Data Analysis and Personal Financial Literacy	4	49,183	1.59	1.08	0.40	0.84	39.85
RLA	OVERALL TEST	52	49,157	24.25	10.30	0.91	3.09	53.50
	Reading	26	49,157	14.25	5.26	0.83	2.17	55.29
	Writing	26	49,157	10.00	5.67	0.85	2.18	50.98

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.4.5. Spring 2025 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	190,108	18.86	7.84	0.89	2.62	48.17
	Numerical Representations and Relationships	12	190,108	6.70	2.75	0.72	1.46	54.25
	Computations and Algebraic Relationships	14	190,108	6.20	3.66	0.80	1.65	42.23
	Geometry and Measurement	7	190,108	4.20	1.57	0.52	1.09	54.87
	Data Analysis and Personal Financial Literacy	4	190,108	1.76	1.08	0.40	0.84	43.92
RLA	OVERALL TEST	52	177,518	25.60	10.28	0.91	3.09	55.95
	Reading	26	177,518	14.80	5.13	0.82	2.17	57.37
	Writing	26	177,518	10.81	5.79	0.86	2.19	53.95

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 2025 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	97,357	22.55	8.00	0.90	2.54	58.53
	Numerical Representations and Relationships	12	97,357	7.86	2.67	0.73	1.39	64.73
	Computations and Algebraic Relationships	14	97,357	7.78	3.77	0.81	1.63	53.36
	Geometry and Measurement	7	97,357	4.73	1.53	0.54	1.04	62.87
	Data Analysis and Personal Financial Literacy	4	97,357	2.19	1.16	0.48	0.84	54.67
RLA	OVERALL TEST	52	97,269	30.17	10.35	0.92	2.97	65.18
	Reading	26	97,269	17.13	5.01	0.83	2.05	66.47
	Writing	26	97,269	13.04	5.97	0.87	2.13	63.36

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.4.7. Spring 2025 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	312,795	19.70	8.23	0.90	2.59	50.56
	Numerical Representations and Relationships	12	312,795	6.95	2.83	0.74	1.44	56.52
	Computations and Algebraic Relationships	14	312,795	6.57	3.81	0.82	1.64	44.82
	Geometry and Measurement	7	312,795	4.33	1.61	0.55	1.07	56.84
	Data Analysis and Personal Financial Literacy	4	312,795	1.86	1.13	0.45	0.84	46.54
RLA	OVERALL TEST	52	130,846	20.15	8.57	0.87	3.11	45.24
	Reading	26	130,846	12.30	4.64	0.77	2.24	47.71
	Writing	26	130,846	7.85	4.66	0.79	2.14	41.77

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.4.8. Spring 2025 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	63,712	22.09	7.87	0.90	2.55	57.11
	Numerical Representations and Relationships	12	63,712	7.71	2.65	0.73	1.38	63.13
	Computations and Algebraic Relationships	14	63,712	7.63	3.70	0.80	1.64	52.20
	Geometry and Measurement	7	63,712	4.64	1.54	0.54	1.04	61.56
	Data Analysis and Personal Financial Literacy	4	63,712	2.11	1.13	0.46	0.83	52.73
RLA	OVERALL TEST	52	232,818	31.32	9.62	0.90	2.97	67.25
	Reading	26	232,818	17.52	4.63	0.80	2.05	67.96
	Writing	26	232,818	13.80	5.66	0.86	2.14	66.25

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.4.9. Spring 2025 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	381,077	22.45	9.15	0.91	2.73	53.83
	Numerical Representations and Relationships	10	381,077	5.84	2.75	0.77	1.32	56.71
	Computations and Algebraic Relationships	12	381,077	6.64	3.11	0.76	1.53	52.90
	Geometry and Measurement	13	381,077	6.73	3.07	0.75	1.55	49.31
	Data Analysis and Personal Financial Literacy	5	381,077	3.24	1.42	0.54	0.96	59.85
RLA	OVERALL TEST	52	371,478	28.23	11.02	0.92	3.06	59.01
	Reading	26	371,478	14.65	5.38	0.83	2.21	56.40
	Writing	26	371,478	13.57	6.31	0.89	2.11	62.69

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 2025 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	187,978	21.62	8.83	0.90	2.76	51.52
	Numerical Representations and Relationships	10	187,978	5.63	2.67	0.75	1.34	54.43
	Computations and Algebraic Relationships	12	187,978	6.34	3.06	0.74	1.55	50.30
	Geometry and Measurement	13	187,978	6.43	2.96	0.72	1.56	46.49
	Data Analysis and Personal Financial Literacy	5	187,978	3.22	1.39	0.52	0.97	59.38
RLA	OVERALL TEST	52	182,836	29.29	10.91	0.92	3.05	60.64
	Reading	26	182,836	15.06	5.29	0.83	2.20	57.88
	Writing	26	182,836	14.23	6.31	0.89	2.10	64.54

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.4.11. Spring 2025 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	193,021	23.26	9.39	0.92	2.70	56.08
	Numerical Representations and Relationships	10	193,021	6.04	2.80	0.78	1.30	58.93
	Computations and Algebraic Relationships	12	193,021	6.94	3.13	0.77	1.51	55.45
	Geometry and Measurement	13	193,021	7.02	3.15	0.76	1.54	52.07
	Data Analysis and Personal Financial Literacy	5	193,021	3.25	1.45	0.56	0.96	60.30
RLA	OVERALL TEST	52	188,567	27.19	11.02	0.92	3.07	57.43
	Reading	26	188,567	14.25	5.45	0.84	2.21	54.97
	Writing	26	188,567	12.94	6.26	0.89	2.11	60.90

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.4.12. Spring 2025 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	49,696	18.86	8.47	0.89	2.79	44.66
	Numerical Representations and Relationships	10	49,696	4.94	2.68	0.75	1.35	47.40
	Computations and Algebraic Relationships	12	49,696	5.63	2.95	0.72	1.57	44.42
	Geometry and Measurement	13	49,696	5.51	2.74	0.68	1.55	39.62
	Data Analysis and Personal Financial Literacy	5	49,696	2.78	1.46	0.49	1.04	50.40
RLA	OVERALL TEST	52	49,766	25.05	10.83	0.92	3.10	53.39
	Reading	26	49,766	13.32	5.35	0.82	2.24	51.35
	Writing	26	49,766	11.73	6.16	0.88	2.13	56.27

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.4.13. Spring 2025 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	195,760	21.33	8.76	0.90	2.77	50.92
	Numerical Representations and Relationships	10	195,760	5.61	2.70	0.76	1.34	54.38
	Computations and Algebraic Relationships	12	195,760	6.25	3.01	0.73	1.56	49.55
	Geometry and Measurement	13	195,760	6.35	2.93	0.71	1.57	46.19
	Data Analysis and Personal Financial Literacy	5	195,760	3.12	1.41	0.51	0.98	57.14
RLA	OVERALL TEST	52	185,283	26.66	10.67	0.92	3.10	55.95
	Reading	26	185,283	13.79	5.19	0.81	2.24	53.07
	Writing	26	185,283	12.87	6.20	0.88	2.13	60.02

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.4.14. Spring 2025 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	96,650	24.87	8.96	0.91	2.67	60.06
	Numerical Representations and Relationships	10	96,650	6.37	2.67	0.76	1.30	62.19
	Computations and Algebraic Relationships	12	96,650	7.40	3.04	0.76	1.49	59.24
	Geometry and Measurement	13	96,650	7.58	3.05	0.75	1.54	56.15
	Data Analysis and Personal Financial Literacy	5	96,650	3.52	1.35	0.54	0.92	66.09
RLA	OVERALL TEST	52	97,040	31.01	10.53	0.92	3.00	64.52
	Reading	26	97,040	16.17	5.19	0.83	2.15	62.28
	Writing	26	97,040	14.85	6.03	0.88	2.07	67.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.15. Spring 2025 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	143,374	17.34	7.73	0.87	2.82	40.80
	Numerical Representations and Relationships	10	143,374	4.50	2.54	0.71	1.37	43.08
	Computations and Algebraic Relationships	12	143,374	5.06	2.69	0.65	1.60	39.78
	Geometry and Measurement	13	143,374	5.21	2.55	0.63	1.55	37.28
	Data Analysis and Personal Financial Literacy	5	143,374	2.58	1.43	0.44	1.07	46.13
RLA	OVERALL TEST	52	135,934	20.53	9.14	0.88	3.17	44.74
	Reading	26	135,934	11.11	4.61	0.75	2.29	42.83
	Writing	26	135,934	9.42	5.33	0.83	2.17	47.44

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.4.16. Spring 2025 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	237,703	25.53	8.55	0.90	2.66	61.69
	Numerical Representations and Relationships	10	237,703	6.65	2.54	0.74	1.29	64.92
	Computations and Algebraic Relationships	12	237,703	7.60	2.95	0.75	1.48	60.82
	Geometry and Measurement	13	237,703	7.65	3.00	0.74	1.54	56.57
	Data Analysis and Personal Financial Literacy	5	237,703	3.63	1.26	0.50	0.89	68.12
RLA	OVERALL TEST	52	235,544	32.67	9.45	0.90	2.98	67.24
	Reading	26	235,544	16.70	4.69	0.79	2.15	64.23
	Writing	26	235,544	15.97	5.55	0.86	2.05	71.50

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.4.17. Spring 2025 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	383,295	24.26	9.67	0.92	2.77	55.14
	Numerical Representations and Relationships	9	383,295	5.80	2.33	0.69	1.30	64.02
	Computations and Algebraic Relationships	20	383,295	11.49	4.92	0.85	1.87	54.05
	Geometry and Measurement	9	383,295	4.80	2.31	0.70	1.27	49.31
	Data Analysis and Personal Financial Literacy	4	383,295	2.17	1.29	0.55	0.86	54.16
RLA	OVERALL TEST	52	378,262	30.89	11.01	0.93	2.95	65.89
	Reading	26	378,262	17.13	5.27	0.84	2.08	66.33
	Writing	26	378,262	13.76	6.32	0.89	2.07	65.26
Science	OVERALL TEST	39	383,147	20.96	7.45	0.85	2.86	52.58
	Matter and Energy	6	383,147	2.83	1.46	0.38	1.15	47.23
	Force, Motion, and Energy	9	383,147	5.04	2.27	0.56	1.50	54.77
	Earth and Space	10	383,147	4.81	2.45	0.63	1.50	48.32
	Organisms and Environments	14	383,147	8.28	2.98	0.73	1.55	58.08

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.4.18. Spring 2025 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	187,940	23.89	9.39	0.91	2.78	54.09
	Numerical Representations and Relationships	9	187,940	5.77	2.30	0.67	1.31	63.71
	Computations and Algebraic Relationships	20	187,940	11.40	4.80	0.85	1.87	53.38
	Geometry and Measurement	9	187,940	4.62	2.23	0.67	1.28	46.94
	Data Analysis and Personal Financial Literacy	4	187,940	2.10	1.29	0.55	0.87	52.57
RLA	OVERALL TEST	52	185,113	32.07	10.68	0.93	2.93	68.11
	Reading	26	185,113	17.70	5.06	0.84	2.05	68.49
	Writing	26	185,113	14.37	6.20	0.89	2.07	67.56
Science	OVERALL TEST	39	187,649	20.53	7.31	0.85	2.87	51.40
	Matter and Energy	6	187,649	2.86	1.46	0.38	1.15	47.69
	Force, Motion, and Energy	9	187,649	4.80	2.20	0.53	1.50	52.22
	Earth and Space	10	187,649	4.66	2.42	0.61	1.50	46.51
	Organisms and Environments	14	187,649	8.21	2.97	0.73	1.55	57.45

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.19. Spring 2025 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	195,287	24.62	9.91	0.92	2.75	56.16
	Numerical Representations and Relationships	9	195,287	5.84	2.36	0.71	1.28	64.34
	Computations and Algebraic Relationships	20	195,287	11.58	5.04	0.86	1.87	54.71
	Geometry and Measurement	9	195,287	4.98	2.37	0.72	1.26	51.59
	Data Analysis and Personal Financial Literacy	4	195,287	2.23	1.29	0.56	0.86	55.70
RLA	OVERALL TEST	52	193,090	29.77	11.20	0.93	2.96	63.77
	Reading	26	193,090	16.58	5.41	0.85	2.10	64.27
	Writing	26	193,090	13.19	6.38	0.89	2.07	63.06
Science	OVERALL TEST	39	195,435	21.38	7.55	0.86	2.85	53.71
	Matter and Energy	6	195,435	2.81	1.46	0.38	1.15	46.79
	Force, Motion, and Energy	9	195,435	5.27	2.31	0.59	1.48	57.23
	Earth and Space	10	195,435	4.96	2.47	0.63	1.49	50.05
	Organisms and Environments	14	195,435	8.35	2.99	0.73	1.55	58.68

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.20. Spring 2025 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	49,200	20.59	9.07	0.90	2.84	46.28
	Numerical Representations and Relationships	9	49,200	4.99	2.32	0.65	1.36	54.86
	Computations and Algebraic Relationships	20	49,200	9.82	4.68	0.83	1.92	45.76
	Geometry and Measurement	9	49,200	4.03	2.11	0.62	1.30	40.36
	Data Analysis and Personal Financial Literacy	4	49,200	1.75	1.27	0.53	0.87	43.71
RLA	OVERALL TEST	52	49,462	27.48	10.82	0.92	3.03	59.51
	Reading	26	49,462	15.70	5.27	0.83	2.15	60.71
	Writing	26	49,462	11.79	6.17	0.88	2.12	57.80
Science	OVERALL TEST	39	49,431	18.14	6.87	0.82	2.89	45.71
	Matter and Energy	6	49,431	2.52	1.39	0.31	1.16	42.02
	Force, Motion, and Energy	9	49,431	4.30	2.14	0.50	1.52	46.80
	Earth and Space	10	49,431	4.03	2.26	0.56	1.50	40.75
	Organisms and Environments	14	49,431	7.29	2.93	0.71	1.58	51.63

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 2025 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	197,424	23.03	9.30	0.91	2.81	52.11
	Numerical Representations and Relationships	9	197,424	5.60	2.30	0.67	1.32	61.73
	Computations and Algebraic Relationships	20	197,424	10.93	4.77	0.84	1.91	51.11
	Geometry and Measurement	9	197,424	4.46	2.21	0.66	1.30	45.32
	Data Analysis and Personal Financial Literacy	4	197,424	2.05	1.29	0.54	0.87	51.14
RLA	OVERALL TEST	52	190,884	29.36	10.76	0.92	3.00	62.93
	Reading	26	190,884	16.42	5.21	0.83	2.13	63.53
	Writing	26	190,884	12.94	6.17	0.88	2.10	62.09
Science	OVERALL TEST	39	195,764	19.70	6.99	0.83	2.89	49.41
	Matter and Energy	6	195,764	2.67	1.41	0.33	1.15	44.43
	Force, Motion, and Energy	9	195,764	4.77	2.21	0.53	1.51	51.86
	Earth and Space	10	195,764	4.44	2.33	0.58	1.51	44.90
	Organisms and Environments	14	195,764	7.82	2.88	0.70	1.57	54.75

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.4.22. Spring 2025 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	97,808	26.66	9.32	0.92	2.71	60.92
	Numerical Representations and Relationships	9	97,808	6.27	2.23	0.69	1.25	69.26
	Computations and Algebraic Relationships	20	97,808	12.50	4.79	0.85	1.85	59.21
	Geometry and Measurement	9	97,808	5.44	2.26	0.70	1.24	56.48
	Data Analysis and Personal Financial Literacy	4	97,808	2.44	1.22	0.50	0.86	60.97
RLA	OVERALL TEST	52	98,546	33.70	10.39	0.93	2.84	71.53
	Reading	26	98,546	18.44	4.97	0.84	1.99	71.58
	Writing	26	98,546	15.25	6.01	0.89	2.01	71.46
Science	OVERALL TEST	39	98,529	23.56	7.26	0.85	2.81	58.95
	Matter and Energy	6	98,529	3.14	1.48	0.41	1.14	52.26
	Force, Motion, and Energy	9	98,529	5.64	2.22	0.56	1.47	61.37
	Earth and Space	10	98,529	5.56	2.44	0.64	1.47	55.29
	Organisms and Environments	14	98,529	9.22	2.82	0.71	1.51	64.57

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.4.23. Spring 2025 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	140,058	18.47	8.09	0.87	2.89	41.34
	Numerical Representations and Relationships	9	140,058	4.61	2.20	0.60	1.39	50.67
	Computations and Algebraic Relationships	20	140,058	8.74	4.24	0.79	1.96	40.52
	Geometry and Measurement	9	140,058	3.58	1.91	0.52	1.32	35.51
	Data Analysis and Personal Financial Literacy	4	140,058	1.54	1.22	0.49	0.87	38.53
RLA	OVERALL TEST	52	134,512	22.68	9.50	0.89	3.11	50.16
	Reading	26	134,512	13.48	4.95	0.79	2.25	51.99
	Writing	26	134,512	9.20	5.26	0.83	2.14	47.58
Science	OVERALL TEST	39	136,841	16.69	6.25	0.78	2.90	42.07
	Matter and Energy	6	136,841	2.38	1.36	0.29	1.15	39.74
	Force, Motion, and Energy	9	136,841	4.03	2.07	0.45	1.53	43.91
	Earth and Space	10	136,841	3.64	2.07	0.49	1.48	37.08
	Organisms and Environments	14	136,841	6.64	2.72	0.66	1.60	46.66

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.4.24. Spring 2025 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	243,237	27.59	8.89	0.91	2.68	63.09
	Numerical Representations and Relationships	9	243,237	6.49	2.12	0.66	1.23	71.71
	Computations and Algebraic Relationships	20	243,237	13.07	4.58	0.84	1.82	61.85
	Geometry and Measurement	9	243,237	5.50	2.23	0.69	1.24	57.25
	Data Analysis and Personal Financial Literacy	4	243,237	2.53	1.19	0.48	0.86	63.17
RLA	OVERALL TEST	52	243,750	35.43	8.98	0.90	2.83	74.57
	Reading	26	243,750	19.14	4.26	0.79	1.98	74.25
	Writing	26	243,750	16.28	5.38	0.86	2.01	75.02
Science	OVERALL TEST	39	246,306	23.34	6.99	0.84	2.83	58.41
	Matter and Energy	6	246,306	3.08	1.46	0.39	1.14	51.39
	Force, Motion, and Energy	9	246,306	5.60	2.18	0.54	1.47	60.81
	Earth and Space	10	246,306	5.46	2.40	0.61	1.49	54.56
	Organisms and Environments	14	246,306	9.19	2.72	0.69	1.52	64.42

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.25. Spring 2025 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	390,062	22.38	9.10	0.90	2.84	51.30
	Numerical Representations and Relationships	11	390,062	6.13	2.53	0.67	1.46	56.87
	Computations and Algebraic Relationships	15	390,062	7.59	3.44	0.76	1.67	49.13
	Geometry and Measurement	7	390,062	3.42	1.69	0.58	1.10	45.61
	Data Analysis and Personal Financial Literacy	10	390,062	5.24	2.66	0.71	1.43	52.27
RLA	OVERALL TEST	56	396,497	31.21	12.53	0.94	3.19	57.47
	Reading	29	396,497	16.91	6.01	0.85	2.31	57.73
	Writing	27	396,497	14.30	7.12	0.91	2.19	57.09

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 2025 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	192,015	22.00	8.94	0.90	2.85	50.45
	Numerical Representations and Relationships	11	192,015	6.03	2.47	0.65	1.45	56.21
	Computations and Algebraic Relationships	15	192,015	7.48	3.41	0.76	1.68	48.10
	Geometry and Measurement	7	192,015	3.35	1.68	0.57	1.11	44.89
	Data Analysis and Personal Financial Literacy	10	192,015	5.15	2.64	0.71	1.43	51.36
RLA	OVERALL TEST	56	194,717	32.73	12.24	0.93	3.18	59.61
	Reading	29	194,717	17.36	5.81	0.84	2.31	59.09
	Writing	27	194,717	15.37	7.02	0.90	2.17	60.40

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.27. Spring 2025 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	197,966	22.74	9.23	0.91	2.82	52.13
	Numerical Representations and Relationships	11	197,966	6.23	2.59	0.68	1.45	57.52
	Computations and Algebraic Relationships	15	197,966	7.69	3.47	0.77	1.66	50.12
	Geometry and Measurement	7	197,966	3.49	1.70	0.59	1.09	46.32
	Data Analysis and Personal Financial Literacy	10	197,966	5.32	2.67	0.72	1.43	53.16
RLA	OVERALL TEST	56	201,704	29.74	12.63	0.94	3.20	55.41
	Reading	29	201,704	16.48	6.16	0.86	2.32	56.42
	Writing	27	201,704	13.27	7.07	0.90	2.20	53.90

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.4.28. Spring 2025 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	49,115	19.55	8.14	0.88	2.88	44.82
	Numerical Representations and Relationships	11	49,115	5.48	2.37	0.61	1.47	51.15
	Computations and Algebraic Relationships	15	49,115	6.64	3.14	0.71	1.70	42.59
	Geometry and Measurement	7	49,115	2.91	1.57	0.51	1.10	38.67
	Data Analysis and Personal Financial Literacy	10	49,115	4.53	2.43	0.65	1.44	45.16
RLA	OVERALL TEST	56	49,290	28.21	11.91	0.93	3.25	52.71
	Reading	29	49,290	15.68	5.82	0.84	2.36	53.51
	Writing	27	49,290	12.53	6.75	0.89	2.21	51.51

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.29. Spring 2025 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	207,919	20.63	8.30	0.88	2.88	47.35
	Numerical Representations and Relationships	11	207,919	5.74	2.40	0.62	1.47	53.41
	Computations and Algebraic Relationships	15	207,919	6.98	3.16	0.71	1.70	45.11
	Geometry and Measurement	7	207,919	3.17	1.58	0.51	1.10	41.99
	Data Analysis and Personal Financial Literacy	10	207,919	4.74	2.51	0.66	1.46	47.47
RLA	OVERALL TEST	56	209,719	28.76	12.17	0.93	3.25	53.27
	Reading	29	209,719	15.74	5.88	0.84	2.36	53.68
	Writing	27	209,719	13.02	6.92	0.90	2.22	52.66

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.4.30. Spring 2025 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	96,116	25.39	9.06	0.90	2.80	58.11
	Numerical Representations and Relationships	11	96,116	6.80	2.51	0.68	1.43	62.74
	Computations and Algebraic Relationships	15	96,116	8.58	3.46	0.77	1.64	55.80
	Geometry and Measurement	7	96,116	3.89	1.69	0.58	1.09	52.03
	Data Analysis and Personal Financial Literacy	10	96,116	6.13	2.63	0.71	1.40	60.99
RLA	OVERALL TEST	56	98,216	35.32	11.63	0.93	3.10	64.59
	Reading	29	98,216	18.87	5.57	0.84	2.23	64.55
	Writing	27	98,216	16.45	6.68	0.90	2.14	64.65

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.4.31. Spring 2025 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	128,039	16.96	6.54	0.80	2.90	39.15
	Numerical Representations and Relationships	11	128,039	4.83	2.10	0.50	1.49	45.21
	Computations and Algebraic Relationships	15	128,039	5.75	2.59	0.56	1.72	37.38
	Geometry and Measurement	7	128,039	2.63	1.40	0.39	1.09	34.78
	Data Analysis and Personal Financial Literacy	10	128,039	3.74	2.09	0.52	1.45	37.54
RLA	OVERALL TEST	56	128,168	21.40	9.97	0.89	3.29	41.34
	Reading	29	128,168	12.57	5.17	0.78	2.42	42.86
	Writing	27	128,168	8.83	5.54	0.84	2.22	39.05

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.4.32. Spring 2025 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	262,023	25.02	8.99	0.90	2.80	57.24
	Numerical Representations and Relationships	11	262,023	6.77	2.48	0.67	1.43	62.56
	Computations and Algebraic Relationships	15	262,023	8.48	3.45	0.77	1.64	54.86
	Geometry and Measurement	7	262,023	3.81	1.69	0.58	1.09	50.90
	Data Analysis and Personal Financial Literacy	10	262,023	5.97	2.60	0.70	1.41	59.47
RLA	OVERALL TEST	56	268,329	35.89	10.80	0.92	3.12	65.18
	Reading	29	268,329	18.98	5.22	0.81	2.25	64.83
	Writing	27	268,329	16.91	6.26	0.88	2.14	65.70

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.33. Spring 2025 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	299,218	21.45	9.92	0.91	3.03	46.83
	Probablility and Numerical Representations	7	299,218	3.46	1.79	0.54	1.21	49.41
	Computations and Algebraic Relationships	17	299,218	7.55	3.67	0.74	1.86	46.18
	Geometry and Measurement	15	299,218	6.66	3.91	0.82	1.67	44.32
	Data Analysis and Personal Financial Literacy	7	299,218	3.77	1.89	0.58	1.22	52.11
RLA	OVERALL TEST	56	398,245	31.89	12.34	0.93	3.21	60.34
	Reading	29	398,245	16.95	6.11	0.86	2.29	57.88
	Writing	27	398,245	14.93	6.81	0.89	2.24	64.03

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.4.34. Spring 2025 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	147,755	21.00	9.62	0.90	3.05	45.84
	Probablility and Numerical Representations	7	147,755	3.31	1.75	0.51	1.23	47.54
	Computations and Algebraic Relationships	17	147,755	7.40	3.58	0.73	1.86	45.18
	Geometry and Measurement	15	147,755	6.57	3.84	0.81	1.68	43.69
	Data Analysis and Personal Financial Literacy	7	147,755	3.72	1.85	0.56	1.23	51.25
RLA	OVERALL TEST	56	194,890	33.57	11.95	0.93	3.20	62.75
	Reading	29	194,890	17.58	5.91	0.85	2.28	59.95
	Writing	27	194,890	16.00	6.62	0.89	2.23	66.96

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.4.35. Spring 2025 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	151,360	21.88	10.18	0.91	3.01	47.81
	Probablility and Numerical Representations	7	151,360	3.61	1.82	0.57	1.19	51.24
	Computations and Algebraic Relationships	17	151,360	7.70	3.75	0.76	1.85	47.16
	Geometry and Measurement	15	151,360	6.74	3.98	0.82	1.67	44.94
	Data Analysis and Personal Financial Literacy	7	151,360	3.83	1.92	0.61	1.21	52.95
RLA	OVERALL TEST	56	203,258	30.27	12.49	0.93	3.21	58.04
	Reading	29	203,258	16.36	6.24	0.86	2.30	55.91
	Writing	27	203,258	13.91	6.83	0.89	2.23	61.23

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.36. Spring 2025 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	40,304	18.36	8.46	0.87	3.06	40.09
	Probablility and Numerical Representations	7	40,304	3.03	1.65	0.44	1.23	43.80
	Computations and Algebraic Relationships	17	40,304	6.51	3.24	0.67	1.85	39.79
	Geometry and Measurement	15	40,304	5.49	3.37	0.75	1.68	36.48
	Data Analysis and Personal Financial Literacy	7	40,304	3.33	1.76	0.48	1.26	45.82
RLA	OVERALL TEST	56	49,716	29.07	11.81	0.92	3.27	55.87
	Reading	29	49,716	15.74	5.88	0.84	2.34	53.72
	Writing	27	49,716	13.33	6.55	0.88	2.27	59.09

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 2025 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	161,588	19.77	8.99	0.88	3.07	43.20
	Probablility and Numerical Representations	7	161,588	3.19	1.67	0.46	1.23	45.57
	Computations and Algebraic Relationships	17	161,588	7.01	3.40	0.70	1.86	42.87
	Geometry and Measurement	15	161,588	6.07	3.61	0.78	1.69	40.46
	Data Analysis and Personal Financial Literacy	7	161,588	3.51	1.81	0.52	1.25	48.38
RLA	OVERALL TEST	56	211,061	29.61	12.06	0.93	3.27	56.43
	Reading	29	211,061	15.85	6.03	0.85	2.33	54.07
	Writing	27	211,061	13.75	6.63	0.88	2.28	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.4.38. Spring 2025 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	71,091	24.49	10.02	0.91	3.01	53.44
	Probablility and Numerical Representations	7	71,091	3.96	1.80	0.56	1.19	56.09
	Computations and Algebraic Relationships	17	71,091	8.52	3.64	0.74	1.86	52.18
	Geometry and Measurement	15	71,091	7.75	4.00	0.83	1.67	51.48
	Data Analysis and Personal Financial Literacy	7	71,091	4.26	1.88	0.61	1.18	58.88
RLA	OVERALL TEST	56	98,711	35.67	11.43	0.93	3.13	66.90
	Reading	29	98,711	18.81	5.67	0.85	2.23	64.31
	Writing	27	98,711	16.86	6.37	0.88	2.18	70.80

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.4.39. Spring 2025 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	107,394	16.48	7.26	0.82	3.07	36.05
	Probablility and Numerical Representations	7	107,394	2.73	1.50	0.34	1.22	39.28
	Computations and Algebraic Relationships	17	107,394	5.87	2.90	0.58	1.87	35.77
	Geometry and Measurement	15	107,394	4.94	3.01	0.69	1.69	33.17
	Data Analysis and Personal Financial Literacy	7	107,394	2.94	1.63	0.40	1.26	40.47
RLA	OVERALL TEST	56	120,535	21.89	10.14	0.89	3.31	43.37
	Reading	29	120,535	12.35	5.39	0.81	2.37	42.05
	Writing	27	120,535	9.54	5.44	0.82	2.29	45.36

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.4.40. Spring 2025 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	191,824	24.22	10.12	0.91	3.00	52.87
	Probablility and Numerical Representations	7	191,824	3.88	1.81	0.56	1.20	55.08
	Computations and Algebraic Relationships	17	191,824	8.49	3.72	0.76	1.84	52.01
	Geometry and Measurement	15	191,824	7.62	4.03	0.83	1.66	50.56
	Data Analysis and Personal Financial Literacy	7	191,824	4.24	1.86	0.59	1.19	58.63
RLA	OVERALL TEST	56	277,710	36.23	10.56	0.91	3.14	67.71
	Reading	29	277,710	18.95	5.26	0.82	2.24	64.76
	Writing	27	277,710	17.27	5.96	0.87	2.19	72.14

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.4.41. Spring 2025 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	372,214	24.16	11.00	0.92	3.08	50.32
	Numerical Representations and Relationships	3	372,214	1.51	1.00	0.43	0.76	50.43
	Computations and Algebraic Relationships	17	372,214	8.75	4.29	0.82	1.82	51.97
	Geometry and Measurement	20	372,214	9.18	4.88	0.84	1.95	46.18
	Data Analysis and Personal Financial Literacy	8	372,214	4.72	2.17	0.68	1.23	57.17
RLA	OVERALL TEST	56	397,825	32.13	12.30	0.93	3.22	58.18
	Reading	30	397,825	17.68	5.99	0.85	2.31	59.48
	Writing	26	397,825	14.46	6.90	0.89	2.24	56.03
Science	OVERALL TEST	46	393,051	24.37	9.60	0.90	3.06	53.85
	Matter and Energy	11	393,051	5.97	2.92	0.76	1.43	58.66
	Force, Motion, and Energy	11	393,051	4.56	2.45	0.68	1.38	43.50
	Earth and Space	12	393,051	7.06	2.84	0.69	1.58	56.27
	Organisms and Environments	12	393,051	6.77	2.94	0.74	1.51	56.40
Social Studies	OVERALL TEST	49	401,346	23.63	10.18	0.91	3.14	49.02
	History	19	401,346	10.09	4.39	0.84	1.78	52.23
	Geography and Culture	10	401,346	4.89	2.34	0.70	1.28	51.33
	Government and Citizenship	13	401,346	5.16	3.03	0.66	1.77	41.86
	Economics, Science, Technology, and Society	7	401,346	3.48	1.78	0.57	1.17	49.30

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.4.42. Spring 2025 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	180,934	24.27	10.74	0.92	3.08	50.45
	Numerical Representations and Relationships	3	180,934	1.50	0.98	0.41	0.76	49.85
	Computations and Algebraic Relationships	17	180,934	8.92	4.20	0.81	1.81	52.81
	Geometry and Measurement	20	180,934	9.13	4.82	0.84	1.95	45.80
	Data Analysis and Personal Financial Literacy	8	180,934	4.73	2.10	0.66	1.22	57.25
RLA	OVERALL TEST	56	193,864	34.01	11.88	0.93	3.20	60.93
	Reading	30	193,864	18.49	5.80	0.85	2.28	62.13
	Writing	26	193,864	15.53	6.67	0.89	2.23	58.93
Science	OVERALL TEST	46	191,944	23.93	9.37	0.89	3.08	52.78
	Matter and Energy	11	191,944	6.07	2.88	0.75	1.44	59.43
	Force, Motion, and Energy	11	191,944	4.28	2.35	0.65	1.38	40.68
	Earth and Space	12	191,944	6.91	2.81	0.68	1.60	54.85
	Organisms and Environments	12	191,944	6.67	2.91	0.73	1.52	55.61
Social Studies	OVERALL TEST	49	196,051	23.21	9.92	0.90	3.14	48.10
	History	19	196,051	9.86	4.32	0.83	1.79	50.96
	Geography and Culture	10	196,051	4.88	2.30	0.69	1.29	51.39
	Government and Citizenship	13	196,051	5.11	2.99	0.65	1.77	41.40
	Economics, Science, Technology, and Society	7	196,051	3.36	1.72	0.53	1.18	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.4.43. Spring 2025 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	191,190	24.06	11.25	0.93	3.07	50.20
	Numerical Representations and Relationships	3	191,190	1.53	1.01	0.45	0.75	50.99
	Computations and Algebraic Relationships	17	191,190	8.59	4.37	0.83	1.82	51.20
	Geometry and Measurement	20	191,190	9.23	4.93	0.84	1.94	46.54
	Data Analysis and Personal Financial Literacy	8	191,190	4.71	2.23	0.69	1.23	57.10
RLA	OVERALL TEST	56	203,876	30.35	12.43	0.93	3.24	55.57
	Reading	30	203,876	16.91	6.06	0.85	2.33	56.97
	Writing	26	203,876	13.44	6.97	0.90	2.25	53.27
Science	OVERALL TEST	46	201,023	24.79	9.81	0.90	3.03	54.87
	Matter and Energy	11	201,023	5.88	2.95	0.77	1.42	57.94
	Force, Motion, and Energy	11	201,023	4.82	2.51	0.70	1.38	46.19
	Earth and Space	12	201,023	7.21	2.85	0.70	1.56	57.64
	Organisms and Environments	12	201,023	6.88	2.96	0.74	1.50	57.16
Social Studies	OVERALL TEST	49	205,200	24.03	10.40	0.91	3.12	49.90
	History	19	205,200	10.32	4.44	0.84	1.77	53.44
	Geography and Culture	10	205,200	4.91	2.39	0.71	1.28	51.27
	Government and Citizenship	13	205,200	5.21	3.07	0.67	1.77	42.32
	Economics, Science, Technology, and Society	7	205,200	3.59	1.84	0.61	1.15	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.4.44. Spring 2025 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	48,320	20.65	9.75	0.90	3.12	42.86
	Numerical Representations and Relationships	3	48,320	1.26	0.95	0.34	0.77	42.10
	Computations and Algebraic Relationships	17	48,320	7.64	4.06	0.80	1.83	45.14
	Geometry and Measurement	20	48,320	7.65	4.21	0.78	1.96	38.61
	Data Analysis and Personal Financial Literacy	8	48,320	4.10	2.05	0.62	1.27	48.89
RLA	OVERALL TEST	56	50,394	29.18	11.86	0.92	3.28	53.62
	Reading	30	50,394	16.44	5.83	0.84	2.35	55.41
	Writing	26	50,394	12.75	6.69	0.88	2.28	50.67
Science	OVERALL TEST	46	50,388	21.02	8.84	0.88	3.12	46.61
	Matter and Energy	11	50,388	5.04	2.72	0.71	1.47	50.14
	Force, Motion, and Energy	11	50,388	3.81	2.16	0.59	1.38	36.42
	Earth and Space	12	50,388	6.28	2.79	0.65	1.64	50.04
	Organisms and Environments	12	50,388	5.90	2.85	0.70	1.55	49.19
Social Studies	OVERALL TEST	49	50,788	21.08	9.35	0.89	3.15	43.76
	History	19	50,788	9.02	4.09	0.80	1.82	46.49
	Geography and Culture	10	50,788	4.35	2.21	0.66	1.29	46.14
	Government and Citizenship	13	50,788	4.59	2.81	0.61	1.76	37.27
	Economics, Science, Technology, and Society	7	50,788	3.11	1.70	0.52	1.18	44.12

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.45. Spring 2025 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	199,455	22.35	10.18	0.91	3.12	46.46
	Numerical Representations and Relationships	3	199,455	1.42	0.97	0.38	0.77	47.45
	Computations and Algebraic Relationships	17	199,455	8.20	4.13	0.80	1.84	48.54
	Geometry and Measurement	20	199,455	8.32	4.43	0.80	1.97	41.81
	Data Analysis and Personal Financial Literacy	8	199,455	4.40	2.11	0.65	1.25	53.21
RLA	OVERALL TEST	56	210,775	29.82	12.10	0.93	3.27	54.33
	Reading	30	210,775	16.60	5.93	0.84	2.35	55.91
	Writing	26	210,775	13.21	6.78	0.89	2.27	51.73
Science	OVERALL TEST	46	207,558	22.56	9.08	0.88	3.11	49.90
	Matter and Energy	11	207,558	5.48	2.81	0.73	1.45	54.32
	Force, Motion, and Energy	11	207,558	4.15	2.26	0.62	1.39	39.48
	Earth and Space	12	207,558	6.69	2.79	0.67	1.61	53.33
	Organisms and Environments	12	207,558	6.25	2.86	0.71	1.54	51.88
Social Studies	OVERALL TEST	49	213,178	21.51	9.34	0.89	3.15	44.66
	History	19	213,178	9.19	4.10	0.81	1.81	47.41
	Geography and Culture	10	213,178	4.51	2.22	0.66	1.30	47.51
	Government and Citizenship	13	213,178	4.64	2.78	0.60	1.77	37.76
	Economics, Science, Technology, and Society	7	213,178	3.17	1.71	0.53	1.18	45.01

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.4.46. Spring 2025 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	90,966	27.47	10.96	0.92	3.04	57.43
	Numerical Representations and Relationships	3	90,966	1.70	1.00	0.45	0.74	56.62
	Computations and Algebraic Relationships	17	90,966	9.69	4.16	0.81	1.81	58.08
	Geometry and Measurement	20	90,966	10.73	4.98	0.85	1.93	53.97
	Data Analysis and Personal Financial Literacy	8	90,966	5.35	2.11	0.68	1.19	65.44
RLA	OVERALL TEST	56	98,639	36.20	11.19	0.92	3.14	64.85
	Reading	30	98,639	19.57	5.48	0.83	2.23	65.77
	Writing	26	98,639	16.63	6.31	0.88	2.20	63.35
Science	OVERALL TEST	46	97,852	27.84	9.11	0.89	2.98	61.37
	Matter and Energy	11	97,852	6.96	2.80	0.76	1.38	67.40
	Force, Motion, and Energy	11	97,852	5.31	2.49	0.69	1.37	50.74
	Earth and Space	12	97,852	7.76	2.67	0.67	1.53	61.69
	Organisms and Environments	12	97,852	7.81	2.71	0.71	1.45	65.18
Social Studies	OVERALL TEST	49	99,129	27.14	10.12	0.90	3.12	56.25
	History	19	99,129	11.62	4.27	0.84	1.73	60.31
	Geography and Culture	10	99,129	5.53	2.32	0.70	1.27	57.64
	Government and Citizenship	13	99,129	5.98	3.15	0.68	1.79	48.34
	Economics, Science, Technology, and Society	7	99,129	4.01	1.75	0.58	1.14	56.75

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.4.47. Spring 2025 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	111,972	17.81	8.51	0.87	3.12	36.93
	Numerical Representations and Relationships	3	111,972	1.15	0.93	0.30	0.78	38.19
	Computations and Algebraic Relationships	17	111,972	6.52	3.67	0.75	1.83	38.54
	Geometry and Measurement	20	111,972	6.60	3.61	0.71	1.95	33.12
	Data Analysis and Personal Financial Literacy	8	111,972	3.55	1.96	0.57	1.28	42.41
RLA	OVERALL TEST	56	114,420	22.01	10.44	0.90	3.35	41.88
	Reading	30	114,420	13.16	5.51	0.81	2.43	44.41
	Writing	26	114,420	8.85	5.67	0.84	2.30	37.70
Science	OVERALL TEST	46	108,697	17.86	7.80	0.84	3.13	39.65
	Matter and Energy	11	108,697	4.12	2.50	0.65	1.47	41.41
	Force, Motion, and Energy	11	108,697	3.30	1.88	0.47	1.37	31.21
	Earth and Space	12	108,697	5.55	2.67	0.61	1.66	44.60
	Organisms and Environments	12	108,697	4.88	2.60	0.64	1.56	40.72
Social Studies	OVERALL TEST	49	109,619	17.37	7.65	0.83	3.13	36.11
	History	19	109,619	7.45	3.51	0.72	1.84	38.25
	Geography and Culture	10	109,619	3.54	1.93	0.55	1.30	37.73
	Government and Citizenship	13	109,619	3.81	2.33	0.45	1.72	31.13
	Economics, Science, Technology, and Society	7	109,619	2.58	1.58	0.45	1.17	36.54

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.4.48. Spring 2025 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	260,242	26.89	10.83	0.92	3.06	56.08
	Numerical Representations and Relationships	3	260,242	1.67	0.99	0.43	0.75	55.70
	Computations and Algebraic Relationships	17	260,242	9.71	4.18	0.81	1.81	57.76
	Geometry and Measurement	20	260,242	10.29	4.93	0.85	1.94	51.79
	Data Analysis and Personal Financial Literacy	8	260,242	5.22	2.06	0.66	1.20	63.51
RLA	OVERALL TEST	56	283,405	36.22	10.51	0.91	3.16	64.76
	Reading	30	283,405	19.50	5.15	0.81	2.25	65.57
	Writing	26	283,405	16.72	6.00	0.87	2.20	63.42
Science	OVERALL TEST	46	284,354	26.85	9.05	0.89	3.02	59.27
	Matter and Energy	11	284,354	6.68	2.75	0.74	1.41	65.25
	Force, Motion, and Energy	11	284,354	5.04	2.47	0.69	1.38	48.19
	Earth and Space	12	284,354	7.64	2.69	0.67	1.55	60.73
	Organisms and Environments	12	284,354	7.50	2.73	0.71	1.48	62.40
Social Studies	OVERALL TEST	49	291,727	25.98	10.01	0.90	3.13	53.87
	History	19	291,727	11.09	4.27	0.83	1.75	57.48
	Geography and Culture	10	291,727	5.40	2.28	0.69	1.27	56.44
	Government and Citizenship	13	291,727	5.67	3.11	0.67	1.78	45.90
	Economics, Science, Technology, and Society	7	291,727	3.81	1.74	0.56	1.15	54.09

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 2025 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	21,318	16.93	7.32	0.87	2.64	42.95
	Numerical Representations and Relationships	12	21,318	6.03	2.67	0.69	1.49	48.24
	Computations and Algebraic Relationships	14	21,318	5.42	3.50	0.78	1.65	37.32
	Geometry and Measurement	7	21,318	3.95	1.48	0.47	1.08	50.71
	Data Analysis and Personal Financial Literacy	4	21,318	1.53	1.00	0.29	0.84	38.26
RLA	OVERALL TEST	52	34,164	23.46	10.04	0.90	3.10	50.46
	Reading	26	34,164	14.05	5.55	0.82	2.32	53.64
	Writing	26	34,164	9.40	5.21	0.85	2.04	45.96

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.4.50. Spring 2025 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	10,796	16.20	7.03	0.86	2.65	41.09
	Numerical Representations and Relationships	12	10,796	5.68	2.62	0.67	1.51	45.65
	Computations and Algebraic Relationships	14	10,796	5.15	3.36	0.76	1.64	35.42
	Geometry and Measurement	7	10,796	3.87	1.44	0.43	1.09	49.42
	Data Analysis and Personal Financial Literacy	4	10,796	1.50	0.98	0.27	0.84	37.45
RLA	OVERALL TEST	52	17,348	24.05	10.02	0.90	3.11	51.57
	Reading	26	17,348	14.30	5.48	0.82	2.32	54.59
	Writing	26	17,348	9.75	5.27	0.85	2.06	47.31

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.51. Spring 2025 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	10,510	17.69	7.53	0.88	2.63	44.88
	Numerical Representations and Relationships	12	10,510	6.39	2.66	0.70	1.46	50.92
	Computations and Algebraic Relationships	14	10,510	5.71	3.60	0.79	1.66	39.28
	Geometry and Measurement	7	10,510	4.03	1.52	0.50	1.07	52.05
	Data Analysis and Personal Financial Literacy	4	10,510	1.56	1.01	0.30	0.84	39.10
RLA	OVERALL TEST	52	16,804	22.85	10.02	0.90	3.09	49.32
	Reading	26	16,804	13.80	5.61	0.83	2.33	52.67
	Writing	26	16,804	9.05	5.13	0.85	2.02	44.58

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.4.52. Spring 2025 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	23	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	12	23	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	14	23	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	7	23	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	4	23	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	30	N/A	N/A	N/A	N/A	N/A
	Reading	26	30	N/A	N/A	N/A	N/A	N/A
	Writing	26	30	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 2025 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	20,840	16.94	7.30	0.87	2.64	42.96
	Numerical Representations and Relationships	12	20,840	6.03	2.66	0.69	1.49	48.27
	Computations and Algebraic Relationships	14	20,840	5.42	3.49	0.78	1.65	37.32
	Geometry and Measurement	7	20,840	3.95	1.48	0.47	1.08	50.72
	Data Analysis and Personal Financial Literacy	4	20,840	1.53	1.00	0.29	0.84	38.27
RLA	OVERALL TEST	52	33,385	23.47	10.04	0.90	3.10	50.49
	Reading	26	33,385	14.07	5.55	0.82	2.32	53.68
	Writing	26	33,385	9.41	5.21	0.85	2.04	45.98

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.54. Spring 2025 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	288	17.45	8.08	0.90	2.60	44.30
	Numerical Representations and Relationships	12	288	6.13	2.80	0.72	1.48	48.82
	Computations and Algebraic Relationships	14	288	5.66	3.81	0.82	1.61	39.03
	Geometry and Measurement	7	288	4.06	1.55	0.48	1.12	52.15
	Data Analysis and Personal Financial Literacy	4	288	1.60	1.09	0.44	0.82	40.10
RLA	OVERALL TEST	52	430	23.14	10.43	0.91	3.10	49.53
	Reading	26	430	13.67	5.73	0.83	2.33	52.28
	Writing	26	430	9.47	5.44	0.86	2.03	45.64

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.55. Spring 2025 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	17,140	16.70	7.29	0.87	2.64	42.37
	Numerical Representations and Relationships	12	17,140	5.94	2.66	0.69	1.49	47.54
	Computations and Algebraic Relationships	14	17,140	5.33	3.48	0.78	1.65	36.73
	Geometry and Measurement	7	17,140	3.92	1.48	0.46	1.08	50.26
	Data Analysis and Personal Financial Literacy	4	17,140	1.51	1.00	0.29	0.84	37.78
RLA	OVERALL TEST	52	9,317	17.52	7.80	0.85	3.07	38.76
	Reading	26	9,317	10.96	4.73	0.75	2.37	41.81
	Writing	26	9,317	6.55	3.88	0.75	1.93	34.45

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.56. Spring 2025 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	4,178	17.89	7.36	0.87	2.63	45.35
	Numerical Representations and Relationships	12	4,178	6.39	2.65	0.69	1.47	51.11
	Computations and Algebraic Relationships	14	4,178	5.81	3.53	0.78	1.67	39.74
	Geometry and Measurement	7	4,178	4.08	1.47	0.48	1.06	52.55
	Data Analysis and Personal Financial Literacy	4	4,178	1.61	1.00	0.29	0.84	40.23
RLA	OVERALL TEST	52	24,847	25.69	9.87	0.90	3.10	54.85
	Reading	26	24,847	15.21	5.39	0.82	2.30	58.08
	Writing	26	24,847	10.47	5.25	0.84	2.07	50.28

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.4.57. Spring 2025 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	15,379	17.56	7.70	0.87	2.81	41.28
	Numerical Representations and Relationships	10	15,379	4.42	2.45	0.69	1.36	41.87
	Computations and Algebraic Relationships	12	15,379	5.29	2.76	0.66	1.60	41.86
	Geometry and Measurement	13	15,379	5.17	2.54	0.63	1.55	37.17
	Data Analysis and Personal Financial Literacy	5	15,379	2.68	1.40	0.44	1.04	47.76
RLA	OVERALL TEST	52	26,357	25.10	10.35	0.91	3.09	52.52
	Reading	26	26,357	13.94	5.44	0.82	2.31	53.29
	Writing	26	26,357	11.16	5.58	0.87	2.04	51.44

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.4.58. Spring 2025 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	7,727	16.81	7.26	0.85	2.81	39.20
	Numerical Representations and Relationships	10	7,727	4.21	2.36	0.67	1.36	39.71
	Computations and Algebraic Relationships	12	7,727	4.98	2.67	0.64	1.60	39.22
	Geometry and Measurement	13	7,727	4.93	2.39	0.58	1.55	34.97
	Data Analysis and Personal Financial Literacy	5	7,727	2.68	1.35	0.41	1.04	47.53
RLA	OVERALL TEST	52	13,402	25.88	10.18	0.91	3.10	53.80
	Reading	26	13,402	14.21	5.31	0.81	2.31	54.38
	Writing	26	13,402	11.66	5.56	0.86	2.05	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.4.59. Spring 2025 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	7,642	18.33	8.05	0.88	2.81	43.40
	Numerical Representations and Relationships	10	7,642	4.63	2.52	0.71	1.36	44.07
	Computations and Algebraic Relationships	12	7,642	5.62	2.82	0.68	1.60	44.55
	Geometry and Measurement	13	7,642	5.41	2.66	0.66	1.56	39.41
	Data Analysis and Personal Financial Literacy	5	7,642	2.68	1.44	0.47	1.05	48.02
RLA	OVERALL TEST	52	12,945	24.30	10.45	0.91	3.07	51.20
	Reading	26	12,945	13.66	5.56	0.83	2.30	52.16
	Writing	26	12,945	10.64	5.55	0.87	2.02	49.85

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.4.60. Spring 2025 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	13	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	10	13	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	12	13	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	13	13	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	5	13	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	11	N/A	N/A	N/A	N/A	N/A
	Reading	26	11	N/A	N/A	N/A	N/A	N/A
	Writing	26	11	N/A	N/A	N/A	N/A	N/A

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Statistics not calculated due to small n-counts.

Table B.4.61. Spring 2025 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	15,023	17.59	7.70	0.87	2.81	41.35
	Numerical Representations and Relationships	10	15,023	4.43	2.45	0.69	1.36	41.94
	Computations and Algebraic Relationships	12	15,023	5.30	2.76	0.66	1.60	41.93
	Geometry and Measurement	13	15,023	5.17	2.54	0.63	1.55	37.22
	Data Analysis and Personal Financial Literacy	5	15,023	2.69	1.40	0.44	1.04	47.89
RLA	OVERALL TEST	52	25,806	25.15	10.34	0.91	3.09	52.62
	Reading	26	25,806	13.97	5.44	0.82	2.31	53.38
	Writing	26	25,806	11.19	5.58	0.87	2.04	51.55

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.62. Spring 2025 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	196	16.80	7.98	0.88	2.82	39.45
	Numerical Representations and Relationships	10	196	4.30	2.46	0.70	1.35	40.70
	Computations and Algebraic Relationships	12	196	5.04	2.92	0.70	1.59	39.62
	Geometry and Measurement	13	196	5.01	2.55	0.61	1.59	35.97
	Data Analysis and Personal Financial Literacy	5	196	2.46	1.51	0.52	1.05	44.01
RLA	OVERALL TEST	52	335	22.98	10.19	0.91	3.09	48.51
	Reading	26	335	13.02	5.58	0.83	2.31	49.75
	Writing	26	335	9.96	5.25	0.85	2.03	46.76

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.63. Spring 2025 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	4,345	14.18	5.87	0.77	2.80	32.84
	Numerical Representations and Relationships	10	4,345	3.54	2.08	0.57	1.35	33.25
	Computations and Algebraic Relationships	12	4,345	4.25	2.23	0.48	1.61	33.18
	Geometry and Measurement	13	4,345	4.26	1.98	0.43	1.50	30.07
	Data Analysis and Personal Financial Literacy	5	4,345	2.14	1.34	0.32	1.11	37.30
RLA	OVERALL TEST	52	7,696	19.15	8.66	0.87	3.09	41.09
	Reading	26	7,696	10.95	4.75	0.76	2.34	41.83
	Writing	26	7,696	8.20	4.66	0.81	2.01	40.05

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.4.64. Spring 2025 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	11,034	18.89	7.92	0.87	2.81	44.61
	Numerical Representations and Relationships	10	11,034	4.76	2.50	0.70	1.36	45.26
	Computations and Algebraic Relationships	12	11,034	5.71	2.84	0.69	1.59	45.28
	Geometry and Measurement	13	11,034	5.53	2.64	0.65	1.57	39.96
	Data Analysis and Personal Financial Literacy	5	11,034	2.89	1.36	0.45	1.01	51.88
RLA	OVERALL TEST	52	18,661	27.55	9.98	0.90	3.08	57.23
	Reading	26	18,661	15.18	5.22	0.81	2.29	58.01
	Writing	26	18,661	12.38	5.47	0.86	2.04	56.14

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.4.65. Spring 2025 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	12,208	17.79	8.15	0.88	2.86	39.72
	Numerical Representations and Relationships	9	12,208	4.34	2.24	0.63	1.36	48.08
	Computations and Algebraic Relationships	20	12,208	8.60	4.35	0.80	1.93	39.81
	Geometry and Measurement	9	12,208	3.41	1.80	0.47	1.31	33.24
	Data Analysis and Personal Financial Literacy	4	12,208	1.44	1.21	0.50	0.85	36.05
RLA	OVERALL TEST	52	19,392	27.32	10.61	0.92	3.05	56.33
	Reading	26	19,392	15.18	5.32	0.82	2.26	58.09
	Writing	26	19,392	12.14	5.96	0.88	2.04	53.84
Science	OVERALL TEST	39	14,380	15.65	5.77	0.75	2.88	39.19
	Matter and Energy	6	14,380	2.15	1.31	0.27	1.11	35.79
	Force, Motion, and Energy	9	14,380	3.67	1.93	0.38	1.52	40.09
	Earth and Space	10	14,380	3.40	1.91	0.39	1.50	33.74
	Organisms and Environments	14	14,380	6.43	2.68	0.65	1.59	45.52

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.66. Spring 2025 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	6,126	17.50	7.85	0.87	2.86	38.91
	Numerical Representations and Relationships	9	6,126	4.28	2.21	0.62	1.36	47.41
	Computations and Algebraic Relationships	20	6,126	8.54	4.24	0.79	1.93	39.38
	Geometry and Measurement	9	6,126	3.32	1.74	0.44	1.30	32.00
	Data Analysis and Personal Financial Literacy	4	6,126	1.37	1.18	0.49	0.84	34.30
RLA	OVERALL TEST	52	9,835	28.46	10.56	0.92	3.05	58.10
	Reading	26	9,835	15.52	5.24	0.82	2.25	59.35
	Writing	26	9,835	12.94	6.00	0.88	2.04	56.33
Science	OVERALL TEST	39	7,207	15.28	5.50	0.73	2.86	38.28
	Matter and Energy	6	7,207	2.15	1.33	0.31	1.11	35.75
	Force, Motion, and Energy	9	7,207	3.55	1.82	0.31	1.51	38.67
	Earth and Space	10	7,207	3.26	1.84	0.34	1.49	32.31
	Organisms and Environments	14	7,207	6.33	2.63	0.64	1.58	44.90

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.4.67. Spring 2025 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	6,071	18.09	8.43	0.88	2.86	40.54
	Numerical Representations and Relationships	9	6,071	4.40	2.27	0.64	1.37	48.77
	Computations and Algebraic Relationships	20	6,071	8.67	4.46	0.81	1.93	40.25
	Geometry and Measurement	9	6,071	3.51	1.86	0.50	1.32	34.51
	Data Analysis and Personal Financial Literacy	4	6,071	1.51	1.23	0.51	0.86	37.86
RLA	OVERALL TEST	52	9,547	26.16	10.54	0.92	3.05	54.53
	Reading	26	9,547	14.84	5.39	0.82	2.27	56.82
	Writing	26	9,547	11.32	5.80	0.88	2.03	51.30
Science	OVERALL TEST	39	7,162	16.02	6.00	0.77	2.89	40.12
	Matter and Energy	6	7,162	2.15	1.28	0.24	1.12	35.84
	Force, Motion, and Energy	9	7,162	3.79	2.03	0.43	1.53	41.52
	Earth and Space	10	7,162	3.54	1.97	0.42	1.50	35.19
	Organisms and Environments	14	7,162	6.54	2.73	0.66	1.60	46.15

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.68. Spring 2025 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	12	N/A	N/A	N/A	N/A	N/A
	Numerical Representations and Relationships	9	12	N/A	N/A	N/A	N/A	N/A
	Computations and Algebraic Relationships	20	12	N/A	N/A	N/A	N/A	N/A
	Geometry and Measurement	9	12	N/A	N/A	N/A	N/A	N/A
	Data Analysis and Personal Financial Literacy	4	12	N/A	N/A	N/A	N/A	N/A
RLA	OVERALL TEST	52	15	N/A	N/A	N/A	N/A	N/A
	Reading	26	15	N/A	N/A	N/A	N/A	N/A
	Writing	26	15	N/A	N/A	N/A	N/A	N/A
Science	OVERALL TEST	39	15	N/A	N/A	N/A	N/A	N/A
	Matter and Energy	6	15	N/A	N/A	N/A	N/A	N/A
	Force, Motion, and Energy	9	15	N/A	N/A	N/A	N/A	N/A
	Earth and Space	10	15	N/A	N/A	N/A	N/A	N/A
	Organisms and Environments	14	15	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 2025 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	11,917	17.82	8.15	0.88	2.86	39.77
	Numerical Representations and Relationships	9	11,917	4.35	2.24	0.63	1.36	48.21
	Computations and Algebraic Relationships	20	11,917	8.61	4.36	0.80	1.93	39.85
	Geometry and Measurement	9	11,917	3.41	1.80	0.47	1.31	33.24
	Data Analysis and Personal Financial Literacy	4	11,917	1.45	1.21	0.50	0.85	36.13
RLA	OVERALL TEST	52	18,998	27.35	10.61	0.92	3.05	56.38
	Reading	26	18,998	15.20	5.32	0.82	2.26	58.14
	Writing	26	18,998	12.16	5.96	0.88	2.04	53.90
Science	OVERALL TEST	39	14,025	15.66	5.77	0.75	2.88	39.21
	Matter and Energy	6	14,025	2.14	1.31	0.27	1.11	35.72
	Force, Motion, and Energy	9	14,025	3.67	1.93	0.38	1.52	40.12
	Earth and Space	10	14,025	3.40	1.91	0.38	1.50	33.76
	Organisms and Environments	14	14,025	6.44	2.69	0.65	1.59	45.57

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.4.70. Spring 2025 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	157	17.46	8.28	0.88	2.82	38.96
	Numerical Representations and Relationships	9	157	3.99	2.27	0.65	1.35	43.99
	Computations and Algebraic Relationships	20	157	8.65	4.25	0.80	1.91	40.23
	Geometry and Measurement	9	157	3.38	1.96	0.56	1.30	32.71
	Data Analysis and Personal Financial Literacy	4	157	1.44	1.25	0.55	0.84	35.99
RLA	OVERALL TEST	52	199	26.18	10.19	0.91	3.11	54.42
	Reading	26	199	14.74	5.15	0.80	2.29	56.56
	Writing	26	199	11.43	5.75	0.87	2.10	51.39
Science	OVERALL TEST	39	205	15.73	5.96	0.76	2.89	39.66
	Matter and Energy	6	205	2.32	1.28	0.21	1.14	38.62
	Force, Motion, and Energy	9	205	3.72	2.09	0.47	1.53	40.73
	Earth and Space	10	205	3.45	2.08	0.48	1.50	34.31
	Organisms and Environments	14	205	6.25	2.63	0.62	1.62	44.34

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.4.71. Spring 2025 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	3,075	14.90	6.36	0.80	2.84	32.97
	Numerical Representations and Relationships	9	3,075	3.70	1.95	0.50	1.38	40.89
	Computations and Algebraic Relationships	20	3,075	7.15	3.61	0.71	1.93	33.07
	Geometry and Measurement	9	3,075	2.98	1.50	0.26	1.29	28.38
	Data Analysis and Personal Financial Literacy	4	3,075	1.07	1.03	0.37	0.82	26.72
RLA	OVERALL TEST	52	4,981	21.82	9.37	0.89	3.13	45.93
	Reading	26	4,981	12.54	4.97	0.77	2.36	47.93
	Writing	26	4,981	9.28	5.17	0.84	2.04	43.11
Science	OVERALL TEST	39	3,618	13.93	5.13	0.69	2.84	35.12
	Matter and Energy	6	3,618	1.98	1.25	0.22	1.10	32.92
	Force, Motion, and Energy	9	3,618	3.37	1.81	0.31	1.50	36.97
	Earth and Space	10	3,618	2.98	1.75	0.31	1.46	29.97
	Organisms and Environments	14	3,618	5.60	2.44	0.58	1.58	39.79

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.4.72. Spring 2025 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	9,133	18.77	8.45	0.89	2.86	41.99
	Numerical Representations and Relationships	9	9,133	4.55	2.29	0.65	1.36	50.50
	Computations and Algebraic Relationships	20	9,133	9.09	4.47	0.81	1.93	42.07
	Geometry and Measurement	9	9,133	3.56	1.87	0.51	1.32	34.87
	Data Analysis and Personal Financial Literacy	4	9,133	1.57	1.23	0.51	0.86	39.20
RLA	OVERALL TEST	52	14,411	29.22	10.35	0.92	3.02	59.93
	Reading	26	14,411	16.10	5.13	0.81	2.22	61.61
	Writing	26	14,411	13.13	5.89	0.88	2.04	57.56
Science	OVERALL TEST	39	10,762	16.23	5.86	0.76	2.89	40.56
	Matter and Energy	6	10,762	2.21	1.32	0.29	1.12	36.75
	Force, Motion, and Energy	9	10,762	3.77	1.96	0.39	1.53	41.14
	Earth and Space	10	10,762	3.53	1.95	0.40	1.51	35.00
	Organisms and Environments	14	10,762	6.72	2.71	0.65	1.59	47.44

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.4.73. Spring 2025 STAAR 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	455,622	31.57	13.42	0.94	3.38	53.90
	Number and Algebraic Methods	12	455,622	6.08	2.95	0.74	1.49	49.57
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	455,622	6.94	3.41	0.79	1.57	55.33
	Writing and Solving Linear Functions, Equations, and Inequalities	15	455,622	8.16	3.75	0.79	1.73	53.72
	Quadratic Functions and Equations	12	455,622	6.72	3.04	0.74	1.54	56.83
	Exponential Functions and Equations	7	455,622	3.67	1.96	0.67	1.13	53.98

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.4.74. Spring 2025 STAAR 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	219,684	32.20	13.06	0.93	3.38	54.98
	Number and Algebraic Methods	12	219,684	6.27	2.91	0.74	1.49	51.04
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	219,684	7.03	3.37	0.78	1.58	56.19
	Writing and Solving Linear Functions, Equations, and Inequalities	15	219,684	8.41	3.63	0.77	1.73	55.25
	Quadratic Functions and Equations	12	219,684	6.85	2.98	0.73	1.55	57.92
	Exponential Functions and Equations	7	219,684	3.63	1.90	0.65	1.13	53.86

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.4.75. Spring 2025 STAAR 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	235,836	31.00	13.71	0.94	3.37	52.89
	Number and Algebraic Methods	12	235,836	5.90	2.97	0.75	1.49	48.20
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	235,836	6.86	3.44	0.79	1.56	54.53
	Writing and Solving Linear Functions, Equations, and Inequalities	15	235,836	7.93	3.84	0.80	1.72	52.30
	Quadratic Functions and Equations	12	235,836	6.60	3.08	0.75	1.54	55.81
	Exponential Functions and Equations	7	235,836	3.70	2.02	0.68	1.14	54.09

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 2025 STAAR 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	61,140	27.61	12.30	0.92	3.44	47.21
	Number and Algebraic Methods	12	61,140	5.36	2.78	0.71	1.50	43.48
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	61,140	5.92	3.11	0.74	1.58	47.71
	Writing and Solving Linear Functions, Equations, and Inequalities	15	61,140	7.20	3.53	0.75	1.78	47.52
	Quadratic Functions and Equations	12	61,140	6.00	2.88	0.70	1.57	50.75
	Exponential Functions and Equations	7	61,140	3.13	1.85	0.62	1.15	45.96

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.4.77. Spring 2025 STAAR 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	245,746	29.64	12.65	0.93	3.43	50.68
	Number and Algebraic Methods	12	245,746	5.74	2.80	0.71	1.50	46.73
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	245,746	6.42	3.24	0.76	1.58	51.48
	Writing and Solving Linear Functions, Equations, and Inequalities	15	245,746	7.70	3.61	0.76	1.76	50.77
	Quadratic Functions and Equations	12	245,746	6.35	2.93	0.71	1.57	53.75
	Exponential Functions and Equations	7	245,746	3.42	1.88	0.63	1.15	50.53

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.4.78. Spring 2025 STAAR 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	107,309	35.07	13.12	0.94	3.34	59.69
	Number and Algebraic Methods	12	107,309	6.62	2.90	0.74	1.49	54.03
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	107,309	7.98	3.35	0.78	1.56	62.97
	Writing and Solving Linear Functions, Equations, and Inequalities	15	107,309	8.97	3.66	0.79	1.69	58.74
	Quadratic Functions and Equations	12	107,309	7.37	2.97	0.74	1.52	62.28
	Exponential Functions and Equations	7	107,309	4.14	1.95	0.67	1.12	60.86

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.4.79. Spring 2025 STAAR 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	126,502	23.99	10.49	0.89	3.46	41.17
	Number and Algebraic Methods	12	126,502	4.78	2.49	0.64	1.50	38.76
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	126,502	5.08	2.71	0.66	1.57	41.15
	Writing and Solving Linear Functions, Equations, and Inequalities	15	126,502	6.23	3.20	0.68	1.80	41.46
	Quadratic Functions and Equations	12	126,502	5.17	2.55	0.61	1.59	43.84
	Exponential Functions and Equations	7	126,502	2.74	1.66	0.51	1.17	40.13

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.4.80. Spring 2025 STAAR 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	329,120	34.49	13.28	0.94	3.34	58.79
	Number and Algebraic Methods	12	329,120	6.58	2.96	0.75	1.48	53.72
	Describing and Graphing Linear Functions, Equations, and Inequalities	13	329,120	7.66	3.38	0.79	1.57	60.77
	Writing and Solving Linear Functions, Equations, and Inequalities	15	329,120	8.90	3.68	0.79	1.70	58.43
	Quadratic Functions and Equations	12	329,120	7.32	3.00	0.74	1.53	61.82
	Exponential Functions and Equations	7	329,120	4.03	1.95	0.67	1.12	59.30

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.4.81. Spring 2025 STAAR 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	483,644	36.16	15.08	0.95	3.37	57.83
	Reading	32	483,644	18.04	6.97	0.88	2.41	55.56
	Writing	32	483,644	18.12	8.67	0.93	2.35	60.70

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.82. Spring 2025 STAAR 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	227,584	38.58	14.79	0.95	3.34	61.08
	Reading	32	227,584	18.99	6.88	0.88	2.38	58.49
	Writing	32	227,584	19.58	8.45	0.92	2.33	64.34

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.83. Spring 2025 STAAR 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	255,951	34.02	15.00	0.95	3.40	54.95
	Reading	32	255,951	17.20	6.94	0.88	2.43	52.97
	Writing	32	255,951	16.82	8.64	0.93	2.36	57.46

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.84. Spring 2025 STAAR 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	63,803	32.59	14.13	0.94	3.44	52.78
	Reading	32	63,803	16.55	6.58	0.86	2.46	50.90
	Writing	32	63,803	16.05	8.19	0.91	2.39	55.14

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.4.85. Spring 2025 STAAR 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	269,673	33.17	14.57	0.94	3.44	53.41
	Reading	32	269,673	16.68	6.75	0.87	2.45	51.26
	Writing	32	269,673	16.49	8.41	0.92	2.40	56.12

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.86. Spring 2025 STAAR 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	108,625	42.54	13.67	0.94	3.24	67.21
	Reading	32	108,625	20.92	6.35	0.87	2.33	64.66
	Writing	32	108,625	21.62	7.89	0.92	2.24	70.41

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.4.87. Spring 2025 STAAR 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	154,640	24.93	11.55	0.91	3.52	41.47
	Reading	32	154,640	13.20	5.65	0.80	2.50	40.29
	Writing	32	154,640	11.73	6.65	0.86	2.46	42.95

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 2025 STAAR 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	329,004	41.44	13.58	0.94	3.29	65.53
	Reading	32	329,004	20.32	6.34	0.86	2.36	62.75
	Writing	32	329,004	21.12	7.84	0.92	2.27	69.04

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 2025 STAAR 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	464,246	36.46	14.17	0.94	3.39	59.21
	Reading	33	464,246	19.87	6.80	0.88	2.40	59.92
	Writing	31	464,246	16.59	8.00	0.91	2.38	58.24

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.90. Spring 2025 STAAR 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	222,170	38.17	13.72	0.94	3.38	61.37
	Reading	33	222,170	20.48	6.51	0.87	2.39	61.68
	Writing	31	222,170	17.69	7.82	0.91	2.38	60.93

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.4.91. Spring 2025 STAAR 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	241,978	34.90	14.39	0.94	3.40	57.23
	Reading	33	241,978	19.31	7.01	0.88	2.42	58.31
	Writing	31	241,978	15.59	8.02	0.91	2.38	55.76

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.92. Spring 2025 STAAR 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	60,987	32.99	13.23	0.93	3.47	54.34
	Reading	33	60,987	18.37	6.49	0.86	2.47	55.38
	Writing	31	60,987	14.62	7.45	0.89	2.42	52.92

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.4.93. Spring 2025 STAAR 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	257,583	33.78	13.79	0.94	3.45	55.29
	Reading	33	257,583	18.68	6.77	0.87	2.45	56.28
	Writing	31	257,583	15.10	7.68	0.90	2.41	53.95

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 2025 STAAR 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	106,068	42.22	12.93	0.94	3.27	67.61
	Reading	33	106,068	22.48	6.05	0.86	2.29	67.90
	Writing	31	106,068	19.74	7.52	0.90	2.32	67.21

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.4.95. Spring 2025 STAAR 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	134,001	25.13	11.02	0.90	3.53	42.88
	Reading	33	134,001	14.75	6.02	0.82	2.54	44.46
	Writing	31	134,001	10.38	5.78	0.82	2.43	40.72

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.4.96. Spring 2025 STAAR 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	330,245	41.06	12.64	0.93	3.32	65.84
	Reading	33	330,245	21.95	5.94	0.85	2.33	66.20
	Writing	31	330,245	19.12	7.37	0.90	2.35	65.35

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.4.97. Spring 2025 STAAR 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	423,312	29.64	9.91	0.90	3.19	55.75
	Cell Structure and Function	13	423,312	6.35	2.92	0.75	1.46	46.39
	Mechanisms of Genetics	11	423,312	6.64	2.38	0.60	1.51	61.94
	Biological Evolution and Classification	10	423,312	7.16	2.36	0.76	1.14	71.97
	Biological Processes and Systems	10	423,312	5.01	2.16	0.58	1.40	48.87
	Interdependence within Environmental Systems	9	423,312	4.48	2.12	0.56	1.41	51.24

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.98. Spring 2025 STAAR 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	206,093	29.93	9.52	0.89	3.18	56.23
	Cell Structure and Function	13	206,093	6.39	2.82	0.73	1.46	46.57
	Mechanisms of Genetics	11	206,093	6.76	2.31	0.58	1.50	63.07
	Biological Evolution and Classification	10	206,093	7.27	2.27	0.74	1.15	72.88
	Biological Processes and Systems	10	206,093	5.02	2.13	0.58	1.39	48.93
	Interdependence within Environmental Systems	9	206,093	4.47	2.07	0.54	1.41	51.33

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.4.99. Spring 2025 STAAR 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	217,117	29.37	10.26	0.90	3.20	55.30
	Cell Structure and Function	13	217,117	6.30	3.02	0.76	1.47	46.21
	Mechanisms of Genetics	11	217,117	6.53	2.45	0.61	1.52	60.86
	Biological Evolution and Classification	10	217,117	7.06	2.43	0.78	1.14	71.12
	Biological Processes and Systems	10	217,117	5.00	2.19	0.59	1.40	48.81
	Interdependence within Environmental Systems	9	217,117	4.48	2.16	0.57	1.41	51.16

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.100. Spring 2025 STAAR 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	54,655	26.78	9.45	0.88	3.26	50.38
	Cell Structure and Function	13	54,655	5.66	2.78	0.72	1.47	40.99
	Mechanisms of Genetics	11	54,655	6.00	2.33	0.57	1.53	56.29
	Biological Evolution and Classification	10	54,655	6.52	2.43	0.75	1.21	65.43
	Biological Processes and Systems	10	54,655	4.59	2.06	0.53	1.42	44.68
	Interdependence within Environmental Systems	9	54,655	4.01	2.03	0.51	1.42	46.09

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 2025 STAAR 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	225,651	27.56	9.40	0.88	3.24	51.95
	Cell Structure and Function	13	225,651	5.80	2.80	0.72	1.47	42.23
	Mechanisms of Genetics	11	225,651	6.24	2.30	0.56	1.52	58.51
	Biological Evolution and Classification	10	225,651	6.78	2.42	0.76	1.18	68.16
	Biological Processes and Systems	10	225,651	4.65	2.05	0.52	1.42	45.30
	Interdependence within Environmental Systems	9	225,651	4.08	2.01	0.51	1.41	47.17

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.4.102. Spring 2025 STAAR 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	103,452	33.48	9.06	0.88	3.11	62.74
	Cell Structure and Function	13	103,452	7.32	2.77	0.73	1.45	53.65
	Mechanisms of Genetics	11	103,452	7.43	2.24	0.57	1.47	68.60
	Biological Evolution and Classification	10	103,452	7.97	1.92	0.69	1.06	80.17
	Biological Processes and Systems	10	103,452	5.58	2.13	0.59	1.36	54.50
	Interdependence within Environmental Systems	9	103,452	5.18	2.06	0.53	1.40	58.62

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.4.103. Spring 2025 STAAR 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	114,739	22.80	8.42	0.85	3.30	43.15
	Cell Structure and Function	13	114,739	4.57	2.52	0.66	1.46	33.17
	Mechanisms of Genetics	11	114,739	5.34	2.21	0.51	1.54	50.22
	Biological Evolution and Classification	10	114,739	5.64	2.52	0.75	1.25	56.79
	Biological Processes and Systems	10	114,739	3.95	1.83	0.38	1.45	38.40
	Interdependence within Environmental Systems	9	114,739	3.31	1.82	0.41	1.40	38.50

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.4.104. Spring 2025 STAAR 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	308,573	32.18	9.20	0.88	3.14	60.44
	Cell Structure and Function	13	308,573	7.01	2.79	0.73	1.45	51.30
	Mechanisms of Genetics	11	308,573	7.13	2.26	0.56	1.49	66.29
	Biological Evolution and Classification	10	308,573	7.73	2.02	0.70	1.10	77.62
	Biological Processes and Systems	10	308,573	5.41	2.14	0.59	1.37	52.76
	Interdependence within Environmental Systems	9	308,573	4.91	2.06	0.53	1.41	55.98

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.4.105. Spring 2025 STAAR 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	393,624	44.76	14.55	0.93	3.82	57.18
	History	34	393,624	18.76	6.67	0.87	2.42	54.96
	Geography and Culture	13	393,624	8.08	2.79	0.66	1.62	62.71
	Government and Citizenship	12	393,624	6.43	2.71	0.74	1.40	55.39
	Economics, Science, Technology, and Society	19	393,624	11.49	3.85	0.76	1.90	58.96

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.4.106. Spring 2025 STAAR 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	194,172	43.95	13.99	0.92	3.84	56.08
	History	34	194,172	18.20	6.39	0.85	2.44	53.28
	Geography and Culture	13	194,172	7.99	2.76	0.65	1.63	62.01
	Government and Citizenship	12	194,172	6.49	2.66	0.72	1.40	55.70
	Economics, Science, Technology, and Society	19	194,172	11.27	3.69	0.73	1.91	57.78

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.4.107. Spring 2025 STAAR 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	199,258	45.55	15.03	0.94	3.79	58.24
	History	34	199,258	19.31	6.89	0.88	2.40	56.59
	Geography and Culture	13	199,258	8.15	2.81	0.67	1.61	63.40
	Government and Citizenship	12	199,258	6.38	2.76	0.75	1.39	55.08
	Economics, Science, Technology, and Society	19	199,258	11.71	3.98	0.78	1.88	60.11

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.108. Spring 2025 STAAR 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	49,969	41.16	13.89	0.92	3.89	52.53
	History	34	49,969	17.27	6.31	0.85	2.46	50.71
	Geography and Culture	13	49,969	7.43	2.73	0.63	1.66	57.54
	Government and Citizenship	12	49,969	5.93	2.64	0.71	1.43	50.86
	Economics, Science, Technology, and Society	19	49,969	10.53	3.76	0.73	1.95	53.85

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.4.109. Spring 2025 STAAR 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	210,259	42.07	13.93	0.92	3.88	53.71
	History	34	210,259	17.58	6.40	0.85	2.46	51.52
	Geography and Culture	13	210,259	7.66	2.74	0.64	1.66	59.50
	Government and Citizenship	12	210,259	6.02	2.64	0.71	1.43	51.83
	Economics, Science, Technology, and Society	19	210,259	10.81	3.72	0.73	1.94	55.34

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.4.110. Spring 2025 STAAR 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	98,581	49.72	13.88	0.93	3.70	63.53
	History	34	98,581	20.94	6.47	0.87	2.36	61.20
	Geography and Culture	13	98,581	8.85	2.63	0.65	1.56	68.82
	Government and Citizenship	12	98,581	7.13	2.62	0.74	1.34	61.50
	Economics, Science, Technology, and Society	19	98,581	12.80	3.62	0.75	1.81	65.88

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Stratified alpha computed using item types (binary, two-point non-handscored, two-point handscored, essay question conventions domain, essay question ideas domain) as strata. Interrater correlation was used as stratum reliability where applicable for the purpose of computing stratified alpha. If stratified alpha was not viable because a stratum had only one item, KR_{20} was computed instead.
3. Mean of percent correct (0–100%) across all items.

Table B.4.111. Spring 2025 STAAR 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	83,558	34.47	12.66	0.90	3.96	44.00
	History	34	83,558	14.53	5.88	0.82	2.51	42.62
	Geography and Culture	13	83,558	6.33	2.59	0.57	1.70	49.01
	Government and Citizenship	12	83,558	4.64	2.36	0.60	1.49	40.13
	Economics, Science, Technology, and Society	19	83,558	8.97	3.58	0.69	2.00	45.89

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.4.112. Spring 2025 STAAR 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	310,066	47.53	13.76	0.92	3.77	60.73
	History	34	310,066	19.90	6.40	0.86	2.39	58.28
	Geography and Culture	13	310,066	8.55	2.65	0.64	1.60	66.41
	Government and Citizenship	12	310,066	6.92	2.60	0.72	1.37	59.51
	Economics, Science, Technology, and Society	19	310,066	12.17	3.63	0.74	1.86	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.

Spring 2025 STAAR
Progress Measure

Table B.5.1. Spring 2025 STAAR Grade 4 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	111567	36756	42514	2866	-	17809	5144	-	791	5030	-	-	657	36756	61114	13697
		33	38	3	-	16	5	-	1	5	-	-	1	33	55	12
Approaches	103832	25905	952	-	16365	24286	-	-	23188	5763	-	-	7373	42270	48426	13136
		25	1	-	16	23	-	-	22	6	-	-	7	41	47	13
Meets	94270	3553	-	-	17410	-	-	17822	20461	-	-	18920	16104	38785	39381	16104
		4	-	-	18	-	-	19	22	-	-	20	17	41	42	17
Masters	53759	156	-	-	1171	-	-	7655	-	-	-	28684	16093	8982	28684	16093
		-	-	-	2	-	-	14	-	-	-	53	30	17	53	30
2024 Total	363428	66370	43466	2866	34946	42095	5144	25477	44440	10793	-	47604	40227	126793	177605	59030
		18	12	1	10	12	1	7	12	3	-	13	11	35	49	16

Notes:

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 2025 STAAR Grade 4 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	84912	29790	17536	2593	524	17754	11647	-	-	4832	-	-	236	30314	35290	19308
		35	21	3	1	21	14	-	-	6	-	-	-	36	42	23
Approaches	88908	9718	-	-	23421	22296	336	-	22111	7619	-	-	3407	33139	44407	11362
		11	-	-	26	25	-	-	25	9	-	-	4	37	50	13
Meets	96898	693	-	-	17099	-	-	30361	22779	-	689	18310	6967	48842	41089	6967
		1	-	-	18	-	-	31	24	-	1	19	7	50	42	7
Masters	72788	40	-	-	1108	-	-	16434	-	-	-	49923	5283	17582	49923	5283
		-	-	-	2	-	-	23	-	-	-	69	7	24	69	7
2024 Total	343506	40241	17536	2593	42152	40050	11983	46795	44890	12451	689	68233	15893	129877	170709	42920
		12	5	1	12	12	3	14	13	4	-	20	5	38	50	12

Notes:

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 2025 STAAR Grade 5 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	117852	45497	30654	1024	769	29045	4347	-	1296	4795	-	-	425	46266	60995	10591
		39	26	1	1	25	4	-	1	4	-	-	-	39	52	9
Approaches	85175	16430	-	-	21241	20069	-	-	20632	3546	-	-	3257	37671	40701	6803
		19	-	-	25	24	-	-	24	4	-	-	4	44	48	8
Meets	89013	2593	-	-	21168	-	-	17333	26126	-	-	13043	8750	41094	39169	8750
		3	-	-	24	-	-	19	29	-	-	15	10	46	44	10
Masters	74945	142	-	-	2095	-	-	16626	-	-	-	38180	17902	18863	38180	17902
		-	-	-	3	-	-	22	-	-	-	51	24	25	51	24
2024 Total	366985	64662	30654	1024	45273	49114	4347	33959	48054	8341	-	51223	30334	143894	179045	44046
		18	8	-	12	13	1	9	13	2	-	14	8	39	49	12

Notes:

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 2025 STAAR Grade 5 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	64856	27101	20782	2813	-	5335	6111	-	-	2545	-	-	169	27101	26117	11638
		42	32	4	-	8	9	-	-	4	-	-	-	42	40	18
Approaches	107954	24648	1529	-	17363	23656	553	-	24334	10255	-	-	5616	42011	49519	16424
		23	1	-	16	22	1	-	23	9	-	-	5	39	46	15
Meets	98922	1749	-	-	12430	-	-	21486	27503	-	-	20757	14997	35665	48260	14997
		2	-	-	13	-	-	22	28	-	-	21	15	36	49	15
Masters	81604	83	-	-	706	-	-	12613	-	-	-	57517	10685	13402	57517	10685
		-	-	-	1	-	-	15	-	-	-	70	13	16	70	13
2024 Total	353336	53581	22311	2813	30499	28991	6664	34099	51837	12800	-	78274	31467	118179	181413	53744
		15	6	1	9	8	2	10	15	4	-	22	9	33	51	15

Notes:

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 2025 STAAR Grade 6 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	91578	34109	26559	1022	1035	24211	2751	-	369	1423	-	-	99	35144	51139	5295
		37	29	1	1	26	3	-	-	2	-	-	-	38	56	6
Approaches	101005	30804	-	-	38507	16262	-	-	12836	1544	-	-	1052	69311	29098	2596
		30	-	-	38	16	-	-	13	2	-	-	1	69	29	3
Meets	110457	7081	-	-	41281	-	-	28725	18705	-	-	10074	4591	77087	28779	4591
		6	-	-	37	-	-	26	17	-	-	9	4	70	26	4
Masters	66842	326	-	-	4161	-	-	20375	-	-	-	30777	11203	24862	30777	11203
		-	-	-	6	-	-	30	-	-	-	46	17	37	46	17
2024 Total	369882	72320	26559	1022	84984	40473	2751	49100	31910	2967	-	40851	16945	206404	139793	23685
		20	7	-	23	11	1	13	9	1	-	11	5	56	38	6

Notes:

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 2025 STAAR Grade 6 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	76396	24422	25003	5998	-	8574	9041	-	-	3193	-	-	165	24422	33577	18397
		32	33	8	-	11	12	-	-	4	-	-	-	32	44	24
Approaches	87964	18608	454	-	14321	25093	264	-	19037	7126	-	-	3061	32929	44584	10451
		21	1	-	16	29	-	-	22	8	-	-	3	37	51	12
Meets	91013	2629	-	-	17853	-	-	20678	25763	-	-	16874	7216	41160	42637	7216
		3	-	-	20	-	-	23	28	-	-	19	8	45	47	8
Masters	103698	246	-	-	2114	-	-	22119	-	-	-	71743	7476	24479	71743	7476
		-	-	-	2	-	-	21	-	-	-	69	7	24	69	7
2024 Total	359071	45905	25457	5998	34288	33667	9305	42797	44800	10319	-	88617	17918	122990	192541	43540
		13	7	2	10	9	3	12	12	3	-	25	5	34	54	12

Notes:

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 2025 STAAR Grade 7 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	100551	31329	52650	743	-	11852	1118	-	873	1850	-	-	136	31329	65375	3847
		31	52	1	-	12	1	-	1	2	-	-	-	31	65	4
Approaches	99273	39714	3836	-	12061	22551	-	-	18214	1798	-	-	1099	51775	44601	2897
		40	4	-	12	23	-	-	18	2	-	-	1	52	45	3
Meets	56727	3818	-	-	11611	-	-	10121	20823	-	-	7328	3026	25550	28151	3026
		7	-	-	20	-	-	18	37	-	-	13	5	45	50	5
Masters	24083	104	-	-	409	-	-	5313	-	-	-	12139	6118	5826	12139	6118
		-	-	-	2	-	-	22	-	-	-	50	25	24	50	25
2024 Total	280634	74965	56486	743	24081	34403	1118	15434	39910	3648	-	19467	10379	114480	150266	15888
		27	20	-	9	12	-	5	14	1	-	7	4	41	54	6

Notes:

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 2025 STAAR Grade 7 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	93924	48345	19831	2490	532	13187	6055	-	188	3080	-	-	216	48877	33206	11841
		51	21	3	1	14	6	-	-	3	-	-	-	52	35	13
Approaches	75875	18580	1	-	19759	17150	1	-	14714	3681	-	-	1989	38339	31865	5671
		24	-	-	26	23	-	-	19	5	-	-	3	51	42	7
Meets	113701	3503	-	-	26776	-	-	27582	29287	-	725	17791	8037	58586	47078	8037
		3	-	-	24	-	-	24	26	-	1	16	7	52	41	7
Masters	93532	106	-	-	1540	-	-	18243	-	-	-	66015	7628	19889	66015	7628
		-	-	-	2	-	-	20	-	-	-	71	8	21	71	8
2024 Total	377032	70534	19832	2490	48607	30337	6056	45825	44189	6761	725	83806	17870	165691	178164	33177
		19	5	1	13	8	2	12	12	2	-	22	5	44	47	9

Notes:

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 2025 STAAR Grade 8 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	143078	45204	39020	862	606	35246	4119	-	5373	11362	-	-	1286	45810	79639	17629
		32	27	1	-	25	3	-	4	8	-	-	1	32	56	12
Approaches	79257	15759	259	-	14890	14692	-	-	24826	4788	-	-	4043	30649	39777	8831
		20	-	-	19	19	-	-	31	6	-	-	5	39	50	11
Meets	79081	2477	-	-	12697	-	-	18492	23147	-	140	13855	8273	33806	37002	8273
		3	-	-	16	-	-	23	29	-	-	18	10	43	47	10
Masters	35498	65	-	-	560	-	-	7134	-	-	-	20196	7543	7759	20196	7543
		-	-	-	2	-	-	20	-	-	-	57	21	22	57	21
2024 Total	336914	63505	39279	862	28753	49938	4119	25626	53346	16150	140	34051	21145	118024	176614	42276
		19	12	-	9	15	1	8	16	5	-	10	6	35	52	13

Notes:

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 2025 STAAR Grade 8 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	103185	30164	24519	7912	-	18115	15974	-	1	6078	-	-	422	30164	42635	30386
		29	24	8	-	18	15	-	-	6	-	-	-	29	41	29
Approaches	76531	7851	-	-	13488	24415	-	-	19331	7537	-	-	3909	21339	43746	11446
		10	-	-	18	32	-	-	25	10	-	-	5	28	57	15
Meets	92483	950	-	-	15442	-	-	17844	30083	-	-	18249	9915	34236	48332	9915
		1	-	-	17	-	-	19	33	-	-	20	11	37	52	11
Masters	105156	92	-	-	1276	-	-	16984	-	-	-	78556	8248	18352	78556	8248
		-	-	-	1	-	-	16	-	-	-	75	8	17	75	8
2024 Total	377355	39057	24519	7912	30206	42530	15974	34828	49415	13615	-	96805	22494	104091	213269	59995
		10	6	2	8	11	4	9	13	4	-	26	6	28	57	16

Notes:

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 2025 STAAR Spanish Grade 4 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	6592	1906	2737	232	-	870	443	-	41	318	-	-	45	1906	3648	1038
		29	42	4	-	13	7	-	1	5	-	-	1	29	55	16
Approaches	3229	814	39	-	467	763	-	-	694	205	-	-	247	1281	1496	452
		25	1	-	14	24	-	-	21	6	-	-	8	40	46	14
Meets	1524	86	-	-	318	-	-	304	318	-	-	286	212	708	604	212
		6	-	-	21	-	-	20	21	-	-	19	14	46	40	14
Masters	411	3	-	-	33	-	-	90	-	-	-	204	81	126	204	81
		1	-	-	8	-	-	22	-	-	-	50	20	31	50	20
2024 Total	11756	2809	2776	232	818	1633	443	394	1053	523	-	490	585	4021	5952	1783
		24	24	2	7	14	4	3	9	4	-	4	5	34	51	15

Notes:

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 2025 STAAR Spanish Grade 4 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	11409	2788	3254	2482	-	224	1626	-	-	906	-	-	129	2788	3478	5143
		24	29	22	-	2	14	-	-	8	-	-	1	24	30	45
Approaches	6237	887	284	-	340	1250	111	-	965	1509	-	-	891	1227	2499	2511
		14	5	-	5	20	2	-	15	24	-	-	14	20	40	40
Meets	1963	45	-	-	213	-	-	169	624	-	-	242	670	427	866	670
		2	-	-	11	-	-	9	32	-	-	12	34	22	44	34
Masters	1744	18	-	-	38	-	-	355	-	-	-	948	385	411	948	385
		1	-	-	2	-	-	20	-	-	-	54	22	24	54	22
2024 Total	21353	3738	3538	2482	591	1474	1737	524	1589	2415	-	1190	2075	4853	7791	8709
		18	17	12	3	7	8	2	7	11	-	6	10	23	36	41

Notes:

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 2025 STAAR Spanish Grade 5 Mathematics Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	5744	2129	1738	59	18	1107	334	-	39	289	-	-	31	2147	2884	713
		37	30	1	-	19	6	-	1	5	-	-	1	37	50	12
Approaches	1718	376	-	-	364	396	-	-	402	87	-	-	93	740	798	180
		22	-	-	21	23	-	-	23	5	-	-	5	43	46	10
Meets	956	40	-	-	223	-	-	132	320	-	-	124	117	395	444	117
		4	-	-	23	-	-	14	33	-	-	13	12	41	46	12
Masters	371	8	-	-	14	-	-	106	-	-	-	175	68	128	175	68
		2	-	-	4	-	-	29	-	-	-	47	18	35	47	18
2024 Total	8789	2553	1738	59	619	1503	334	238	761	376	-	299	309	3410	4301	1078
		29	20	1	7	17	4	3	9	4	-	3	4	39	49	12

Notes:

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 2025 STAAR Spanish Grade 5 RLA Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	7940	2608	1564	255	167	1592	1038	-	49	590	-	-	77	2775	3205	1960
		33	20	3	2	20	13	-	1	7	-	-	1	35	40	25
Approaches	2672	235	-	-	656	563	-	-	740	279	-	-	199	891	1303	478
		9	-	-	25	21	-	-	28	10	-	-	7	33	49	18
Meets	2531	34	-	-	579	-	-	445	776	-	-	334	363	1058	1110	363
		1	-	-	23	-	-	18	31	-	-	13	14	42	44	14
Masters	1687	5	-	-	87	-	-	513	-	-	-	924	158	605	924	158
		-	-	-	5	-	-	30	-	-	-	55	9	36	55	9
2024 Total	14830	2882	1564	255	1489	2155	1038	958	1565	869	-	1258	797	5329	6542	2959
		19	11	2	10	15	7	6	11	6	-	8	5	36	44	20

Notes:

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 2025 STAAR Algebra I Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	100134	46701	64	-	29832	9650	78	-	7596	2748	-	-	3465	76533	17310	6291
		47	-	-	30	10	-	-	8	3	-	-	3	76	17	6
Approaches	105104	17956	-	-	41750	5747	-	-	24826	1892	-	-	12933	59706	30573	14825
		17	-	-	40	5	-	-	24	2	-	-	12	57	29	14
Meets	109117	3252	-	-	23968	-	-	5307	28845	-	-	9610	38135	32527	38455	38135
		3	-	-	22	-	-	5	26	-	-	9	35	30	35	35
Masters	69729	141	-	-	1861	-	-	6745	-	-	-	10623	50359	8747	10623	50359
		-	-	-	3	-	-	10	-	-	-	15	72	13	15	72
2024 Total	384084	68050	64	-	97411	15397	78	12052	61267	4640	-	20233	104892	177513	96961	109610
		18	-	-	25	4	-	3	16	1	-	5	27	46	25	29

Notes:

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 2025 STAAR English I Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	75728	58191	385	-	5063	7575	46	-	4288	140	-	-	40	63254	12248	226
		77	1	-	7	10	-	-	6	-	-	-	-	84	16	-
Approaches	91507	23976	-	-	25214	7564	-	-	34281	81	-	-	391	49190	41845	472
		26	-	-	28	8	-	-	37	-	-	-	-	54	46	1
Meets	99765	2715	-	-	12767	-	-	4457	72376	-	-	3549	3901	19939	75925	3901
		3	-	-	13	-	-	4	73	-	-	4	4	20	76	4
Masters	111361	123	-	-	712	-	-	46553	-	-	-	33280	30693	47388	33280	30693
		-	-	-	1	-	-	42	-	-	-	30	28	43	30	28
2024 Total	378361	85005	385	-	43756	15139	46	51010	110945	221	-	36829	35025	179771	163298	35292
		22	-	-	12	4	-	13	29	-	-	10	9	48	43	9

Notes:

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 2025 STAAR English II Progress Measure

		2025														
		Did Not Meet			Approaches			Meets			Masters			2025 Total		
2024	**N**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**	**Limited**	**Expected**	**Accelerated**
Did Not Meet	85490	31857	28783	127	-	17263	196	-	6651	604	-	-	9	31857	52697	936
		37	34	-	-	20	-	-	8	1	-	-	-	37	62	1
Approaches	51929	13051	-	-	10799	9350	-	-	18693	27	-	-	9	23850	28043	36
		25	-	-	21	18	-	-	36	-	-	-	-	46	54	-
Meets	169001	4818	-	-	20021	-	-	80098	59328	-	-	4617	119	104937	63945	119
		3	-	-	12	-	-	47	35	-	-	3	-	62	38	-
Masters	79609	42	-	-	137	-	-	45842	-	-	-	32276	1312	46021	32276	1312
		-	-	-	-	-	-	58	-	-	-	41	2	58	41	2
2024 Total	386029	49768	28783	127	30957	26613	196	125940	84672	631	-	36893	1449	206665	176961	2403
		13	7	-	8	7	-	33	22	-	-	10	-	54	46	1

Notes:

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 2025 STAAR
On-Track Measure

Table B.6.1. Spring 2025 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	363428	192400	53
5	Mathematics	8	366985	208534	57
6	Mathematics	8	369882	146462	40
7	Mathematics	8	280634	120614	43
8	Mathematics	Algebra I	336914	179353	53
4	Mathematics Spanish	5	11756	4450	38
4	RLA	5	343506	198736	58
5	RLA	8	353336	233365	66
6	RLA	8	359071	229626	64
7	RLA	8	377032	206927	55
8	RLA	English I	377355	250630	66
English I	RLA	English II	378361	256456	68
4	RLA Spanish	5	21353	13024	61

Spring 2025 STAAR
Scale Score Descriptive Statistics

Table B.7.1. Spring 2025 STAAR Grades 3–8 Scale Score Descriptive Statistics

	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	376,507	1459.89	1446	1294	1210	246	167.51	28058.71	0.45	0.20
Grade 4	Mathematics	381,077	1565.08	1538	1389	1220	252	176.18	31041.15	0.64	0.32
Grade 5	Mathematics	383,295	1641.24	1616	1457	1200	245	181.39	32900.58	0.61	0.23
Grade 6	Mathematics	390,062	1720.75	1689	1603	1280	203	163.96	26883.16	1.00	1.45
Grade 7	Mathematics	299,218	1752.50	1703	1649	1250	180	159.32	25381.79	1.18	2.17
Grade 8	Mathematics	372,214	1858.83	1836	1707	1230	224	169.36	28683.50	0.84	1.00
Grade 3	RLA	363,664	1476.74	1467	1493	1480	228	162.30	26340.47	0.01	0.14
Grade 4	RLA	371,478	1554.79	1564	1564	1390	208	154.11	23749.67	-0.05	0.48
Grade 5	RLA	378,262	1610.42	1616	1700	1390	253	170.69	29135.59	-0.07	0.16
Grade 6	RLA	396,497	1650.28	1654	1741	1400	231	159.49	25437.41	-0.03	0.21
Grade 7	RLA	398,245	1667.11	1669	1771	1400	216	164.85	27177.11	-0.09	0.29
Grade 8	RLA	397,825	1720.80	1726	1839	1380	208	154.94	24007.93	-0.15	0.26
Grade 5	Science	383,147	3768.37	3751	3550	5060	736	532.65	283717.81	0.43	1.09
Grade 8	Science	393,051	3973.31	3905	4085	6980	983	698.40	487765.80	0.44	0.73
Grade 8	Social Studies	401,346	3719.25	3642	3186	5500	848	621.79	386627.88	0.54	0.77

Table B.7.2. Spring 2025 STAAR Spanish Grades 3–5 Scale Score Descriptive Statistics

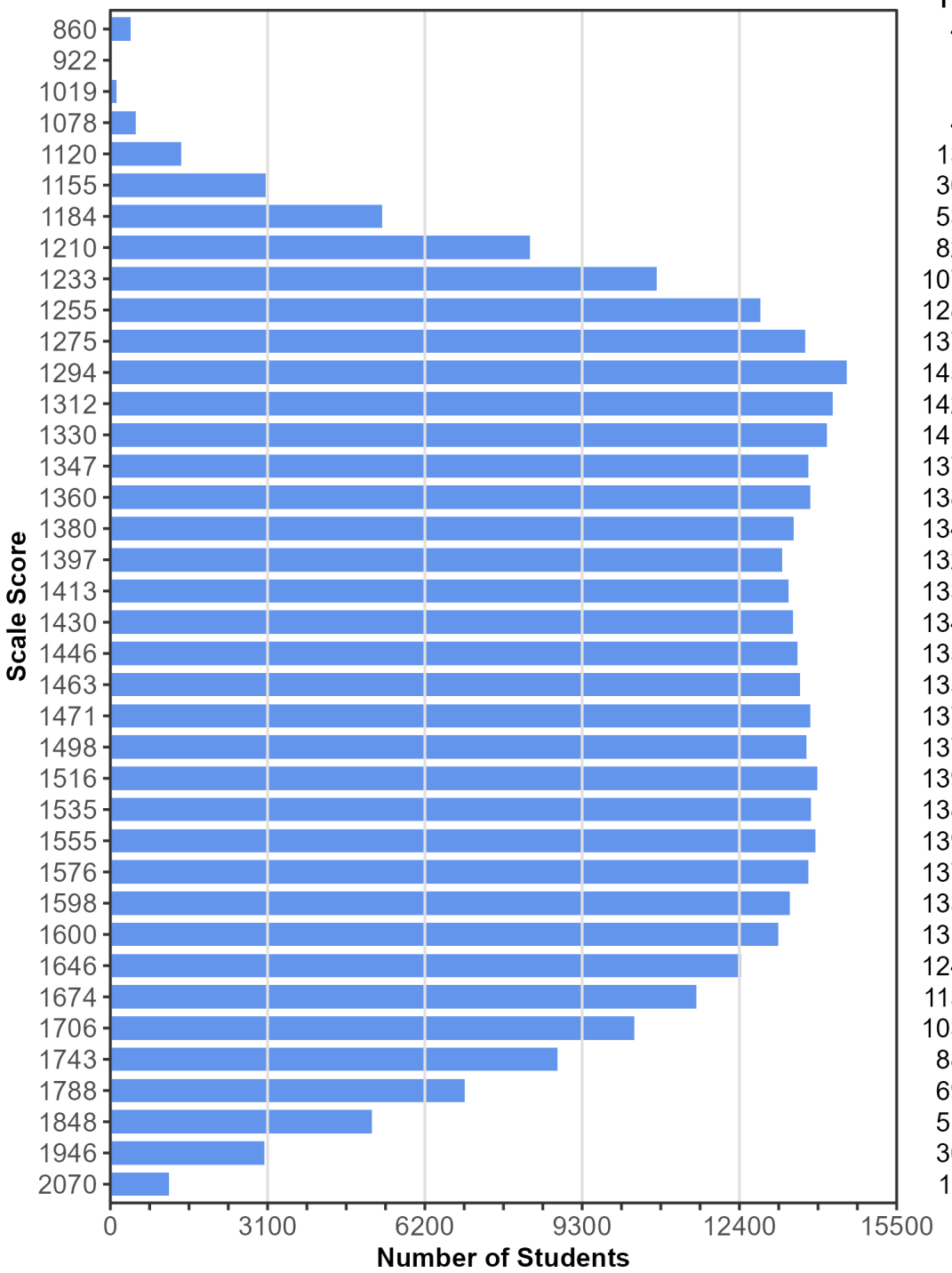
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	21,318	1396.58	1380	1275	1210	177	138.49	19178.87	0.64	0.40
Grade 4	Mathematics	15,379	1475.34	1450	1373	1220	181	130.39	17002.03	0.96	1.12
Grade 5	Mathematics	12,208	1527.99	1502	1441	1200	178	135.16	18268.25	1.09	1.57
Grade 3	RLA	34,164	1332.21	1318	1220	1378	214	143.88	20701.34	0.23	-0.23
Grade 4	RLA	26,357	1414.58	1408	1503	1430	221	152.08	23128.79	0.13	-0.28
Grade 5	RLA	19,392	1486.48	1492	1556	1460	224	152.04	23117.09	0.10	-0.19
Grade 5	Science	14,380	3408.46	3391	3191	4225	515	384.67	147973.47	0.48	0.79

Table B.7.3. Spring 2025 STAAR EOC Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	455,622	4051.50	3952	3467	4930	841	619.22	383439.2	0.88	1.36
English I	483,644	4038.72	4000	4570	4250	774	519.98	270383.6	0.17	-0.15
English II	464,246	4060.98	4075	4556	4400	730	495.10	245122.4	0.01	0.05
Biology	423,312	4151.08	4149	4230	4360	620	467.77	218812.2	0.17	0.68
U.S. History	393,624	4281.78	4257	4318	5330	717	515.69	265939.0	0.32	0.77

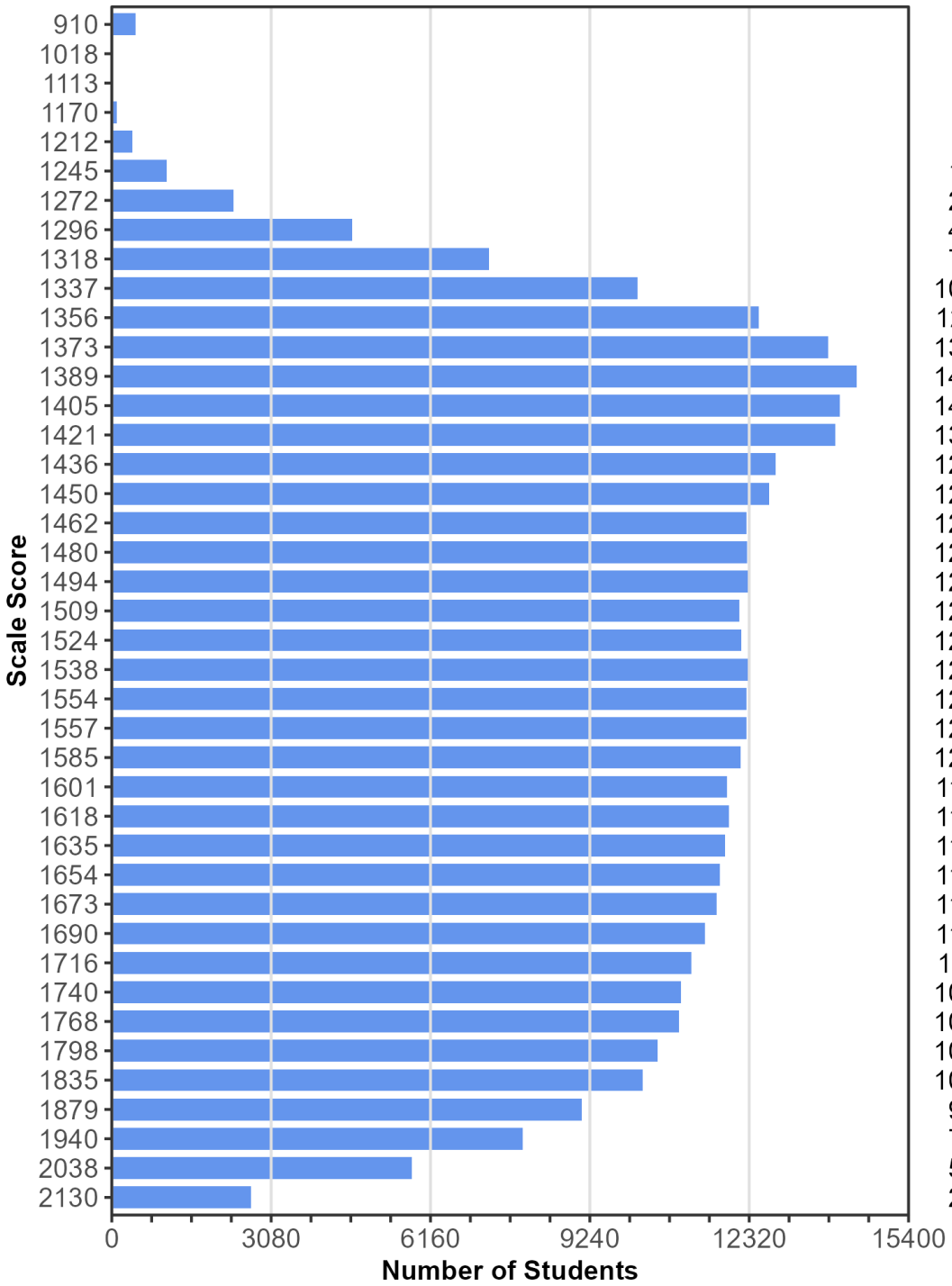
Spring 2025 STAAR
Frequency Distribution of Scale Scores

Figure B.7.1. Spring 2025 STAAR Grade 3 Mathematics
Frequency Distribution of Scale Scores



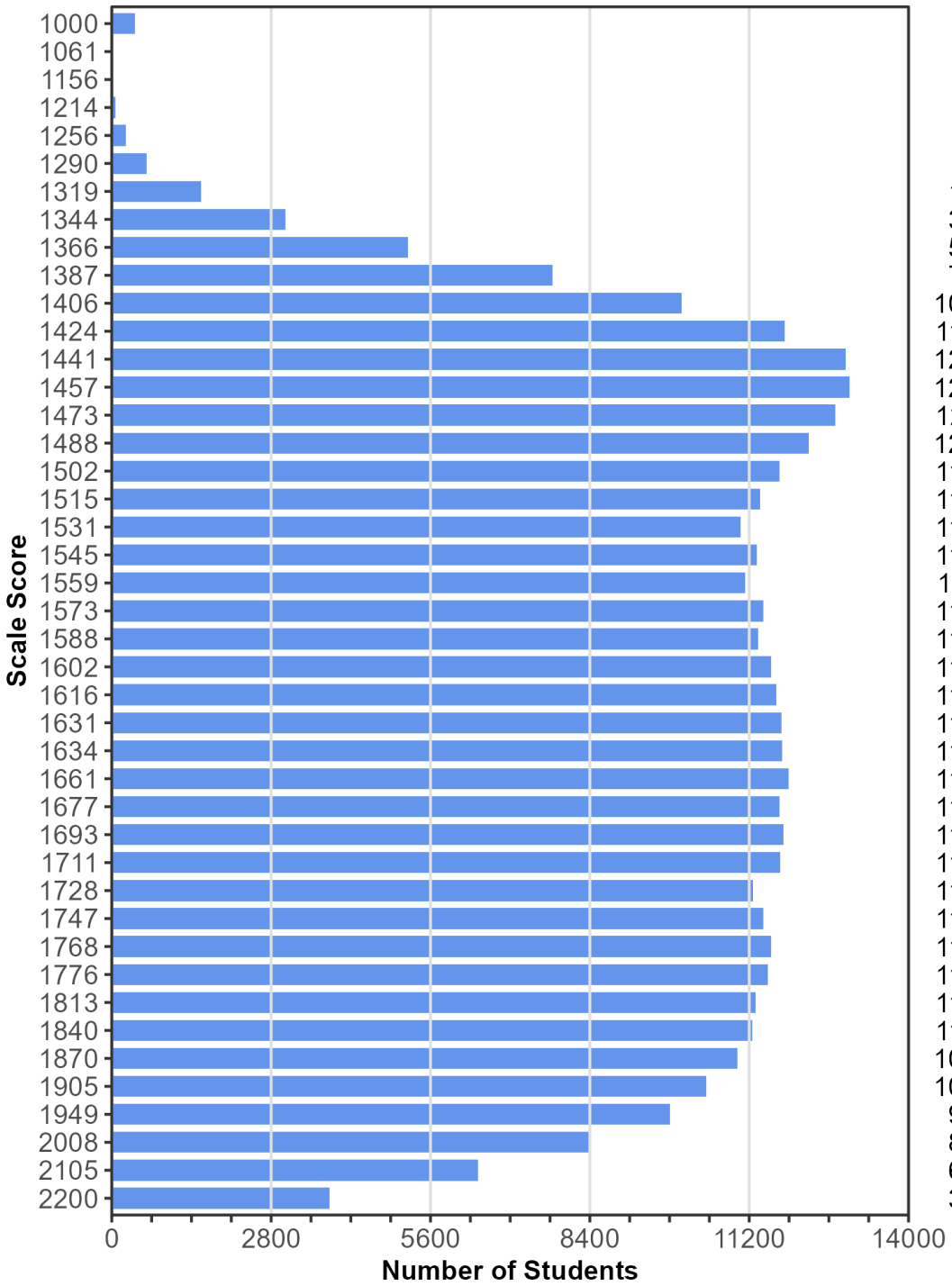
	Freq	Cum Freq	Pct	Cum Pct
860	401	401	0.11	0.11
922	23	424	0.01	0.11
1019	125	549	0.03	0.15
1078	498	1047	0.13	0.28
1120	1396	2443	0.37	0.65
1155	3056	5499	0.81	1.46
1184	5356	10855	1.42	2.88
1210	8275	19130	2.20	5.08
1233	10770	29900	2.86	7.94
1255	12815	42715	3.40	11.35
1275	13702	56417	3.64	14.98
1294	14517	70934	3.86	18.84
1312	14235	85169	3.78	22.62
1330	14128	99297	3.75	26.37
1347	13754	113051	3.65	30.03
1360	13800	126851	3.67	33.69
1380	13467	140318	3.58	37.27
1397	13248	153566	3.52	40.79
1413	13368	166934	3.55	44.34
1430	13454	180388	3.57	47.91
1446	13551	193939	3.60	51.51
1463	13594	207533	3.61	55.12
1471	13794	221327	3.66	58.78
1498	13728	235055	3.65	62.43
1516	13931	248986	3.70	66.13
1535	13804	262790	3.67	69.80
1555	13904	276694	3.69	73.49
1576	13765	290459	3.66	77.15
1598	13397	303856	3.56	80.70
1600	13166	317022	3.50	84.20
1646	12437	329459	3.30	87.50
1674	11558	341017	3.07	90.57
1706	10324	351341	2.74	93.32
1743	8820	360161	2.34	95.66
1788	6989	367150	1.86	97.51
1848	5155	372305	1.37	98.88
1946	3042	375347	0.81	99.69
2070	1160	376507	0.31	100.00

Figure B.7.2. Spring 2025 STAAR Grade 4 Mathematics
Frequency Distribution of Scale Scores



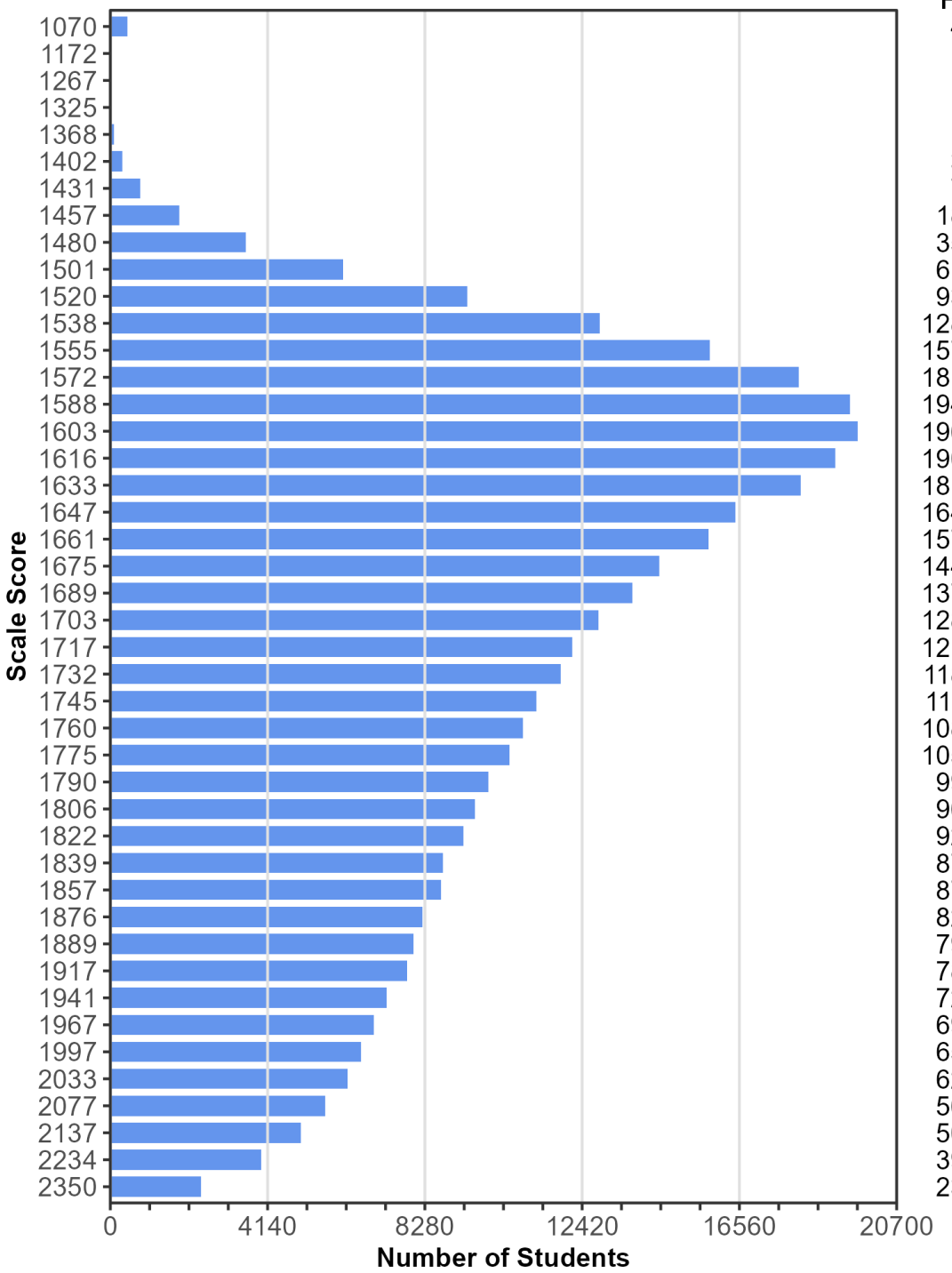
	Freq	Cum Freq	Pct	Cum Pct
910	455	455	0.12	0.12
1018	14	469	0.00	0.12
1113	24	493	0.01	0.13
1170	97	590	0.03	0.15
1212	400	990	0.10	0.26
1245	1067	2057	0.28	0.54
1272	2357	4414	0.62	1.16
1296	4652	9066	1.22	2.38
1318	7286	16352	1.91	4.29
1337	10156	26508	2.67	6.96
1356	12511	39019	3.28	10.24
1373	13842	52861	3.63	13.87
1389	14397	67258	3.78	17.65
1405	14075	81333	3.69	21.34
1421	13978	95311	3.67	25.01
1436	12836	108147	3.37	28.38
1450	12706	120853	3.33	31.71
1462	12262	133115	3.22	34.93
1480	12276	145391	3.22	38.15
1494	12296	157687	3.23	41.38
1509	12126	169813	3.18	44.56
1524	12173	181986	3.19	47.76
1538	12334	194320	3.24	50.99
1554	12273	206593	3.22	54.21
1557	12273	218866	3.22	57.43
1585	12152	231018	3.19	60.62
1601	11896	242914	3.12	63.74
1618	11927	254841	3.13	66.87
1635	11850	266691	3.11	69.98
1654	11747	278438	3.08	73.07
1673	11689	290127	3.07	76.13
1690	11459	301586	3.01	79.14
1716	11196	312782	2.94	82.08
1740	10999	323781	2.89	84.96
1768	10969	334750	2.88	87.84
1798	10548	345298	2.77	90.61
1835	10261	355559	2.69	93.30
1879	9084	364643	2.38	95.69
1940	7944	372587	2.08	97.77
2038	5802	378389	1.52	99.29
2130	2688	381077	0.71	100.00

Figure B.7.3. Spring 2025 STAAR Grade 5 Mathematics
Frequency Distribution of Scale Scores



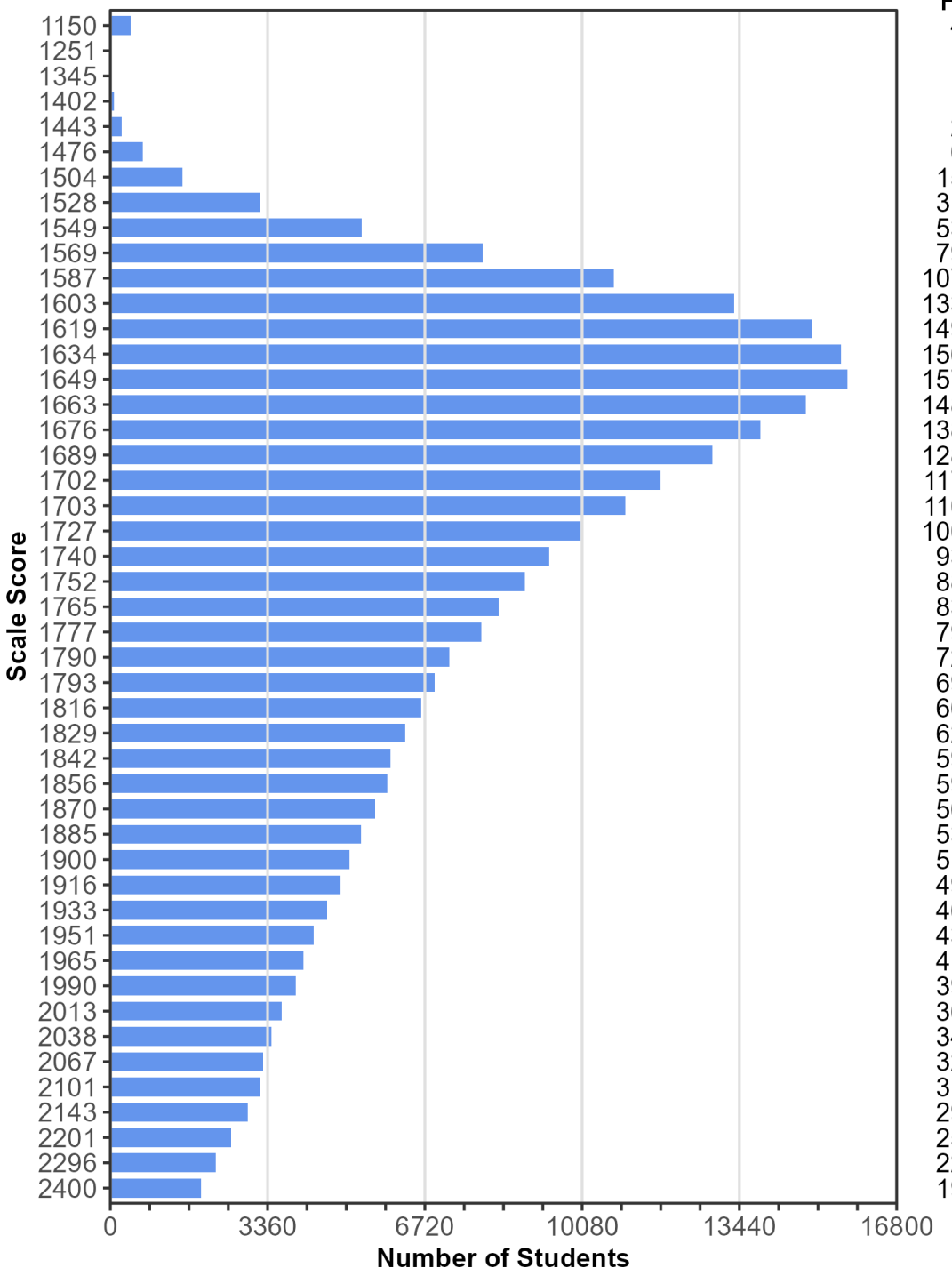
	Freq	Cum Freq	Pct	Cum Pct
1000	407	407	0.11	0.11
1061	9	416	0.00	0.11
1156	16	432	0.00	0.11
1214	62	494	0.02	0.13
1256	242	736	0.06	0.19
1290	611	1347	0.16	0.35
1319	1567	2914	0.41	0.76
1344	3048	5962	0.80	1.56
1366	5199	11161	1.36	2.91
1387	7748	18909	2.02	4.93
1406	10010	28919	2.61	7.54
1424	11828	40747	3.09	10.63
1441	12893	53640	3.36	13.99
1457	12967	66607	3.38	17.38
1473	12711	79318	3.32	20.69
1488	12249	91567	3.20	23.89
1502	11738	103305	3.06	26.95
1515	11386	114691	2.97	29.92
1531	11049	125740	2.88	32.81
1545	11335	137075	2.96	35.76
1559	11132	148207	2.90	38.67
1573	11453	159660	2.99	41.65
1588	11362	171022	2.96	44.62
1602	11580	182602	3.02	47.64
1616	11673	194275	3.05	50.69
1631	11768	206043	3.07	53.76
1634	11777	217820	3.07	56.83
1661	11891	229711	3.10	59.93
1677	11731	241442	3.06	62.99
1693	11799	253241	3.08	66.07
1711	11747	264988	3.06	69.13
1728	11267	276255	2.94	72.07
1747	11448	287703	2.99	75.06
1768	11581	299284	3.02	78.08
1776	11530	310814	3.01	81.09
1813	11313	322127	2.95	84.04
1840	11260	333387	2.94	86.98
1870	10989	344376	2.87	89.85
1905	10445	354821	2.73	92.57
1949	9813	364634	2.56	95.13
2008	8405	373039	2.19	97.32
2105	6433	379472	1.68	99.00
2200	3823	383295	1.00	100.00

Figure B.7.4. Spring 2025 STAAR Grade 6 Mathematics
Frequency Distribution of Scale Scores



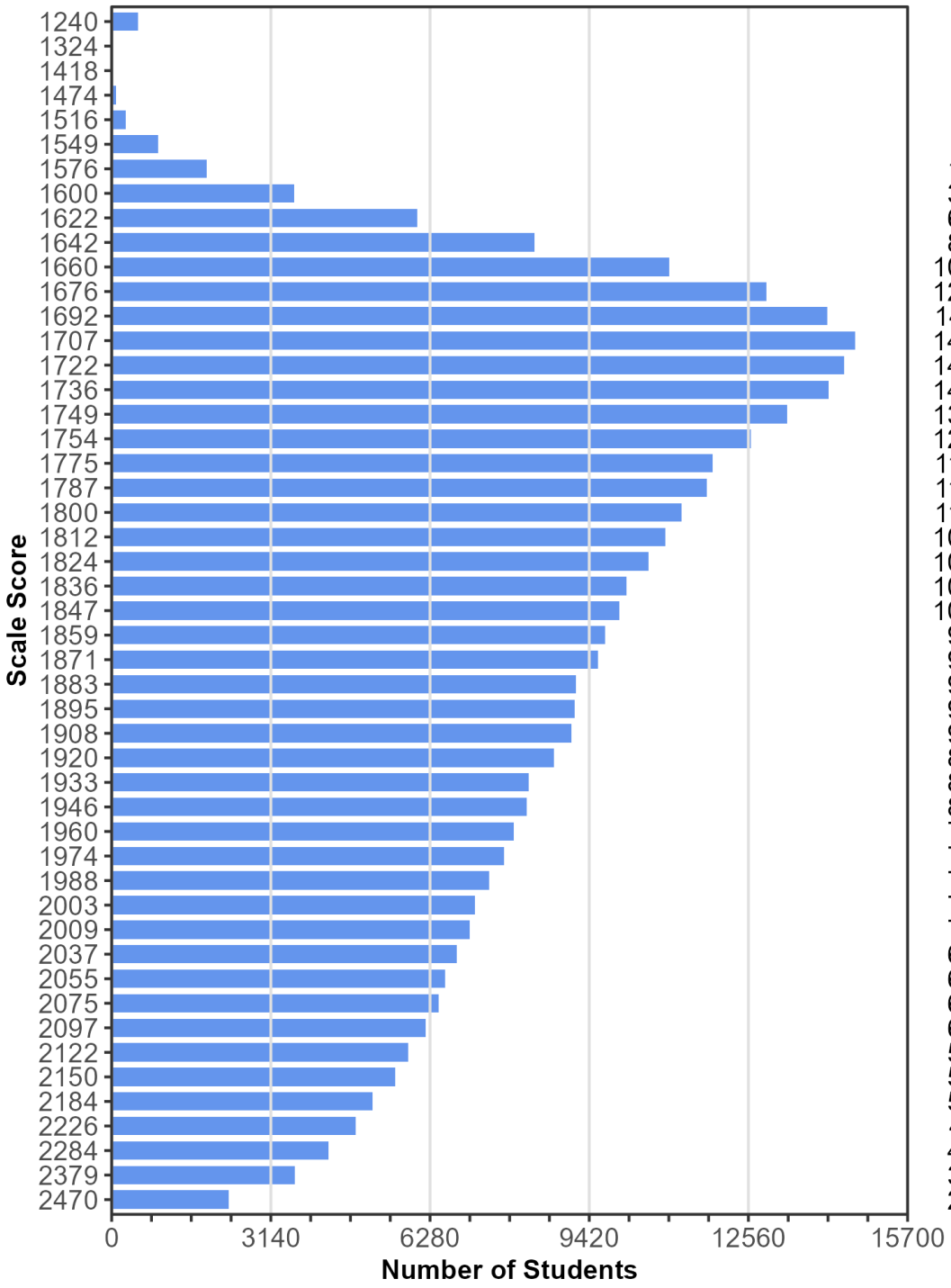
Freq	Cum Freq	Pct	Cum Pct
449	449	0.12	0.12
18	467	0.00	0.12
16	483	0.00	0.12
28	511	0.01	0.13
94	605	0.02	0.16
320	925	0.08	0.24
793	1718	0.20	0.44
1823	3541	0.47	0.91
3575	7116	0.92	1.82
6127	13243	1.57	3.40
9390	22633	2.41	5.80
12887	35520	3.30	9.11
15780	51300	4.05	13.15
18123	69423	4.65	17.80
19471	88894	4.99	22.79
19677	108571	5.04	27.83
19084	127655	4.89	32.73
18169	145824	4.66	37.38
16459	162283	4.22	41.60
15755	178038	4.04	45.64
14447	192485	3.70	49.35
13742	206227	3.52	52.87
12844	219071	3.29	56.16
12165	231236	3.12	59.28
11849	243085	3.04	62.32
11211	254296	2.87	65.19
10859	265155	2.78	67.98
10506	275661	2.69	70.67
9957	285618	2.55	73.22
9603	295221	2.46	75.69
9294	304515	2.38	78.07
8759	313274	2.25	80.31
8710	321984	2.23	82.55
8219	330203	2.11	84.65
7982	338185	2.05	86.70
7806	345991	2.00	88.70
7276	353267	1.87	90.57
6928	360195	1.78	92.34
6593	366788	1.69	94.03
6254	373042	1.60	95.64
5656	378698	1.45	97.09
5008	383706	1.28	98.37
3970	387676	1.02	99.39
2386	390062	0.61	100.00

Figure B.7.5. Spring 2025 STAAR Grade 7 Mathematics
Frequency Distribution of Scale Scores



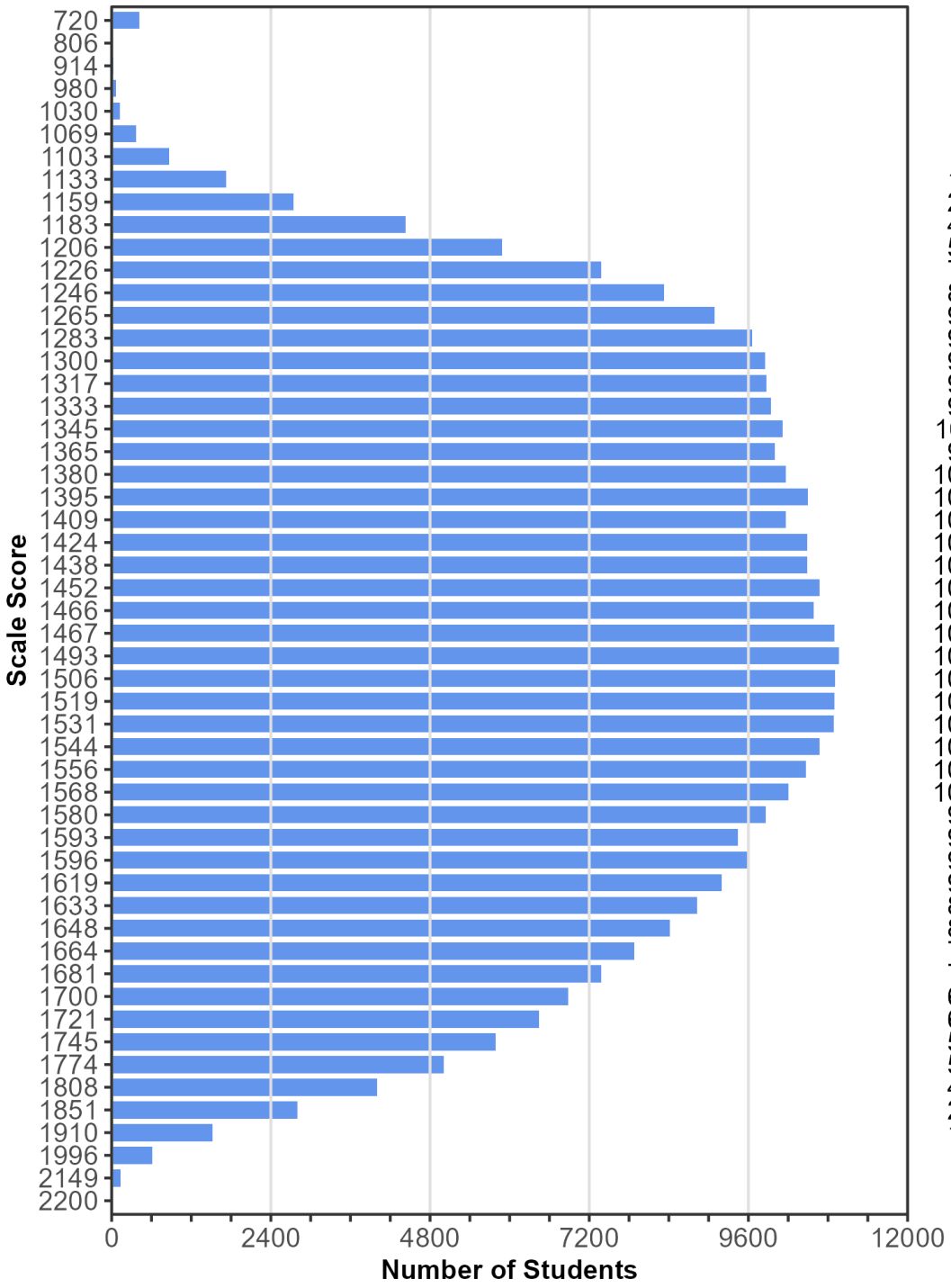
Freq	Cum Freq	Pct	Cum Pct
427	427	0.14	0.14
20	447	0.01	0.15
25	472	0.01	0.16
79	551	0.03	0.18
249	800	0.08	0.27
695	1495	0.23	0.50
1539	3034	0.51	1.01
3195	6229	1.07	2.08
5376	11605	1.80	3.88
7950	19555	2.66	6.54
10757	30312	3.60	10.13
13333	43645	4.46	14.59
14981	58626	5.01	19.59
15605	74231	5.22	24.81
15750	89981	5.26	30.07
14857	104838	4.97	35.04
13893	118731	4.64	39.68
12865	131596	4.30	43.98
11755	143351	3.93	47.91
11003	154354	3.68	51.59
10040	164394	3.36	54.94
9371	173765	3.13	58.07
8855	182620	2.96	61.03
8301	190921	2.77	63.81
7920	198841	2.65	66.45
7242	206083	2.42	68.87
6929	213012	2.32	71.19
6644	219656	2.22	73.41
6296	225952	2.10	75.51
5988	231940	2.00	77.52
5920	237860	1.98	79.49
5659	243519	1.89	81.39
5352	248871	1.79	83.17
5106	253977	1.71	84.88
4921	258898	1.64	86.52
4625	263523	1.55	88.07
4345	267868	1.45	89.52
4121	271989	1.38	90.90
3965	275954	1.33	92.23
3660	279614	1.22	93.45
3447	283061	1.15	94.60
3263	286324	1.09	95.69
3192	289516	1.07	96.76
2935	292451	0.98	97.74
2578	295029	0.86	98.60
2247	297276	0.75	99.35
1942	299218	0.65	100.00

Figure B.7.6. Spring 2025 STAAR Grade 8 Mathematics
Frequency Distribution of Scale Scores



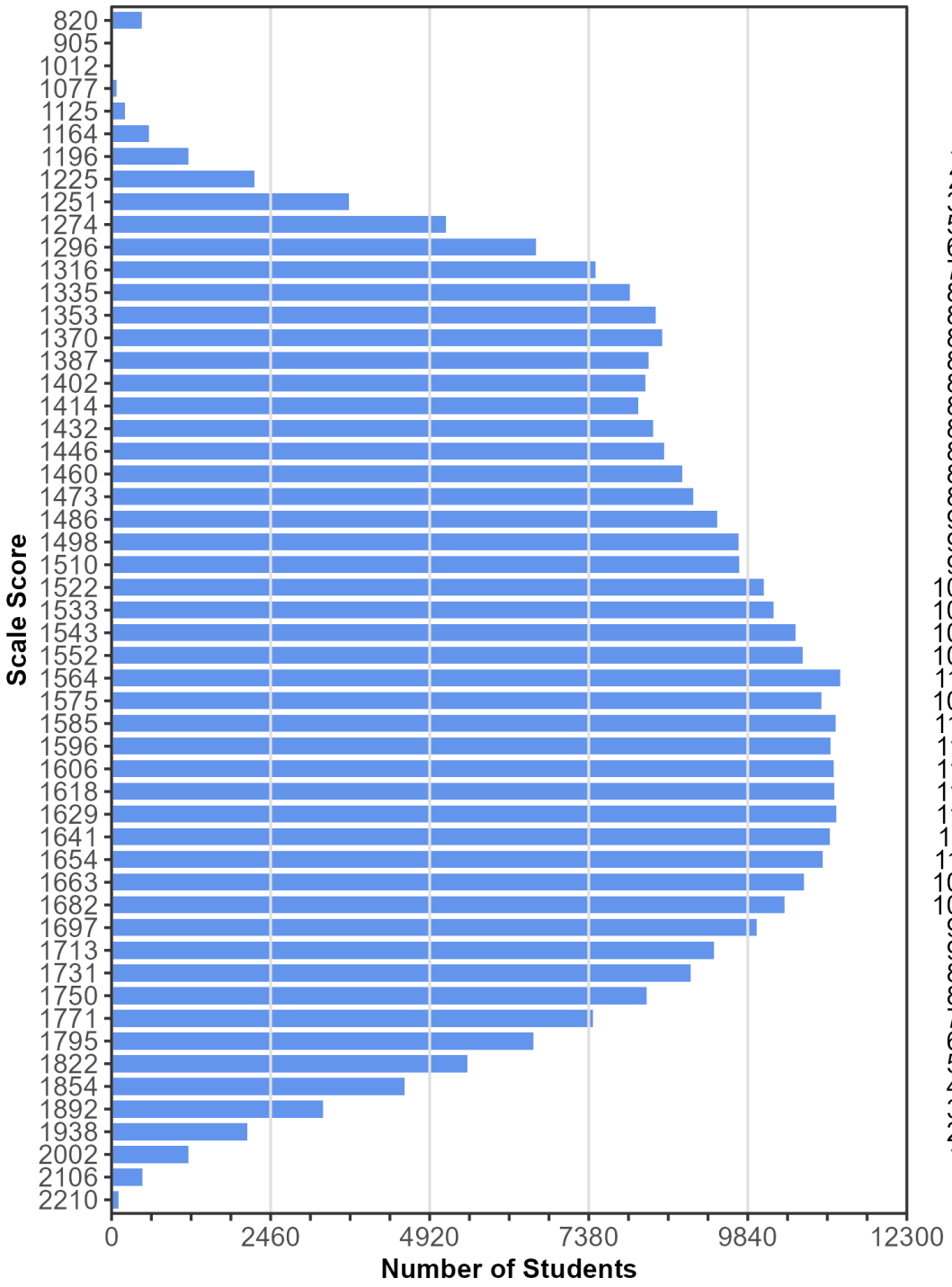
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1240	520	520	0.14	0.14
1324	14	534	0.00	0.14
1418	24	558	0.01	0.15
1474	91	649	0.02	0.17
1516	282	931	0.08	0.25
1549	918	1849	0.25	0.50
1576	1880	3729	0.51	1.00
1600	3597	7326	0.97	1.97
1622	6029	13355	1.62	3.59
1642	8346	21701	2.24	5.83
1660	10993	32694	2.95	8.78
1676	12918	45612	3.47	12.25
1692	14112	59724	3.79	16.05
1707	14669	74393	3.94	19.99
1722	14442	88835	3.88	23.87
1736	14146	102981	3.80	27.67
1749	13329	116310	3.58	31.25
1754	12613	128923	3.39	34.64
1775	11851	140774	3.18	37.82
1787	11735	152509	3.15	40.97
1800	11242	163751	3.02	43.99
1812	10921	174672	2.93	46.93
1824	10593	185265	2.85	49.77
1836	10151	195416	2.73	52.50
1847	10010	205426	2.69	55.19
1859	9737	215163	2.62	57.81
1871	9590	224753	2.58	60.38
1883	9159	233912	2.46	62.84
1895	9135	243047	2.45	65.30
1908	9063	252110	2.43	67.73
1920	8718	260828	2.34	70.07
1933	8227	269055	2.21	72.29
1946	8190	277245	2.20	74.49
1960	7930	285175	2.13	76.62
1974	7742	292917	2.08	78.70
1988	7443	300360	2.00	80.70
2003	7165	307525	1.92	82.62
2009	7062	314587	1.90	84.52
2037	6806	321393	1.83	86.35
2055	6583	327976	1.77	88.11
2075	6444	334420	1.73	89.85
2097	6190	340610	1.66	91.51
2122	5851	346461	1.57	93.08
2150	5594	352055	1.50	94.58
2184	5143	357198	1.38	95.97
2226	4811	362009	1.29	97.26
2284	4280	366289	1.15	98.41
2379	3615	369904	0.97	99.38
2470	2310	372214	0.62	100.00

Figure B.7.7. Spring 2025 STAAR Grade 3 RLA
Frequency Distribution of Scale Scores



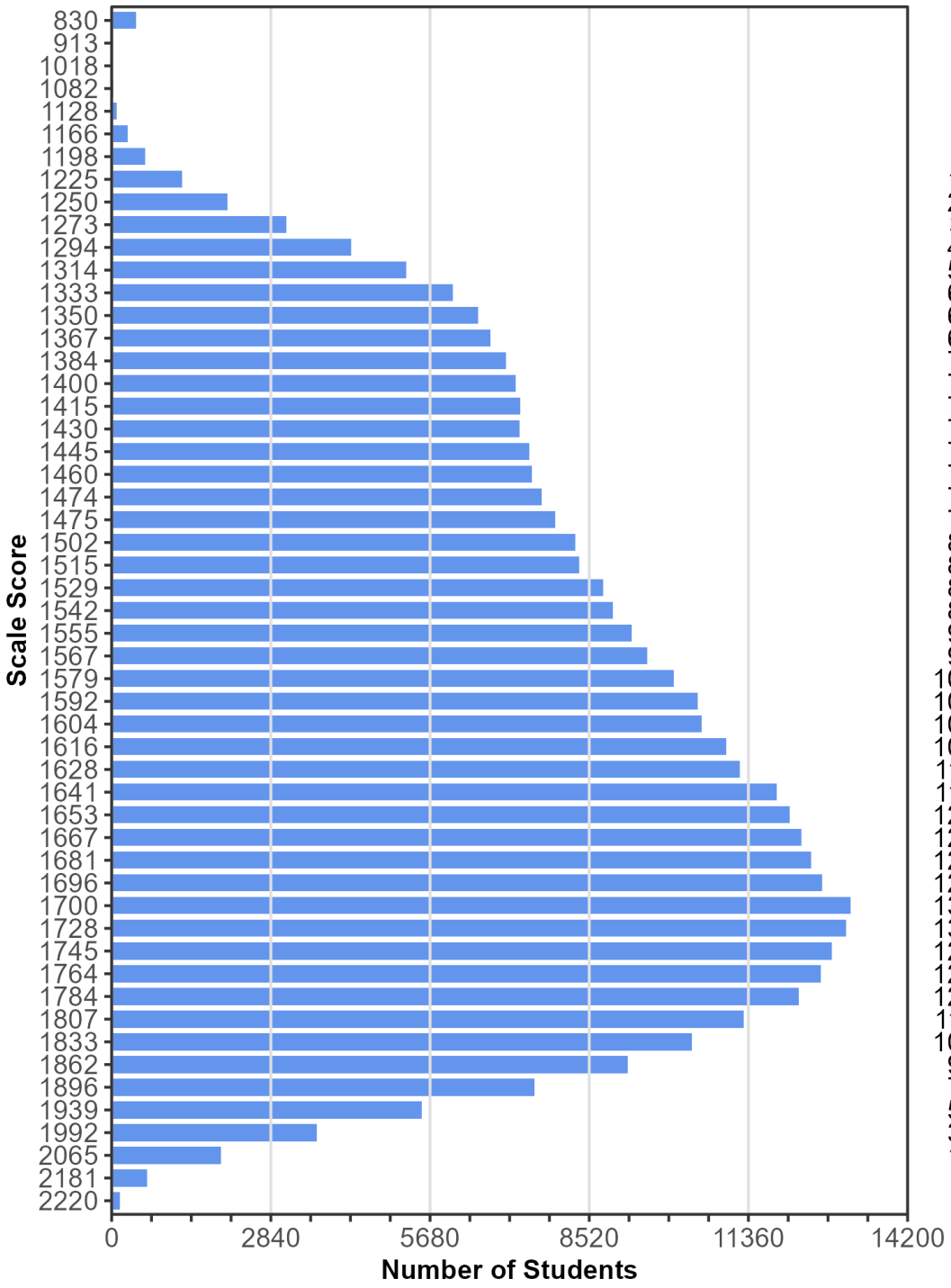
Freq	Cum Freq	Pct	Cum Pct
414	414	0.11	0.11
17	431	0.00	0.12
29	460	0.01	0.13
66	526	0.02	0.14
127	653	0.03	0.18
365	1018	0.10	0.28
862	1880	0.24	0.52
1723	3603	0.47	0.99
2745	6348	0.75	1.75
4431	10779	1.22	2.96
5883	16662	1.62	4.58
7383	24045	2.03	6.61
8333	32378	2.29	8.90
9091	41469	2.50	11.40
9653	51122	2.65	14.06
9856	60978	2.71	16.77
9876	70854	2.72	19.48
9935	80789	2.73	22.22
10113	90902	2.78	25.00
9995	100897	2.75	27.74
10164	111061	2.79	30.54
10497	121558	2.89	33.43
10169	131727	2.80	36.22
10485	142212	2.88	39.11
10490	152702	2.88	41.99
10668	163370	2.93	44.92
10589	173959	2.91	47.84
10901	184860	3.00	50.83
10962	195822	3.01	53.85
10910	206732	3.00	56.85
10901	217633	3.00	59.84
10883	228516	2.99	62.84
10674	239190	2.94	65.77
10470	249660	2.88	68.65
10204	259864	2.81	71.46
9860	269724	2.71	74.17
9442	279166	2.60	76.76
9613	288779	2.64	79.41
9195	297974	2.53	81.94
8822	306796	2.43	84.36
8412	315208	2.31	86.68
7874	323082	2.17	88.84
7384	330466	2.03	90.87
6881	337347	1.89	92.76
6447	343794	1.77	94.54
5786	349580	1.59	96.13
5009	354589	1.38	97.50
4000	358589	1.10	98.60
2796	361385	0.77	99.37
1517	362902	0.42	99.79
616	363518	0.17	99.96
137	363655	0.04	100.00
9	363664	0.00	100.00

Figure B.7.8. Spring 2025 STAAR Grade 4 RLA
Frequency Distribution of Scale Scores



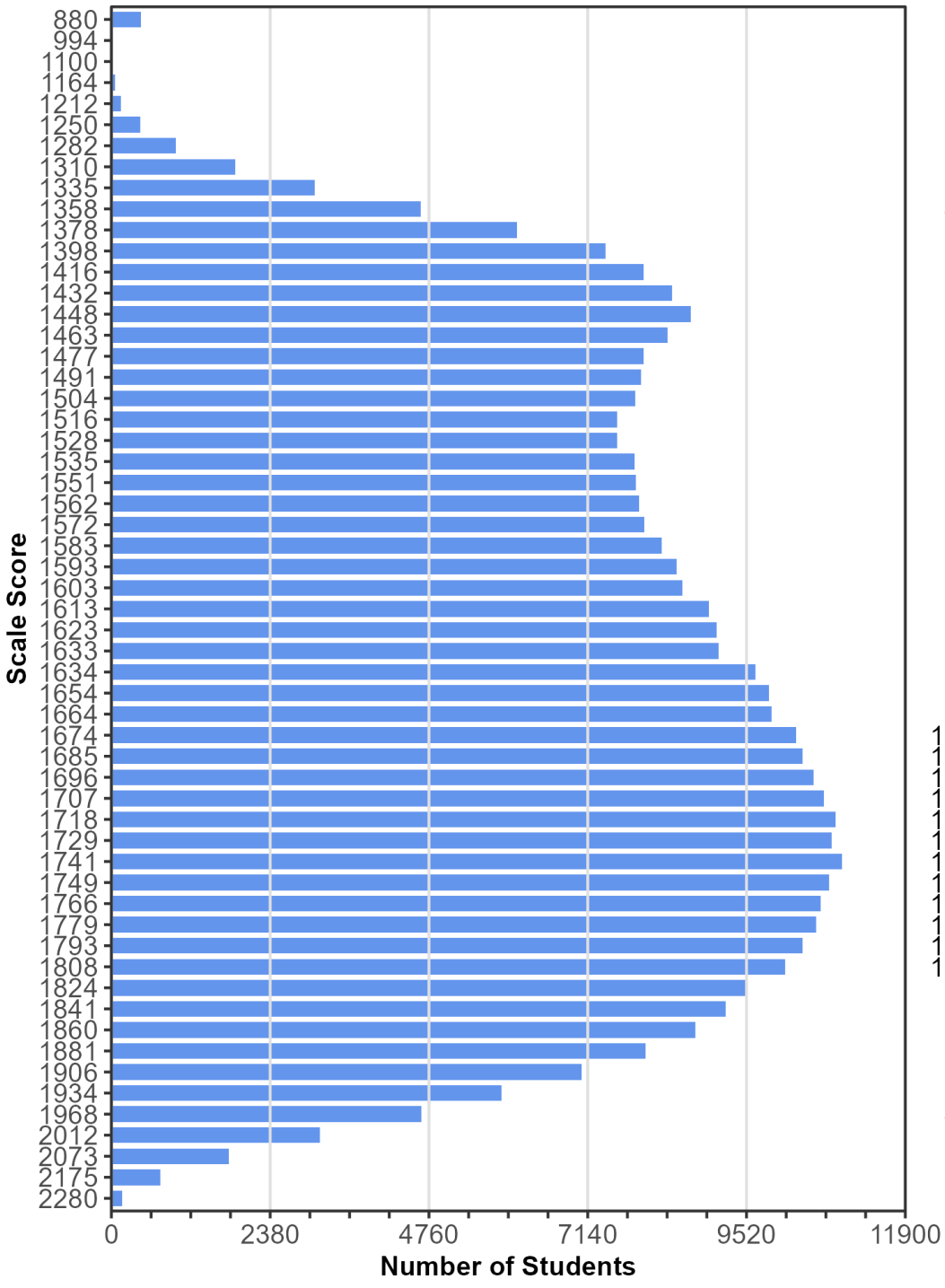
Freq	Cum Freq	Pct	Cum Pct
468	468	0.13	0.13
13	481	0.00	0.13
19	500	0.01	0.13
73	573	0.02	0.15
212	785	0.06	0.21
580	1365	0.16	0.37
1189	2554	0.32	0.69
2211	4765	0.60	1.28
3675	8440	0.99	2.27
5174	13614	1.39	3.66
6569	20183	1.77	5.43
7488	27671	2.02	7.45
8016	35687	2.16	9.61
8416	44103	2.27	11.87
8515	52618	2.29	14.16
8305	60923	2.24	16.40
8260	69183	2.22	18.62
8148	77331	2.19	20.82
8379	85710	2.26	23.07
8548	94258	2.30	25.37
8823	103081	2.38	27.75
8993	112074	2.42	30.17
9366	121440	2.52	32.69
9693	131133	2.61	35.30
9703	140836	2.61	37.91
10089	150925	2.72	40.63
10240	161165	2.76	43.38
10578	171743	2.85	46.23
10690	182433	2.88	49.11
11273	193706	3.03	52.14
10975	204681	2.95	55.10
11202	215883	3.02	58.11
11121	227004	2.99	61.11
11169	238173	3.01	64.11
11177	249350	3.01	67.12
11211	260561	3.02	70.14
11112	271673	2.99	73.13
11003	282676	2.96	76.09
10707	293383	2.88	78.98
10409	303792	2.80	81.78
9973	313765	2.68	84.46
9321	323086	2.51	86.97
8954	332040	2.41	89.38
8277	340317	2.23	91.61
7447	347764	2.00	93.62
6526	354290	1.76	95.37
5502	359792	1.48	96.85
4532	364324	1.22	98.07
3274	367598	0.88	98.96
2103	369701	0.57	99.52
1185	370886	0.32	99.84
480	371366	0.13	99.97
112	371478	0.03	100.00

Figure B.7.9. Spring 2025 STAAR Grade 5 RLA
Frequency Distribution of Scale Scores



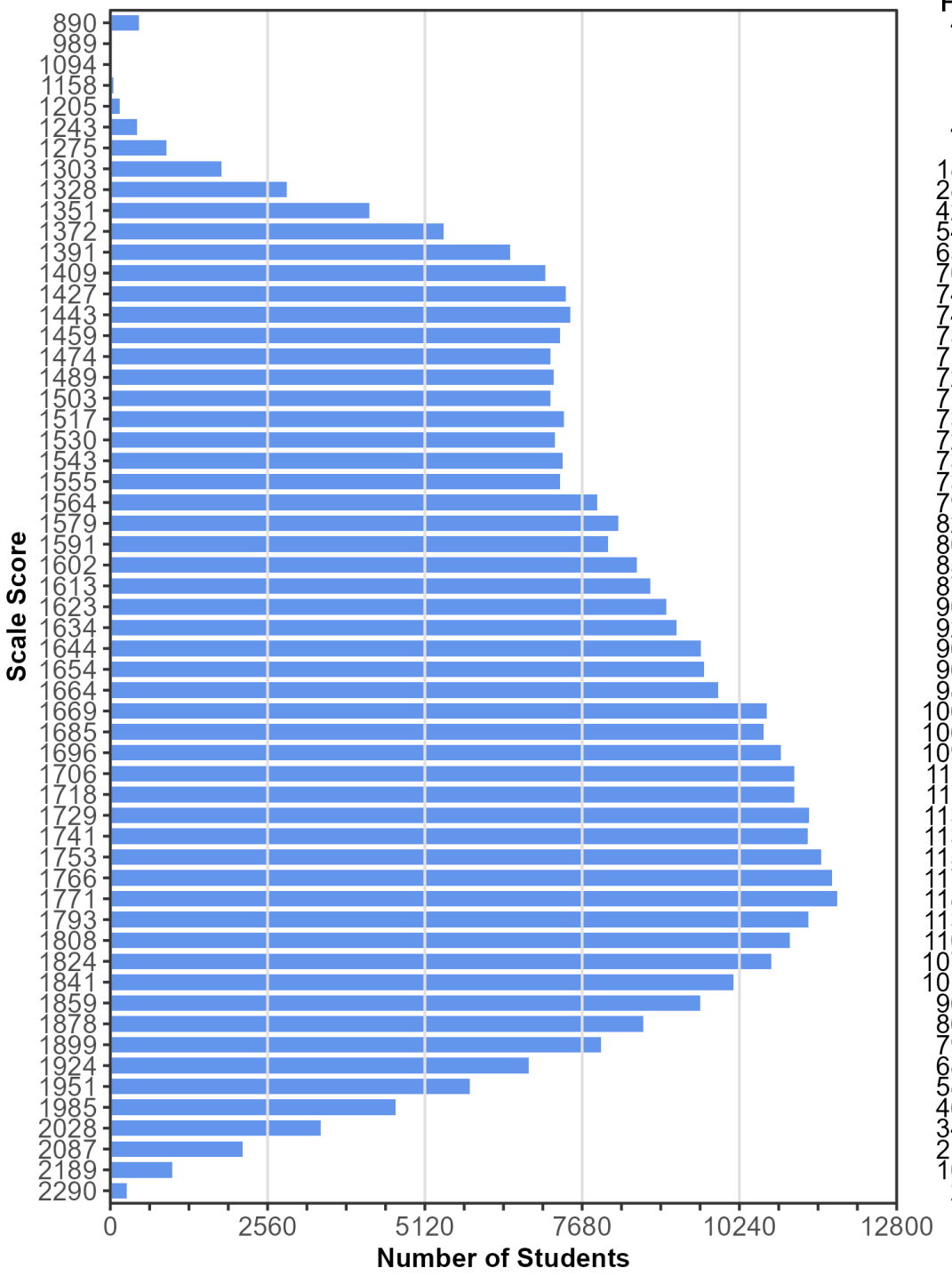
Freq	Cum Freq	Pct	Cum Pct
438	438	0.12	0.12
7	445	0.00	0.12
15	460	0.00	0.12
31	491	0.01	0.13
92	583	0.02	0.15
285	868	0.08	0.23
596	1464	0.16	0.39
1255	2719	0.33	0.72
2061	4780	0.54	1.26
3114	7894	0.82	2.09
4274	12168	1.13	3.22
5257	17425	1.39	4.61
6086	23511	1.61	6.22
6537	30048	1.73	7.94
6761	36809	1.79	9.73
7038	43847	1.86	11.59
7204	51051	1.90	13.50
7290	58341	1.93	15.42
7283	65624	1.93	17.35
7448	73072	1.97	19.32
7498	80570	1.98	21.30
7671	88241	2.03	23.33
7910	96151	2.09	25.42
8268	104419	2.19	27.60
8336	112755	2.20	29.81
8764	121519	2.32	32.13
8944	130463	2.36	34.49
9272	139735	2.45	36.94
9558	149293	2.53	39.47
10029	159322	2.65	42.12
10458	169780	2.76	44.88
10526	180306	2.78	47.67
10970	191276	2.90	50.57
11210	202486	2.96	53.53
11867	214353	3.14	56.67
12102	226455	3.20	59.87
12300	238755	3.25	63.12
12483	251238	3.30	66.42
12680	263918	3.35	69.77
13184	277102	3.49	73.26
13103	290205	3.46	76.72
12844	303049	3.40	80.12
12653	315702	3.35	83.46
12258	327960	3.24	86.70
11280	339240	2.98	89.68
10356	349596	2.74	92.42
9209	358805	2.43	94.86
7546	366351	1.99	96.85
5530	371881	1.46	98.31
3660	375541	0.97	99.28
1946	377487	0.51	99.80
631	378118	0.17	99.96
144	378262	0.04	100.00

Figure B.7.10. Spring 2025 STAAR Grade 6 RLA
Frequency Distribution of Scale Scores



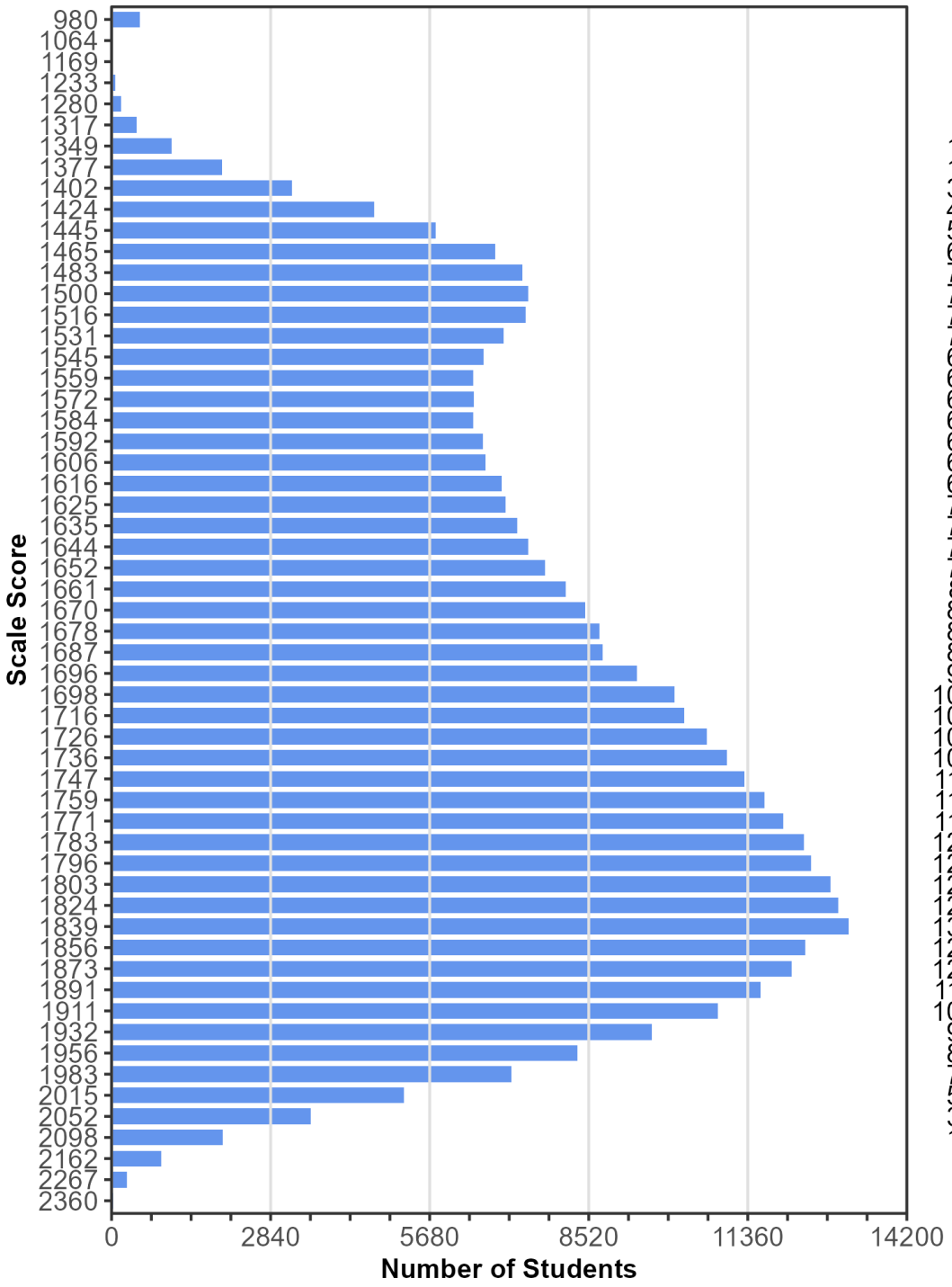
Freq	Cum Freq	Pct	Cum Pct
446	446	0.11	0.11
9	455	0.00	0.11
13	468	0.00	0.12
53	521	0.01	0.13
139	660	0.04	0.17
432	1092	0.11	0.28
968	2060	0.24	0.52
1858	3918	0.47	0.99
3048	6966	0.77	1.76
4634	11600	1.17	2.93
6081	17681	1.53	4.46
7410	25091	1.87	6.33
7974	33065	2.01	8.34
8402	41467	2.12	10.46
8688	50155	2.19	12.65
8332	58487	2.10	14.75
7978	66465	2.01	16.76
7941	74406	2.00	18.77
7850	82256	1.98	20.75
7584	89840	1.91	22.66
7584	97424	1.91	24.57
7839	105263	1.98	26.55
7860	113123	1.98	28.53
7907	121030	1.99	30.52
7989	129019	2.01	32.54
8248	137267	2.08	34.62
8476	145743	2.14	36.76
8557	154300	2.16	38.92
8957	163257	2.26	41.17
9075	172332	2.29	43.46
9098	181430	2.29	45.76
9656	191086	2.44	48.19
9860	200946	2.49	50.68
9892	210838	2.49	53.18
10263	221101	2.59	55.76
10364	231465	2.61	58.38
10524	241989	2.65	61.03
10683	252672	2.69	63.73
10850	263522	2.74	66.46
10800	274322	2.72	69.19
10949	285271	2.76	71.95
10761	296032	2.71	74.66
10627	306659	2.68	77.34
10564	317223	2.66	80.01
10362	327585	2.61	82.62
10095	337680	2.55	85.17
9524	347204	2.40	87.57
9206	356410	2.32	89.89
8753	365163	2.21	92.10
8011	373174	2.02	94.12
7050	380224	1.78	95.90
5844	386068	1.47	97.37
4645	390713	1.17	98.54
3128	393841	0.79	99.33
1762	395603	0.44	99.77
735	396338	0.19	99.96
159	396497	0.04	100.00

Figure B.7.11. Spring 2025 STAAR Grade 7 RLA
Frequency Distribution of Scale Scores



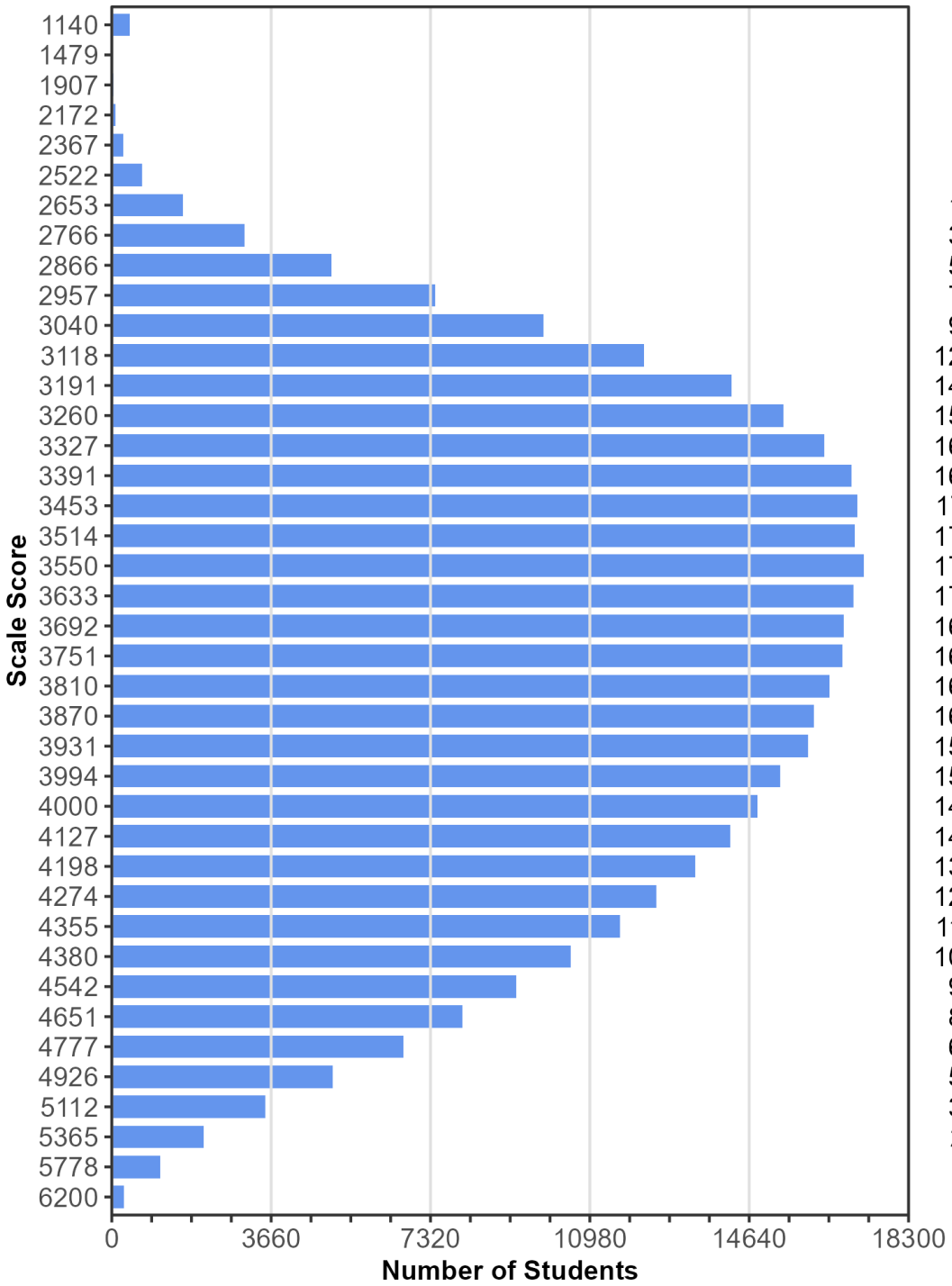
Freq	Cum Freq	Pct	Cum Pct
466	466	0.12	0.12
11	477	0.00	0.12
14	491	0.00	0.12
50	541	0.01	0.14
156	697	0.04	0.18
430	1127	0.11	0.28
911	2038	0.23	0.51
1808	3846	0.45	0.97
2869	6715	0.72	1.69
4212	10927	1.06	2.74
5424	16351	1.36	4.11
6508	22859	1.63	5.74
7084	29943	1.78	7.52
7410	37353	1.86	9.38
7492	44845	1.88	11.26
7319	52164	1.84	13.10
7160	59324	1.80	14.90
7219	66543	1.81	16.71
7161	73704	1.80	18.51
7385	81089	1.85	20.36
7240	88329	1.82	22.18
7359	95688	1.85	24.03
7325	103013	1.84	25.87
7924	110937	1.99	27.86
8265	119202	2.08	29.93
8099	127301	2.03	31.97
8568	135869	2.15	34.12
8793	144662	2.21	36.32
9049	153711	2.27	38.60
9218	162929	2.31	40.91
9608	172537	2.41	43.32
9668	182205	2.43	45.75
9899	192104	2.49	48.24
10685	202789	2.68	50.92
10629	213418	2.67	53.59
10913	224331	2.74	56.33
11130	235461	2.79	59.12
11137	246598	2.80	61.92
11369	257967	2.85	64.78
11351	269318	2.85	67.63
11575	280893	2.91	70.53
11748	292641	2.95	73.48
11829	304470	2.97	76.45
11365	315835	2.85	79.31
11063	326898	2.78	82.08
10763	337661	2.70	84.79
10146	347807	2.55	87.33
9600	357407	2.41	89.75
8677	366084	2.18	91.92
7992	374076	2.01	93.93
6814	380890	1.71	95.64
5853	386743	1.47	97.11
4640	391383	1.17	98.28
3429	394812	0.86	99.14
2158	396970	0.54	99.68
1005	397975	0.25	99.93
270	398245	0.07	100.00

Figure B.7.12. Spring 2025 STAAR Grade 8 RLA
Frequency Distribution of Scale Scores



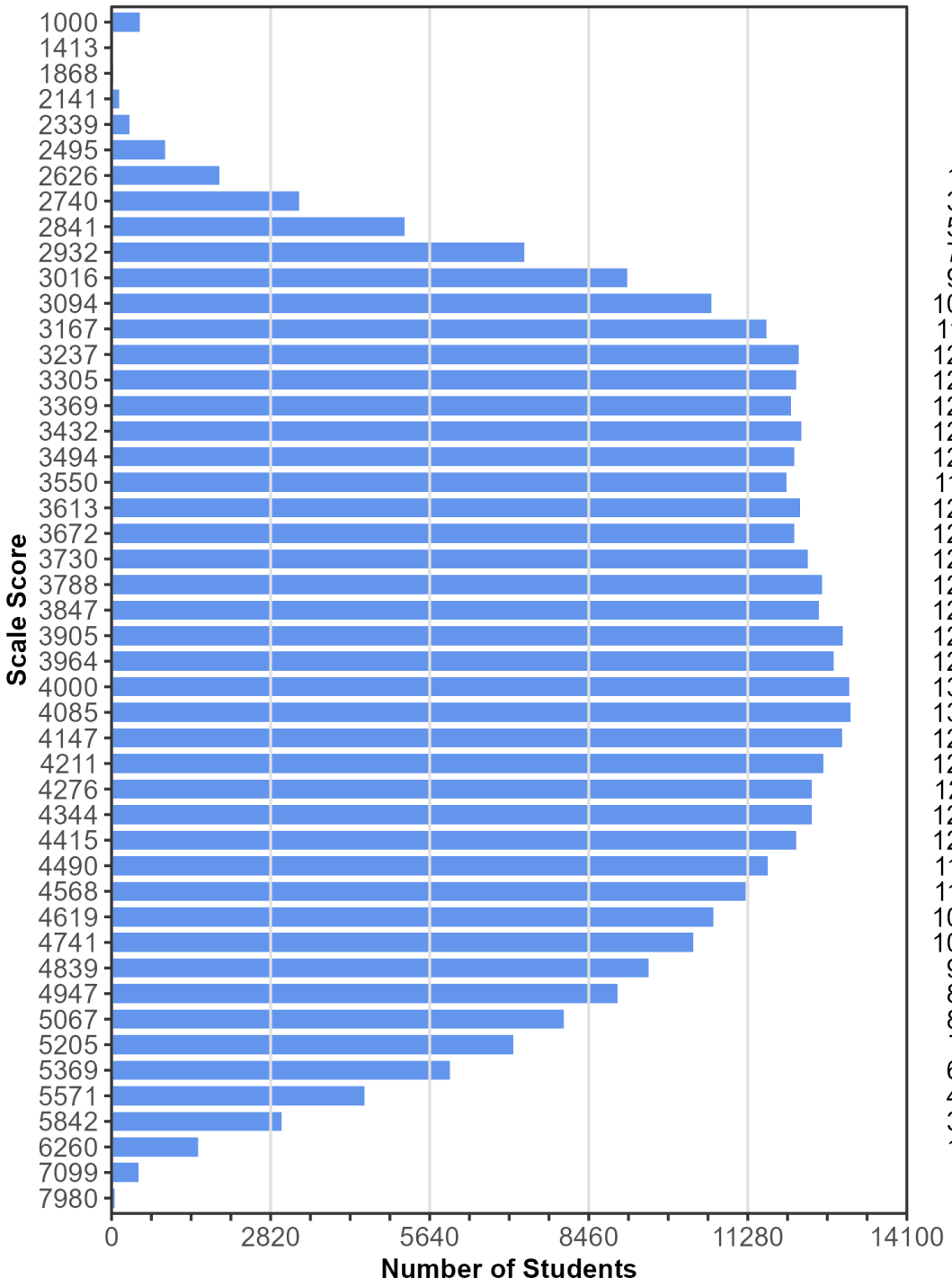
Freq	Cum Freq	Pct	Cum Pct
508	508	0.13	0.13
9	517	0.00	0.13
21	538	0.01	0.14
62	600	0.02	0.15
175	775	0.04	0.19
452	1227	0.11	0.31
1069	2296	0.27	0.58
1968	4264	0.49	1.07
3223	7487	0.81	1.88
4692	12179	1.18	3.06
5787	17966	1.45	4.52
6852	24818	1.72	6.24
7335	32153	1.84	8.08
7444	39597	1.87	9.95
7389	46986	1.86	11.81
7004	53990	1.76	13.57
6645	60635	1.67	15.24
6457	67092	1.62	16.86
6465	73557	1.63	18.49
6458	80015	1.62	20.11
6634	86649	1.67	21.78
6677	93326	1.68	23.46
6965	100291	1.75	25.21
7033	107324	1.77	26.98
7240	114564	1.82	28.80
7441	122005	1.87	30.67
7744	129749	1.95	32.61
8115	137864	2.04	34.65
8457	146321	2.13	36.78
8708	155029	2.19	38.97
8766	163795	2.20	41.17
9378	173173	2.36	43.53
10048	183221	2.53	46.06
10227	193448	2.57	48.63
10631	204079	2.67	51.30
10985	215064	2.76	54.06
11300	226364	2.84	56.90
11654	238018	2.93	59.83
11998	250016	3.02	62.85
12360	262376	3.11	65.95
12489	274865	3.14	69.09
12835	287700	3.23	72.32
12973	300673	3.26	75.58
13160	313833	3.31	78.89
12388	326221	3.11	82.00
12142	338363	3.05	85.05
11592	349955	2.91	87.97
10827	360782	2.72	90.69
9649	370431	2.43	93.11
8314	378745	2.09	95.20
7141	385886	1.80	97.00
5218	391104	1.31	98.31
3552	394656	0.89	99.20
1980	396636	0.50	99.70
887	397523	0.22	99.92
274	397797	0.07	99.99
28	397825	0.01	100.00

Figure B.7.13. Spring 2025 STAAR Grade 5 Science
Frequency Distribution of Scale Scores



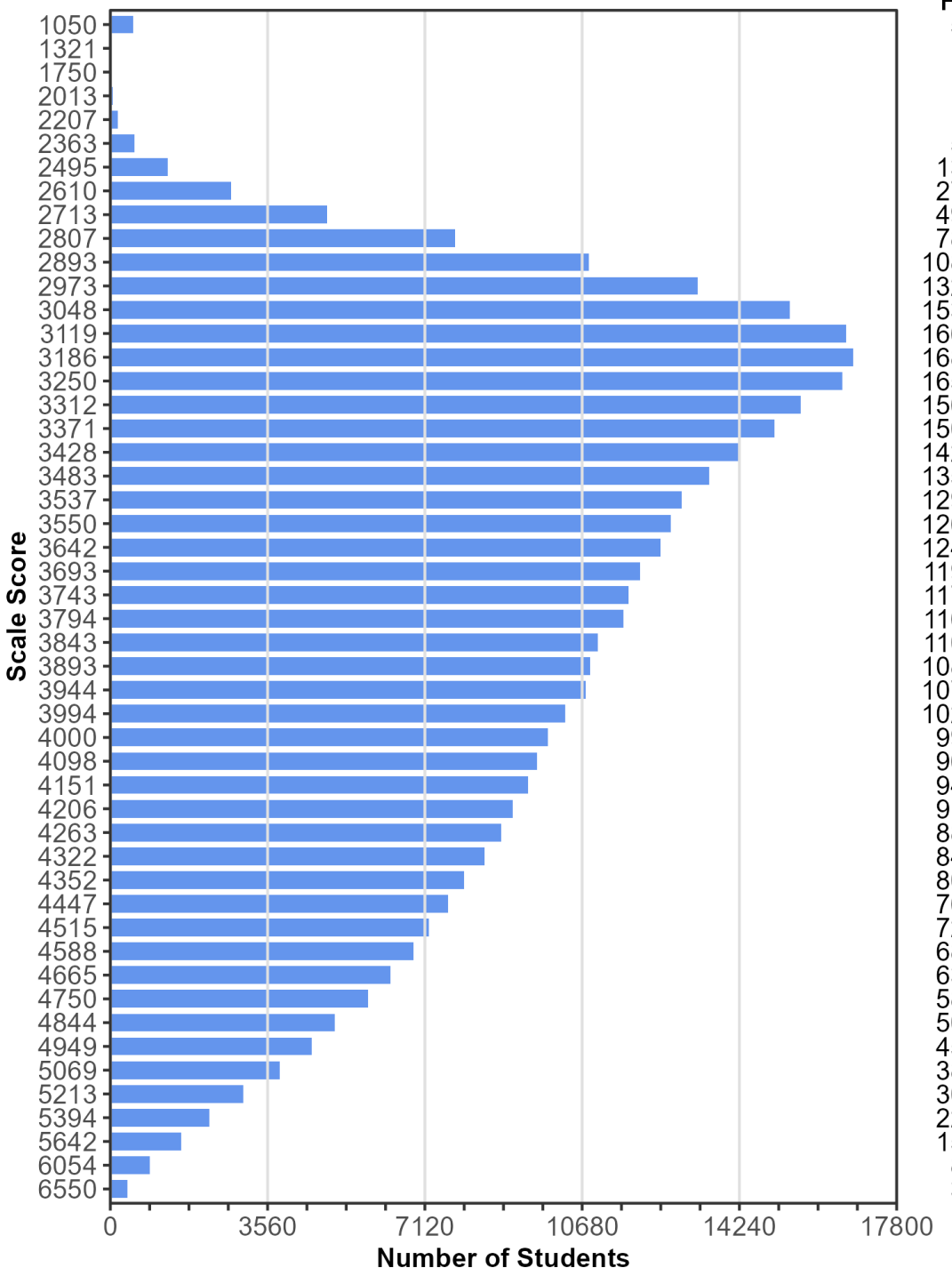
	Freq	Cum Freq	Pct	Cum Pct
1140	408	408	0.11	0.11
1479	11	419	0.00	0.11
1907	35	454	0.01	0.12
2172	90	544	0.02	0.14
2367	259	803	0.07	0.21
2522	697	1500	0.18	0.39
2653	1633	3133	0.43	0.82
2766	3049	6182	0.80	1.61
2866	5039	11221	1.32	2.93
2957	7425	18646	1.94	4.87
3040	9920	28566	2.59	7.46
3118	12225	40791	3.19	10.65
3191	14239	55030	3.72	14.36
3260	15427	70457	4.03	18.39
3327	16358	86815	4.27	22.66
3391	16985	103800	4.43	27.09
3453	17117	120917	4.47	31.56
3514	17064	137981	4.45	36.01
3550	17271	155252	4.51	40.52
3633	17033	172285	4.45	44.97
3692	16813	189098	4.39	49.35
3751	16780	205878	4.38	53.73
3810	16488	222366	4.30	58.04
3870	16133	238499	4.21	62.25
3931	15990	254489	4.17	66.42
3994	15356	269845	4.01	70.43
4000	14826	284671	3.87	74.30
4127	14203	298874	3.71	78.01
4198	13397	312271	3.50	81.50
4274	12506	324777	3.26	84.77
4355	11674	336451	3.05	87.81
4380	10535	346986	2.75	90.56
4542	9293	356279	2.43	92.99
4651	8060	364339	2.10	95.09
4777	6699	371038	1.75	96.84
4926	5079	376117	1.33	98.17
5112	3533	379650	0.92	99.09
5365	2113	381763	0.55	99.64
5778	1109	382872	0.29	99.93
6200	275	383147	0.07	100.00

Figure B.7.14. Spring 2025 STAAR Grade 8 Science
Frequency Distribution of Scale Scores



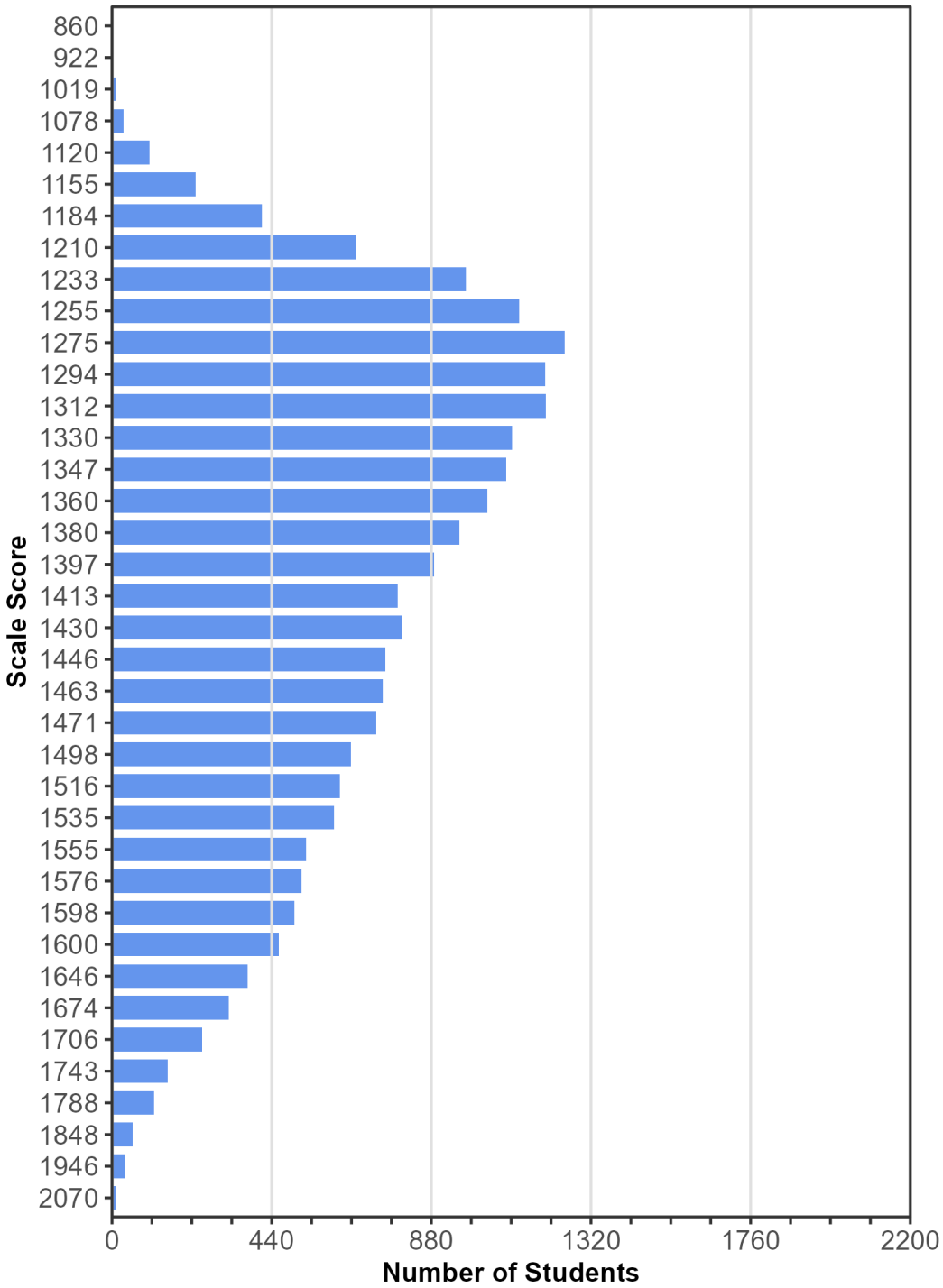
	Freq	Cum Freq	Pct	Cum Pct
1000	497	497	0.13	0.13
1413	21	518	0.01	0.13
1868	28	546	0.01	0.14
2141	130	676	0.03	0.17
2339	321	997	0.08	0.25
2495	945	1942	0.24	0.49
2626	1909	3851	0.49	0.98
2740	3327	7178	0.85	1.83
2841	5196	12374	1.32	3.15
2932	7324	19698	1.86	5.01
3016	9138	28836	2.32	7.34
3094	10640	39476	2.71	10.04
3167	11611	51087	2.95	13.00
3237	12179	63266	3.10	16.10
3305	12138	75404	3.09	19.18
3369	12041	87445	3.06	22.25
3432	12228	99673	3.11	25.36
3494	12098	111771	3.08	28.44
3550	11965	123736	3.04	31.48
3613	12203	135939	3.10	34.59
3672	12099	148038	3.08	37.66
3730	12341	160379	3.14	40.80
3788	12594	172973	3.20	44.01
3847	12538	185511	3.19	47.20
3905	12962	198473	3.30	50.50
3964	12808	211281	3.26	53.75
4000	13084	224365	3.33	57.08
4085	13098	237463	3.33	60.42
4147	12950	250413	3.29	63.71
4211	12625	263038	3.21	66.92
4276	12411	275449	3.16	70.08
4344	12413	287862	3.16	73.24
4415	12134	299996	3.09	76.32
4490	11628	311624	2.96	79.28
4568	11239	322863	2.86	82.14
4619	10672	333535	2.72	84.86
4741	10314	343849	2.62	87.48
4839	9521	353370	2.42	89.90
4947	8967	362337	2.28	92.19
5067	8016	370353	2.04	94.23
5205	7119	377472	1.81	96.04
5369	6001	383473	1.53	97.56
5571	4487	387960	1.14	98.70
5842	3018	390978	0.77	99.47
6260	1540	392518	0.39	99.86
7099	481	392999	0.12	99.99
7980	52	393051	0.01	100.00

Figure B.7.15. Spring 2025 STAAR Grade 8 Social Studies
Frequency Distribution of Scale Scores



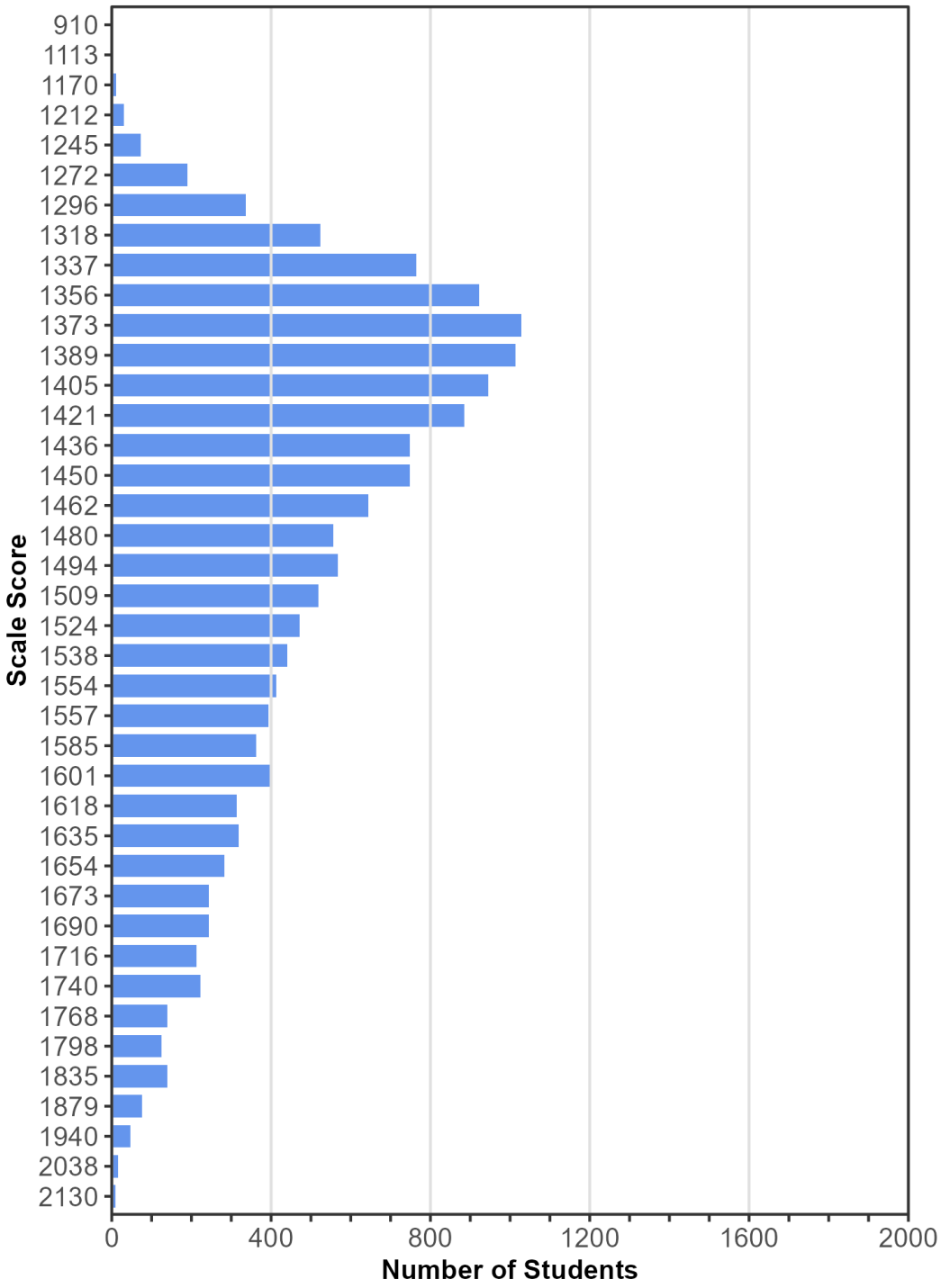
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1050	512	512	0.13	0.13
1321	11	523	0.00	0.13
1750	15	538	0.00	0.13
2013	53	591	0.01	0.15
2207	166	757	0.04	0.19
2363	542	1299	0.14	0.32
2495	1306	2605	0.33	0.65
2610	2741	5346	0.68	1.33
2713	4909	10255	1.22	2.56
2807	7803	18058	1.94	4.50
2893	10830	28888	2.70	7.20
2973	13289	42177	3.31	10.51
3048	15376	57553	3.83	14.34
3119	16652	74205	4.15	18.49
3186	16820	91025	4.19	22.68
3250	16573	107598	4.13	26.81
3312	15627	123225	3.89	30.70
3371	15031	138256	3.75	34.45
3428	14204	152460	3.54	37.99
3483	13563	166023	3.38	41.37
3537	12932	178955	3.22	44.59
3550	12692	191647	3.16	47.75
3642	12461	204108	3.10	50.86
3693	11986	216094	2.99	53.84
3743	11724	227818	2.92	56.76
3794	11612	239430	2.89	59.66
3843	11037	250467	2.75	62.41
3893	10859	261326	2.71	65.11
3944	10759	272085	2.68	67.79
3994	10297	282382	2.57	70.36
4000	9910	292292	2.47	72.83
4098	9664	301956	2.41	75.24
4151	9461	311417	2.36	77.59
4206	9109	320526	2.27	79.86
4263	8853	329379	2.21	82.07
4322	8467	337846	2.11	84.18
4352	8004	345850	1.99	86.17
4447	7650	353500	1.91	88.08
4515	7217	360717	1.80	89.88
4588	6864	367581	1.71	91.59
4665	6337	373918	1.58	93.17
4750	5834	379752	1.45	94.62
4844	5081	384833	1.27	95.89
4949	4554	389387	1.13	97.02
5069	3837	393224	0.96	97.98
5213	3010	396234	0.75	98.73
5394	2242	398476	0.56	99.28
5642	1599	400075	0.40	99.68
6054	890	400965	0.22	99.91
6550	381	401346	0.09	100.00

Figure B.7.16. Spring 2025 STAAR Spanish Grade 3 Mathematics
Frequency Distribution of Scale Scores



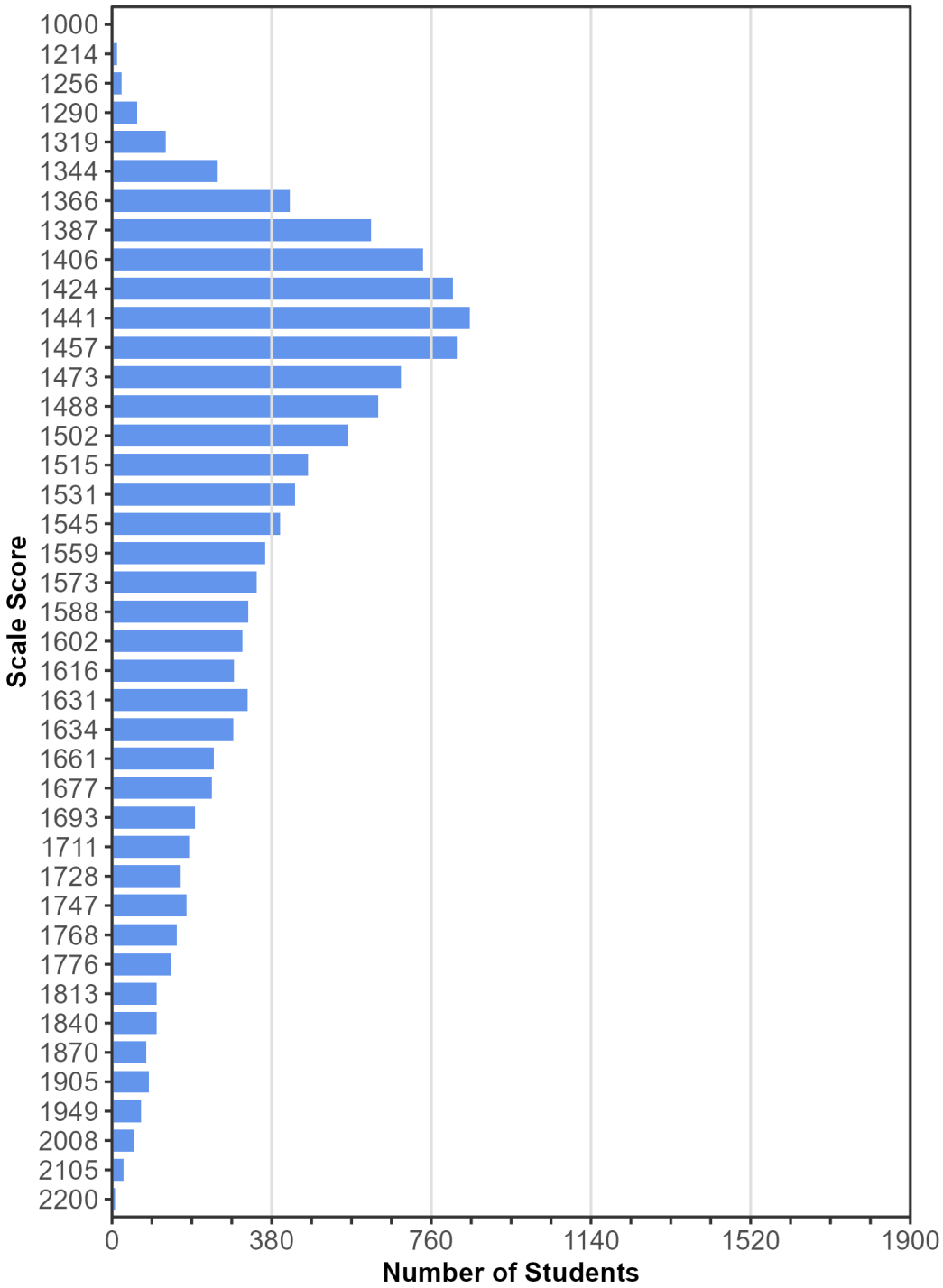
	Freq	Cum Freq	Pct	Cum Pct
860	3	3	0.01	0.01
922	2	5	0.01	0.02
1019	12	17	0.06	0.08
1078	32	49	0.15	0.23
1120	103	152	0.48	0.71
1155	231	383	1.08	1.80
1184	413	796	1.94	3.73
1210	673	1469	3.16	6.89
1233	976	2445	4.58	11.47
1255	1122	3567	5.26	16.73
1275	1248	4815	5.85	22.59
1294	1193	6008	5.60	28.18
1312	1196	7204	5.61	33.79
1330	1102	8306	5.17	38.96
1347	1086	9392	5.09	44.06
1360	1035	10427	4.86	48.91
1380	957	11384	4.49	53.40
1397	887	12271	4.16	57.56
1413	788	13059	3.70	61.26
1430	799	13858	3.75	65.01
1446	753	14611	3.53	68.54
1463	747	15358	3.50	72.04
1471	728	16086	3.41	75.46
1498	659	16745	3.09	78.55
1516	628	17373	2.95	81.49
1535	611	17984	2.87	84.36
1555	535	18519	2.51	86.87
1576	523	19042	2.45	89.32
1598	503	19545	2.36	91.68
1600	460	20005	2.16	93.84
1646	374	20379	1.75	95.60
1674	321	20700	1.51	97.10
1706	248	20948	1.16	98.26
1743	154	21102	0.72	98.99
1788	115	21217	0.54	99.53
1848	56	21273	0.26	99.79
1946	35	21308	0.16	99.95
2070	10	21318	0.05	100.00

Figure B.7.17. Spring 2025 STAAR Spanish Grade 4 Mathematics
Frequency Distribution of Scale Scores



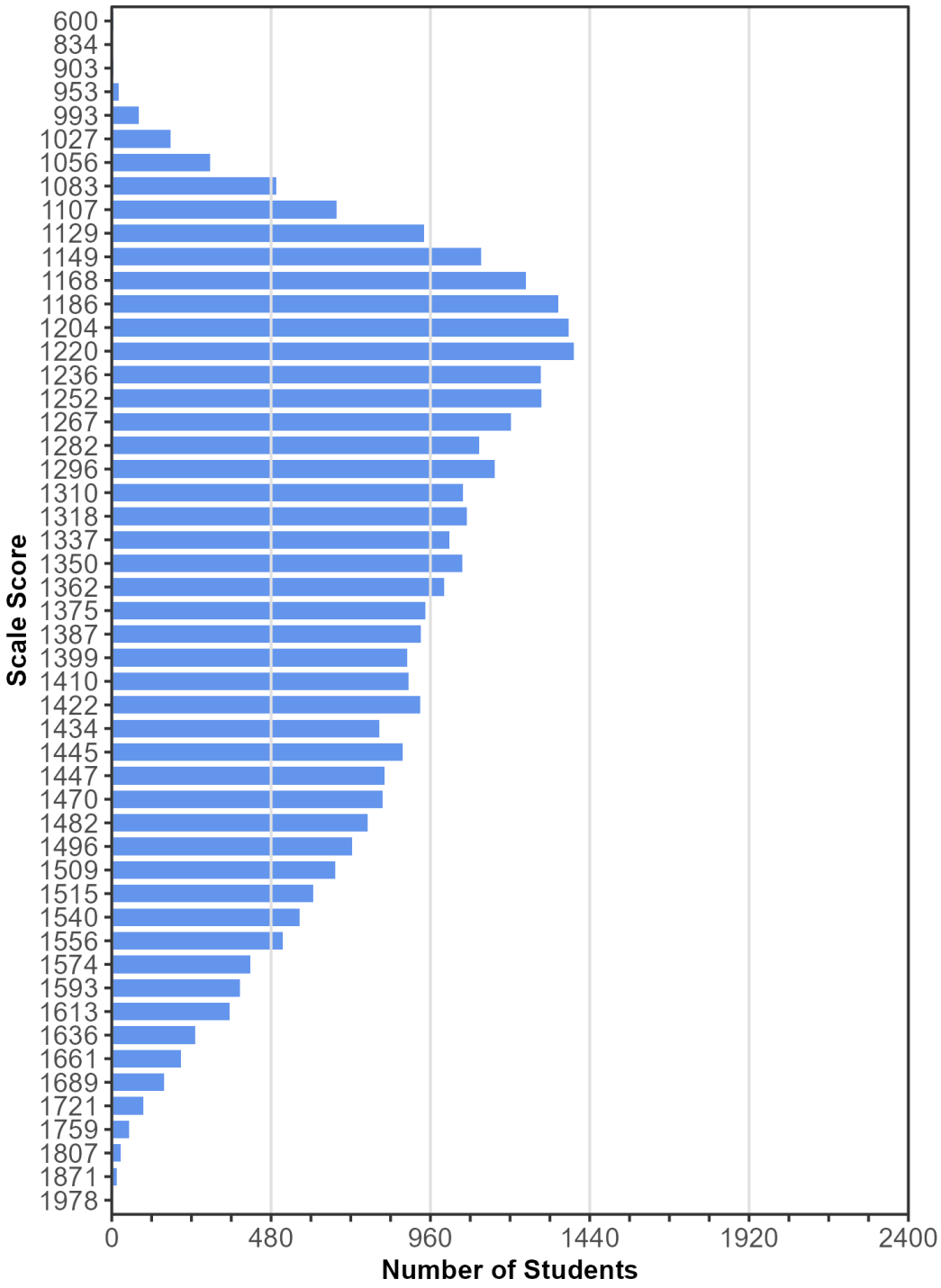
	Freq	Cum Freq	Pct	Cum Pct
910	3	3	0.02	0.02
1113	1	4	0.01	0.03
1170	11	15	0.07	0.10
1212	30	45	0.20	0.29
1245	73	118	0.47	0.77
1272	190	308	1.24	2.00
1296	337	645	2.19	4.19
1318	523	1168	3.40	7.59
1337	765	1933	4.97	12.57
1356	922	2855	6.00	18.56
1373	1029	3884	6.69	25.26
1389	1013	4897	6.59	31.84
1405	945	5842	6.14	37.99
1421	885	6727	5.75	43.74
1436	748	7475	4.86	48.61
1450	749	8224	4.87	53.48
1462	644	8868	4.19	57.66
1480	556	9424	3.62	61.28
1494	568	9992	3.69	64.97
1509	519	10511	3.37	68.35
1524	472	10983	3.07	71.42
1538	440	11423	2.86	74.28
1554	413	11836	2.69	76.96
1557	393	12229	2.56	79.52
1585	362	12591	2.35	81.87
1601	397	12988	2.58	84.45
1618	314	13302	2.04	86.49
1635	318	13620	2.07	88.56
1654	283	13903	1.84	90.40
1673	243	14146	1.58	91.98
1690	244	14390	1.59	93.57
1716	212	14602	1.38	94.95
1740	223	14825	1.45	96.40
1768	140	14965	0.91	97.31
1798	125	15090	0.81	98.12
1835	140	15230	0.91	99.03
1879	76	15306	0.49	99.53
1940	47	15353	0.31	99.83
2038	16	15369	0.10	99.93
2130	10	15379	0.07	100.00

Figure B.7.18. Spring 2025 STAAR Spanish Grade 5 Mathematics
Frequency Distribution of Scale Scores



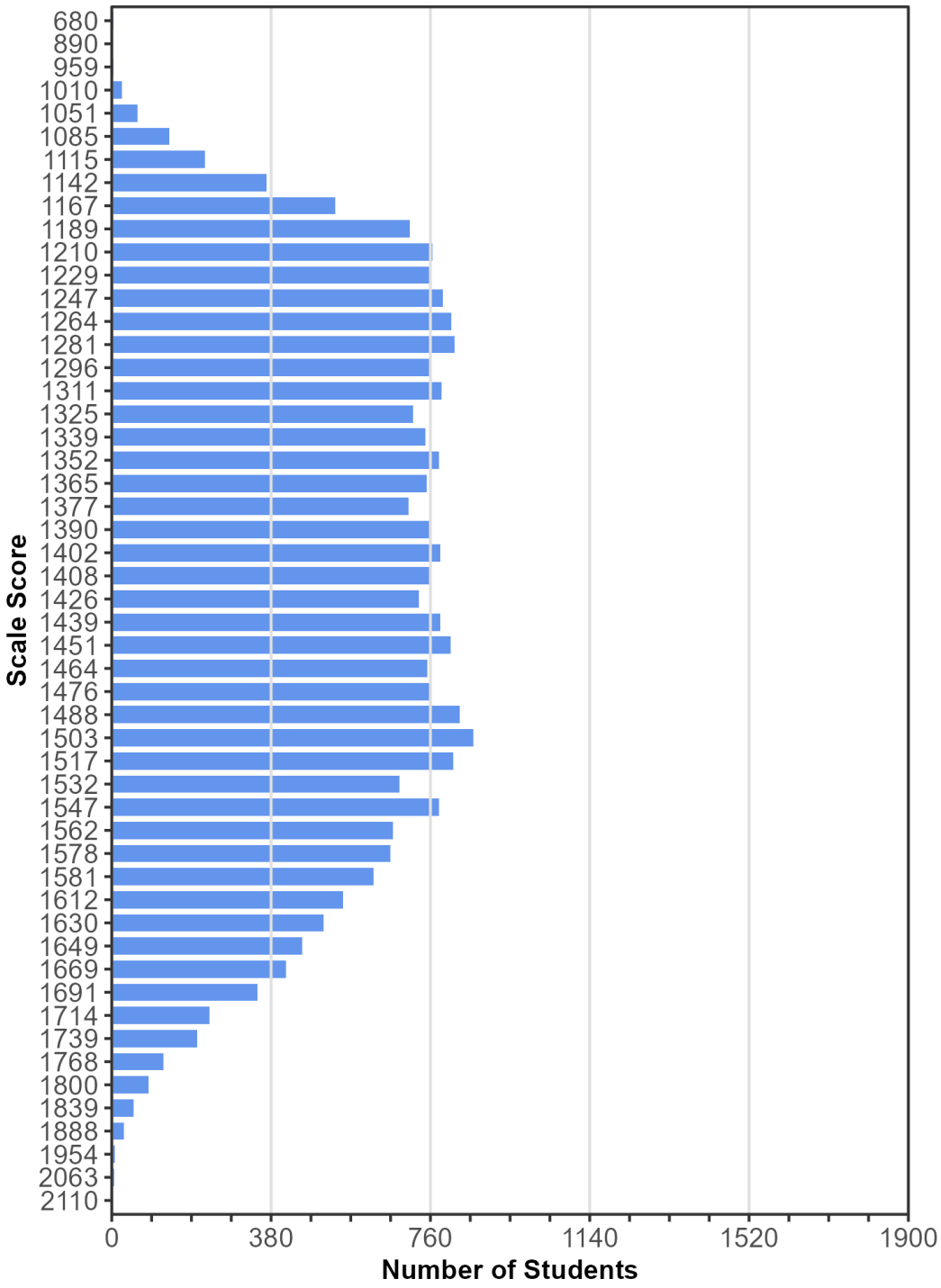
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1000	2	2	0.02	0.02
1214	12	14	0.10	0.11
1256	22	36	0.18	0.29
1290	60	96	0.49	0.79
1319	128	224	1.05	1.83
1344	252	476	2.06	3.90
1366	423	899	3.46	7.36
1387	617	1516	5.05	12.42
1406	740	2256	6.06	18.48
1424	811	3067	6.64	25.12
1441	851	3918	6.97	32.09
1457	821	4739	6.73	38.82
1473	688	5427	5.64	44.45
1488	633	6060	5.19	49.64
1502	562	6622	4.60	54.24
1515	466	7088	3.82	58.06
1531	436	7524	3.57	61.63
1545	400	7924	3.28	64.91
1559	364	8288	2.98	67.89
1573	345	8633	2.83	70.72
1588	324	8957	2.65	73.37
1602	310	9267	2.54	75.91
1616	291	9558	2.38	78.29
1631	323	9881	2.65	80.94
1634	289	10170	2.37	83.31
1661	243	10413	1.99	85.30
1677	238	10651	1.95	87.25
1693	198	10849	1.62	88.87
1711	184	11033	1.51	90.38
1728	164	11197	1.34	91.72
1747	178	11375	1.46	93.18
1768	154	11529	1.26	94.44
1776	140	11669	1.15	95.58
1813	106	11775	0.87	96.45
1840	107	11882	0.88	97.33
1870	81	11963	0.66	97.99
1905	88	12051	0.72	98.71
1949	69	12120	0.57	99.28
2008	52	12172	0.43	99.71
2105	28	12200	0.23	99.93
2200	8	12208	0.07	100.00

Figure B.7.19. Spring 2025 STAAR Spanish Grade 3 RLA
Frequency Distribution of Scale Scores



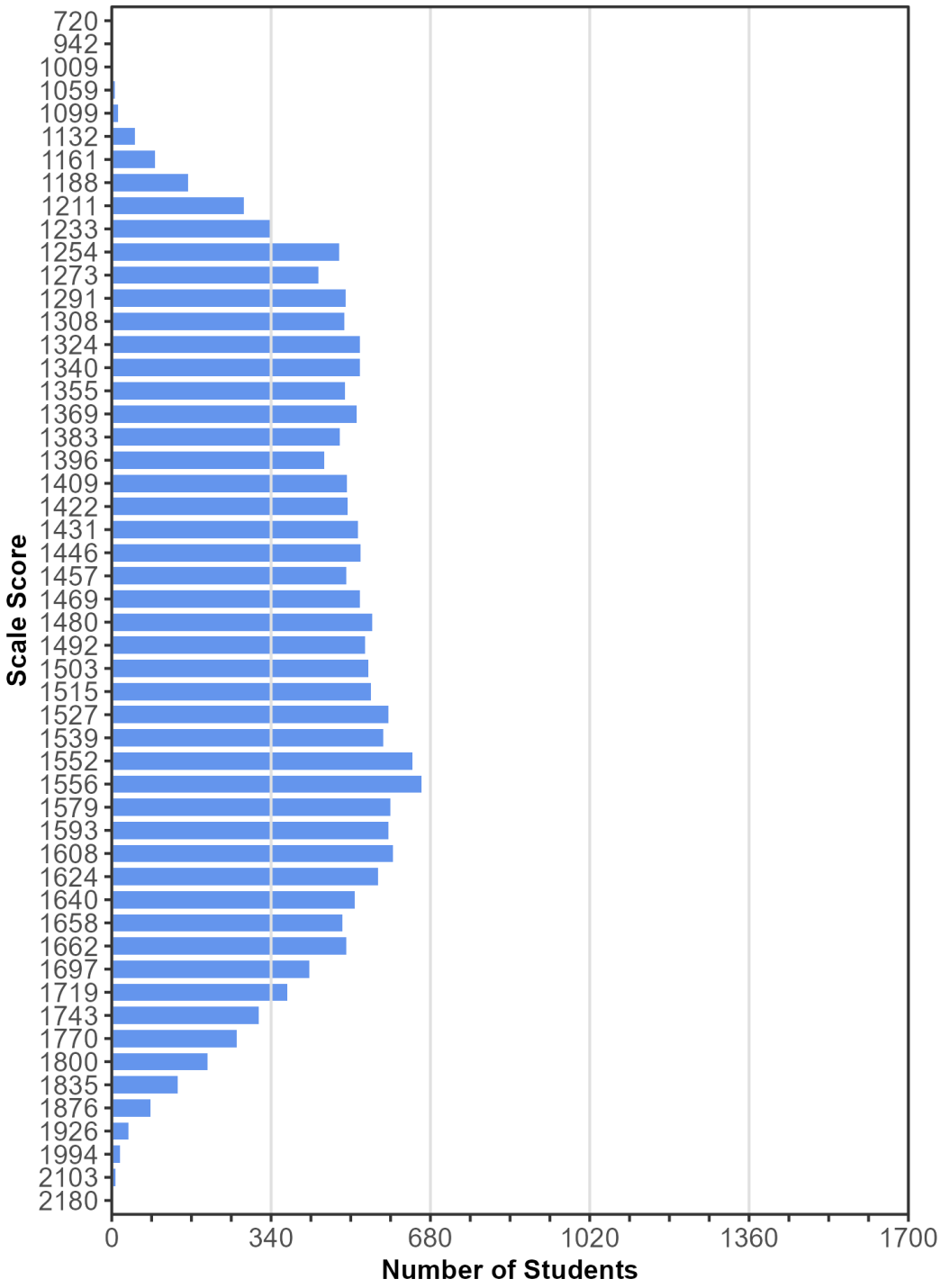
Scale Score	Freq	Cum Freq	Pct	Cum Pct
600	5	5	0.01	0.01
834	3	8	0.01	0.02
903	6	14	0.02	0.04
953	20	34	0.06	0.10
993	81	115	0.24	0.34
1027	177	292	0.52	0.85
1056	296	588	0.87	1.72
1083	496	1084	1.45	3.17
1107	677	1761	1.98	5.15
1129	941	2702	2.75	7.91
1149	1112	3814	3.25	11.16
1168	1248	5062	3.65	14.82
1186	1345	6407	3.94	18.75
1204	1376	7783	4.03	22.78
1220	1393	9176	4.08	26.86
1236	1293	10469	3.78	30.64
1252	1295	11764	3.79	34.43
1267	1203	12967	3.52	37.96
1282	1107	14074	3.24	41.20
1296	1154	15228	3.38	44.57
1310	1058	16286	3.10	47.67
1318	1069	17355	3.13	50.80
1337	1018	18373	2.98	53.78
1350	1056	19429	3.09	56.87
1362	1002	20431	2.93	59.80
1375	944	21375	2.76	62.57
1387	931	22306	2.73	65.29
1399	890	23196	2.61	67.90
1410	894	24090	2.62	70.51
1422	930	25020	2.72	73.23
1434	806	25826	2.36	75.59
1445	876	26702	2.56	78.16
1447	822	27524	2.41	80.56
1470	815	28339	2.39	82.95
1482	771	29110	2.26	85.21
1496	725	29835	2.12	87.33
1509	674	30509	1.97	89.30
1515	606	31115	1.77	91.08
1540	566	31681	1.66	92.73
1556	515	32196	1.51	94.24
1574	417	32613	1.22	95.46
1593	387	33000	1.13	96.59
1613	354	33354	1.04	97.63
1636	251	33605	0.73	98.36
1661	208	33813	0.61	98.97
1689	157	33970	0.46	99.43
1721	96	34066	0.28	99.71
1759	52	34118	0.15	99.87
1807	27	34145	0.08	99.94
1871	15	34160	0.04	99.99
1978	4	34164	0.01	100.00

Figure B.7.20. Spring 2025 STAAR Spanish Grade 4 RLA
Frequency Distribution of Scale Scores



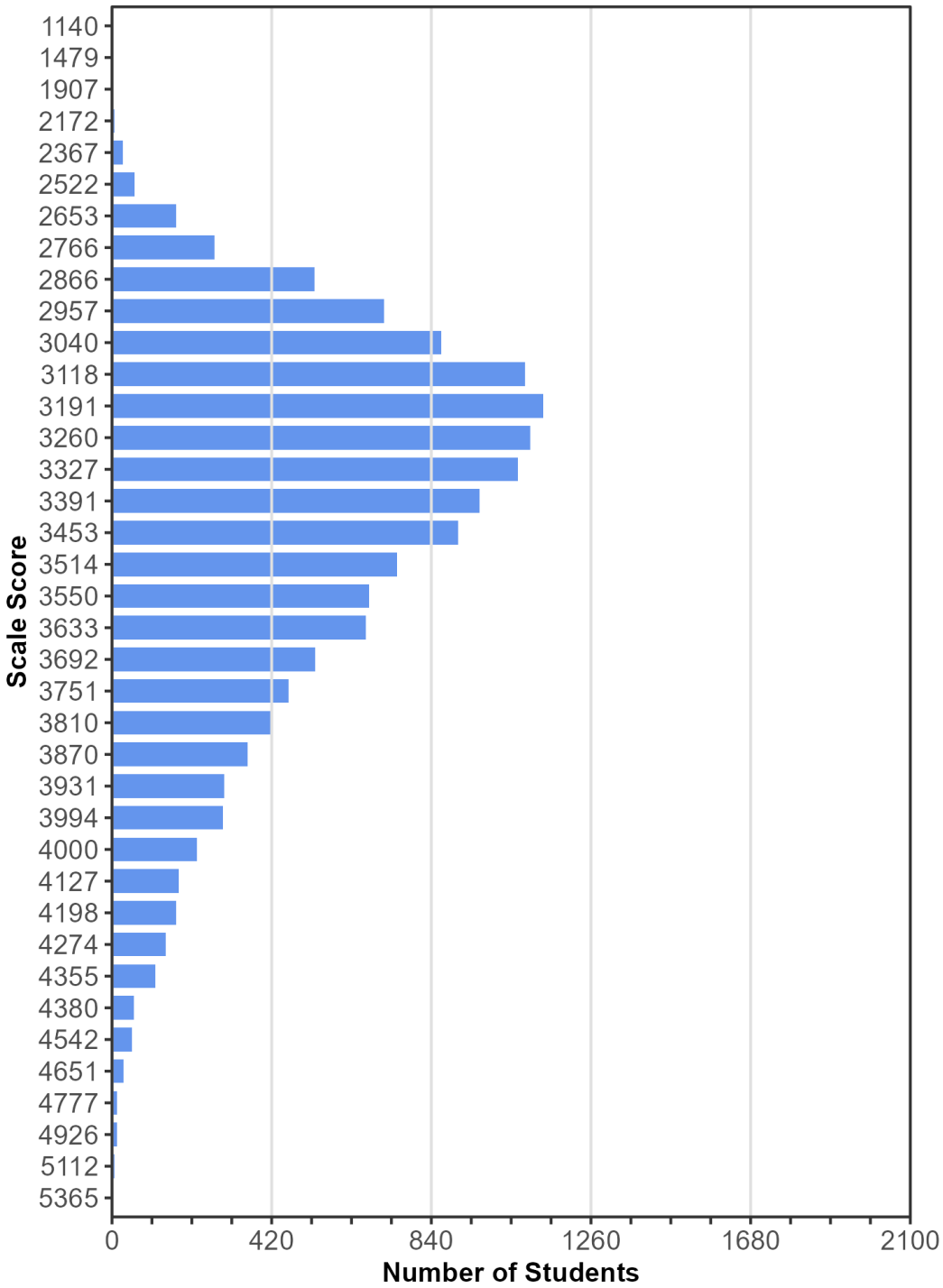
Freq	Cum Freq	Pct	Cum Pct
4	4	0.02	0.02
1	5	0.00	0.02
4	9	0.02	0.03
24	33	0.09	0.13
61	94	0.23	0.36
137	231	0.52	0.88
223	454	0.85	1.72
369	823	1.40	3.12
533	1356	2.02	5.14
711	2067	2.70	7.84
765	2832	2.90	10.74
758	3590	2.88	13.62
790	4380	3.00	16.62
810	5190	3.07	19.69
817	6007	3.10	22.79
757	6764	2.87	25.66
787	7551	2.99	28.65
718	8269	2.72	31.37
748	9017	2.84	34.21
780	9797	2.96	37.17
751	10548	2.85	40.02
708	11256	2.69	42.71
763	12019	2.89	45.60
784	12803	2.97	48.58
762	13565	2.89	51.47
732	14297	2.78	54.24
784	15081	2.97	57.22
808	15889	3.07	60.28
753	16642	2.86	63.14
763	17405	2.89	66.04
830	18235	3.15	69.18
863	19098	3.27	72.46
814	19912	3.09	75.55
686	20598	2.60	78.15
781	21379	2.96	81.11
671	22050	2.55	83.66
664	22714	2.52	86.18
625	23339	2.37	88.55
552	23891	2.09	90.64
506	24397	1.92	92.56
454	24851	1.72	94.29
415	25266	1.57	95.86
347	25613	1.32	97.18
233	25846	0.88	98.06
204	26050	0.77	98.84
123	26173	0.47	99.30
88	26261	0.33	99.64
52	26313	0.20	99.83
29	26342	0.11	99.94
8	26350	0.03	99.97
5	26355	0.02	99.99
2	26357	0.01	100.00

Figure B.7.21. Spring 2025 STAAR Spanish Grade 5 RLA
Frequency Distribution of Scale Scores



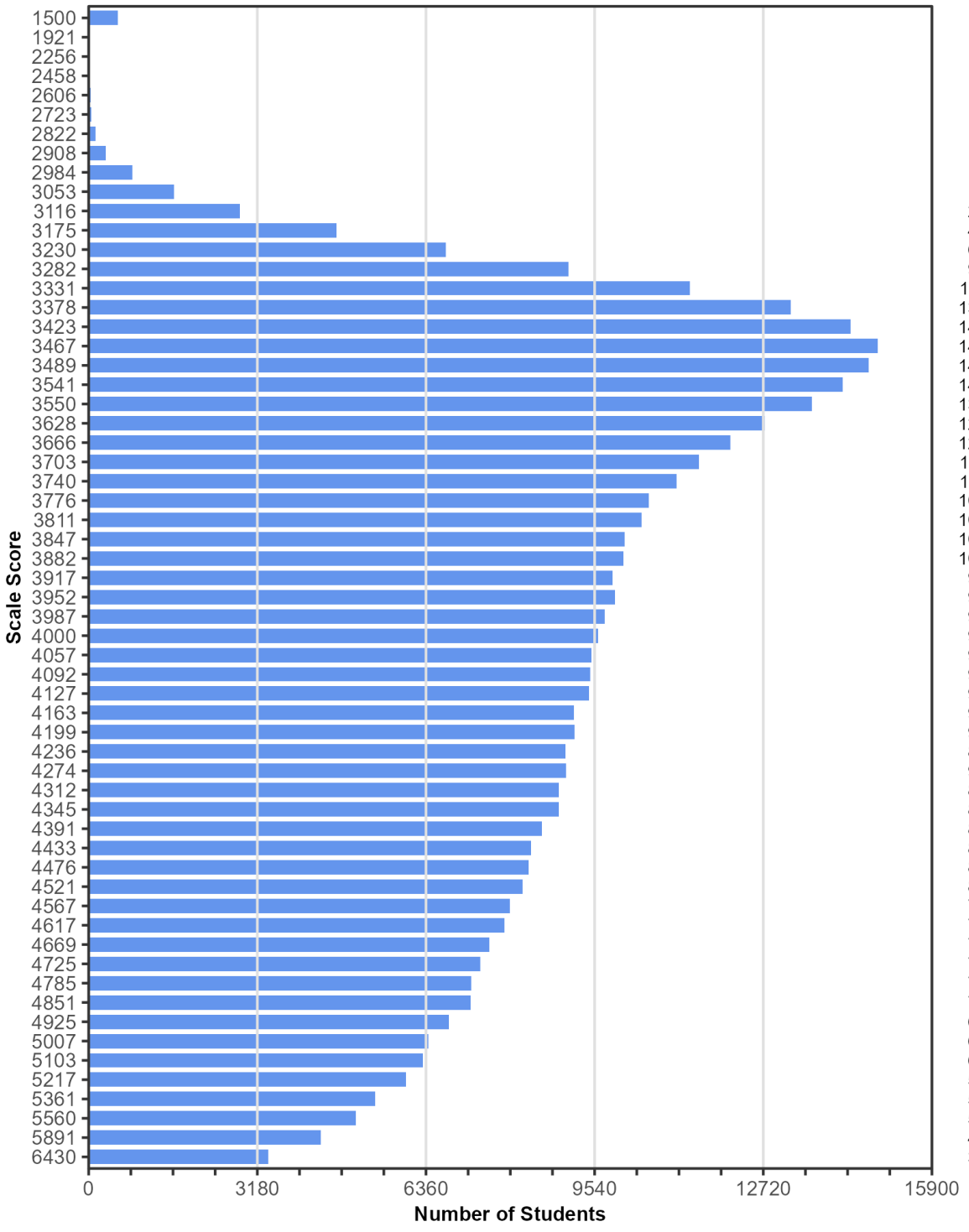
Freq	Cum Freq	Pct	Cum Pct
3	3	0.02	0.02
2	5	0.01	0.03
3	8	0.02	0.04
7	15	0.04	0.08
14	29	0.07	0.15
50	79	0.26	0.41
93	172	0.48	0.89
163	335	0.84	1.73
282	617	1.45	3.18
343	960	1.77	4.95
485	1445	2.50	7.45
441	1886	2.27	9.73
499	2385	2.57	12.30
496	2881	2.56	14.86
529	3410	2.73	17.58
529	3939	2.73	20.31
498	4437	2.57	22.88
523	4960	2.70	25.58
486	5446	2.51	28.08
453	5899	2.34	30.42
502	6401	2.59	33.01
503	6904	2.59	35.60
526	7430	2.71	38.31
531	7961	2.74	41.05
500	8461	2.58	43.63
529	8990	2.73	46.36
556	9546	2.87	49.23
541	10087	2.79	52.02
547	10634	2.82	54.84
553	11187	2.85	57.69
590	11777	3.04	60.73
579	12356	2.99	63.72
642	12998	3.31	67.03
661	13659	3.41	70.44
594	14253	3.06	73.50
591	14844	3.05	76.55
600	15444	3.09	79.64
568	16012	2.93	82.57
519	16531	2.68	85.25
492	17023	2.54	87.78
501	17524	2.58	90.37
422	17946	2.18	92.54
375	18321	1.93	94.48
314	18635	1.62	96.10
267	18902	1.38	97.47
205	19107	1.06	98.53
141	19248	0.73	99.26
83	19331	0.43	99.69
35	19366	0.18	99.87
17	19383	0.09	99.95
8	19391	0.04	99.99
1	19392	0.01	100.00

Figure B.7.22. Spring 2025 STAAR Spanish Grade 5 Science
Frequency Distribution of Scale Scores



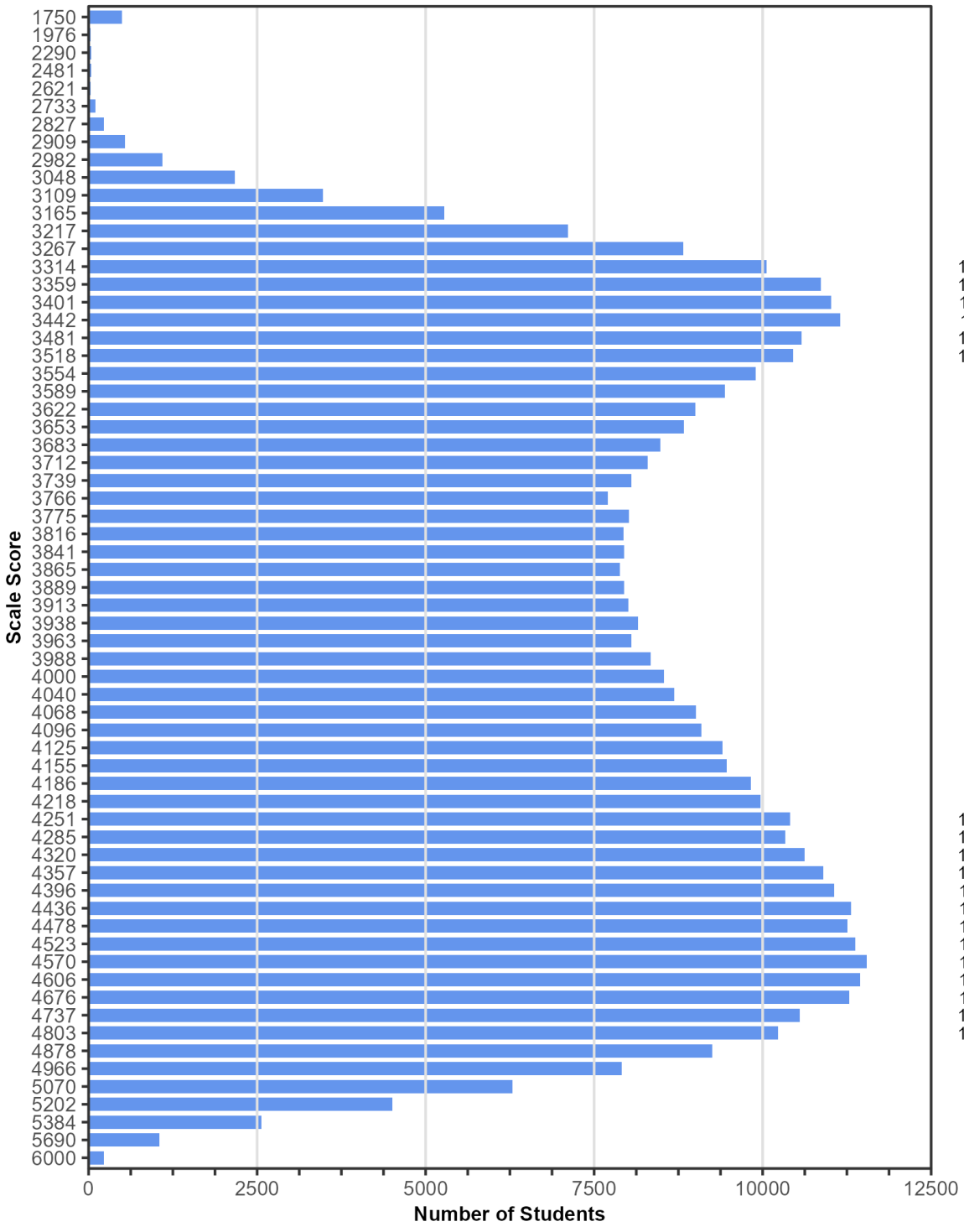
	Cum		Cum
Freq	Freq	Pct	Pct
2	2	0.01	0.01
1	3	0.01	0.02
3	6	0.02	0.04
7	13	0.05	0.09
28	41	0.19	0.29
60	101	0.42	0.70
169	270	1.18	1.88
270	540	1.88	3.76
533	1073	3.71	7.46
716	1789	4.98	12.44
866	2655	6.02	18.46
1087	3742	7.56	26.02
1135	4877	7.89	33.92
1100	5977	7.65	41.56
1068	7045	7.43	48.99
967	8012	6.72	55.72
911	8923	6.34	62.05
750	9673	5.22	67.27
676	10349	4.70	71.97
667	11016	4.64	76.61
534	11550	3.71	80.32
464	12014	3.23	83.55
423	12437	2.94	86.49
357	12794	2.48	88.97
296	13090	2.06	91.03
292	13382	2.03	93.06
223	13605	1.55	94.61
175	13780	1.22	95.83
168	13948	1.17	97.00
141	14089	0.98	97.98
114	14203	0.79	98.77
58	14261	0.40	99.17
53	14314	0.37	99.54
30	14344	0.21	99.75
14	14358	0.10	99.85
14	14372	0.10	99.94
6	14378	0.04	99.99
2	14380	0.01	100.00

Figure B.7.23. Spring 2025 STAAR Algebra I
Frequency Distribution of Scale Scores



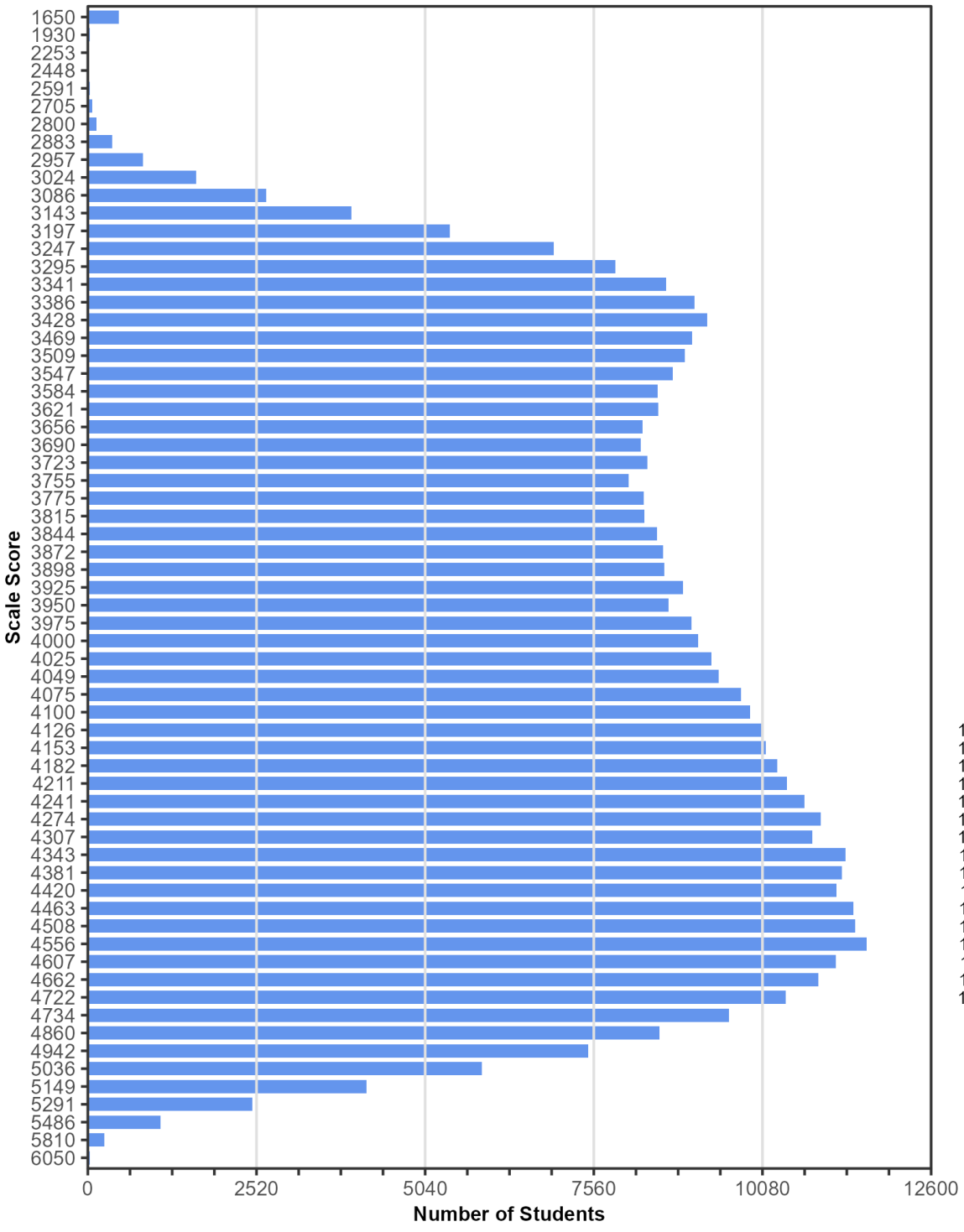
	Freq	Cum Freq	Pct	Cum Pct
1500	546	546	0.12	0.12
1921	20	566	0.00	0.12
2256	19	585	0.00	0.13
2458	15	600	0.00	0.13
2606	34	634	0.01	0.14
2723	53	687	0.01	0.15
2822	132	819	0.03	0.18
2908	317	1136	0.07	0.25
2984	822	1958	0.18	0.43
3053	1612	3570	0.35	0.78
3116	2847	6417	0.62	1.41
3175	4670	11087	1.02	2.43
3230	6739	17826	1.48	3.91
3282	9050	26876	1.99	5.90
3331	11334	38210	2.49	8.39
3378	13241	51451	2.91	11.29
3423	14368	65819	3.15	14.45
3467	14876	80695	3.26	17.71
3489	14709	95404	3.23	20.94
3541	14215	109619	3.12	24.06
3550	13638	123257	2.99	27.05
3628	12745	136002	2.80	29.85
3666	12102	148104	2.66	32.51
3703	11503	159607	2.52	35.03
3740	11091	170698	2.43	37.46
3776	10566	181264	2.32	39.78
3811	10421	191685	2.29	42.07
3847	10103	201788	2.22	44.29
3882	10085	211873	2.21	46.50
3917	9878	221751	2.17	48.67
3952	9923	231674	2.18	50.85
3987	9729	241403	2.14	52.98
4000	9600	251003	2.11	55.09
4057	9485	260488	2.08	57.17
4092	9462	269950	2.08	59.25
4127	9432	279382	2.07	61.32
4163	9147	288529	2.01	63.33
4199	9163	297692	2.01	65.34
4236	8992	306684	1.97	67.31
4274	9000	315684	1.98	69.29
4312	8862	324546	1.95	71.23
4345	8866	333412	1.95	73.18
4391	8540	341952	1.87	75.05
4433	8341	350293	1.83	76.88
4476	8293	358586	1.82	78.70
4521	8182	366768	1.80	80.50
4567	7944	374712	1.74	82.24
4617	7842	382554	1.72	83.96
4669	7558	390112	1.66	85.62
4725	7389	397501	1.62	87.24
4785	7215	404716	1.58	88.83
4851	7202	411918	1.58	90.41
4925	6796	418714	1.49	91.90
5007	6404	425118	1.41	93.30
5103	6306	431424	1.38	94.69
5217	5988	437412	1.31	96.00
5361	5402	442814	1.19	97.19
5560	5038	447852	1.11	98.29
5891	4381	452233	0.96	99.26
6430	3389	455622	0.74	100.00

Figure B.7.24. Spring 2025 STAAR English I
Frequency Distribution of Scale Scores



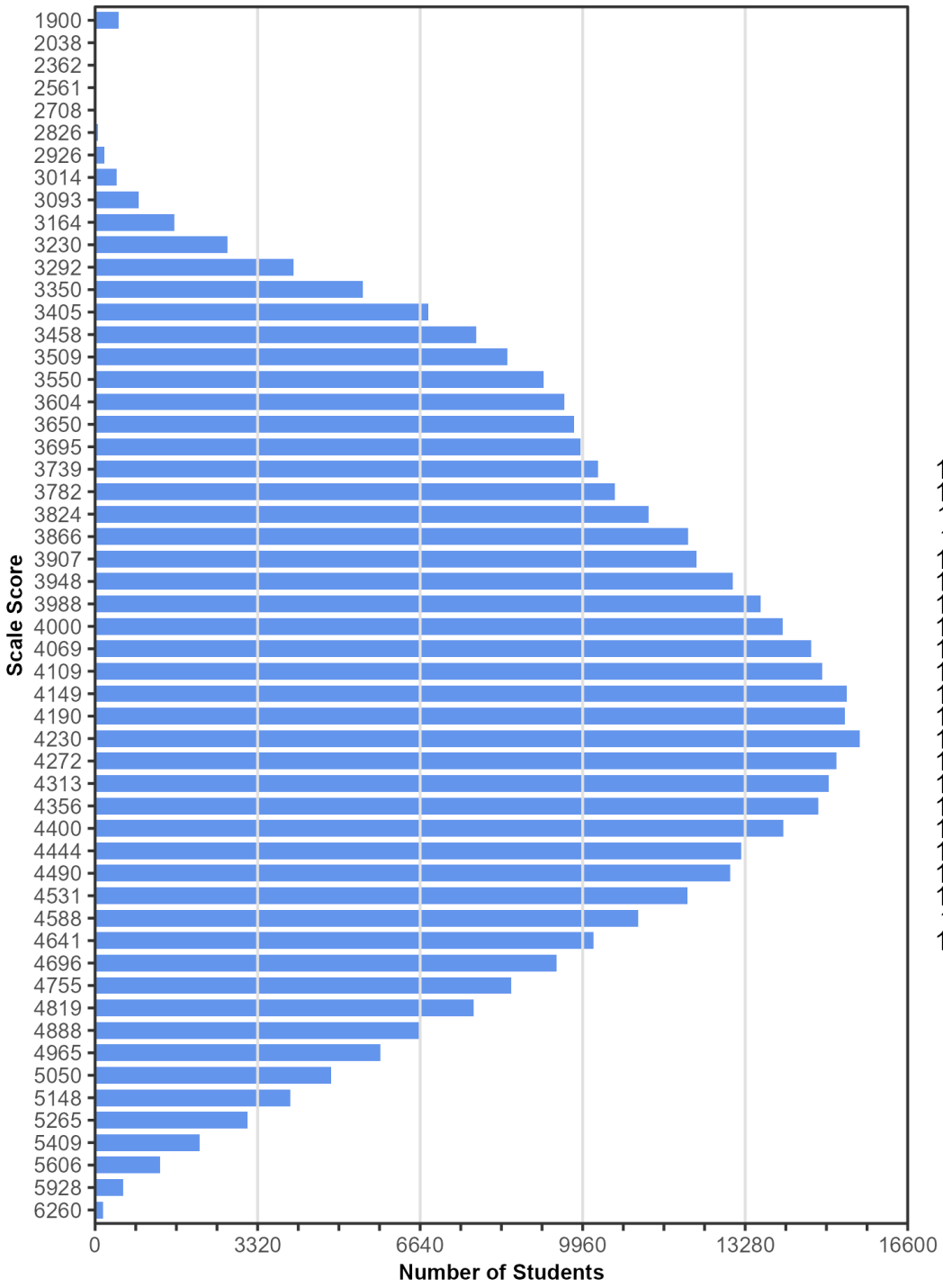
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1750	494	494	0.10	0.10
1976	26	520	0.01	0.11
2290	37	557	0.01	0.12
2481	37	594	0.01	0.12
2621	32	626	0.01	0.13
2733	99	725	0.02	0.15
2827	227	952	0.05	0.20
2909	538	1490	0.11	0.31
2982	1097	2587	0.23	0.53
3048	2168	4755	0.45	0.98
3109	3481	8236	0.72	1.70
3165	5278	13514	1.09	2.79
3217	7113	20627	1.47	4.26
3267	8819	29446	1.82	6.09
3314	10054	39500	2.08	8.17
3359	10866	50366	2.25	10.41
3401	11012	61378	2.28	12.69
3442	11154	72532	2.31	15.00
3481	10582	83114	2.19	17.18
3518	10453	93567	2.16	19.35
3554	9895	103462	2.05	21.39
3589	9438	112900	1.95	23.34
3622	8998	121898	1.86	25.20
3653	8830	130728	1.83	27.03
3683	8478	139206	1.75	28.78
3712	8294	147500	1.71	30.50
3739	8048	155548	1.66	32.16
3766	7707	163255	1.59	33.76
3775	8020	171275	1.66	35.41
3816	7940	179215	1.64	37.06
3841	7946	187161	1.64	38.70
3865	7882	195043	1.63	40.33
3889	7945	202988	1.64	41.97
3913	8008	210996	1.66	43.63
3938	8154	219150	1.69	45.31
3963	8048	227198	1.66	46.98
3988	8337	235535	1.72	48.70
4000	8535	244070	1.76	50.46
4040	8687	252757	1.80	52.26
4068	9012	261769	1.86	54.12
4096	9095	270864	1.88	56.00
4125	9405	280269	1.94	57.95
4155	9463	289732	1.96	59.91
4186	9829	299561	2.03	61.94
4218	9969	309530	2.06	64.00
4251	10403	319933	2.15	66.15
4285	10340	330273	2.14	68.29
4320	10618	340891	2.20	70.48
4357	10899	351790	2.25	72.74
4396	11059	362849	2.29	75.02
4436	11316	374165	2.34	77.36
4478	11260	385425	2.33	79.69
4523	11374	396799	2.35	82.04
4570	11546	408345	2.39	84.43
4606	11444	419789	2.37	86.80
4676	11287	431076	2.33	89.13
4737	10548	441624	2.18	91.31
4803	10225	451849	2.11	93.43
4878	9256	461105	1.91	95.34
4966	7907	469012	1.63	96.97
5070	6285	475297	1.30	98.27
5202	4505	479802	0.93	99.21
5384	2566	482368	0.53	99.74
5690	1054	483422	0.22	99.95
6000	222	483644	0.05	100.00

Figure B.7.25. Spring 2025 STAAR English II
Frequency Distribution of Scale Scores



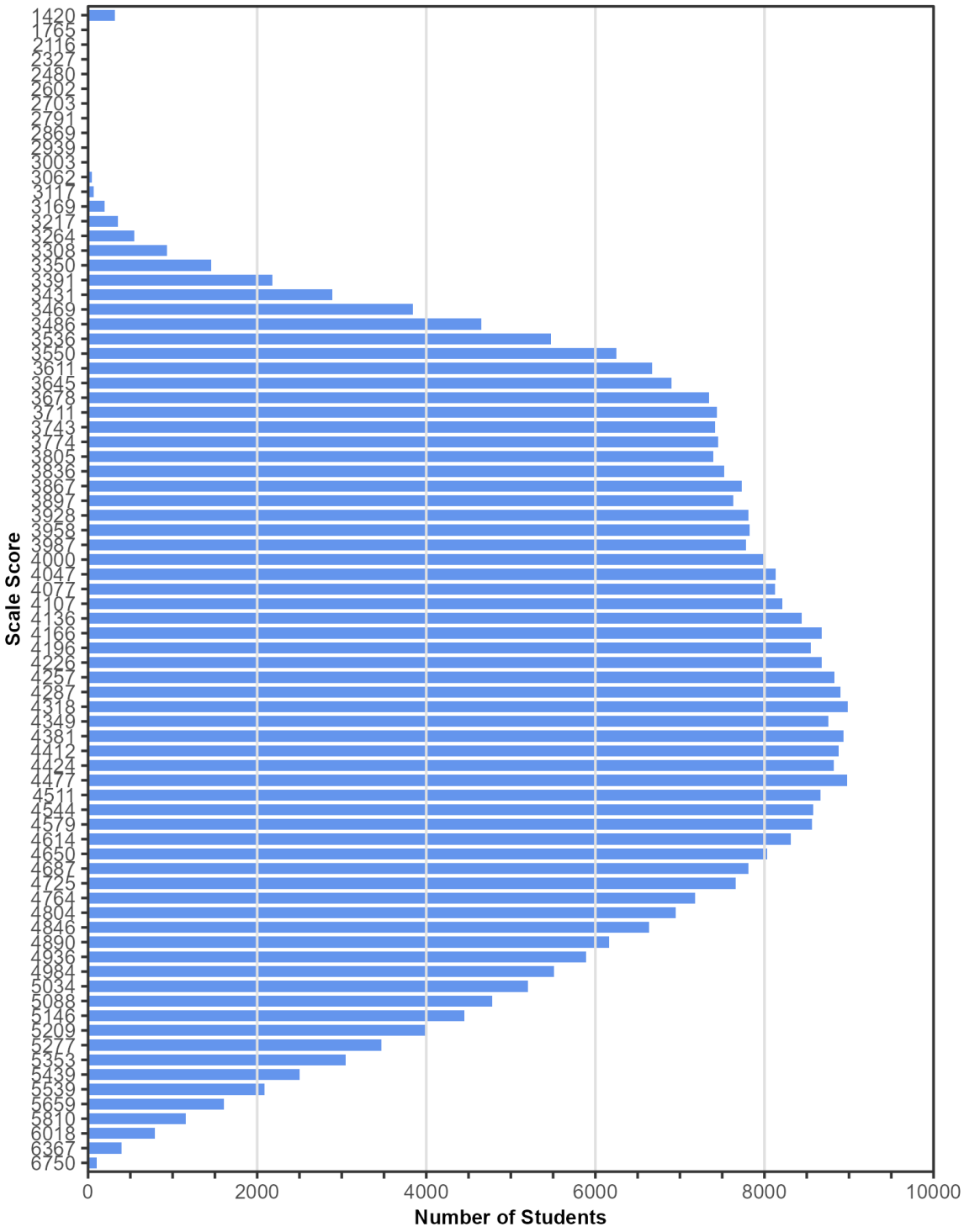
	Freq	Cum Freq	Pct	Cum Pct
1650	466	466	0.10	0.10
1930	29	495	0.01	0.11
2253	19	514	0.00	0.11
2448	25	539	0.01	0.12
2591	26	565	0.01	0.12
2705	66	631	0.01	0.14
2800	129	760	0.03	0.16
2883	362	1122	0.08	0.24
2957	822	1944	0.18	0.42
3024	1614	3558	0.35	0.77
3086	2666	6224	0.57	1.34
3143	3938	10162	0.85	2.19
3197	5407	15569	1.16	3.35
3247	6965	22534	1.50	4.85
3295	7885	30419	1.70	6.55
3341	8638	39057	1.86	8.41
3386	9068	48125	1.95	10.37
3428	9258	57383	1.99	12.36
3469	9024	66407	1.94	14.30
3509	8921	75328	1.92	16.23
3547	8742	84070	1.88	18.11
3584	8514	92584	1.83	19.94
3621	8524	101108	1.84	21.78
3656	8288	109396	1.79	23.56
3690	8262	117658	1.78	25.34
3723	8363	126021	1.80	27.15
3755	8079	134100	1.74	28.89
3775	8304	142404	1.79	30.67
3815	8312	150716	1.79	32.46
3844	8501	159217	1.83	34.30
3872	8596	167813	1.85	36.15
3898	8613	176426	1.86	38.00
3925	8891	185317	1.92	39.92
3950	8676	193993	1.87	41.79
3975	9023	203016	1.94	43.73
4000	9116	212132	1.96	45.69
4025	9320	221452	2.01	47.70
4049	9424	230876	2.03	49.73
4075	9757	240633	2.10	51.83
4100	9893	250526	2.13	53.96
4126	10074	260600	2.17	56.13
4153	10131	270731	2.18	58.32
4182	10301	281032	2.22	60.54
4211	10444	291476	2.25	62.78
4241	10709	302185	2.31	65.09
4274	10953	313138	2.36	67.45
4307	10821	323959	2.33	69.78
4343	11321	335280	2.44	72.22
4381	11264	346544	2.43	74.65
4420	11188	357732	2.41	77.06
4463	11441	369173	2.46	79.52
4508	11462	380635	2.47	81.99
4556	11633	392268	2.51	84.50
4607	11175	403443	2.41	86.90
4662	10911	414354	2.35	89.25
4722	10425	424779	2.25	91.50
4734	9579	434358	2.06	93.56
4860	8538	442896	1.84	95.40
4942	7476	450372	1.61	97.01
5036	5889	456261	1.27	98.28
5149	4160	460421	0.90	99.18
5291	2459	462880	0.53	99.71
5486	1088	463968	0.23	99.94
5810	245	464213	0.05	99.99
6050	33	464246	0.01	100.00

Figure B.7.26. Spring 2025 STAAR Biology
Frequency Distribution of Scale Scores



Freq	Cum Freq	Pct	Cum Pct
482	482	0.11	0.11
22	504	0.01	0.12
22	526	0.01	0.12
22	548	0.01	0.13
23	571	0.01	0.13
56	627	0.01	0.15
192	819	0.05	0.19
446	1265	0.11	0.30
891	2156	0.21	0.51
1617	3773	0.38	0.89
2709	6482	0.64	1.53
4048	10530	0.96	2.49
5470	16000	1.29	3.78
6802	22802	1.61	5.39
7792	30594	1.84	7.23
8427	39021	1.99	9.22
9156	48177	2.16	11.38
9587	57764	2.26	13.65
9785	67549	2.31	15.96
9919	77468	2.34	18.30
10268	87736	2.43	20.73
10622	98358	2.51	23.24
11309	109667	2.67	25.91
12111	121778	2.86	28.77
12285	134063	2.90	31.67
13022	147085	3.08	34.75
13596	160681	3.21	37.96
14047	174728	3.32	41.28
14624	189352	3.45	44.73
14857	204209	3.51	48.24
15359	219568	3.63	51.87
15323	234891	3.62	55.49
15627	250518	3.69	59.18
15149	265667	3.58	62.76
14982	280649	3.54	66.30
14774	295423	3.49	69.79
14062	309485	3.32	73.11
13196	322681	3.12	76.23
12980	335661	3.07	79.29
12104	347765	2.86	82.15
11101	358866	2.62	84.78
10180	369046	2.40	87.18
9428	378474	2.23	89.41
8504	386978	2.01	91.42
7731	394709	1.83	93.24
6653	401362	1.57	94.81
5830	407192	1.38	96.19
4821	412013	1.14	97.33
3985	415998	0.94	98.27
3112	419110	0.74	99.01
2138	421248	0.51	99.51
1332	422580	0.31	99.83
576	423156	0.14	99.96
156	423312	0.04	100.00

Figure B.7.27. Spring 2025 STAAR U.S. History Frequency Distribution of Scale Scores



Freq	Cum Freq	Pct	Cum Pct
320	320	0.08	0.08
9	329	0.00	0.08
6	335	0.00	0.09
11	346	0.00	0.09
5	351	0.00	0.09
5	356	0.00	0.09
6	362	0.00	0.09
7	369	0.00	0.09
4	373	0.00	0.09
9	382	0.00	0.10
18	400	0.00	0.10
47	447	0.01	0.11
69	516	0.02	0.13
196	712	0.05	0.18
356	1068	0.09	0.27
544	1612	0.14	0.41
936	2548	0.24	0.65
1453	4001	0.37	1.02
2181	6182	0.55	1.57
2888	9070	0.73	2.30
3841	12911	0.98	3.28
4648	17559	1.18	4.46
5476	23035	1.39	5.85
6246	29281	1.59	7.44
6670	35951	1.69	9.13
6899	42850	1.75	10.89
7342	50192	1.87	12.75
7437	57629	1.89	14.64
7416	65045	1.88	16.52
7451	72496	1.89	18.42
7393	79889	1.88	20.30
7523	87412	1.91	22.21
7734	95146	1.96	24.17
7634	102780	1.94	26.11
7813	110593	1.98	28.10
7823	118416	1.99	30.08
7784	126200	1.98	32.06
8005	134205	2.03	34.09
8129	142334	2.07	36.16
8125	150459	2.06	38.22
8214	158673	2.09	40.31
8437	167110	2.14	42.45
8677	175787	2.20	44.66
8550	184337	2.17	46.83
8679	193016	2.20	49.04
8825	201841	2.24	51.28
8902	210743	2.26	53.54
8985	219728	2.28	55.82
8755	228483	2.22	58.05
8935	237418	2.27	60.32
8879	246297	2.26	62.57
8823	255120	2.24	64.81
8976	264096	2.28	67.09
8666	272762	2.20	69.30
8579	281341	2.18	71.47
8559	289900	2.17	73.65
8311	298211	2.11	75.76
8032	306243	2.04	77.80
7807	314050	1.98	79.78
7662	321712	1.95	81.73
7178	328890	1.82	83.55
6948	335838	1.77	85.32
6637	342475	1.69	87.01
6161	348636	1.57	88.57
5888	354524	1.50	90.07
5508	360032	1.40	91.47
5204	365236	1.32	92.79
4777	370013	1.21	94.00
4451	374464	1.13	95.13
4011	378475	1.02	96.15
3466	381941	0.88	97.03
3048	384989	0.77	97.81
2500	387489	0.64	98.44
2085	389574	0.53	98.97
1610	391184	0.41	99.38
1155	392339	0.29	99.67
791	393130	0.20	99.87
393	393523	0.10	99.97
101	393624	0.03	100.00

Spring 2025 STAAR
English Learner Performance Measure

Table B.8.1. Spring 2025 STAAR Grades 3–8 Mathematics English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 3	6,773	1506 22%	2283 34%	1700 25%	410 6%	874 13%
Grade 4	3,793	660 17%	1653 44%	708 19%	178 5%	594 16%
Grade 5	3,500	947 27%	1174 34%	719 21%	158 5%	502 14%
Grade 6	6,177	1387 22%	3423 55%	909 15%	130 2%	328 5%
Grade 7	5,432	2212 41%	2234 41%	660 12%	120 2%	206 4%
Grade 8	6,048	1660 27%	2622 43%	1125 19%	210 3%	431 7%

Table B.8.2. Spring 2025 STAAR Grades 3–8 RLA English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 3	5,543	1262 23%	1654 30%	1487 27%	377 7%	763 14%
Grade 4	2,966	967 33%	796 27%	782 26%	96 3%	325 11%
Grade 5	2,888	1106 38%	725 25%	591 20%	98 3%	368 13%
Grade 6	6,219	4011 64%	1348 22%	601 10%	70 1%	189 3%
Grade 7	6,271	4161 66%	1253 20%	594 9%	63 1%	200 3%
Grade 8	6,132	4101 67%	1195 19%	586 10%	74 1%	176 3%

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 2025 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	3,305	732 22%	1854 56%	504 15%	63 2%	152 5%
Grade 8	5,912	2499 42%	2297 39%	816 14%	132 2%	168 3%

Table B.8.4. Spring 2025 STAAR Grade 8 Social Studies English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Grade 8	6,093	3366 55%	2198 36%	345 6%	45 1%	139 2%

Table B.8.5. Spring 2025 STAAR EOC Assessments English Learner Performance Measure

Assessment	Total	Did Not Meet	Approaches	Meets	Masters	Masters based on STAAR
Algebra I	8,487	1086 13%	4141 49%	1786 21%	360 4%	1114 13%
English I	9,679	5417 56%	3138 32%	901 9%	154 2%	69 1%
English II	10,460	4948 47%	3712 35%	1583 15%	182 2%	35 0%
Biology	8,050	1936 24%	3183 40%	2197 27%	343 4%	391 5%
U.S. History	4,775	163 3%	2465 52%	1215 25%	272 6%	660 14%

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 2024–2025

Appendix C

STAAR Alternate 2 Statistical Tables and Figures

List of Tables

Table C.1.1. 2025 STAAR Alternate 2 Grades 3–8 Classification Consistency and Accuracy

Table C.1.2. 2025 STAAR Alternate 2 EOC Assessments Classification Consistency and Accuracy

Table C.2.1. 2025 STAAR Alternate 2 Grades 3–8 Scale Score Correlations

Table C.2.2. 2025 STAAR Alternate 2 EOC Assessments Scale Score Correlations

Table C.3.1. 2025 STAAR Alternate 2 Grades 3–5 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table C.3.2. 2025 STAAR Alternate 2 Grades 6–8 Mathematics Conditional Standard Error of Measurement for Scale Scores

Table C.3.3. 2025 STAAR Alternate 2 Grades 3–5 RLA Conditional Standard Error of Measurement for Scale Scores

Table C.3.4. 2025 STAAR Alternate 2 Grades 6–8 RLA Conditional Standard Error of Measurement for Scale Scores

Table C.3.5. 2025 STAAR Alternate 2 Grade 5 and Grade 8 Science Conditional Standard Error of Measurement for Scale Scores

Table C.3.6. 2025 STAAR Alternate 2 Grade 8 Social Studies Conditional Standard Error of Measurement for Scale Scores

Table C.3.7. 2025 STAAR Alternate 2 Algebra I Conditional Standard Error of Measurement for Scale Scores

Table C.3.8. 2025 STAAR Alternate 2 English I and English II Conditional Standard Error of Measurement for Scale Scores

Table C.3.9. 2025 STAAR Alternate 2 Biology Conditional Standard Error of Measurement for Scale Scores

Table C.3.10. 2025 STAAR Alternate 2 U.S. History Conditional Standard Error of Measurement for Scale Scores

Table C.4.1. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.2. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.3. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.4. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.5. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.6. 2025 STAAR Alternate 2 Grade 3 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.7. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.8. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.9. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.10. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.11. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.12. 2025 STAAR Alternate 2 Grade 4 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.13. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.14. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.15. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.16. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.17. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.18. 2025 STAAR Alternate 2 Grade 5 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.19. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.20. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.21. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.22. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.23. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.24. 2025 STAAR Alternate 2 Grade 6 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.25. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.26. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.27. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.28. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.29. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.30. 2025 STAAR Alternate 2 Grade 7 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.31. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.32. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.33. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.34. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.35. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.36. 2025 STAAR Alternate 2 Grade 8 Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.37. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.38. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.39. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.40. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.41. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.42. 2025 STAAR Alternate 2 Algebra I Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.43. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.44. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.45. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.46. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.47. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.48. 2025 STAAR Alternate 2 English I and English II Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.49. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.50. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.51. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.52. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.53. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.54. 2025 STAAR Alternate 2 Biology Mean P-Values, Raw Score Summary, and Reliability White

Table C.4.55. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Total Group

Table C.4.56. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Female

Table C.4.57. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Male

Table C.4.58. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Black or African American

Table C.4.59. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability Hispanic or Latino

Table C.4.60. 2025 STAAR Alternate 2 U.S. History Mean P-Values, Raw Score Summary, and Reliability White

Table C.5.1. 2025 STAAR Alternate 2 All Mathematics Progress Measure

Table C.5.2. 2025 STAAR Alternate 2 All RLA Progress Measure

Table C.5.3. 2025 STAAR Alternate 2 Grade 4 Mathematics Progress Measure

Table C.5.4. 2025 STAAR Alternate 2 Grade 5 Mathematics Progress Measure

Table C.5.5. 2025 STAAR Alternate 2 Grade 6 Mathematics Progress Measure

Table C.5.6. 2025 STAAR Alternate 2 Grade 7 Mathematics Progress Measure

Table C.5.7. 2025 STAAR Alternate 2 Grade 8 Mathematics Progress Measure

Table C.5.8. 2025 STAAR Alternate 2 Algebra I Progress Measure

Table C.5.9. 2025 STAAR Alternate 2 Grade 4 RLA Progress Measure

Table C.5.10. 2025 STAAR Alternate 2 Grade 5 RLA Progress Measure

Table C.5.11. 2025 STAAR Alternate 2 Grade 6 RLA Progress Measure

Table C.5.12. 2025 STAAR Alternate 2 Grade 7 RLA Progress Measure

Table C.5.13. 2025 STAAR Alternate 2 Grade 8 RLA Progress Measure

Table C.5.14. 2025 STAAR Alternate 2 English I Progress Measure

Table C.5.15. 2025 STAAR Alternate 2 English II Progress Measure

Table C.6.1. 2025 STAAR Alternate 2 Grades 3–8 Scale Score Descriptive Statistics

Table C.6.2. 2025 STAAR Alternate 2 EOC Assessments Scale Score Descriptive Statistics

List of Figures

Figure C.6.1. 2025 STAAR Alternate 2 Grade 3 Mathematics Frequency Distribution of Scale Scores

Figure C.6.2. 2025 STAAR Alternate 2 Grade 4 Mathematics Frequency Distribution of Scale Scores

Figure C.6.3. 2025 STAAR Alternate 2 Grade 5 Mathematics Frequency Distribution of Scale Scores

Figure C.6.4. 2025 STAAR Alternate 2 Grade 6 Mathematics Frequency Distribution of Scale Scores

Figure C.6.5. 2025 STAAR Alternate 2 Grade 7 Mathematics Frequency Distribution of Scale Scores

Figure C.6.6. 2025 STAAR Alternate 2 Grade 8 Mathematics Frequency Distribution of Scale Scores

Figure C.6.7. 2025 STAAR Alternate 2 Grade 3 RLA Frequency Distribution of Scale Scores

Figure C.6.8. 2025 STAAR Alternate 2 Grade 4 RLA Frequency Distribution of Scale Scores

Figure C.6.9. 2025 STAAR Alternate 2 Grade 5 RLA Frequency Distribution of Scale Scores

Figure C.6.10. 2025 STAAR Alternate 2 Grade 6 RLA Frequency Distribution of Scale Scores

Figure C.6.11. 2025 STAAR Alternate 2 Grade 7 RLA Frequency Distribution of Scale Scores

Figure C.6.12. 2025 STAAR Alternate 2 Grade 8 RLA Frequency Distribution of Scale Scores

Figure C.6.13. 2025 STAAR Alternate 2 Grade 5 Science Frequency Distribution of Scale Scores

Figure C.6.14. 2025 STAAR Alternate 2 Grade 8 Science Frequency Distribution of Scale Scores

Figure C.6.15. 2025 STAAR Alternate 2 Grade 8 Social Studies Frequency Distribution of Scale Scores

Figure C.6.16. 2025 STAAR Alternate 2 Algebra I Frequency Distribution of Scale Scores

Figure C.6.17. 2025 STAAR Alternate 2 English I Frequency Distribution of Scale Scores

Figure C.6.18. 2025 STAAR Alternate 2 English II Frequency Distribution of Scale Scores

Figure C.6.19. 2025 STAAR Alternate 2 Biology Frequency Distribution of Scale Scores

Figure C.6.20. 2025 STAAR Alternate 2 U.S. History Frequency Distribution of Scale Scores

2025 STAAR Alternate 2 Classification Consistency and Accuracy

Table C.1.1. 2025 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	79.4	80.2			85.3	86.0		
4	83.2	78.4			88.0	84.7		
5	79.7	74.5	78.7		85.5	81.8	84.7	
6	76.0	74.2			82.9	81.7		
7	78.7	75.5			84.9	82.6		
8	75.1	72.9	80.6	74.9	82.2	80.7	86.1	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.

Table C.1.2. 2025 STAAR Alternate 2 EOC Assessments
Classification Consistency and Accuracy

Test	Decision Consistency	Decision Accuracy
Algebra I	73.6	81.3
English I	72.8	80.6
English II	74.3	81.7
Biology	78.9	84.8
U.S. History	76.3	83.2

Notes:
1. Consistency indicates the proportion of students that would be classified into the same performance levels if they were administered a parallel test form. The proportions are converted to a 0–100% scale.
2. Accuracy indicates the proportion of students that are accurately classified. The proportions are converted to a 0–100% scale.

2025 STAAR Alternate 2
Scale Score Correlations

Table C.2.1. 2025 STAAR Alternate 2 Grades 3–8 Scale Score Correlations

Grade	STAAR Alternate 2	STAAR Alternate 2	N	Correlation
3	Mathematics	RLA	7,507	0.79
4	Mathematics	RLA	6,660	0.81
5	Mathematics	RLA	6,318	0.75
	Mathematics	Science	6,313	0.77
	RLA	Science	6,313	0.79
6	Mathematics	RLA	5,870	0.75
7	Mathematics	RLA	5,381	0.72
8	Mathematics	RLA	4,894	0.74
	Mathematics	Science	4,891	0.74
	Mathematics	Social Studies	4,891	0.73
	RLA	Science	4,897	0.76
	RLA	Social Studies	4,898	0.75
	Science	Social Studies	4,896	0.78

Table C.2.2. 2025 STAAR Alternate 2 EOC Assessments
Scale Score Correlations

STAAR Alternate 2	STAAR Alternate 2	N	Correlation
Algebra I	English I	4,836	0.68
Algebra I	English II	101	0.75
Algebra I	Biology	3,685	0.69
Algebra I	US History	71	0.63
English I	English II	63	0.79
English I	Biology	3,677	0.69
English I	US History	70	0.73
English II	Biology	1,236	0.69
English II	US History	455	0.79
Biology	US History	218	0.79

2025 STAAR Alternate 2 Conditional Standard Error of Measurement for Scale Scores

Table C.3.1. 2025 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. 2025 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. 2025 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. 2025 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. 2025 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	108		132	
1	153	40	173	36
2	177	28	196	25
3	192	23	210	21
4	203	21	220	19
5	212	19	228	17
6	220	18	235	16
7	228	17	241	15
8	234	17	247	15
9	241	16	253	14
10	247	16	258	14
11	253	16	262	13
12	258	16	267	13
13	264	15	271	13
14	269	15	276	13
15	274	15	280	13
16	279	15	284	12
17	284	14	288	12
18	288	14	292	12
19	293	14	296	12
20	298	14	**300	12
21	**300	14	303	12
22	306	14	307	12
23	311	14	311	12
24	315	14	315	12
25	320	14	319	13
26	324	14	323	13
27	329	14	328	13
28	334	15	332	13
29	338	15	337	13
30	344	15	341	14
31	349	16	346	14
32	355	16	352	15

(Continued)

	Grade 5		Grade 8	
Raw	**SS**	**CSEM**	**SS**	**CSEM**
33	361	17	357	15
34	368	18	364	16
35	376	20	371	18
36	***387	22	***382	19
37	398	25	391	22
38	415	30	406	27
39	443	43	432	38
40	494		477	

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. 2025 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. 2025 STAAR Alternate 2 Algebra I 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. 2025 STAAR Alternate 2 English I and English II 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. 2025 STAAR Alternate 2 Biology
Conditional Standard Error of Measurement for Scale Scores

		Biology
Raw	**SS**	**CSEM**
0	133	
1	173	35
2	195	24
3	207	20
4	217	18
5	224	16
6	231	15
7	237	15
8	242	14
9	247	14
10	252	14
11	257	13
12	262	13
13	266	13
14	271	13
15	275	13
16	279	13
17	284	13
18	288	13
19	292	13
20	296	13
21	**300	13
22	305	13
23	309	13
24	313	13
25	317	13
26	322	13
27	326	13
28	331	13
29	336	14
30	340	14
31	346	14
32	351	15

(Continued)

		Biology
Raw	**SS**	**CSEM**
33	357	16
34	364	17
35	372	18
36	***383	20
37	392	22
38	408	27
39	434	38
40	480	

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. 2025 STAAR Alternate 2 U.S. History 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.

2025 STAAR Alternate 2
Mean P-Values, Raw Score Summary, and Reliability

Table C.4.1. 2025 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	7,510	27.96	7.98	0.86	2.94	1.40
	Numerical Representations and Relationships	16	7,510	11.01	3.50	0.71	1.88	1.38
	Computations and Algebraic Relationships	8	7,510	5.35	2.17	0.59	1.39	1.34
	Geometry and Measurement	8	7,510	6.15	1.91	0.63	1.17	1.54
	Data Analysis and Personal Financial Literacy	8	7,510	5.45	1.89	0.51	1.31	1.36
RLA	OVERALL TEST	40	7,516	25.56	7.84	0.84	3.09	1.28
	Reading	24	7,516	16.00	4.85	0.78	2.28	1.33
	Writing	16	7,516	9.55	3.60	0.67	2.08	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.4.2. 2025 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	2,189	27.70	7.74	0.85	2.95	1.39
	Numerical Representations and Relationships	16	2,189	10.91	3.43	0.70	1.89	1.36
	Computations and Algebraic Relationships	8	2,189	5.29	2.14	0.57	1.40	1.32
	Geometry and Measurement	8	2,189	6.11	1.83	0.58	1.19	1.53
	Data Analysis and Personal Financial Literacy	8	2,189	5.39	1.87	0.51	1.31	1.35
RLA	OVERALL TEST	40	2,192	25.72	7.61	0.84	3.08	1.29
	Reading	24	2,192	16.10	4.69	0.76	2.27	1.34
	Writing	16	2,192	9.62	3.53	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.4.3. 2025 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	5,320	28.07	8.08	0.87	2.93	1.40
	Numerical Representations and Relationships	16	5,320	11.05	3.53	0.72	1.87	1.38
	Computations and Algebraic Relationships	8	5,320	5.38	2.18	0.59	1.39	1.35
	Geometry and Measurement	8	5,320	6.16	1.94	0.64	1.16	1.54
	Data Analysis and Personal Financial Literacy	8	5,320	5.48	1.89	0.52	1.31	1.37
RLA	OVERALL TEST	40	5,323	25.49	7.94	0.85	3.09	1.27
	Reading	24	5,323	15.96	4.92	0.78	2.29	1.33
	Writing	16	5,323	9.53	3.63	0.67	2.08	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.4.4. 2025 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	1,292	27.31	8.07	0.86	2.98	1.37
	Numerical Representations and Relationships	16	1,292	10.76	3.56	0.71	1.90	1.35
	Computations and Algebraic Relationships	8	1,292	5.17	2.18	0.59	1.40	1.29
	Geometry and Measurement	8	1,292	6.00	1.95	0.63	1.19	1.50
	Data Analysis and Personal Financial Literacy	8	1,292	5.38	1.91	0.51	1.33	1.35
RLA	OVERALL TEST	40	1,293	25.06	8.09	0.85	3.11	1.25
	Reading	24	1,293	15.73	5.02	0.79	2.30	1.31
	Writing	16	1,293	9.33	3.63	0.66	2.10	1.17

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. 2025 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	4,191	28.70	7.69	0.86	2.88	1.44
	Numerical Representations and Relationships	16	4,191	11.28	3.37	0.70	1.85	1.41
	Computations and Algebraic Relationships	8	4,191	5.52	2.11	0.58	1.37	1.38
	Geometry and Measurement	8	4,191	6.31	1.83	0.61	1.15	1.58
	Data Analysis and Personal Financial Literacy	8	4,191	5.60	1.84	0.51	1.29	1.40
RLA	OVERALL TEST	40	4,195	26.22	7.61	0.84	3.04	1.31
	Reading	24	4,195	16.40	4.69	0.77	2.25	1.37
	Writing	16	4,195	9.82	3.53	0.66	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.4.6. 2025 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	1,260	27.72	8.02	0.86	2.97	1.39
	Numerical Representations and Relationships	16	1,260	10.97	3.50	0.70	1.91	1.37
	Computations and Algebraic Relationships	8	1,260	5.29	2.21	0.59	1.41	1.32
	Geometry and Measurement	8	1,260	6.11	1.94	0.64	1.17	1.53
	Data Analysis and Personal Financial Literacy	8	1,260	5.36	1.87	0.49	1.34	1.34
RLA	OVERALL TEST	40	1,261	25.58	7.65	0.83	3.13	1.28
	Reading	24	1,261	15.99	4.73	0.76	2.31	1.33
	Writing	16	1,261	9.59	3.55	0.65	2.11	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.4.7. 2025 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	6,666	28.77	7.11	0.84	2.81	1.44
	Numerical Representations and Relationships	16	6,666	11.36	3.31	0.70	1.80	1.42
	Computations and Algebraic Relationships	8	6,666	5.58	1.80	0.47	1.31	1.40
	Geometry and Measurement	8	6,666	6.13	1.62	0.53	1.11	1.53
	Data Analysis and Personal Financial Literacy	8	6,666	5.70	1.86	0.50	1.31	1.42
RLA	OVERALL TEST	40	6,668	27.48	7.44	0.83	3.04	1.37
	Reading	24	6,668	17.44	4.66	0.77	2.24	1.45
	Writing	16	6,668	10.04	3.44	0.65	2.04	1.25

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. 2025 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	1,976	28.89	6.92	0.84	2.78	1.44
	Numerical Representations and Relationships	16	1,976	11.40	3.21	0.69	1.79	1.42
	Computations and Algebraic Relationships	8	1,976	5.57	1.75	0.45	1.30	1.39
	Geometry and Measurement	8	1,976	6.15	1.57	0.51	1.11	1.54
	Data Analysis and Personal Financial Literacy	8	1,976	5.77	1.85	0.52	1.28	1.44
RLA	OVERALL TEST	40	1,979	27.81	7.31	0.83	3.01	1.39
	Reading	24	1,979	17.70	4.56	0.76	2.22	1.47
	Writing	16	1,979	10.12	3.40	0.65	2.03	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.4.9. 2025 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	4,690	28.72	7.19	0.85	2.82	1.44
	Numerical Representations and Relationships	16	4,690	11.35	3.35	0.71	1.81	1.42
	Computations and Algebraic Relationships	8	4,690	5.59	1.81	0.48	1.31	1.40
	Geometry and Measurement	8	4,690	6.12	1.63	0.54	1.11	1.53
	Data Analysis and Personal Financial Literacy	8	4,690	5.67	1.86	0.50	1.32	1.42
RLA	OVERALL TEST	40	4,689	27.34	7.49	0.83	3.05	1.37
	Reading	24	4,689	17.33	4.70	0.77	2.25	1.44
	Writing	16	4,689	10.01	3.45	0.65	2.04	1.25

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. 2025 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	1,186	28.30	7.03	0.83	2.86	1.41
	Numerical Representations and Relationships	16	1,186	11.13	3.33	0.70	1.82	1.39
	Computations and Algebraic Relationships	8	1,186	5.46	1.82	0.46	1.34	1.36
	Geometry and Measurement	8	1,186	6.07	1.61	0.52	1.12	1.52
	Data Analysis and Personal Financial Literacy	8	1,186	5.64	1.84	0.48	1.33	1.41
RLA	OVERALL TEST	40	1,184	26.96	7.43	0.83	3.08	1.35
	Reading	24	1,184	17.13	4.70	0.76	2.29	1.43
	Writing	16	1,184	9.83	3.41	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.4.11. 2025 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	3,647	29.42	6.98	0.85	2.74	1.47
	Numerical Representations and Relationships	16	3,647	11.64	3.22	0.70	1.76	1.46
	Computations and Algebraic Relationships	8	3,647	5.70	1.75	0.47	1.28	1.42
	Geometry and Measurement	8	3,647	6.27	1.57	0.53	1.08	1.57
	Data Analysis and Personal Financial Literacy	8	3,647	5.82	1.83	0.51	1.28	1.45
RLA	OVERALL TEST	40	3,652	28.12	7.27	0.83	2.97	1.41
	Reading	24	3,652	17.79	4.53	0.77	2.19	1.48
	Writing	16	3,652	10.33	3.38	0.65	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.4.12. 2025 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	1,224	28.24	7.19	0.84	2.86	1.41
	Numerical Representations and Relationships	16	1,224	11.09	3.37	0.70	1.84	1.39
	Computations and Algebraic Relationships	8	1,224	5.52	1.82	0.46	1.33	1.38
	Geometry and Measurement	8	1,224	5.96	1.63	0.51	1.14	1.49
	Data Analysis and Personal Financial Literacy	8	1,224	5.67	1.88	0.52	1.30	1.42
RLA	OVERALL TEST	40	1,225	27.09	7.39	0.83	3.09	1.35
	Reading	24	1,225	17.31	4.68	0.76	2.27	1.44
	Writing	16	1,225	9.77	3.40	0.63	2.08	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.4.13. 2025 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	6,324	30.26	7.11	0.85	2.77	1.51
	Numerical Representations and Relationships	8	6,324	6.01	1.79	0.49	1.28	1.50
	Computations and Algebraic Relationships	16	6,324	12.36	3.17	0.71	1.70	1.55
	Geometry and Measurement	8	6,324	5.74	1.79	0.47	1.30	1.44
	Data Analysis and Personal Financial Literacy	8	6,324	6.15	1.74	0.48	1.26	1.54
RLA	OVERALL TEST	40	6,321	28.27	7.05	0.82	3.00	1.41
	Reading	24	6,321	17.38	4.35	0.73	2.27	1.45
	Writing	16	6,321	10.89	3.32	0.65	1.96	1.36
Science	OVERALL TEST	40	6,315	30.58	6.81	0.84	2.72	1.53
	Matter and Energy	8	6,315	6.38	1.67	0.46	1.22	1.59
	Force, Motion, and Energy	8	6,315	6.41	1.69	0.54	1.15	1.60
	Earth and Space	16	6,315	11.53	2.88	0.61	1.79	1.44
	Organisms and Environments	8	6,315	6.27	1.87	0.58	1.21	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.4.14. 2025 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	1,965	30.05	6.82	0.83	2.81	1.50
	Numerical Representations and Relationships	8	1,965	5.95	1.74	0.44	1.30	1.49
	Computations and Algebraic Relationships	16	1,965	12.27	3.07	0.68	1.73	1.53
	Geometry and Measurement	8	1,965	5.71	1.76	0.46	1.30	1.43
	Data Analysis and Personal Financial Literacy	8	1,965	6.11	1.74	0.48	1.26	1.53
RLA	OVERALL TEST	40	1,964	28.41	6.91	0.81	2.99	1.42
	Reading	24	1,964	17.52	4.23	0.72	2.25	1.46
	Writing	16	1,964	10.89	3.29	0.64	1.97	1.36
Science	OVERALL TEST	40	1,962	30.76	6.67	0.84	2.71	1.54
	Matter and Energy	8	1,962	6.44	1.65	0.46	1.21	1.61
	Force, Motion, and Energy	8	1,962	6.42	1.68	0.53	1.15	1.61
	Earth and Space	16	1,962	11.58	2.84	0.61	1.78	1.45
	Organisms and Environments	8	1,962	6.32	1.84	0.58	1.20	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.4.15. 2025 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	4,358	30.36	7.24	0.86	2.75	1.52
	Numerical Representations and Relationships	8	4,358	6.03	1.81	0.51	1.27	1.51
	Computations and Algebraic Relationships	16	4,358	12.40	3.21	0.72	1.69	1.55
	Geometry and Measurement	8	4,358	5.75	1.80	0.48	1.30	1.44
	Data Analysis and Personal Financial Literacy	8	4,358	6.17	1.74	0.48	1.26	1.54
RLA	OVERALL TEST	40	4,356	28.21	7.11	0.82	3.00	1.41
	Reading	24	4,356	17.32	4.40	0.73	2.27	1.44
	Writing	16	4,356	10.89	3.34	0.66	1.95	1.36
Science	OVERALL TEST	40	4,352	30.50	6.87	0.84	2.72	1.53
	Matter and Energy	8	4,352	6.35	1.68	0.46	1.23	1.59
	Force, Motion, and Energy	8	4,352	6.40	1.69	0.54	1.15	1.60
	Earth and Space	16	4,352	11.51	2.90	0.62	1.80	1.44
	Organisms and Environments	8	4,352	6.24	1.88	0.59	1.21	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.4.16. 2025 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	1,085	30.01	6.93	0.84	2.80	1.50
	Numerical Representations and Relationships	8	1,085	5.95	1.74	0.42	1.32	1.49
	Computations and Algebraic Relationships	16	1,085	12.20	3.11	0.69	1.73	1.53
	Geometry and Measurement	8	1,085	5.69	1.76	0.44	1.31	1.42
	Data Analysis and Personal Financial Literacy	8	1,085	6.18	1.71	0.47	1.25	1.54
RLA	OVERALL TEST	40	1,085	28.25	6.93	0.81	3.00	1.41
	Reading	24	1,085	17.37	4.32	0.72	2.27	1.45
	Writing	16	1,085	10.88	3.24	0.64	1.95	1.36
Science	OVERALL TEST	40	1,084	30.66	6.72	0.83	2.73	1.53
	Matter and Energy	8	1,084	6.43	1.67	0.48	1.20	1.61
	Force, Motion, and Energy	8	1,084	6.34	1.69	0.51	1.19	1.59
	Earth and Space	16	1,084	11.59	2.84	0.60	1.80	1.45
	Organisms and Environments	8	1,084	6.30	1.85	0.57	1.21	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.4.17. 2025 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	3,511	30.91	6.96	0.85	2.69	1.55
	Numerical Representations and Relationships	8	3,511	6.12	1.75	0.49	1.25	1.53
	Computations and Algebraic Relationships	16	3,511	12.61	3.09	0.71	1.66	1.58
	Geometry and Measurement	8	3,511	5.87	1.77	0.49	1.26	1.47
	Data Analysis and Personal Financial Literacy	8	3,511	6.31	1.69	0.49	1.21	1.58
RLA	OVERALL TEST	40	3,510	28.85	6.81	0.81	2.94	1.44
	Reading	24	3,510	17.73	4.18	0.72	2.22	1.48
	Writing	16	3,510	11.12	3.24	0.65	1.93	1.39
Science	OVERALL TEST	40	3,507	31.14	6.53	0.84	2.65	1.56
	Matter and Energy	8	3,507	6.45	1.62	0.45	1.20	1.61
	Force, Motion, and Energy	8	3,507	6.56	1.61	0.53	1.10	1.64
	Earth and Space	16	3,507	11.74	2.78	0.60	1.76	1.47
	Organisms and Environments	8	3,507	6.39	1.80	0.58	1.16	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.4.18. 2025 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	1,128	29.31	7.55	0.86	2.85	1.47
	Numerical Representations and Relationships	8	1,128	5.82	1.90	0.53	1.31	1.46
	Computations and Algebraic Relationships	16	1,128	11.99	3.38	0.73	1.74	1.50
	Geometry and Measurement	8	1,128	5.56	1.81	0.44	1.35	1.39
	Data Analysis and Personal Financial Literacy	8	1,128	5.92	1.83	0.48	1.32	1.48
RLA	OVERALL TEST	40	1,126	27.62	7.60	0.84	3.06	1.38
	Reading	24	1,126	16.95	4.75	0.77	2.30	1.41
	Writing	16	1,126	10.67	3.47	0.66	2.01	1.33
Science	OVERALL TEST	40	1,124	29.90	7.44	0.86	2.77	1.50
	Matter and Energy	8	1,124	6.29	1.79	0.52	1.24	1.57
	Force, Motion, and Energy	8	1,124	6.28	1.82	0.58	1.18	1.57
	Earth and Space	16	1,124	11.23	3.09	0.65	1.82	1.40
	Organisms and Environments	8	1,124	6.10	2.01	0.62	1.24	1.53

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. 2025 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	5,877	30.03	6.64	0.82	2.78	1.50
	Numerical Representations and Relationships	8	5,877	6.61	1.65	0.55	1.10	1.65
	Computations and Algebraic Relationships	16	5,877	11.82	3.03	0.64	1.81	1.48
	Geometry and Measurement	8	5,877	5.96	1.77	0.54	1.20	1.49
	Data Analysis and Personal Financial Literacy	8	5,877	5.65	1.75	0.43	1.32	1.41
RLA	OVERALL TEST	40	5,874	28.77	6.92	0.82	2.94	1.44
	Reading	24	5,874	17.61	4.31	0.74	2.22	1.47
	Writing	16	5,874	11.16	3.24	0.65	1.92	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. 2025 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	1,862	29.71	6.45	0.81	2.80	1.49
	Numerical Representations and Relationships	8	1,862	6.54	1.63	0.53	1.11	1.64
	Computations and Algebraic Relationships	16	1,862	11.65	2.95	0.62	1.83	1.46
	Geometry and Measurement	8	1,862	5.91	1.73	0.51	1.21	1.48
	Data Analysis and Personal Financial Literacy	8	1,862	5.61	1.73	0.41	1.33	1.40
RLA	OVERALL TEST	40	1,858	28.73	6.84	0.82	2.94	1.44
	Reading	24	1,858	17.63	4.20	0.72	2.22	1.47
	Writing	16	1,858	11.10	3.27	0.65	1.92	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.4.21. 2025 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	4,015	30.18	6.72	0.83	2.77	1.51
	Numerical Representations and Relationships	8	4,015	6.65	1.66	0.56	1.10	1.66
	Computations and Algebraic Relationships	16	4,015	11.89	3.06	0.65	1.80	1.49
	Geometry and Measurement	8	4,015	5.98	1.78	0.55	1.20	1.50
	Data Analysis and Personal Financial Literacy	8	4,015	5.66	1.76	0.44	1.32	1.42
RLA	OVERALL TEST	40	4,016	28.79	6.96	0.82	2.94	1.44
	Reading	24	4,016	17.60	4.37	0.74	2.22	1.47
	Writing	16	4,016	11.19	3.23	0.65	1.92	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.4.22. 2025 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	1,009	29.53	6.82	0.83	2.83	1.48
	Numerical Representations and Relationships	8	1,009	6.54	1.64	0.50	1.16	1.63
	Computations and Algebraic Relationships	16	1,009	11.67	3.10	0.65	1.82	1.46
	Geometry and Measurement	8	1,009	5.78	1.86	0.58	1.21	1.45
	Data Analysis and Personal Financial Literacy	8	1,009	5.54	1.81	0.45	1.35	1.38
RLA	OVERALL TEST	40	1,008	28.39	7.16	0.83	2.97	1.42
	Reading	24	1,008	17.36	4.47	0.75	2.25	1.45
	Writing	16	1,008	11.03	3.33	0.66	1.93	1.38

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.23. 2025 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	3,182	30.68	6.42	0.82	2.71	1.53
	Numerical Representations and Relationships	8	3,182	6.70	1.58	0.55	1.07	1.67
	Computations and Algebraic Relationships	16	3,182	12.04	2.93	0.63	1.77	1.50
	Geometry and Measurement	8	3,182	6.13	1.71	0.53	1.17	1.53
	Data Analysis and Personal Financial Literacy	8	3,182	5.80	1.72	0.44	1.29	1.45
RLA	OVERALL TEST	40	3,182	29.34	6.70	0.82	2.88	1.47
	Reading	24	3,182	17.98	4.15	0.73	2.17	1.50
	Writing	16	3,182	11.36	3.17	0.65	1.89	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. 2025 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	1,088	29.16	6.93	0.83	2.87	1.46
	Numerical Representations and Relationships	8	1,088	6.45	1.77	0.58	1.15	1.61
	Computations and Algebraic Relationships	16	1,088	11.51	3.21	0.67	1.84	1.44
	Geometry and Measurement	8	1,088	5.77	1.80	0.52	1.25	1.44
	Data Analysis and Personal Financial Literacy	8	1,088	5.44	1.78	0.40	1.38	1.36
RLA	OVERALL TEST	40	1,087	28.11	7.28	0.83	3.01	1.41
	Reading	24	1,087	17.20	4.54	0.75	2.28	1.43
	Writing	16	1,087	10.92	3.37	0.66	1.96	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. 2025 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	5,385	31.20	6.28	0.82	2.65	1.56
	Numerical Representations and Relationships	16	5,385	12.75	2.88	0.66	1.67	1.59
	Computations and Algebraic Relationships	8	5,385	5.98	1.71	0.48	1.24	1.49
	Geometry and Measurement	8	5,385	6.69	1.48	0.52	1.03	1.67
	Data Analysis and Personal Financial Literacy	8	5,385	5.78	1.70	0.45	1.26	1.45
RLA	OVERALL TEST	40	5,385	29.62	6.51	0.81	2.82	1.48
	Reading	24	5,385	18.32	4.08	0.73	2.11	1.53
	Writing	16	5,385	11.30	3.06	0.63	1.87	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.4.26. 2025 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	1,728	30.88	6.15	0.81	2.68	1.54
	Numerical Representations and Relationships	16	1,728	12.53	2.82	0.64	1.71	1.57
	Computations and Algebraic Relationships	8	1,728	5.92	1.72	0.48	1.24	1.48
	Geometry and Measurement	8	1,728	6.68	1.45	0.49	1.03	1.67
	Data Analysis and Personal Financial Literacy	8	1,728	5.76	1.68	0.42	1.28	1.44
RLA	OVERALL TEST	40	1,725	29.72	6.43	0.81	2.80	1.49
	Reading	24	1,725	18.43	4.02	0.73	2.08	1.54
	Writing	16	1,725	11.28	3.02	0.62	1.87	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.4.27. 2025 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	3,656	31.35	6.33	0.83	2.64	1.57
	Numerical Representations and Relationships	16	3,656	12.85	2.90	0.67	1.66	1.61
	Computations and Algebraic Relationships	8	3,656	6.01	1.70	0.47	1.24	1.50
	Geometry and Measurement	8	3,656	6.70	1.50	0.53	1.03	1.67
	Data Analysis and Personal Financial Literacy	8	3,656	5.79	1.71	0.46	1.26	1.45
RLA	OVERALL TEST	40	3,659	29.58	6.54	0.81	2.83	1.48
	Reading	24	3,659	18.27	4.12	0.73	2.12	1.52
	Writing	16	3,659	11.31	3.07	0.63	1.87	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.4.28. 2025 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	928	30.77	6.84	0.85	2.67	1.54
	Numerical Representations and Relationships	16	928	12.55	3.06	0.70	1.69	1.57
	Computations and Algebraic Relationships	8	928	5.95	1.76	0.49	1.25	1.49
	Geometry and Measurement	8	928	6.58	1.64	0.61	1.03	1.65
	Data Analysis and Personal Financial Literacy	8	928	5.69	1.79	0.49	1.28	1.42
RLA	OVERALL TEST	40	928	29.48	6.95	0.84	2.80	1.47
	Reading	24	928	18.26	4.31	0.76	2.10	1.52
	Writing	16	928	11.22	3.22	0.67	1.85	1.40

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.29. 2025 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	2,974	31.74	6.02	0.82	2.59	1.59
	Numerical Representations and Relationships	16	2,974	12.94	2.78	0.65	1.64	1.62
	Computations and Algebraic Relationships	8	2,974	6.08	1.68	0.49	1.20	1.52
	Geometry and Measurement	8	2,974	6.81	1.37	0.46	1.01	1.70
	Data Analysis and Personal Financial Literacy	8	2,974	5.91	1.66	0.44	1.24	1.48
RLA	OVERALL TEST	40	2,973	30.19	6.23	0.80	2.77	1.51
	Reading	24	2,973	18.65	3.88	0.71	2.07	1.55
	Writing	16	2,973	11.54	2.99	0.62	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.4.30. 2025 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	1,013	30.42	6.35	0.81	2.74	1.52
	Numerical Representations and Relationships	16	1,013	12.52	2.94	0.66	1.71	1.57
	Computations and Algebraic Relationships	8	1,013	5.77	1.75	0.45	1.30	1.44
	Geometry and Measurement	8	1,013	6.50	1.59	0.55	1.07	1.63
	Data Analysis and Personal Financial Literacy	8	1,013	5.62	1.69	0.41	1.30	1.41
RLA	OVERALL TEST	40	1,013	28.94	6.61	0.81	2.90	1.45
	Reading	24	1,013	17.97	4.25	0.74	2.15	1.50
	Writing	16	1,013	10.97	3.00	0.58	1.93	1.37

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.31. 2025 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	4,900	30.97	6.20	0.81	2.67	1.55
	Numerical Representations and Relationships	16	4,900	12.72	2.84	0.67	1.64	1.59
	Computations and Algebraic Relationships	8	4,900	6.09	1.62	0.44	1.22	1.52
	Geometry and Measurement	8	4,900	6.00	1.58	0.36	1.27	1.50
	Data Analysis and Personal Financial Literacy	8	4,900	6.17	1.68	0.52	1.16	1.54
RLA	OVERALL TEST	40	4,906	29.84	6.37	0.81	2.79	1.49
	Reading	24	4,906	17.83	4.04	0.72	2.14	1.49
	Writing	16	4,906	12.01	2.99	0.65	1.77	1.50
Science	OVERALL TEST	40	4,901	31.50	6.18	0.82	2.61	1.57
	Matter and Energy	8	4,901	6.43	1.53	0.43	1.15	1.61
	Force, Motion, and Energy	8	4,901	6.52	1.54	0.48	1.12	1.63
	Earth and Space	8	4,901	6.04	1.69	0.44	1.27	1.51
	Organisms and Environments	16	4,901	12.51	2.73	0.63	1.66	1.56
Social Studies	OVERALL TEST	40	4,904	31.51	5.89	0.81	2.57	1.58
	History	16	4,904	12.33	2.65	0.60	1.68	1.54
	Geography and Culture	8	4,904	6.48	1.51	0.44	1.13	1.62
	Government and Citizenship	8	4,904	6.21	1.56	0.40	1.21	1.55
	Economics, Science, Technology, and Society	8	4,904	6.50	1.43	0.40	1.11	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. 2025 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	1,630	30.52	6.12	0.81	2.69	1.53
	Numerical Representations and Relationships	16	1,630	12.45	2.85	0.65	1.68	1.56
	Computations and Algebraic Relationships	8	1,630	6.02	1.61	0.42	1.22	1.51
	Geometry and Measurement	8	1,630	5.96	1.55	0.34	1.26	1.49
	Data Analysis and Personal Financial Literacy	8	1,630	6.09	1.67	0.50	1.17	1.52
RLA	OVERALL TEST	40	1,629	29.80	6.31	0.81	2.78	1.49
	Reading	24	1,629	17.81	4.01	0.72	2.14	1.48
	Writing	16	1,629	12.00	2.95	0.64	1.77	1.50
Science	OVERALL TEST	40	1,629	31.46	6.12	0.82	2.59	1.57
	Matter and Energy	8	1,629	6.46	1.50	0.42	1.14	1.61
	Force, Motion, and Energy	8	1,629	6.53	1.53	0.48	1.10	1.63
	Earth and Space	8	1,629	5.99	1.65	0.40	1.28	1.50
	Organisms and Environments	16	1,629	12.49	2.74	0.64	1.64	1.56
Social Studies	OVERALL TEST	40	1,632	31.46	5.80	0.81	2.56	1.57
	History	16	1,632	12.30	2.62	0.59	1.68	1.54
	Geography and Culture	8	1,632	6.47	1.48	0.43	1.12	1.62
	Government and Citizenship	8	1,632	6.21	1.56	0.41	1.19	1.55
	Economics, Science, Technology, and Society	8	1,632	6.48	1.42	0.40	1.10	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.4.33. 2025 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	3,270	31.19	6.22	0.82	2.65	1.56
	Numerical Representations and Relationships	16	3,270	12.85	2.83	0.67	1.62	1.61
	Computations and Algebraic Relationships	8	3,270	6.12	1.63	0.45	1.21	1.53
	Geometry and Measurement	8	3,270	6.01	1.59	0.36	1.27	1.50
	Data Analysis and Personal Financial Literacy	8	3,270	6.21	1.68	0.53	1.16	1.55
RLA	OVERALL TEST	40	3,277	29.85	6.40	0.81	2.79	1.49
	Reading	24	3,277	17.84	4.06	0.72	2.14	1.49
	Writing	16	3,277	12.01	3.01	0.65	1.77	1.50
Science	OVERALL TEST	40	3,272	31.51	6.22	0.82	2.61	1.58
	Matter and Energy	8	3,272	6.42	1.54	0.44	1.16	1.60
	Force, Motion, and Energy	8	3,272	6.52	1.55	0.47	1.12	1.63
	Earth and Space	8	3,272	6.07	1.71	0.46	1.26	1.52
	Organisms and Environments	16	3,272	12.52	2.72	0.63	1.66	1.56
Social Studies	OVERALL TEST	40	3,272	31.53	5.93	0.81	2.57	1.58
	History	16	3,272	12.34	2.66	0.60	1.68	1.54
	Geography and Culture	8	3,272	6.48	1.52	0.45	1.13	1.62
	Government and Citizenship	8	3,272	6.20	1.56	0.39	1.22	1.55
	Economics, Science, Technology, and Society	8	3,272	6.51	1.43	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. 2025 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	848	30.57	6.35	0.82	2.69	1.53
	Numerical Representations and Relationships	16	848	12.57	2.90	0.67	1.66	1.57
	Computations and Algebraic Relationships	8	848	6.01	1.66	0.45	1.23	1.50
	Geometry and Measurement	8	848	5.92	1.61	0.38	1.27	1.48
	Data Analysis and Personal Financial Literacy	8	848	6.08	1.72	0.55	1.16	1.52
RLA	OVERALL TEST	40	851	29.97	6.46	0.82	2.78	1.50
	Reading	24	851	17.93	4.15	0.74	2.13	1.49
	Writing	16	851	12.04	2.96	0.65	1.76	1.50
Science	OVERALL TEST	40	849	31.43	6.32	0.83	2.61	1.57
	Matter and Energy	8	849	6.38	1.60	0.48	1.15	1.59
	Force, Motion, and Energy	8	849	6.55	1.58	0.52	1.09	1.64
	Earth and Space	8	849	6.04	1.72	0.46	1.26	1.51
	Organisms and Environments	16	849	12.46	2.79	0.64	1.67	1.56
Social Studies	OVERALL TEST	40	849	31.61	6.12	0.83	2.55	1.58
	History	16	849	12.44	2.78	0.65	1.65	1.56
	Geography and Culture	8	849	6.49	1.53	0.47	1.11	1.62
	Government and Citizenship	8	849	6.23	1.57	0.40	1.21	1.56
	Economics, Science, Technology, and Society	8	849	6.45	1.48	0.44	1.11	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. 2025 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	2,680	31.47	6.02	0.81	2.61	1.57
	Numerical Representations and Relationships	16	2,680	12.93	2.74	0.66	1.60	1.62
	Computations and Algebraic Relationships	8	2,680	6.18	1.59	0.43	1.20	1.54
	Geometry and Measurement	8	2,680	6.07	1.53	0.34	1.24	1.52
	Data Analysis and Personal Financial Literacy	8	2,680	6.29	1.63	0.51	1.15	1.57
RLA	OVERALL TEST	40	2,684	30.23	6.14	0.80	2.75	1.51
	Reading	24	2,684	18.10	3.89	0.71	2.11	1.51
	Writing	16	2,684	12.13	2.91	0.64	1.75	1.52
Science	OVERALL TEST	40	2,681	31.95	5.88	0.81	2.56	1.60
	Matter and Energy	8	2,681	6.52	1.47	0.42	1.12	1.63
	Force, Motion, and Energy	8	2,681	6.61	1.48	0.46	1.09	1.65
	Earth and Space	8	2,681	6.12	1.63	0.41	1.25	1.53
	Organisms and Environments	16	2,681	12.69	2.62	0.62	1.63	1.59
Social Studies	OVERALL TEST	40	2,685	31.92	5.63	0.80	2.52	1.60
	History	16	2,685	12.47	2.54	0.58	1.65	1.56
	Geography and Culture	8	2,685	6.57	1.45	0.43	1.10	1.64
	Government and Citizenship	8	2,685	6.30	1.53	0.40	1.18	1.57
	Economics, Science, Technology, and Society	8	2,685	6.59	1.37	0.38	1.08	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.4.36. 2025 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	922	30.62	6.22	0.81	2.70	1.53
	Numerical Representations and Relationships	16	922	12.58	2.91	0.68	1.65	1.57
	Computations and Algebraic Relationships	8	922	6.02	1.63	0.42	1.24	1.51
	Geometry and Measurement	8	922	5.99	1.58	0.35	1.28	1.50
	Data Analysis and Personal Financial Literacy	8	922	6.03	1.70	0.52	1.18	1.51
RLA	OVERALL TEST	40	921	29.32	6.60	0.81	2.84	1.47
	Reading	24	921	17.48	4.14	0.72	2.18	1.46
	Writing	16	921	11.84	3.10	0.66	1.81	1.48
Science	OVERALL TEST	40	921	31.14	6.40	0.83	2.65	1.56
	Matter and Energy	8	921	6.38	1.57	0.44	1.18	1.60
	Force, Motion, and Energy	8	921	6.40	1.56	0.45	1.16	1.60
	Earth and Space	8	921	5.97	1.77	0.47	1.28	1.49
	Organisms and Environments	16	921	12.39	2.74	0.62	1.68	1.55
Social Studies	OVERALL TEST	40	921	31.02	6.00	0.81	2.61	1.55
	History	16	921	12.15	2.71	0.60	1.71	1.52
	Geography and Culture	8	921	6.36	1.56	0.45	1.16	1.59
	Government and Citizenship	8	921	6.10	1.54	0.36	1.23	1.53
	Economics, Science, Technology, and Society	8	921	6.42	1.43	0.38	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.4.37. 2025 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	4,932	29.91	6.35	0.81	2.76	1.50
	Number and Algebraic Methods	8	4,932	5.81	1.60	0.41	1.23	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	4,932	5.74	1.68	0.39	1.32	1.43
	Writing and Solving Linear Functions, Equations, and Inequalities	8	4,932	6.33	1.63	0.47	1.19	1.58
	Quadratic Functions and Equations	8	4,932	6.09	1.84	0.60	1.16	1.52
	Exponential Functions and Equations	8	4,932	5.95	1.66	0.41	1.27	1.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.38. 2025 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	1,676	29.88	6.03	0.79	2.78	1.49
	Number and Algebraic Methods	8	1,676	5.80	1.56	0.37	1.24	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	1,676	5.76	1.62	0.34	1.32	1.44
	Writing and Solving Linear Functions, Equations, and Inequalities	8	1,676	6.32	1.59	0.43	1.20	1.58
	Quadratic Functions and Equations	8	1,676	6.07	1.78	0.56	1.18	1.52
	Exponential Functions and Equations	8	1,676	5.93	1.65	0.40	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.4.39. 2025 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	3,256	29.92	6.51	0.82	2.75	1.50
	Number and Algebraic Methods	8	3,256	5.81	1.62	0.43	1.23	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	3,256	5.72	1.71	0.41	1.32	1.43
	Writing and Solving Linear Functions, Equations, and Inequalities	8	3,256	6.33	1.65	0.49	1.19	1.58
	Quadratic Functions and Equations	8	3,256	6.10	1.87	0.62	1.15	1.52
	Exponential Functions and Equations	8	3,256	5.96	1.66	0.42	1.26	1.49

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.40. 2025 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	875	29.54	6.48	0.81	2.79	1.48
	Number and Algebraic Methods	8	875	5.79	1.66	0.45	1.24	1.45
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	875	5.64	1.61	0.30	1.34	1.41
	Writing and Solving Linear Functions, Equations, and Inequalities	8	875	6.30	1.64	0.47	1.20	1.58
	Quadratic Functions and Equations	8	875	5.94	1.97	0.65	1.16	1.49
	Exponential Functions and Equations	8	875	5.87	1.66	0.41	1.28	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. 2025 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	2,661	30.59	6.06	0.80	2.70	1.53
	Number and Algebraic Methods	8	2,661	5.95	1.54	0.38	1.21	1.49
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	2,661	5.87	1.64	0.38	1.29	1.47
	Writing and Solving Linear Functions, Equations, and Inequalities	8	2,661	6.45	1.58	0.47	1.15	1.61
	Quadratic Functions and Equations	8	2,661	6.24	1.74	0.57	1.14	1.56
	Exponential Functions and Equations	8	2,661	6.08	1.61	0.41	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.4.42. 2025 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	973	29.17	6.22	0.79	2.85	1.46
	Number and Algebraic Methods	8	973	5.55	1.60	0.40	1.24	1.39
	Describing and Graphing Linear Functions, Equations, and Inequalities	8	973	5.59	1.72	0.39	1.34	1.40
	Writing and Solving Linear Functions, Equations, and Inequalities	8	973	6.21	1.66	0.44	1.24	1.55
	Quadratic Functions and Equations	8	973	5.98	1.84	0.57	1.21	1.50
	Exponential Functions and Equations	8	973	5.84	1.64	0.36	1.31	1.46

Notes:

1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.43. 2025 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	4,924	29.31	6.38	0.81	2.81	1.47
	Reading	24	4,924	18.15	3.99	0.72	2.10	1.51
	Writing	16	4,924	11.16	3.04	0.63	1.85	1.40
English II	OVERALL TEST	40	4,731	29.89	6.60	0.82	2.79	1.49
	Reading	24	4,731	18.40	4.17	0.75	2.08	1.53
	Writing	16	4,731	11.49	3.04	0.63	1.85	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.4.44. 2025 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	1,680	29.36	6.19	0.79	2.81	1.47
	Reading	24	1,680	18.22	3.88	0.71	2.10	1.52
	Writing	16	1,680	11.14	3.00	0.61	1.86	1.39
English II	OVERALL TEST	40	1,624	29.70	6.67	0.83	2.79	1.48
	Reading	24	1,624	18.27	4.22	0.76	2.07	1.52
	Writing	16	1,624	11.43	3.06	0.63	1.85	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.4.45. 2025 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	3,243	29.28	6.48	0.81	2.80	1.46
	Reading	24	3,243	18.11	4.04	0.73	2.10	1.51
	Writing	16	3,243	11.17	3.06	0.64	1.85	1.40
English II	OVERALL TEST	40	3,105	30.01	6.55	0.82	2.79	1.50
	Reading	24	3,105	18.48	4.13	0.74	2.09	1.54
	Writing	16	3,105	11.53	3.03	0.63	1.85	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.4.46. 2025 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	877	28.89	6.54	0.81	2.86	1.44
	Reading	24	877	17.85	4.07	0.72	2.16	1.49
	Writing	16	877	11.04	3.08	0.63	1.88	1.38
English II	OVERALL TEST	40	796	29.80	6.73	0.83	2.80	1.49
	Reading	24	796	18.39	4.34	0.77	2.07	1.53
	Writing	16	796	11.42	3.06	0.63	1.86	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.4.47. 2025 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	2,649	29.88	6.08	0.79	2.75	1.49
	Reading	24	2,649	18.50	3.80	0.71	2.05	1.54
	Writing	16	2,649	11.37	2.95	0.62	1.83	1.42
English II	OVERALL TEST	40	2,585	30.47	6.38	0.82	2.72	1.52
	Reading	24	2,585	18.74	4.00	0.74	2.03	1.56
	Writing	16	2,585	11.74	2.97	0.63	1.81	1.47

Notes:
1. Maximum possible score points (may exceed the number of items because of multiple-point items).
2. Coefficient alpha computed for polytomous items.
3. Mean of student scores (0, 1, or 2) on items.

Table C.4.48. 2025 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	980	28.95	6.32	0.80	2.86	1.45
	Reading	24	980	17.92	3.96	0.70	2.15	1.49
	Writing	16	980	11.03	2.99	0.61	1.88	1.38
English II	OVERALL TEST	40	1,014	29.08	6.75	0.82	2.88	1.45
	Reading	24	1,014	17.92	4.28	0.75	2.15	1.49
	Writing	16	1,014	11.16	3.07	0.62	1.90	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.4.49. 2025 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	4,992	32.24	5.91	0.82	2.49	1.61
	Cell Structure and Function	8	4,992	6.73	1.41	0.45	1.05	1.68
	Mechanisms of Genetics	8	4,992	6.27	1.55	0.39	1.21	1.57
	Biological Evolution and Classification	8	4,992	6.40	1.57	0.52	1.09	1.60
	Biological Processes and Systems	8	4,992	6.49	1.50	0.47	1.10	1.62
	Interdependence within Environmental Systems	8	4,992	6.34	1.59	0.48	1.14	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.4.50. 2025 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	1,717	32.12	5.80	0.81	2.50	1.61
	Cell Structure and Function	8	1,717	6.71	1.40	0.44	1.05	1.68
	Mechanisms of Genetics	8	1,717	6.26	1.56	0.41	1.20	1.56
	Biological Evolution and Classification	8	1,717	6.41	1.52	0.48	1.10	1.60
	Biological Processes and Systems	8	1,717	6.44	1.49	0.44	1.12	1.61
	Interdependence within Environmental Systems	8	1,717	6.31	1.58	0.45	1.17	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.4.51. 2025 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	3,273	32.30	5.96	0.83	2.48	1.62
	Cell Structure and Function	8	3,273	6.75	1.42	0.46	1.04	1.69
	Mechanisms of Genetics	8	3,273	6.27	1.54	0.39	1.21	1.57
	Biological Evolution and Classification	8	3,273	6.40	1.60	0.54	1.09	1.60
	Biological Processes and Systems	8	3,273	6.52	1.51	0.48	1.09	1.63
	Interdependence within Environmental Systems	8	3,273	6.35	1.59	0.50	1.13	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.4.52. 2025 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	844	31.81	6.38	0.84	2.52	1.59
	Cell Structure and Function	8	844	6.61	1.55	0.52	1.08	1.65
	Mechanisms of Genetics	8	844	6.14	1.65	0.46	1.22	1.53
	Biological Evolution and Classification	8	844	6.39	1.64	0.56	1.09	1.60
	Biological Processes and Systems	8	844	6.43	1.59	0.52	1.11	1.61
	Interdependence within Environmental Systems	8	844	6.25	1.68	0.52	1.16	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.4.53. 2025 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	2,728	32.76	5.49	0.80	2.43	1.64
	Cell Structure and Function	8	2,728	6.82	1.32	0.40	1.02	1.71
	Mechanisms of Genetics	8	2,728	6.38	1.48	0.36	1.18	1.59
	Biological Evolution and Classification	8	2,728	6.54	1.47	0.48	1.06	1.64
	Biological Processes and Systems	8	2,728	6.58	1.43	0.42	1.08	1.65
	Interdependence within Environmental Systems	8	2,728	6.44	1.52	0.46	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.4.54. 2025 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	1,013	31.89	6.02	0.82	2.54	1.59
	Cell Structure and Function	8	1,013	6.70	1.45	0.46	1.07	1.68
	Mechanisms of Genetics	8	1,013	6.23	1.55	0.36	1.24	1.56
	Biological Evolution and Classification	8	1,013	6.21	1.62	0.51	1.14	1.55
	Biological Processes and Systems	8	1,013	6.45	1.55	0.49	1.11	1.61
	Interdependence within Environmental Systems	8	1,013	6.30	1.61	0.49	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.4.55. 2025 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	4,343	31.93	5.89	0.82	2.50	1.60
	History	8	4,343	5.96	1.54	0.35	1.25	1.49
	Geography and Culture	8	4,343	6.54	1.43	0.44	1.07	1.64
	Government and Citizenship	8	4,343	6.32	1.56	0.42	1.19	1.58
	Economics, Science, Technology, and Society	16	4,343	13.11	2.61	0.66	1.53	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.4.56. 2025 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	1,447	31.63	6.09	0.83	2.51	1.58
	History	8	1,447	5.85	1.58	0.38	1.24	1.46
	Geography and Culture	8	1,447	6.49	1.47	0.47	1.07	1.62
	Government and Citizenship	8	1,447	6.30	1.57	0.43	1.19	1.57
	Economics, Science, Technology, and Society	16	1,447	12.99	2.69	0.67	1.54	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.4.57. 2025 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	2,893	32.08	5.78	0.81	2.50	1.60
	History	8	2,893	6.01	1.52	0.33	1.25	1.50
	Geography and Culture	8	2,893	6.56	1.41	0.42	1.07	1.64
	Government and Citizenship	8	2,893	6.33	1.55	0.42	1.18	1.58
	Economics, Science, Technology, and Society	16	2,893	13.17	2.56	0.64	1.53	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.4.58. 2025 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	752	31.82	5.75	0.81	2.49	1.59
	History	8	752	5.88	1.53	0.34	1.24	1.47
	Geography and Culture	8	752	6.47	1.42	0.42	1.08	1.62
	Government and Citizenship	8	752	6.36	1.50	0.38	1.19	1.59
	Economics, Science, Technology, and Society	16	752	13.11	2.54	0.65	1.51	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.4.59. 2025 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	2,322	32.36	5.78	0.82	2.45	1.62
	History	8	2,322	6.07	1.52	0.35	1.23	1.52
	Geography and Culture	8	2,322	6.60	1.41	0.44	1.05	1.65
	Government and Citizenship	8	2,322	6.40	1.53	0.43	1.16	1.60
	Economics, Science, Technology, and Society	16	2,322	13.28	2.53	0.65	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.4.60. 2025 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	947	31.53	5.79	0.80	2.57	1.58
	History	8	947	5.84	1.52	0.30	1.27	1.46
	Geography and Culture	8	947	6.52	1.42	0.41	1.09	1.63
	Government and Citizenship	8	947	6.21	1.56	0.39	1.22	1.55
	Economics, Science, Technology, and Society	16	947	12.95	2.59	0.62	1.58	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.

2025 STAAR Alternate 2
Progress Measure

Table C.5.1. 2025 STAAR Alternate 2 All Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	386	114	29	76	90	65	12	0	143	243
		30%	8%	20%	23%	17%	3%	0%	37%	63%
Stage 2	279	37	26	70	77	59	10	37	96	146
		13%	9%	25%	28%	21%	4%	13%	34%	52%
Stage 3	1565	87	66	351	572	422	67	153	923	489
		6%	4%	22%	37%	27%	4%	10%	59%	31%
Stage 4	6375	76	75	513	2225	2823	663	664	5048	663
		1%	1%	8%	35%	44%	10%	10%	79%	10%
Stage 5	13929	65	46	411	2748	7191	3468	3270	10659	0
		0%	0%	3%	20%	52%	25%	23%	77%	0%
Stage 6	7497	14	6	57	610	2488	4322	3175	0	4322
		0%	0%	1%	8%	33%	58%	42%	0%	58%
2024 Total*	30031	393	248	1478	6322	13048	8542	7299	16869	5863
		1%	1%	5%	21%	43%	28%	24%	56%	20%

Notes:

* Due to rounding, percentages may not sum to 100.

Table C.5.2. 2025 STAAR Alternate 2 All RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	597	196	57	137	104	78	25	0	253	344
		33%	10%	23%	17%	13%	4%	0%	42%	58%
Stage 2	437	46	43	145	132	56	15	46	188	203
		11%	10%	33%	30%	13%	3%	11%	43%	46%
Stage 3	3092	130	133	787	1160	688	194	263	1947	882
		4%	4%	25%	38%	22%	6%	9%	63%	29%
Stage 4	8802	116	92	1055	3306	3144	1089	1263	6450	1089
		1%	1%	12%	38%	36%	12%	14%	73%	12%
Stage 5	13777	69	78	747	3285	6296	3302	4179	9598	0
		1%	1%	5%	24%	46%	24%	30%	70%	0%
Stage 6	7763	14	11	149	761	2538	4290	3473	0	4290
		0%	0%	2%	10%	33%	55%	45%	0%	55%
2024 Total*	34468	571	414	3020	8748	12800	8915	9224	18436	6808
		2%	1%	9%	25%	37%	26%	27%	53%	20%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.3. 2025 STAAR Alternate 2 Grade 4 Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	114	45	5	26	23	14	1	0	50	64
		39%	4%	23%	20%	12%	1%	0%	44%	56%
Stage 2	92	10	10	28	28	15	1	10	38	44
		11%	11%	30%	30%	16%	1%	11%	41%	48%
Stage 3	422	11	20	102	149	129	11	31	251	140
		3%	5%	24%	35%	31%	3%	7%	59%	33%
Stage 4	1547	14	19	120	513	812	69	153	1325	69
		1%	1%	8%	33%	52%	4%	10%	86%	4%
Stage 5	2517	11	9	51	379	1750	317	450	2067	0
		0%	0%	2%	15%	70%	13%	18%	82%	0%
Stage 6	1086	2	0	9	78	596	401	685	0	401
		0%	0%	1%	7%	55%	37%	63%	0%	37%
2024 Total*	5778	93	63	336	1170	3316	800	1329	3731	718
		2%	1%	6%	20%	57%	14%	23%	65%	12%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.4. 2025 STAAR Alternate 2 Grade 5 Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	71	24	6	16	12	11	2	0	30	41
		34%	8%	23%	17%	15%	3%	0%	42%	58%
Stage 2	64	13	8	16	16	8	3	13	24	27
		20%	13%	25%	25%	13%	5%	20%	38%	42%
Stage 3	305	21	8	91	114	68	3	29	205	71
		7%	3%	30%	37%	22%	1%	10%	67%	23%
Stage 4	1106	9	11	115	372	540	59	135	912	59
		1%	1%	10%	34%	49%	5%	12%	82%	5%
Stage 5	3267	15	11	102	511	1827	801	639	2628	0
		0%	0%	3%	16%	56%	25%	20%	80%	0%
Stage 6	793	1	0	6	32	269	485	308	0	485
		0%	0%	1%	4%	34%	61%	39%	0%	61%
2024 Total*	5606	83	44	346	1057	2723	1353	1124	3799	683
		1%	1%	6%	19%	49%	24%	20%	68%	12%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.5. 2025 STAAR Alternate 2 Grade 6 Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	58	12	2	13	20	10	1	0	14	44
		21%	3%	22%	34%	17%	2%	0%	24%	76%
Stage 2	31	1	0	11	8	10	1	1	11	19
		3%	0%	35%	26%	32%	3%	3%	35%	61%
Stage 3	291	17	12	50	106	92	14	29	156	106
		6%	4%	17%	36%	32%	5%	10%	54%	36%
Stage 4	1032	19	20	81	350	408	154	120	758	154
		2%	2%	8%	34%	40%	15%	12%	73%	15%
Stage 5	2544	14	11	85	542	1163	729	652	1892	0
		1%	0%	3%	21%	46%	29%	26%	74%	0%
Stage 6	1142	3	3	6	72	378	680	462	0	680
		0%	0%	1%	6%	33%	60%	40%	0%	60%
2024 Total*	5098	66	48	246	1098	2061	1579	1264	2831	1003
		1%	1%	5%	22%	40%	31%	25%	56%	20%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.6. 2025 STAAR Alternate 2 Grade 7 Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	78	15	13	9	19	19	3	0	28	50
		19%	17%	12%	24%	24%	4%	0%	36%	64%
Stage 2	43	7	3	7	8	16	2	7	10	26
		16%	7%	16%	19%	37%	5%	16%	23%	60%
Stage 3	221	13	13	29	83	71	12	26	112	83
		6%	6%	13%	38%	32%	5%	12%	51%	38%
Stage 4	977	8	11	60	306	495	97	79	801	97
		1%	1%	6%	31%	51%	10%	8%	82%	10%
Stage 5	2004	3	5	35	336	1085	540	379	1625	0
		0%	0%	2%	17%	54%	27%	19%	81%	0%
Stage 6	1549	0	0	7	79	525	938	611	0	938
		0%	0%	0%	5%	34%	61%	39%	0%	61%
2024 Total*	4872	46	45	147	831	2211	1592	1102	2576	1194
		1%	1%	3%	17%	45%	33%	23%	53%	25%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.7. 2025 STAAR Alternate 2 Grade 8 Mathematics Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	34	11	2	5	7	7	2	0	13	21
		32%	6%	15%	21%	21%	6%	0%	38%	62%
Stage 2	27	2	4	7	7	5	2	2	11	14
		7%	15%	26%	26%	19%	7%	7%	41%	52%
Stage 3	166	12	9	34	70	33	8	21	104	41
		7%	5%	20%	42%	20%	5%	13%	63%	25%
Stage 4	748	12	3	65	309	281	78	80	590	78
		2%	0%	9%	41%	38%	10%	11%	79%	10%
Stage 5	2102	11	7	68	485	896	635	571	1531	0
		1%	0%	3%	23%	43%	30%	27%	73%	0%
Stage 6	1442	2	2	8	104	351	975	467	0	975
		0%	0%	1%	7%	24%	68%	32%	0%	68%
2024 Total*	4519	50	27	187	982	1573	1700	1141	2249	1129
		1%	1%	4%	22%	35%	38%	25%	50%	25%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.8. 2025 STAAR Alternate 2 Algebra I Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	31	7	1	7	9	4	3	0	8	23
		23%	3%	23%	29%	13%	10%	0%	26%	74%
Stage 2	22	4	1	1	10	5	1	4	2	16
		18%	5%	5%	45%	23%	5%	18%	9%	73%
Stage 3	160	13	4	45	50	29	19	17	95	48
		8%	3%	28%	31%	18%	12%	11%	59%	30%
Stage 4	965	14	11	72	375	287	206	97	662	206
		1%	1%	7%	39%	30%	21%	10%	69%	21%
Stage 5	1495	11	3	70	495	470	446	579	916	0
		1%	0%	5%	33%	31%	30%	39%	61%	0%
Stage 6	1485	6	1	21	245	369	843	642	0	843
		0%	0%	1%	16%	25%	57%	43%	0%	57%
2024 Total*	4158	55	21	216	1184	1164	1518	1339	1683	1136
		1%	1%	5%	28%	28%	37%	32%	40%	27%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.9. 2025 STAAR Alternate 2 Grade 4 RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	140	56	17	28	22	16	1	0	73	67
		40%	12%	20%	16%	11%	1%	0%	52%	48%
Stage 2	97	9	12	36	25	13	2	9	48	40
		9%	12%	37%	26%	13%	2%	9%	49%	41%
Stage 3	503	32	23	148	197	95	8	55	345	103
		6%	5%	29%	39%	19%	2%	11%	69%	20%
Stage 4	1196	29	14	162	511	418	62	205	929	62
		2%	1%	14%	43%	35%	5%	17%	78%	5%
Stage 5	3077	20	18	176	824	1598	441	1038	2039	0
		1%	1%	6%	27%	52%	14%	34%	66%	0%
Stage 6	769	1	2	20	77	352	317	452	0	317
		0%	0%	3%	10%	46%	41%	59%	0%	41%
2024 Total*	5782	147	86	570	1656	2492	831	1759	3434	589
		3%	1%	10%	29%	43%	14%	30%	59%	10%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.10. 2025 STAAR Alternate 2 Grade 5 RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	122	46	12	38	14	11	1	0	58	64
		38%	10%	31%	11%	9%	1%	0%	48%	52%
Stage 2	78	8	5	39	22	3	1	8	44	26
		10%	6%	50%	28%	4%	1%	10%	56%	33%
Stage 3	496	23	23	183	184	72	11	46	367	83
		5%	5%	37%	37%	15%	2%	9%	74%	17%
Stage 4	1544	24	20	311	582	500	107	355	1082	107
		2%	1%	20%	38%	32%	7%	23%	70%	7%
Stage 5	2501	10	21	178	645	1172	475	854	1647	0
		0%	1%	7%	26%	47%	19%	34%	66%	0%
Stage 6	860	3	1	13	87	371	385	475	0	385
		0%	0%	2%	10%	43%	45%	55%	0%	45%
2024 Total*	5601	114	82	762	1534	2129	980	1738	3198	665
		2%	1%	14%	27%	38%	17%	31%	57%	12%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.11. 2025 STAAR Alternate 2 Grade 6 RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	71	17	9	12	21	7	5	0	26	45
		24%	13%	17%	30%	10%	7%	0%	37%	63%
Stage 2	81	4	12	19	30	14	2	4	31	46
		5%	15%	23%	37%	17%	2%	5%	38%	57%
Stage 3	687	29	25	150	278	160	45	54	428	205
		4%	4%	22%	40%	23%	7%	8%	62%	30%
Stage 4	1358	18	17	122	550	425	226	157	975	226
		1%	1%	9%	41%	31%	17%	12%	72%	17%
Stage 5	1987	14	19	128	534	709	583	695	1292	0
		1%	1%	6%	27%	36%	29%	35%	65%	0%
Stage 6	911	2	1	32	158	281	437	474	0	437
		0%	0%	4%	17%	31%	48%	52%	0%	48%
2024 Total*	5095	84	83	463	1571	1596	1298	1384	2752	959
		2%	2%	9%	31%	31%	25%	27%	54%	19%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.12. 2025 STAAR Alternate 2 Grade 7 RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	84	30	6	16	16	13	3	0	36	48
		36%	7%	19%	19%	15%	4%	0%	43%	57%
Stage 2	59	10	7	19	15	4	4	10	26	23
		17%	12%	32%	25%	7%	7%	17%	44%	39%
Stage 3	410	13	21	107	156	97	16	34	263	113
		3%	5%	26%	38%	24%	4%	8%	64%	28%
Stage 4	1527	17	20	149	608	605	128	186	1213	128
		1%	1%	10%	40%	40%	8%	12%	79%	8%
Stage 5	1552	7	8	82	367	773	315	464	1088	0
		0%	1%	5%	24%	50%	20%	30%	70%	0%
Stage 6	1240	1	0	18	124	447	650	590	0	650
		0%	0%	1%	10%	36%	52%	48%	0%	52%
2024 Total*	4872	78	62	391	1286	1939	1116	1284	2626	962
		2%	1%	8%	26%	40%	23%	26%	54%	20%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.13. 2025 STAAR Alternate 2 Grade 8 RLA Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	60	16	3	18	12	7	4	0	19	41
		27%	5%	30%	20%	12%	7%	0%	32%	68%
Stage 2	50	5	5	16	16	7	1	5	21	24
		10%	10%	32%	32%	14%	2%	10%	42%	48%
Stage 3	390	12	16	98	151	88	25	28	249	113
		3%	4%	25%	39%	23%	6%	7%	64%	29%
Stage 4	1243	13	7	144	509	430	140	164	939	140
		1%	1%	12%	41%	35%	11%	13%	76%	11%
Stage 5	1783	9	7	64	419	774	510	499	1284	0
		1%	0%	4%	23%	43%	29%	28%	72%	0%
Stage 6	1001	2	1	13	79	241	665	336	0	665
		0%	0%	1%	8%	24%	66%	34%	0%	66%
2024 Total*	4527	57	39	353	1186	1547	1345	1032	2512	983
		1%	1%	8%	26%	34%	30%	23%	55%	22%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.14. 2025 STAAR Alternate 2 English I Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	48	7	4	8	8	15	6	0	11	37
		15%	8%	17%	17%	31%	13%	0%	23%	77%
Stage 2	40	5	0	10	11	11	3	5	10	25
		13%	0%	25%	28%	28%	8%	13%	25%	63%
Stage 3	363	12	13	59	106	108	65	25	165	173
		3%	4%	16%	29%	30%	18%	7%	45%	48%
Stage 4	1146	10	11	87	285	453	300	108	738	300
		1%	1%	8%	25%	40%	26%	9%	64%	26%
Stage 5	1378	7	4	64	230	549	524	305	1073	0
		1%	0%	5%	17%	40%	38%	22%	78%	0%
Stage 6	1216	2	3	18	109	341	743	473	0	743
		0%	0%	1%	9%	28%	61%	39%	0%	61%
2024 Total*	4191	43	35	246	749	1477	1641	916	1997	1278
		1%	1%	6%	18%	35%	39%	22%	48%	30%

Notes:
* Due to rounding, percentages may not sum to 100.

Table C.5.15. 2025 STAAR Alternate 2 English II Progress Measure

		2025						2025 Total*		
2024	**N**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Stage 5**	**Stage 6**	**Did Not Meet**	**Met**	**Exceeded**
Stage 1	72	24	6	17	11	9	5	0	30	42
		33%	8%	24%	15%	13%	7%	0%	42%	58%
Stage 2	32	5	2	6	13	4	2	5	8	19
		16%	6%	19%	41%	13%	6%	16%	25%	59%
Stage 3	243	9	12	42	88	68	24	21	130	92
		4%	5%	17%	36%	28%	10%	9%	53%	38%
Stage 4	788	5	3	80	261	313	126	88	574	126
		1%	0%	10%	33%	40%	16%	11%	73%	16%
Stage 5	1499	2	1	55	266	721	454	324	1175	0
		0%	0%	4%	18%	48%	30%	22%	78%	0%
Stage 6	1766	3	3	35	127	505	1093	673	0	1093
		0%	0%	2%	7%	29%	62%	38%	0%	62%
2024 Total*	4400	48	27	235	766	1620	1704	1111	1917	1372
		1%	1%	5%	17%	37%	39%	25%	44%	31%

Notes:
* Due to rounding, percentages may not sum to 100.

2025 STAAR Alternate 2
Scale Score Descriptive Statistics

Table C.6.1. 2025 STAAR Alternate 2 Grades 3–8
Scale Score Descriptive Statistics

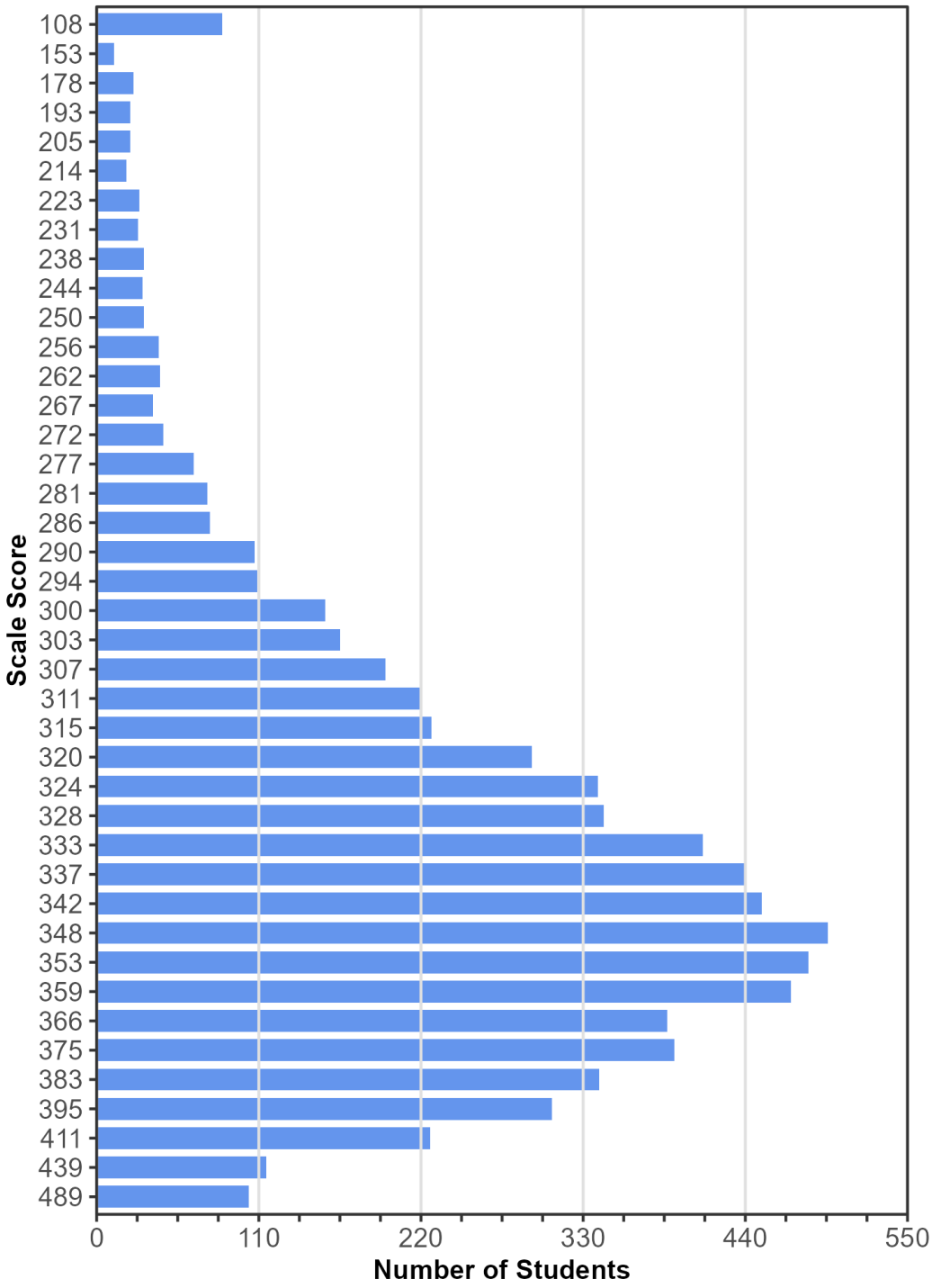
	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 3	Mathematics	7,510	337.52	342	348	381	44	50.43	2543.50	-1.02	5.07
Grade 4	Mathematics	6,666	345.84	347	347	372	45	45.73	2090.84	-0.88	5.64
Grade 5	Mathematics	6,324	350.59	345	357	369	46	49.26	2426.71	-0.18	3.50
Grade 6	Mathematics	5,877	352.68	350	356	408	52	50.13	2513.51	-0.42	5.02
Grade 7	Mathematics	5,385	361.65	356	350	400	55	50.26	2525.83	0.00	3.57
Grade 8	Mathematics	4,900	355.43	352	352	402	47	50.17	2516.99	-0.08	4.18
Grade 3	RLA	7,516	334.79	340	340	432	47	52.69	2776.45	-1.22	5.25
Grade 4	RLA	6,668	335.11	338	350	422	51	50.12	2511.78	-0.87	5.12
Grade 5	RLA	6,321	333.63	331	349	412	45	49.08	2408.86	-0.37	4.73
Grade 6	RLA	5,874	341.08	340	340	420	56	51.19	2620.51	-0.27	4.42
Grade 7	RLA	5,385	347.97	342	348	428	54	52.51	2756.93	0.09	3.50
Grade 8	RLA	4,906	345.95	344	338	414	50	50.63	2563.05	0.02	4.29
Grade 5	Science	6,315	355.53	355	368	386	47	48.35	2337.70	-0.58	4.60
Grade 8	Science	4,901	358.52	357	364	345	45	43.86	1923.59	-0.07	4.00
Grade 8	Social Studies	4,904	354.22	349	362	375	49	45.40	2061.50	-0.21	4.44

Table C.6.2. 2025 STAAR Alternate 2 EOC Assessments Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Algebra I	4,932	346.27	345	345	401	42	47.14	2222.01	-0.18	4.88
English I	4,924	352.21	349	349	426	52	48.74	2375.81	-0.10	5.05
English II	4,731	355.36	352	359	423	54	51.77	2680.07	-0.10	4.38
Biology	4,992	363.04	357	364	347	43	44.26	1958.75	-0.14	3.76
US History	4,343	357.11	354	380	382	50	48.42	2344.21	-0.23	4.17

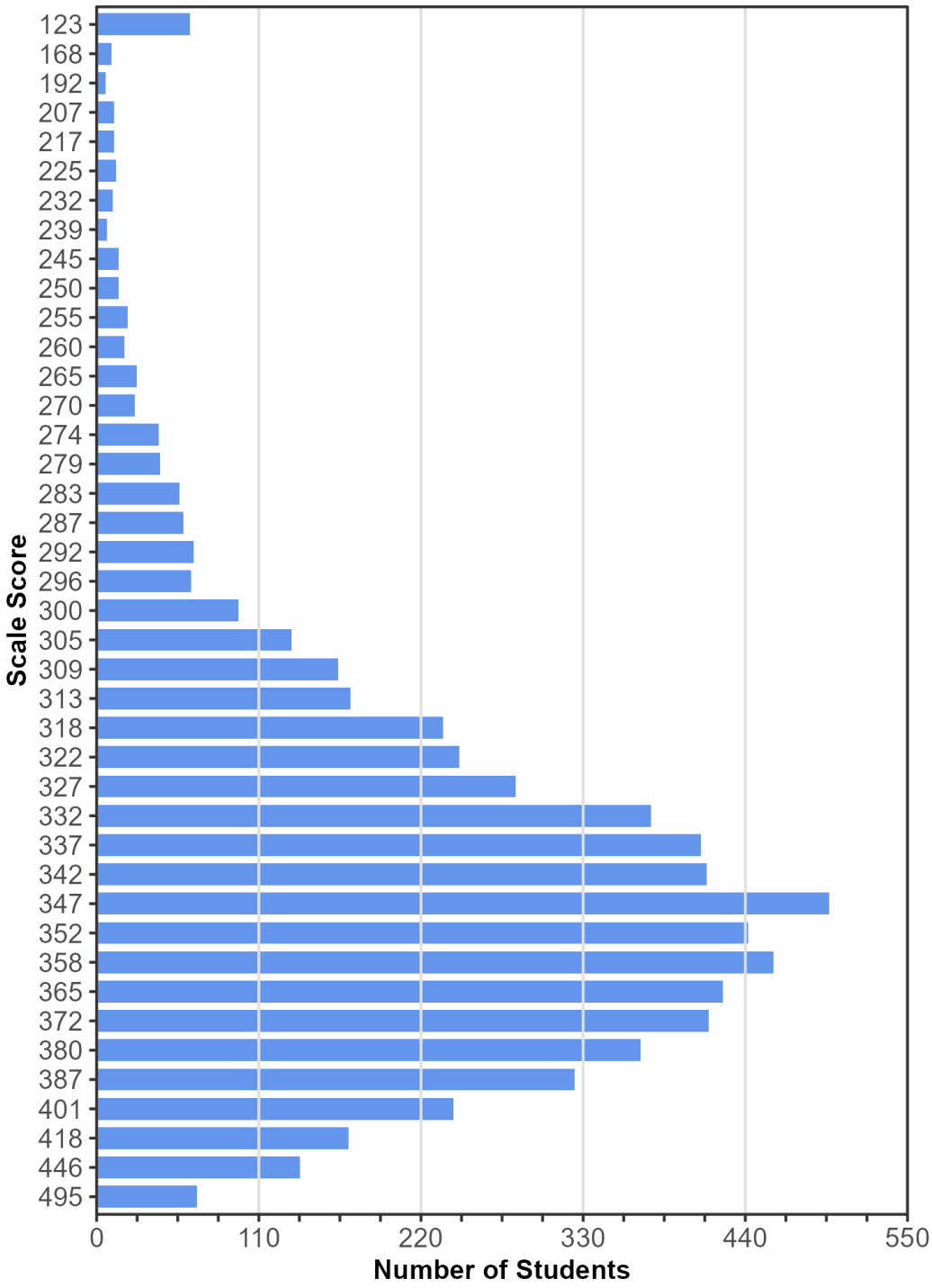
2025 STAAR Alternate 2
Frequency Distribution of Scale Scores

Figure C.6.1. 2025 STAAR Alternate 2 Grade 3 Mathematics
Frequency Distribution of Scale Scores



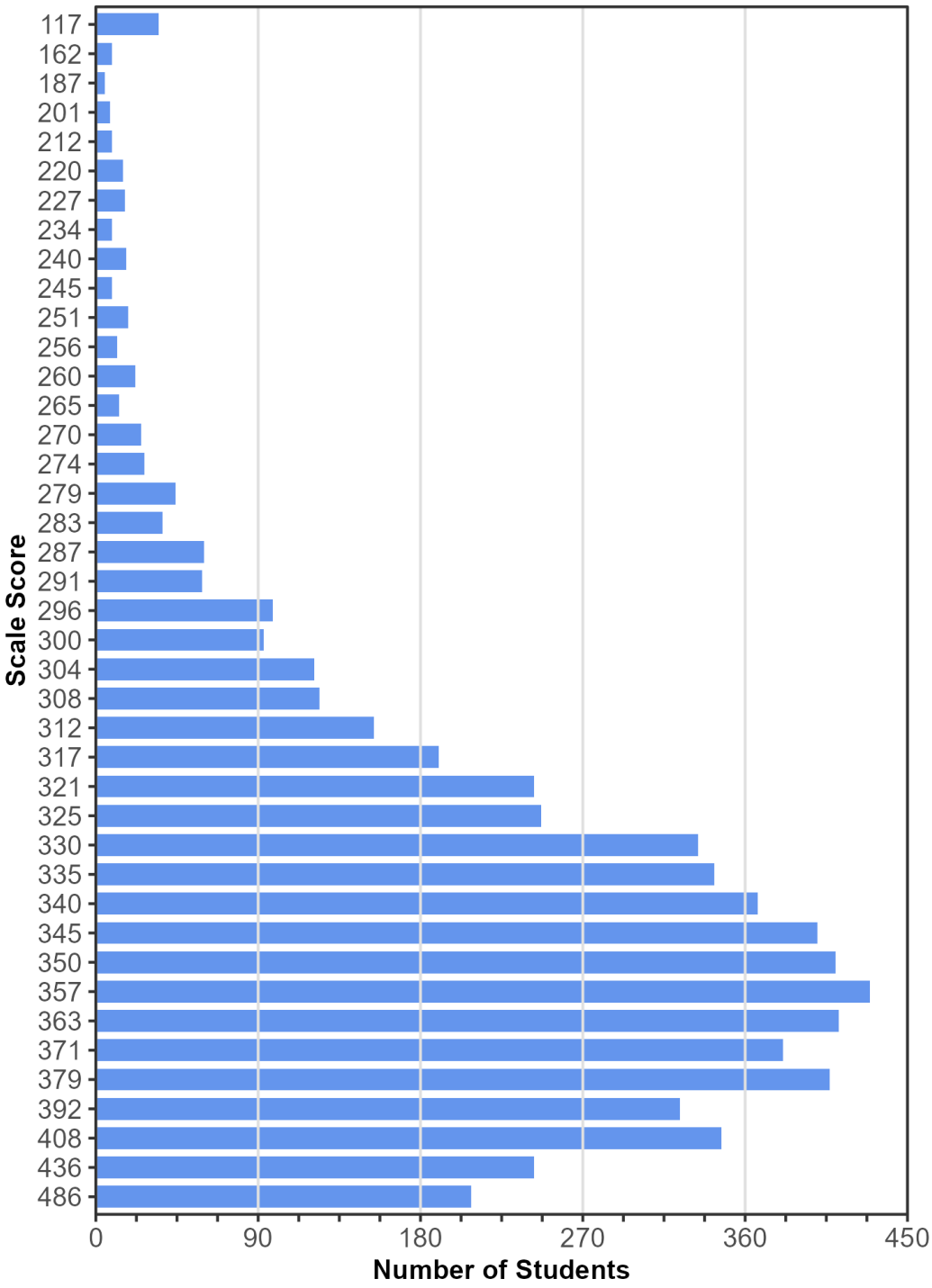
Freq	Cum Freq	Pct	Cum Pct
85	85	1.13	1.13
12	97	0.16	1.29
25	122	0.33	1.62
23	145	0.31	1.93
23	168	0.31	2.24
20	188	0.27	2.50
29	217	0.39	2.89
28	245	0.37	3.26
32	277	0.43	3.69
31	308	0.41	4.10
32	340	0.43	4.53
42	382	0.56	5.09
43	425	0.57	5.66
38	463	0.51	6.17
45	508	0.60	6.76
66	574	0.88	7.64
75	649	1.00	8.64
77	726	1.03	9.67
107	833	1.42	11.09
109	942	1.45	12.54
155	1097	2.06	14.61
165	1262	2.20	16.80
196	1458	2.61	19.41
221	1679	2.94	22.36
227	1906	3.02	25.38
295	2201	3.93	29.31
340	2541	4.53	33.83
344	2885	4.58	38.42
411	3296	5.47	43.89
440	3736	5.86	49.75
451	4187	6.01	55.75
496	4683	6.60	62.36
483	5166	6.43	68.79
471	5637	6.27	75.06
387	6024	5.15	80.21
392	6416	5.22	85.43
341	6757	4.54	89.97
309	7066	4.11	94.09
226	7292	3.01	97.10
115	7407	1.53	98.63
103	7510	1.37	100.00

Figure C.6.2. 2025 STAAR Alternate 2 Grade 4 Mathematics
Frequency Distribution of Scale Scores



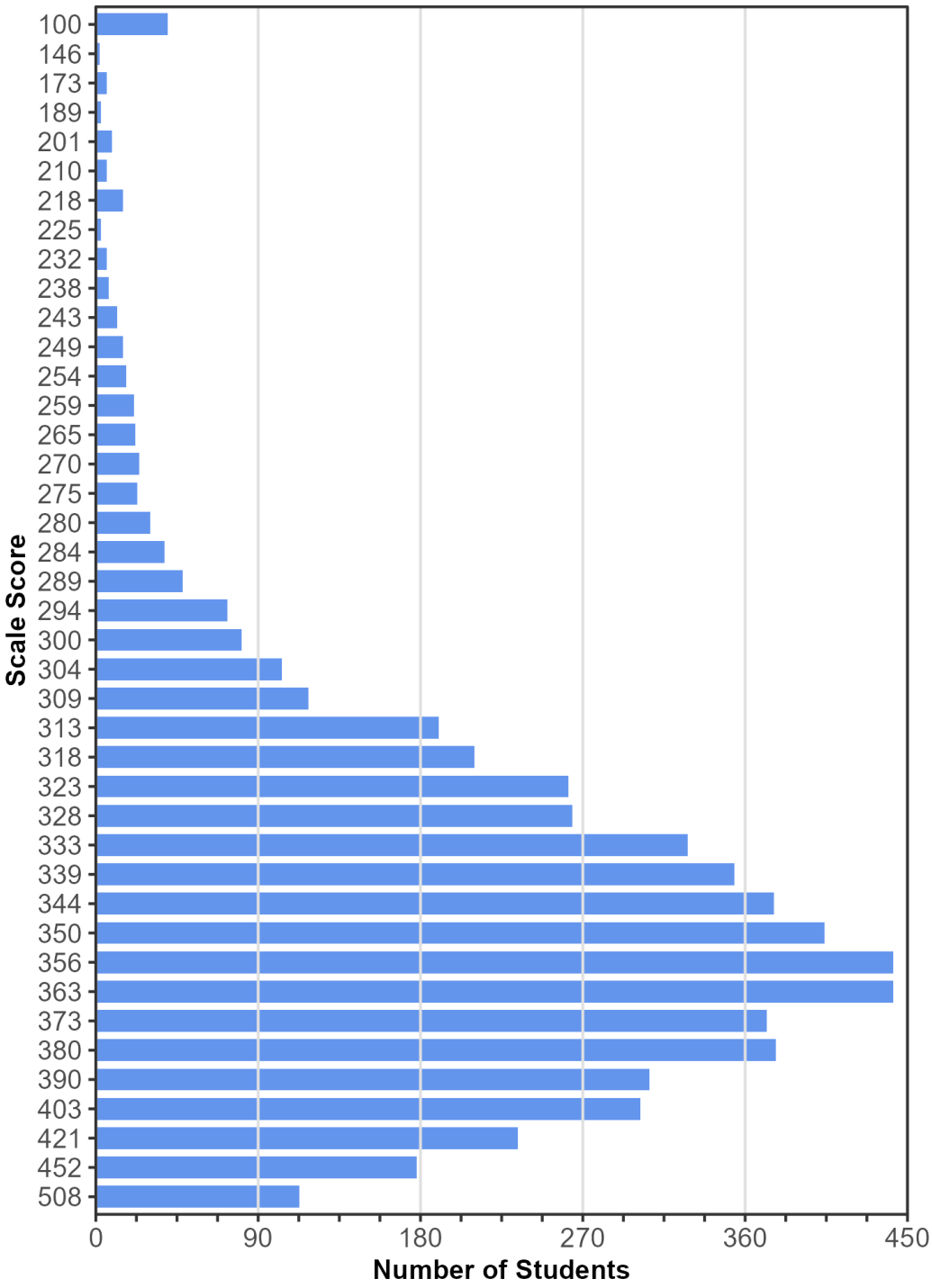
Scale Score	Freq	Cum Freq	Pct	Cum Pct
123	63	63	0.95	0.95
168	10	73	0.15	1.10
192	6	79	0.09	1.19
207	12	91	0.18	1.37
217	12	103	0.18	1.55
225	13	116	0.20	1.74
232	11	127	0.17	1.91
239	7	134	0.11	2.01
245	15	149	0.23	2.24
250	15	164	0.23	2.46
255	21	185	0.32	2.78
260	19	204	0.29	3.06
265	27	231	0.41	3.47
270	26	257	0.39	3.86
274	42	299	0.63	4.49
279	43	342	0.65	5.13
283	56	398	0.84	5.97
287	59	457	0.89	6.86
292	66	523	0.99	7.85
296	64	587	0.96	8.81
300	96	683	1.44	10.25
305	132	815	1.98	12.23
309	164	979	2.46	14.69
313	172	1151	2.58	17.27
318	235	1386	3.53	20.79
322	246	1632	3.69	24.48
327	284	1916	4.26	28.74
332	376	2292	5.64	34.38
337	410	2702	6.15	40.53
342	414	3116	6.21	46.74
347	497	3613	7.46	54.20
352	442	4055	6.63	60.83
358	459	4514	6.89	67.72
365	425	4939	6.38	74.09
372	415	5354	6.23	80.32
380	369	5723	5.54	85.85
387	324	6047	4.86	90.71
401	242	6289	3.63	94.34
418	171	6460	2.57	96.91
446	138	6598	2.07	98.98
495	68	6666	1.02	100.00

Figure C.6.3. 2025 STAAR Alternate 2 Grade 5 Mathematics
Frequency Distribution of Scale Scores



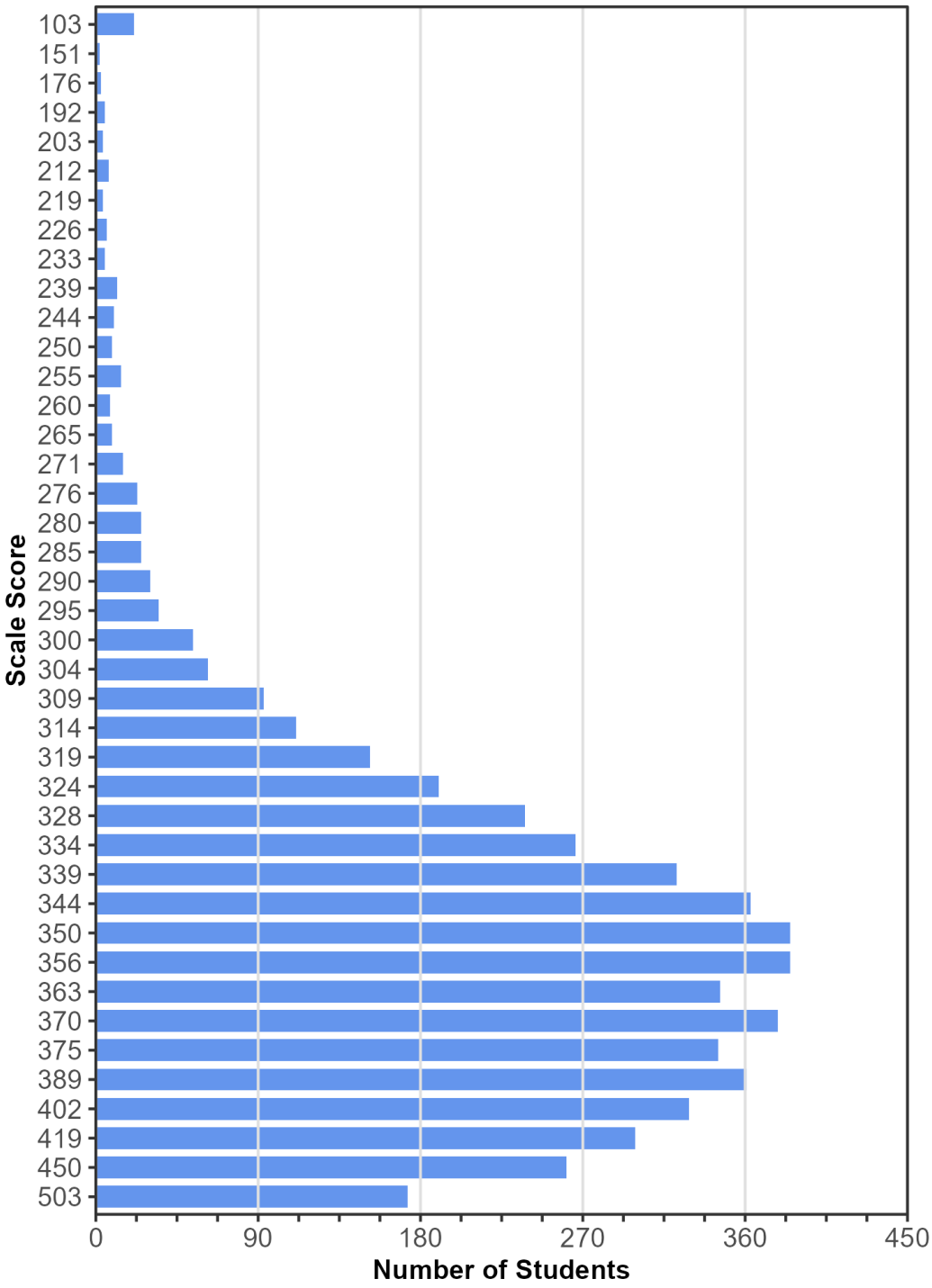
Scale Score	Freq	Cum Freq	Pct	Cum Pct
117	35	35	0.55	0.55
162	9	44	0.14	0.70
187	5	49	0.08	0.77
201	8	57	0.13	0.90
212	9	66	0.14	1.04
220	15	81	0.24	1.28
227	16	97	0.25	1.53
234	9	106	0.14	1.68
240	17	123	0.27	1.94
245	9	132	0.14	2.09
251	18	150	0.28	2.37
256	12	162	0.19	2.56
260	22	184	0.35	2.91
265	13	197	0.21	3.12
270	25	222	0.40	3.51
274	27	249	0.43	3.94
279	44	293	0.70	4.63
283	37	330	0.59	5.22
287	60	390	0.95	6.17
291	59	449	0.93	7.10
296	98	547	1.55	8.65
300	93	640	1.47	10.12
304	121	761	1.91	12.03
308	124	885	1.96	13.99
312	154	1039	2.44	16.43
317	190	1229	3.00	19.43
321	243	1472	3.84	23.28
325	247	1719	3.91	27.18
330	334	2053	5.28	32.46
335	343	2396	5.42	37.89
340	367	2763	5.80	43.69
345	400	3163	6.33	50.02
350	410	3573	6.48	56.50
357	429	4002	6.78	63.28
363	412	4414	6.51	69.80
371	381	4795	6.02	75.82
379	407	5202	6.44	82.26
392	324	5526	5.12	87.38
408	347	5873	5.49	92.87
436	243	6116	3.84	96.71
486	208	6324	3.29	100.00

Figure C.6.4. 2025 STAAR Alternate 2 Grade 6 Mathematics
Frequency Distribution of Scale Scores



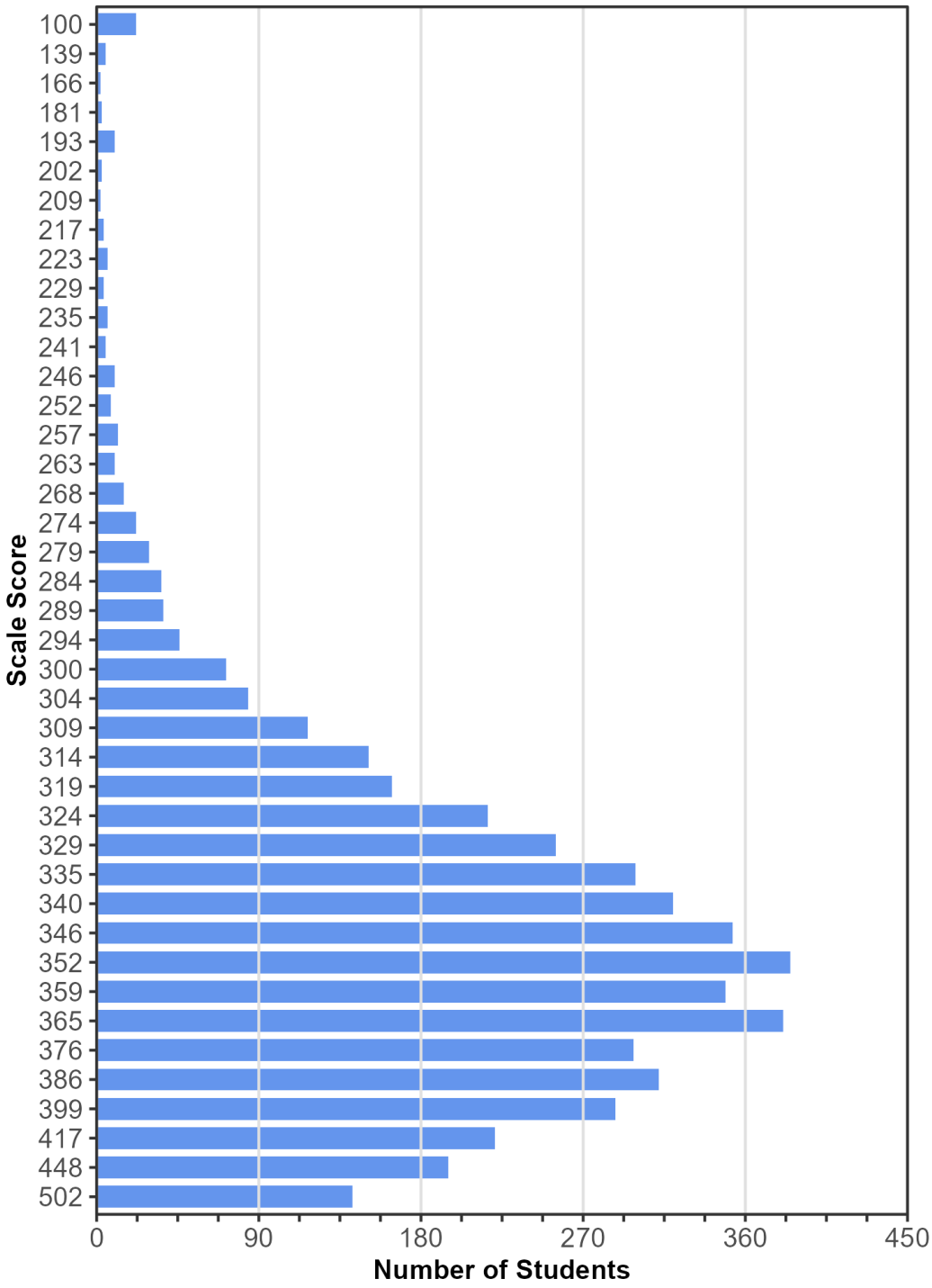
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	40	40	0.68	0.68
146	2	42	0.03	0.71
173	6	48	0.10	0.82
189	3	51	0.05	0.87
201	9	60	0.15	1.02
210	6	66	0.10	1.12
218	15	81	0.26	1.38
225	3	84	0.05	1.43
232	6	90	0.10	1.53
238	7	97	0.12	1.65
243	12	109	0.20	1.85
249	15	124	0.26	2.11
254	17	141	0.29	2.40
259	21	162	0.36	2.76
265	22	184	0.37	3.13
270	24	208	0.41	3.54
275	23	231	0.39	3.93
280	30	261	0.51	4.44
284	38	299	0.65	5.09
289	48	347	0.82	5.90
294	73	420	1.24	7.15
300	81	501	1.38	8.52
304	103	604	1.75	10.28
309	118	722	2.01	12.29
313	190	912	3.23	15.52
318	210	1122	3.57	19.09
323	262	1384	4.46	23.55
328	264	1648	4.49	28.04
333	328	1976	5.58	33.62
339	354	2330	6.02	39.65
344	376	2706	6.40	46.04
350	404	3110	6.87	52.92
356	442	3552	7.52	60.44
363	442	3994	7.52	67.96
373	372	4366	6.33	74.29
380	377	4743	6.41	80.70
390	307	5050	5.22	85.93
403	302	5352	5.14	91.07
421	234	5586	3.98	95.05
452	178	5764	3.03	98.08
508	113	5877	1.92	100.00

Figure C.6.5. 2025 STAAR Alternate 2 Grade 7 Mathematics
Frequency Distribution of Scale Scores



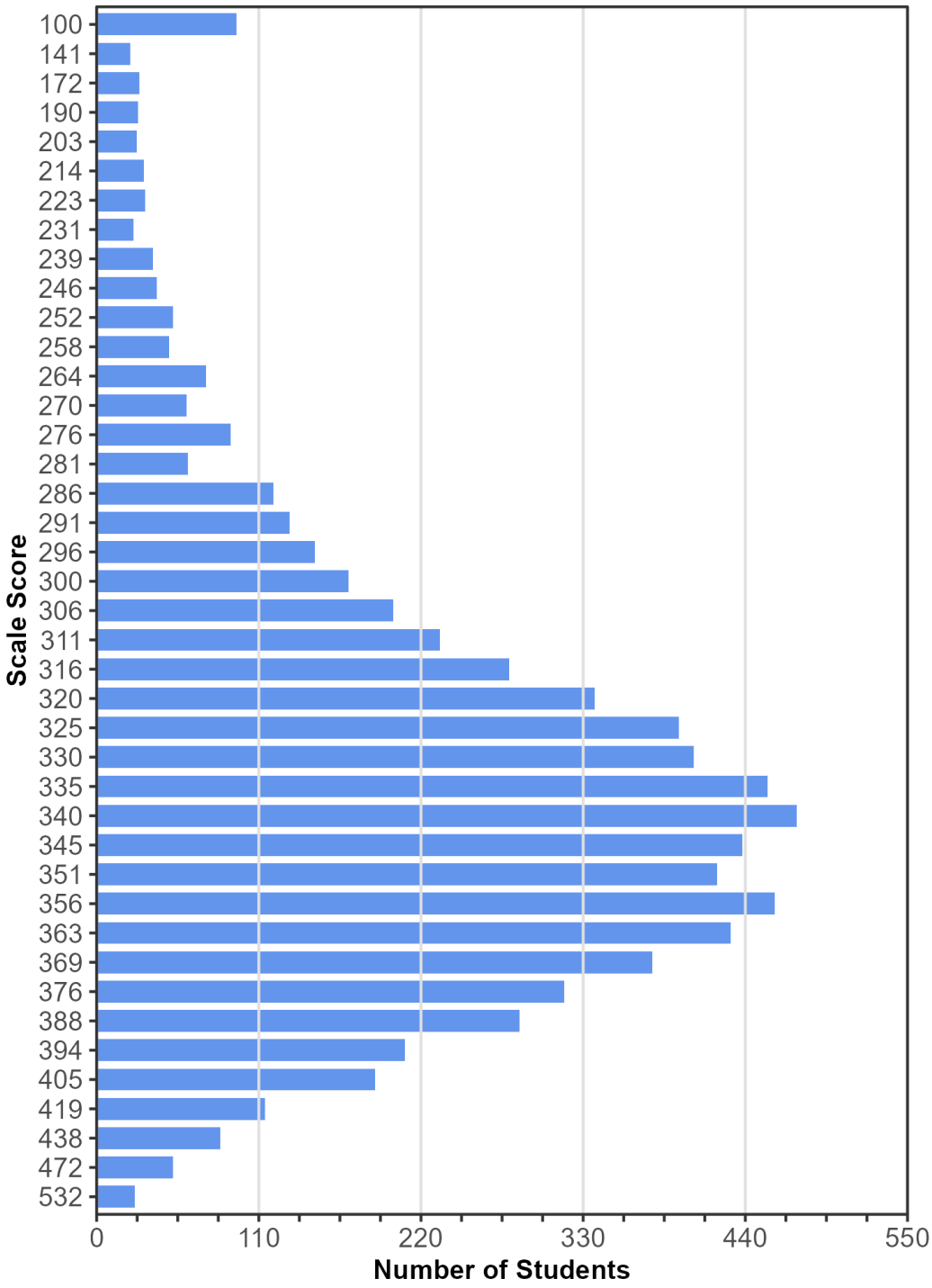
Scale Score	Freq	Cum Freq	Pct	Cum Pct
103	21	21	0.39	0.39
151	2	23	0.04	0.43
176	3	26	0.06	0.48
192	5	31	0.09	0.58
203	4	35	0.07	0.65
212	7	42	0.13	0.78
219	4	46	0.07	0.85
226	6	52	0.11	0.97
233	5	57	0.09	1.06
239	12	69	0.22	1.28
244	10	79	0.19	1.47
250	9	88	0.17	1.63
255	14	102	0.26	1.89
260	8	110	0.15	2.04
265	9	119	0.17	2.21
271	15	134	0.28	2.49
276	23	157	0.43	2.92
280	25	182	0.46	3.38
285	25	207	0.46	3.84
290	30	237	0.56	4.40
295	35	272	0.65	5.05
300	54	326	1.00	6.05
304	62	388	1.15	7.21
309	93	481	1.73	8.93
314	111	592	2.06	10.99
319	152	744	2.82	13.82
324	190	934	3.53	17.34
328	238	1172	4.42	21.76
334	266	1438	4.94	26.70
339	322	1760	5.98	32.68
344	363	2123	6.74	39.42
350	385	2508	7.15	46.57
356	385	2893	7.15	53.72
363	346	3239	6.43	60.15
370	378	3617	7.02	67.17
375	345	3962	6.41	73.57
389	361	4323	6.70	80.28
402	329	4652	6.11	86.39
419	299	4951	5.55	91.94
450	261	5212	4.85	96.79
503	173	5385	3.21	100.00

Figure C.6.6. 2025 STAAR Alternate 2 Grade 8 Mathematics
Frequency Distribution of Scale Scores



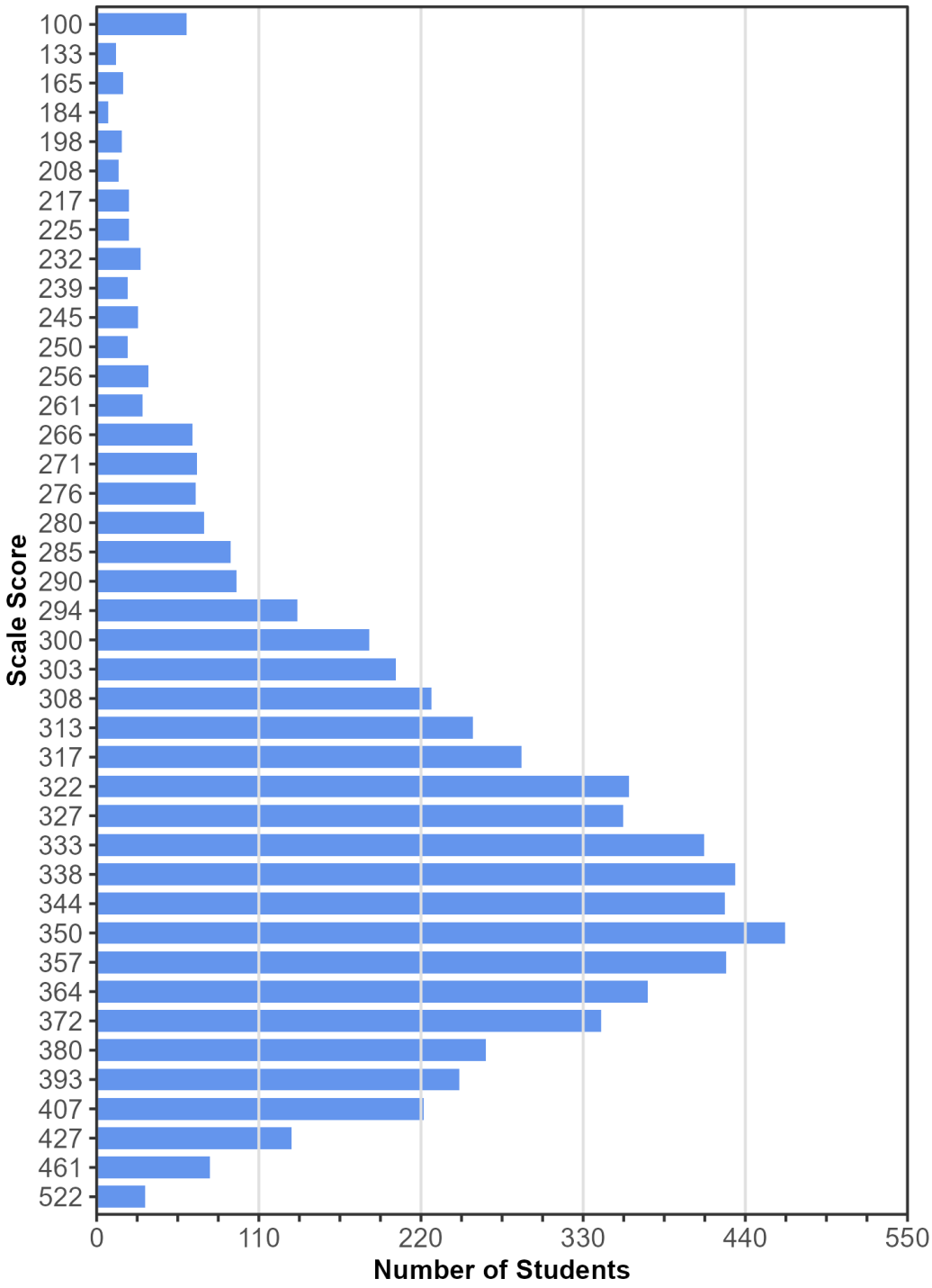
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	22	22	0.45	0.45
139	5	27	0.10	0.55
166	2	29	0.04	0.59
181	3	32	0.06	0.65
193	10	42	0.20	0.86
202	3	45	0.06	0.92
209	2	47	0.04	0.96
217	4	51	0.08	1.04
223	6	57	0.12	1.16
229	4	61	0.08	1.24
235	6	67	0.12	1.37
241	5	72	0.10	1.47
246	10	82	0.20	1.67
252	8	90	0.16	1.84
257	12	102	0.24	2.08
263	10	112	0.20	2.29
268	15	127	0.31	2.59
274	22	149	0.45	3.04
279	29	178	0.59	3.63
284	36	214	0.73	4.37
289	37	251	0.76	5.12
294	46	297	0.94	6.06
300	72	369	1.47	7.53
304	84	453	1.71	9.24
309	117	570	2.39	11.63
314	151	721	3.08	14.71
319	164	885	3.35	18.06
324	217	1102	4.43	22.49
329	255	1357	5.20	27.69
335	299	1656	6.10	33.80
340	320	1976	6.53	40.33
346	353	2329	7.20	47.53
352	385	2714	7.86	55.39
359	349	3063	7.12	62.51
365	381	3444	7.78	70.29
376	298	3742	6.08	76.37
386	312	4054	6.37	82.73
399	288	4342	5.88	88.61
417	221	4563	4.51	93.12
448	195	4758	3.98	97.10
502	142	4900	2.90	100.00

Figure C.6.7. 2025 STAAR Alternate 2 Grade 3 RLA
Frequency Distribution of Scale Scores



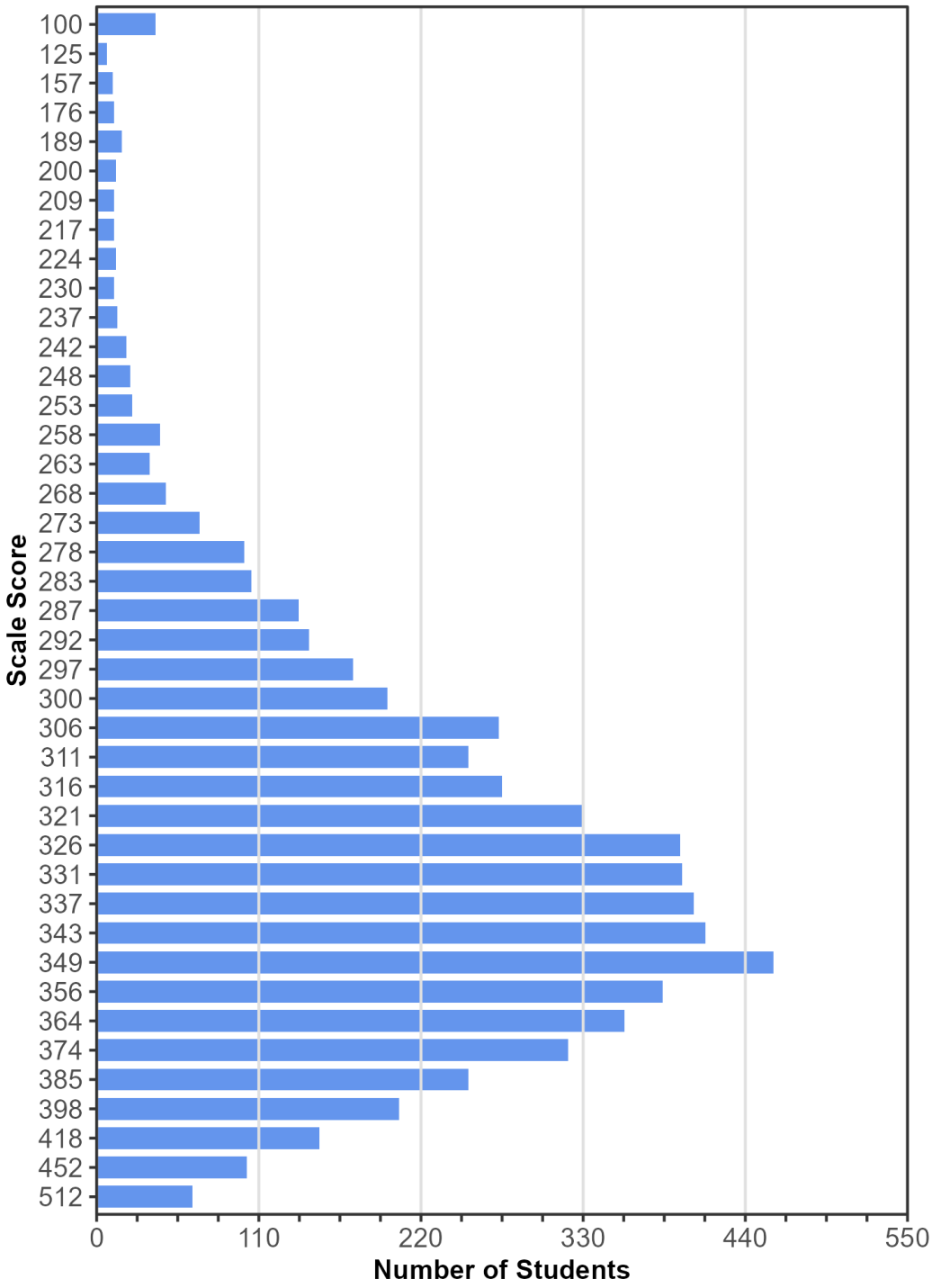
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	95	95	1.26	1.26
141	23	118	0.31	1.57
172	29	147	0.39	1.96
190	28	175	0.37	2.33
203	27	202	0.36	2.69
214	32	234	0.43	3.11
223	33	267	0.44	3.55
231	25	292	0.33	3.89
239	38	330	0.51	4.39
246	41	371	0.55	4.94
252	52	423	0.69	5.63
258	49	472	0.65	6.28
264	74	546	0.98	7.26
270	61	607	0.81	8.08
276	91	698	1.21	9.29
281	62	760	0.82	10.11
286	120	880	1.60	11.71
291	131	1011	1.74	13.45
296	148	1159	1.97	15.42
300	171	1330	2.28	17.70
306	201	1531	2.67	20.37
311	233	1764	3.10	23.47
316	280	2044	3.73	27.20
320	338	2382	4.50	31.69
325	395	2777	5.26	36.95
330	405	3182	5.39	42.34
335	455	3637	6.05	48.39
340	475	4112	6.32	54.71
345	438	4550	5.83	60.54
351	421	4971	5.60	66.14
356	460	5431	6.12	72.26
363	430	5861	5.72	77.98
369	377	6238	5.02	83.00
376	317	6555	4.22	87.21
388	287	6842	3.82	91.03
394	209	7051	2.78	93.81
405	189	7240	2.51	96.33
419	114	7354	1.52	97.84
438	84	7438	1.12	98.96
472	52	7490	0.69	99.65
532	26	7516	0.35	100.00

Figure C.6.8. 2025 STAAR Alternate 2 Grade 4 RLA
Frequency Distribution of Scale Scores



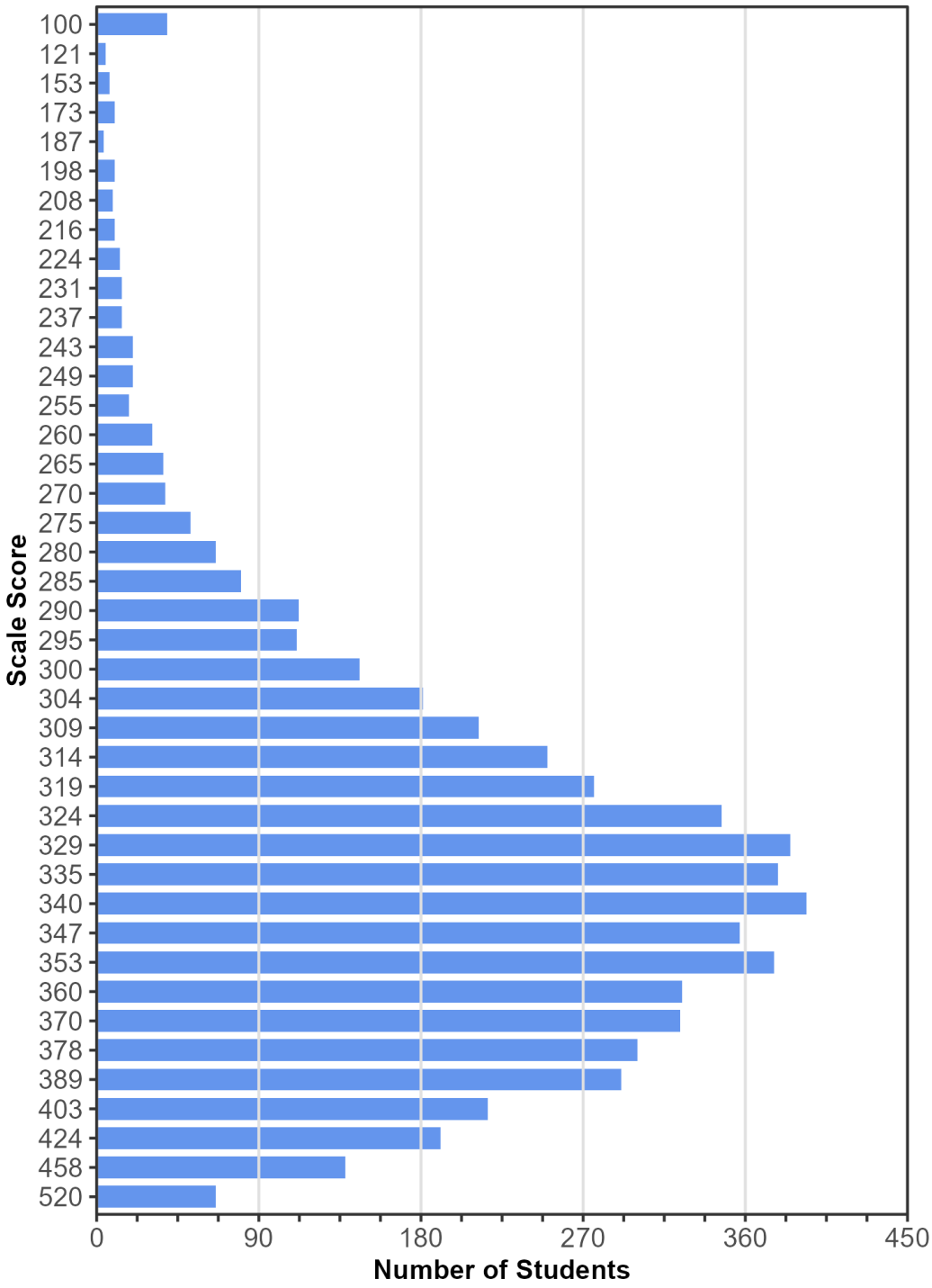
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	61	61	0.91	0.91
133	13	74	0.19	1.11
165	18	92	0.27	1.38
184	8	100	0.12	1.50
198	17	117	0.25	1.75
208	15	132	0.22	1.98
217	22	154	0.33	2.31
225	22	176	0.33	2.64
232	30	206	0.45	3.09
239	21	227	0.31	3.40
245	28	255	0.42	3.82
250	21	276	0.31	4.14
256	35	311	0.52	4.66
261	31	342	0.46	5.13
266	65	407	0.97	6.10
271	68	475	1.02	7.12
276	67	542	1.00	8.13
280	73	615	1.09	9.22
285	91	706	1.36	10.59
290	95	801	1.42	12.01
294	136	937	2.04	14.05
300	185	1122	2.77	16.83
303	203	1325	3.04	19.87
308	227	1552	3.40	23.28
313	255	1807	3.82	27.10
317	288	2095	4.32	31.42
322	361	2456	5.41	36.83
327	357	2813	5.35	42.19
333	412	3225	6.18	48.37
338	433	3658	6.49	54.86
344	426	4084	6.39	61.25
350	467	4551	7.00	68.25
357	427	4978	6.40	74.66
364	374	5352	5.61	80.26
372	342	5694	5.13	85.39
380	264	5958	3.96	89.35
393	246	6204	3.69	93.04
407	222	6426	3.33	96.37
427	132	6558	1.98	98.35
461	77	6635	1.15	99.51
522	33	6668	0.49	100.00

Figure C.6.9. 2025 STAAR Alternate 2 Grade 5 RLA
Frequency Distribution of Scale Scores



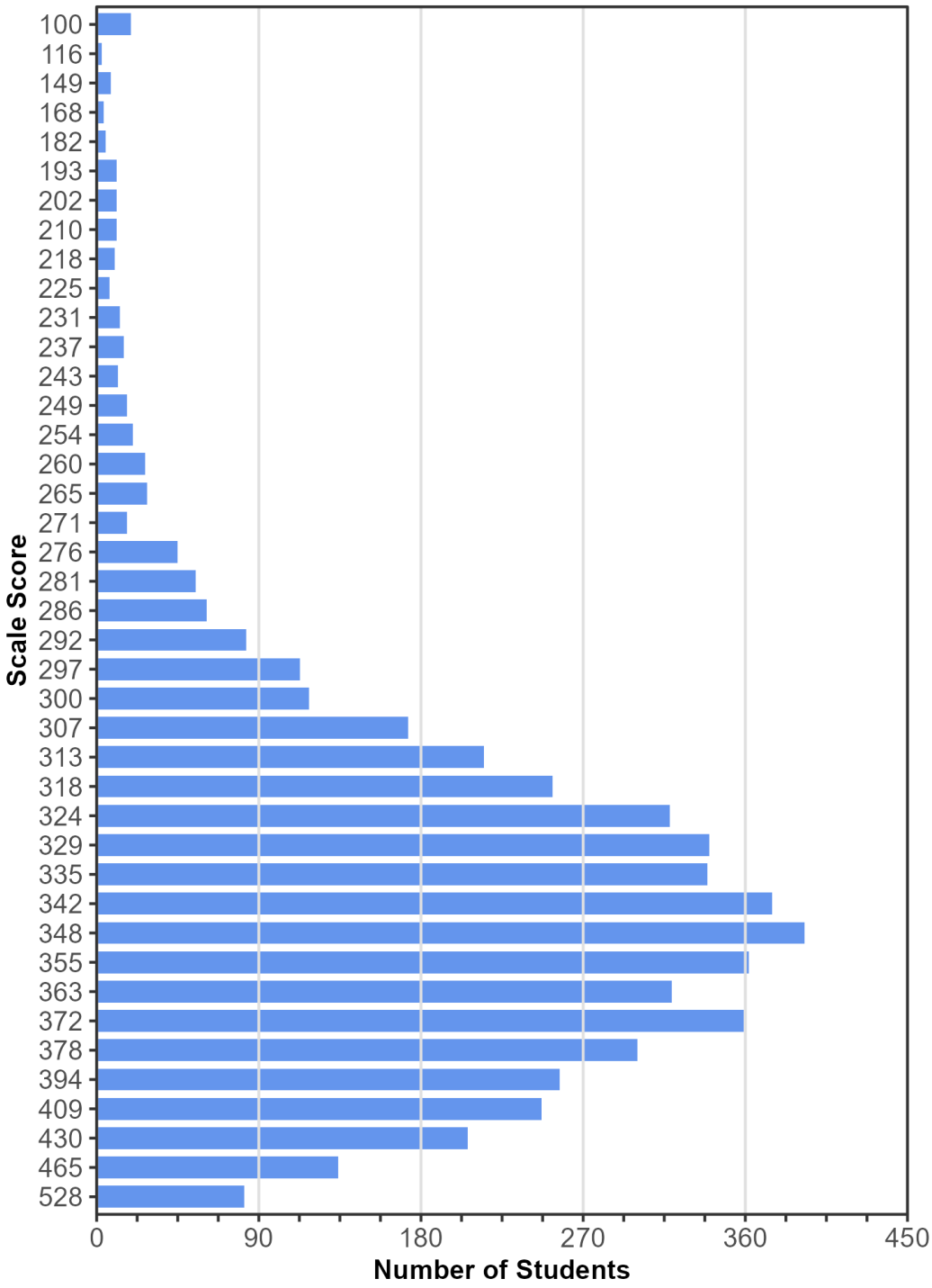
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	40	40	0.63	0.63
125	7	47	0.11	0.74
157	11	58	0.17	0.92
176	12	70	0.19	1.11
189	17	87	0.27	1.38
200	13	100	0.21	1.58
209	12	112	0.19	1.77
217	12	124	0.19	1.96
224	13	137	0.21	2.17
230	12	149	0.19	2.36
237	14	163	0.22	2.58
242	20	183	0.32	2.90
248	23	206	0.36	3.26
253	24	230	0.38	3.64
258	43	273	0.68	4.32
263	36	309	0.57	4.89
268	47	356	0.74	5.63
273	70	426	1.11	6.74
278	100	526	1.58	8.32
283	105	631	1.66	9.98
287	137	768	2.17	12.15
292	144	912	2.28	14.43
297	174	1086	2.75	17.18
300	197	1283	3.12	20.30
306	273	1556	4.32	24.62
311	252	1808	3.99	28.60
316	275	2083	4.35	32.95
321	331	2414	5.24	38.19
326	396	2810	6.26	44.45
331	397	3207	6.28	50.74
337	405	3612	6.41	57.14
343	413	4025	6.53	63.68
349	459	4484	7.26	70.94
356	384	4868	6.07	77.01
364	358	5226	5.66	82.68
374	320	5546	5.06	87.74
385	252	5798	3.99	91.73
398	205	6003	3.24	94.97
418	151	6154	2.39	97.36
452	102	6256	1.61	98.97
512	65	6321	1.03	100.00

Figure C.6.10. 2025 STAAR Alternate 2 Grade 6 RLA
Frequency Distribution of Scale Scores



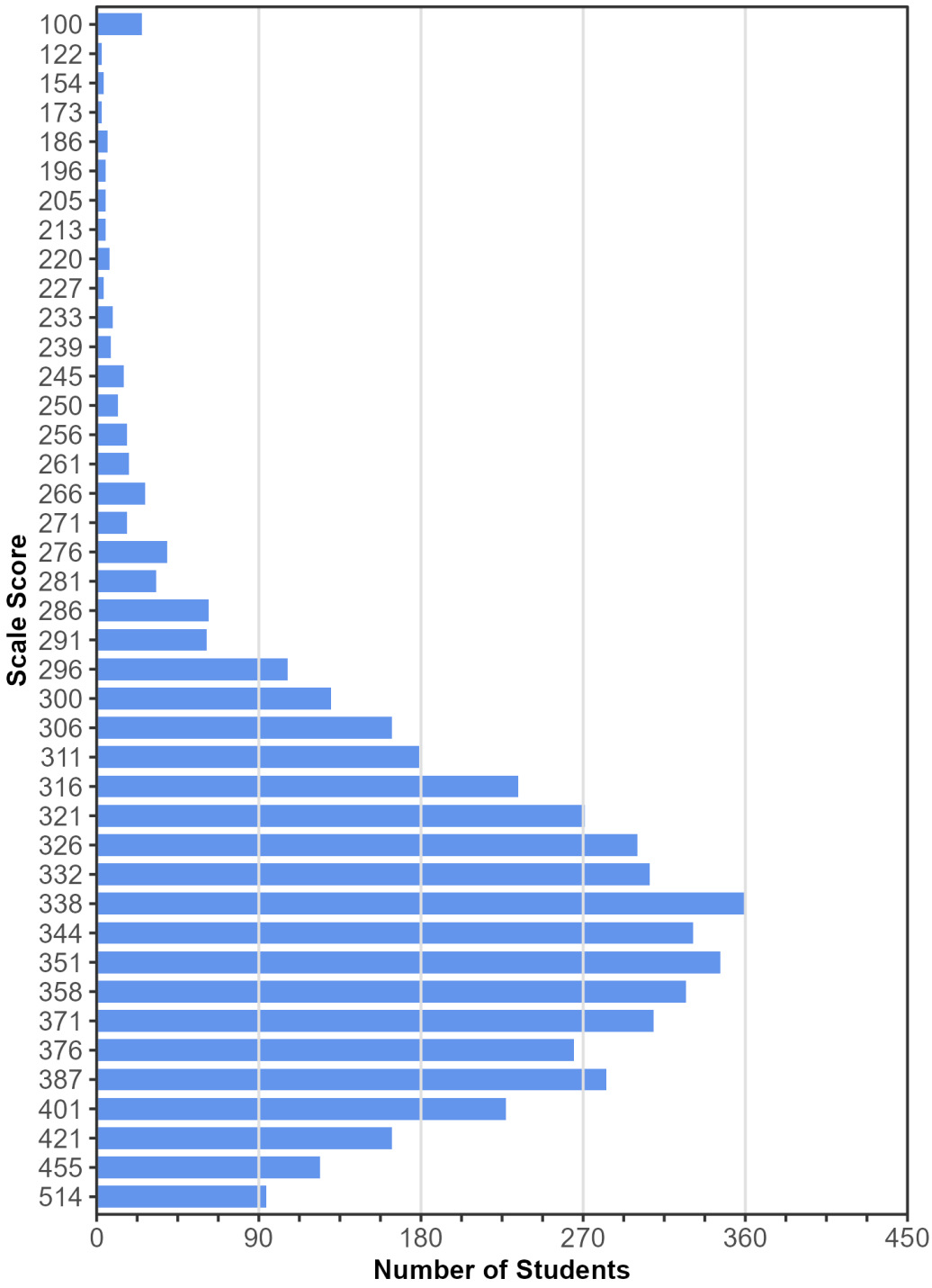
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	39	39	0.66	0.66
121	5	44	0.09	0.75
153	7	51	0.12	0.87
173	10	61	0.17	1.04
187	4	65	0.07	1.11
198	10	75	0.17	1.28
208	9	84	0.15	1.43
216	10	94	0.17	1.60
224	13	107	0.22	1.82
231	14	121	0.24	2.06
237	14	135	0.24	2.30
243	20	155	0.34	2.64
249	20	175	0.34	2.98
255	18	193	0.31	3.29
260	31	224	0.53	3.81
265	37	261	0.63	4.44
270	38	299	0.65	5.09
275	52	351	0.89	5.98
280	66	417	1.12	7.10
285	80	497	1.36	8.46
290	112	609	1.91	10.37
295	111	720	1.89	12.26
300	146	866	2.49	14.74
304	181	1047	3.08	17.82
309	212	1259	3.61	21.43
314	250	1509	4.26	25.69
319	276	1785	4.70	30.39
324	347	2132	5.91	36.30
329	385	2517	6.55	42.85
335	378	2895	6.44	49.28
340	394	3289	6.71	55.99
347	357	3646	6.08	62.07
353	376	4022	6.40	68.47
360	325	4347	5.53	74.00
370	324	4671	5.52	79.52
378	300	4971	5.11	84.63
389	291	5262	4.95	89.58
403	217	5479	3.69	93.28
424	191	5670	3.25	96.53
458	138	5808	2.35	98.88
520	66	5874	1.12	100.00

Figure C.6.11. 2025 STAAR Alternate 2 Grade 7 RLA
Frequency Distribution of Scale Scores



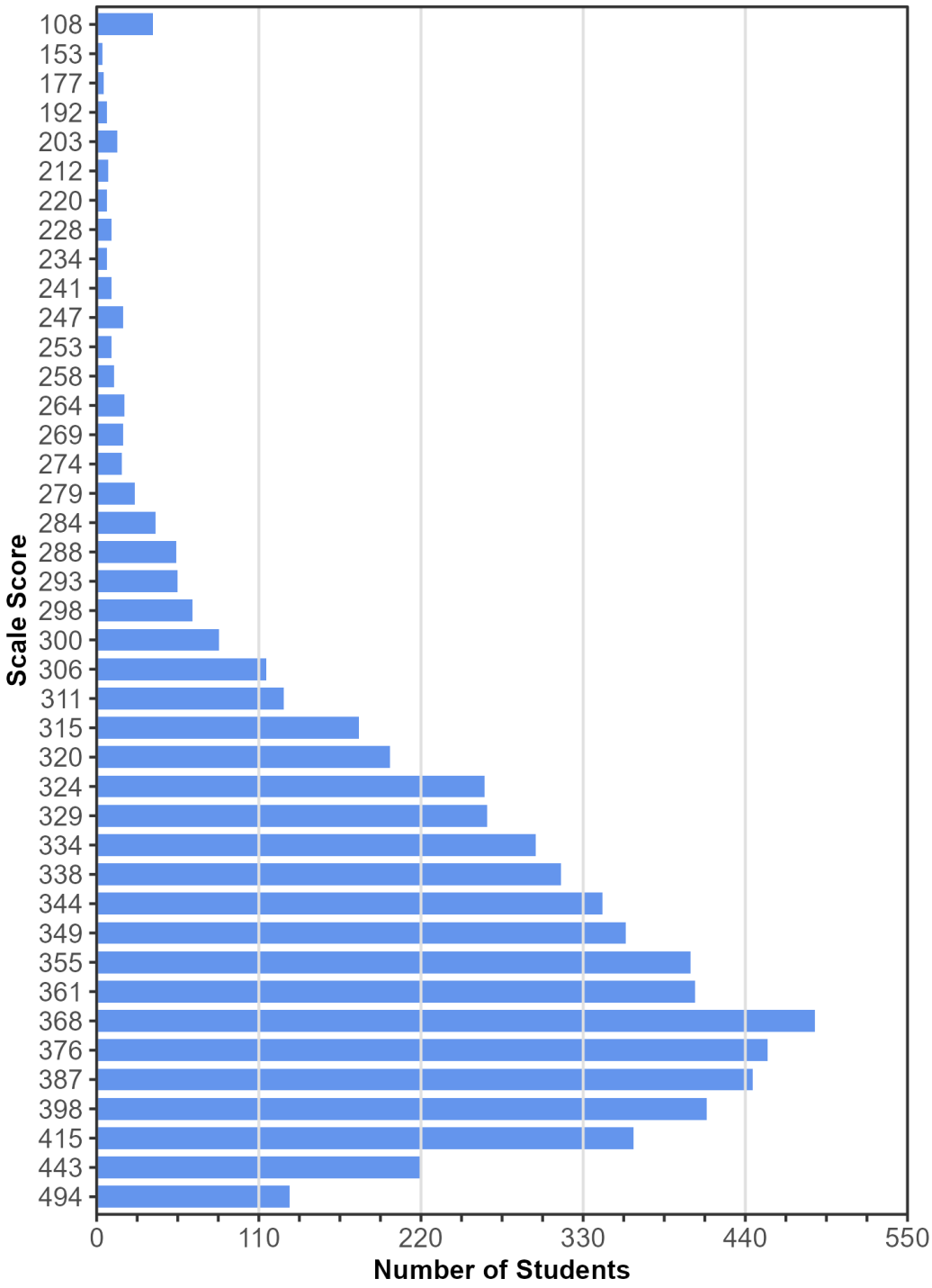
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	19	19	0.35	0.35
116	3	22	0.06	0.41
149	8	30	0.15	0.56
168	4	34	0.07	0.63
182	5	39	0.09	0.72
193	11	50	0.20	0.93
202	11	61	0.20	1.13
210	11	72	0.20	1.34
218	10	82	0.19	1.52
225	7	89	0.13	1.65
231	13	102	0.24	1.89
237	15	117	0.28	2.17
243	12	129	0.22	2.40
249	17	146	0.32	2.71
254	20	166	0.37	3.08
260	27	193	0.50	3.58
265	28	221	0.52	4.10
271	17	238	0.32	4.42
276	45	283	0.84	5.26
281	55	338	1.02	6.28
286	61	399	1.13	7.41
292	83	482	1.54	8.95
297	113	595	2.10	11.05
300	118	713	2.19	13.24
307	173	886	3.21	16.45
313	215	1101	3.99	20.45
318	253	1354	4.70	25.14
324	318	1672	5.91	31.05
329	340	2012	6.31	37.36
335	339	2351	6.30	43.66
342	375	2726	6.96	50.62
348	393	3119	7.30	57.92
355	362	3481	6.72	64.64
363	319	3800	5.92	70.57
372	359	4159	6.67	77.23
378	300	4459	5.57	82.80
394	257	4716	4.77	87.58
409	247	4963	4.59	92.16
430	206	5169	3.83	95.99
465	134	5303	2.49	98.48
528	82	5385	1.52	100.00

Figure C.6.12. 2025 STAAR Alternate 2 Grade 8 RLA
Frequency Distribution of Scale Scores



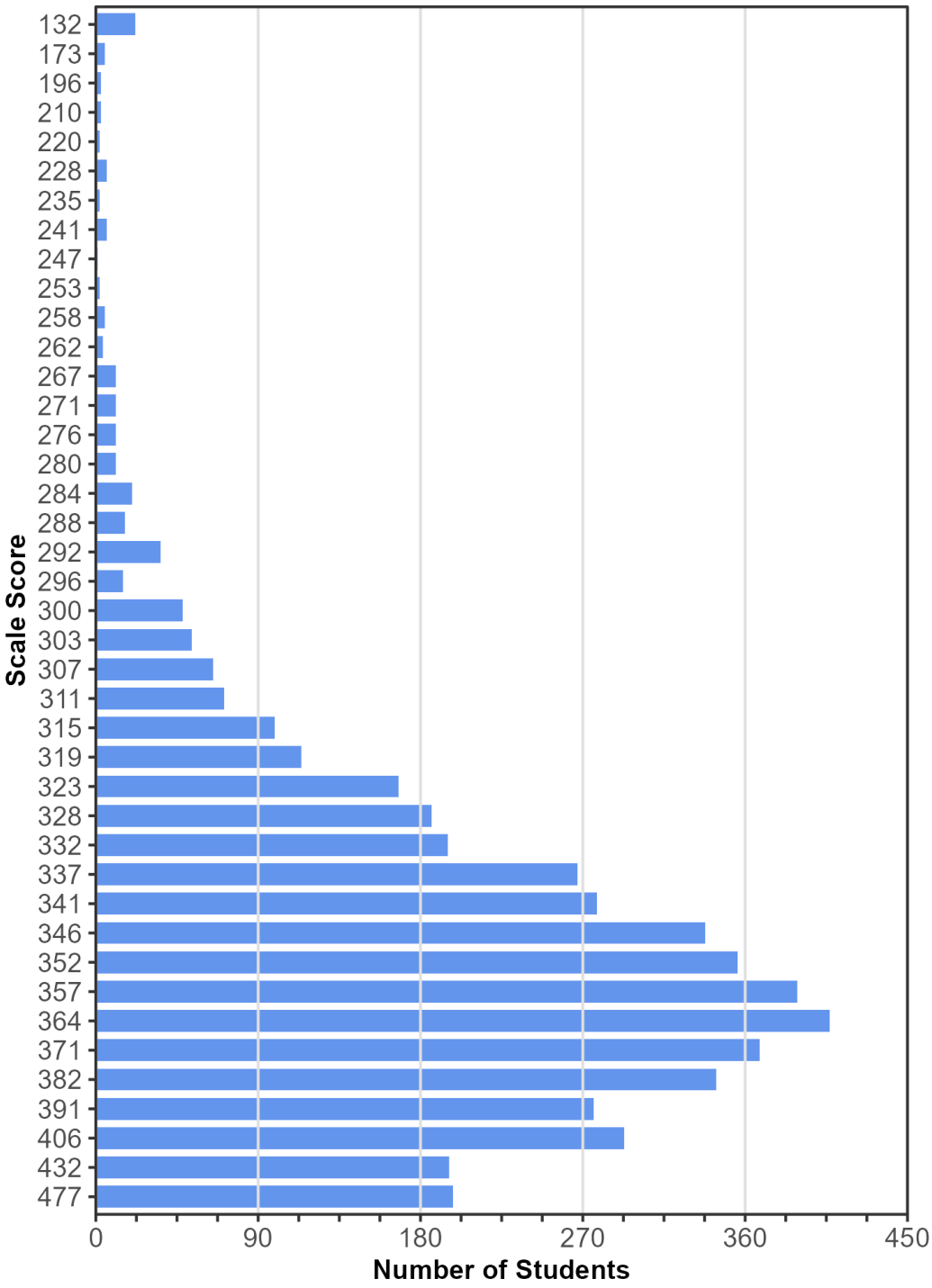
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	25	25	0.51	0.51
122	3	28	0.06	0.57
154	4	32	0.08	0.65
173	3	35	0.06	0.71
186	6	41	0.12	0.84
196	5	46	0.10	0.94
205	5	51	0.10	1.04
213	5	56	0.10	1.14
220	7	63	0.14	1.28
227	4	67	0.08	1.37
233	9	76	0.18	1.55
239	8	84	0.16	1.71
245	15	99	0.31	2.02
250	12	111	0.24	2.26
256	17	128	0.35	2.61
261	18	146	0.37	2.98
266	27	173	0.55	3.53
271	17	190	0.35	3.87
276	39	229	0.79	4.67
281	33	262	0.67	5.34
286	62	324	1.26	6.60
291	61	385	1.24	7.85
296	106	491	2.16	10.01
300	130	621	2.65	12.66
306	164	785	3.34	16.00
311	179	964	3.65	19.65
316	234	1198	4.77	24.42
321	271	1469	5.52	29.94
326	300	1769	6.11	36.06
332	307	2076	6.26	42.32
338	360	2436	7.34	49.65
344	331	2767	6.75	56.40
351	346	3113	7.05	63.45
358	327	3440	6.67	70.12
371	309	3749	6.30	76.42
376	265	4014	5.40	81.82
387	283	4297	5.77	87.59
401	227	4524	4.63	92.21
421	164	4688	3.34	95.56
455	124	4812	2.53	98.08
514	94	4906	1.92	100.00

Figure C.6.13. 2025 STAAR Alternate 2 Grade 5 Science
Frequency Distribution of Scale Scores



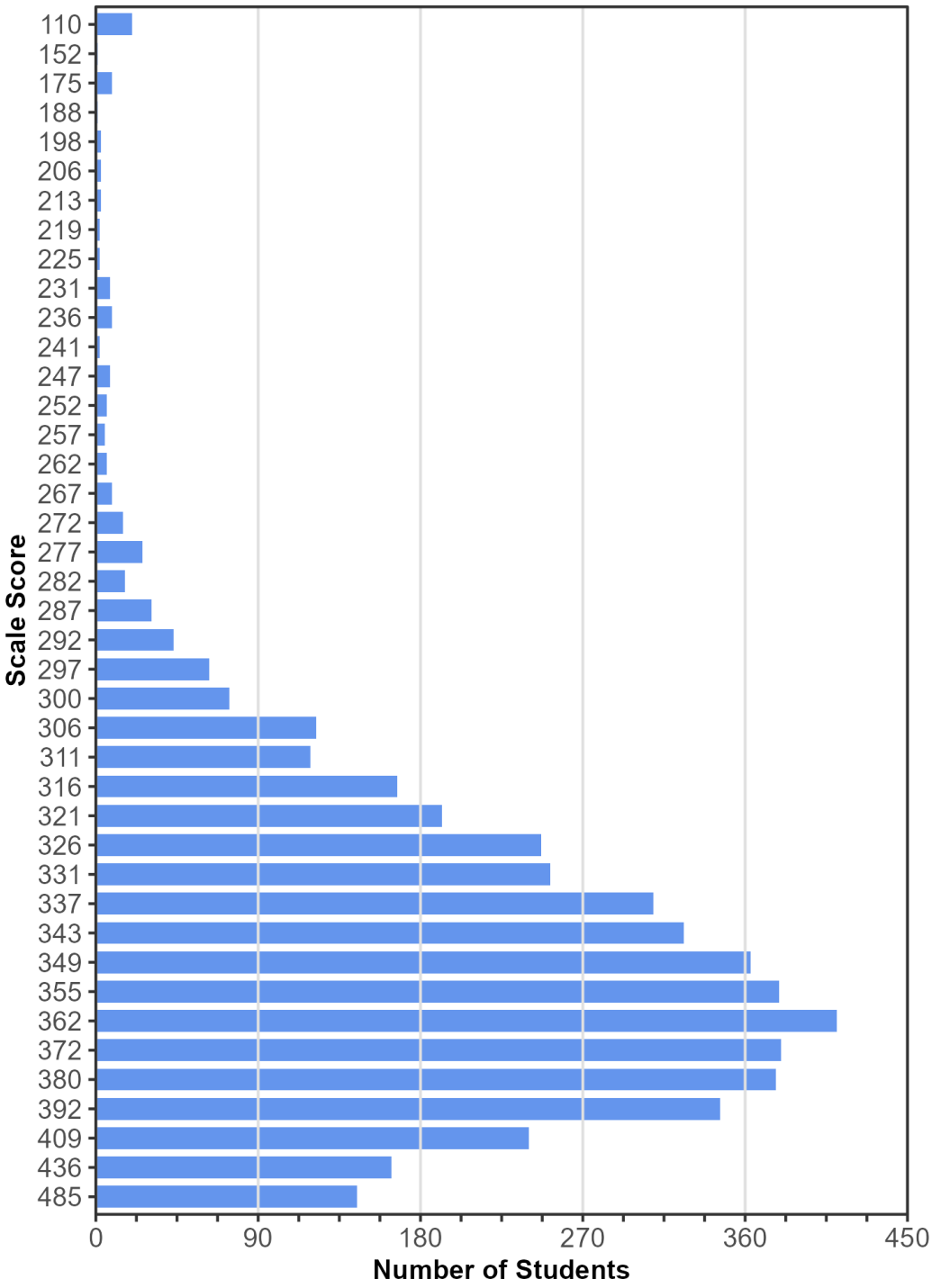
Scale Score	Freq	Cum Freq	Pct	Cum Pct
108	38	38	0.60	0.60
153	4	42	0.06	0.67
177	5	47	0.08	0.74
192	7	54	0.11	0.86
203	14	68	0.22	1.08
212	8	76	0.13	1.20
220	7	83	0.11	1.31
228	10	93	0.16	1.47
234	7	100	0.11	1.58
241	10	110	0.16	1.74
247	18	128	0.29	2.03
253	10	138	0.16	2.19
258	12	150	0.19	2.38
264	19	169	0.30	2.68
269	18	187	0.29	2.96
274	17	204	0.27	3.23
279	26	230	0.41	3.64
284	40	270	0.63	4.28
288	54	324	0.86	5.13
293	55	379	0.87	6.00
298	65	444	1.03	7.03
300	83	527	1.31	8.35
306	115	642	1.82	10.17
311	127	769	2.01	12.18
315	178	947	2.82	15.00
320	199	1146	3.15	18.15
324	263	1409	4.16	22.31
329	265	1674	4.20	26.51
334	298	1972	4.72	31.23
338	315	2287	4.99	36.22
344	343	2630	5.43	41.65
349	359	2989	5.68	47.33
355	403	3392	6.38	53.71
361	406	3798	6.43	60.14
368	487	4285	7.71	67.85
376	455	4740	7.21	75.06
387	445	5185	7.05	82.11
398	414	5599	6.56	88.66
415	364	5963	5.76	94.43
443	221	6184	3.50	97.93
494	131	6315	2.07	100.00

Figure C.6.14. 2025 STAAR Alternate 2 Grade 8 Science Frequency Distribution of Scale Scores



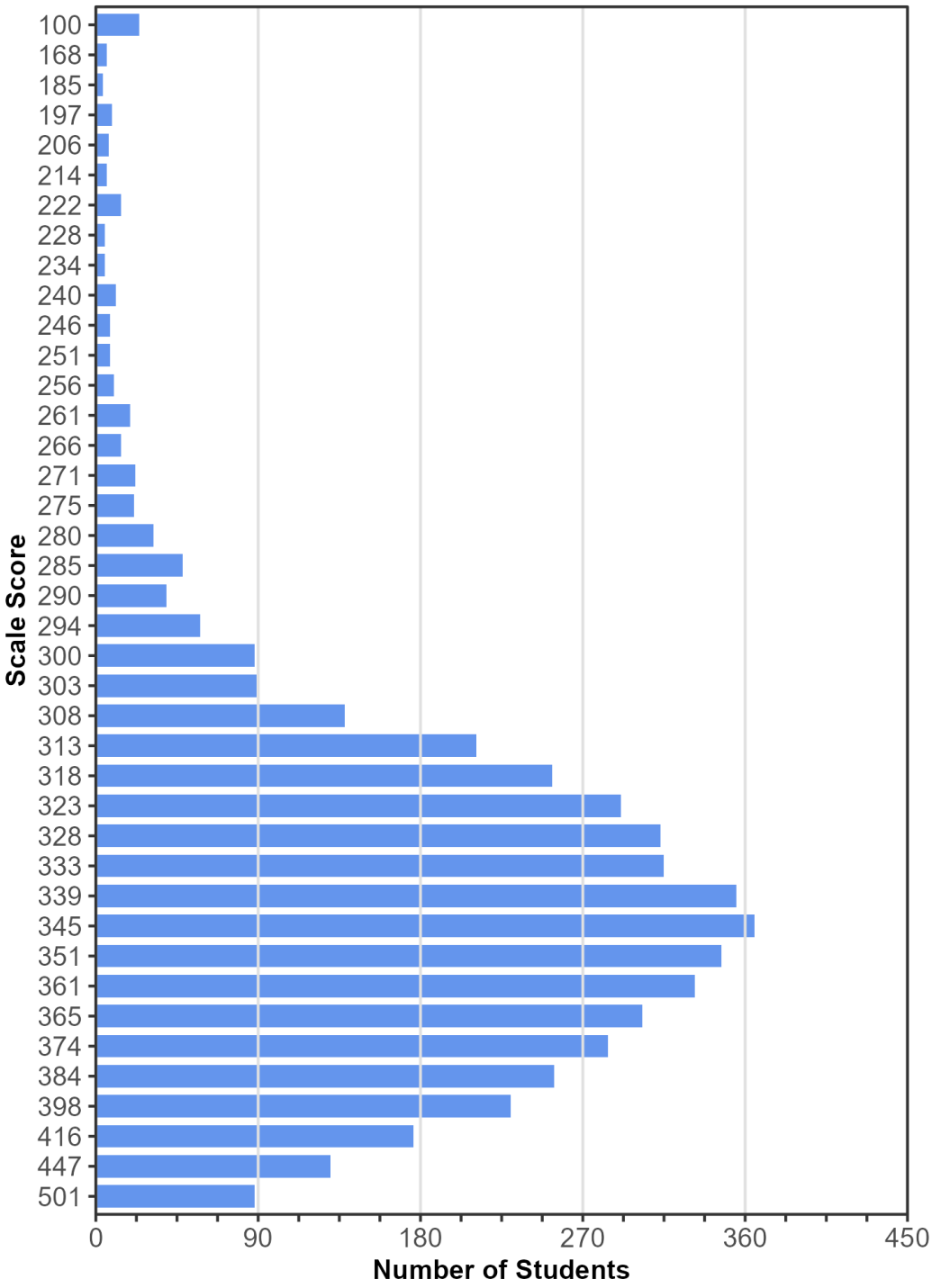
Scale Score	Freq	Cum Freq	Pct	Cum Pct
132	22	22	0.45	0.45
173	5	27	0.10	0.55
196	3	30	0.06	0.61
210	3	33	0.06	0.67
220	2	35	0.04	0.71
228	6	41	0.12	0.84
235	2	43	0.04	0.88
241	6	49	0.12	1.00
247	1	50	0.02	1.02
253	2	52	0.04	1.06
258	5	57	0.10	1.16
262	4	61	0.08	1.24
267	11	72	0.22	1.47
271	11	83	0.22	1.69
276	11	94	0.22	1.92
280	11	105	0.22	2.14
284	20	125	0.41	2.55
288	16	141	0.33	2.88
292	36	177	0.73	3.61
296	15	192	0.31	3.92
300	48	240	0.98	4.90
303	53	293	1.08	5.98
307	65	358	1.33	7.30
311	71	429	1.45	8.75
315	99	528	2.02	10.77
319	114	642	2.33	13.10
323	168	810	3.43	16.53
328	186	996	3.80	20.32
332	195	1191	3.98	24.30
337	267	1458	5.45	29.75
341	278	1736	5.67	35.42
346	338	2074	6.90	42.32
352	356	2430	7.26	49.58
357	389	2819	7.94	57.52
364	407	3226	8.30	65.82
371	368	3594	7.51	73.33
382	344	3938	7.02	80.35
391	276	4214	5.63	85.98
406	293	4507	5.98	91.96
432	196	4703	4.00	95.96
477	198	4901	4.04	100.00

Figure C.6.15. 2025 STAAR Alternate 2 Grade 8 Social Studies
Frequency Distribution of Scale Scores



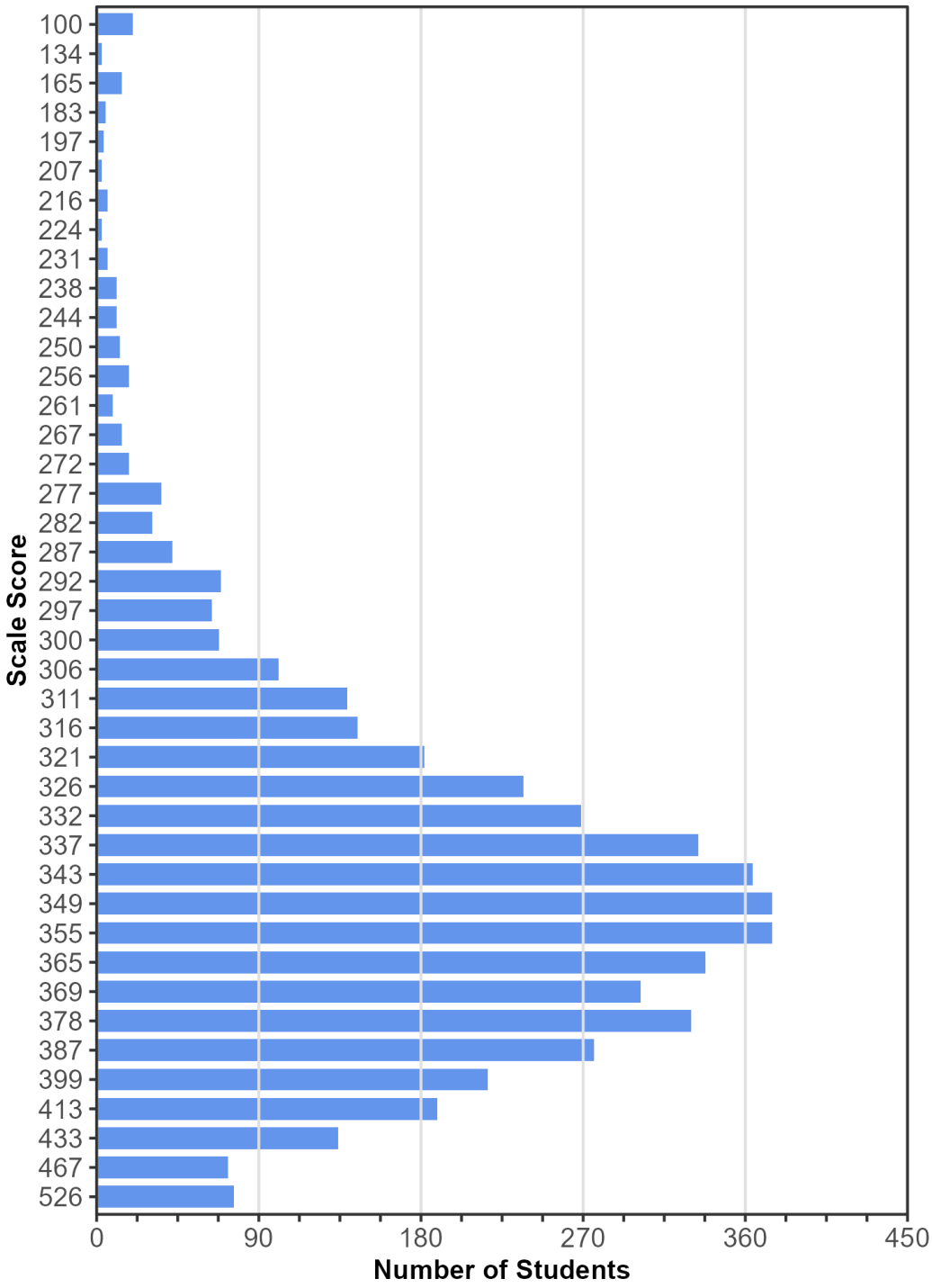
Scale Score	Freq	Cum Freq	Pct	Cum Pct
110	20	20	0.41	0.41
152	1	21	0.02	0.43
175	9	30	0.18	0.61
188	1	31	0.02	0.63
198	3	34	0.06	0.69
206	3	37	0.06	0.75
213	3	40	0.06	0.82
219	2	42	0.04	0.86
225	2	44	0.04	0.90
231	8	52	0.16	1.06
236	9	61	0.18	1.24
241	2	63	0.04	1.28
247	8	71	0.16	1.45
252	6	77	0.12	1.57
257	5	82	0.10	1.67
262	6	88	0.12	1.79
267	9	97	0.18	1.98
272	15	112	0.31	2.28
277	26	138	0.53	2.81
282	16	154	0.33	3.14
287	31	185	0.63	3.77
292	43	228	0.88	4.65
297	63	291	1.28	5.93
300	74	365	1.51	7.44
306	122	487	2.49	9.93
311	119	606	2.43	12.36
316	167	773	3.41	15.76
321	192	965	3.92	19.68
326	247	1212	5.04	24.71
331	252	1464	5.14	29.85
337	309	1773	6.30	36.15
343	326	2099	6.65	42.80
349	363	2462	7.40	50.20
355	379	2841	7.73	57.93
362	411	3252	8.38	66.31
372	380	3632	7.75	74.06
380	377	4009	7.69	81.75
392	346	4355	7.06	88.81
409	240	4595	4.89	93.70
436	164	4759	3.34	97.04
485	145	4904	2.96	100.00

Figure C.6.16. 2025 STAAR Alternate 2 Algebra I
Frequency Distribution of Scale Scores



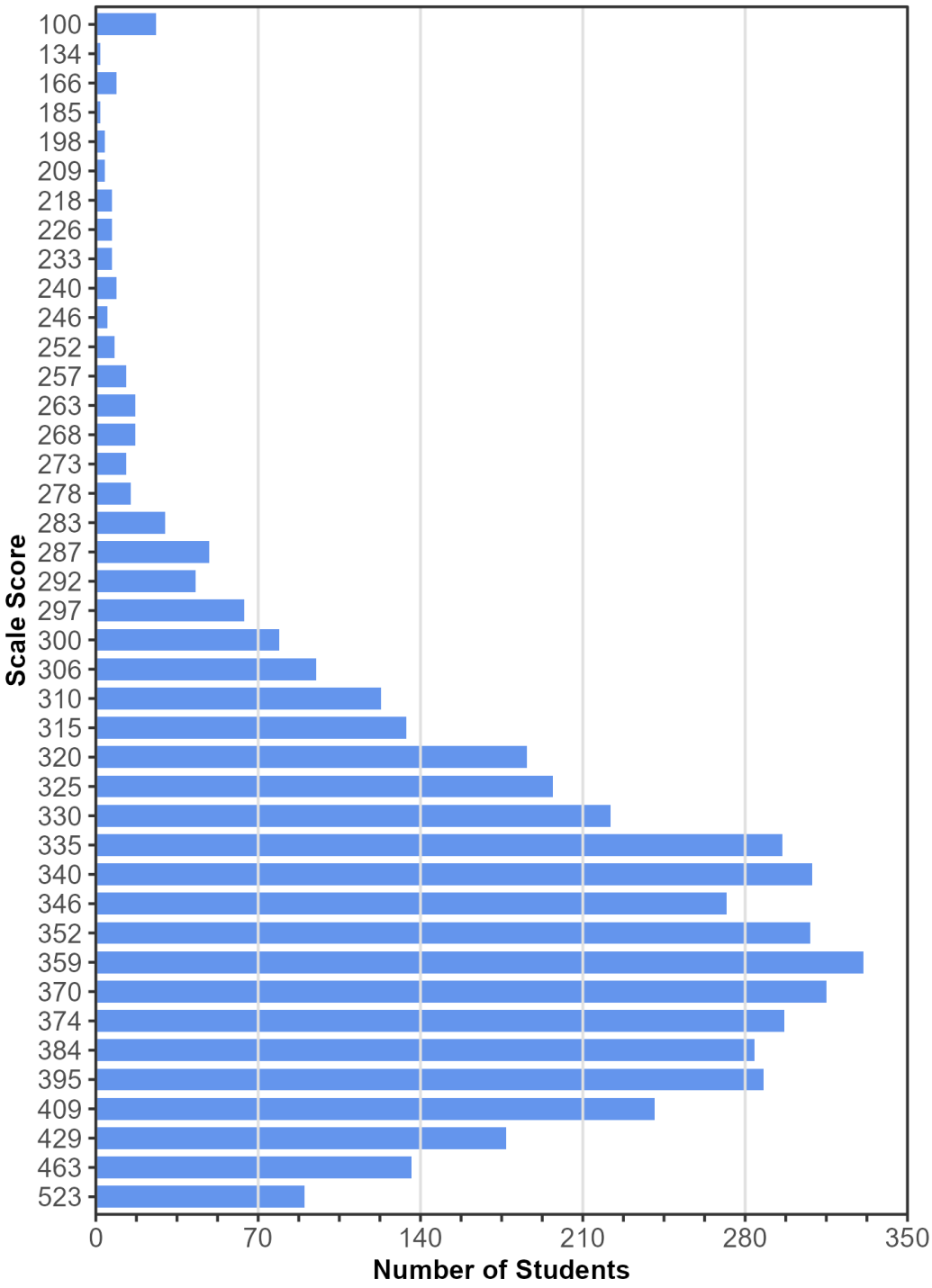
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	24	24	0.49	0.49
168	6	30	0.12	0.61
185	4	34	0.08	0.69
197	9	43	0.18	0.87
206	7	50	0.14	1.01
214	6	56	0.12	1.14
222	14	70	0.28	1.42
228	5	75	0.10	1.52
234	5	80	0.10	1.62
240	11	91	0.22	1.85
246	8	99	0.16	2.01
251	8	107	0.16	2.17
256	10	117	0.20	2.37
261	19	136	0.39	2.76
266	14	150	0.28	3.04
271	22	172	0.45	3.49
275	21	193	0.43	3.91
280	32	225	0.65	4.56
285	48	273	0.97	5.54
290	39	312	0.79	6.33
294	58	370	1.18	7.50
300	88	458	1.78	9.29
303	89	547	1.80	11.09
308	138	685	2.80	13.89
313	211	896	4.28	18.17
318	253	1149	5.13	23.30
323	291	1440	5.90	29.20
328	313	1753	6.35	35.54
333	315	2068	6.39	41.93
339	355	2423	7.20	49.13
345	365	2788	7.40	56.53
351	347	3135	7.04	63.56
361	332	3467	6.73	70.30
365	303	3770	6.14	76.44
374	284	4054	5.76	82.20
384	254	4308	5.15	87.35
398	230	4538	4.66	92.01
416	176	4714	3.57	95.58
447	130	4844	2.64	98.22
501	88	4932	1.78	100.00

Figure C.6.17. 2025 STAAR Alternate 2 English I
Frequency Distribution of Scale Scores



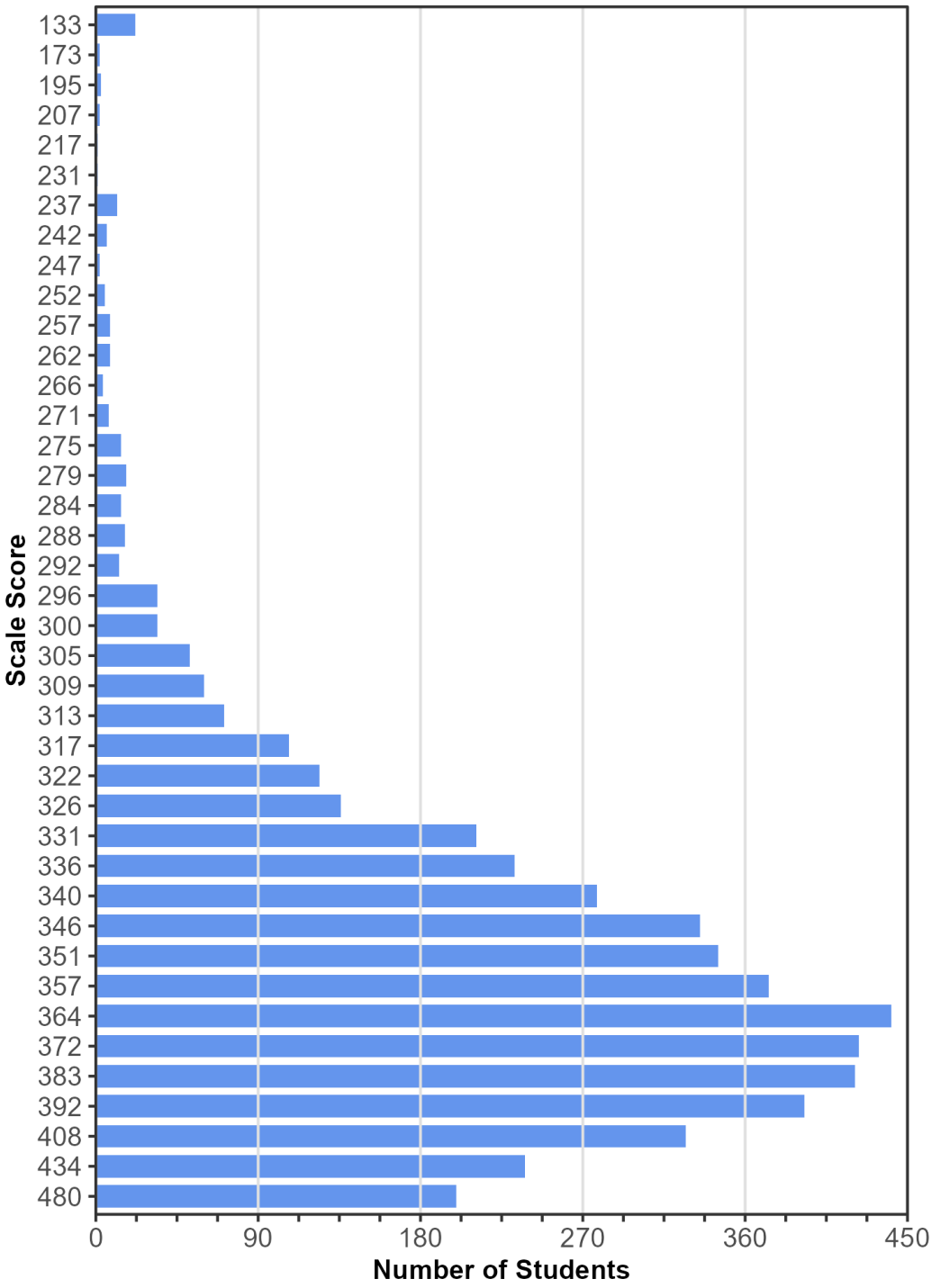
Scale Score	Freq	Cum Freq	Pct	Cum Pct
100	20	20	0.41	0.41
134	3	23	0.06	0.47
165	14	37	0.28	0.75
183	5	42	0.10	0.85
197	4	46	0.08	0.93
207	3	49	0.06	1.00
216	6	55	0.12	1.12
224	3	58	0.06	1.18
231	6	64	0.12	1.30
238	11	75	0.22	1.52
244	11	86	0.22	1.75
250	13	99	0.26	2.01
256	18	117	0.37	2.38
261	9	126	0.18	2.56
267	14	140	0.28	2.84
272	18	158	0.37	3.21
277	36	194	0.73	3.94
282	31	225	0.63	4.57
287	42	267	0.85	5.42
292	69	336	1.40	6.82
297	64	400	1.30	8.12
300	68	468	1.38	9.50
306	101	569	2.05	11.56
311	139	708	2.82	14.38
316	145	853	2.94	17.32
321	182	1035	3.70	21.02
326	237	1272	4.81	25.83
332	269	1541	5.46	31.30
337	334	1875	6.78	38.08
343	364	2239	7.39	45.47
349	375	2614	7.62	53.09
355	375	2989	7.62	60.70
365	338	3327	6.86	67.57
369	302	3629	6.13	73.70
378	330	3959	6.70	80.40
387	276	4235	5.61	86.01
399	217	4452	4.41	90.41
413	189	4641	3.84	94.25
433	134	4775	2.72	96.97
467	73	4848	1.48	98.46
526	76	4924	1.54	100.00

Figure C.6.18. 2025 STAAR Alternate 2 English II
Frequency Distribution of Scale Scores



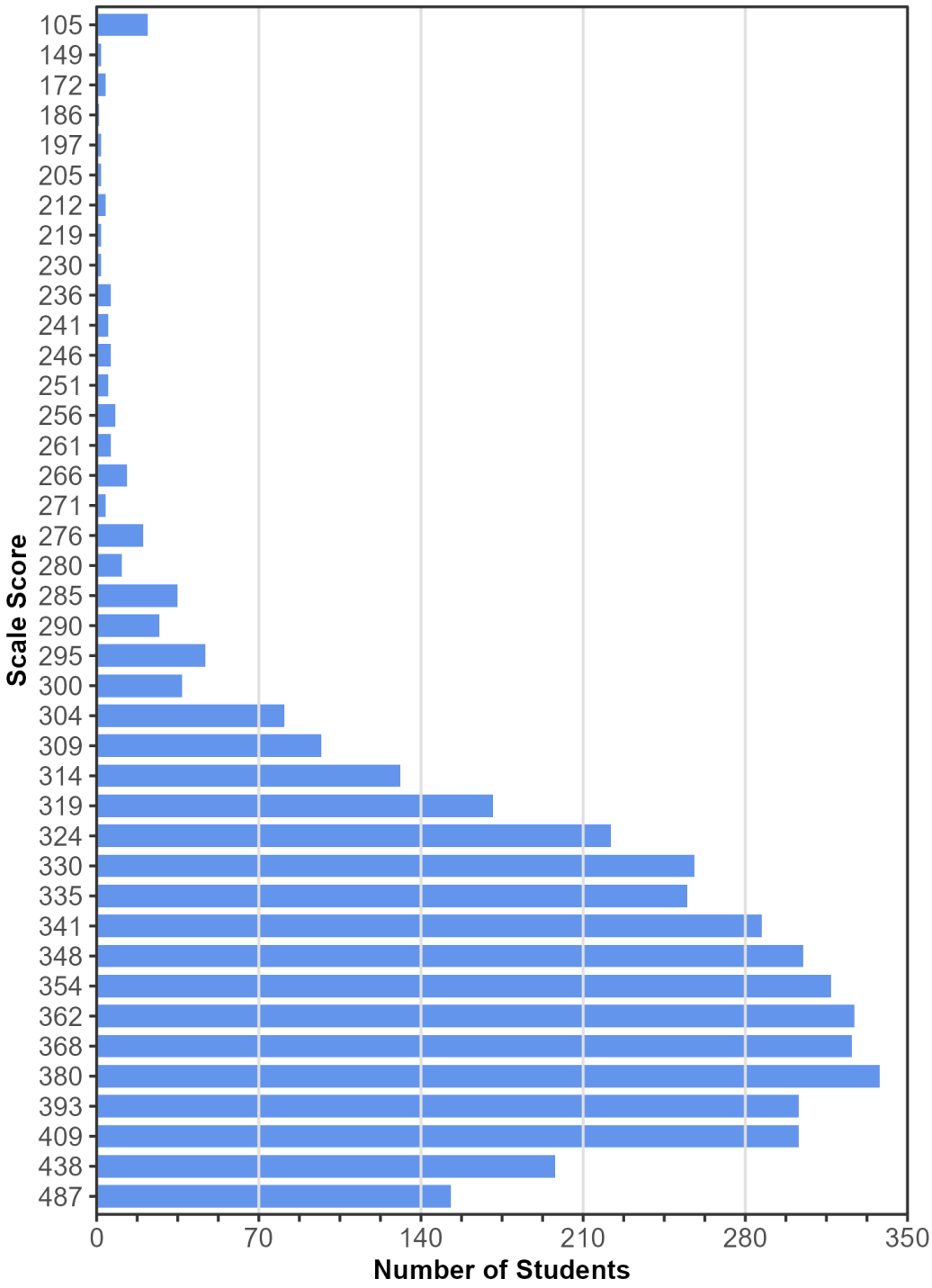
	Freq	Cum Freq	Pct	Cum Pct
100	26	26	0.55	0.55
134	2	28	0.04	0.59
166	9	37	0.19	0.78
185	2	39	0.04	0.82
198	4	43	0.08	0.91
209	4	47	0.08	0.99
218	7	54	0.15	1.14
226	7	61	0.15	1.29
233	7	68	0.15	1.44
240	9	77	0.19	1.63
246	5	82	0.11	1.73
252	8	90	0.17	1.90
257	13	103	0.27	2.18
263	17	120	0.36	2.54
268	17	137	0.36	2.90
273	13	150	0.27	3.17
278	15	165	0.32	3.49
283	30	195	0.63	4.12
287	49	244	1.04	5.16
292	43	287	0.91	6.07
297	64	351	1.35	7.42
300	79	430	1.67	9.09
306	95	525	2.01	11.10
310	123	648	2.60	13.70
315	134	782	2.83	16.53
320	186	968	3.93	20.46
325	197	1165	4.16	24.62
330	222	1387	4.69	29.32
335	296	1683	6.26	35.57
340	309	1992	6.53	42.11
346	272	2264	5.75	47.85
352	308	2572	6.51	54.36
359	331	2903	7.00	61.36
370	315	3218	6.66	68.02
374	297	3515	6.28	74.30
384	284	3799	6.00	80.30
395	288	4087	6.09	86.39
409	241	4328	5.09	91.48
429	177	4505	3.74	95.22
463	136	4641	2.87	98.10
523	90	4731	1.90	100.00

Figure C.6.19. 2025 STAAR Alternate 2 Biology
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
133	22	22	0.44	0.44
173	2	24	0.04	0.48
195	3	27	0.06	0.54
207	2	29	0.04	0.58
217	1	30	0.02	0.60
231	1	31	0.02	0.62
237	12	43	0.24	0.86
242	6	49	0.12	0.98
247	2	51	0.04	1.02
252	5	56	0.10	1.12
257	8	64	0.16	1.28
262	8	72	0.16	1.44
266	4	76	0.08	1.52
271	7	83	0.14	1.66
275	14	97	0.28	1.94
279	17	114	0.34	2.28
284	14	128	0.28	2.56
288	16	144	0.32	2.88
292	13	157	0.26	3.15
296	34	191	0.68	3.83
300	34	225	0.68	4.51
305	52	277	1.04	5.55
309	60	337	1.20	6.75
313	71	408	1.42	8.17
317	107	515	2.14	10.32
322	124	639	2.48	12.80
326	136	775	2.72	15.52
331	211	986	4.23	19.75
336	232	1218	4.65	24.40
340	278	1496	5.57	29.97
346	335	1831	6.71	36.68
351	345	2176	6.91	43.59
357	373	2549	7.47	51.06
364	441	2990	8.83	59.90
372	423	3413	8.47	68.37
383	421	3834	8.43	76.80
392	393	4227	7.87	84.68
408	327	4554	6.55	91.23
434	238	4792	4.77	95.99
480	200	4992	4.01	100.00

Figure C.6.20. 2025 STAAR Alternate 2 U.S. History
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
105	22	22	0.51	0.51
149	2	24	0.05	0.55
172	4	28	0.09	0.64
186	1	29	0.02	0.67
197	2	31	0.05	0.71
205	2	33	0.05	0.76
212	4	37	0.09	0.85
219	2	39	0.05	0.90
230	2	41	0.05	0.94
236	6	47	0.14	1.08
241	5	52	0.12	1.20
246	6	58	0.14	1.34
251	5	63	0.12	1.45
256	8	71	0.18	1.63
261	6	77	0.14	1.77
266	13	90	0.30	2.07
271	4	94	0.09	2.16
276	20	114	0.46	2.62
280	11	125	0.25	2.88
285	35	160	0.81	3.68
290	27	187	0.62	4.31
295	47	234	1.08	5.39
300	37	271	0.85	6.24
304	81	352	1.87	8.10
309	97	449	2.23	10.34
314	131	580	3.02	13.35
319	171	751	3.94	17.29
324	222	973	5.11	22.40
330	258	1231	5.94	28.34
335	255	1486	5.87	34.22
341	287	1773	6.61	40.82
348	305	2078	7.02	47.85
354	317	2395	7.30	55.15
362	327	2722	7.53	62.68
368	326	3048	7.51	70.18
380	338	3386	7.78	77.96
393	303	3689	6.98	84.94
409	303	3992	6.98	91.92
438	198	4190	4.56	96.48
487	153	4343	3.52	100.00

TECHNICAL DIGEST 2024–2025

Appendix D

TELPAS Statistical Tables and Figures

List of Tables

Table D.1.1. 2025 TELPAS Classification Consistency and Accuracy

Table D.2.1. 2025 TELPAS Scale Score Correlations

Table D.3.1. 2025 TELPAS Listening Grades 2–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.2. 2025 TELPAS Speaking Grades 2–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.3. 2025 TELPAS Reading Grades 2–5 Conditional Standard Error of Measurement for Scale Scores

Table D.3.4. 2025 TELPAS Reading Grades 6–12 Conditional Standard Error of Measurement for Scale Scores

Table D.3.5. 2025 TELPAS Writing Grades 2–5 Conditional Standard Error of Measurement for Scale Scores

Table D.3.6. 2025 TELPAS Writing Grades 6–12 Conditional Standard Error of Measurement for Scale Scores

Table D.4.1. 2025 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.2. 2025 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.3. 2025 TELPAS Listening and Speaking Grades 2–3 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.4. 2025 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.5. 2025 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.6. 2025 TELPAS Listening and Speaking Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.7. 2025 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.8. 2025 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.9. 2025 TELPAS Listening and Speaking Grades 6–8 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.10. 2025 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.11. 2025 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.12. 2025 TELPAS Listening and Speaking Grades 9–12 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.13. 2025 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.14. 2025 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.15. 2025 TELPAS Reading and Writing Grade 2 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.16. 2025 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.17. 2025 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.18. 2025 TELPAS Reading and Writing Grade 3 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.19. 2025 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.20. 2025 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.21. 2025 TELPAS Reading and Writing Grades 4–5 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.22. 2025 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.23. 2025 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.24. 2025 TELPAS Reading and Writing Grades 6–7 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.25. 2025 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.26. 2025 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.27. 2025 TELPAS Reading and Writing Grades 8–9 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.4.28. 2025 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Total Group

Table D.4.29. 2025 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Female

Table D.4.30. 2025 TELPAS Reading and Writing Grades 10–12 Mean P-Values, Raw Score Summary, and Reliability Male

Table D.5.1. 2025 TELPAS Listening Scale Score Descriptive Statistics

Table D.5.2. 2025 TELPAS Speaking Scale Score Descriptive Statistics

Table D.5.3. 2025 TELPAS Reading Scale Score Descriptive Statistics

Table D.5.4. 2025 TELPAS Writing 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. 2025 TELPAS Grade 2 Listening Frequency Distribution of Scale Scores

Figure D.5.2. 2025 TELPAS Grade 3 Listening Frequency Distribution of Scale Scores

Figure D.5.3. 2025 TELPAS Grade 4 Listening Frequency Distribution of Scale Scores

Figure D.5.4. 2025 TELPAS Grade 5 Listening Frequency Distribution of Scale Scores

Figure D.5.5. 2025 TELPAS Grade 6 Listening Frequency Distribution of Scale Scores

Figure D.5.6. 2025 TELPAS Grade 7 Listening Frequency Distribution of Scale Scores

Figure D.5.7. 2025 TELPAS Grade 8 Listening Frequency Distribution of Scale Scores

Figure D.5.8. 2025 TELPAS Grade 9 Listening Frequency Distribution of Scale Scores

Figure D.5.9. 2025 TELPAS Grade 10 Listening Frequency Distribution of Scale Scores

Figure D.5.10. 2025 TELPAS Grade 11 Listening Frequency Distribution of Scale Scores

Figure D.5.11. 2025 TELPAS Grade 12 Listening Frequency Distribution of Scale Scores

Figure D.5.12. 2025 TELPAS Grade 2 Speaking Frequency Distribution of Scale Scores

Figure D.5.13. 2025 TELPAS Grade 3 Speaking Frequency Distribution of Scale Scores

Figure D.5.14. 2025 TELPAS Grade 4 Speaking Frequency Distribution of Scale Scores

Figure D.5.15. 2025 TELPAS Grade 5 Speaking Frequency Distribution of Scale Scores

Figure D.5.16. 2025 TELPAS Grade 6 Speaking Frequency Distribution of Scale Scores

Figure D.5.17. 2025 TELPAS Grade 7 Speaking Frequency Distribution of Scale Scores

Figure D.5.18. 2025 TELPAS Grade 8 Speaking Frequency Distribution of Scale Scores

Figure D.5.19. 2025 TELPAS Grade 9 Speaking Frequency Distribution of Scale Scores

Figure D.5.20. 2025 TELPAS Grade 10 Speaking Frequency Distribution of Scale Scores

Figure D.5.21. 2025 TELPAS Grade 11 Speaking Frequency Distribution of Scale Scores

Figure D.5.22. 2025 TELPAS Grade 12 Speaking Frequency Distribution of Scale Scores

Figure D.5.23. 2025 TELPAS Grade 2 Reading Frequency Distribution of Scale Scores

Figure D.5.24. 2025 TELPAS Grade 3 Reading Frequency Distribution of Scale Scores

Figure D.5.25. 2025 TELPAS Grade 4 Reading Frequency Distribution of Scale Scores

Figure D.5.26. 2025 TELPAS Grade 5 Reading Frequency Distribution of Scale Scores

Figure D.5.27. 2025 TELPAS Grade 6 Reading Frequency Distribution of Scale Scores

Figure D.5.28. 2025 TELPAS Grade 7 Reading Frequency Distribution of Scale Scores
Figure D.5.29. 2025 TELPAS Grade 8 Reading Frequency Distribution of Scale Scores
Figure D.5.30. 2025 TELPAS Grade 9 Reading Frequency Distribution of Scale Scores
Figure D.5.31. 2025 TELPAS Grade 10 Reading Frequency Distribution of Scale Scores
Figure D.5.32. 2025 TELPAS Grade 11 Reading Frequency Distribution of Scale Scores
Figure D.5.33. 2025 TELPAS Grade 12 Reading Frequency Distribution of Scale Scores
Figure D.5.34. 2025 TELPAS Grade 2 Writing Frequency Distribution of Scale Scores
Figure D.5.35. 2025 TELPAS Grade 3 Writing Frequency Distribution of Scale Scores
Figure D.5.36. 2025 TELPAS Grade 4 Writing Frequency Distribution of Scale Scores
Figure D.5.37. 2025 TELPAS Grade 5 Writing Frequency Distribution of Scale Scores
Figure D.5.38. 2025 TELPAS Grade 6 Writing Frequency Distribution of Scale Scores
Figure D.5.39. 2025 TELPAS Grade 7 Writing Frequency Distribution of Scale Scores
Figure D.5.40. 2025 TELPAS Grade 8 Writing Frequency Distribution of Scale Scores
Figure D.5.41. 2025 TELPAS Grade 9 Writing Frequency Distribution of Scale Scores
Figure D.5.42. 2025 TELPAS Grade 10 Writing Frequency Distribution of Scale Scores
Figure D.5.43. 2025 TELPAS Grade 11 Writing Frequency Distribution of Scale Scores
Figure D.5.44. 2025 TELPAS Grade 12 Writing Frequency Distribution of Scale Scores

2025 TELPAS
Classification Consistency and Accuracy

Table D.1.1. 2025 TELPAS Classification Consistency and Accuracy

Grade	Domain	Decision Consistency[1]	Decision Accuracy[2]
2–3	Listening	68.1	76.8
4–5	Listening	64.5	73.9
6–8	Listening	64.2	73.6
9–12	Listening	66.4	75.4
2–3	Speaking	79.3	84.6
4–5	Speaking	75.9	82.4
6–8	Speaking	80.8	86.0
9–12	Speaking	81.3	87.0
2	Reading	68.4	76.8
3	Reading	67.5	76.1
4–5	Reading	67.6	76.3
6–7	Reading	66.0	75.1
8–9	Reading	69.8	78.2
10–12	Reading	70.3	78.5
2	Writing	75.5	83.1
3	Writing	78.1	84.3
4–5	Writing	74.3	82.1
6–7	Writing	71.4	79.9
8–9	Writing	73.8	81.7
10–12	Writing	70.4	79.1

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.

2025 TELPAS
Scale Score Correlations

Table D.2.1. 2025 TELPAS Scale Score Correlations

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
2	Listening	Speaking	103,607	0.50
	Listening	Reading	103,493	0.69
	Listening	Writing	103,493	0.66
	Speaking	Reading	103,493	0.42
	Speaking	Writing	103,493	0.48
	Reading	Writing	103,751	0.77
3	Listening	Speaking	107,526	0.52
	Listening	Reading	107,411	0.74
	Listening	Writing	107,411	0.71
	Speaking	Reading	107,411	0.46
	Speaking	Writing	107,411	0.53
	Reading	Writing	107,683	0.78
4	Listening	Speaking	103,773	0.50
	Listening	Reading	103,663	0.77
	Listening	Writing	103,663	0.68
	Speaking	Reading	103,663	0.45
	Speaking	Writing	103,663	0.53
	Reading	Writing	103,898	0.76
5	Listening	Speaking	99,683	0.50
	Listening	Reading	99,593	0.78
	Listening	Writing	99,593	0.71
	Speaking	Reading	99,593	0.46
	Speaking	Writing	99,593	0.55
	Reading	Writing	99,812	0.77
6	Listening	Speaking	92,082	0.45
	Listening	Reading	91,958	0.77
	Listening	Writing	91,958	0.71
	Speaking	Reading	91,958	0.39
	Speaking	Writing	91,958	0.48
	Reading	Writing	92,175	0.75
7	Listening	Speaking	97,041	0.45
	Listening	Reading	96,861	0.78
	Listening	Writing	96,861	0.72
	Speaking	Reading	96,861	0.40
	Speaking	Writing	96,861	0.50
	Reading	Writing	97,086	0.76

(Continued)

Grade	TELPAS Domain	TELPAS Domain	N	Correlation
8	Listening	Speaking	93,751	0.46
	Listening	Reading	93,539	0.77
	Listening	Writing	93,539	0.71
	Speaking	Reading	93,539	0.39
	Speaking	Writing	93,539	0.51
	Reading	Writing	93,809	0.71
9	Listening	Speaking	103,008	0.50
	Listening	Reading	102,601	0.76
	Listening	Writing	102,601	0.72
	Speaking	Reading	102,601	0.47
	Speaking	Writing	102,601	0.53
	Reading	Writing	103,265	0.73
10	Listening	Speaking	88,115	0.47
	Listening	Reading	87,748	0.74
	Listening	Writing	87,748	0.68
	Speaking	Reading	87,748	0.43
	Speaking	Writing	87,748	0.50
	Reading	Writing	88,339	0.74
11	Listening	Speaking	67,254	0.44
	Listening	Reading	66,899	0.72
	Listening	Writing	66,899	0.65
	Speaking	Reading	66,899	0.40
	Speaking	Writing	66,899	0.47
	Reading	Writing	67,356	0.72
12	Listening	Speaking	57,704	0.43
	Listening	Reading	57,318	0.70
	Listening	Writing	57,318	0.62
	Speaking	Reading	57,318	0.37
	Speaking	Writing	57,318	0.44
	Reading	Writing	57,877	0.70

2025 TELPAS
Conditional Standard Error of Measurement for Scale Scores

Table D.3.1. 2025 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	1190		1166		1174		1219	
1	1275	70	1247	67	1258	70	1286	55
2	1326	51	1295	48	1309	51	1327	40
3	1357	43	1326	41	1340	42	1352	34
4	1381	38	1349	36	1364	38	1371	30
5	1401	35	1367	34	1383	35	1387	28
6	1418	33	1384	32	1400	33	1401	26
7	1433	31	1398	30	1415	31	1414	25
8	*1447	30	1412	29	1429	30	1425	24
9	1460	29	1425	28	*1441	29	1436	24
10	1473	29	1437	28	1454	28	1446	23
11	1485	28	1448	27	1465	28	*1457	23
12	1497	28	*1460	27	1477	28	1466	23
13	1509	28	1471	27	1488	28	1476	23
14	1520	28	1482	27	1499	28	1486	23
15	**1532	28	1493	27	1510	28	1496	23
16	1544	28	1504	27	1522	28	1506	23
17	1556	29	1516	28	**1533	28	1516	24
18	1569	29	**1528	28	1546	29	**1526	24
19	1582	30	1541	29	1558	30	1537	25
20	1596	31	1554	30	1572	31	1549	26
21	***1611	33	1569	31	1586	32	1562	27
22	1628	35	1585	33	***1603	34	1576	28
23	1648	38	***1604	36	1622	38	1592	31
24	1672	43	1626	41	1645	42	***1611	34
25	1703	51	1657	48	1676	50	1637	41
26	1754	69	1705	66	1727	69	1678	56
27	1838		1786		1811		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.3.2. 2025 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	1257		1359		1381		1415	
11	1302	37	1391	26	1407	21	1441	21
12	1331	28	1411	19	1424	16	1456	15
13	1350	24	1423	17	1435	14	1465	13
14	1364	22	1433	15	1445	13	1471	11
15	1377	21	1442	14	1453	12	1477	10
16	1389	20	1450	13	*1461	12	1481	9
17	1400	20	1457	13	1468	12	*1485	9
18	1410	19	1463	13	1475	12	1489	9
19	*1421	20	*1470	12	1482	12	1493	9
20	1433	20	1476	13	1489	12	1496	9
21	1444	21	1482	13	1497	12	1500	9
22	1457	21	1489	13	1504	12	1504	9
23	1471	22	1497	14	1512	13	1508	10
24	1486	23	1505	15	1520	13	1512	10
25	1502	24	1514	16	**1529	14	1518	11
26	1520	26	**1525	17	1540	15	**1525	13
27	**1541	28	1538	19	1551	15	1534	15
28	1565	30	1554	21	1564	16	1546	17
29	1590	30	1574	23	1576	16	1561	17
30	***1615	30	1595	22	1588	15	1574	16
31	1640	30	***1613	21	1599	14	1586	15
32	1666	31	1630	20	***1610	14	1597	15
33	1694	32	1646	20	1621	15	***1608	16
34	1724	33	1664	22	1633	16	1621	18
35	1761	40	1688	27	1650	21	1639	23
36	1811		1721		1676		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.

Table D.3.3. 2025 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.3.4. 2025 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.3.5. 2025 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	1349		1262			
7	1399	41	1320	46		
8	1430	30	1357	34		
9	*1449	25	1381	29	1258	
10	1464	22	1398	26	1320	52
11	1475	21	*1413	24	1350	35
12	1486	20	1427	24	1367	28
13	1497	20	1441	24	1380	25
14	1508	20	1455	24	1390	24
15	1519	21	1470	25	1399	23
16	**1532	22	1485	26	*1409	23
17	1546	23	1502	26	1418	24
18	1559	22	1519	26	1429	25
19	1572	22	**1535	26	1440	27
20	1584	21	1550	25	1454	29
21	1596	21	1565	24	1470	30
22	***1608	21	1579	24	1485	29
23	1620	22	1593	24	1498	26
24	1633	22	***1607	25	1510	24
25	1646	23	1622	25	1519	23
26	1660	23	1637	26	**1528	22
27	1675	24	1654	27	1537	22
28	1692	28	1674	31	1545	22
29	1718	37	1704	41	1553	22
30	1761		1752		1562	22
31					1571	23
32					1580	23
33					1589	23
34					1598	22
35					***1606	22
36					1614	22
37					1623	22
38					1632	23
39					1642	26
40					1656	30
41					1677	43
42					1721	

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. 2025 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		1250	
10	1345	47	1304	56	1336	68
11	1370	32	1338	37	1378	43
12	1384	27	1356	29	1399	34
13	1395	24	1369	26	1413	29
14	1405	23	1379	24	1424	27
15	1414	23	1389	23	1434	25
16	1422	23	1397	23	1443	24
17	*1432	24	1406	23	*1451	23
18	1442	25	*1415	24	1459	23
19	1453	27	1426	25	1466	22
20	1466	28	1437	27	1473	22
21	1480	28	1451	30	1480	22
22	1492	26	1467	31	1487	22
23	1503	24	1482	29	1494	22
24	1512	23	1495	26	1501	22
25	1521	22	1506	24	1508	22
26	**1529	22	1516	23	1515	22
27	1537	21	1524	22	1522	23
28	1544	22	**1532	22	**1530	23
29	1553	22	1540	22	1538	24
30	1561	23	1548	22	1546	24
31	1570	23	1556	22	1555	25
32	1579	23	1565	22	1564	25
33	1588	22	1573	22	1572	25
34	1596	22	1581	22	1581	25
35	***1604	22	1589	22	1590	25
36	1612	22	1597	22	1599	25
37	1621	22	***1605	22	***1609	26
38	1630	24	1614	23	1619	28
39	1640	26	1624	25	1632	32
40	1654	32	1637	30	1650	39
41	1678	47	1657	41	1684	61
42	1730		1696		1759	

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.

2025 TELPAS
Mean P-Values, Raw Score Summary, and Reliability

Table D.4.1. 2025 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	211,133	17.75	6.15	0.88	2.09	65.74
	Understand spoken words and language structures	5	211,133	3.34	1.43	0.64	0.86	66.73
	Basic understanding of spoken English	16	211,133	10.41	3.59	0.79	1.64	65.08
	Analyze and evaluate information in spoken English	6	211,133	4.00	1.74	0.70	0.95	66.64
Speaking	OVERALL TEST	36	211,133	21.01	6.23	0.93	1.66	47.01
	Provide and summarize information	16	211,133	9.70	2.55	0.82	1.09	51.95
	Share opinions and analyze information	20	211,133	11.31	3.87	0.91	1.17	42.08

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.4.2. 2025 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	102,590	17.86	5.97	0.88	2.08	66.15
	Understand spoken words and language structures	5	102,590	3.35	1.41	0.63	0.85	66.94
	Basic understanding of spoken English	16	102,590	10.47	3.48	0.78	1.64	65.42
	Analyze and evaluate information in spoken English	6	102,590	4.05	1.70	0.69	0.95	67.42
Speaking	OVERALL TEST	36	102,590	21.55	6.17	0.93	1.67	49.03
	Provide and summarize information	16	102,590	9.89	2.53	0.81	1.10	53.70
	Share opinions and analyze information	20	102,590	11.65	3.84	0.90	1.19	44.36

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous.
3. Mean of percent correct (0–100%) for the items.

Table D.4.3. 2025 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	108,483	17.65	6.31	0.89	2.09	65.35
	Understand spoken words and language structures	5	108,483	3.33	1.44	0.64	0.86	66.55
	Basic understanding of spoken English	16	108,483	10.36	3.70	0.80	1.64	64.77
	Analyze and evaluate information in spoken English	6	108,483	3.95	1.77	0.71	0.96	65.92
Speaking	OVERALL TEST	36	108,483	20.50	6.24	0.93	1.64	45.12
	Provide and summarize information	16	108,483	9.51	2.57	0.82	1.08	50.31
	Share opinions and analyze information	20	108,483	10.99	3.86	0.91	1.16	39.92

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.4.4. 2025 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	203,456	18.14	6.01	0.88	2.08	67.20
	Understand spoken words and language structures	5	203,456	3.70	1.37	0.63	0.83	74.03
	Basic understanding of spoken English	16	203,456	10.85	3.50	0.79	1.60	67.84
	Analyze and evaluate information in spoken English	6	203,456	3.59	1.79	0.68	1.02	59.78
Speaking	OVERALL TEST	36	203,456	23.55	6.05	0.92	1.73	56.42
	Provide and summarize information	16	203,456	11.47	2.84	0.80	1.29	65.62
	Share opinions and analyze information	20	203,456	12.08	3.42	0.89	1.15	47.22

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.4.5. 2025 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	98,192	18.11	5.89	0.88	2.08	67.08
	Understand spoken words and language structures	5	98,192	3.67	1.36	0.62	0.84	73.34
	Basic understanding of spoken English	16	98,192	10.84	3.41	0.78	1.59	67.77
	Analyze and evaluate information in spoken English	6	98,192	3.60	1.78	0.67	1.02	60.05
Speaking	OVERALL TEST	36	98,192	23.74	6.02	0.92	1.75	57.05
	Provide and summarize information	16	98,192	11.53	2.84	0.79	1.30	66.04
	Share opinions and analyze information	20	98,192	12.21	3.42	0.88	1.16	48.06

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.4.6. 2025 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	105,234	18.17	6.12	0.89	2.07	67.31
	Understand spoken words and language structures	5	105,234	3.73	1.38	0.64	0.82	74.68
	Basic understanding of spoken English	16	105,234	10.87	3.58	0.80	1.59	67.92
	Analyze and evaluate information in spoken English	6	105,234	3.57	1.80	0.68	1.02	59.53
Speaking	OVERALL TEST	36	105,234	23.38	6.06	0.92	1.70	55.85
	Provide and summarize information	16	105,234	11.41	2.85	0.80	1.27	65.24
	Share opinions and analyze information	20	105,234	11.97	3.42	0.89	1.13	46.45

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.4.7. 2025 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	282,874	17.93	5.76	0.86	2.15	66.42
	Understand spoken words and language structures	5	282,874	3.39	1.46	0.64	0.87	67.81
	Basic understanding of spoken English	16	282,874	10.83	3.51	0.78	1.63	67.66
	Analyze and evaluate information in spoken English	6	282,874	3.72	1.55	0.52	1.08	61.94
Speaking	OVERALL TEST	36	282,874	22.11	6.10	0.92	1.67	51.65
	Provide and summarize information	16	282,874	10.16	2.64	0.82	1.11	56.92
	Share opinions and analyze information	20	282,874	11.96	3.63	0.88	1.24	46.37

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.4.8. 2025 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	134,002	18.28	5.67	0.86	2.12	67.71
	Understand spoken words and language structures	5	134,002	3.40	1.45	0.64	0.87	67.90
	Basic understanding of spoken English	16	134,002	11.09	3.45	0.79	1.60	69.28
	Analyze and evaluate information in spoken English	6	134,002	3.80	1.52	0.50	1.07	63.34
Speaking	OVERALL TEST	36	134,002	22.25	6.09	0.92	1.70	52.20
	Provide and summarize information	16	134,002	10.19	2.63	0.81	1.13	57.27
	Share opinions and analyze information	20	134,002	12.07	3.63	0.88	1.26	47.13

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.4.9. 2025 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	148,806	17.62	5.81	0.86	2.17	65.27
	Understand spoken words and language structures	5	148,806	3.39	1.46	0.64	0.88	67.75
	Basic understanding of spoken English	16	148,806	10.59	3.54	0.78	1.65	66.22
	Analyze and evaluate information in spoken English	6	148,806	3.64	1.58	0.53	1.08	60.69
Speaking	OVERALL TEST	36	148,806	21.99	6.11	0.93	1.65	51.16
	Provide and summarize information	16	148,806	10.14	2.64	0.83	1.09	56.61
	Share opinions and analyze information	20	148,806	11.86	3.64	0.89	1.22	45.70

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.4.10. 2025 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	316,081	18.54	6.00	0.89	2.01	68.65
	Understand spoken words and language structures	5	316,081	3.68	1.39	0.66	0.81	73.67
	Basic understanding of spoken English	16	316,081	10.81	3.62	0.81	1.56	67.59
	Analyze and evaluate information in spoken English	6	316,081	4.04	1.58	0.62	0.98	67.31
Speaking	OVERALL TEST	36	316,081	21.08	8.67	0.96	1.76	45.19
	Provide and summarize information	16	316,081	9.99	3.88	0.89	1.29	49.78
	Share opinions and analyze information	20	316,081	11.09	4.94	0.94	1.16	40.59

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.4.11. 2025 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	146,566	18.70	5.77	0.88	2.00	69.25
	Understand spoken words and language structures	5	146,566	3.71	1.36	0.65	0.81	74.23
	Basic understanding of spoken English	16	146,566	10.95	3.51	0.81	1.55	68.43
	Analyze and evaluate information in spoken English	6	146,566	4.04	1.51	0.59	0.97	67.29
Speaking	OVERALL TEST	36	146,566	21.43	8.71	0.96	1.81	46.51
	Provide and summarize information	16	146,566	10.12	3.89	0.88	1.33	50.96
	Share opinions and analyze information	20	146,566	11.31	4.99	0.94	1.19	42.06

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.4.12. 2025 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	169,408	18.40	6.18	0.89	2.02	68.14
	Understand spoken words and language structures	5	169,408	3.66	1.41	0.67	0.81	73.19
	Basic understanding of spoken English	16	169,408	10.70	3.71	0.82	1.57	66.87
	Analyze and evaluate information in spoken English	6	169,408	4.04	1.63	0.64	0.98	67.33
Speaking	OVERALL TEST	36	169,408	20.77	8.62	0.96	1.71	44.05
	Provide and summarize information	16	169,408	9.87	3.87	0.89	1.26	48.78
	Share opinions and analyze information	20	169,408	10.90	4.90	0.95	1.13	39.33

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.4.13. 2025 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	103,751	12.76	6.29	0.86	2.39	42.54
	Understand words and language structures	12	103,751	5.37	2.85	0.73	1.49	44.72
	Basic understanding of variety of texts written in English	7	103,751	2.60	1.72	0.55	1.16	37.14
	Analyze and evaluate information and ideas in a variety of texts written in English	11	103,751	4.80	2.58	0.68	1.45	43.60
Writing	OVERALL TEST	30	103,751	10.26	4.70	0.92	1.36	26.16
	Express ideas in writing and address writing assignments	24	103,751	8.29	3.09	0.88	1.07	12.71
	Use standard grammar, usage, and spelling to edit writing tasks	6	103,751	1.97	1.95	0.81	0.85	32.89

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. 2025 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	50,415	12.86	6.22	0.85	2.39	42.87
	Understand words and language structures	12	50,415	5.30	2.79	0.71	1.50	44.17
	Basic understanding of variety of texts written in English	7	50,415	2.67	1.73	0.55	1.17	38.10
	Analyze and evaluate information and ideas in a variety of texts written in English	11	50,415	4.89	2.57	0.68	1.45	44.49
Writing	OVERALL TEST	30	50,415	10.55	4.86	0.92	1.40	27.25
	Express ideas in writing and address writing assignments	24	50,415	8.52	3.23	0.88	1.11	13.97
	Use standard grammar, usage, and spelling to edit writing tasks	6	50,415	2.03	1.95	0.81	0.85	33.88

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. 2025 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	53,299	12.67	6.35	0.86	2.38	42.23
	Understand words and language structures	12	53,299	5.43	2.91	0.74	1.49	45.24
	Basic understanding of variety of texts written in English	7	53,299	2.54	1.71	0.54	1.16	36.23
	Analyze and evaluate information and ideas in a variety of texts written in English	11	53,299	4.70	2.59	0.69	1.45	42.76
Writing	OVERALL TEST	30	53,299	9.99	4.54	0.91	1.33	25.14
	Express ideas in writing and address writing assignments	24	53,299	8.07	2.93	0.88	1.04	11.53
	Use standard grammar, usage, and spelling to edit writing tasks	6	53,299	1.92	1.95	0.82	0.84	31.95

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.4.16. 2025 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	107,683	16.80	7.62	0.89	2.51	49.42
	Understand words and language structures	12	107,683	7.39	3.20	0.81	1.39	61.61
	Basic understanding of variety of texts written in English	9	107,683	3.96	2.27	0.66	1.32	44.00
	Analyze and evaluate information and ideas in a variety of texts written in English	13	107,683	5.45	3.16	0.75	1.57	41.91
Writing	OVERALL TEST	30	107,683	13.56	5.61	0.92	1.59	44.55
	Express ideas in writing and address writing assignments	24	107,683	10.27	3.84	0.88	1.34	23.70
	Use standard grammar, usage, and spelling to edit writing tasks	6	107,683	3.30	2.09	0.83	0.86	54.98

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. 2025 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	52,279	16.95	7.64	0.89	2.50	49.87
	Understand words and language structures	12	52,279	7.40	3.17	0.81	1.39	61.70
	Basic understanding of variety of texts written in English	9	52,279	4.01	2.27	0.66	1.32	44.60
	Analyze and evaluate information and ideas in a variety of texts written in English	13	52,279	5.54	3.19	0.76	1.57	42.58
Writing	OVERALL TEST	30	52,279	14.02	5.74	0.92	1.64	46.11
	Express ideas in writing and address writing assignments	24	52,279	10.64	3.97	0.88	1.39	25.79
	Use standard grammar, usage, and spelling to edit writing tasks	6	52,279	3.38	2.08	0.83	0.86	56.27

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.4.18. 2025 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	55,380	16.66	7.60	0.89	2.51	49.01
	Understand words and language structures	12	55,380	7.39	3.23	0.81	1.39	61.54
	Basic understanding of variety of texts written in English	9	55,380	3.91	2.27	0.66	1.32	43.45
	Analyze and evaluate information and ideas in a variety of texts written in English	13	55,380	5.37	3.13	0.75	1.57	41.28
Writing	OVERALL TEST	30	55,380	13.14	5.46	0.92	1.55	43.09
	Express ideas in writing and address writing assignments	24	55,380	9.91	3.67	0.88	1.29	21.73
	Use standard grammar, usage, and spelling to edit writing tasks	6	55,380	3.23	2.10	0.83	0.86	53.77

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. 2025 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	203,710	19.67	7.97	0.89	2.58	54.64
	Understand words and language structures	12	203,710	7.13	3.03	0.77	1.44	59.44
	Basic understanding of variety of texts written in English	10	203,710	5.14	2.55	0.71	1.37	51.36
	Analyze and evaluate information and ideas in a variety of texts written in English	14	203,710	7.40	3.23	0.74	1.64	52.88
Writing	OVERALL TEST	42	203,710	22.31	8.72	0.93	2.34	47.82
	Express ideas in writing and address writing assignments	36	203,710	19.13	7.08	0.91	2.15	37.53
	Use standard grammar, usage, and spelling to edit writing tasks	6	203,710	3.18	2.12	0.82	0.91	52.97

Notes:

1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.20. 2025 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	98,288	19.77	7.90	0.89	2.58	54.93
	Understand words and language structures	12	98,288	7.02	2.98	0.76	1.45	58.53
	Basic understanding of variety of texts written in English	10	98,288	5.24	2.56	0.71	1.37	52.37
	Analyze and evaluate information and ideas in a variety of texts written in English	14	98,288	7.51	3.20	0.74	1.64	53.67
Writing	OVERALL TEST	42	98,288	23.37	8.92	0.93	2.39	49.74
	Express ideas in writing and address writing assignments	36	98,288	20.13	7.28	0.91	2.22	41.24
	Use standard grammar, usage, and spelling to edit writing tasks	6	98,288	3.24	2.11	0.81	0.91	53.99

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.21. 2025 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	105,391	19.58	8.04	0.90	2.58	54.38
	Understand words and language structures	12	105,391	7.24	3.08	0.78	1.43	60.29
	Basic understanding of variety of texts written in English	10	105,391	5.04	2.54	0.71	1.37	50.42
	Analyze and evaluate information and ideas in a variety of texts written in English	14	105,391	7.30	3.25	0.75	1.64	52.15
Writing	OVERALL TEST	42	105,391	21.32	8.40	0.93	2.28	46.05
	Express ideas in writing and address writing assignments	36	105,391	18.20	6.75	0.90	2.09	34.08
	Use standard grammar, usage, and spelling to edit writing tasks	6	105,391	3.12	2.12	0.82	0.90	52.03

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.22. 2025 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	189,261	21.10	7.96	0.89	2.66	57.02
	Understand words and language structures	12	189,261	7.51	3.14	0.79	1.43	62.58
	Basic understanding of variety of texts written in English	10	189,261	6.03	2.62	0.74	1.35	60.26
	Analyze and evaluate information and ideas in a variety of texts written in English	15	189,261	7.56	3.15	0.69	1.76	50.42
Writing	OVERALL TEST	42	189,261	23.86	8.28	0.91	2.49	47.72
	Express ideas in writing and address writing assignments	36	189,261	20.89	7.01	0.89	2.27	44.03
	Use standard grammar, usage, and spelling to edit writing tasks	6	189,261	2.97	1.76	0.66	1.02	49.56

Notes:

1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.23. 2025 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	90,102	21.80	7.91	0.89	2.64	58.93
	Understand words and language structures	12	90,102	7.63	3.13	0.79	1.42	63.62
	Basic understanding of variety of texts written in English	10	90,102	6.27	2.58	0.73	1.33	62.68
	Analyze and evaluate information and ideas in a variety of texts written in English	15	90,102	7.90	3.14	0.69	1.75	52.68
Writing	OVERALL TEST	42	90,102	25.20	8.40	0.91	2.52	50.47
	Express ideas in writing and address writing assignments	36	90,102	22.11	7.13	0.90	2.30	48.55
	Use standard grammar, usage, and spelling to edit writing tasks	6	90,102	3.09	1.75	0.66	1.02	51.43

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.24. 2025 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	99,113	20.46	7.95	0.89	2.67	55.30
	Understand words and language structures	12	99,113	7.40	3.15	0.79	1.43	61.65
	Basic understanding of variety of texts written in English	10	99,113	5.81	2.64	0.73	1.37	58.07
	Analyze and evaluate information and ideas in a variety of texts written in English	15	99,113	7.25	3.13	0.68	1.76	48.37
Writing	OVERALL TEST	42	99,113	22.65	7.98	0.91	2.45	45.23
	Express ideas in writing and address writing assignments	36	99,113	19.78	6.70	0.89	2.23	39.93
	Use standard grammar, usage, and spelling to edit writing tasks	6	99,113	2.87	1.77	0.66	1.03	47.88

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.25. 2025 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	197,074	21.09	8.27	0.90	2.60	56.99
	Understand words and language structures	12	197,074	7.27	3.00	0.78	1.41	60.62
	Basic understanding of variety of texts written in English	10	197,074	5.38	2.42	0.66	1.41	53.81
	Analyze and evaluate information and ideas in a variety of texts written in English	15	197,074	8.43	3.78	0.81	1.64	56.21
Writing	OVERALL TEST	42	197,074	23.63	8.04	0.90	2.58	55.95
	Express ideas in writing and address writing assignments	36	197,074	19.79	6.56	0.86	2.43	39.97
	Use standard grammar, usage, and spelling to edit writing tasks	6	197,074	3.84	2.05	0.82	0.87	63.93

Notes:

1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.26. 2025 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	91,319	21.54	8.18	0.90	2.59	58.23
	Understand words and language structures	12	91,319	7.31	2.90	0.76	1.41	60.93
	Basic understanding of variety of texts written in English	10	91,319	5.55	2.46	0.67	1.41	55.53
	Analyze and evaluate information and ideas in a variety of texts written in English	15	91,319	8.68	3.75	0.81	1.64	57.87
Writing	OVERALL TEST	42	91,319	24.94	8.09	0.89	2.68	58.93
	Express ideas in writing and address writing assignments	36	91,319	20.97	6.62	0.85	2.55	44.32
	Use standard grammar, usage, and spelling to edit writing tasks	6	91,319	3.97	2.04	0.83	0.85	66.23

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.27. 2025 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	105,718	20.70	8.33	0.90	2.61	55.93
	Understand words and language structures	12	105,718	7.24	3.09	0.79	1.42	60.37
	Basic understanding of variety of texts written in English	10	105,718	5.23	2.39	0.65	1.42	52.33
	Analyze and evaluate information and ideas in a variety of texts written in English	15	105,718	8.22	3.80	0.81	1.64	54.79
Writing	OVERALL TEST	42	105,718	22.50	7.82	0.90	2.48	53.38
	Express ideas in writing and address writing assignments	36	105,718	18.78	6.34	0.87	2.32	36.22
	Use standard grammar, usage, and spelling to edit writing tasks	6	105,718	3.72	2.05	0.82	0.88	61.96

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.28. 2025 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	213,572	21.32	7.94	0.90	2.55	57.62
	Understand words and language structures	12	213,572	8.77	2.99	0.82	1.26	73.06
	Basic understanding of variety of texts written in English	10	213,572	5.32	2.54	0.71	1.36	53.17
	Analyze and evaluate information and ideas in a variety of texts written in English	15	213,572	7.23	3.39	0.74	1.71	48.23
Writing	OVERALL TEST	42	213,572	24.76	8.97	0.89	3.03	50.12
	Express ideas in writing and address writing assignments	36	213,572	21.65	7.91	0.87	2.87	46.86
	Use standard grammar, usage, and spelling to edit writing tasks	6	213,572	3.10	1.71	0.69	0.95	51.75

Notes:

1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.29. 2025 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	99,581	21.68	7.78	0.89	2.54	58.60
	Understand words and language structures	12	99,581	8.83	2.93	0.82	1.25	73.59
	Basic understanding of variety of texts written in English	10	99,581	5.51	2.51	0.71	1.36	55.11
	Analyze and evaluate information and ideas in a variety of texts written in English	15	99,581	7.34	3.33	0.74	1.71	48.92
Writing	OVERALL TEST	42	99,581	26.48	8.90	0.88	3.09	53.46
	Express ideas in writing and address writing assignments	36	99,581	23.25	7.82	0.86	2.94	52.79
	Use standard grammar, usage, and spelling to edit writing tasks	6	99,581	3.23	1.71	0.70	0.94	53.79

Notes:
1. Total number of score points.
2. Coefficient alpha computed if there are any polytomous items;
(a) KR_{20} is computed instead if all items are dichotomous;
(b) stratified alpha is computed instead if there are multiple item types.
3. Mean of percent correct (0–100%) for the items.

Table D.4.30. 2025 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	113,899	21.00	8.06	0.90	2.55	56.77
	Understand words and language structures	12	113,899	8.71	3.04	0.83	1.26	72.60
	Basic understanding of variety of texts written in English	10	113,899	5.15	2.55	0.71	1.37	51.48
	Analyze and evaluate information and ideas in a variety of texts written in English	15	113,899	7.14	3.44	0.75	1.71	47.62
Writing	OVERALL TEST	42	113,899	23.25	8.75	0.89	2.94	47.21
	Express ideas in writing and address writing assignments	36	113,899	20.26	7.72	0.87	2.78	41.69
	Use standard grammar, usage, and spelling to edit writing tasks	6	113,899	3.00	1.70	0.69	0.95	49.97

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.

2025 TELPAS
Scale Score Descriptive Statistics

Table D.5.1. 2025 TELPAS Listening Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Listening	103,607	1552.30	1544	1569	648	126	89.08	7934.51	0.45	0.55
Grade 3	Listening	107,526	1613.09	1611	1703	648	128	105.40	11109.61	0.19	-0.29
Grade 4	Listening	103,773	1530.35	1528	1585	620	114	87.05	7577.70	0.32	0.17
Grade 5	Listening	99,683	1563.08	1569	1626	620	133	98.55	9712.72	0.12	-0.19
Grade 6	Listening	92,082	1539.88	1546	1586	637	98	79.60	6336.83	0.19	0.32
Grade 7	Listening	97,041	1560.46	1558	1622	637	123	87.12	7589.73	0.10	0.18
Grade 8	Listening	93,751	1572.30	1572	1645	637	112	90.36	8165.16	0.05	0.08
Grade 9	Listening	103,008	1534.74	1537	1592	527	106	79.20	6272.14	-0.20	0.32
Grade 10	Listening	88,115	1550.15	1562	1611	527	115	80.02	6403.62	-0.31	0.50
Grade 11	Listening	67,254	1551.02	1562	1611	527	105	77.94	6074.64	-0.33	0.70
Grade 12	Listening	57,704	1544.87	1549	1592	527	96	79.55	6328.73	-0.43	1.20

Table D.5.2. 2025 TELPAS Speaking Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Speaking	103,607	1434.26	1433	1257	554	125	95.91	9199.48	0.01	-0.09
Grade 3	Speaking	107,526	1469.91	1471	1520	554	131	102.80	10568.68	-0.13	-0.08
Grade 4	Speaking	103,773	1510.27	1514	1554	362	78	63.14	3986.76	-0.16	0.33
Grade 5	Speaking	99,683	1516.86	1514	1554	362	78	67.60	4569.85	-0.09	0.15
Grade 6	Speaking	92,082	1510.85	1512	1529	295	76	55.41	3070.44	-0.13	0.21
Grade 7	Speaking	97,041	1508.79	1512	1540	295	76	57.44	3298.78	-0.11	0.10
Grade 8	Speaking	93,751	1507.38	1512	1529	295	83	60.17	3620.97	-0.12	-0.05
Grade 9	Speaking	103,008	1504.18	1508	1415	252	105	66.50	4422.06	0.20	-0.79
Grade 10	Speaking	88,115	1508.05	1512	1415	252	120	68.34	4669.77	0.15	-0.84
Grade 11	Speaking	67,254	1505.95	1508	1415	252	120	68.89	4746.03	0.18	-0.88
Grade 12	Speaking	57,704	1498.31	1500	1415	252	131	68.78	4730.90	0.33	-0.83

Table D.5.3. 2025 TELPAS Reading Scale Score Descriptive Statistics

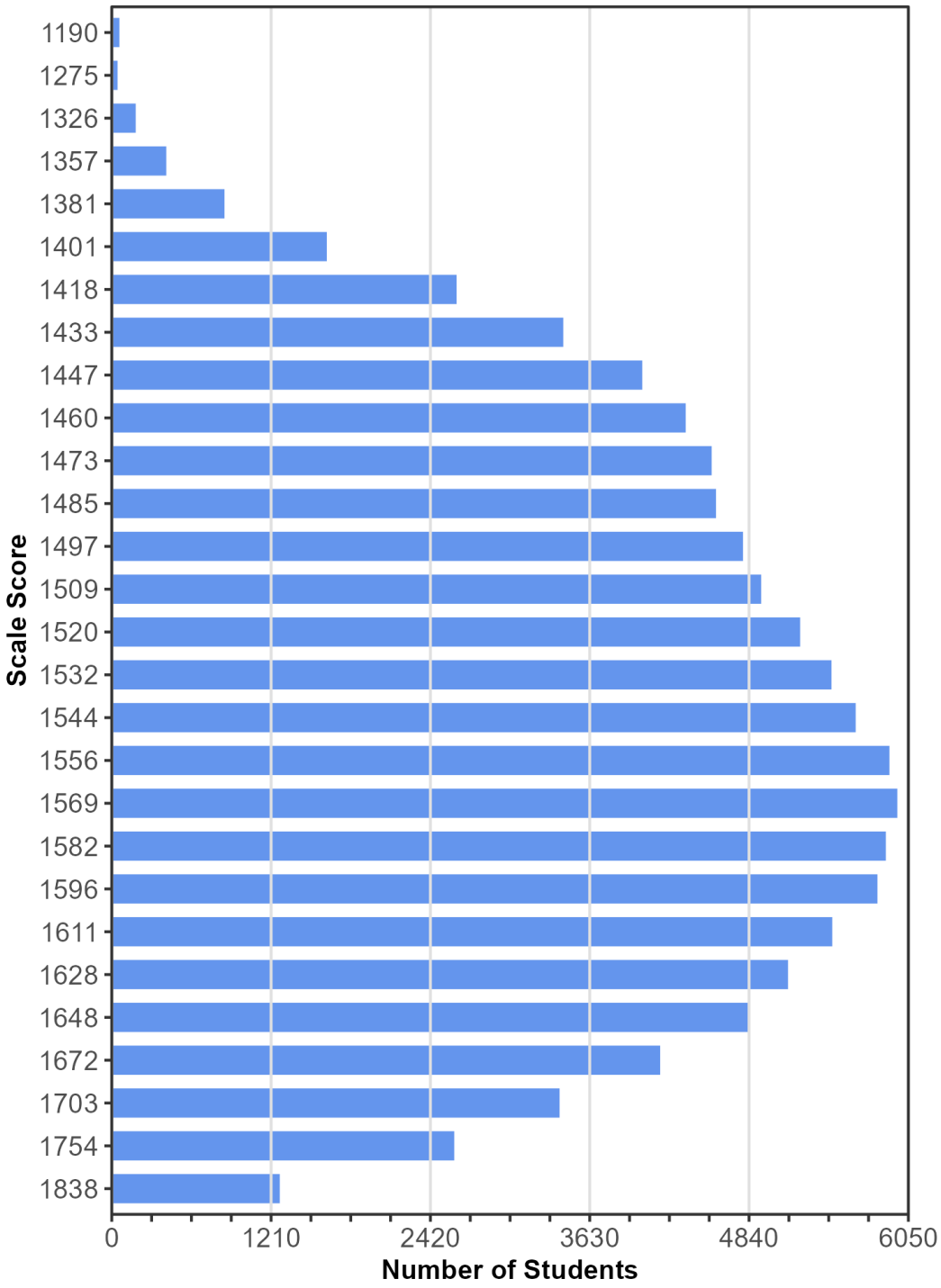
Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Reading	103,751	1466.96	1450	1416	642	94	73.98	5472.74	0.79	0.85
Grade 3	Reading	107,683	1509.38	1496	1420	881	148	112.45	12645.10	0.80	0.80
Grade 4	Reading	103,898	1537.95	1531	1450	876	135	99.46	9892.36	0.59	0.57
Grade 5	Reading	99,812	1574.61	1564	1650	876	164	113.99	12993.17	0.46	0.20
Grade 6	Reading	92,175	1525.83	1521	1605	795	139	85.62	7330.82	0.25	-0.17
Grade 7	Reading	97,086	1549.50	1550	1618	795	144	93.48	8738.56	0.11	-0.23
Grade 8	Reading	93,809	1543.49	1537	1503	698	125	83.62	6991.75	0.44	0.29
Grade 9	Reading	103,265	1545.18	1537	1449	698	136	92.06	8475.32	0.30	0.53
Grade 10	Reading	88,339	1521.93	1516	1516	677	116	84.15	7081.28	0.14	0.89
Grade 11	Reading	67,356	1527.00	1524	1507	677	107	84.92	7211.80	0.15	0.93
Grade 12	Reading	57,877	1521.50	1516	1507	677	116	85.77	7356.19	-0.03	1.79

Table D.5.4. 2025 TELPAS Writing Scale Score Descriptive Statistics

Grade	Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
Grade 2	Writing	103,751	1441.62	1430	1349	412	159	79.19	6270.37	0.39	-0.87
Grade 3	Writing	107,683	1438.61	1455	1320	490	162	102.08	10421.16	-0.16	-0.91
Grade 4	Writing	103,898	1455.12	1470	1258	463	157	107.51	11559.36	-0.38	-0.73
Grade 5	Writing	99,812	1487.54	1510	1258	463	133	109.47	11984.05	-0.62	-0.32
Grade 6	Writing	92,175	1486.94	1503	1291	439	111	90.03	8105.17	-0.61	-0.08
Grade 7	Writing	97,086	1508.11	1529	1291	439	104	92.27	8513.86	-0.73	0.21
Grade 8	Writing	93,809	1477.13	1495	1495	462	114	95.64	9147.93	-0.91	0.52
Grade 9	Writing	103,265	1470.50	1495	1234	462	133	103.41	10693.29	-0.82	0.06
Grade 10	Writing	88,339	1498.14	1515	1250	509	105	97.24	9455.04	-0.78	0.76
Grade 11	Writing	67,356	1501.50	1515	1250	509	105	93.95	8827.35	-0.78	0.97
Grade 12	Writing	57,877	1499.34	1515	1250	509	105	92.09	8480.63	-0.77	0.93

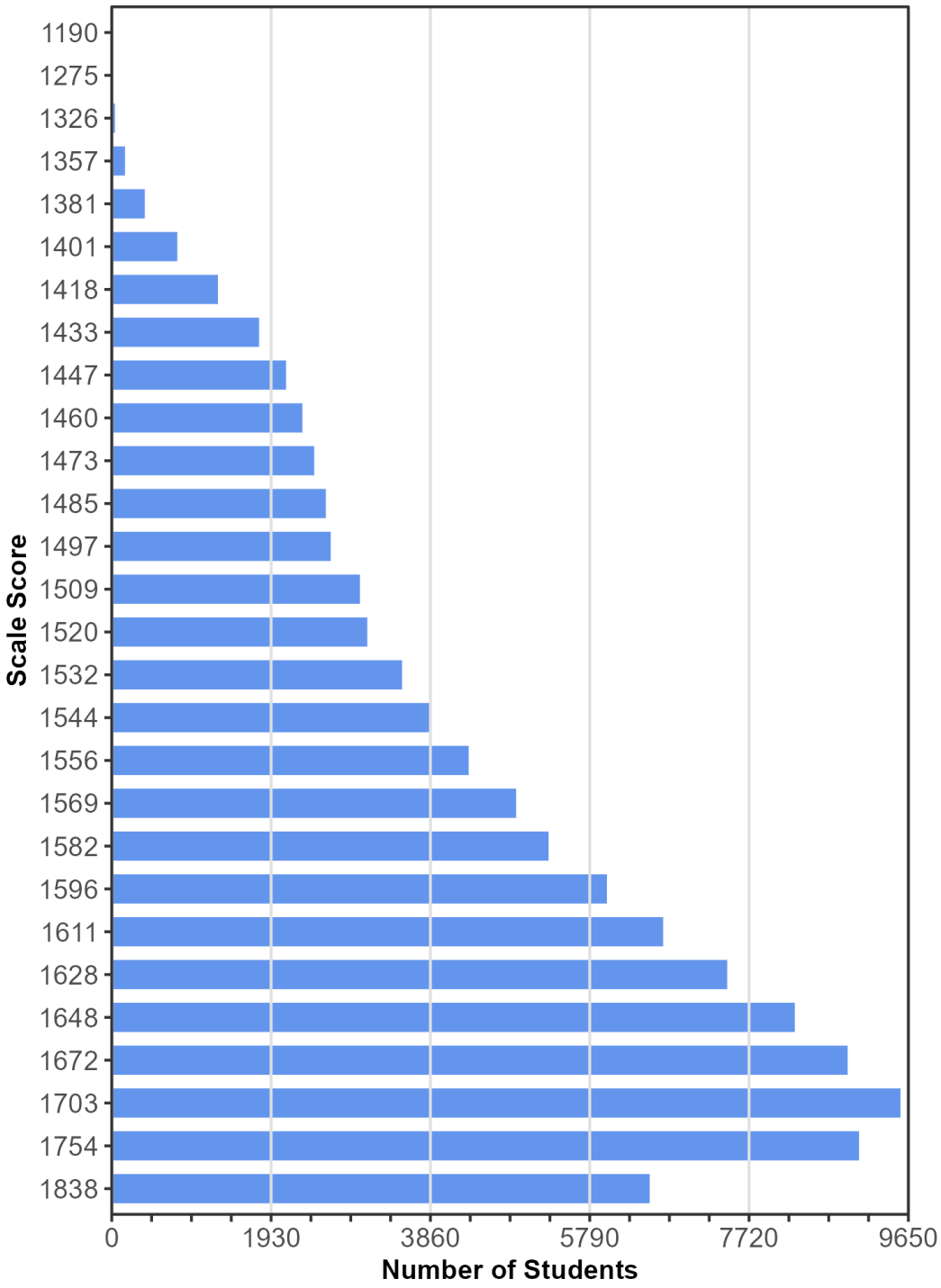
2025 TELPAS
Frequency Distribution of Scale Scores

Figure D.5.1. 2025 TELPAS Grade 2 Listening
Frequency Distribution of Scale Scores



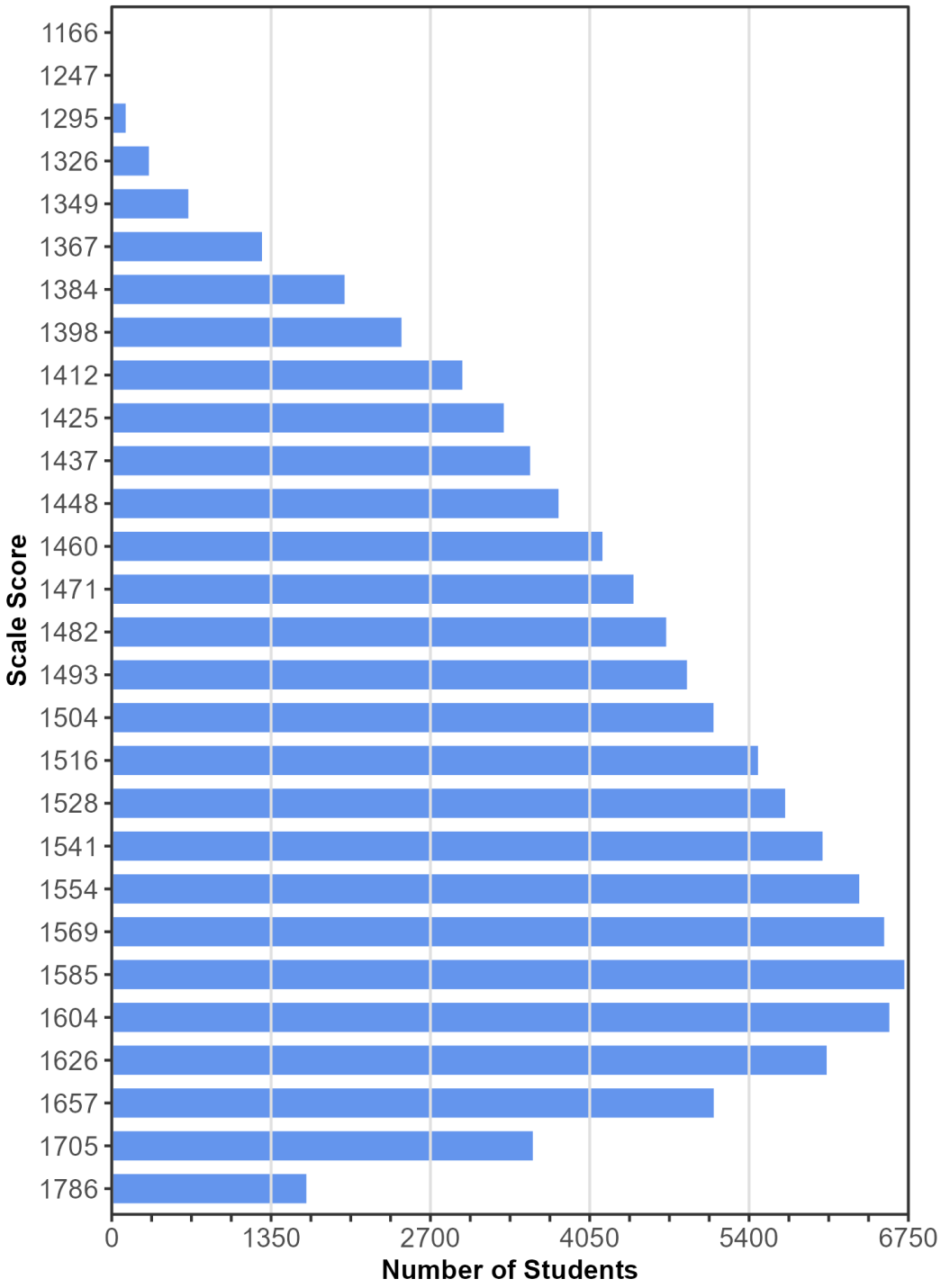
	Freq	Cum Freq	Pct	Cum Pct
1190	69	69	0.07	0.07
1275	55	124	0.05	0.12
1326	193	317	0.19	0.31
1357	425	742	0.41	0.72
1381	867	1609	0.84	1.55
1401	1644	3253	1.59	3.14
1418	2630	5883	2.54	5.68
1433	3440	9323	3.32	9.00
1447	4040	13363	3.90	12.90
1460	4370	17733	4.22	17.12
1473	4567	22300	4.41	21.52
1485	4600	26900	4.44	25.96
1497	4805	31705	4.64	30.60
1509	4943	36648	4.77	35.37
1520	5239	41887	5.06	40.43
1532	5477	47364	5.29	45.72
1544	5661	53025	5.46	51.18
1556	5917	58942	5.71	56.89
1569	5977	64919	5.77	62.66
1582	5890	70809	5.68	68.34
1596	5826	76635	5.62	73.97
1611	5483	82118	5.29	79.26
1628	5147	87265	4.97	84.23
1648	4855	92120	4.69	88.91
1672	4176	96296	4.03	92.94
1703	3412	99708	3.29	96.24
1754	2612	102320	2.52	98.76
1838	1287	103607	1.24	100.00

Figure D.5.2. 2025 TELPAS Grade 3 Listening
Frequency Distribution of Scale Scores



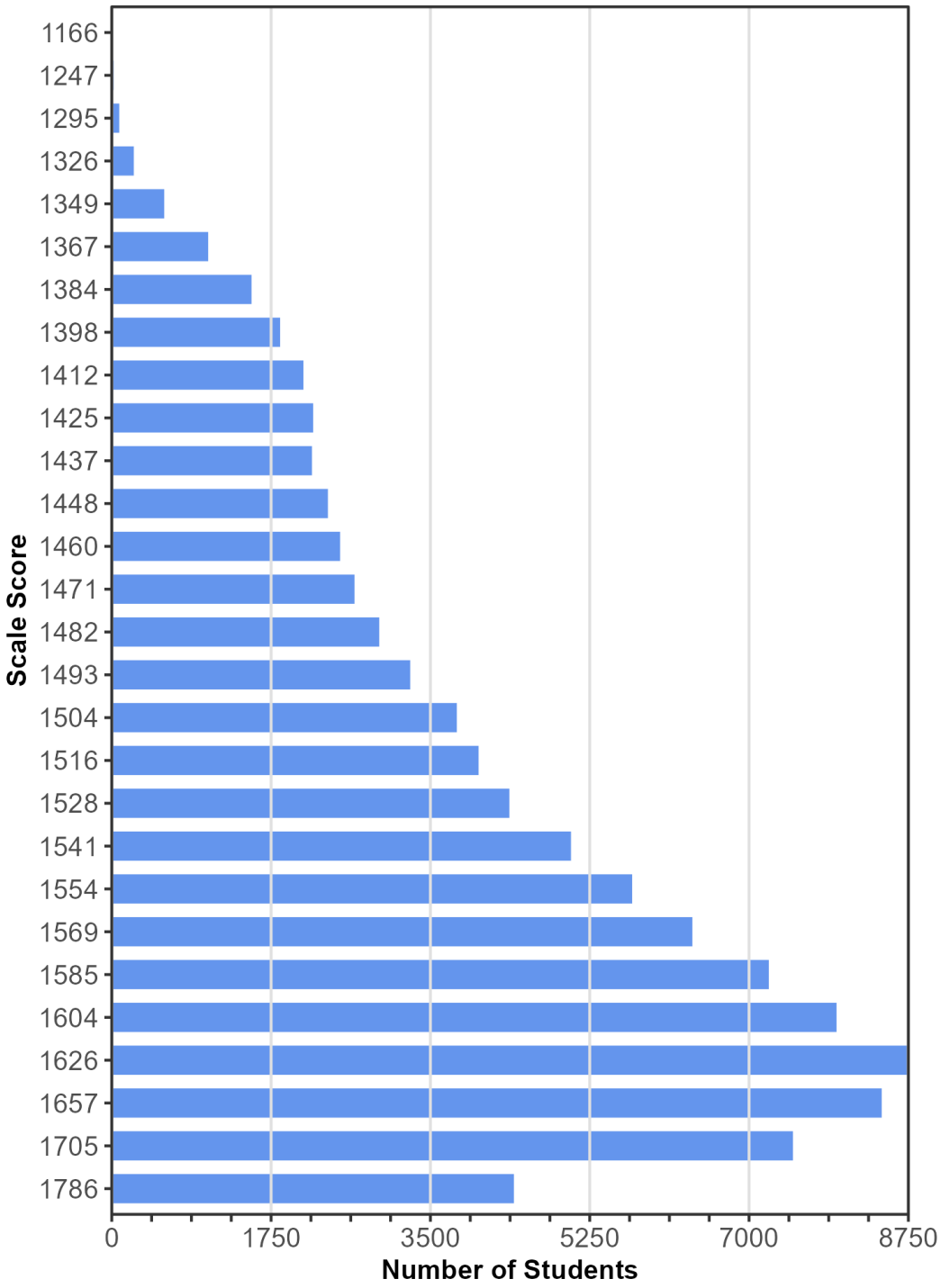
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1190	25	25	0.02	0.02
1275	29	54	0.03	0.05
1326	56	110	0.05	0.10
1357	178	288	0.17	0.27
1381	418	706	0.39	0.66
1401	812	1518	0.76	1.41
1418	1303	2821	1.21	2.62
1433	1802	4623	1.68	4.30
1447	2129	6752	1.98	6.28
1460	2327	9079	2.16	8.44
1473	2470	11549	2.30	10.74
1485	2611	14160	2.43	13.17
1497	2670	16830	2.48	15.65
1509	3024	19854	2.81	18.46
1520	3113	22967	2.90	21.36
1532	3534	26501	3.29	24.65
1544	3873	30374	3.60	28.25
1556	4341	34715	4.04	32.29
1569	4916	39631	4.57	36.86
1582	5309	44940	4.94	41.79
1596	6016	50956	5.59	47.39
1611	6697	57653	6.23	53.62
1628	7473	65126	6.95	60.57
1648	8292	73418	7.71	68.28
1672	8932	82350	8.31	76.59
1703	9572	91922	8.90	85.49
1754	9070	100992	8.44	93.92
1838	6534	107526	6.08	100.00

Figure D.5.3. 2025 TELPAS Grade 4 Listening
Frequency Distribution of Scale Scores



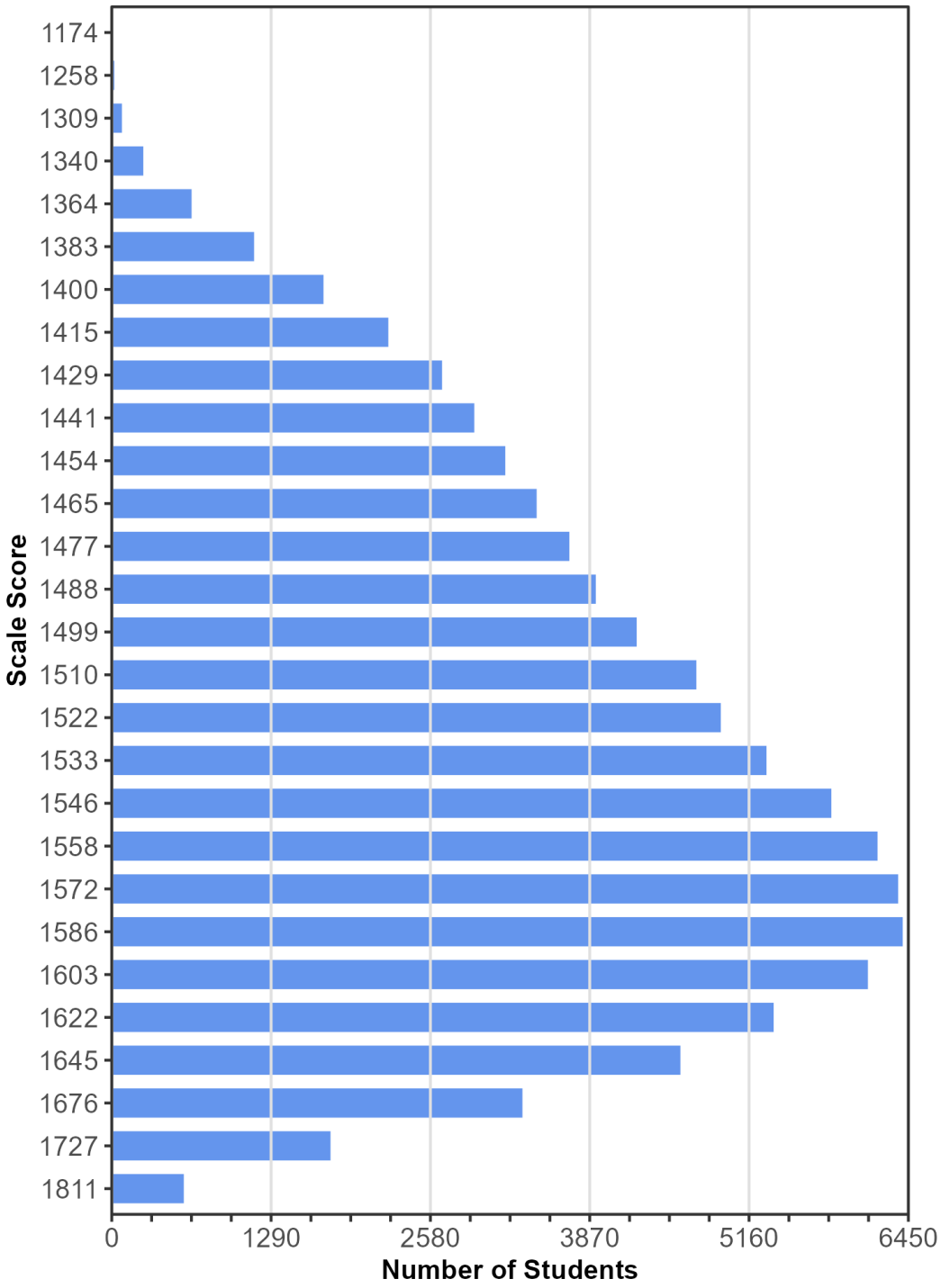
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1166	20	20	0.02	0.02
1247	19	39	0.02	0.04
1295	130	169	0.13	0.16
1326	327	496	0.32	0.48
1349	661	1157	0.64	1.11
1367	1285	2442	1.24	2.35
1384	1985	4427	1.91	4.27
1398	2467	6894	2.38	6.64
1412	2983	9877	2.87	9.52
1425	3334	13211	3.21	12.73
1437	3557	16768	3.43	16.16
1448	3798	20566	3.66	19.82
1460	4170	24736	4.02	23.84
1471	4433	29169	4.27	28.11
1482	4710	33879	4.54	32.65
1493	4886	38765	4.71	37.36
1504	5111	43876	4.93	42.28
1516	5488	49364	5.29	47.57
1528	5718	55082	5.51	53.08
1541	6035	61117	5.82	58.89
1554	6346	67463	6.12	65.01
1569	6557	74020	6.32	71.33
1585	6728	80748	6.48	77.81
1604	6601	87349	6.36	84.17
1626	6070	93419	5.85	90.02
1657	5113	98532	4.93	94.95
1705	3580	102112	3.45	98.40
1786	1661	103773	1.60	100.00

Figure D.5.4. 2025 TELPAS Grade 5 Listening
Frequency Distribution of Scale Scores



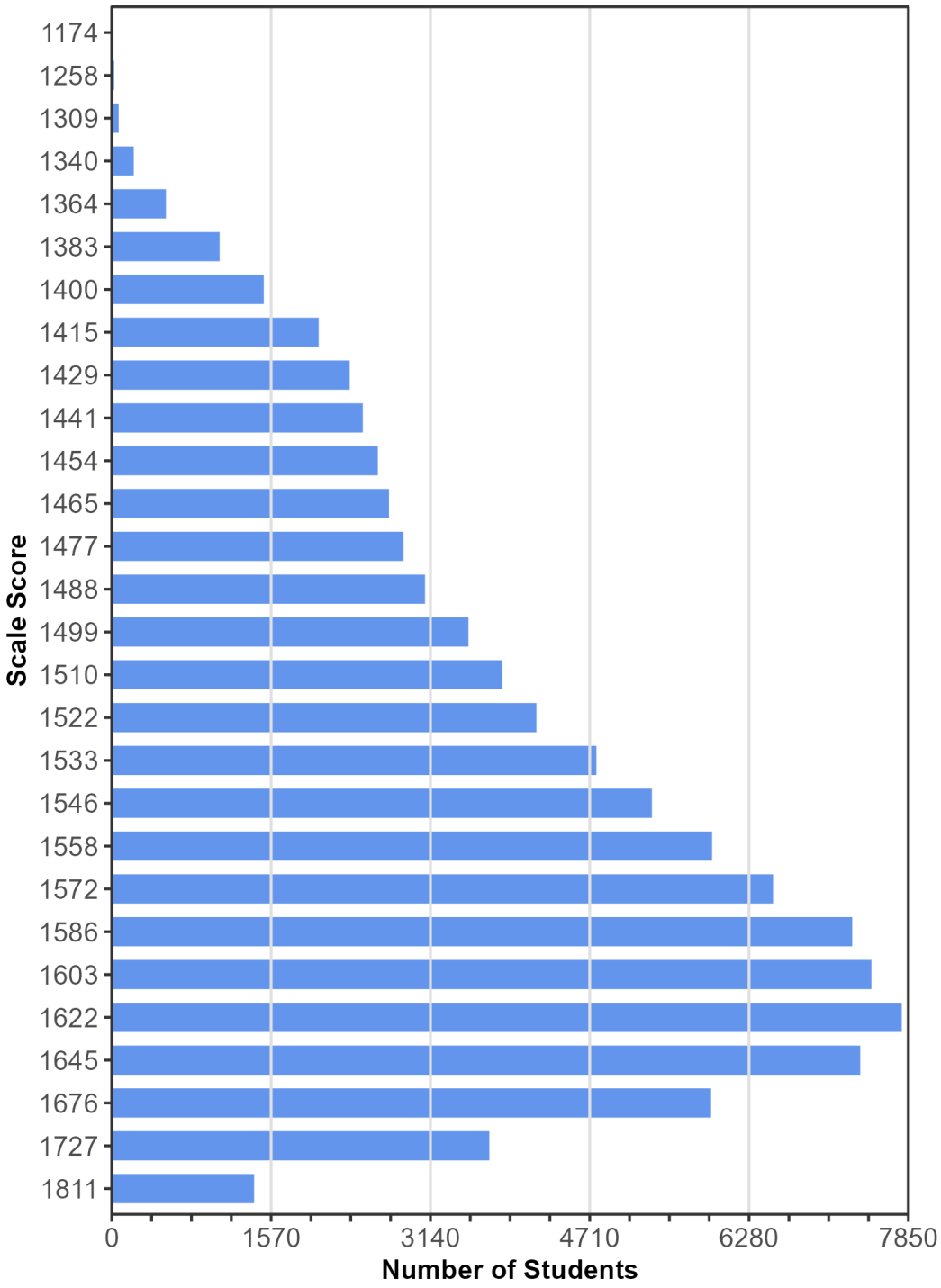
	Freq	Cum Freq	Pct	Cum Pct
1166	17	17	0.02	0.02
1247	37	54	0.04	0.05
1295	98	152	0.10	0.15
1326	258	410	0.26	0.41
1349	592	1002	0.59	1.01
1367	1075	2077	1.08	2.08
1384	1551	3628	1.56	3.64
1398	1865	5493	1.87	5.51
1412	2121	7614	2.13	7.64
1425	2228	9842	2.24	9.87
1437	2215	12057	2.22	12.10
1448	2390	14447	2.40	14.49
1460	2524	16971	2.53	17.02
1471	2682	19653	2.69	19.72
1482	2954	22607	2.96	22.68
1493	3294	25901	3.30	25.98
1504	3806	29707	3.82	29.80
1516	4045	33752	4.06	33.86
1528	4383	38135	4.40	38.26
1541	5060	43195	5.08	43.33
1554	5732	48927	5.75	49.08
1569	6393	55320	6.41	55.50
1585	7233	62553	7.26	62.75
1604	7977	70530	8.00	70.75
1626	8749	79279	8.78	79.53
1657	8473	87752	8.50	88.03
1705	7498	95250	7.52	95.55
1786	4433	99683	4.45	100.00

Figure D.5.5. 2025 TELPAS Grade 6 Listening
Frequency Distribution of Scale Scores



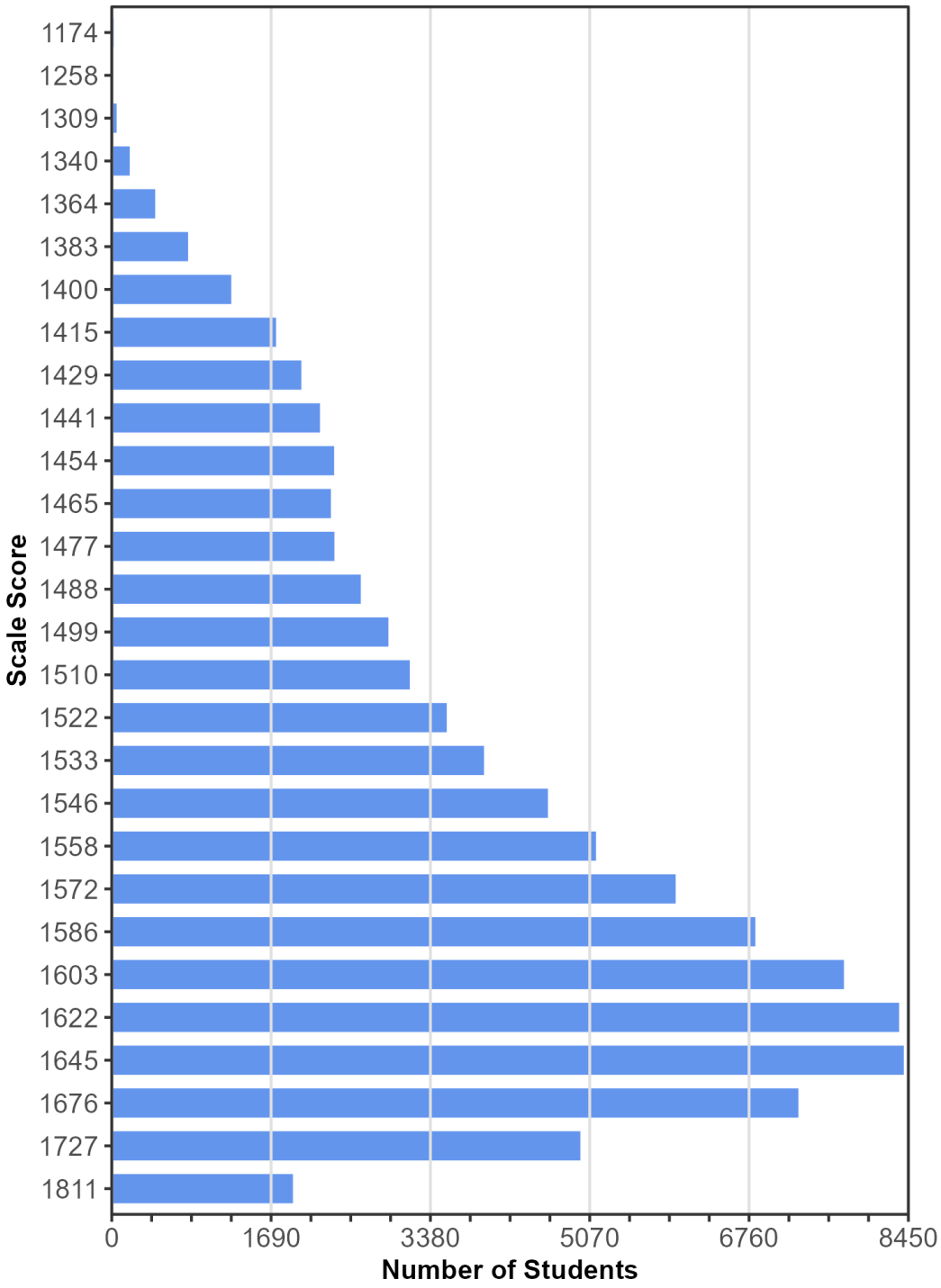
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1174	15	15	0.02	0.02
1258	32	47	0.03	0.05
1309	94	141	0.10	0.15
1340	267	408	0.29	0.44
1364	658	1066	0.71	1.16
1383	1164	2230	1.26	2.42
1400	1726	3956	1.87	4.30
1415	2251	6207	2.44	6.74
1429	2686	8893	2.92	9.66
1441	2947	11840	3.20	12.86
1454	3198	15038	3.47	16.33
1465	3452	18490	3.75	20.08
1477	3717	22207	4.04	24.12
1488	3931	26138	4.27	28.39
1499	4262	30400	4.63	33.01
1510	4745	35145	5.15	38.17
1522	4943	40088	5.37	43.54
1533	5313	45401	5.77	49.30
1546	5838	51239	6.34	55.64
1558	6212	57451	6.75	62.39
1572	6379	63830	6.93	69.32
1586	6416	70246	6.97	76.29
1603	6134	76380	6.66	82.95
1622	5371	81751	5.83	88.78
1645	4616	86367	5.01	93.79
1676	3337	89704	3.62	97.42
1727	1783	91487	1.94	99.35
1811	595	92082	0.65	100.00

Figure D.5.6. 2025 TELPAS Grade 7 Listening
Frequency Distribution of Scale Scores



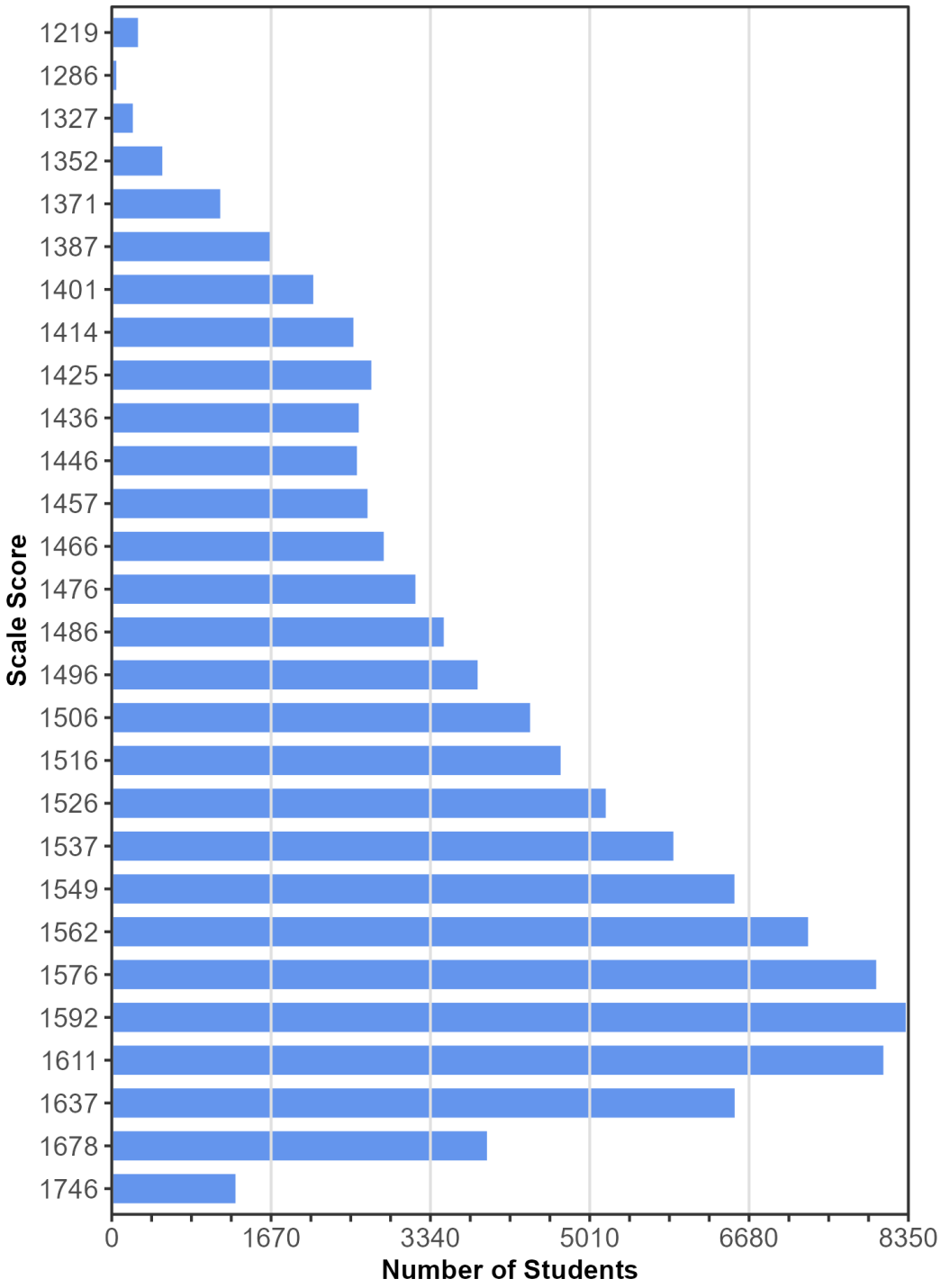
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1174	29	29	0.03	0.03
1258	38	67	0.04	0.07
1309	82	149	0.08	0.15
1340	231	380	0.24	0.39
1364	548	928	0.56	0.96
1383	1077	2005	1.11	2.07
1400	1512	3517	1.56	3.62
1415	2053	5570	2.12	5.74
1429	2359	7929	2.43	8.17
1441	2488	10417	2.56	10.73
1454	2636	13053	2.72	13.45
1465	2746	15799	2.83	16.28
1477	2889	18688	2.98	19.26
1488	3101	21789	3.20	22.45
1499	3529	25318	3.64	26.09
1510	3864	29182	3.98	30.07
1522	4199	33381	4.33	34.40
1533	4790	38171	4.94	39.33
1546	5337	43508	5.50	44.83
1558	5930	49438	6.11	50.95
1572	6531	55969	6.73	57.68
1586	7311	63280	7.53	65.21
1603	7500	70780	7.73	72.94
1622	7799	78579	8.04	80.98
1645	7390	85969	7.62	88.59
1676	5920	91889	6.10	94.69
1727	3735	95624	3.85	98.54
1811	1417	97041	1.46	100.00

Figure D.5.7. 2025 TELPAS Grade 8 Listening
Frequency Distribution of Scale Scores



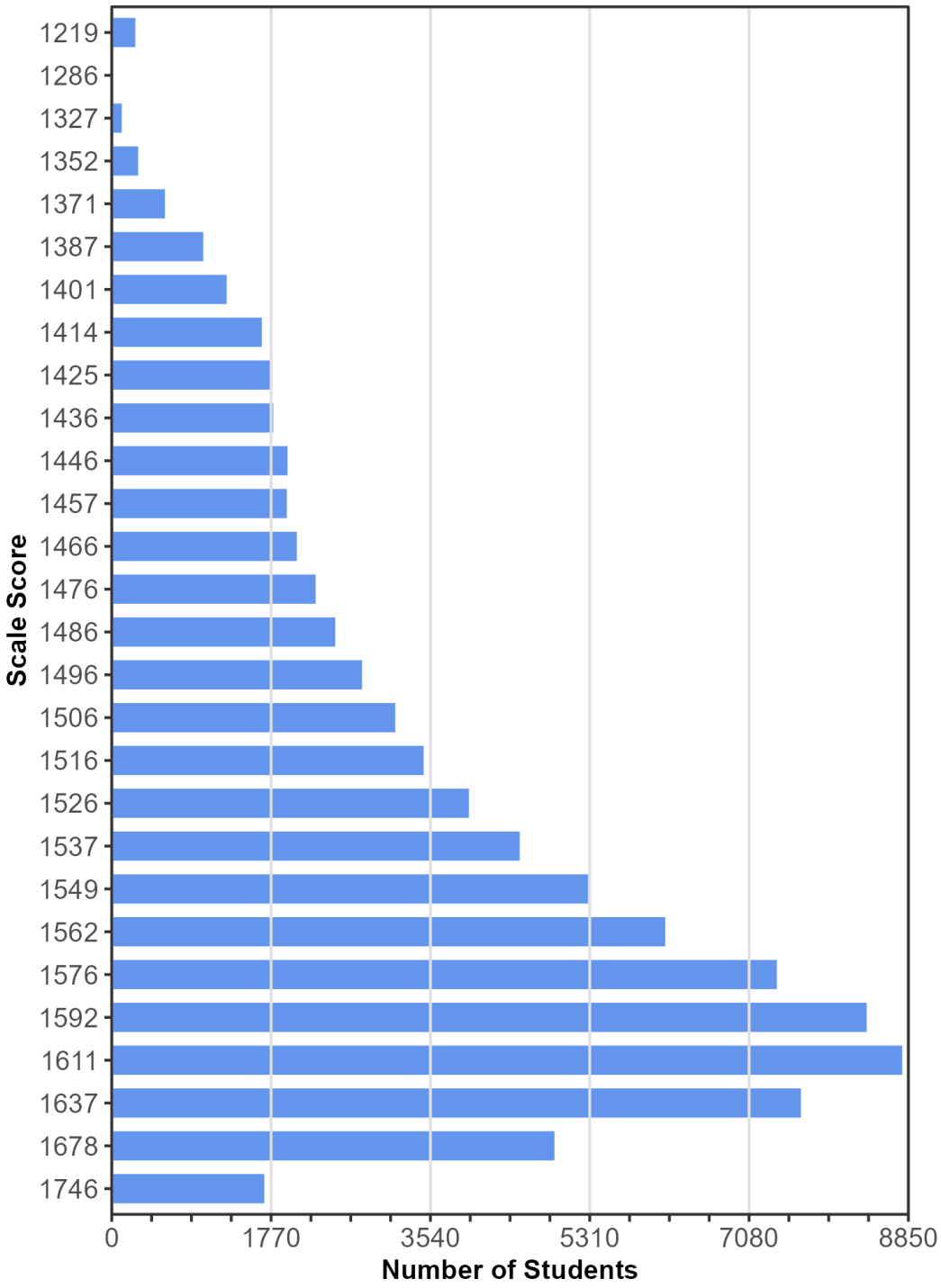
	Freq	Cum Freq	Pct	Cum Pct
1174	36	36	0.04	0.04
1258	21	57	0.02	0.06
1309	67	124	0.07	0.13
1340	206	330	0.22	0.35
1364	476	806	0.51	0.86
1383	825	1631	0.88	1.74
1400	1283	2914	1.37	3.11
1415	1758	4672	1.88	4.98
1429	2026	6698	2.16	7.14
1441	2224	8922	2.37	9.52
1454	2374	11296	2.53	12.05
1465	2340	13636	2.50	14.54
1477	2377	16013	2.54	17.08
1488	2657	18670	2.83	19.91
1499	2950	21620	3.15	23.06
1510	3177	24797	3.39	26.45
1522	3569	28366	3.81	30.26
1533	3964	32330	4.23	34.48
1546	4642	36972	4.95	39.44
1558	5152	42124	5.50	44.93
1572	5997	48121	6.40	51.33
1586	6842	54963	7.30	58.63
1603	7782	62745	8.30	66.93
1622	8367	71112	8.92	75.85
1645	8417	79529	8.98	84.83
1676	7299	86828	7.79	92.62
1727	4986	91814	5.32	97.93
1811	1937	93751	2.07	100.00

Figure D.5.8. 2025 TELPAS Grade 9 Listening
Frequency Distribution of Scale Scores



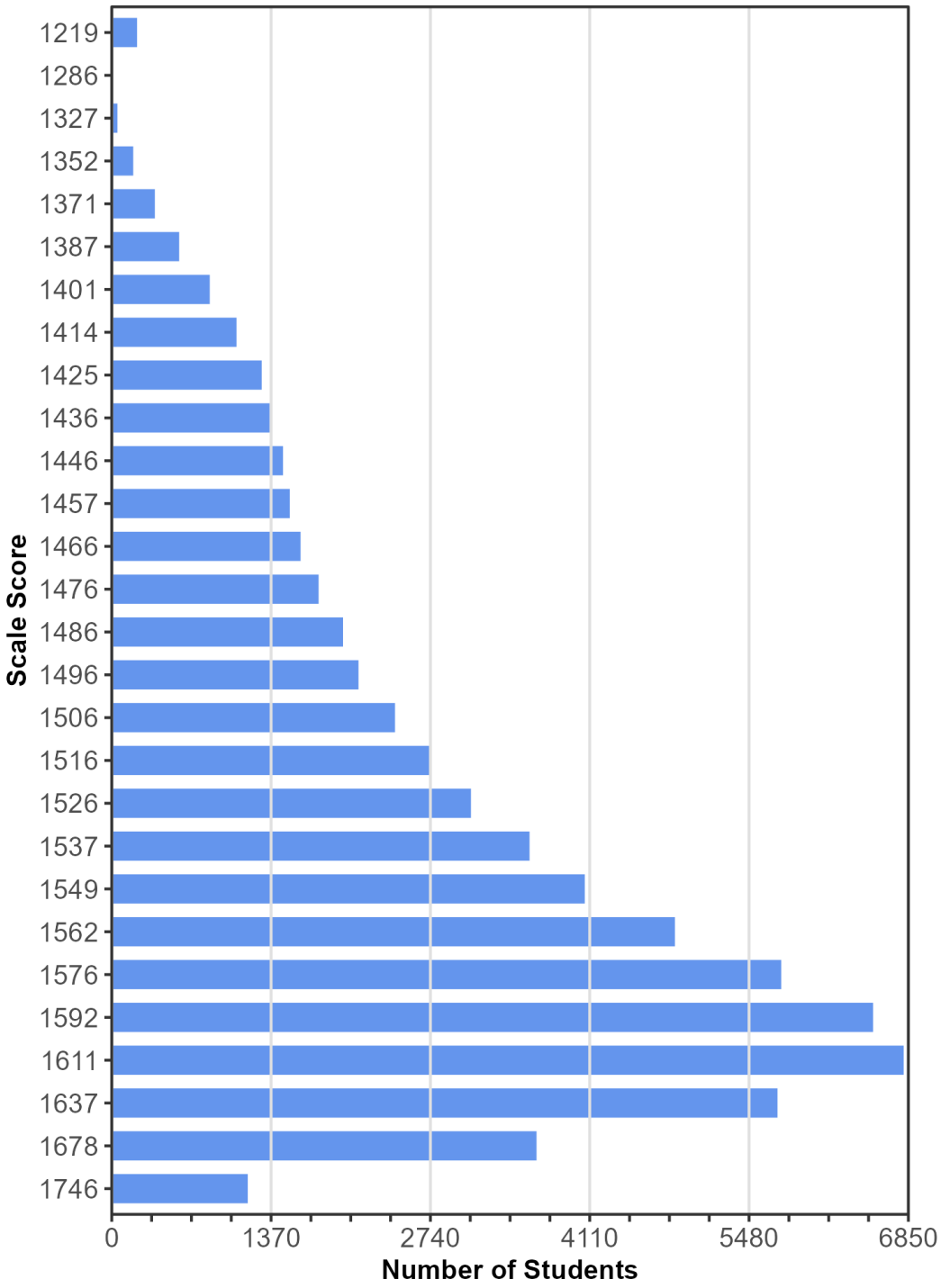
	Freq	Cum Freq	Pct	Cum Pct
1219	290	290	0.28	0.28
1286	63	353	0.06	0.34
1327	236	589	0.23	0.57
1352	545	1134	0.53	1.10
1371	1152	2286	1.12	2.22
1387	1674	3960	1.63	3.84
1401	2127	6087	2.06	5.91
1414	2548	8635	2.47	8.38
1425	2736	11371	2.66	11.04
1436	2604	13975	2.53	13.57
1446	2585	16560	2.51	16.08
1457	2696	19256	2.62	18.69
1466	2866	22122	2.78	21.48
1476	3198	25320	3.10	24.58
1486	3495	28815	3.39	27.97
1496	3850	32665	3.74	31.71
1506	4400	37065	4.27	35.98
1516	4720	41785	4.58	40.56
1526	5193	46978	5.04	45.61
1537	5902	52880	5.73	51.34
1549	6542	59422	6.35	57.69
1562	7314	66736	7.10	64.79
1576	8028	74764	7.79	72.58
1592	8337	83101	8.09	80.67
1611	8103	91204	7.87	88.54
1637	6544	97748	6.35	94.89
1678	3948	101696	3.83	98.73
1746	1312	103008	1.27	100.00

Figure D.5.9. 2025 TELPAS Grade 10 Listening
Frequency Distribution of Scale Scores



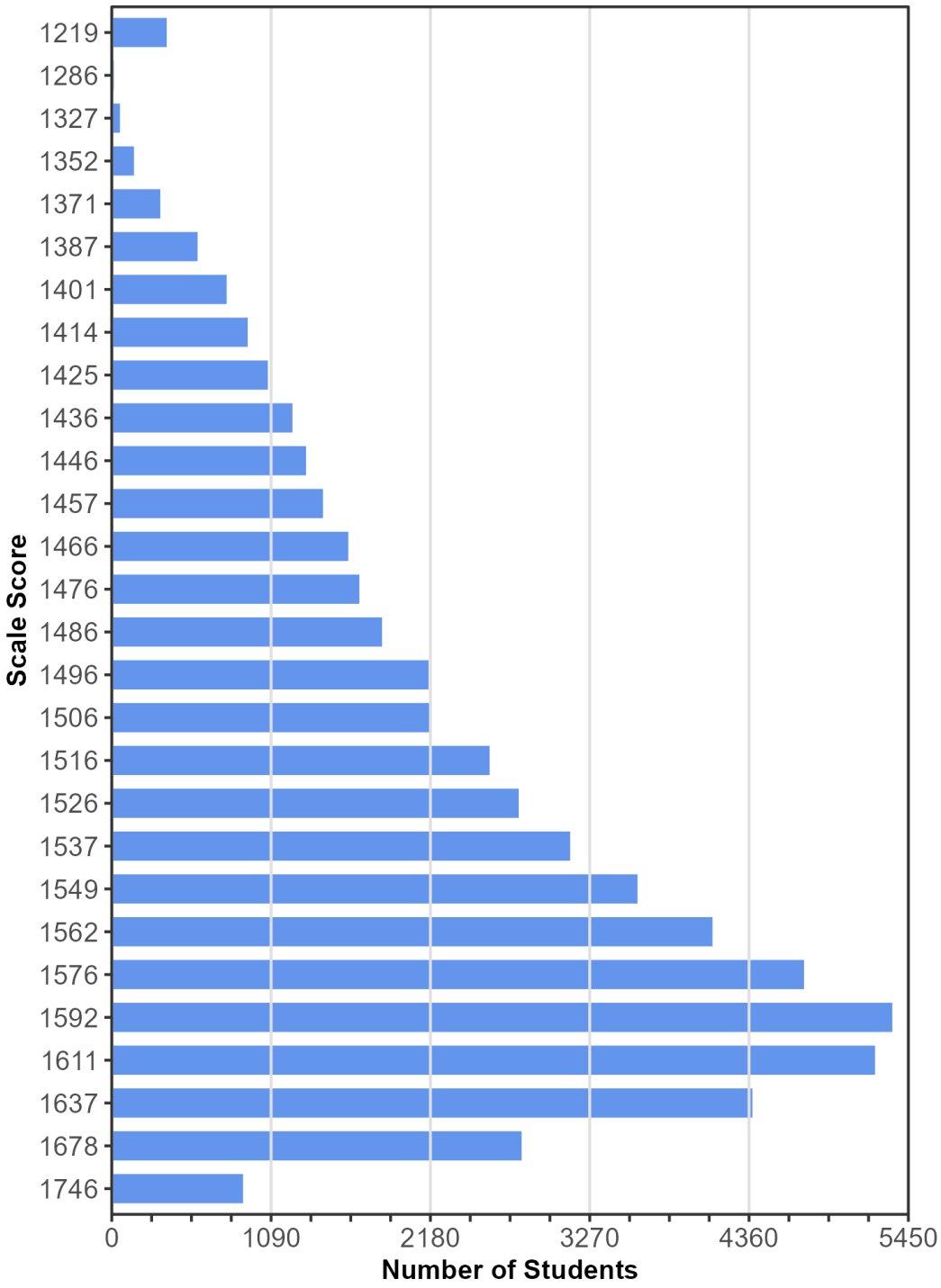
	Freq	Cum Freq	Pct	Cum Pct
1219	278	278	0.32	0.32
1286	29	307	0.03	0.35
1327	127	434	0.14	0.49
1352	309	743	0.35	0.84
1371	607	1350	0.69	1.53
1387	1033	2383	1.17	2.70
1401	1293	3676	1.47	4.17
1414	1683	5359	1.91	6.08
1425	1793	7152	2.03	8.12
1436	1810	8962	2.05	10.17
1446	1969	10931	2.23	12.41
1457	1961	12892	2.23	14.63
1466	2072	14964	2.35	16.98
1476	2281	17245	2.59	19.57
1486	2499	19744	2.84	22.41
1496	2797	22541	3.17	25.58
1506	3166	25707	3.59	29.17
1516	3481	29188	3.95	33.12
1526	3984	33172	4.52	37.65
1537	4548	37720	5.16	42.81
1549	5307	43027	6.02	48.83
1562	6165	49192	7.00	55.83
1576	7405	56597	8.40	64.23
1592	8403	65000	9.54	73.77
1611	8798	73798	9.98	83.75
1637	7672	81470	8.71	92.46
1678	4934	86404	5.60	98.06
1746	1711	88115	1.94	100.00

Figure D.5.10. 2025 TELPAS Grade 11 Listening
Frequency Distribution of Scale Scores



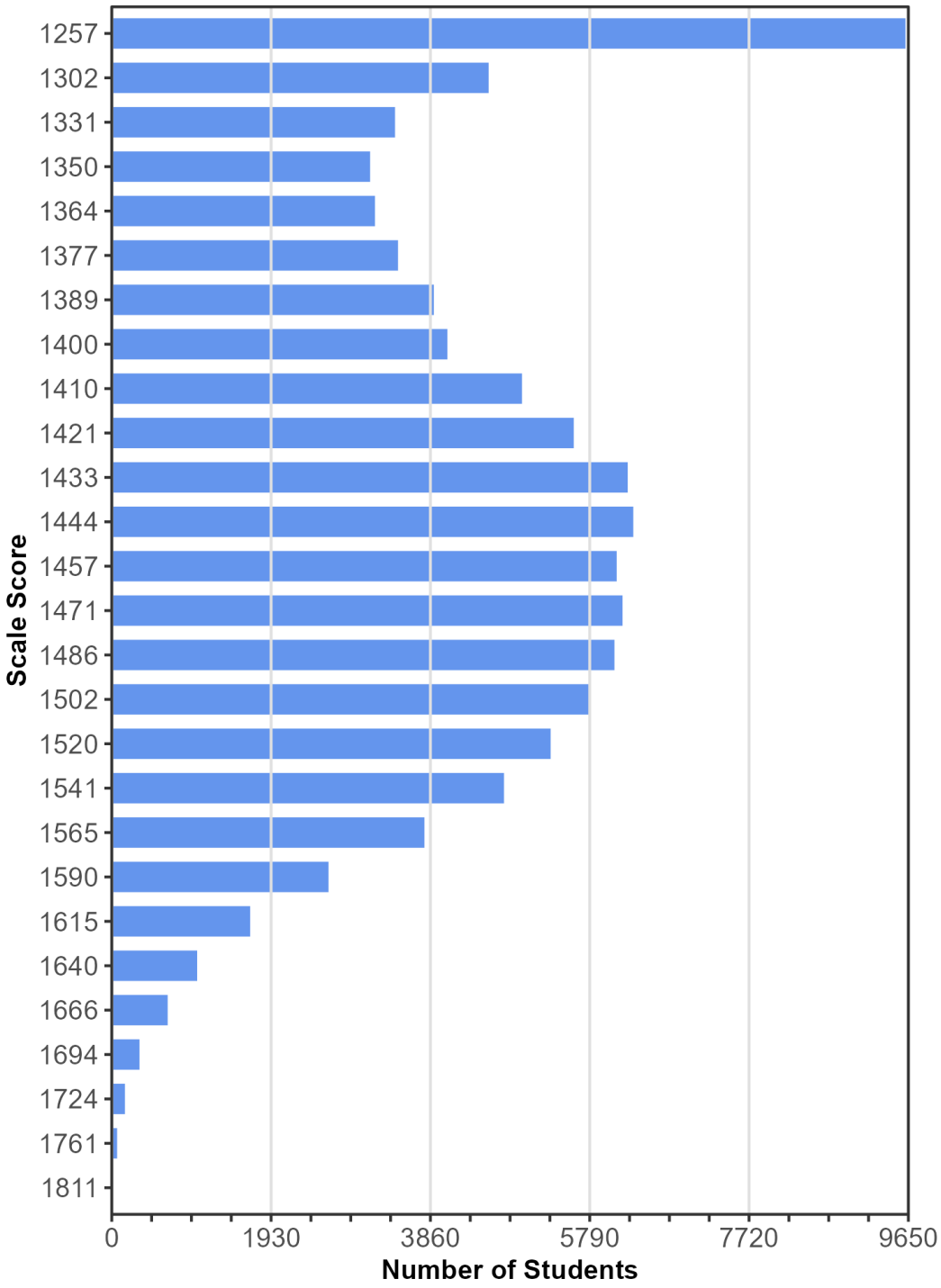
	Freq	Cum Freq	Pct	Cum Pct
1219	230	230	0.34	0.34
1286	22	252	0.03	0.37
1327	61	313	0.09	0.47
1352	197	510	0.29	0.76
1371	383	893	0.57	1.33
1387	592	1485	0.88	2.21
1401	855	2340	1.27	3.48
1414	1085	3425	1.61	5.09
1425	1302	4727	1.94	7.03
1436	1370	6097	2.04	9.07
1446	1485	7582	2.21	11.27
1457	1543	9125	2.29	13.57
1466	1636	10761	2.43	16.00
1476	1791	12552	2.66	18.66
1486	2001	14553	2.98	21.64
1496	2134	16687	3.17	24.81
1506	2448	19135	3.64	28.45
1516	2744	21879	4.08	32.53
1526	3101	24980	4.61	37.14
1537	3605	28585	5.36	42.50
1549	4080	32665	6.07	48.57
1562	4855	37520	7.22	55.79
1576	5769	43289	8.58	64.37
1592	6559	49848	9.75	74.12
1611	6822	56670	10.14	84.26
1637	5737	62407	8.53	92.79
1678	3665	66072	5.45	98.24
1746	1182	67254	1.76	100.00

Figure D.5.11. 2025 TELPAS Grade 12 Listening
Frequency Distribution of Scale Scores



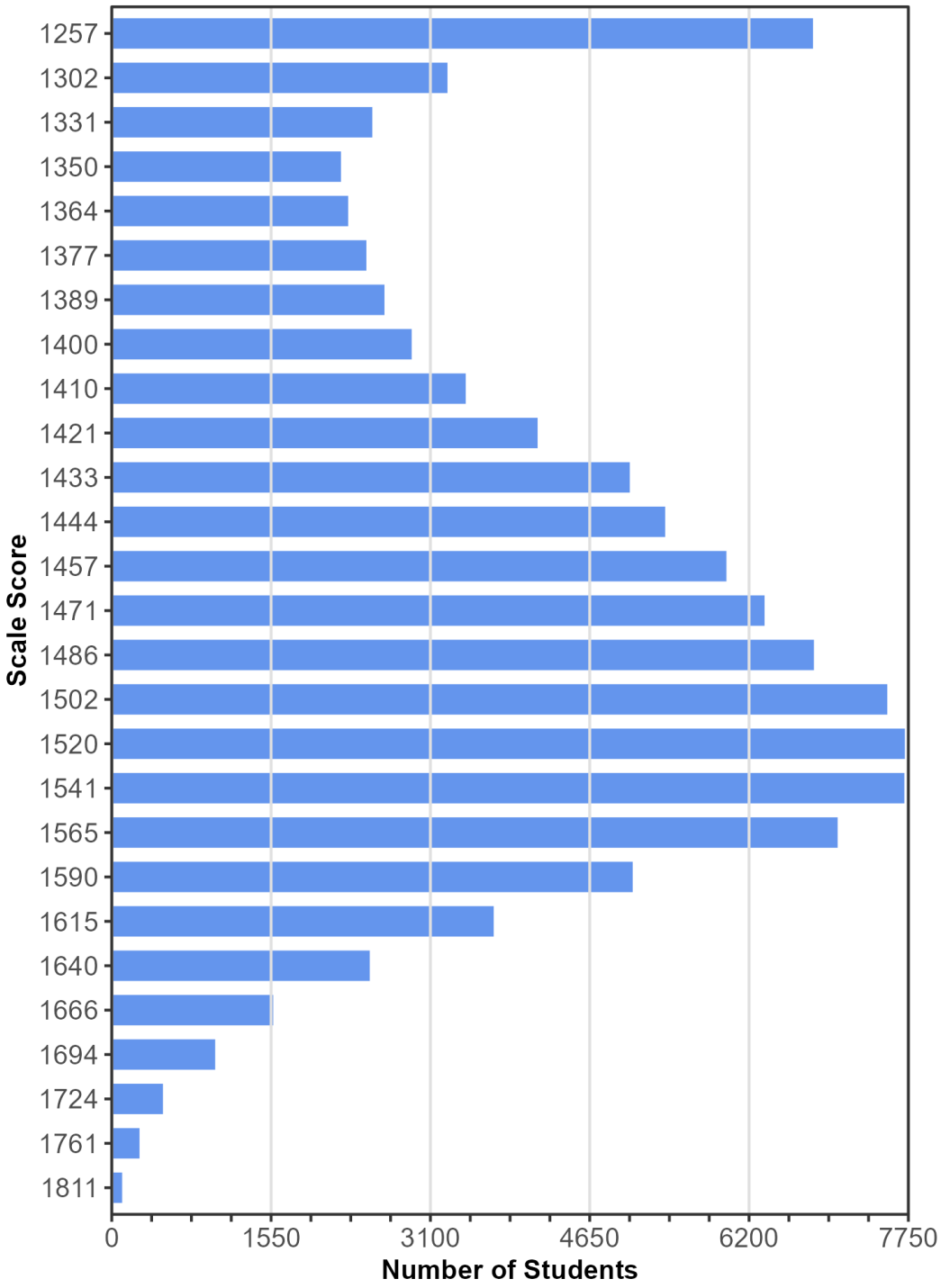
	Freq	Cum Freq	Pct	Cum Pct
1219	386	386	0.67	0.67
1286	24	410	0.04	0.71
1327	66	476	0.11	0.82
1352	162	638	0.28	1.11
1371	342	980	0.59	1.70
1387	597	1577	1.03	2.73
1401	796	2373	1.38	4.11
1414	940	3313	1.63	5.74
1425	1078	4391	1.87	7.61
1436	1246	5637	2.16	9.77
1446	1339	6976	2.32	12.09
1457	1455	8431	2.52	14.61
1466	1628	10059	2.82	17.43
1476	1704	11763	2.95	20.39
1486	1859	13622	3.22	23.61
1496	2177	15799	3.77	27.38
1506	2200	17999	3.81	31.19
1516	2595	20594	4.50	35.69
1526	2794	23388	4.84	40.53
1537	3146	26534	5.45	45.98
1549	3607	30141	6.25	52.23
1562	4120	34261	7.14	59.37
1576	4746	39007	8.22	67.60
1592	5350	44357	9.27	76.87
1611	5232	49589	9.07	85.94
1637	4393	53982	7.61	93.55
1678	2814	56796	4.88	98.43
1746	908	57704	1.57	100.00

Figure D.5.12. 2025 TELPAS Grade 2 Speaking
Frequency Distribution of Scale Scores



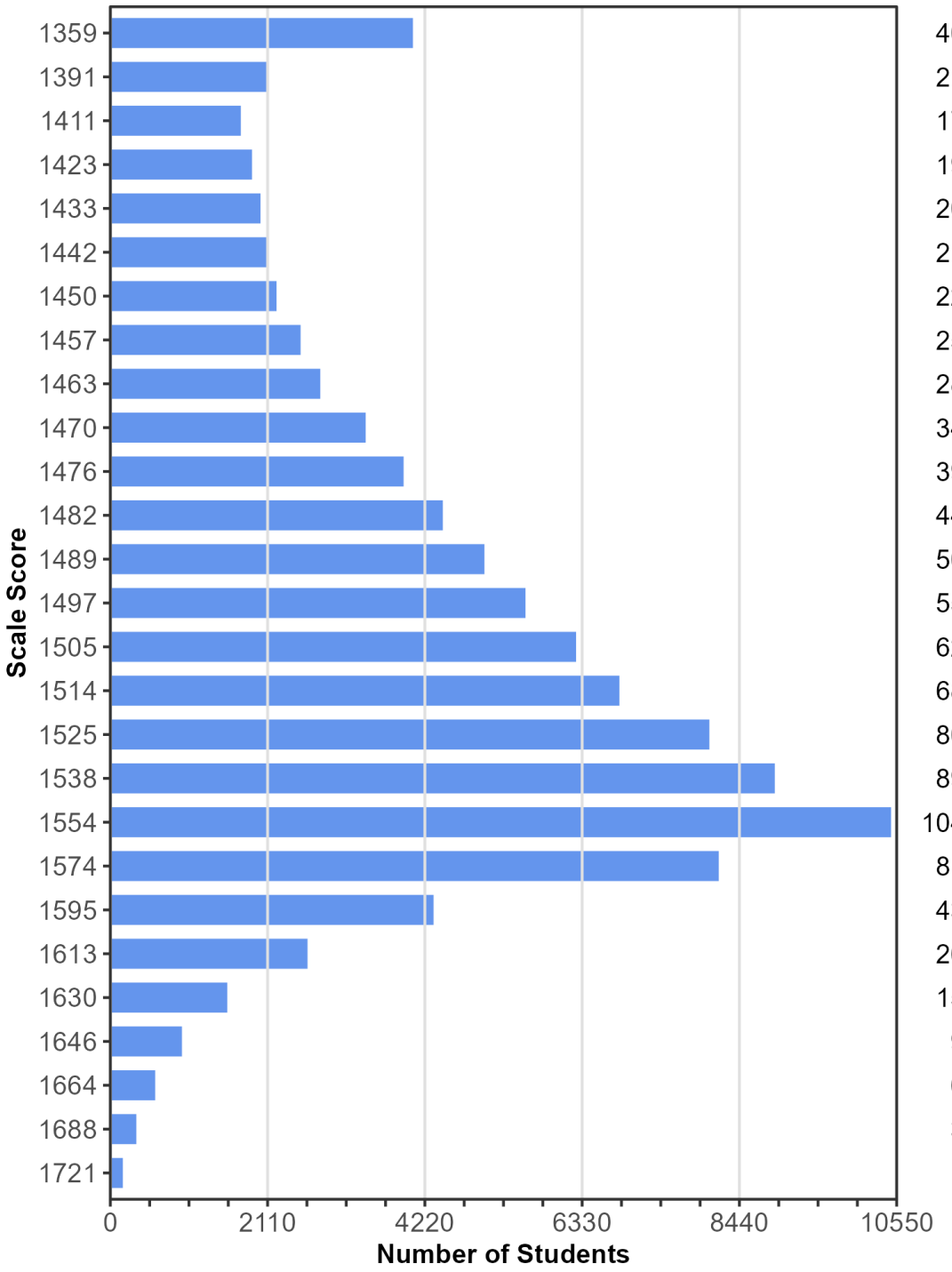
	Freq	Cum Freq	Pct	Cum Pct
1257	9632	9632	9.30	9.30
1302	4583	14215	4.42	13.72
1331	3449	17664	3.33	17.05
1350	3147	20811	3.04	20.09
1364	3206	24017	3.09	23.18
1377	3485	27502	3.36	26.54
1389	3919	31421	3.78	30.33
1400	4083	35504	3.94	34.27
1410	4987	40491	4.81	39.08
1421	5614	46105	5.42	44.50
1433	6268	52373	6.05	50.55
1444	6335	58708	6.11	56.66
1457	6134	64842	5.92	62.58
1471	6205	71047	5.99	68.57
1486	6107	77154	5.89	74.47
1502	5816	82970	5.61	80.08
1520	5334	88304	5.15	85.23
1541	4769	93073	4.60	89.83
1565	3805	96878	3.67	93.51
1590	2643	99521	2.55	96.06
1615	1694	101215	1.64	97.69
1640	1052	102267	1.02	98.71
1666	695	102962	0.67	99.38
1694	353	103315	0.34	99.72
1724	177	103492	0.17	99.89
1761	82	103574	0.08	99.97
1811	33	103607	0.03	100.00

Figure D.5.13. 2025 TELPAS Grade 3 Speaking
Frequency Distribution of Scale Scores



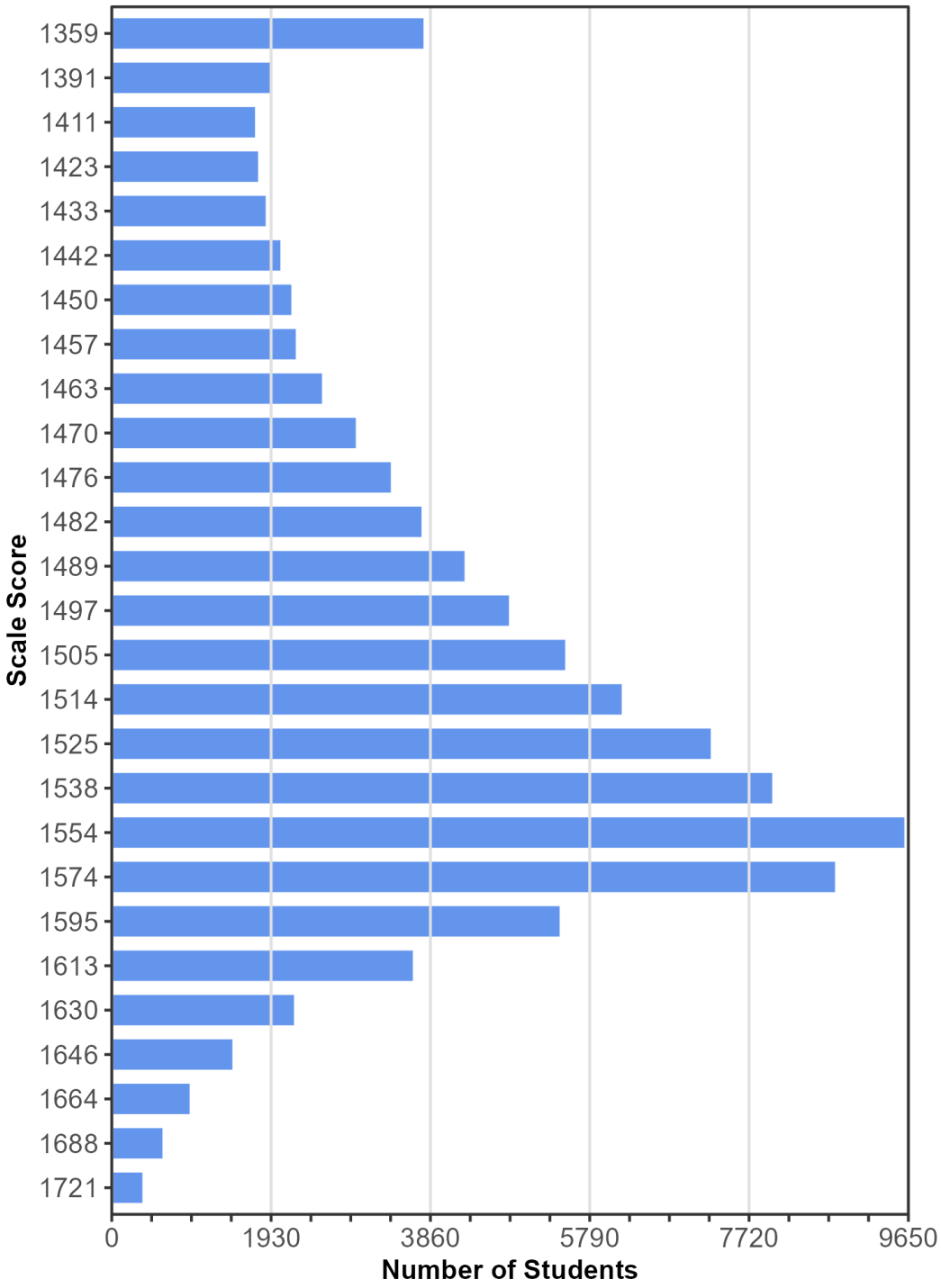
	Freq	Cum Freq	Pct	Cum Pct
1257	6836	6836	6.36	6.36
1302	3280	10116	3.05	9.41
1331	2549	12665	2.37	11.78
1350	2244	14909	2.09	13.87
1364	2314	17223	2.15	16.02
1377	2491	19714	2.32	18.33
1389	2667	22381	2.48	20.81
1400	2932	25313	2.73	23.54
1410	3458	28771	3.22	26.76
1421	4157	32928	3.87	30.62
1433	5054	37982	4.70	35.32
1444	5398	43380	5.02	40.34
1457	5994	49374	5.57	45.92
1471	6365	55739	5.92	51.84
1486	6845	62584	6.37	58.20
1502	7559	70143	7.03	65.23
1520	7730	77873	7.19	72.42
1541	7725	85598	7.18	79.61
1565	7076	92674	6.58	86.19
1590	5082	97756	4.73	90.91
1615	3730	101486	3.47	94.38
1640	2524	104010	2.35	96.73
1666	1585	105595	1.47	98.20
1694	1020	106615	0.95	99.15
1724	512	107127	0.48	99.63
1761	284	107411	0.26	99.89
1811	115	107526	0.11	100.00

Figure D.5.14. 2025 TELPAS Grade 4 Speaking
Frequency Distribution of Scale Scores



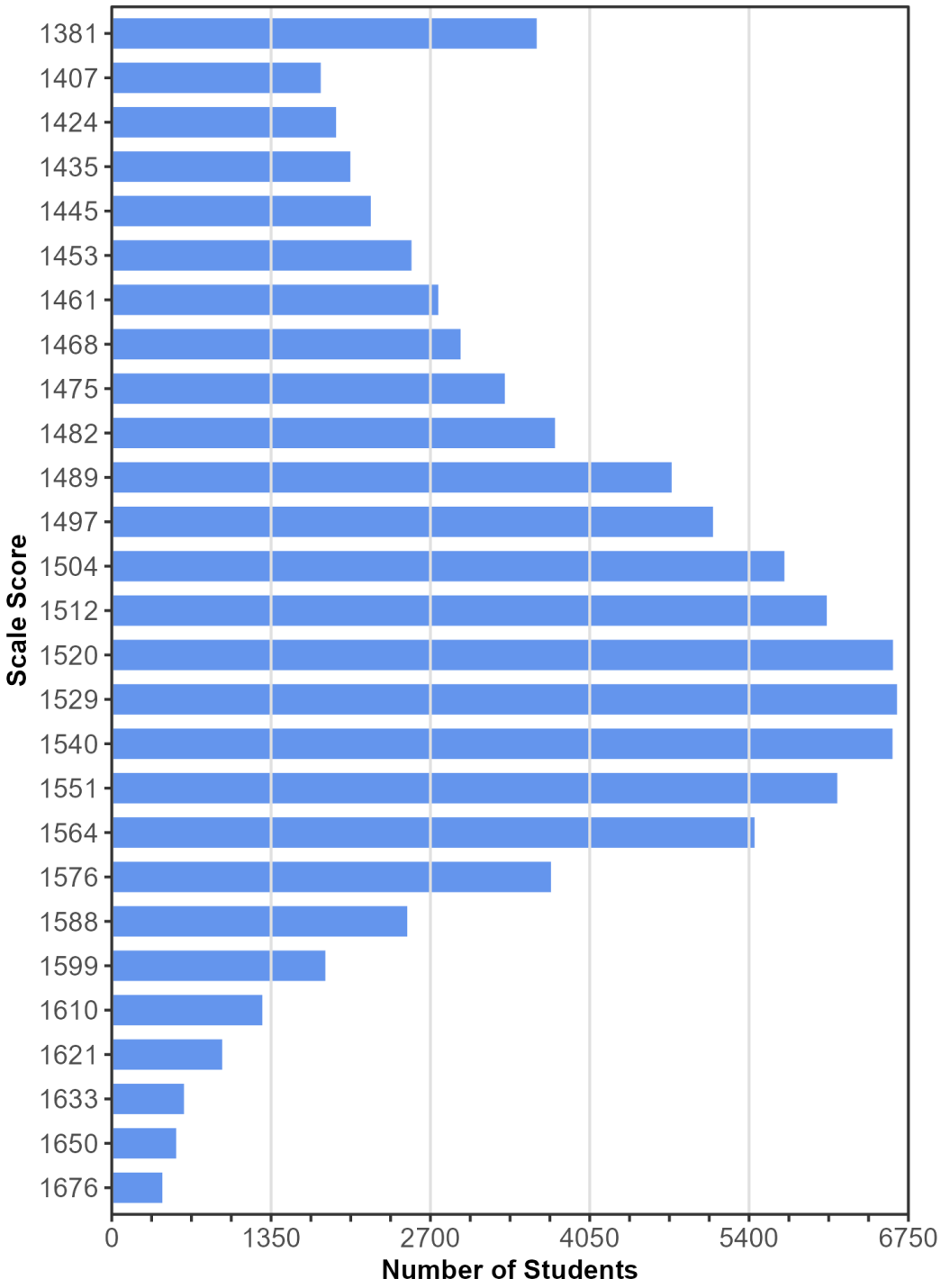
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1359	4079	4079	3.93	3.93
1391	2136	6215	2.06	5.99
1411	1770	7985	1.71	7.69
1423	1920	9905	1.85	9.54
1433	2034	11939	1.96	11.50
1442	2123	14062	2.05	13.55
1450	2249	16311	2.17	15.72
1457	2572	18883	2.48	18.20
1463	2836	21719	2.73	20.93
1470	3445	25164	3.32	24.25
1476	3954	29118	3.81	28.06
1482	4480	33598	4.32	32.38
1489	5038	38636	4.85	37.23
1497	5589	44225	5.39	42.62
1505	6268	50493	6.04	48.66
1514	6849	57342	6.60	55.26
1525	8056	65398	7.76	63.02
1538	8934	74332	8.61	71.63
1554	10494	84826	10.11	81.74
1574	8182	93008	7.88	89.63
1595	4356	97364	4.20	93.82
1613	2665	100029	2.57	96.39
1630	1588	101617	1.53	97.92
1646	980	102597	0.94	98.87
1664	621	103218	0.60	99.47
1688	368	103586	0.35	99.82
1721	187	103773	0.18	100.00

Figure D.5.15. 2025 TELPAS Grade 5 Speaking
Frequency Distribution of Scale Scores



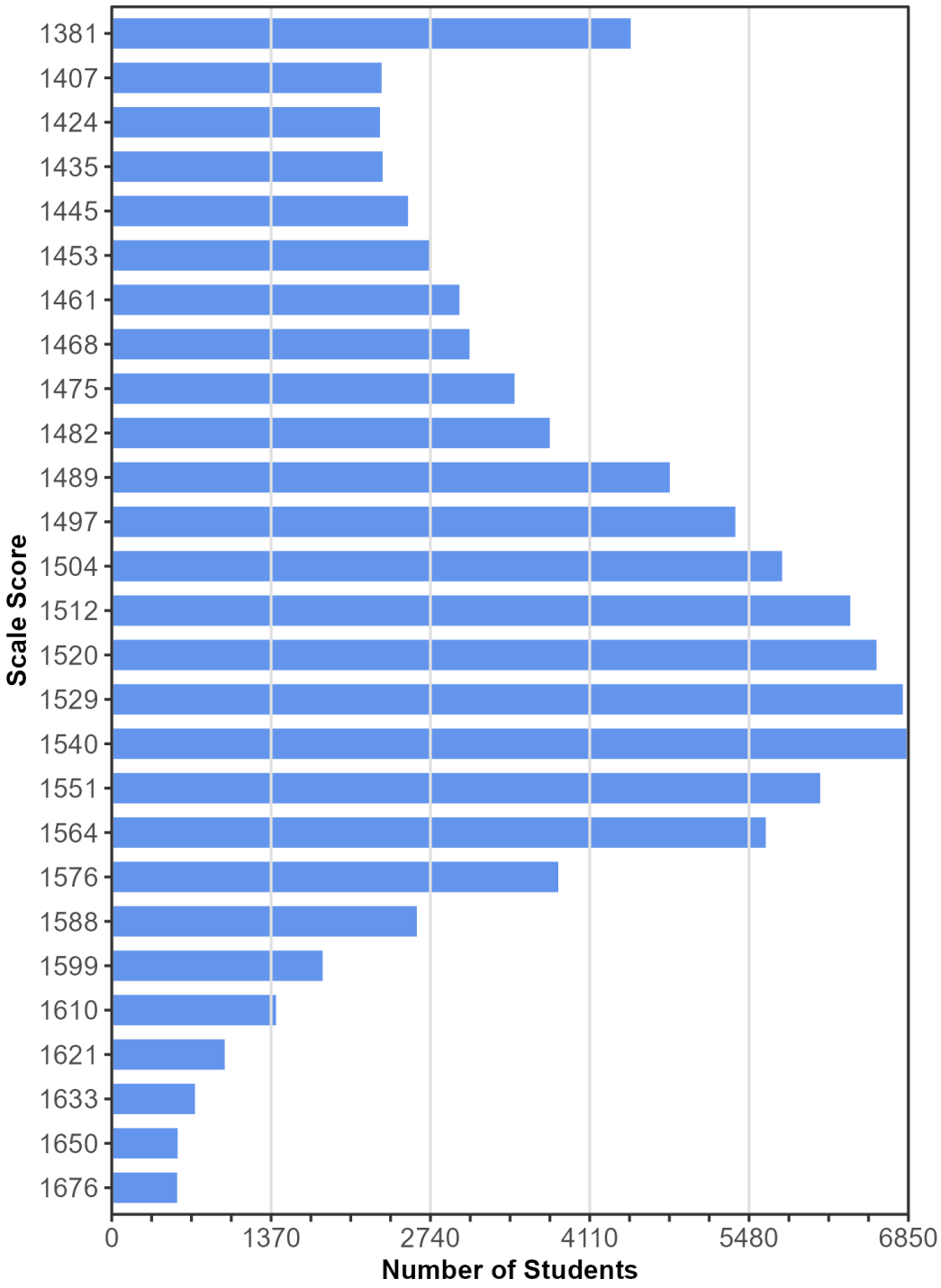
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1359	3794	3794	3.81	3.81
1391	1949	5743	1.96	5.76
1411	1753	7496	1.76	7.52
1423	1790	9286	1.80	9.32
1433	1882	11168	1.89	11.20
1442	2060	13228	2.07	13.27
1450	2193	15421	2.20	15.47
1457	2247	17668	2.25	17.72
1463	2564	20232	2.57	20.30
1470	2974	23206	2.98	23.28
1476	3398	26604	3.41	26.69
1482	3770	30374	3.78	30.47
1489	4291	34665	4.30	34.78
1497	4829	39494	4.84	39.62
1505	5510	45004	5.53	45.15
1514	6195	51199	6.21	51.36
1525	7274	58473	7.30	58.66
1538	8018	66491	8.04	66.70
1554	9619	76110	9.65	76.35
1574	8779	84889	8.81	85.16
1595	5443	90332	5.46	90.62
1613	3665	93997	3.68	94.30
1630	2225	96222	2.23	96.53
1646	1479	97701	1.48	98.01
1664	961	98662	0.96	98.98
1688	632	99294	0.63	99.61
1721	389	99683	0.39	100.00

Figure D.5.16. 2025 TELPAS Grade 6 Speaking
Frequency Distribution of Scale Scores



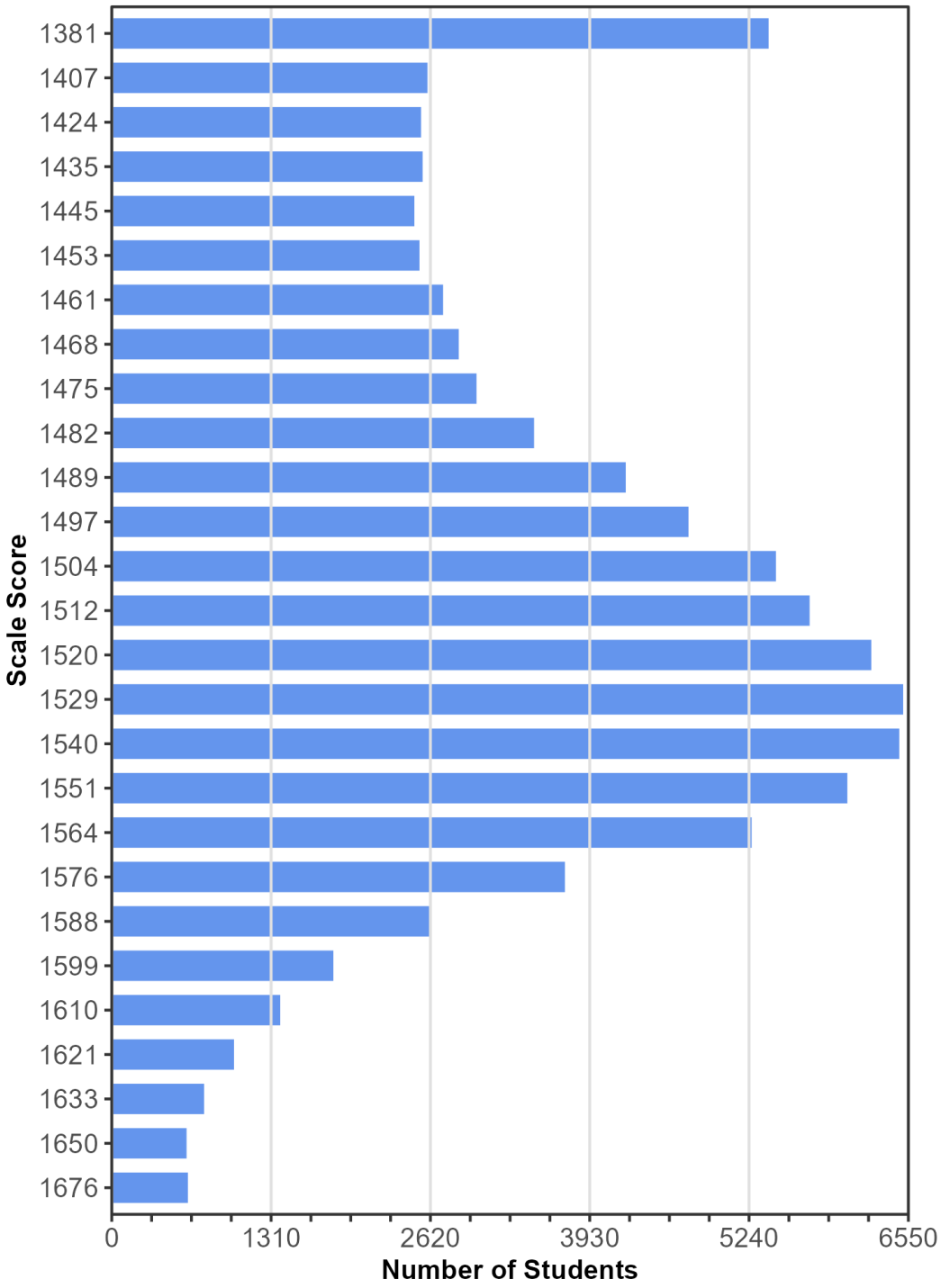
	Freq	Cum Freq	Pct	Cum Pct
1381	3613	3613	3.92	3.92
1407	1783	5396	1.94	5.86
1424	1913	7309	2.08	7.94
1435	2034	9343	2.21	10.15
1445	2207	11550	2.40	12.54
1453	2552	14102	2.77	15.31
1461	2779	16881	3.02	18.33
1468	2968	19849	3.22	21.56
1475	3343	23192	3.63	25.19
1482	3768	26960	4.09	29.28
1489	4756	31716	5.16	34.44
1497	5107	36823	5.55	39.99
1504	5712	42535	6.20	46.19
1512	6071	48606	6.59	52.79
1520	6632	55238	7.20	59.99
1529	6666	61904	7.24	67.23
1540	6628	68532	7.20	74.42
1551	6160	74692	6.69	81.11
1564	5459	80151	5.93	87.04
1576	3734	83885	4.06	91.10
1588	2516	86401	2.73	93.83
1599	1822	88223	1.98	95.81
1610	1288	89511	1.40	97.21
1621	948	90459	1.03	98.24
1633	624	91083	0.68	98.92
1650	558	91641	0.61	99.52
1676	441	92082	0.48	100.00

Figure D.5.17. 2025 TELPAS Grade 7 Speaking
Frequency Distribution of Scale Scores



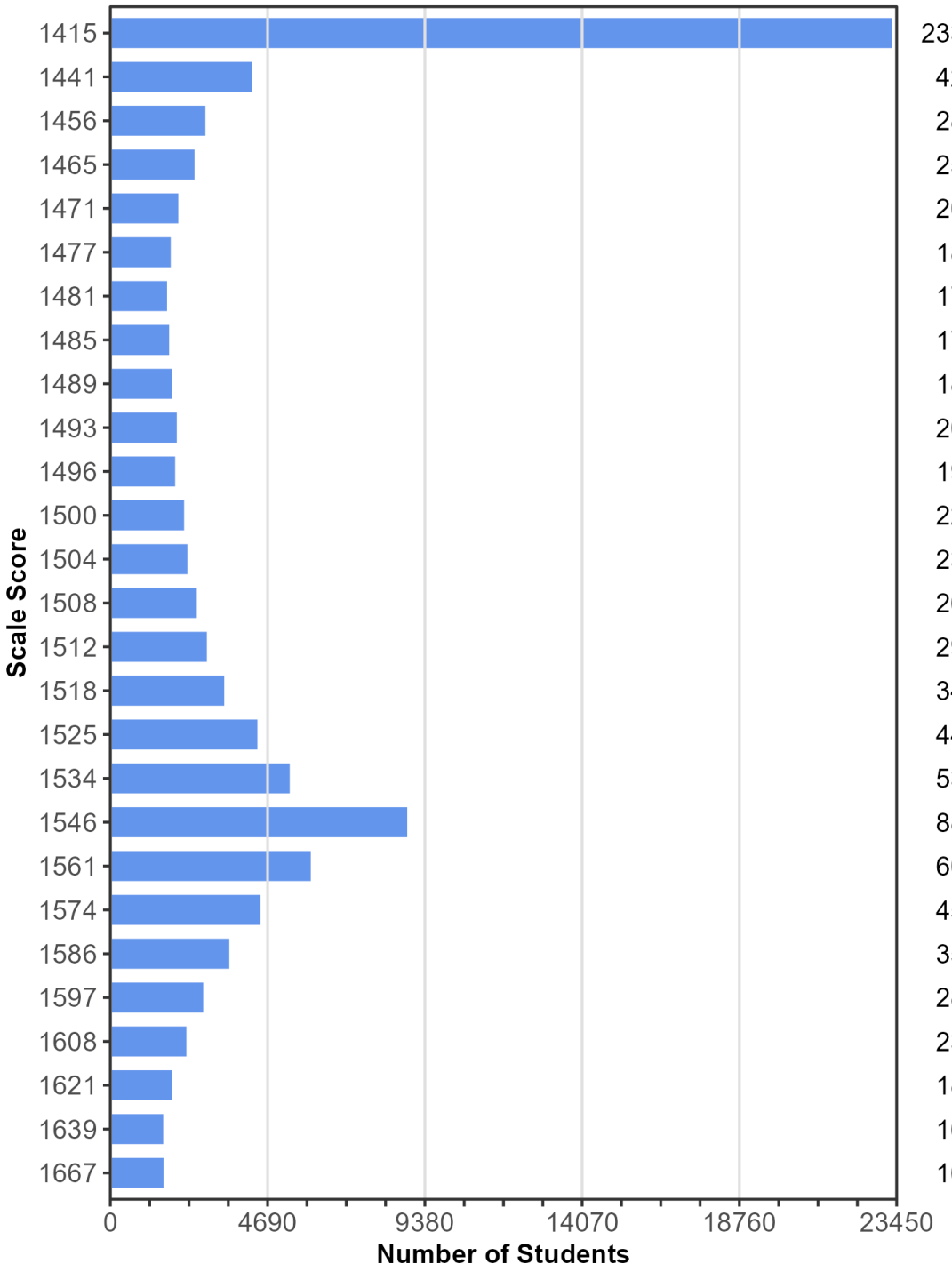
	Freq	Cum Freq	Pct	Cum Pct
1381	4475	4475	4.61	4.61
1407	2333	6808	2.40	7.02
1424	2319	9127	2.39	9.41
1435	2342	11469	2.41	11.82
1445	2560	14029	2.64	14.46
1453	2744	16773	2.83	17.28
1461	3002	19775	3.09	20.38
1468	3089	22864	3.18	23.56
1475	3476	26340	3.58	27.14
1482	3779	30119	3.89	31.04
1489	4811	34930	4.96	36.00
1497	5375	40305	5.54	41.53
1504	5777	46082	5.95	47.49
1512	6363	52445	6.56	54.04
1520	6589	59034	6.79	60.83
1529	6815	65849	7.02	67.86
1540	6847	72696	7.06	74.91
1551	6104	78800	6.29	81.20
1564	5636	84436	5.81	87.01
1576	3852	88288	3.97	90.98
1588	2636	90924	2.72	93.70
1599	1826	92750	1.88	95.58
1610	1425	94175	1.47	97.05
1621	984	95159	1.01	98.06
1633	728	95887	0.75	98.81
1650	579	96466	0.60	99.41
1676	575	97041	0.59	100.00

Figure D.5.18. 2025 TELPAS Grade 8 Speaking
Frequency Distribution of Scale Scores



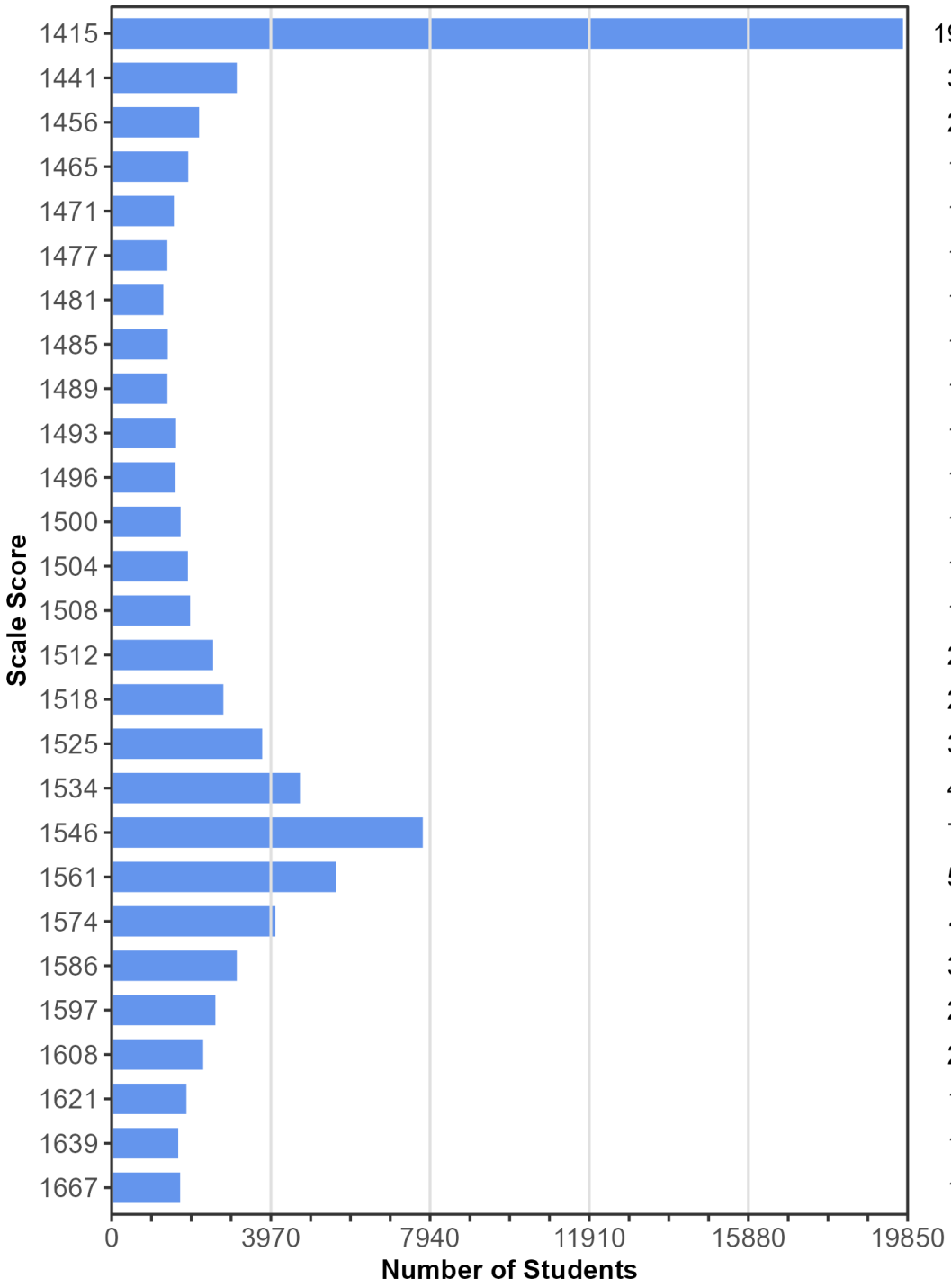
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1381	5413	5413	5.77	5.77
1407	2608	8021	2.78	8.56
1424	2555	10576	2.73	11.28
1435	2568	13144	2.74	14.02
1445	2500	15644	2.67	16.69
1453	2542	18186	2.71	19.40
1461	2736	20922	2.92	22.32
1468	2865	23787	3.06	25.37
1475	3011	26798	3.21	28.58
1482	3484	30282	3.72	32.30
1489	4238	34520	4.52	36.82
1497	4754	39274	5.07	41.89
1504	5473	44747	5.84	47.73
1512	5750	50497	6.13	53.86
1520	6257	56754	6.67	60.54
1529	6519	63273	6.95	67.49
1540	6487	69760	6.92	74.41
1551	6060	75820	6.46	80.87
1564	5273	81093	5.62	86.50
1576	3738	84831	3.99	90.49
1588	2636	87467	2.81	93.30
1599	1834	89301	1.96	95.25
1610	1397	90698	1.49	96.74
1621	1017	91715	1.08	97.83
1633	771	92486	0.82	98.65
1650	627	93113	0.67	99.32
1676	638	93751	0.68	100.00

Figure D.5.19. 2025 TELPAS Grade 9 Speaking
Frequency Distribution of Scale Scores



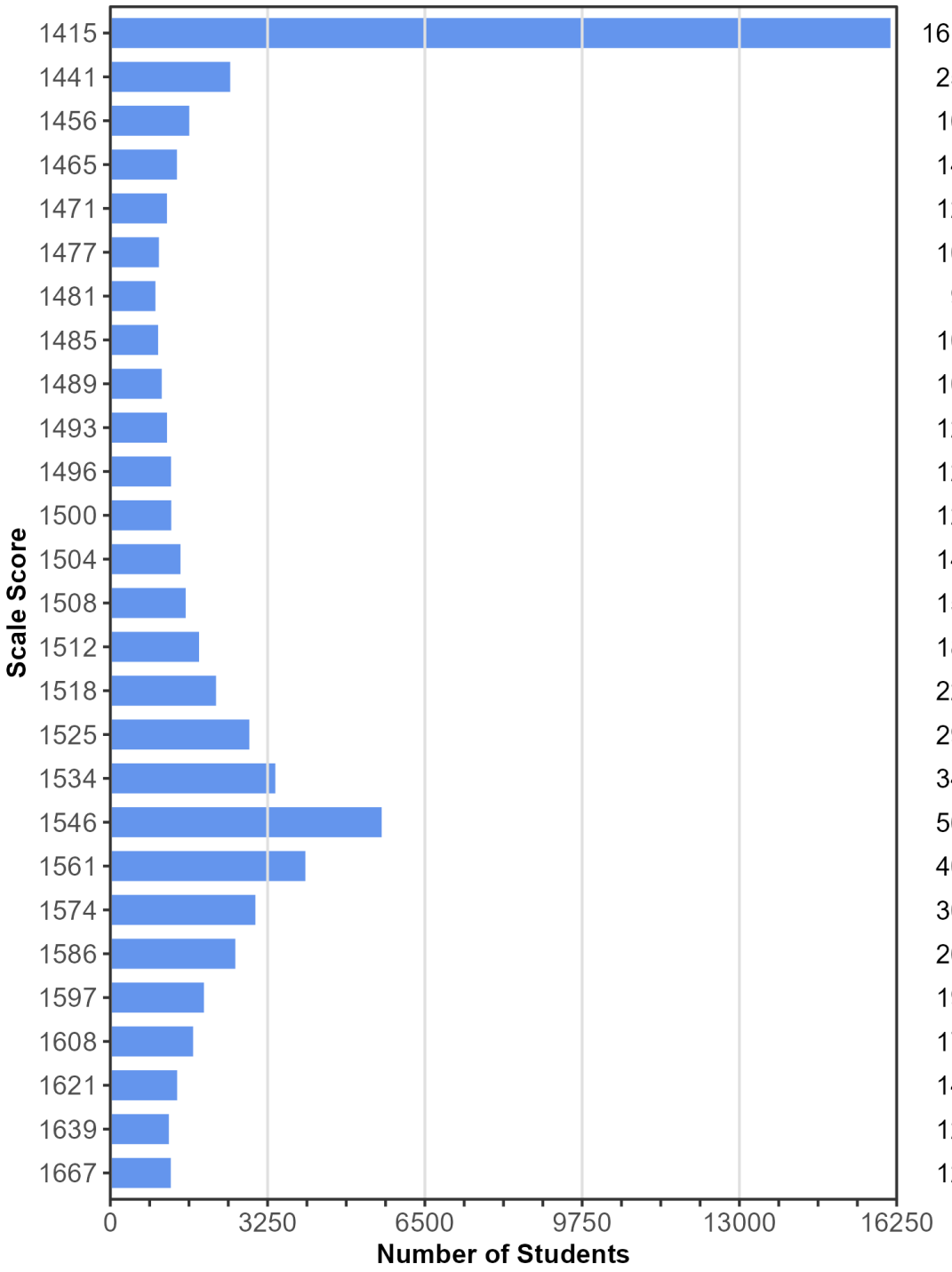
	Freq	Cum Freq	Pct	Cum Pct
1415	23354	23354	22.67	22.67
1441	4255	27609	4.13	26.80
1456	2877	30486	2.79	29.60
1465	2552	33038	2.48	32.07
1471	2069	35107	2.01	34.08
1477	1844	36951	1.79	35.87
1481	1733	38684	1.68	37.55
1485	1796	40480	1.74	39.30
1489	1870	42350	1.82	41.11
1493	2024	44374	1.96	43.08
1496	1975	46349	1.92	45.00
1500	2242	48591	2.18	47.17
1504	2340	50931	2.27	49.44
1508	2621	53552	2.54	51.99
1512	2921	56473	2.84	54.82
1518	3439	59912	3.34	58.16
1525	4428	64340	4.30	62.46
1534	5392	69732	5.23	67.70
1546	8896	78628	8.64	76.33
1561	6020	84648	5.84	82.18
1574	4521	89169	4.39	86.57
1586	3590	92759	3.49	90.05
1597	2812	95571	2.73	92.78
1608	2310	97881	2.24	95.02
1621	1874	99755	1.82	96.84
1639	1619	101374	1.57	98.41
1667	1634	103008	1.59	100.00

Figure D.5.20. 2025 TELPAS Grade 10 Speaking
Frequency Distribution of Scale Scores



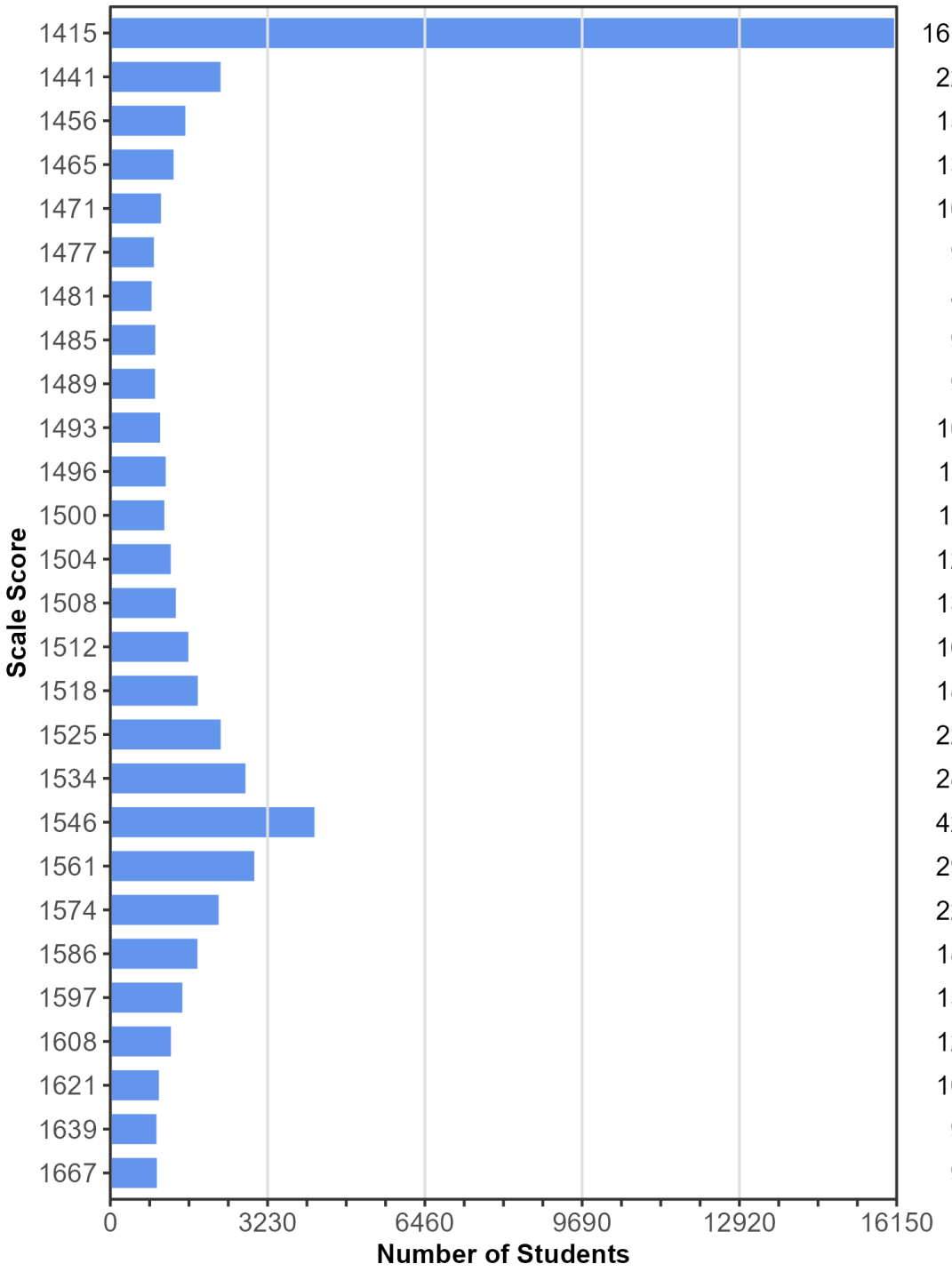
	Freq	Cum Freq	Pct	Cum Pct
1415	19770	19770	22.44	22.44
1441	3156	22926	3.58	26.02
1456	2216	25142	2.51	28.53
1465	1945	27087	2.21	30.74
1471	1588	28675	1.80	32.54
1477	1424	30099	1.62	34.16
1481	1324	31423	1.50	35.66
1485	1434	32857	1.63	37.29
1489	1426	34283	1.62	38.91
1493	1641	35924	1.86	40.77
1496	1624	37548	1.84	42.61
1500	1754	39302	1.99	44.60
1504	1936	41238	2.20	46.80
1508	1991	43229	2.26	49.06
1512	2565	45794	2.91	51.97
1518	2821	48615	3.20	55.17
1525	3790	52405	4.30	59.47
1534	4731	57136	5.37	64.84
1546	7795	64931	8.85	73.69
1561	5631	70562	6.39	80.08
1574	4119	74681	4.67	84.75
1586	3156	77837	3.58	88.34
1597	2622	80459	2.98	91.31
1608	2317	82776	2.63	93.94
1621	1901	84677	2.16	96.10
1639	1695	86372	1.92	98.02
1667	1743	88115	1.98	100.00

Figure D.5.21. 2025 TELPAS Grade 11 Speaking
Frequency Distribution of Scale Scores



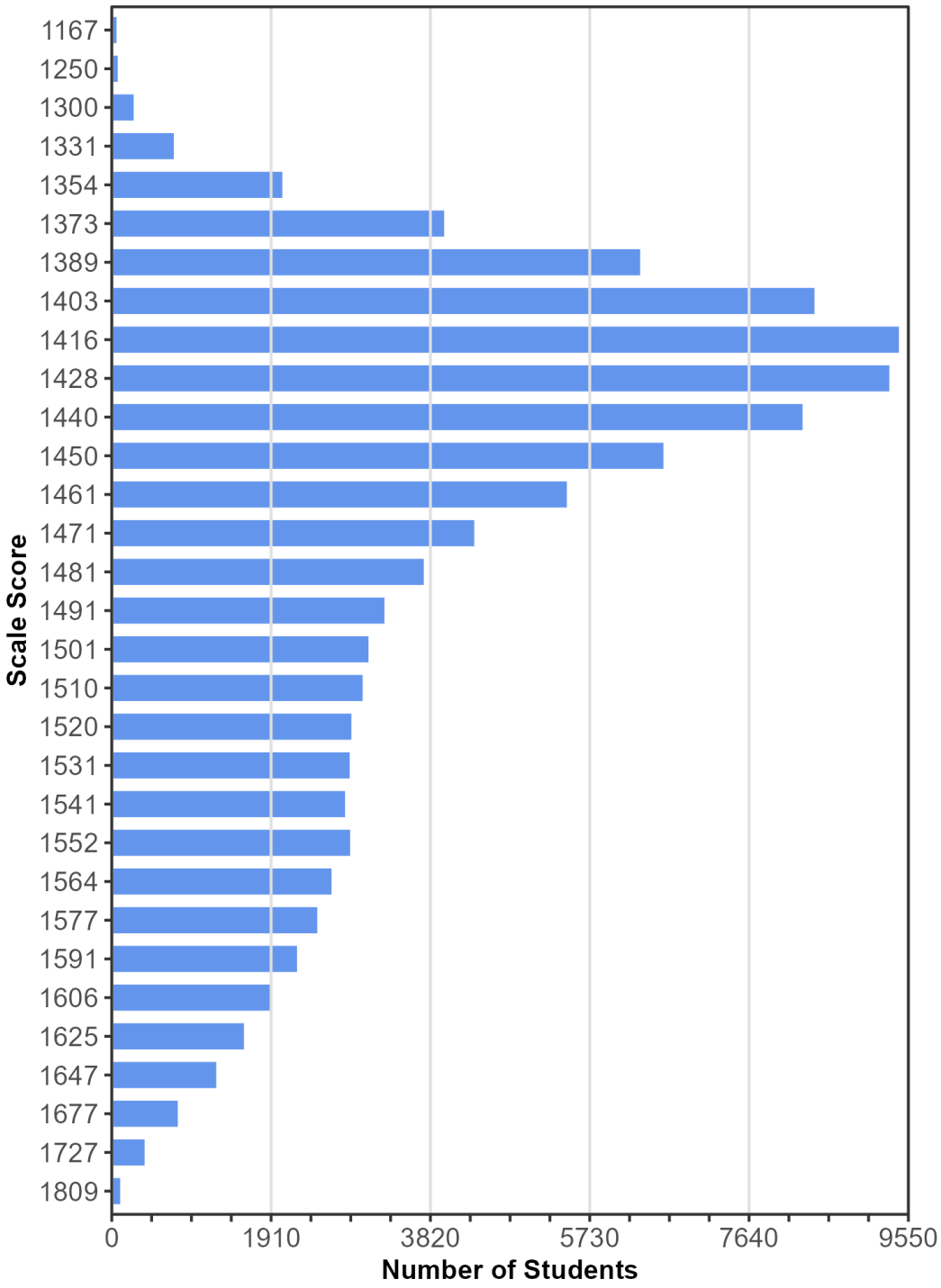
	Freq	Cum Freq	Pct	Cum Pct
1415	16152	16152	24.02	24.02
1441	2507	18659	3.73	27.74
1456	1661	20320	2.47	30.21
1465	1406	21726	2.09	32.30
1471	1200	22926	1.78	34.09
1477	1034	23960	1.54	35.63
1481	961	24921	1.43	37.06
1485	1017	25938	1.51	38.57
1489	1092	27030	1.62	40.19
1493	1201	28231	1.79	41.98
1496	1284	29515	1.91	43.89
1500	1288	30803	1.92	45.80
1504	1479	32282	2.20	48.00
1508	1586	33868	2.36	50.36
1512	1863	35731	2.77	53.13
1518	2214	37945	3.29	56.42
1525	2902	40847	4.31	60.74
1534	3439	44286	5.11	65.85
1546	5637	49923	8.38	74.23
1561	4060	53983	6.04	80.27
1574	3027	57010	4.50	84.77
1586	2612	59622	3.88	88.65
1597	1964	61586	2.92	91.57
1608	1740	63326	2.59	94.16
1621	1410	64736	2.10	96.26
1639	1239	65975	1.84	98.10
1667	1279	67254	1.90	100.00

Figure D.5.22. 2025 TELPAS Grade 12 Speaking
Frequency Distribution of Scale Scores



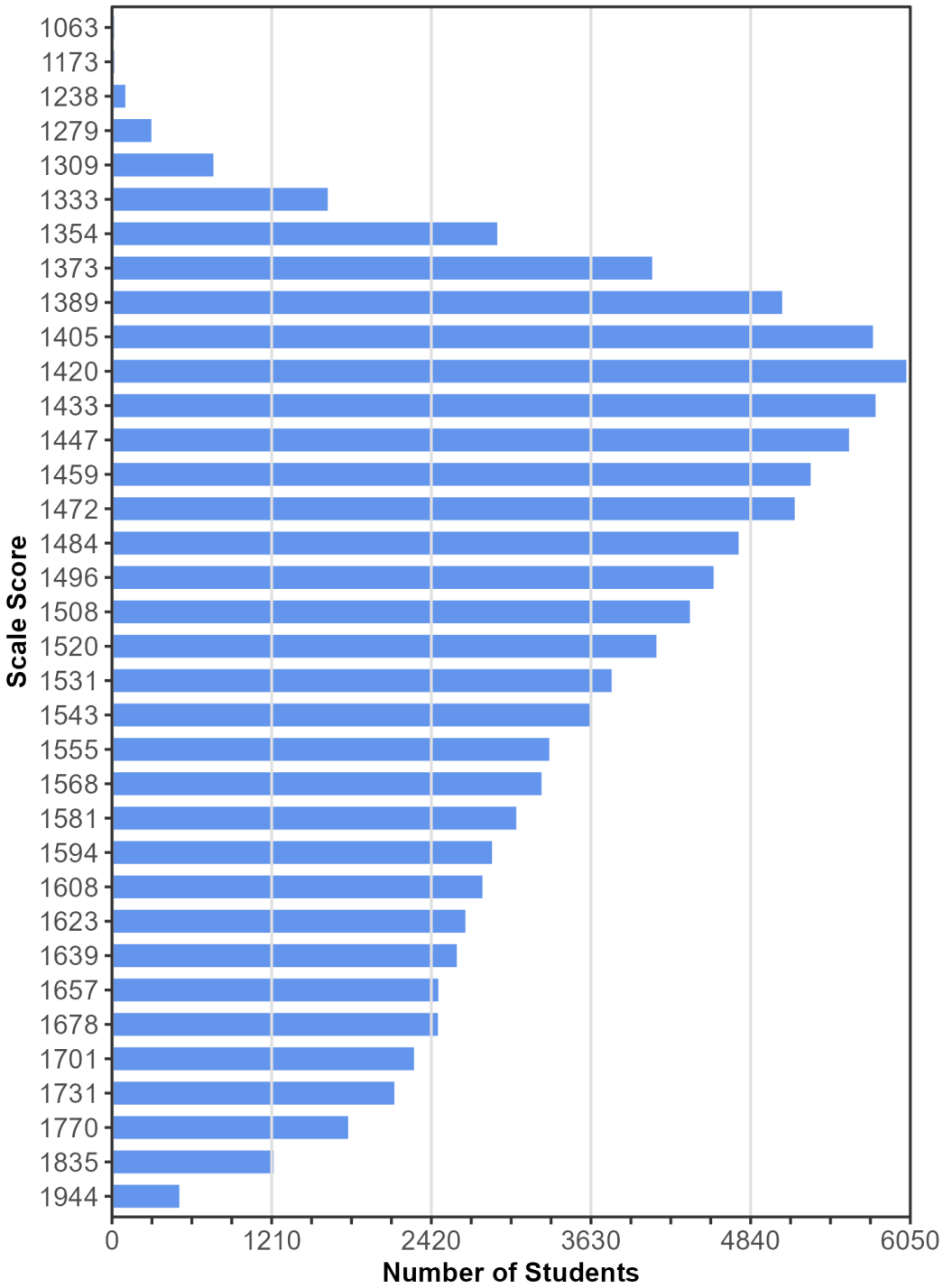
	Freq	Cum Freq	Pct	Cum Pct
1415	16125	16125	27.94	27.94
1441	2294	18419	3.98	31.92
1456	1569	19988	2.72	34.64
1465	1329	21317	2.30	36.94
1471	1069	22386	1.85	38.79
1477	924	23310	1.60	40.40
1481	877	24187	1.52	41.92
1485	953	25140	1.65	43.57
1489	948	26088	1.64	45.21
1493	1052	27140	1.82	47.03
1496	1167	28307	2.02	49.06
1500	1139	29446	1.97	51.03
1504	1271	30717	2.20	53.23
1508	1375	32092	2.38	55.61
1512	1632	33724	2.83	58.44
1518	1826	35550	3.16	61.61
1525	2296	37846	3.98	65.59
1534	2805	40651	4.86	70.45
1546	4222	44873	7.32	77.76
1561	2986	47859	5.17	82.94
1574	2254	50113	3.91	86.84
1586	1819	51932	3.15	90.00
1597	1509	53441	2.62	92.61
1608	1274	54715	2.21	94.82
1621	1027	55742	1.78	96.60
1639	977	56719	1.69	98.29
1667	985	57704	1.71	100.00

Figure D.5.23. 2025 TELPAS Grade 2 Reading
Frequency Distribution of Scale Scores



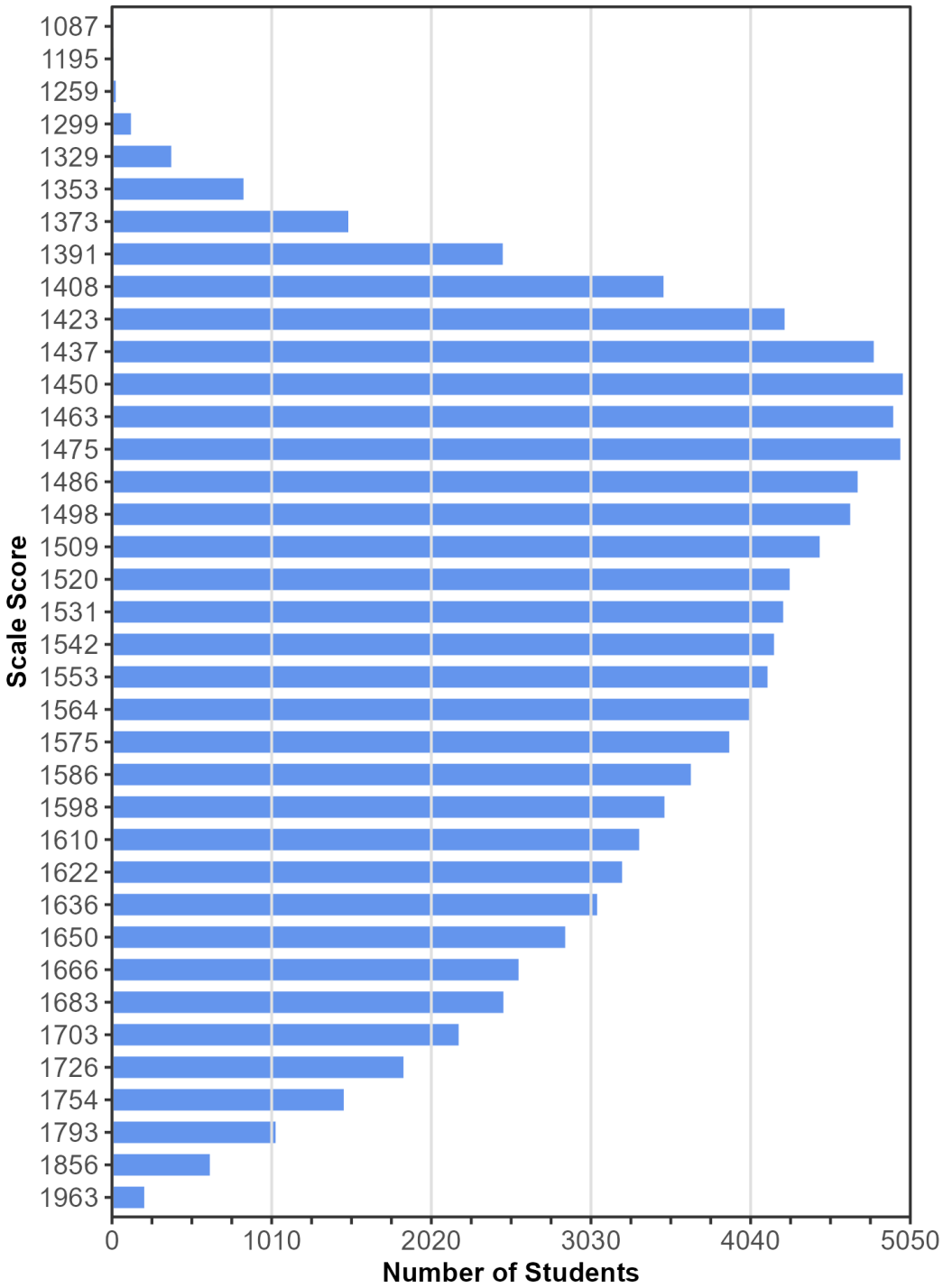
	Freq	Cum Freq	Pct	Cum Pct
1167	75	75	0.07	0.07
1250	89	164	0.09	0.16
1300	280	444	0.27	0.43
1331	761	1205	0.73	1.16
1354	2062	3267	1.99	3.15
1373	4002	7269	3.86	7.01
1389	6352	13621	6.12	13.13
1403	8442	22063	8.14	21.27
1416	9454	31517	9.11	30.38
1428	9339	40856	9.00	39.38
1440	8299	49155	8.00	47.38
1450	6631	55786	6.39	53.77
1461	5473	61259	5.28	59.04
1471	4363	65622	4.21	63.25
1481	3758	69380	3.62	66.87
1491	3286	72666	3.17	70.04
1501	3095	75761	2.98	73.02
1510	3025	78786	2.92	75.94
1520	2890	81676	2.79	78.72
1531	2870	84546	2.77	81.49
1541	2814	87360	2.71	84.20
1552	2875	90235	2.77	86.97
1564	2652	92887	2.56	89.53
1577	2481	95368	2.39	91.92
1591	2238	97606	2.16	94.08
1606	1935	99541	1.87	95.94
1625	1602	101143	1.54	97.49
1647	1270	102413	1.22	98.71
1677	809	103222	0.78	99.49
1727	411	103633	0.40	99.89
1809	118	103751	0.11	100.00

Figure D.5.24. 2025 TELPAS Grade 3 Reading
Frequency Distribution of Scale Scores



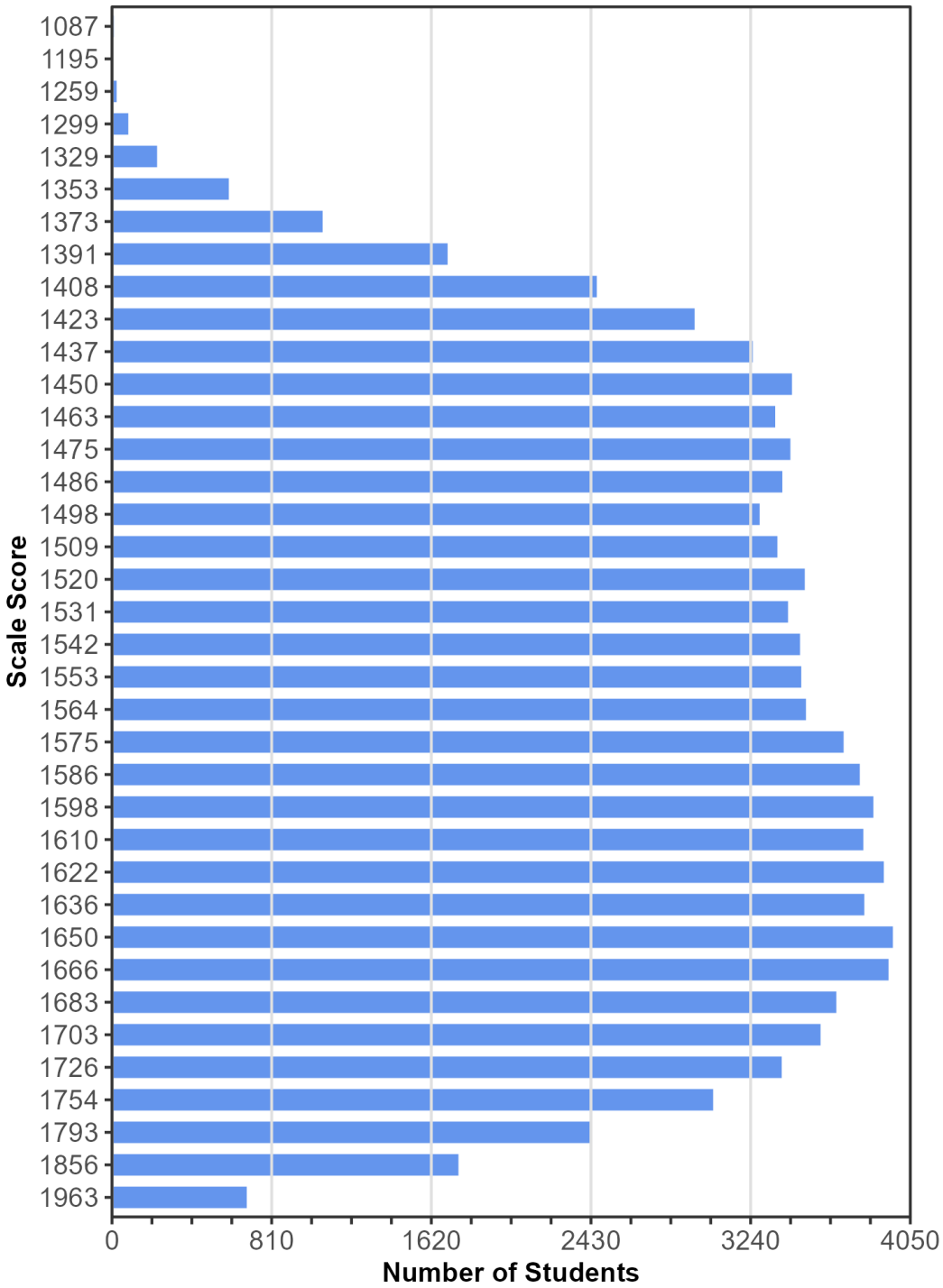
	Freq	Cum Freq	Pct	Cum Pct
1063	27	27	0.03	0.03
1173	28	55	0.03	0.05
1238	112	167	0.10	0.16
1279	309	476	0.29	0.44
1309	779	1255	0.72	1.17
1333	1646	2901	1.53	2.69
1354	2931	5832	2.72	5.42
1373	4105	9937	3.81	9.23
1389	5090	15027	4.73	13.95
1405	5778	20805	5.37	19.32
1420	6031	26836	5.60	24.92
1433	5798	32634	5.38	30.31
1447	5597	38231	5.20	35.50
1459	5306	43537	4.93	40.43
1472	5185	48722	4.82	45.25
1484	4760	53482	4.42	49.67
1496	4570	58052	4.24	53.91
1508	4391	62443	4.08	57.99
1520	4137	66580	3.84	61.83
1531	3797	70377	3.53	65.36
1543	3644	74021	3.38	68.74
1555	3325	77346	3.09	71.83
1568	3266	80612	3.03	74.86
1581	3075	83687	2.86	77.72
1594	2891	86578	2.68	80.40
1608	2818	89396	2.62	83.02
1623	2689	92085	2.50	85.51
1639	2624	94709	2.44	87.95
1657	2484	97193	2.31	90.26
1678	2481	99674	2.30	92.56
1701	2300	101974	2.14	94.70
1731	2151	104125	2.00	96.70
1770	1801	105926	1.67	98.37
1835	1236	107162	1.15	99.52
1944	521	107683	0.48	100.00

Figure D.5.25. 2025 TELPAS Grade 4 Reading
Frequency Distribution of Scale Scores



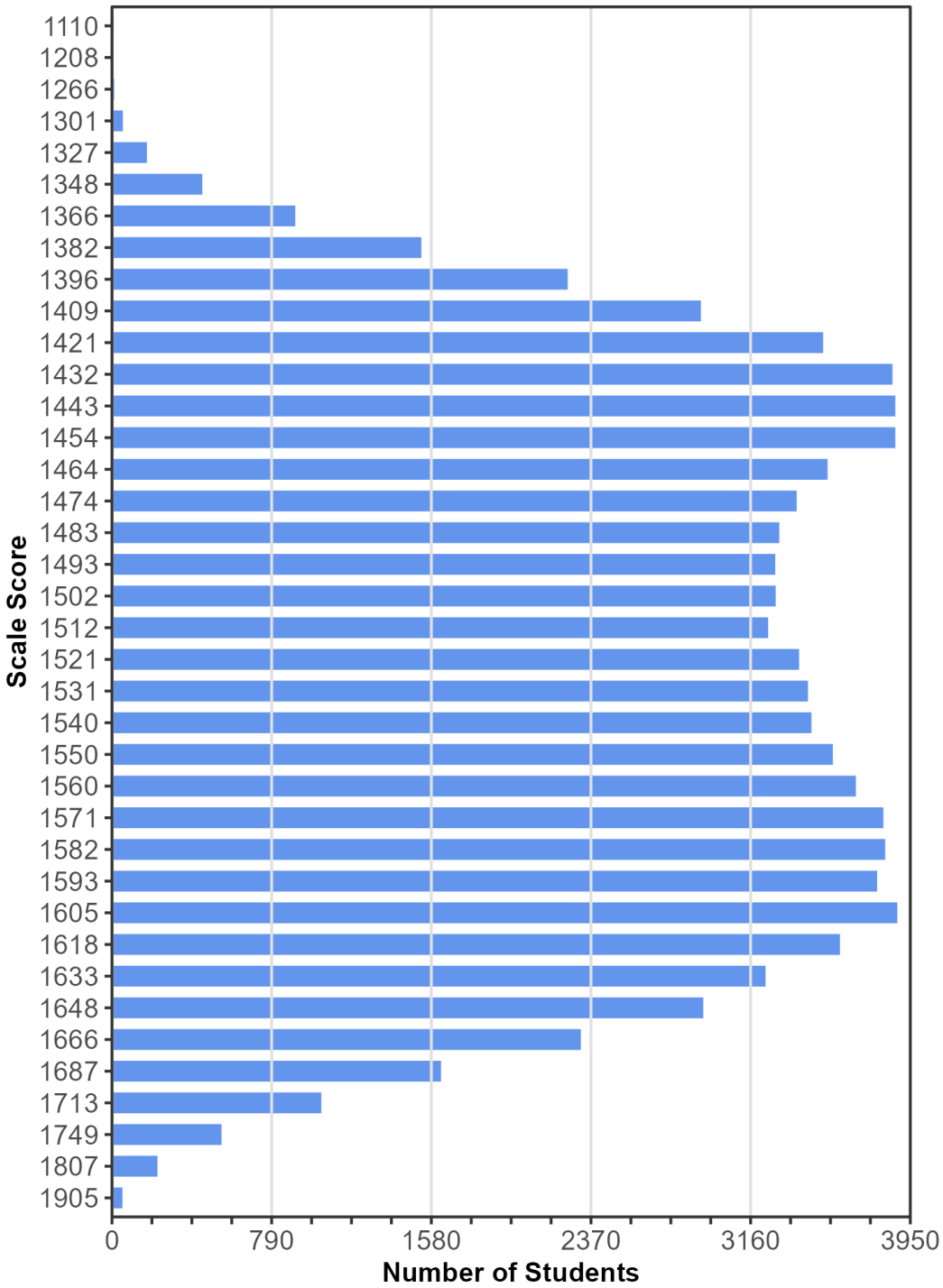
	Freq	Cum Freq	Pct	Cum Pct
1087	19	19	0.02	0.02
1195	20	39	0.02	0.04
1259	33	72	0.03	0.07
1299	129	201	0.12	0.19
1329	384	585	0.37	0.56
1353	841	1426	0.81	1.37
1373	1504	2930	1.45	2.82
1391	2481	5411	2.39	5.21
1408	3498	8909	3.37	8.57
1423	4264	13173	4.10	12.68
1437	4828	18001	4.65	17.33
1450	5012	23013	4.82	22.15
1463	4951	27964	4.77	26.91
1475	4996	32960	4.81	31.72
1486	4726	37686	4.55	36.27
1498	4679	42365	4.50	40.78
1509	4486	46851	4.32	45.09
1520	4296	51147	4.13	49.23
1531	4255	55402	4.10	53.32
1542	4197	59599	4.04	57.36
1553	4156	63755	4.00	61.36
1564	4039	67794	3.89	65.25
1575	3914	71708	3.77	69.02
1586	3671	75379	3.53	72.55
1598	3504	78883	3.37	75.92
1610	3344	82227	3.22	79.14
1622	3236	85463	3.11	82.26
1636	3078	88541	2.96	85.22
1650	2876	91417	2.77	87.99
1666	2581	93998	2.48	90.47
1683	2485	96483	2.39	92.86
1703	2202	98685	2.12	94.98
1726	1853	100538	1.78	96.77
1754	1475	102013	1.42	98.19
1793	1044	103057	1.00	99.19
1856	628	103685	0.60	99.79
1963	213	103898	0.21	100.00

Figure D.5.26. 2025 TELPAS Grade 5 Reading
Frequency Distribution of Scale Scores



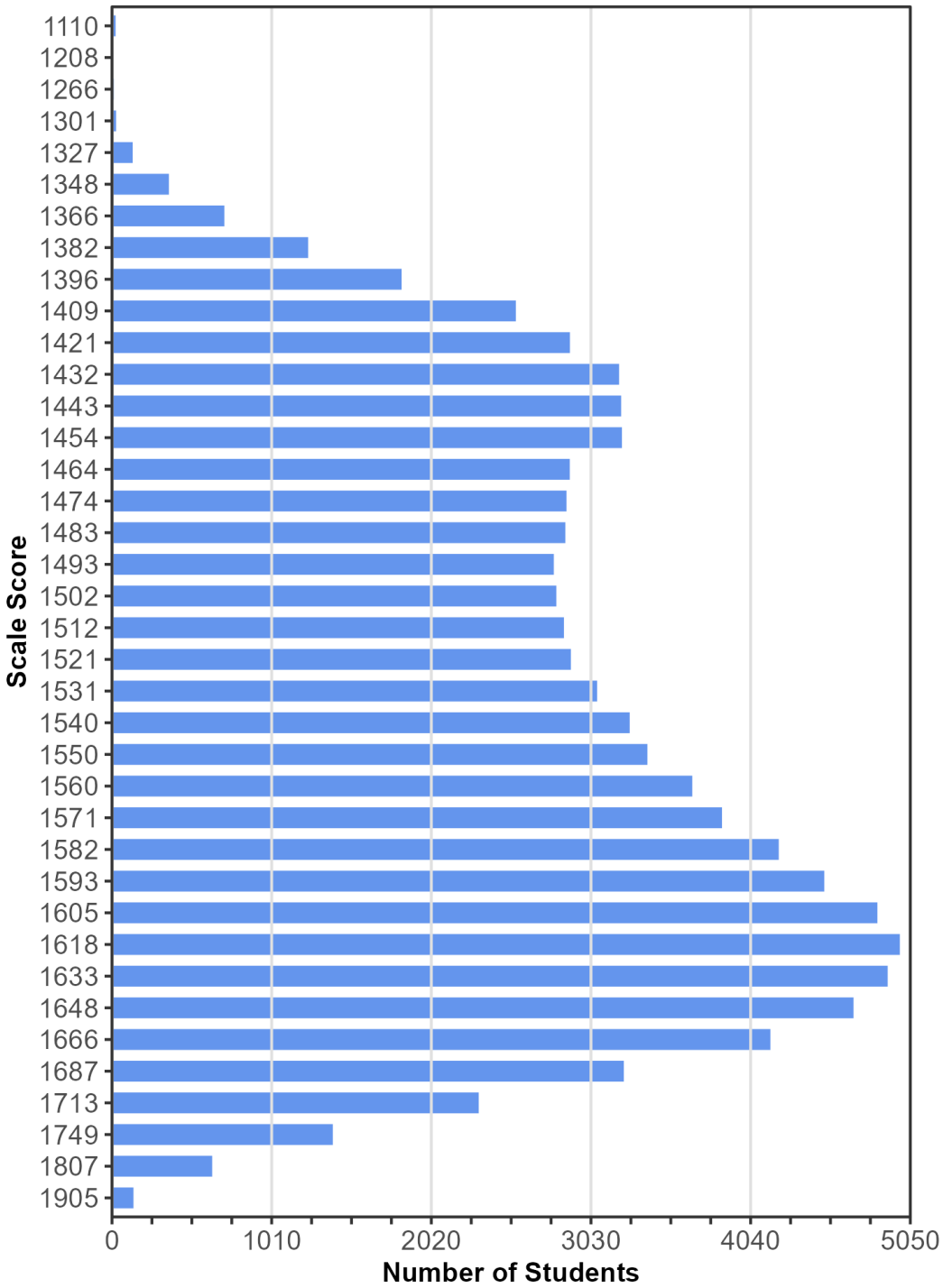
	Freq	Cum Freq	Pct	Cum Pct
1087	17	17	0.02	0.02
1195	7	24	0.01	0.02
1259	31	55	0.03	0.06
1299	90	145	0.09	0.15
1329	236	381	0.24	0.38
1353	600	981	0.60	0.98
1373	1076	2057	1.08	2.06
1391	1710	3767	1.71	3.77
1408	2467	6234	2.47	6.25
1423	2963	9197	2.97	9.21
1437	3258	12455	3.26	12.48
1450	3457	15912	3.46	15.94
1463	3372	19284	3.38	19.32
1475	3449	22733	3.46	22.78
1486	3408	26141	3.41	26.19
1498	3293	29434	3.30	29.49
1509	3383	32817	3.39	32.88
1520	3522	36339	3.53	36.41
1531	3437	39776	3.44	39.85
1542	3498	43274	3.50	43.36
1553	3504	46778	3.51	46.87
1564	3528	50306	3.53	50.40
1575	3719	54025	3.73	54.13
1586	3801	57826	3.81	57.93
1598	3870	61696	3.88	61.81
1610	3819	65515	3.83	65.64
1622	3923	69438	3.93	69.57
1636	3824	73262	3.83	73.40
1650	3969	77231	3.98	77.38
1666	3947	81178	3.95	81.33
1683	3682	84860	3.69	85.02
1703	3602	88462	3.61	88.63
1726	3405	91867	3.41	92.04
1754	3057	94924	3.06	95.10
1793	2432	97356	2.44	97.54
1856	1765	99121	1.77	99.31
1963	691	99812	0.69	100.00

Figure D.5.27. 2025 TELPAS Grade 6 Reading
Frequency Distribution of Scale Scores



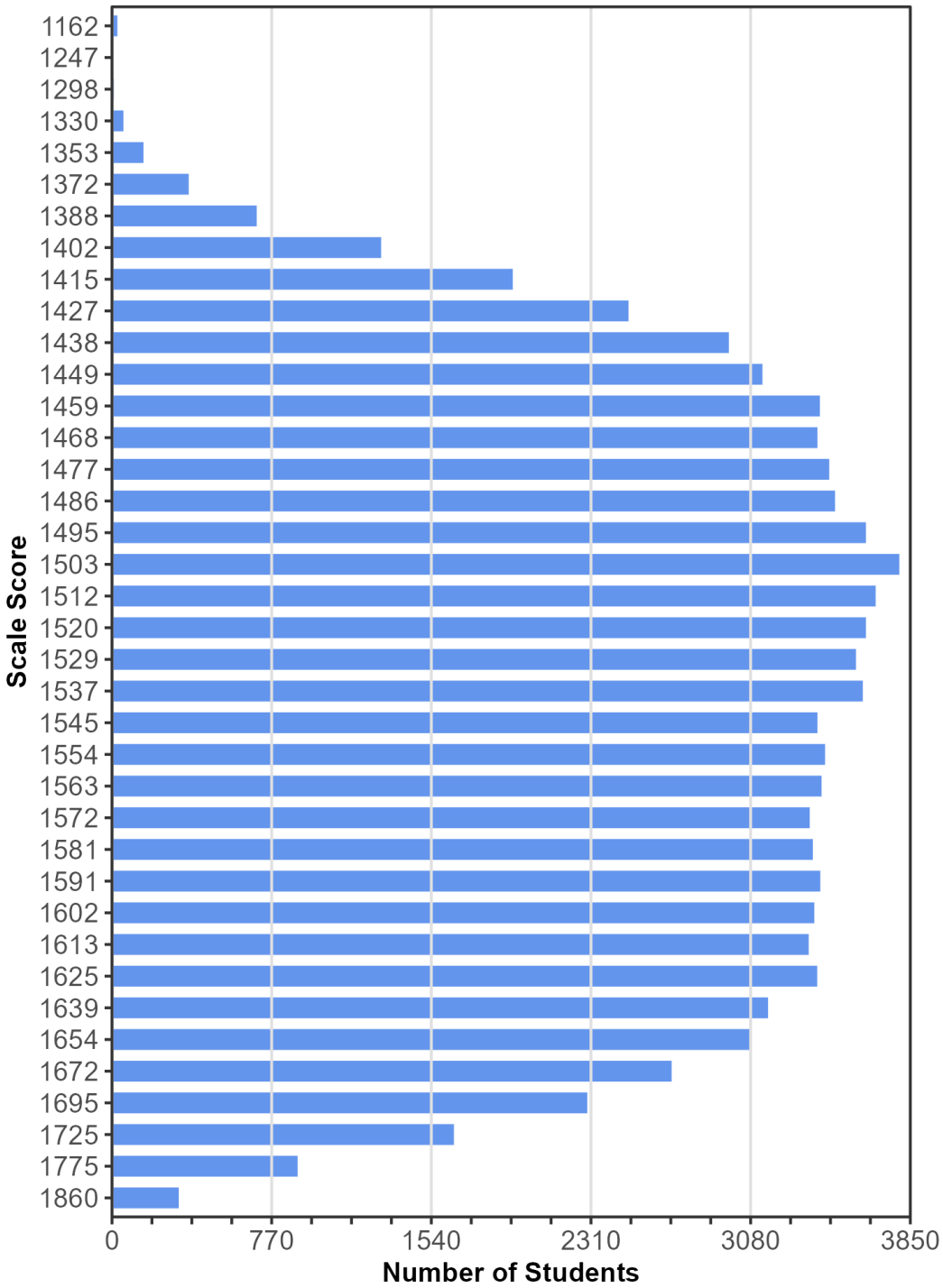
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1110	16	16	0.02	0.02
1208	5	21	0.01	0.02
1266	18	39	0.02	0.04
1301	61	100	0.07	0.11
1327	180	280	0.20	0.30
1348	454	734	0.49	0.80
1366	914	1648	0.99	1.79
1382	1538	3186	1.67	3.46
1396	2262	5448	2.45	5.91
1409	2921	8369	3.17	9.08
1421	3526	11895	3.83	12.90
1432	3869	15764	4.20	17.10
1443	3883	19647	4.21	21.31
1454	3883	23530	4.21	25.53
1464	3548	27078	3.85	29.38
1474	3395	30473	3.68	33.06
1483	3309	33782	3.59	36.65
1493	3289	37071	3.57	40.22
1502	3291	40362	3.57	43.79
1512	3254	43616	3.53	47.32
1521	3407	47023	3.70	51.01
1531	3451	50474	3.74	54.76
1540	3468	53942	3.76	58.52
1550	3574	57516	3.88	62.40
1560	3688	61204	4.00	66.40
1571	3824	65028	4.15	70.55
1582	3833	68861	4.16	74.71
1593	3793	72654	4.11	78.82
1605	3893	76547	4.22	83.05
1618	3609	80156	3.92	86.96
1633	3241	83397	3.52	90.48
1648	2933	86330	3.18	93.66
1666	2327	88657	2.52	96.18
1687	1635	90292	1.77	97.96
1713	1043	91335	1.13	99.09
1749	549	91884	0.60	99.68
1807	232	92116	0.25	99.94
1905	59	92175	0.06	100.00

Figure D.5.28. 2025 TELPAS Grade 7 Reading
Frequency Distribution of Scale Scores



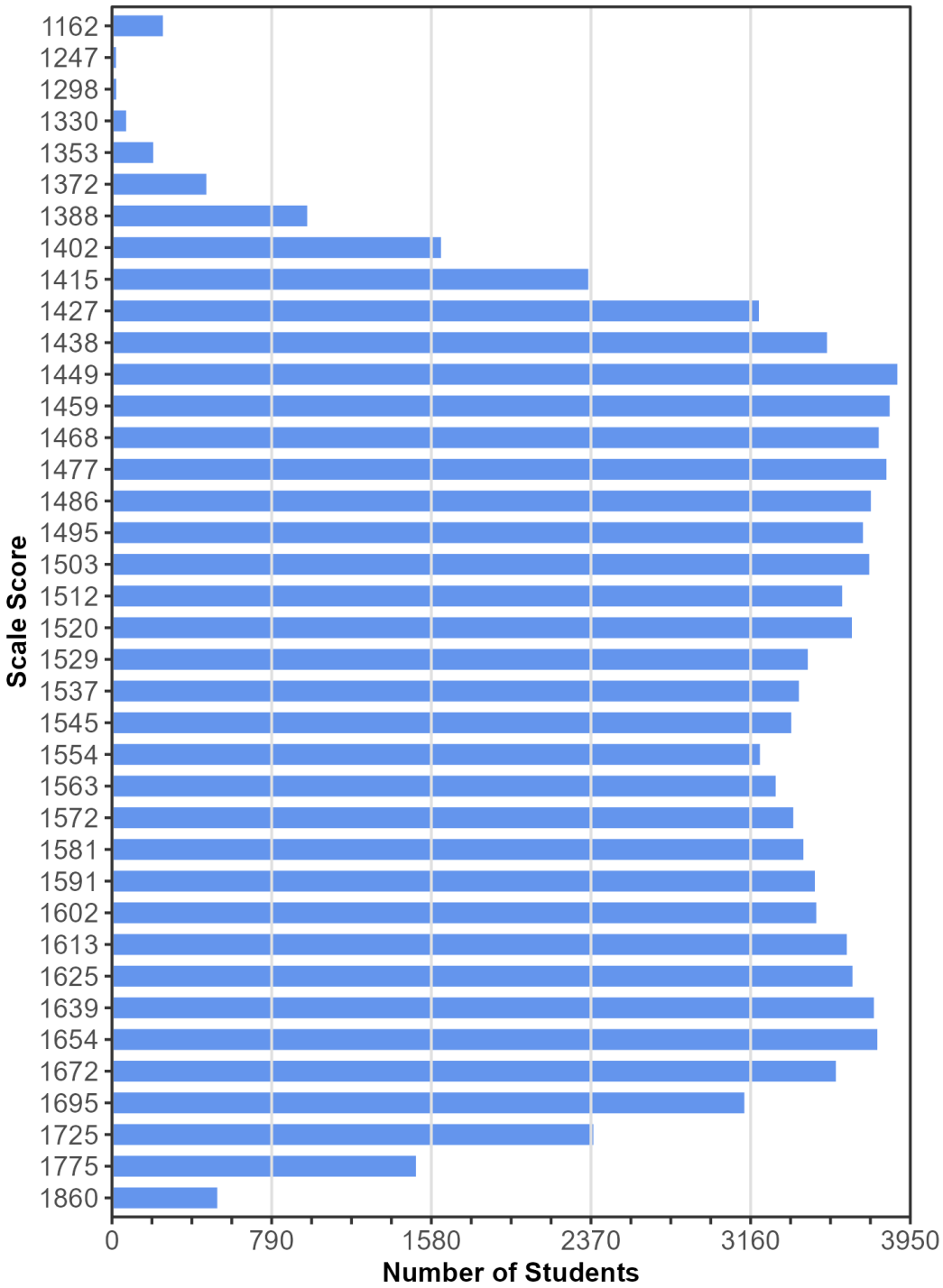
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1110	32	32	0.03	0.03
1208	4	36	0.00	0.04
1266	21	57	0.02	0.06
1301	35	92	0.04	0.09
1327	140	232	0.14	0.24
1348	369	601	0.38	0.62
1366	720	1321	0.74	1.36
1382	1250	2571	1.29	2.65
1396	1841	4412	1.90	4.54
1409	2564	6976	2.64	7.19
1421	2906	9882	2.99	10.18
1432	3217	13099	3.31	13.49
1443	3230	16329	3.33	16.82
1454	3235	19564	3.33	20.15
1464	2905	22469	2.99	23.14
1474	2884	25353	2.97	26.11
1483	2877	28230	2.96	29.08
1493	2804	31034	2.89	31.97
1502	2820	33854	2.90	34.87
1512	2868	36722	2.95	37.82
1521	2912	39634	3.00	40.82
1531	3078	42712	3.17	43.99
1540	3284	45996	3.38	47.38
1550	3396	49392	3.50	50.87
1560	3680	53072	3.79	54.66
1571	3868	56940	3.98	58.65
1582	4227	61167	4.35	63.00
1593	4515	65682	4.65	67.65
1605	4851	70533	5.00	72.65
1618	4993	75526	5.14	77.79
1633	4916	80442	5.06	82.86
1648	4700	85142	4.84	87.70
1666	4174	89316	4.30	92.00
1687	3247	92563	3.34	95.34
1713	2329	94892	2.40	97.74
1749	1406	96298	1.45	99.19
1807	643	96941	0.66	99.85
1905	145	97086	0.15	100.00

Figure D.5.29. 2025 TELPAS Grade 8 Reading
Frequency Distribution of Scale Scores



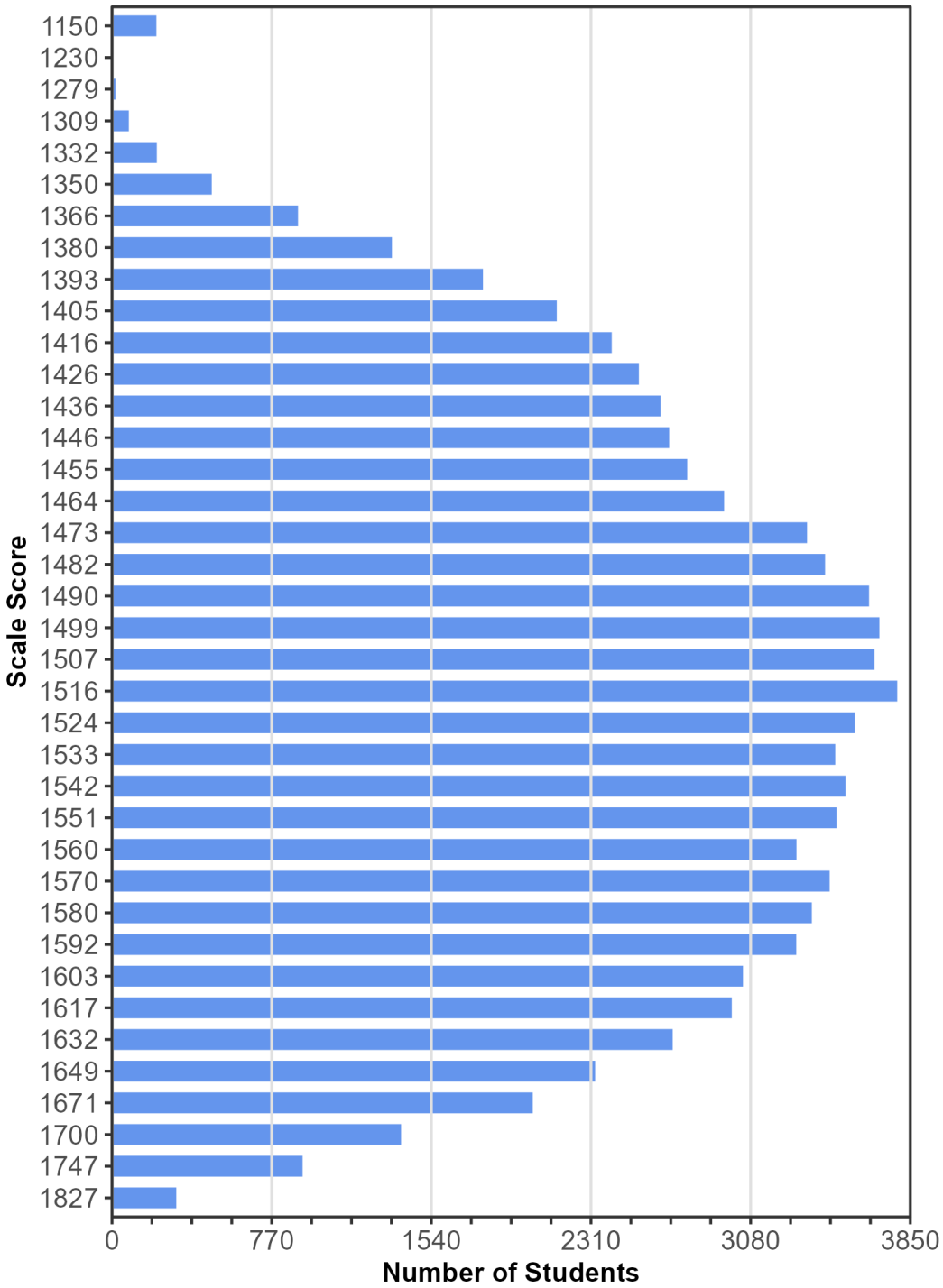
	Freq	Cum Freq	Pct	Cum Pct
1162	33	33	0.04	0.04
1247	6	39	0.01	0.04
1298	16	55	0.02	0.06
1330	62	117	0.07	0.12
1353	159	276	0.17	0.29
1372	377	653	0.40	0.70
1388	705	1358	0.75	1.45
1402	1305	2663	1.39	2.84
1415	1940	4603	2.07	4.91
1427	2498	7101	2.66	7.57
1438	2982	10083	3.18	10.75
1449	3144	13227	3.35	14.10
1459	3421	16648	3.65	17.75
1468	3409	20057	3.63	21.38
1477	3466	23523	3.69	25.08
1486	3494	27017	3.72	28.80
1495	3643	30660	3.88	32.68
1503	3804	34464	4.06	36.74
1512	3690	38154	3.93	40.67
1520	3643	41797	3.88	44.56
1529	3595	45392	3.83	48.39
1537	3628	49020	3.87	52.26
1545	3409	52429	3.63	55.89
1554	3446	55875	3.67	59.56
1563	3429	59304	3.66	63.22
1572	3372	62676	3.59	66.81
1581	3387	66063	3.61	70.42
1591	3423	69486	3.65	74.07
1602	3394	72880	3.62	77.69
1613	3367	76247	3.59	81.28
1625	3408	79655	3.63	84.91
1639	3171	82826	3.38	88.29
1654	3091	85917	3.29	91.59
1672	2706	88623	2.88	94.47
1695	2299	90922	2.45	96.92
1725	1656	92578	1.77	98.69
1775	902	93480	0.96	99.65
1860	329	93809	0.35	100.00

Figure D.5.30. 2025 TELPAS Grade 9 Reading
Frequency Distribution of Scale Scores



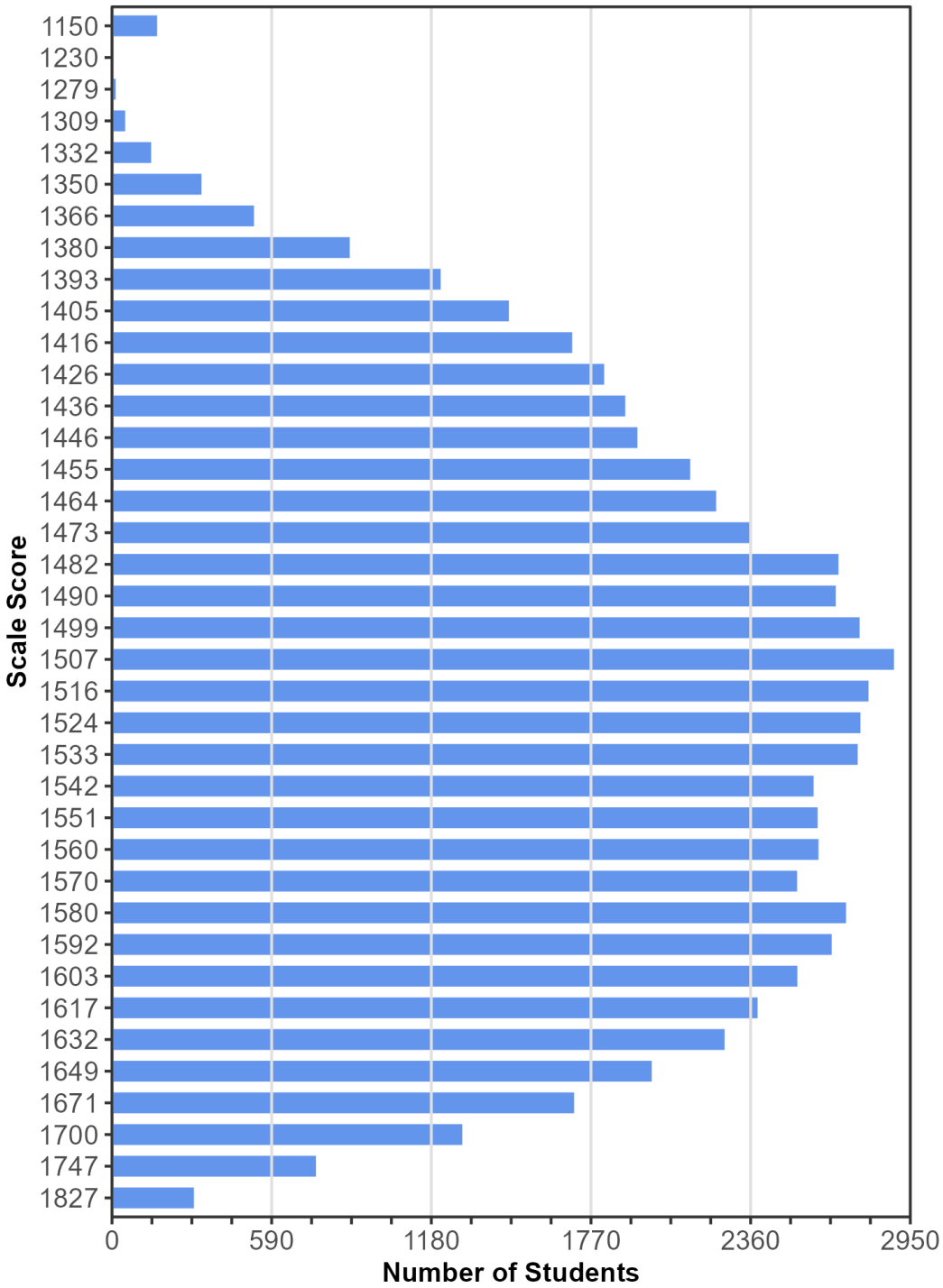
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1162	259	259	0.25	0.25
1247	27	286	0.03	0.28
1298	28	314	0.03	0.30
1330	77	391	0.07	0.38
1353	211	602	0.20	0.58
1372	474	1076	0.46	1.04
1388	973	2049	0.94	1.98
1402	1635	3684	1.58	3.57
1415	2364	6048	2.29	5.86
1427	3208	9256	3.11	8.96
1438	3545	12801	3.43	12.40
1449	3893	16694	3.77	16.17
1459	3855	20549	3.73	19.90
1468	3801	24350	3.68	23.58
1477	3839	28189	3.72	27.30
1486	3762	31951	3.64	30.94
1495	3723	35674	3.61	34.55
1503	3754	39428	3.64	38.18
1512	3620	43048	3.51	41.69
1520	3668	46716	3.55	45.24
1529	3450	50166	3.34	48.58
1537	3406	53572	3.30	51.88
1545	3368	56940	3.26	55.14
1554	3213	60153	3.11	58.25
1563	3291	63444	3.19	61.44
1572	3378	66822	3.27	64.71
1581	3428	70250	3.32	68.03
1591	3485	73735	3.37	71.40
1602	3492	77227	3.38	74.79
1613	3643	80870	3.53	78.31
1625	3671	84541	3.55	81.87
1639	3777	88318	3.66	85.53
1654	3794	92112	3.67	89.20
1672	3589	95701	3.48	92.68
1695	3136	98837	3.04	95.71
1725	2389	101226	2.31	98.03
1775	1511	102737	1.46	99.49
1860	528	103265	0.51	100.00

Figure D.5.31. 2025 TELPAS Grade 10 Reading
Frequency Distribution of Scale Scores



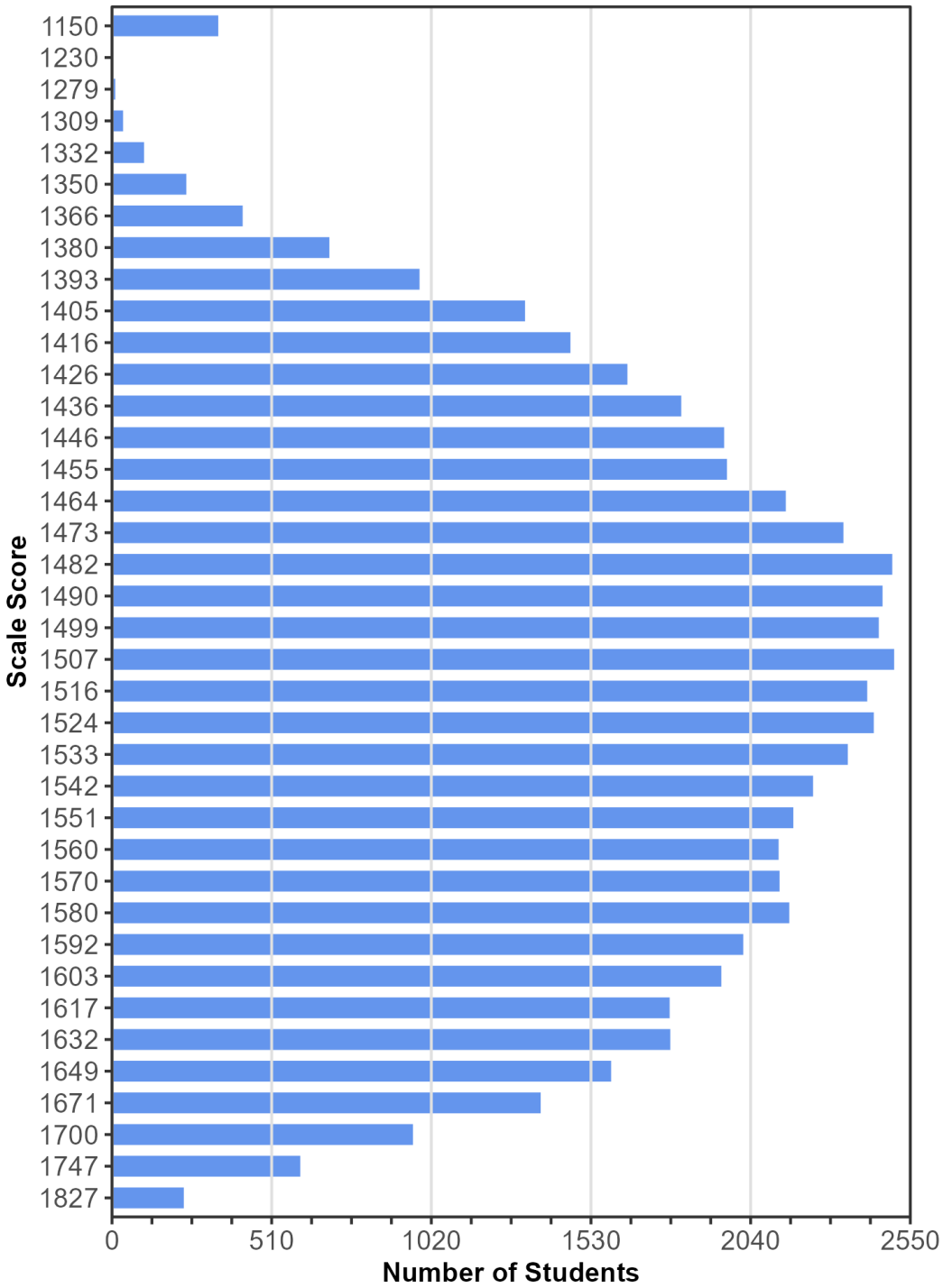
	Freq	Cum Freq	Pct	Cum Pct
1150	221	221	0.25	0.25
1230	14	235	0.02	0.27
1279	24	259	0.03	0.29
1309	88	347	0.10	0.39
1332	223	570	0.25	0.65
1350	488	1058	0.55	1.20
1366	904	1962	1.02	2.22
1380	1357	3319	1.54	3.76
1393	1796	5115	2.03	5.79
1405	2152	7267	2.44	8.23
1416	2417	9684	2.74	10.96
1426	2548	12232	2.88	13.85
1436	2653	14885	3.00	16.85
1446	2694	17579	3.05	19.90
1455	2781	20360	3.15	23.05
1464	2959	23319	3.35	26.40
1473	3359	26678	3.80	30.20
1482	3446	30124	3.90	34.10
1490	3658	33782	4.14	38.24
1499	3708	37490	4.20	42.44
1507	3684	41174	4.17	46.61
1516	3794	44968	4.29	50.90
1524	3590	48558	4.06	54.97
1533	3495	52053	3.96	58.92
1542	3545	55598	4.01	62.94
1551	3502	59100	3.96	66.90
1560	3308	62408	3.74	70.65
1570	3468	65876	3.93	74.57
1580	3382	69258	3.83	78.40
1592	3307	72565	3.74	82.14
1603	3050	75615	3.45	85.60
1617	2996	78611	3.39	88.99
1632	2711	81322	3.07	92.06
1649	2337	83659	2.65	94.70
1671	2036	85695	2.30	97.01
1700	1401	87096	1.59	98.59
1747	926	88022	1.05	99.64
1827	317	88339	0.36	100.00

Figure D.5.32. 2025 TELPAS Grade 11 Reading
Frequency Distribution of Scale Scores



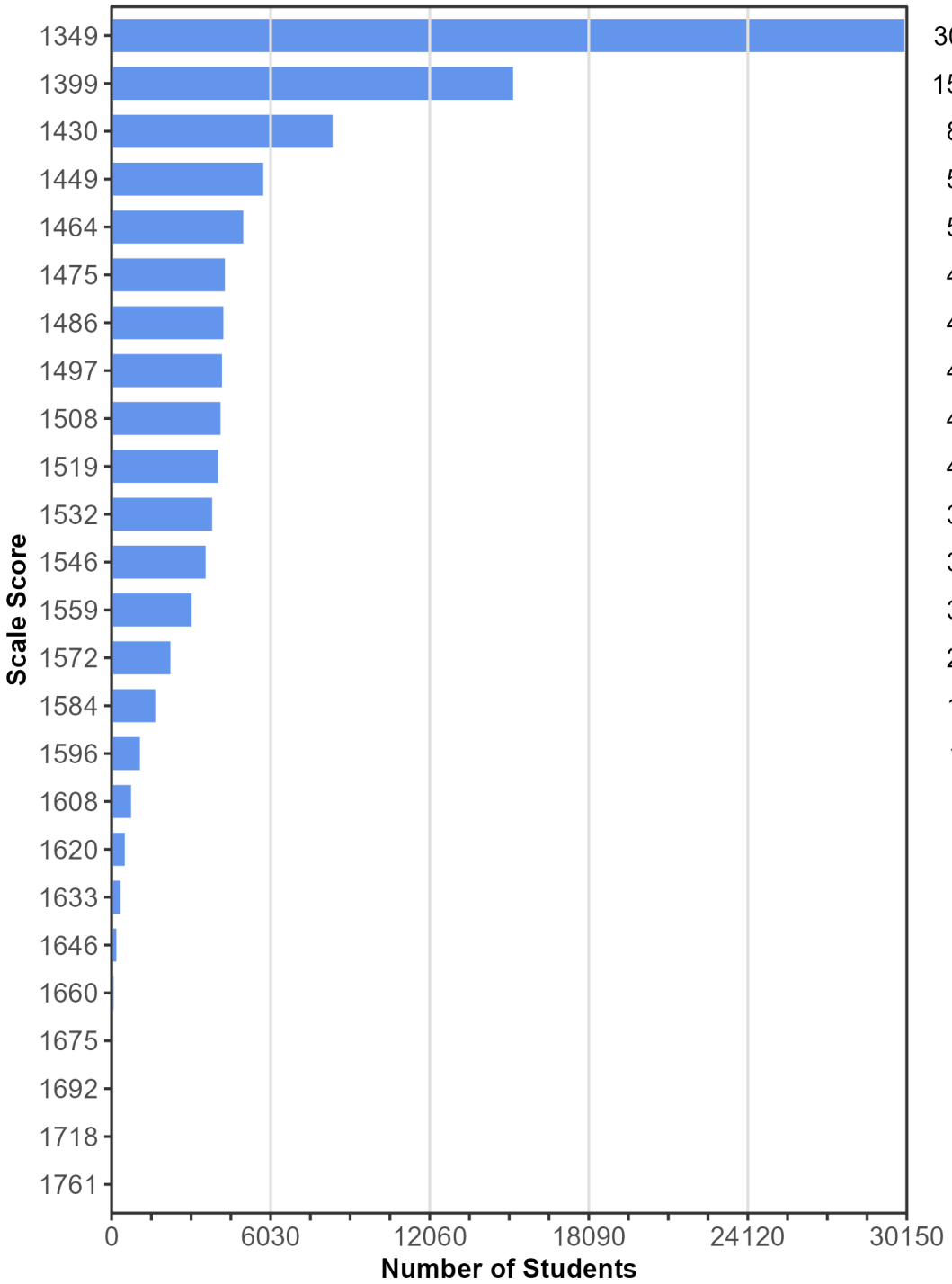
	Freq	Cum Freq	Pct	Cum Pct
1150	172	172	0.26	0.26
1230	6	178	0.01	0.26
1279	19	197	0.03	0.29
1309	54	251	0.08	0.37
1332	150	401	0.22	0.60
1350	336	737	0.50	1.09
1366	530	1267	0.79	1.88
1380	884	2151	1.31	3.19
1393	1220	3371	1.81	5.00
1405	1472	4843	2.19	7.19
1416	1706	6549	2.53	9.72
1426	1824	8373	2.71	12.43
1436	1902	10275	2.82	15.25
1446	1947	12222	2.89	18.15
1455	2142	14364	3.18	21.33
1464	2238	16602	3.32	24.65
1473	2363	18965	3.51	28.16
1482	2690	21655	3.99	32.15
1490	2680	24335	3.98	36.13
1499	2768	27103	4.11	40.24
1507	2895	29998	4.30	44.54
1516	2801	32799	4.16	48.69
1524	2771	35570	4.11	52.81
1533	2761	38331	4.10	56.91
1542	2598	40929	3.86	60.77
1551	2613	43542	3.88	64.64
1560	2616	46158	3.88	68.53
1570	2537	48695	3.77	72.29
1580	2718	51413	4.04	76.33
1592	2665	54078	3.96	80.29
1603	2538	56616	3.77	84.05
1617	2391	59007	3.55	87.60
1632	2269	61276	3.37	90.97
1649	2000	63276	2.97	93.94
1671	1713	64989	2.54	96.49
1700	1300	66289	1.93	98.42
1747	759	67048	1.13	99.54
1827	308	67356	0.46	100.00

Figure D.5.33. 2025 TELPAS Grade 12 Reading
Frequency Distribution of Scale Scores



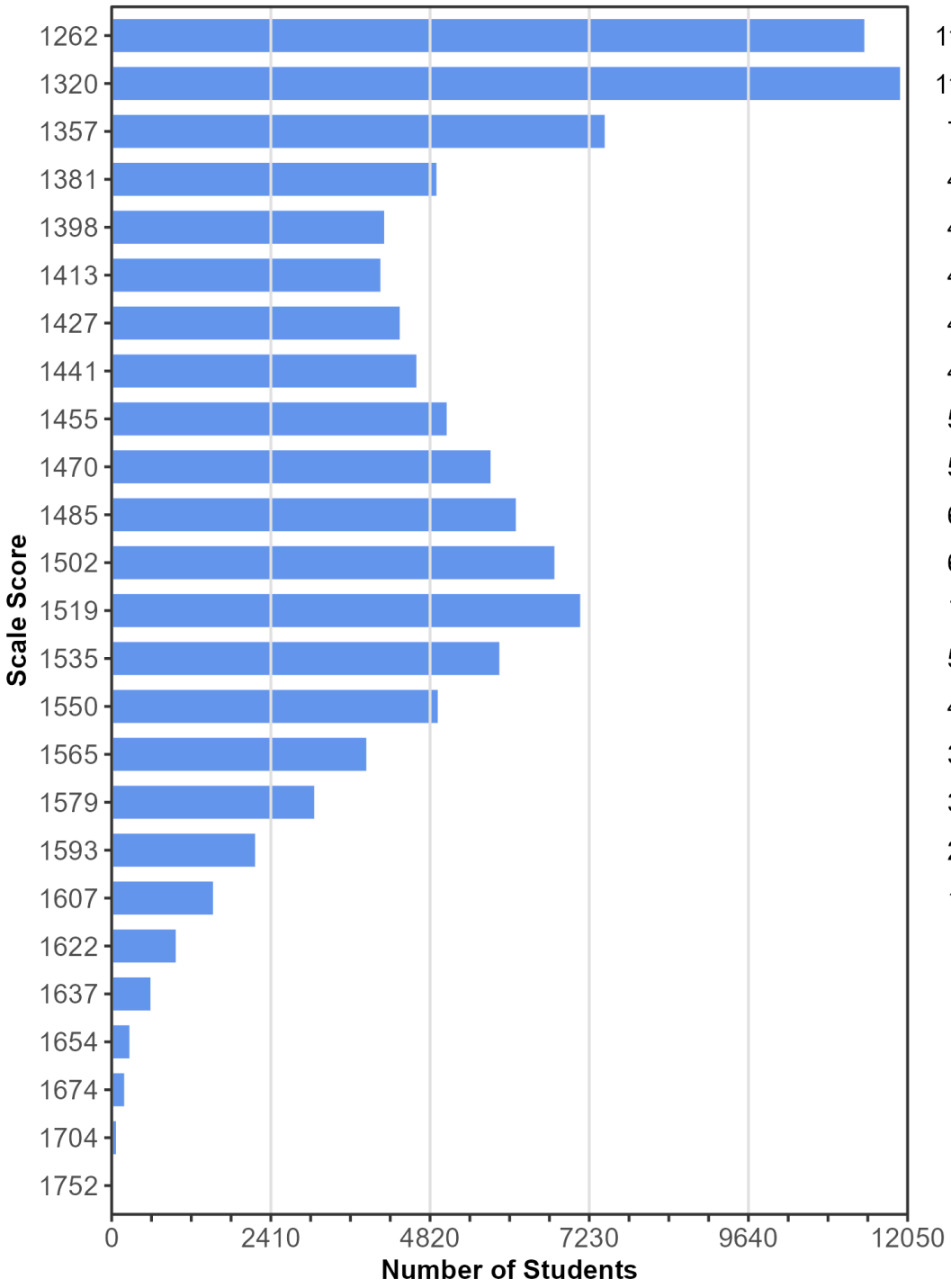
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1150	344	344	0.59	0.59
1230	8	352	0.01	0.61
1279	15	367	0.03	0.63
1309	40	407	0.07	0.70
1332	107	514	0.18	0.89
1350	242	756	0.42	1.31
1366	422	1178	0.73	2.04
1380	699	1877	1.21	3.24
1393	987	2864	1.71	4.95
1405	1324	4188	2.29	7.24
1416	1469	5657	2.54	9.77
1426	1651	7308	2.85	12.63
1436	1823	9131	3.15	15.78
1446	1960	11091	3.39	19.16
1455	1969	13060	3.40	22.57
1464	2157	15217	3.73	26.29
1473	2341	17558	4.04	30.34
1482	2497	20055	4.31	34.65
1490	2466	22521	4.26	38.91
1499	2454	24975	4.24	43.15
1507	2503	27478	4.32	47.48
1516	2417	29895	4.18	51.65
1524	2438	32333	4.21	55.87
1533	2355	34688	4.07	59.93
1542	2244	36932	3.88	63.81
1551	2181	39113	3.77	67.58
1560	2134	41247	3.69	71.27
1570	2137	43384	3.69	74.96
1580	2168	45552	3.75	78.70
1592	2021	47573	3.49	82.20
1603	1951	49524	3.37	85.57
1617	1786	51310	3.09	88.65
1632	1788	53098	3.09	91.74
1649	1599	54697	2.76	94.51
1671	1374	56071	2.37	96.88
1700	966	57037	1.67	98.55
1747	606	57643	1.05	99.60
1827	234	57877	0.40	100.00

Figure D.5.34. 2025 TELPAS Grade 2 Writing
Frequency Distribution of Scale Scores



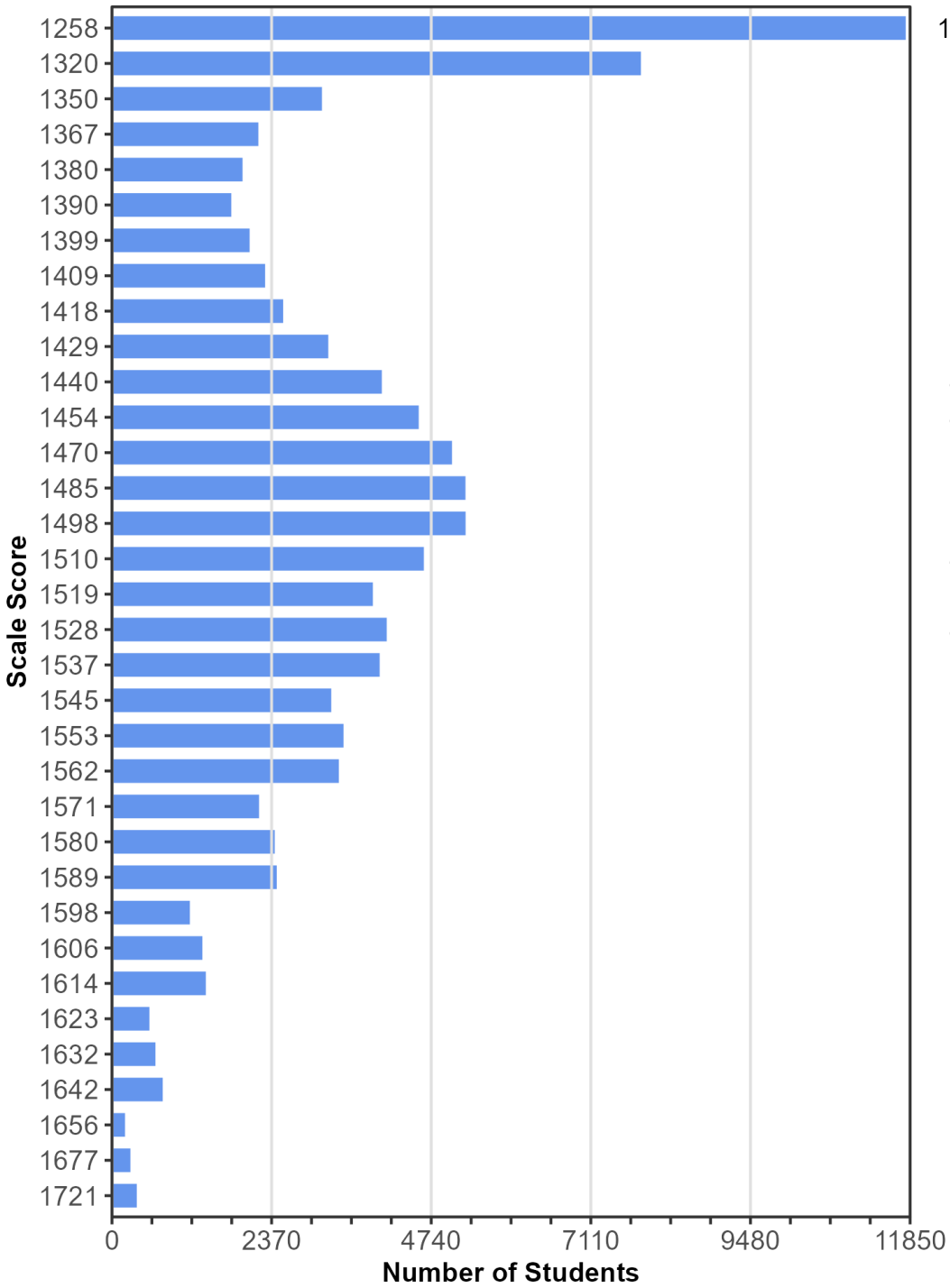
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1349	30113	30113	29.02	29.02
1399	15276	45389	14.72	43.75
1430	8433	53822	8.13	51.88
1449	5804	59626	5.59	57.47
1464	5045	64671	4.86	62.33
1475	4350	69021	4.19	66.53
1486	4293	73314	4.14	70.66
1497	4240	77554	4.09	74.75
1508	4183	81737	4.03	78.78
1519	4093	85830	3.95	82.73
1532	3867	89697	3.73	86.45
1546	3619	93316	3.49	89.94
1559	3087	96403	2.98	92.92
1572	2285	98688	2.20	95.12
1584	1709	100397	1.65	96.77
1596	1128	101525	1.09	97.85
1608	793	102318	0.76	98.62
1620	554	102872	0.53	99.15
1633	396	103268	0.38	99.53
1646	235	103503	0.23	99.76
1660	132	103635	0.13	99.89
1675	67	103702	0.06	99.95
1692	30	103732	0.03	99.98
1718	13	103745	0.01	99.99
1761	6	103751	0.01	100.00

Figure D.5.35. 2025 TELPAS Grade 3 Writing
Frequency Distribution of Scale Scores



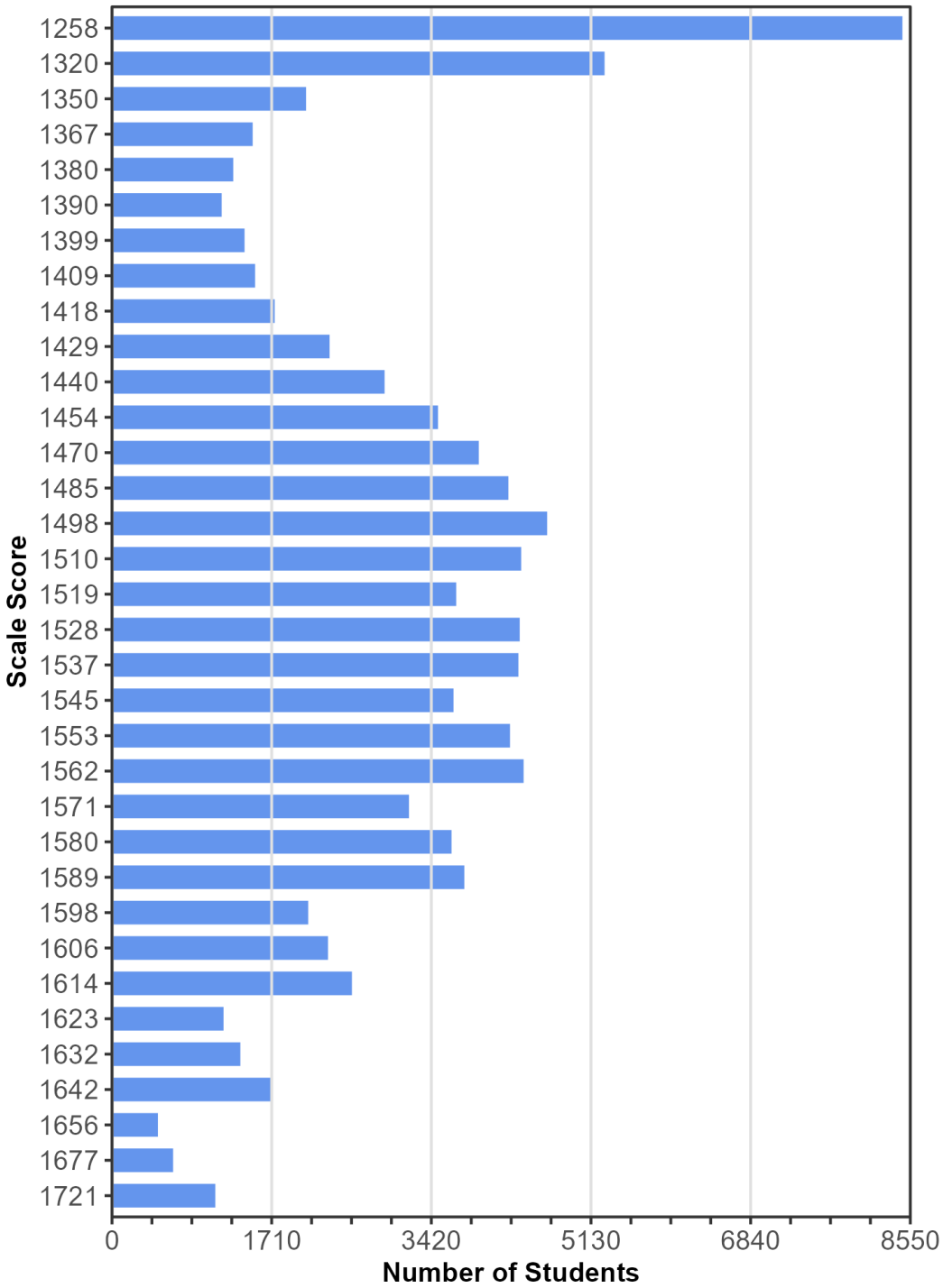
	Freq	Cum Freq	Pct	Cum Pct
1262	11417	11417	10.60	10.60
1320	11958	23375	11.10	21.71
1357	7486	30861	6.95	28.66
1381	4939	35800	4.59	33.25
1398	4147	39947	3.85	37.10
1413	4090	44037	3.80	40.90
1427	4383	48420	4.07	44.97
1441	4635	53055	4.30	49.27
1455	5094	58149	4.73	54.00
1470	5758	63907	5.35	59.35
1485	6141	70048	5.70	65.05
1502	6725	76773	6.25	71.30
1519	7115	83888	6.61	77.90
1535	5891	89779	5.47	83.37
1550	4959	94738	4.61	87.98
1565	3877	98615	3.60	91.58
1579	3088	101703	2.87	94.45
1593	2193	103896	2.04	96.48
1607	1556	105452	1.44	97.93
1622	992	106444	0.92	98.85
1637	609	107053	0.57	99.41
1654	290	107343	0.27	99.68
1674	211	107554	0.20	99.88
1704	87	107641	0.08	99.96
1752	42	107683	0.04	100.00

Figure D.5.36. 2025 TELPAS Grade 4 Writing
Frequency Distribution of Scale Scores



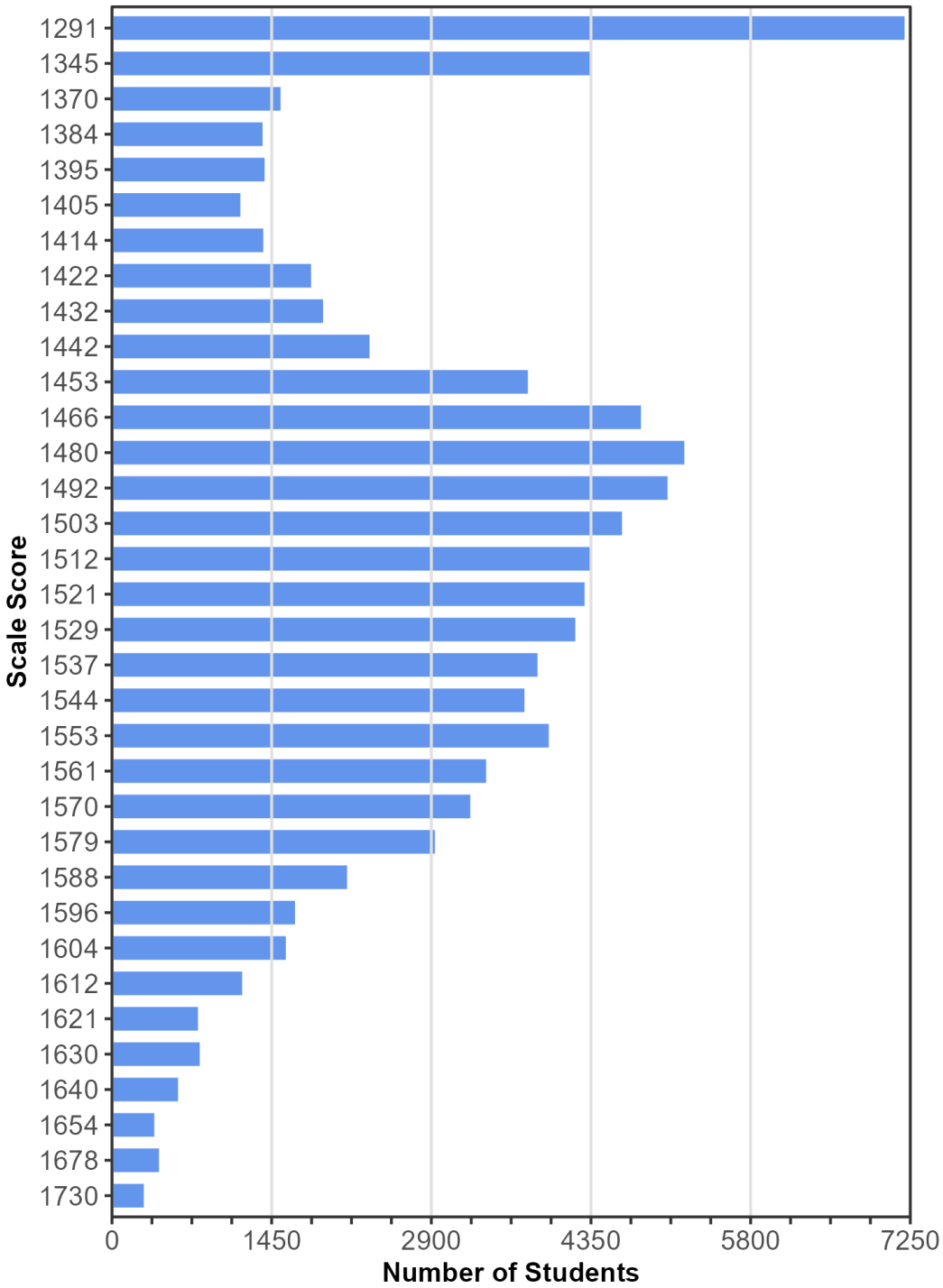
	Freq	Cum Freq	Pct	Cum Pct
1258	11807	11807	11.36	11.36
1320	7875	19682	7.58	18.94
1350	3139	22821	3.02	21.96
1367	2195	25016	2.11	24.08
1380	1960	26976	1.89	25.96
1390	1793	28769	1.73	27.69
1399	2065	30834	1.99	29.68
1409	2295	33129	2.21	31.89
1418	2563	35692	2.47	34.35
1429	3233	38925	3.11	37.46
1440	4028	42953	3.88	41.34
1454	4576	47529	4.40	45.75
1470	5070	52599	4.88	50.63
1485	5270	57869	5.07	55.70
1498	5272	63141	5.07	60.77
1510	4652	67793	4.48	65.25
1519	3895	71688	3.75	69.00
1528	4101	75789	3.95	72.95
1537	3996	79785	3.85	76.79
1545	3277	83062	3.15	79.95
1553	3459	86521	3.33	83.27
1562	3390	89911	3.26	86.54
1571	2205	92116	2.12	88.66
1580	2438	94554	2.35	91.01
1589	2468	97022	2.38	93.38
1598	1178	98200	1.13	94.52
1606	1364	99564	1.31	95.83
1614	1415	100979	1.36	97.19
1623	577	101556	0.56	97.75
1632	667	102223	0.64	98.39
1642	775	102998	0.75	99.13
1656	215	103213	0.21	99.34
1677	295	103508	0.28	99.62
1721	390	103898	0.38	100.00

Figure D.5.37. 2025 TELPAS Grade 5 Writing
Frequency Distribution of Scale Scores



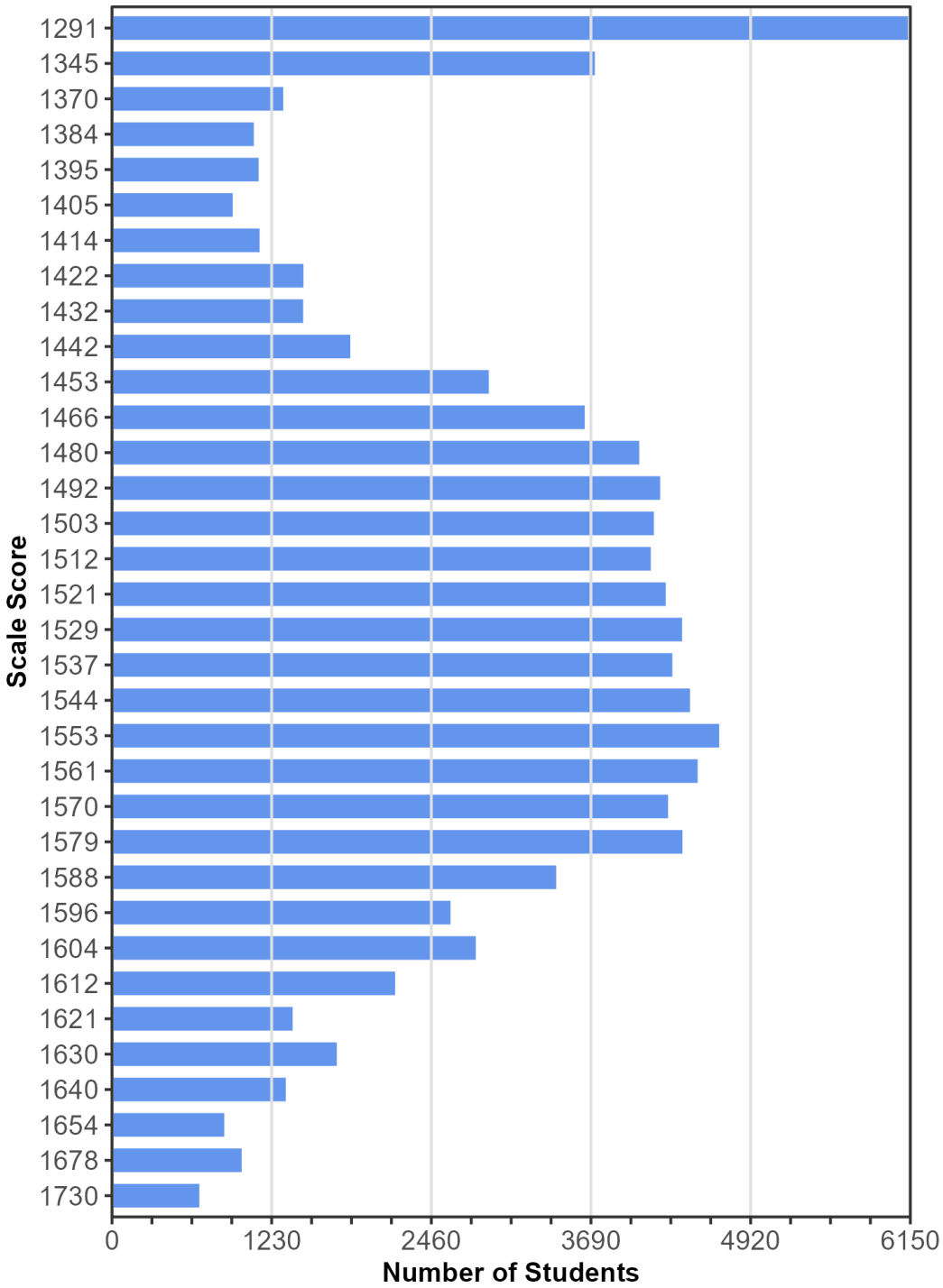
	Freq	Cum Freq	Pct	Cum Pct
1258	8481	8481	8.50	8.50
1320	5291	13772	5.30	13.80
1350	2094	15866	2.10	15.90
1367	1522	17388	1.52	17.42
1380	1314	18702	1.32	18.74
1390	1190	19892	1.19	19.93
1399	1435	21327	1.44	21.37
1409	1548	22875	1.55	22.92
1418	1758	24633	1.76	24.68
1429	2346	26979	2.35	27.03
1440	2935	29914	2.94	29.97
1454	3507	33421	3.51	33.48
1470	3944	37365	3.95	37.44
1485	4261	41626	4.27	41.70
1498	4676	46302	4.68	46.39
1510	4398	50700	4.41	50.80
1519	3702	54402	3.71	54.50
1528	4382	58784	4.39	58.89
1537	4368	63152	4.38	63.27
1545	3673	66825	3.68	66.95
1553	4278	71103	4.29	71.24
1562	4424	75527	4.43	75.67
1571	3196	78723	3.20	78.87
1580	3652	82375	3.66	82.53
1589	3789	86164	3.80	86.33
1598	2117	88281	2.12	88.45
1606	2329	90610	2.33	90.78
1614	2585	93195	2.59	93.37
1623	1210	94405	1.21	94.58
1632	1390	95795	1.39	95.98
1642	1718	97513	1.72	97.70
1656	507	98020	0.51	98.20
1677	670	98690	0.67	98.88
1721	1122	99812	1.12	100.00

Figure D.5.38. 2025 TELPAS Grade 6 Writing
Frequency Distribution of Scale Scores



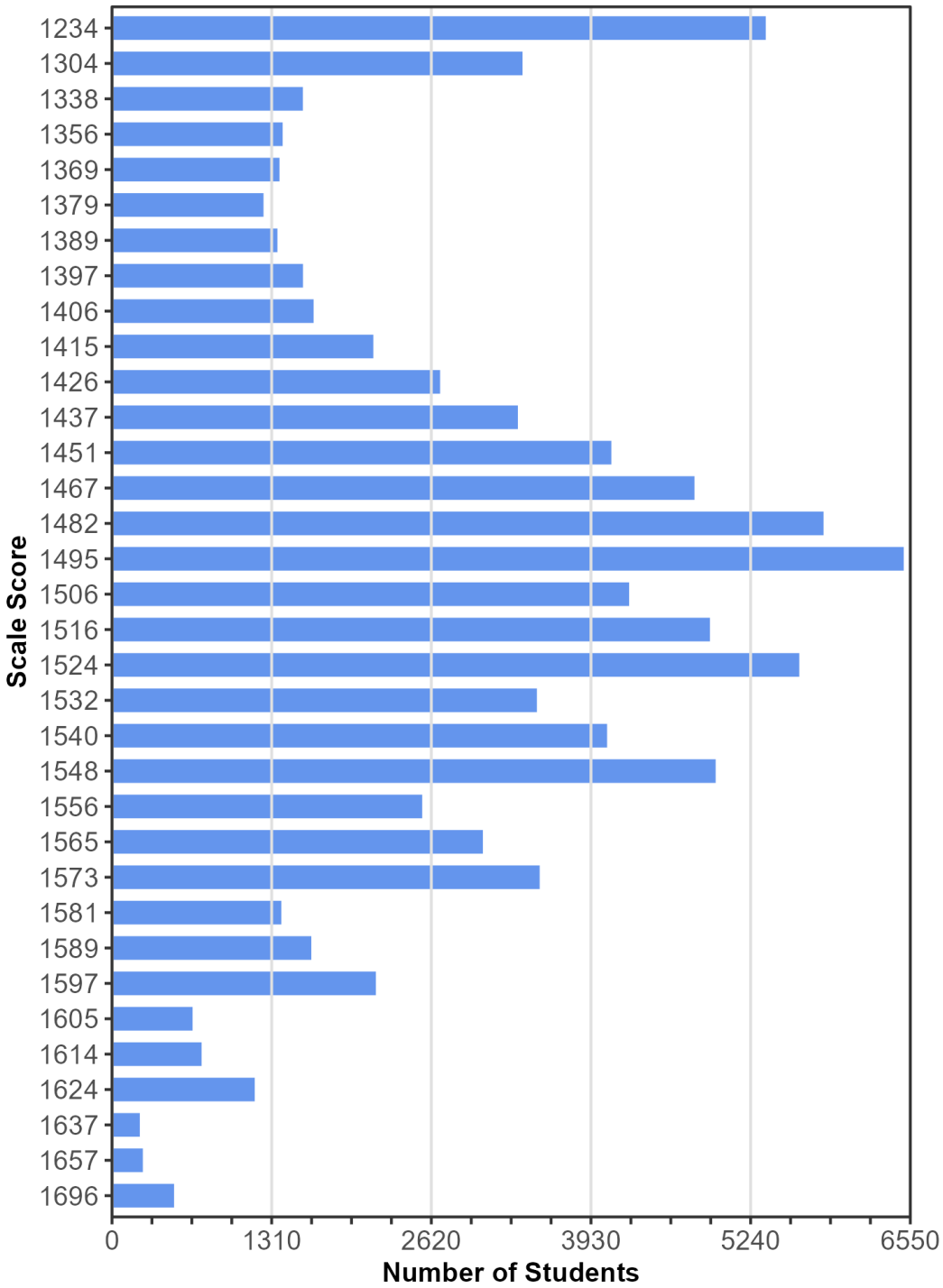
	Freq	Cum Freq	Pct	Cum Pct
1291	7211	7211	7.82	7.82
1345	4356	11567	4.73	12.55
1370	1544	13111	1.68	14.22
1384	1381	14492	1.50	15.72
1395	1399	15891	1.52	17.24
1405	1179	17070	1.28	18.52
1414	1387	18457	1.50	20.02
1422	1821	20278	1.98	22.00
1432	1930	22208	2.09	24.09
1442	2352	24560	2.55	26.64
1453	3790	28350	4.11	30.76
1466	4817	33167	5.23	35.98
1480	5211	38378	5.65	41.64
1492	5059	43437	5.49	47.12
1503	4645	48082	5.04	52.16
1512	4360	52442	4.73	56.89
1521	4305	56747	4.67	61.56
1529	4221	60968	4.58	66.14
1537	3878	64846	4.21	70.35
1544	3759	68605	4.08	74.43
1553	3980	72585	4.32	78.75
1561	3411	75996	3.70	82.45
1570	3267	79263	3.54	85.99
1579	2946	82209	3.20	89.19
1588	2148	84357	2.33	91.52
1596	1675	86032	1.82	93.34
1604	1592	87624	1.73	95.06
1612	1195	88819	1.30	96.36
1621	794	89613	0.86	97.22
1630	810	90423	0.88	98.10
1640	613	91036	0.67	98.76
1654	397	91433	0.43	99.20
1678	440	91873	0.48	99.67
1730	302	92175	0.33	100.00

Figure D.5.39. 2025 TELPAS Grade 7 Writing
Frequency Distribution of Scale Scores



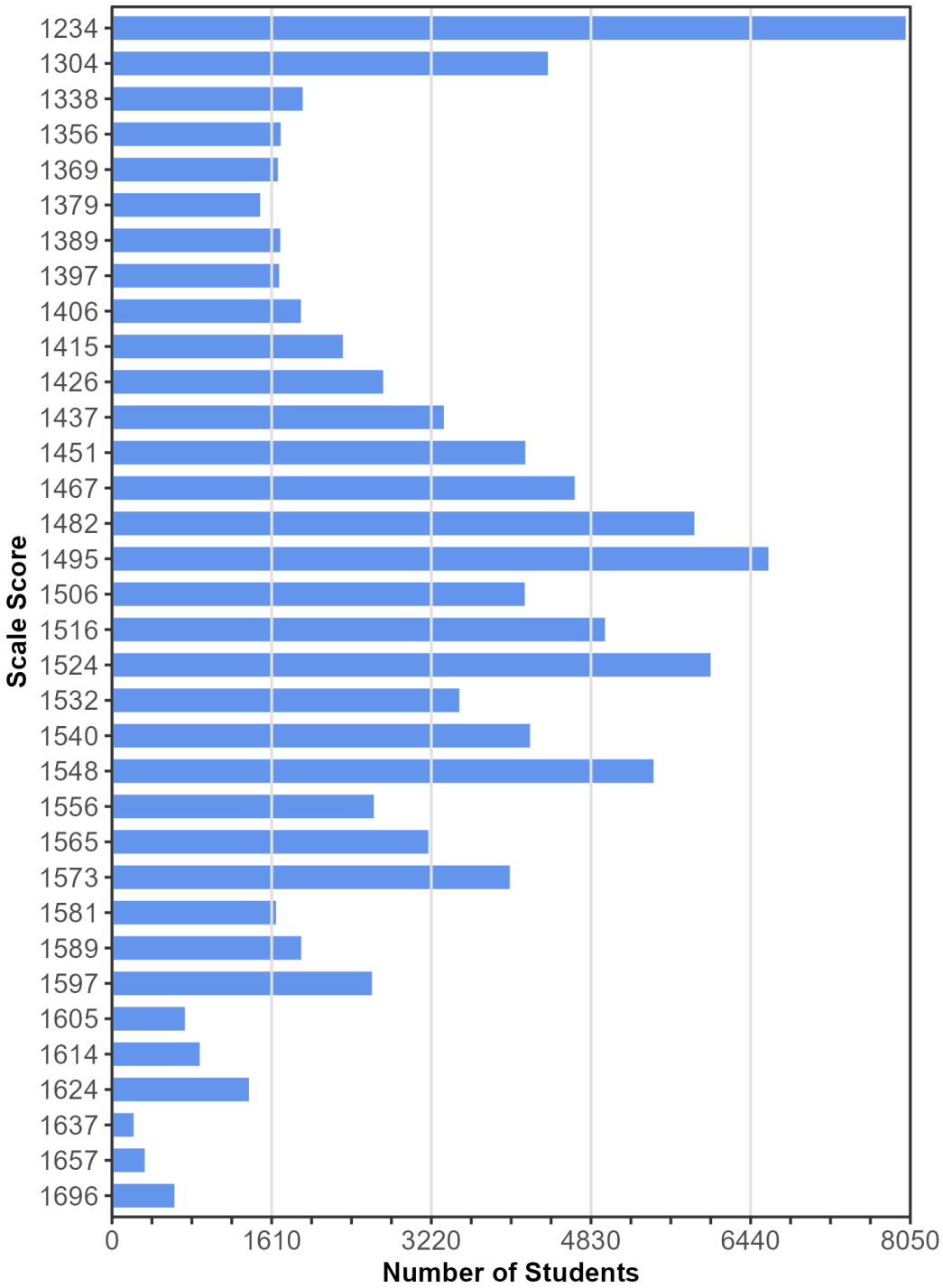
Scale Score	Freq	Cum Freq	Pct	Cum Pct
1291	6143	6143	6.33	6.33
1345	3731	9874	3.84	10.17
1370	1330	11204	1.37	11.54
1384	1104	12308	1.14	12.68
1395	1141	13449	1.18	13.85
1405	941	14390	0.97	14.82
1414	1148	15538	1.18	16.00
1422	1486	17024	1.53	17.53
1432	1484	18508	1.53	19.06
1442	1847	20355	1.90	20.97
1453	2914	23269	3.00	23.97
1466	3653	26922	3.76	27.73
1480	4073	30995	4.20	31.93
1492	4234	35229	4.36	36.29
1503	4186	39415	4.31	40.60
1512	4162	43577	4.29	44.88
1521	4277	47854	4.41	49.29
1529	4403	52257	4.54	53.83
1537	4328	56585	4.46	58.28
1544	4464	61049	4.60	62.88
1553	4689	65738	4.83	67.71
1561	4523	70261	4.66	72.37
1570	4295	74556	4.42	76.79
1579	4406	78962	4.54	81.33
1588	3433	82395	3.54	84.87
1596	2619	85014	2.70	87.57
1604	2814	87828	2.90	90.46
1612	2192	90020	2.26	92.72
1621	1403	91423	1.45	94.17
1630	1743	93166	1.80	95.96
1640	1350	94516	1.39	97.35
1654	876	95392	0.90	98.26
1678	1010	96402	1.04	99.30
1730	684	97086	0.70	100.00

Figure D.5.40. 2025 TELPAS Grade 8 Writing
Frequency Distribution of Scale Scores



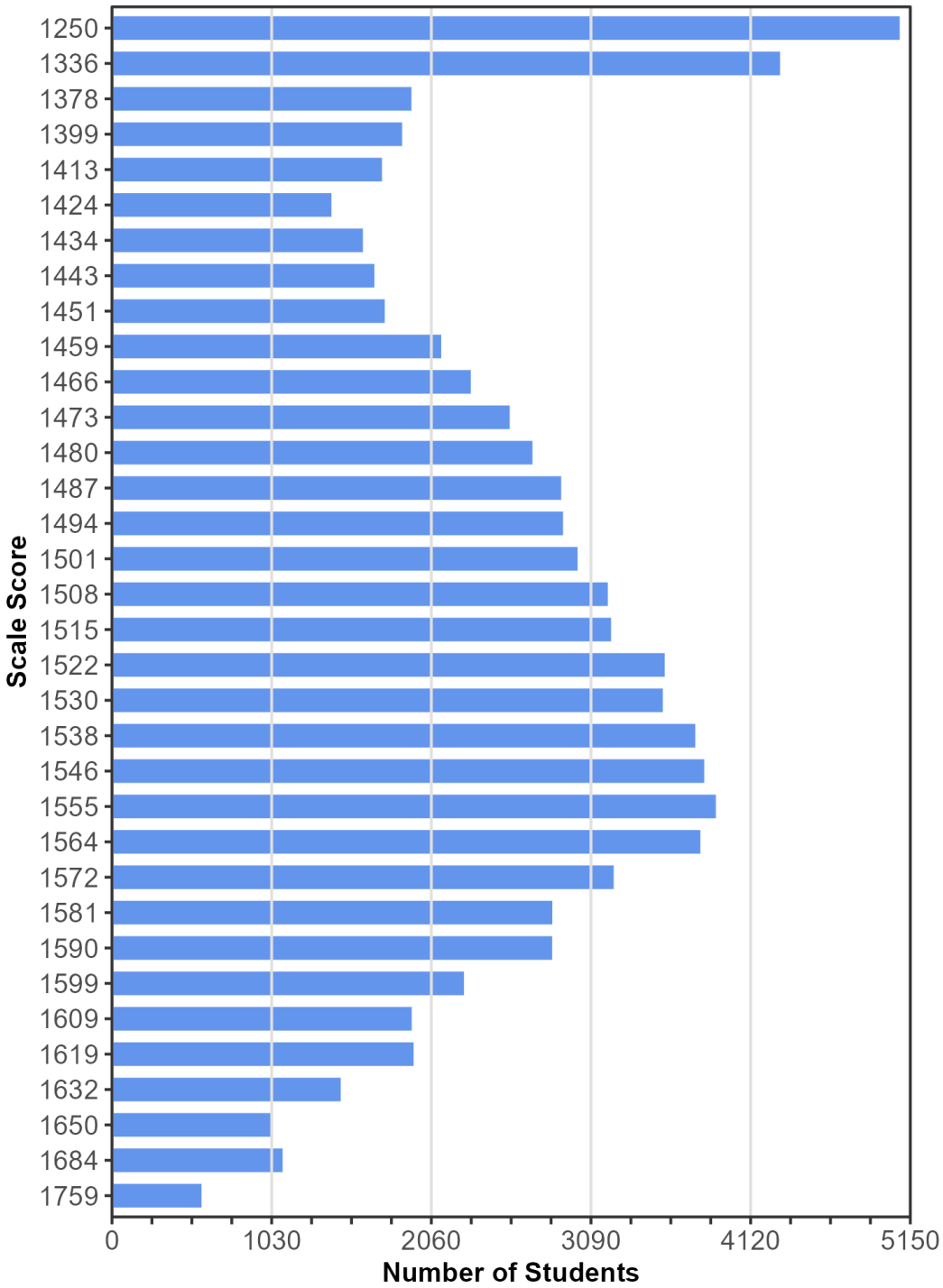
	Freq	Cum Freq	Pct	Cum Pct
1234	5376	5376	5.73	5.73
1304	3380	8756	3.60	9.33
1338	1578	10334	1.68	11.02
1356	1412	11746	1.51	12.52
1369	1386	13132	1.48	14.00
1379	1255	14387	1.34	15.34
1389	1369	15756	1.46	16.80
1397	1579	17335	1.68	18.48
1406	1666	19001	1.78	20.25
1415	2156	21157	2.30	22.55
1426	2704	23861	2.88	25.44
1437	3342	27203	3.56	29.00
1451	4109	31312	4.38	33.38
1467	4791	36103	5.11	38.49
1482	5850	41953	6.24	44.72
1495	6508	48461	6.94	51.66
1506	4255	52716	4.54	56.20
1516	4918	57634	5.24	61.44
1524	5651	63285	6.02	67.46
1532	3498	66783	3.73	71.19
1540	4074	70857	4.34	75.53
1548	4965	75822	5.29	80.83
1556	2557	78379	2.73	83.55
1565	3055	81434	3.26	86.81
1573	3521	84955	3.75	90.56
1581	1401	86356	1.49	92.06
1589	1647	88003	1.76	93.81
1597	2177	90180	2.32	96.13
1605	673	90853	0.72	96.85
1614	747	91600	0.80	97.65
1624	1183	92783	1.26	98.91
1637	240	93023	0.26	99.16
1657	265	93288	0.28	99.44
1696	521	93809	0.56	100.00

Figure D.5.41. 2025 TELPAS Grade 9 Writing
Frequency Distribution of Scale Scores



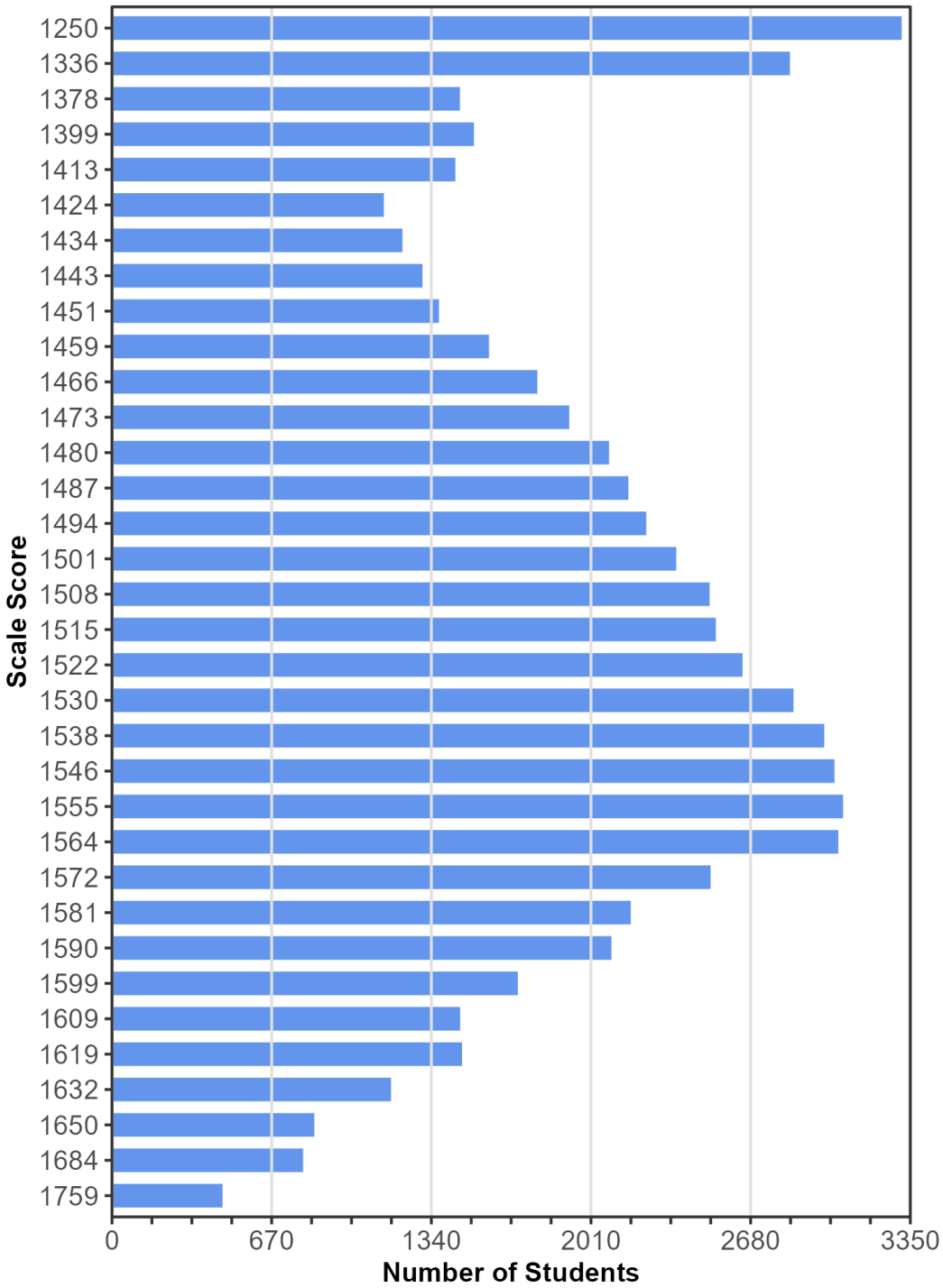
	Freq	Cum Freq	Pct	Cum Pct
1234	8018	8018	7.76	7.76
1304	4410	12428	4.27	12.04
1338	1938	14366	1.88	13.91
1356	1715	16081	1.66	15.57
1369	1688	17769	1.63	17.21
1379	1509	19278	1.46	18.67
1389	1711	20989	1.66	20.33
1397	1700	22689	1.65	21.97
1406	1920	24609	1.86	23.83
1415	2343	26952	2.27	26.10
1426	2749	29701	2.66	28.76
1437	3361	33062	3.25	32.02
1451	4183	37245	4.05	36.07
1467	4681	41926	4.53	40.60
1482	5887	47813	5.70	46.30
1495	6634	54447	6.42	52.73
1506	4177	58624	4.04	56.77
1516	4986	63610	4.83	61.60
1524	6051	69661	5.86	67.46
1532	3517	73178	3.41	70.86
1540	4230	77408	4.10	74.96
1548	5476	82884	5.30	80.26
1556	2655	85539	2.57	82.83
1565	3205	88744	3.10	85.94
1573	4026	92770	3.90	89.84
1581	1669	94439	1.62	91.45
1589	1923	96362	1.86	93.32
1597	2636	98998	2.55	95.87
1605	750	99748	0.73	96.59
1614	899	100647	0.87	97.46
1624	1396	102043	1.35	98.82
1637	233	102276	0.23	99.04
1657	345	102621	0.33	99.38
1696	644	103265	0.62	100.00

Figure D.5.42. 2025 TELPAS Grade 10 Writing
Frequency Distribution of Scale Scores



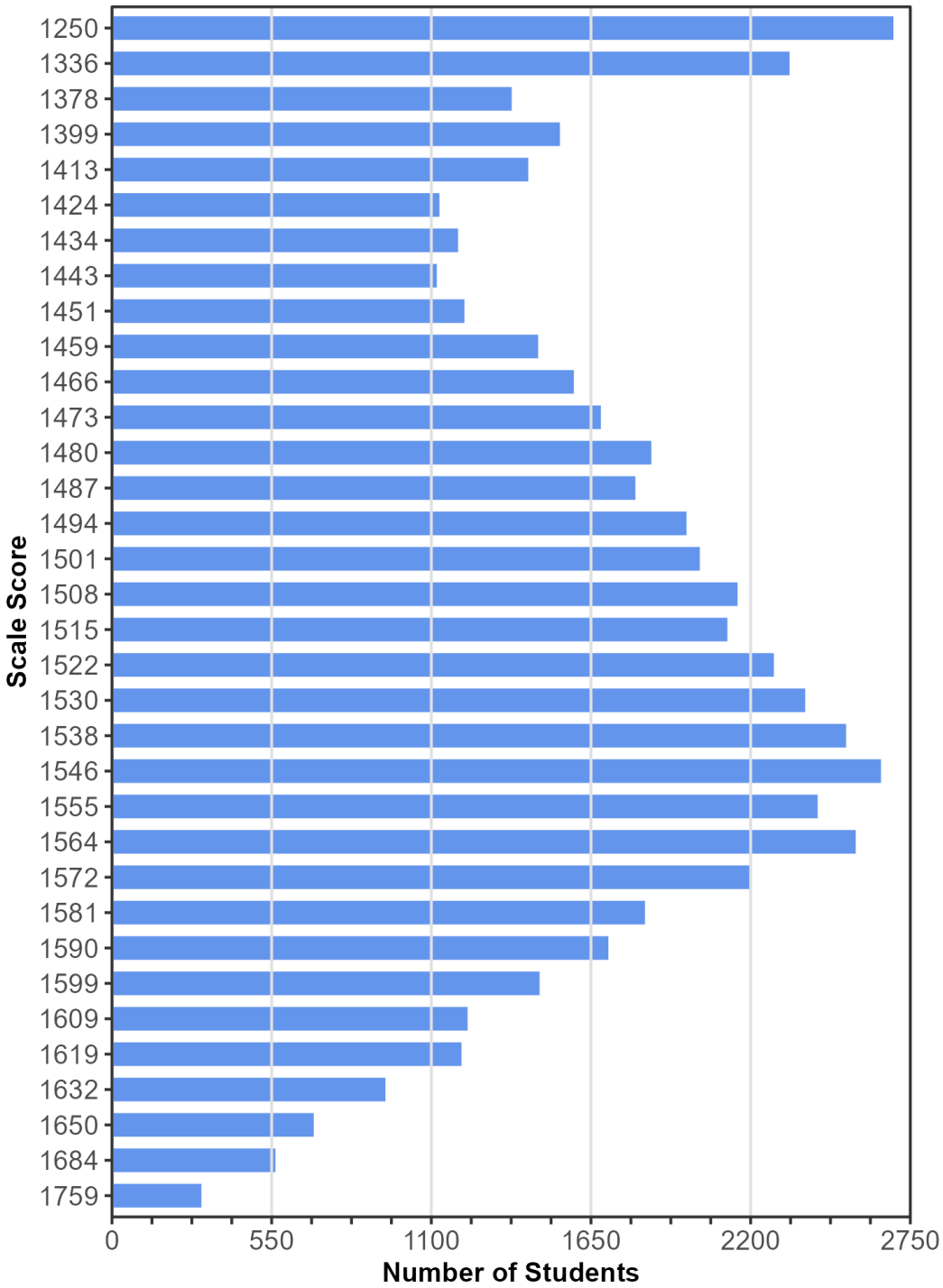
	Freq	Cum Freq	Pct	Cum Pct
1250	5091	5091	5.76	5.76
1336	4319	9410	4.89	10.65
1378	1941	11351	2.20	12.85
1399	1881	13232	2.13	14.98
1413	1751	14983	1.98	16.96
1424	1425	16408	1.61	18.57
1434	1628	18036	1.84	20.42
1443	1702	19738	1.93	22.34
1451	1769	21507	2.00	24.35
1459	2133	23640	2.41	26.76
1466	2324	25964	2.63	29.39
1473	2575	28539	2.91	32.31
1480	2722	31261	3.08	35.39
1487	2907	34168	3.29	38.68
1494	2919	37087	3.30	41.98
1501	3013	40100	3.41	45.39
1508	3208	43308	3.63	49.02
1515	3229	46537	3.66	52.68
1522	3575	50112	4.05	56.73
1530	3563	53675	4.03	60.76
1538	3772	57447	4.27	65.03
1546	3830	61277	4.34	69.37
1555	3905	65182	4.42	73.79
1564	3805	68987	4.31	78.09
1572	3246	72233	3.67	81.77
1581	2850	75083	3.23	84.99
1590	2849	77932	3.23	88.22
1599	2280	80212	2.58	90.80
1609	1943	82155	2.20	93.00
1619	1955	84110	2.21	95.21
1632	1485	85595	1.68	96.89
1650	1047	86642	1.19	98.08
1684	1110	87752	1.26	99.34
1759	587	88339	0.66	100.00

Figure D.5.43. 2025 TELPAS Grade 11 Writing
Frequency Distribution of Scale Scores



	Freq	Cum Freq	Pct	Cum Pct
1250	3320	3320	4.93	4.93
1336	2851	6171	4.23	9.16
1378	1466	7637	2.18	11.34
1399	1525	9162	2.26	13.60
1413	1447	10609	2.15	15.75
1424	1147	11756	1.70	17.45
1434	1225	12981	1.82	19.27
1443	1309	14290	1.94	21.22
1451	1378	15668	2.05	23.26
1459	1588	17256	2.36	25.62
1466	1791	19047	2.66	28.28
1473	1925	20972	2.86	31.14
1480	2092	23064	3.11	34.24
1487	2173	25237	3.23	37.47
1494	2248	27485	3.34	40.81
1501	2374	29859	3.52	44.33
1508	2514	32373	3.73	48.06
1515	2540	34913	3.77	51.83
1522	2652	37565	3.94	55.77
1530	2866	40431	4.26	60.03
1538	2995	43426	4.45	64.47
1546	3038	46464	4.51	68.98
1555	3074	49538	4.56	73.55
1564	3054	52592	4.53	78.08
1572	2518	55110	3.74	81.82
1581	2183	57293	3.24	85.06
1590	2102	59395	3.12	88.18
1599	1709	61104	2.54	90.72
1609	1467	62571	2.18	92.90
1619	1475	64046	2.19	95.09
1632	1177	65223	1.75	96.83
1650	855	66078	1.27	98.10
1684	808	66886	1.20	99.30
1759	470	67356	0.70	100.00

Figure D.5.44. 2025 TELPAS Grade 12 Writing
Frequency Distribution of Scale Scores



	Freq	Cum Freq	Pct	Cum Pct
1250	2697	2697	4.66	4.66
1336	2339	5036	4.04	8.70
1378	1382	6418	2.39	11.09
1399	1548	7966	2.67	13.76
1413	1439	9405	2.49	16.25
1424	1133	10538	1.96	18.21
1434	1197	11735	2.07	20.28
1443	1124	12859	1.94	22.22
1451	1219	14078	2.11	24.32
1459	1473	15551	2.55	26.87
1466	1596	17147	2.76	29.63
1473	1689	18836	2.92	32.54
1480	1863	20699	3.22	35.76
1487	1808	22507	3.12	38.89
1494	1984	24491	3.43	42.32
1501	2030	26521	3.51	45.82
1508	2160	28681	3.73	49.56
1515	2125	30806	3.67	53.23
1522	2285	33091	3.95	57.17
1530	2393	35484	4.13	61.31
1538	2534	38018	4.38	65.69
1546	2654	40672	4.59	70.27
1555	2436	43108	4.21	74.48
1564	2567	45675	4.44	78.92
1572	2201	47876	3.80	82.72
1581	1841	49717	3.18	85.90
1590	1715	51432	2.96	88.86
1599	1478	52910	2.55	91.42
1609	1230	54140	2.13	93.54
1619	1209	55349	2.09	95.63
1632	947	56296	1.64	97.27
1650	700	56996	1.21	98.48
1684	568	57564	0.98	99.46
1759	313	57877	0.54	100.00

2025 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	2025	107156	44	29	17	10	107084	53	24	15	8	106954	65	17	11	7	107004	67	18	9	6	106809	51	28	13	7
	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	2025	110053	24	33	24	19	109906	34	30	20	16	109810	45	25	16	14	109939	49	26	14	11	109697	31	34	20	15
	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

			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**
2	2025	103734	9	31	34	26	103702	39	46	11	4	103783	39	39	15	6	103785	52	31	15	2	103673	24	52	21	3
	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	2025	107643	4	17	26	53	107627	27	46	18	9	107726	30	32	19	20	107724	37	41	19	4	107621	15	44	32	9
	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	2025	103861	20	28	30	22	103857	21	34	39	6	103936	13	37	27	24	103931	30	39	26	5	103819	14	38	37	11
	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	2025	99748	14	19	29	37	99741	20	31	39	9	99843	9	27	25	38	99837	21	33	34	12	99710	11	30	39	20
	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	2025	92128	10	34	33	24	92125	15	45	36	4	92194	21	30	28	21	92191	22	40	32	7	92044	11	38	41	10
	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	2025	97088	8	26	31	35	97085	17	44	35	4	97111	17	24	27	32	97109	18	32	38	12	96955	9	32	43	16
	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

			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**
8	2025	93795	7	23	28	41	93790	19	41	35	5	93839	8	37	30	26	93837	20	47	29	4	93633	7	37	44	12
	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	2025	103255	16	24	40	19	103249	38	21	35	7	103442	9	36	26	29	103439	24	44	28	4	102829	14	40	35	11
	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	2025	88385	12	21	41	26	88383	36	20	36	9	88592	14	41	27	18	88589	22	34	34	9	88006	13	40	35	13
	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	2025	67640	11	21	42	26	67639	37	19	35	8	67725	12	40	27	20	67726	21	35	35	9	67251	11	42	35	12
	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	2025	57743	12	24	41	23	57743	42	20	31	7	57910	13	43	26	18	57910	22	35	34	9	57373	12	46	32	10
	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 2024–2025

Appendix E

TELPAS Alternate Statistical Tables and Figures

List of Tables

Table E.1.1. 2025 TELPAS Alternate Classification Consistency and Accuracy

Table E.2.1. 2025 TELPAS Alternate Scale Score Correlations

Table E.3.1. 2025 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

Table E.4.1. 2025 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Total Group

Table E.4.2. 2025 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Female

Table E.4.3. 2025 TELPAS Alternate Mean P-Values, Raw Score Summary, and Reliability Male

Table E.5.1. 2025 TELPAS Alternate Scale Score Descriptive Statistics

List of Figures

Figure E.5.1. 2025 TELPAS Alternate Listening Frequency Distribution of Scale Scores

Figure E.5.2. 2025 TELPAS Alternate Speaking Frequency Distribution of Scale Scores

Figure E.5.3. 2025 TELPAS Alternate Reading Frequency Distribution of Scale Scores

Figure E.5.4. 2025 TELPAS Alternate Writing Frequency Distribution of Scale Scores

2025 TELPAS Alternate Classification Consistency and Accuracy

Table E.1.1. 2025 TELPAS Alternate Classification Consistency and Accuracy

Domain	Decision Consistency[1]	Decision Accuracy[2]
Listening	79.4	85.3
Speaking	83.9	88.5
Reading	81.0	86.4
Writing	82.8	87.7

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.

2025 TELPAS Alternate
Scale Score Correlations

Table E.2.1. 2025 TELPAS Alternate Scale Score Correlations

TELPAS Alternate Domain	TELPAS Alternate Domain	N	Correlation
Listening	Speaking	11,902	0.89
Listening	Reading	11,902	0.90
Listening	Writing	11,902	0.86
Speaking	Reading	11,902	0.84
Speaking	Writing	11,902	0.83
Reading	Writing	11,902	0.92

2025 TELPAS Alternate Conditional Standard Error of Measurement for Scale Scores

Table E.3.1. 2025 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.

2025 TELPAS Alternate
Mean P-Values, Raw Score Summary, and Reliability

Table E.4.1. 2025 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	11,902	29.26	11.96	0.98	1.87	48.15
	Understand spoken words and language structures	20	11,902	11.57	5.06	0.94	1.19	47.33
	Basic understanding of spoken English	30	11,902	17.69	7.10	0.96	1.38	48.70
Speaking	OVERALL TEST	50	11,902	27.32	12.92	0.98	1.61	43.30
	Provide and summarize information	35	11,902	19.02	9.18	0.98	1.35	42.94
	Share opinions and analyze information	15	11,902	8.30	3.85	0.95	0.86	44.13
Reading	OVERALL TEST	50	11,902	25.15	11.80	0.98	1.71	37.88
	Understand words and language structures	25	11,902	13.27	6.26	0.97	1.12	41.35
	Basic understanding of a variety of texts written in English	15	11,902	7.23	3.54	0.94	0.86	35.26
	Analyze and evaluate information and ideas in a variety of texts written in English	10	11,902	4.65	2.40	0.94	0.60	33.15
Writing	OVERALL TEST	50	11,902	22.15	11.04	0.98	1.62	30.37
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	11,902	9.75	4.83	0.96	1.01	35.94
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	11,902	12.40	6.46	0.97	1.08	26.65

Notes:

1. Total number of score points.
2. Mean of percent correct (0–100%) for the items.

Table E.4.2. 2025 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	3,796	29.12	11.95	0.98	1.87	47.80
	Understand spoken words and language structures	20	3,796	11.47	5.06	0.95	1.18	46.70
	Basic understanding of spoken English	30	3,796	17.65	7.09	0.96	1.38	48.54
Speaking	OVERALL TEST	50	3,796	27.63	12.93	0.98	1.62	44.06
	Provide and summarize information	35	3,796	19.22	9.20	0.98	1.36	43.66
	Share opinions and analyze information	15	3,796	8.40	3.84	0.95	0.85	45.01
Reading	OVERALL TEST	50	3,796	24.70	11.69	0.98	1.65	36.76
	Understand words and language structures	25	3,796	12.94	6.18	0.97	1.08	39.68
	Basic understanding of a variety of texts written in English	15	3,796	7.14	3.49	0.94	0.85	34.46
	Analyze and evaluate information and ideas in a variety of texts written in English	10	3,796	4.63	2.37	0.93	0.61	32.92
Writing	OVERALL TEST	50	3,796	22.04	11.03	0.98	1.58	30.10
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	3,796	9.64	4.81	0.96	0.98	35.24
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	3,796	12.40	6.46	0.97	1.07	26.67

Notes:

1. Total number of score points.
2. Mean of percent correct (0–100%) for the items.

Table E.4.3. 2025 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	8,103	29.33	11.96	0.98	1.87	48.32
	Understand spoken words and language structures	20	8,103	11.62	5.06	0.94	1.19	47.62
	Basic understanding of spoken English	30	8,103	17.71	7.10	0.96	1.38	48.78
Speaking	OVERALL TEST	50	8,103	27.18	12.92	0.98	1.61	42.94
	Provide and summarize information	35	8,103	18.93	9.17	0.98	1.35	42.61
	Share opinions and analyze information	15	8,103	8.25	3.85	0.95	0.86	43.72
Reading	OVERALL TEST	50	8,103	25.36	11.85	0.98	1.73	38.41
	Understand words and language structures	25	8,103	13.43	6.29	0.97	1.13	42.13
	Basic understanding of a variety of texts written in English	15	8,103	7.28	3.56	0.94	0.87	35.64
	Analyze and evaluate information and ideas in a variety of texts written in English	10	8,103	4.66	2.41	0.94	0.60	33.26
Writing	OVERALL TEST	50	8,103	22.20	11.04	0.98	1.64	30.50
	Demonstrate an ability to use English vocabulary and language structures in a variety of academic and social situations	20	8,103	9.80	4.85	0.95	1.03	36.28
	Demonstrate an ability to apply knowledge of English to complete a variety of writing tasks	30	8,103	12.39	6.47	0.97	1.09	26.64

Notes:

1. Total number of score points.
2. Mean of percent correct (0–100%) for the items.

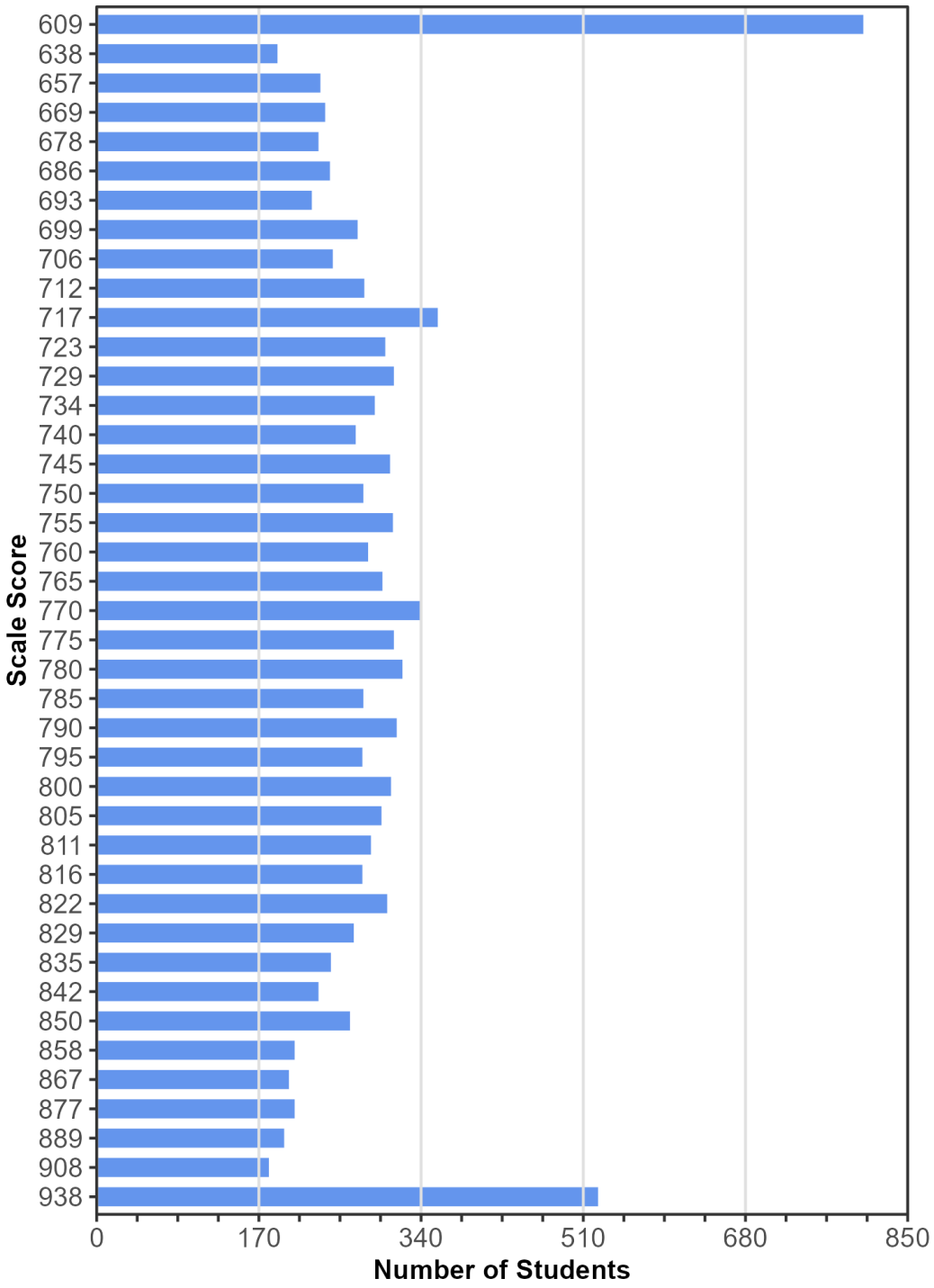
2025 TELPAS Alternate
Scale Score Descriptive Statistics

Table E.5.1. 2025 TELPAS Alternate Scale Score Descriptive Statistics

Subject	N	Mean	Median	Mode	Range	Interquartile Range	SD	Variance	Skewness	Kurtosis
TELPAS Alternate Listening	11,902	765.47	765	609	329	104	79.92	6387.15	0.04	-0.26
TELPAS Alternate Speaking	11,902	752.04	750	600	352	145	98.60	9722.23	0.17	-0.75
TELPAS Alternate Reading	11,902	742.74	743	601	347	126	88.49	7831.11	0.11	-0.60
TELPAS Alternate Writing	11,902	726.22	728	612	343	122	83.65	6997.84	0.33	-0.48

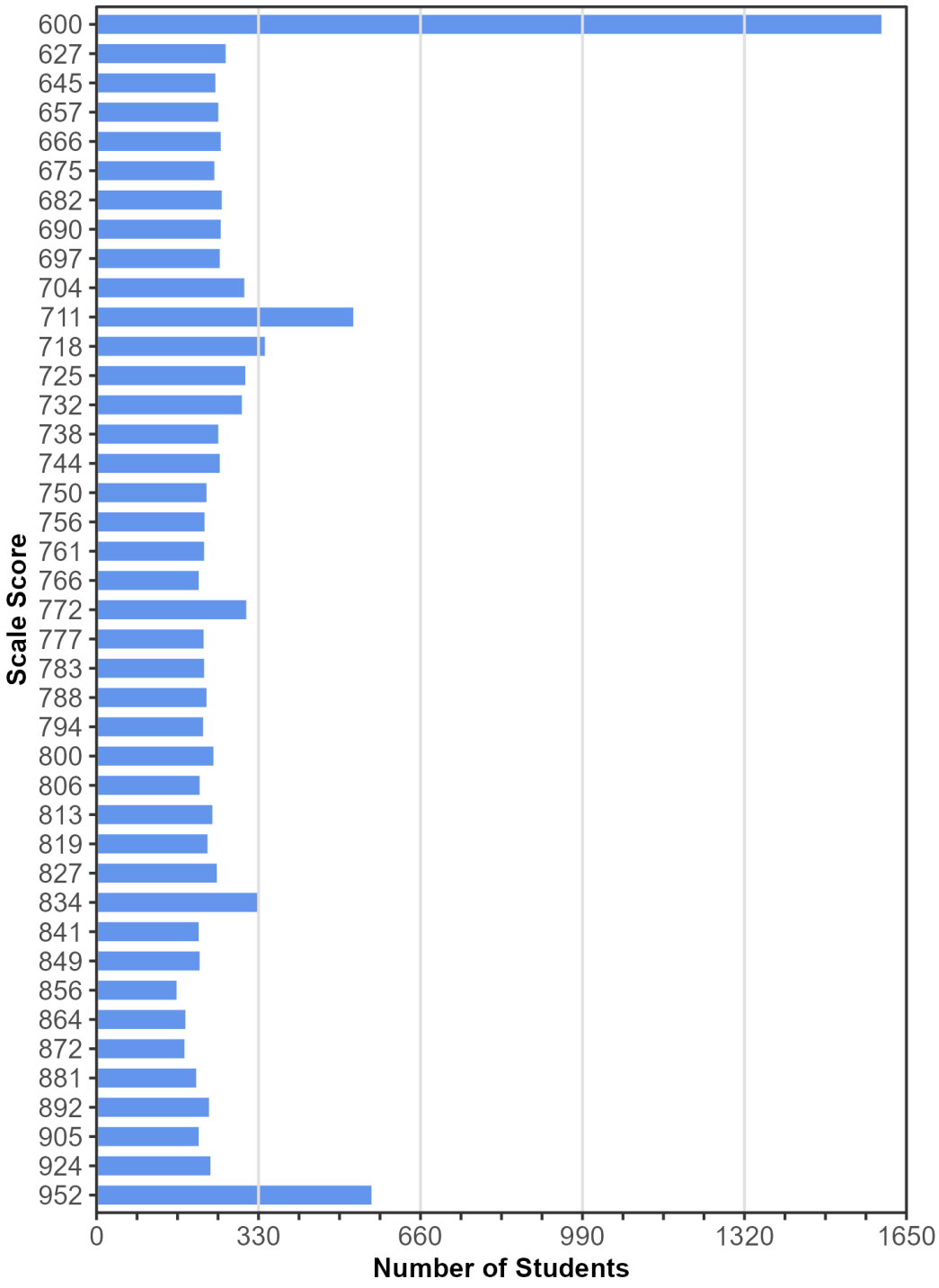
2025 TELPAS Alternate
Frequency Distribution of Scale Scores

Figure E.5.1. 2025 TELPAS Alternate Listening
Frequency Distribution of Scale Scores



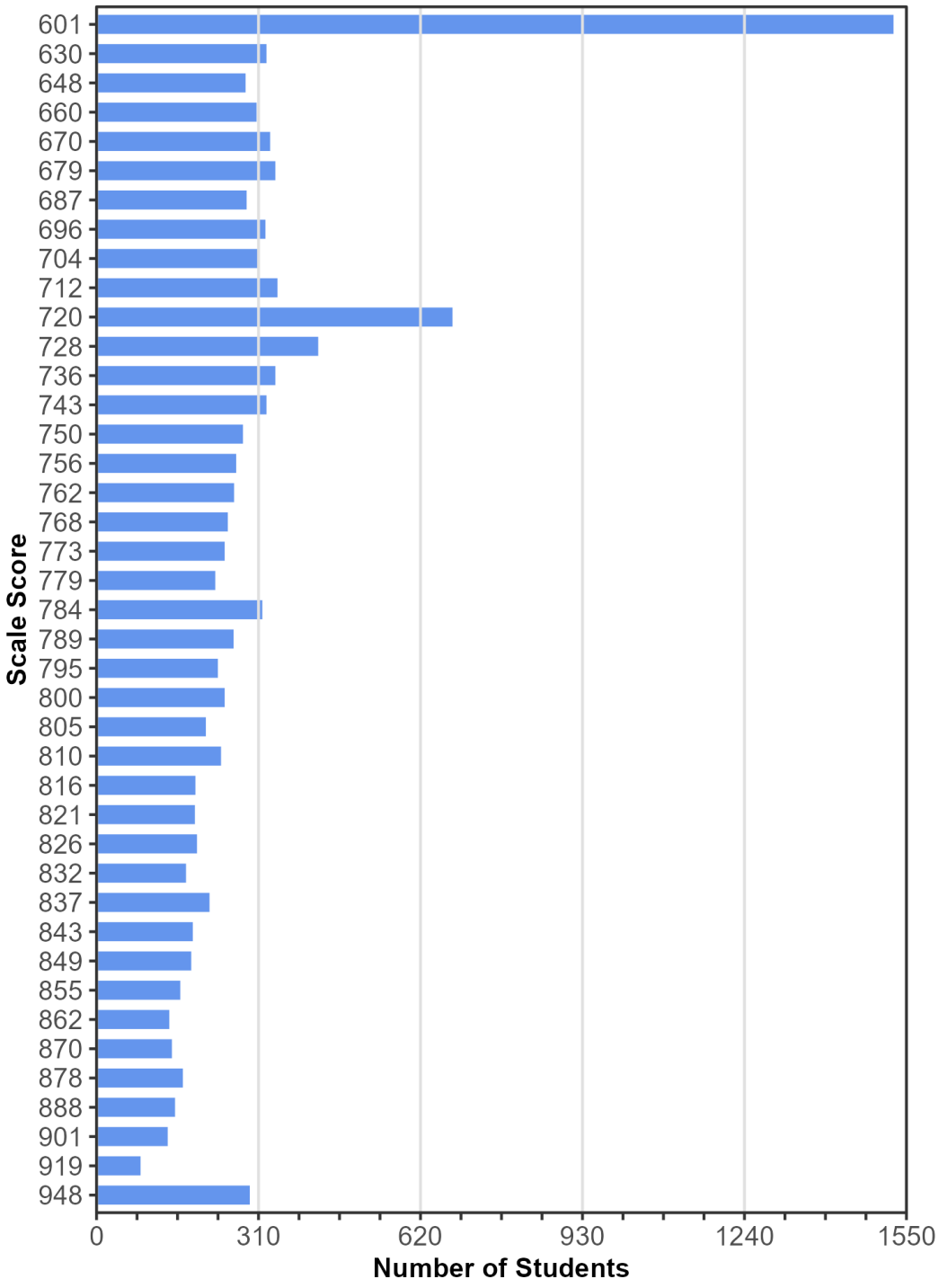
	Freq	Cum Freq	Pct	Cum Pct
609	805	805	6.76	6.76
638	191	996	1.60	8.37
657	236	1232	1.98	10.35
669	241	1473	2.02	12.38
678	234	1707	1.97	14.34
686	246	1953	2.07	16.41
693	227	2180	1.91	18.32
699	275	2455	2.31	20.63
706	249	2704	2.09	22.72
712	282	2986	2.37	25.09
717	359	3345	3.02	28.10
723	304	3649	2.55	30.66
729	313	3962	2.63	33.29
734	293	4255	2.46	35.75
740	273	4528	2.29	38.04
745	309	4837	2.60	40.64
750	281	5118	2.36	43.00
755	312	5430	2.62	45.62
760	286	5716	2.40	48.03
765	301	6017	2.53	50.55
770	341	6358	2.87	53.42
775	313	6671	2.63	56.05
780	322	6993	2.71	58.75
785	281	7274	2.36	61.12
790	316	7590	2.66	63.77
795	280	7870	2.35	66.12
800	310	8180	2.60	68.73
805	300	8480	2.52	71.25
811	289	8769	2.43	73.68
816	280	9049	2.35	76.03
822	306	9355	2.57	78.60
829	271	9626	2.28	80.88
835	247	9873	2.08	82.95
842	234	10107	1.97	84.92
850	267	10374	2.24	87.16
858	209	10583	1.76	88.92
867	203	10786	1.71	90.62
877	209	10995	1.76	92.38
889	198	11193	1.66	94.04
908	182	11375	1.53	95.57
938	527	11902	4.43	100.00

Figure E.5.2. 2025 TELPAS Alternate Speaking
Frequency Distribution of Scale Scores



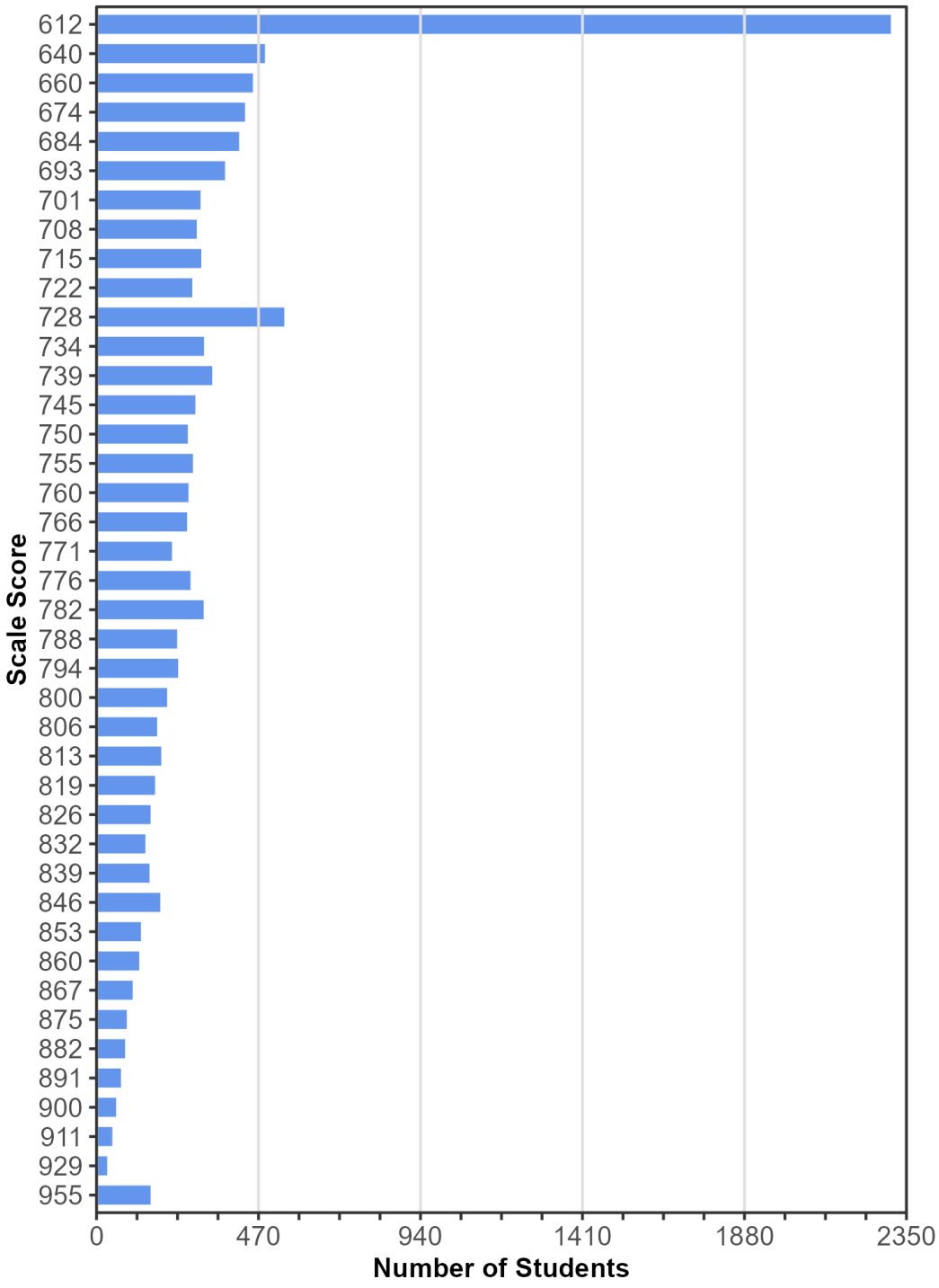
	Freq	Cum Freq	Pct	Cum Pct
600	1602	1602	13.46	13.46
627	266	1868	2.23	15.69
645	245	2113	2.06	17.75
657	251	2364	2.11	19.86
666	256	2620	2.15	22.01
675	243	2863	2.04	24.05
682	258	3121	2.17	26.22
690	256	3377	2.15	28.37
697	254	3631	2.13	30.51
704	304	3935	2.55	33.06
711	526	4461	4.42	37.48
718	346	4807	2.91	40.39
725	306	5113	2.57	42.96
732	299	5412	2.51	45.47
738	251	5663	2.11	47.58
744	254	5917	2.13	49.71
750	227	6144	1.91	51.62
756	223	6367	1.87	53.50
761	222	6589	1.87	55.36
766	211	6800	1.77	57.13
772	308	7108	2.59	59.72
777	221	7329	1.86	61.58
783	222	7551	1.87	63.44
788	227	7778	1.91	65.35
794	220	7998	1.85	67.20
800	241	8239	2.02	69.22
806	213	8452	1.79	71.01
813	239	8691	2.01	73.02
819	229	8920	1.92	74.95
827	248	9168	2.08	77.03
834	331	9499	2.78	79.81
841	211	9710	1.77	81.58
849	213	9923	1.79	83.37
856	166	10089	1.39	84.77
864	184	10273	1.55	86.31
872	182	10455	1.53	87.84
881	206	10661	1.73	89.57
892	232	10893	1.95	91.52
905	211	11104	1.77	93.30
924	235	11339	1.97	95.27
952	563	11902	4.73	100.00

Figure E.5.3. 2025 TELPAS Alternate Reading
Frequency Distribution of Scale Scores



Scale Score	Freq	Cum Freq	Pct	Cum Pct
601	1528	1528	12.84	12.84
630	328	1856	2.76	15.59
648	288	2144	2.42	18.01
660	309	2453	2.60	20.61
670	335	2788	2.81	23.42
679	345	3133	2.90	26.32
687	290	3423	2.44	28.76
696	326	3749	2.74	31.50
704	311	4060	2.61	34.11
712	349	4409	2.93	37.04
720	684	5093	5.75	42.79
728	427	5520	3.59	46.38
736	345	5865	2.90	49.28
743	328	6193	2.76	52.03
750	283	6476	2.38	54.41
756	270	6746	2.27	56.68
762	266	7012	2.23	58.91
768	254	7266	2.13	61.05
773	248	7514	2.08	63.13
779	230	7744	1.93	65.06
784	320	8064	2.69	67.75
789	265	8329	2.23	69.98
795	235	8564	1.97	71.95
800	248	8812	2.08	74.04
805	212	9024	1.78	75.82
810	241	9265	2.02	77.84
816	192	9457	1.61	79.46
821	191	9648	1.60	81.06
826	195	9843	1.64	82.70
832	174	10017	1.46	84.16
837	219	10236	1.84	86.00
843	187	10423	1.57	87.57
849	184	10607	1.55	89.12
855	163	10770	1.37	90.49
862	142	10912	1.19	91.68
870	147	11059	1.24	92.92
878	168	11227	1.41	94.33
888	153	11380	1.29	95.61
901	139	11519	1.17	96.78
919	87	11606	0.73	97.51
948	296	11902	2.49	100.00

Figure E.5.4. 2025 TELPAS Alternate Writing
Frequency Distribution of Scale Scores



	Freq	Cum Freq	Pct	Cum Pct
612	2309	2309	19.40	19.40
640	493	2802	4.14	23.54
660	458	3260	3.85	27.39
674	435	3695	3.65	31.05
684	418	4113	3.51	34.56
693	377	4490	3.17	37.72
701	306	4796	2.57	40.30
708	295	5091	2.48	42.77
715	308	5399	2.59	45.36
722	282	5681	2.37	47.73
728	549	6230	4.61	52.34
734	316	6546	2.66	55.00
739	340	6886	2.86	57.86
745	291	7177	2.44	60.30
750	269	7446	2.26	62.56
755	284	7730	2.39	64.95
760	271	8001	2.28	67.22
766	267	8268	2.24	69.47
771	223	8491	1.87	71.34
776	277	8768	2.33	73.67
782	315	9083	2.65	76.31
788	238	9321	2.00	78.31
794	241	9562	2.02	80.34
800	209	9771	1.76	82.10
806	180	9951	1.51	83.61
813	192	10143	1.61	85.22
819	174	10317	1.46	86.68
826	161	10478	1.35	88.04
832	146	10624	1.23	89.26
839	158	10782	1.33	90.59
846	189	10971	1.59	92.18
853	133	11104	1.12	93.30
860	128	11232	1.08	94.37
867	109	11341	0.92	95.29
875	92	11433	0.77	96.06
882	87	11520	0.73	96.79
891	75	11595	0.63	97.42
900	61	11656	0.51	97.93
911	50	11706	0.42	98.35
929	35	11741	0.29	98.65
955	161	11902	1.35	100.00