KU ACHIEVEMENT & ASSESSMENT INSTITUTE
The University of Kansas

Kansas Assessment Program
Technical Manual
2021

January 2022

KU ACHIEVEMENT & ASSESSMENT INSTITUTE
The University of Kansas

Table of Contents

Table of Tables

Table of Figures

KU ACHIEVEMENT & ASSESSMENT INSTITUTE
The University of Kansas

I. Statewide System of Standards and Assessments

The Kansas Assessment Program (KAP), a program of the Kansas State Board of Education (hereafter "the State Board"), is mandated by the Kansas state legislature. In addition, the English language arts (ELA), mathematics, and science components of KAP also are used to comply with federal elementary- and secondary-education legislation. The three main purposes of KAP, as stated in the *Kansas Assessment Examiner's Manual 2020–2021*, are to

- measure specific claims related to the Kansas standards in grades 3–8 and high school
- report individual student scores along with each student's performance level
- provide subscale and total scores that can be used with local assessment scores to assist in improving a building's or district's programs in ELA, mathematics, and science

The state statutory authority behind KAP is Kan. Stat. Ann. §72-5170 (2020). According to this statute, the State Board is mandated, in part, to

- design and adopt a school performance accreditation system based upon improvement in performance that reflects high academic standards and is measurable
- establish curriculum standards that reflect high academic standards for the core academic areas of mathematics, science, reading, writing, and social studies
- provide for statewide assessments in the core academic areas of mathematics, science, reading, writing and social studies and determine performance levels on the statewide assessments

KAP is the summative assessment for all students in grades 3–8 and high school, except students with significant cognitive disabilities, who are eligible for alternate assessments. The original KAP technical manual (i.e., the *2015 KAP Technical Manual*) for the 2014–2015 school year was developed in 2015 and published in April 2016. The *2015 KAP Technical Manual* was updated in 2016 and published in in January 2017 as the *2016 KAP Technical Manual*. The *2016 KAP Technical Manual* was updated in 2017 and published in November 2017 as the *2017 KAP Technical Manual*. The 2017–2018 and 2018–2019 Technical Manual Addenda were developed in 2018 and 2019, respectively. The *2017 KAP Technical Manual* was updated in 2020 and published in April 2021 as the *2020 KAP Technical Manual*. The current technical manual provides updates, where applicable, for the 2020–2021 school year, including a description of test forms used for 2021 assessment, technical analysis results using 2021 assessment data, and a summary of validity evidence to support the interpretation of test scores for intended test uses.

Important Note on the COVID-19 Pandemic

The 2020–2021 academic school year was significantly affected by the COVID-19 pandemic. After complete school and district closures and halting of assessment administration in spring 2020, the reopening of schools in fall 2020 was characterized by variations of remote, in-person, and hybrid instructional models both within and across states. In many states and districts, the degree to which these instructional models were used changed over the course of the school year and depended on multiple factors, including COVID-19 case counts, district size, ages of students within schools, local policy, student needs, and parent choice. While state and local education agencies made every effort to ensure all students had access to instruction and instructional materials regardless of learning environment, it is well acknowledged that changes to learning inevitably occurred during the 2020–2021 academic year.

On February 22, 2021, in response to the COVID-19 pandemic and the disruption caused by the pandemic on student learning, the United States Department of Education (USDoE) offered states waivers pertaining to Every Student Succeeds Act (ESSA) accountability, school identification, and related reporting requirements for the 2020-2021 school year. On June 29, 2021, the USDoE approved Kansas' request to waive the ESSA accountability, school identification, and related reporting requirements for the 2020-2021 school year.

I.1. State Adoption of Academic Content Standards for All Students

For ELA and mathematics, the State Board adopted the Kansas standards in 2010. The first administration of the operational KAP ELA and mathematics assessments that aligned with the 2010 Kansas standards occurred in 2015. More information about the 2010 Kansas standards and KAP assessments can be found in the *2015 KAP Technical Manual* and the *2016 KAP Technical Manual*. In 2017, the State Board adopted the updated version of the 2010 Kansas standards for ELA and mathematics. More information about the 2017 Kansas standards and KAP assessments can be found in the *2020 KAP Technical Manual*.

For science, the State Board adopted the Kansas standards in 2013. The first administration of the operational KAP science assessments that aligned with the 2013 Kansas standards occurred in 2017. In 2018, the Kansas science standards review committee evaluated the 2013 Kansas science standards and concluded that no updates to the 2013 Kansas science standards were needed. More information about the 2013 Kansas standards and KAP assessments for science can be found in the *2017 KAP Technical Manual*.

I.2. Coherent and Rigorous Academic Content Standards

Committees of Kansas educators and stakeholders provided input on the standards in Kansas. These standards supported the vision of the State Board: to lead the world in the success of each student (refer to http://www.ksde.org/Board). These standards help schools equip students with the academic, cognitive, metacognitive, technical, and employability skills required for postsecondary success, as well as the capacity to positively affect the world around them. The

Kansas standards are Kansas's coherent and rigorous academic content standards, which adhere to the State Board's mission: to prepare Kansas students for lifelong success through rigorous, quality academic instruction, career training, and character development according to each student's gifts and talents. The detailed process and timeline of development of the 2010 Kansas ELA and mathematics standards can be found in the *2015 KAP Technical Manual* and the *2016 KAP Technical Manual*. The detailed process and timeline of review for the 2017 ELA and mathematics Kansas standards can be found in the *2020 KAP Technical Manual*. The detailed process and timeline of the development of the 2013 Kansas science standards can be found in the *2017 KAP Technical Manual*.

I.3. Required Assessments and Intended Population

The KAP assessment tests students in the subject areas of ELA, mathematics, and science. The subject areas and grades tested are ELA in grades 3–8 and 10; mathematics in grades 3–8 and 10; and science in grades 5, 8, and 11.

Kansas is committed to including all students in the KAP assessment. Students enrolled in Kansas public schools must take one of three tests: the KAP assessment, the English language proficiency assessment, or the alternate assessment. In the first year when English learners arrived in the United States, they are required to take the KAP mathematics and science tests. They are not required to take the ELA tests but must take the Kansas English Language Proficiency Assessment. In the second year after English learner arrived in the United States, they are required to take all three KAP assessments.

Qualifying students with significant cognitive disabilities, typically no more than 1% of Kansas students, take the Dynamic Learning Maps® Alternate Assessment for ELA, mathematics, and science. Other special-needs students with Individualized Education Programs, 504 plans, or Student Intervention Team plans take the KAP assessment but can use accommodations consistent with their personal needs profiles. If an unapproved accommodation is used (e.g., reading aloud to a student on the KAP ELA test), the student test record is considered invalid. A detailed accommodation summary can be found in Chapter V Inclusion of All Students.

Exemptions from KAP assessments are granted to students who, during the testing window,

- move into a different school
- experience catastrophic illnesses or accidents
- are serving long-term suspension
- are truant for more than two consecutive weeks
- are incarcerated in an adult facility
- are in a special detention center

A special exemption was added in 2021. With approval of Kansas State Department of Education (KSDE), this special exemption was granted to students who were learning remotely because of COVID-19 and were unable to attend a brick-and-mortar school to test.

II. Assessment System Operations

The development of any test requires making many critical decisions regarding, for example, the content and cognitive complexity, the appropriate scope of that content for particular subject areas, and the number of items associated with each test. These decisions are not made in isolation for different grades but must take into consideration the importance of coherence across all grade levels of the assessment. Together, these decisions guide the test-construction process and products.

II.1. Assessment Framework of the Assessed Grades

The assessment framework groups Kansas standards based on similar content into different categories that follow a hierarchical structure. The hierarchical structure helps the organization of standards. For example, the 2017 Kansas standards for English language arts (ELA) and mathematics are grouped by domain and cluster; clusters are the sublevels of domains. ELA has three domains and each domain has two to three clusters. Mathematics has two domains and each domain has two to eleven clusters. KAP items are assigned to only one domain and one cluster based on their primary alignment. The 2017 Kansas standards assessment framework for ELA and mathematics can be found in the *2020 KAP Technical Manual*. The 2010 Kansas standards assessment framework for ELA and mathematics and the framework for of science standards can be found in the *2017 KAP Technical Manual*.

II.2. Test Design and Development

KAP assessments are all computer based. Accessible Teaching, Learning, and Assessment Systems (ATLAS) worked with the Kansas State Department of Education (KSDE) to determine the content to be assessed by the KAP assessments for each subject area and grade level. In fall 2020, ATLAS, in collaboration with Technical Advisory Committee (TAC) and KSDE, decided to administer 2019 operational forms in 2021. The purpose for using the 2019 forms in 2021 was to evaluate the effect of COVID-19 on students' performance while controlling for known properties of the test. The 2019 ELA operational forms aligned with the 2010 Kansas standards for ELA, and the 2019 science operational forms aligned with the 2013 Kansas standards for science. The 2019 mathematics operational forms were considered as the transitional-year forms. In 2019, items that do not align to 2017 Kansas standards were removed from the mathematics operational forms. The *2020 KAP Technical Manual* provides a detailed test-development timeline for the three subjects for 2019 operational forms.

II.2.1. Test Blueprints

Test blueprints guided the construction of the test forms and provide the range of the proportion of items required for each content category across grades for different subjects. The detailed test blueprint for the three subjects can be found in the *2017 KAP Technical Manual*.

II.2.2. Test Design

The mathematics and science tests utilize a fixed-form design, while the ELA test utilize a two-stage adaptive design. Mathematics and science each have one operational form administered in two sessions: Session 1 and Session 2. For the adaptive test in ELA, students all get the same test

with medium difficulty (i.e. average item difficulty) in the first session (Session 1), including items in a wider range of difficulty levels. Then, assignment of the test of the second test session (the easy or hard test) is determined by the ability estimates based on students' answers to items in the Session 1. For three subjects, each session and difficulty level offers several identical blocks of items that are presented in different orders to deter cheating. All blocks are pre-assembled. Students are randomly assigned to one block in each session and difficulty level; each session and difficulty level has a designated block of items for students who need accommodations. Table II-1 shows the test design of the KAP assessment for each session by subject. The number of operational items per session and per test is also included in the table

Table II-1. Test Design of the KAP Assessment by Subject and Session

Subject	Grade	Session 1	Session 2	Total
English Language Arts	3-8, HS	22 (M)	25 (H) 25 (E)	47 (M&H) 47 (M&E)
Mathematics	3-8, HS	25	30	55
Science	5	18	15	33
	8	20	21	41
	HS	19	19	38

Notes. HS = high school; E = average item difficulty is easy; M = average item difficulty is medium; H = average item difficulty is hard.

II.2.3. Operational Test Construction

The test forms for 2021 administration (i.e. 2019 KAP forms) were constructed in 2019, using the same procedures and guidelines as in previous years. A detailed description of test-construction procedures and guidelines can be found in the *2017 KAP Technical Manual* for ELA and science. The *2020 KAP Technical Manual* describes the modified procedures and guidelines for mathematics fixed test designs.

II.3. Item Development

Anticipating COVID-19's effect on student performance and the impact to item statistics, KSDE with input from ATLAS and advice and suggestions from the TAC decided to not include any field-tested items on 2021 forms. Thus, during 2020–2021, ATLAS did not conduct any item development or item review to put new field-tested items on the 2021 summative forms.

II.4. Test Administration

To keep the test-administration conditions stable across years, especially through the COVID-19 pandemic, KSDE decided to allow only in-person testing in 2021; no remote testing was available. Students learning remotely were asked to return to their school for testing. Chapter V Inclusion of All Students describes the effort KSDE made to encourage in-person testing. In some cases, students who were learning remotely because of COVID-19 and were unable to attend in-person testing were assigned a special circumstance (SC) code to be exempt from testing. Adding this SC code to Kite® Educator Portal for individual students must be approved by KSDE. To encourage participation, the 2021 KAP testing window was postponed by 2 weeks,

opening on April 1 and closing on May 18. However, the length of testing window did not change across years.

The test administration and security training were conducted through online conference or with online training materials. Kansas uses a train-the-trainer model. District test coordinators received training first from KSDE and then trained building-level personnel before the local test. The detailed training procedure can be founded in the *2017 KAP Technical Manual*.

The test-administration conditions were the same as in previous years. Detailed information about test-administration procedures and guidance can be found in the *2017 KAP Technical Manual*. During the testing window, ATLAS sent the 2021 KAP Teacher survey to all test coordinators, administrators, curriculum coordinators, teachers, and other educators who administered the KAP to gather information about testing settings and to investigate whether some districts changed testing settings because of COVID-19. Only 1% of recipients responded to the survey. To maintain social distancing during summative testing, districts used a variety of settings. Among survey responses, 44 of 206 (21%) educators said that at least some of their students tested in different settings. Those settings included gymnasiums, common areas, cafeterias, multipurpose rooms, conference rooms, board rooms, nurse's offices, school district offices, libraries, and district buildings.

II.5. Monitoring Test Administration

Test-administration monitoring includes both testing-data monitoring and on-site visits. The testing-data monitoring was the same in 2021 as in previous years and included looking at test usage and testing irregularities like enrollment errors or missing data. However, because of the impact of COVID-19, the on-site visit from KSDE and members of the Kansas Assessment Advisory Council was halted in 2021. Detailed information about standard procedures and protocols for test-administration monitoring can be found in the *2017 KAP Technical Manual*.

ATLAS also monitored weekly testing rates, that is, the proportion of completed tests to enrolled tests. Testing rates were organized by grade and subject for the state, for districts, as well as for different student groups[1] (gender, race, ethnicity, English learner status, and disability status). Then weekly testing rates were compared with the data from previous weeks. Likewise, the number of students with COVID-19 SC codes was monitored and organized by grade and subject.

II.6. Test Security

Because the test-administration conditions were the same this year, the same procedures and protocols related to test security were applied. Detailed information about these procedures and protocols can be found in the *2017 KAP Technical Manual*.

[1] Economic disadvantaged status is not shared with ATLAS to protect the privacy of students, so this student group is not included in the comparison.

III. Technical Quality: Validity

As defined in the *Standards for Educational and Psychological Testing* (the *Standards* hereafter), *validity* refers to "the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests" (American Psychological Association [APA] et al., 2014, p. 11).

The *Standards* (APA et al., 2014) provide a framework for describing the sources of evidence that should be considered when evaluating test-score validity. These sources include evidence based on test content, response processes, internal test structure, relationships between test scores and other variables, and the consequences of testing. The validation process involves the ongoing collection of a variety of evidence to support the proposed test-score interpretations and uses. This technical manual mainly describes aspects of the KAP assessments that support KAP test-score interpretations and uses.

III.1. Validity Evidence Based on Test Content

Content evidence for the KAP assessments comes from the alignment between KAP items and Kansas standards, from the congruence between the test and the test blueprint, and from the congruence between the test blueprint and the standards (i.e., balance of representation of standards). Chapters I and II of this current technical manual, the *2017 KAP Technical Manual*, and the *2019 KAP Technical Manual Addendum* present validity evidence related to the high-quality item-development process, the alignment between item and standards, the expected item cognitive complexity levels, the correspondence between the test and test blueprint, and the blueprint quality and alignment. An independent alignment study for KAP was conducted by edCount, and the alignment results on items, tests, blueprints, and Kansas Standards can be found in *Kansas Assessment Program Alignment Evaluation Report 2015–2016* (Forte et al., 2016).

III.2. Validity Evidence Based on Response Process

Response-process evidence examines the extent to which the cognitive skills and processes students use to answer an item match those targeted by item writers. The *2017 KAP Technical Manual* describes validity evidence related to how cognitive skills were considered by item writers and item reviewers for each item during item development. Also, during the development of performance-level descriptors, the expectations of students' cognitive process were stated differently in different levels of performance-level descriptors.

III.3. Validity Evidence Based on Internal Structure

As described in the *Standards* (APA et al., 2014), internal-structure evidence refers to "the degree to which the relationships among test items and test components conform to the construct on which the proposed test score interpretations are based" (p. 13). Two sets of validity evidence about the internal structure are provided that include (a) evidence that the KAP assessment is essentially unidimensional; (b) evidence that the item response theory (IRT) model used for each subject showed good fit results; and (c) evidence that the test contains no or few items flagged for significant and large differential item functioning (DIF), which helps support comparable

measurement across groups. For each subject and grade, the KAP assessment is fitted by an IRT unidimensional model. The evidence of all items measuring one primary construct, that is, unidimensionality, is one type of internal-structure validity evidence. Moreover, KAP dichotomous items are fitted by the two-parameter logistic (2PL) model, and the polytomous items are fitted by the graded-response model. The IRT model assumption evaluation, including model-fit, also can provide internal-structure evidence. The evaluation of unidimensionality, as well as IRT assumptions of the KAP English language arts (ELA) and mathematics tests, are described in the *2015 KAP Technical Manual*. The evaluation of unidimensionality and IRT assumptions of science KAP tests are described in the *2017 KAP Technical Manual*. Finally, it is expected internal structure of a test should be consistent for different student groups. DIF analysis can help identify items that are performing differently for student groups. Thus, DIF analysis results are included as further internal-structure validity evidence.

III.3.1. Differential Item Functioning

DIF analysis evaluates items for potential bias and examine whether an item shows statistical difference between two groups of students controlling for student ability. Logistic regression was used to detect items with uniform DIF, i.e. an item is consistently more difficult for one group of students than the other group for all ability levels. The Jodoin and Gierl (2001) DIF classification criteria were used to indicate the degree of DIF (i.e., negligible, moderate, large). When the DIF test is significant, large DIF is identified by a Nagelkerke R^2 change greater than or equal to .070; moderate DIF has a Nagelkerke R^2 change between .035 and .070; and negligible DIF has a Nagelkerke R^2 change less than .035.

For each subject and grade, DIF was examined across gender (i.e., female vs. male), race (i.e. Black vs. White), and ethnicity (i.e., Hispanic vs. non-Hispanic). For all subjects and grades, no items were flagged for moderate or large race-related DIF or ethnicity-related DIF. Only two grade-5 ELA items were flagged for moderate DIF favoring male students. No other items were identified with either moderate or large gender-rated DIF. All results suggest that the item-development process and procedures effectively addressed potential bias-and-sensitivity issues during the development phase. When an item is flagged, test-development teams review the item for potential sources of bias against subgroups of the population. The test-development teams reviewed these two grade 5 ELA items flagged for moderate gender DIF and were unable to identify any content bias that may contribute to gender DIF.

III.4. Validity Evidence Based on Relations to Other Variables

As described in the *Standards*, "evidence based on relationships with other variables provides evidence about the degree to which these relationships are consistent with the construct underlying the proposed test score interpretations" (APA et al., 2014, p. 16).

This kind of evidence refers to external evidence and is classified into two types: convergent and discriminant. *Convergent evidence* is provided by the relationships between students' performance on different assessments intended to measure similar constructs. *Discriminant evidence* is provided by the relationships between students' performance on different tests intended to measure different constructs. The comparison of the KAP and National Assessment

of Educational Progress (NAEP) performance is considered convergent evidence. The correlations among different KAP subject scores is considered discriminant evidence.

III.4.1 Relationships Among KAP Subjects

Discriminant validity can be evaluated using the correlation between subjects, such as ELA and mathematics. Past studies showed high correlations between subjects, which indicates that some common traits are shared across subjects; however, the correlations should not be too high. The correlations and disattenuated correlations after correcting measurement errors are presented in Table III-1 are between subjects of the same grade, and the values range from .69 to .77 for correlations and .75 to .86 for disattenuated correlations. The lowest correlations among subjects are the grade-5 mathematics and science correlation and the grade-10 ELA and mathematics correlation. The highest correlations is between grade-5 ELA and mathematics. After correlation measurement errors, the lowest disattenuated correction is between grade-10 ELA and mathematics and the highest disattenuated correlation is between grade-5 ELA and science. According to Cohen (1988), a correlation larger than .50 is considered a large correlation. All correlations among KAP subjects are large correlations, indicating that some common traits are shared across KAP subjects.

Table III-1. Correlations (C) and Disattenuated Correlations (DC) Among English Language Arts (ELA), Mathematics, and Science Scores

	ELA vs. mathematics		ELA vs. science		Mathematics vs. science	
Grade	C	DC	C	DC	C	DC
3	.77	.83	-	-	-	-
4	.75	.81	-	-	-	-
5	.74	.80	.74	.86	.69	.79
6	.75	.81	-	-	-	-
7	.73	.81	-	-	-	-
8	.72	.79	.74	.85	.70	.80
10	.69	.75	-	-	-	-

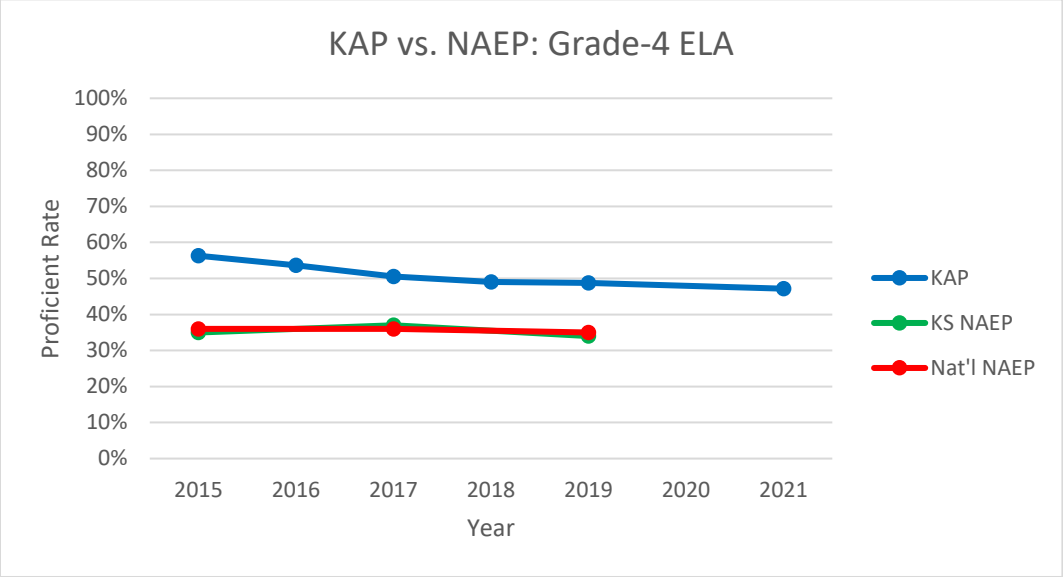
III.4.2. Relationships Between KAP Assessment and National Assessment of Educational Progress

Convergence validity requires that another test measure a similar construct. The state of Kansas participates in the NAEP, also known as the Nation's Report Card. NEAP is the largest nationally representative assessment of what American students know and can do, and it serves a different role than state assessments. NAEP assessments allow each state to be compared to national results and to evaluate progress over time. It informs the public about the academic achievement of elementary (grade 4) and secondary (grade 8) students in Kansas and in the United States on ELA and mathematics. Thus, the relationship between KAP and NAEP performance can be used as one source of convergent evidence. However, individual NAEP scores are not available. Thus, only the trend of proficiency rates across years are compared between the two assessments. The KAP and NAEP are using different achievement standards to

judge whether a student meets proficiency or not. Only comparing the proficiency rates within a year does not have meaning, but comparing the trend of proficiency rates across years between two assessments can indicate the relationship between these two assessments on measuring a similar construct. The Kansas State Department of Education (KSDE) provides more information about NAEP.

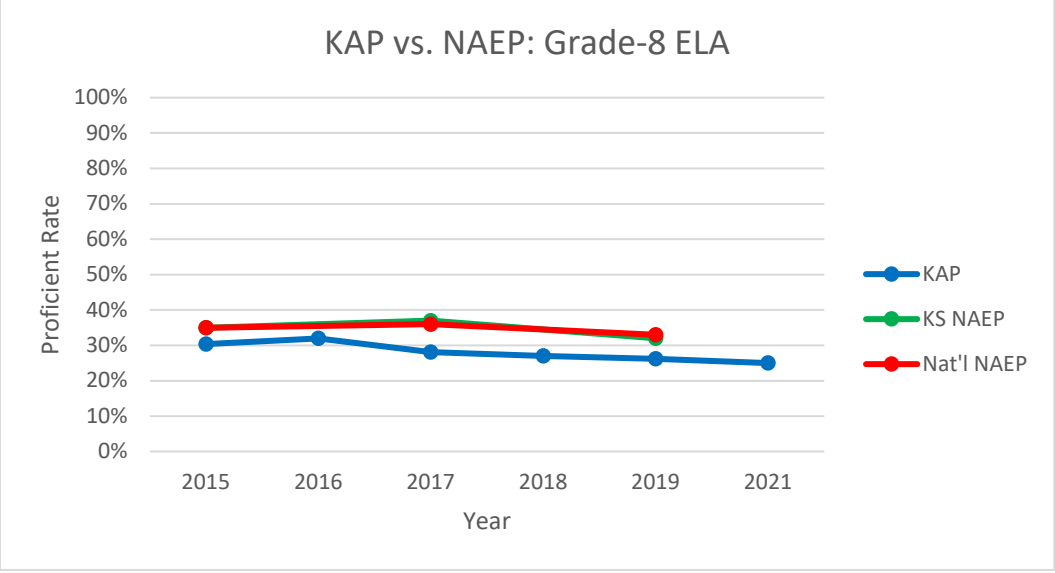
KAP provides student performance in four performance levels (Levels 1, 2, 3, and 4). The proficiency rate of KAP is the percentage of students in Levels 3 and 4. NAEP separates students into three performance levels (Basic, Proficient, and Advanced). The proficiency rate of NAEP is the percentage of students in Proficient and Advanced levels. Comparisons between KAP and NAEP proficiency rates across years for grades 4 and 8 ELA and mathematics are presented in Figures III-1 through III-4. In years 2015 through 2019, KAP proficiency rates ranged from 49% to 56% for grade-4 ELA, from 26% to 32% for grade-8 ELA, from 36% to 40% for grade-4 mathematics, and from 24% to 27% for grade-8 mathematics. The Kansas and national NAEP proficiency rates for both grade-4 and grade-8 ELA and mathematics are very similar across years, ranging from 30% to 40%, with most around 35%. Both KAP and Kansas NAEP have a slight decrease in proficiency rate from 2017 to 2019 for grade 4 and 8 ELA and grade 4 mathematics. The similar trend of proficiency rates between KAP and Kansas NAEP provide one source of convergent validity evidence.

Figure III-1. Grade-4 English Language Arts (ELA) Proficiency-Rate Trend Across Years: KAP vs. NAEP



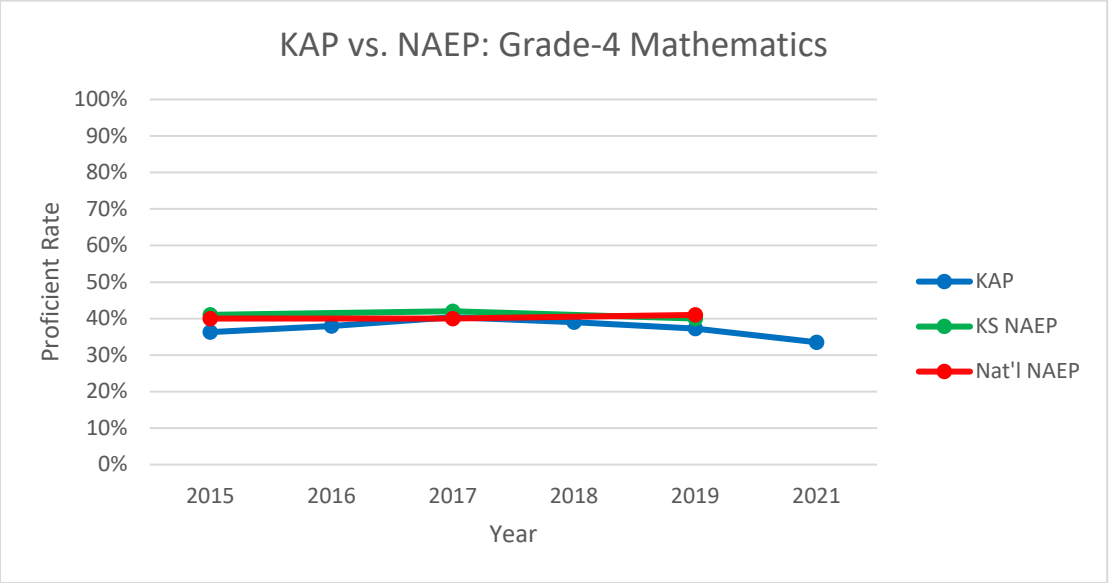
Note. KAP = Kansas Assessment Program; NAEP = National Assessment of Educational Progress.

Figure III-2. Grade-8 English Language Arts (ELA) Proficiency-Rate Trend Across Years: KAP vs. NAEP



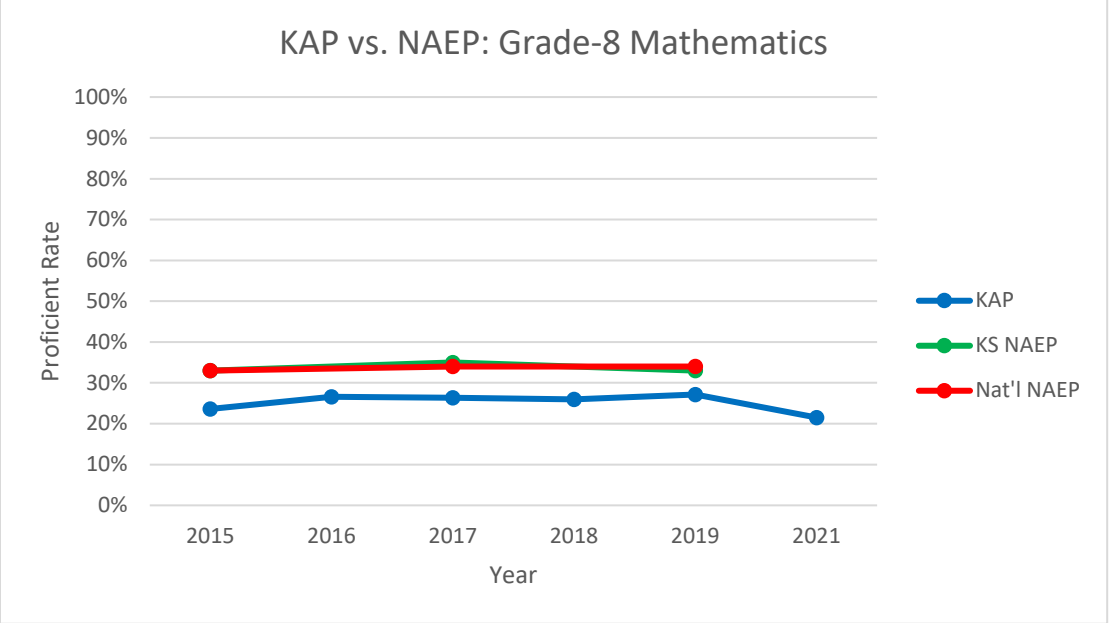
Note. KAP = Kansas Assessment Program; NAEP = National Assessment of Educational Progress.

Figure III-3. Grade-4 Mathematics Proficiency-Rate Trend Across Years: KAP vs. NAEP



Note. KAP = Kansas Assessment Program; NAEP = National Assessment of Educational Progress.

Figure III-4. Grade-8 Mathematics Proficiency-Rate Trend Across Years: KAP vs. NAEP



Note. KAP = Kansas Assessment Program; NAEP = National Assessment of Educational Progress.

III.5. Validity Evidence Based on Consequences of Testing

Validity evidence based on consequences refers to evidence supporting the intended uses and interpretation of test scores. The primary intended use of KAP test scores is to provide scores that can be used with local assessment scores to assist in improving a building's or district's

programs as stated in the Examiner's Manual. Section IV.4. Scoring and Scaling summarizes how items and tests are scored. For a given test score, the performance level is determined by a set of established cut scores. Chapter VI Academic Achievement Standards and Reporting summarizes the cut scores and includes an example of a KAP student score report. To help educators and parents interpret KAP results, Accessible Teaching, Learning, and Assessment Systems (ATLAS) also provides the KAP Educator Guide and the KAP Parent Guide.

In previous years, one of the primary ways to use KAP scores and performance levels in improving building or district programs was in accountability systems for districts and schools; these systems are required by the U.S. Department of Education (ED). In 2021, ED's guidance on assessment offered states the option to apply for a one-year waiver of accountability requirements. ED also recommended that the focus of the 2021 assessment-score use be to provide information for states and allow them to target resources and supports provided to local districts and schools.

Before 2021, the KAP Educator Guide and the KAP Parent Guide described "meet federal and state accountability" as one of the test purposes for KAP. According to ED's guideline, this test purpose was removed from the 2021 KAP Educator Guide and the 2021 KAP Parent Guide. Therefore, as described in both of those guides, the use and interpretation of KAP scores is limited to providing information about student learning.

Moreover, language was added to both the student score report and the KAP Educator Guide to remind students, parents, and educators that learning conditions and student performance may have been affected by COVID-19. This caveat states

> Please note a single test score does not provide a complete or precise measure of student achievement. When interpreting KAP results, please take into consideration other measures of student achievement. Also, consider how the conditions for learning, which may have been disrupted by the pandemic, may influence performance. (KAP Educator Guide, KSDE, 2021)

Although parents and educators can still use test scores to help identify students' relative strengths and limitations, determine students' progress toward meeting state curriculum standards, and compare students' performance to that of other students in the school, district, and state, as stated in the KAP Educator Guide (KSDE, 2021), parents and educators need to consider the impact of the COVID-19 pandemic on learning.

Because of COVID-19, there were limited opportunities to collect data to evaluate validity evidence based on consequences of testing for KAP. In 2022, we plan to collect data through a KAP teacher survey, including questions about how educators use the data and how those uses affect students' learning outcomes.

IV. Technical Quality: Other

Evidence related to technical quality of the Kansas Assessment Program (KAP) is provided in this chapter. The majority of the analysis conducted in this chapter is based on this year's assessment data. Technical quality evidence includes test reliability, fairness and accessibility, item analysis summary, test analysis summary, and trend data. The usage of multiple forms in English language arts (ELA) and technical analysis for ongoing maintenance such as plans for grade 10 mathematics standard setting in 2022 is also summarized.

IV.1. Reliability

Reliability is a test-score consistency index that represents the degree of consistency of test scores across repeated measures. The more stable the test scores across repeated measures, the more reliable the tests are. The factors leading to unstable test scores are called measurement error. The measurement error includes but is not limited to changing of testing conditions; changing of student's knowledge, physical condition, and mental status; and changing of testing content across multiple test administrations. Measurement error cannot be fully removed but can be reduced. For example, standardized testing procedures can reduce the measurement error caused by changing testing conditions. KAP has standardized its testing procedures, and the same procedures are applied to all students; specific accommodations are provided to students with special needs. The testing procedure specifications can be found in the *Kansas Assessment Examiner's Manual 2020–2021*.

In the context of an educational achievement test, factors such as learning, fatigue, and motivation may affect test takers at different rates for repeated measures. It is impractical to test the same content area repeatedly because test takers cannot maintain the same knowledge, physical condition, and mental status across test administrations. Therefore, reliability for educational measures is typically estimated rather than calculated directly. Estimated reliability coefficients usually range from 0 to 1. Higher values indicate more reliable tests with less measurement error.

In this section, we include the KAP reliabilities of both overall scores and subscores. The overall score reliabilities were calculated for the whole tested sample and different student groups. Moreover, ELA path reliabilities were calculated because ELA has adaptive test design. The item response theory (IRT) information functions and the conditional standard error of measurement of the tests are also included. After that, classification consistency and accuracy of the overall score performance classification is presented. For KAP subscores, reliability, classification consistency, and classification accuracy are summarized.

IV.1.1. Test Reliability

Marginal reliability (Green et al., 1984) was used to estimate test reliability. Marginal reliability could be used to estimate reliability for both fixed-form and adaptive tests. The detailed method for marginal-reliability calculation can be found in the *2017 KAP Technical Manual*. As shown in Table IV-1, marginal reliabilities of ELA and mathematics are above .90; science has relatively lower reliabilities because there are fewer test items compared to ELA and mathematics, but values are still greater than or equal to .80.

Table IV-1. Test Reliability by Subject and Grade

Grade	English language arts	Mathematics	Science
3	.93	.93	
4	.91	.94	
5	.91	.93	.82
6	.91	.93	
7	.90	.91	
8	.90	.91	.83
High school	.92	.92	.84

IV.1.1.2. Student Group Reliability

Marginal reliabilities were calculated for gender groups, race groups, ethnicity groups, English learner (EL) status groups, and disability status groups[2]. Student group reliabilities are presented in Table IV-2 through Table IV-4 for ELA, mathematics, and science. For ELA and mathematics, the subgroup marginal reliabilities for each group were close to or above .90 across grades, ranging from .88 to .93 for ELA and from .89 to .94 for mathematics. Science had relatively lower subgroup reliabilities because it had fewer test items compared to ELA and mathematics. Science subgroup marginal reliabilities ranged from .80 to .88 across grades. For all three subjects, the differences in reliabilities among different student groups were small.

[2] Economic disadvantaged status is not shared with ATLAS to protect the privacy of students, so this student group is not included in the comparison.

Table IV-2. Student Group Reliability for English Language Arts

Subgroup	Grade						
	3	4	5	6	7	8	10
Gender							
Male	.93	.91	.91	.91	.91	.90	.92
Female	.93	.91	.91	.91	.90	.90	.91
Race							
AI	.93	.92	.92	.92	.92	.91	.93
Asian	.93	.89	.90	.90	.89	.88	.90
Black	.93	.92	.92	.92	.92	.91	.93
NHPI	.93	.91	.92	.92	.91	.91	.93
Other	.93	.91	.91	.92	.91	.90	.92
White	.93	.91	.91	.91	.90	.90	.92
Hispanic							
Yes	.93	.92	.92	.92	.92	.91	.93
No	.93	.90	.91	.91	.90	.90	.91
SWD							
Yes	.93	.91	.92	.92	.92	.91	.93
No	.93	.91	.91	.91	.90	.90	.91
EL							
Yes	.93	.92	.92	.92	.93	.91	.93
No	.93	.91	.91	.91	.90	.90	.91

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-3. Student Group Reliability for Mathematics

Subgroup	Grade						
	3	4	5	6	7	8	10
Gender							
Male	.93	.94	.93	.92	.91	.91	.91
Female	.93	.94	.94	.93	.92	.92	.92
Race							
AI	.94	.94	.94	.94	.92	.91	.92
Asian	.92	.93	.91	.89	.89	.90	.89
Black	.94	.94	.94	.93	.91	.91	.92
NHPI	.94	.94	.94	.92	.91	.91	.92
Other	.94	.94	.93	.93	.92	.91	.92
White	.93	.94	.93	.93	.91	.91	.92
Hispanic							
Yes	.94	.94	.94	.93	.92	.92	.92
No	.93	.94	.93	.93	.91	.91	.92
SWD							
Yes	.94	.94	.93	.93	.91	.91	.91
No	.93	.94	.93	.93	.91	.91	.92
EL							
Yes	.94	.94	.94	.93	.91	.91	.92
No	.93	.94	.93	.93	.91	.91	.92

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-4. Student Group Reliability for Science

Subgroup	Grade		
	5	8	11
Gender			
Male	.81	.83	.83
Female	.83	.84	.85
Race			
American Indian	.85	.84	.87
Asian	.80	.82	.81
Black	.86	.83	.88
NHPI	.86	.84	.88
Other	.83	.84	.85
White	.82	.83	.84
Hispanic			
Yes	.85	.84	.87
No	.81	.83	.84
Student with Disability			
Yes	.85	.83	.87
No	.82	.83	.84
English Learner			
Yes	.86	.83	.88
No	.81	.83	.84

Note. NHPI = Native Hawaiian and Pacific Islander.

IV.1.1.2. Path Reliability

Path reliability is the product of using a multistage adaptive test design; therefore, it applies only to ELA tests. The multistage adaptive test design dictates that different sets of items are assigned to students at the second stage (i.e., second test section). The different paths mean that students take item sets with different levels of difficulty. Marginal reliabilities were used to calculate the path reliability. Because an ELA test has only one set of items with a medium level of difficulty in the first test section (Section 1) and has two sets of items—easy and hard—in the second test section (Section 2), two path reliabilities (easy and hard) were calculated for each test. Table IV-5 presents the path reliabilities and the number and percentage of students who took each path. As shown in Table IV-5, most grades had a similar number of students who took different paths, except grade 4 and grade 6. For all grades, the path reliability of the easy path was higher than the path reliability of the hard path. The difference was small in grades 3–6 but slightly larger in grades 7–10. All path reliabilities were close to or above .90, ranging from .88 to .94, indicating all paths can provide reliable scores.

Table IV-5. Path Reliability for English Language Arts

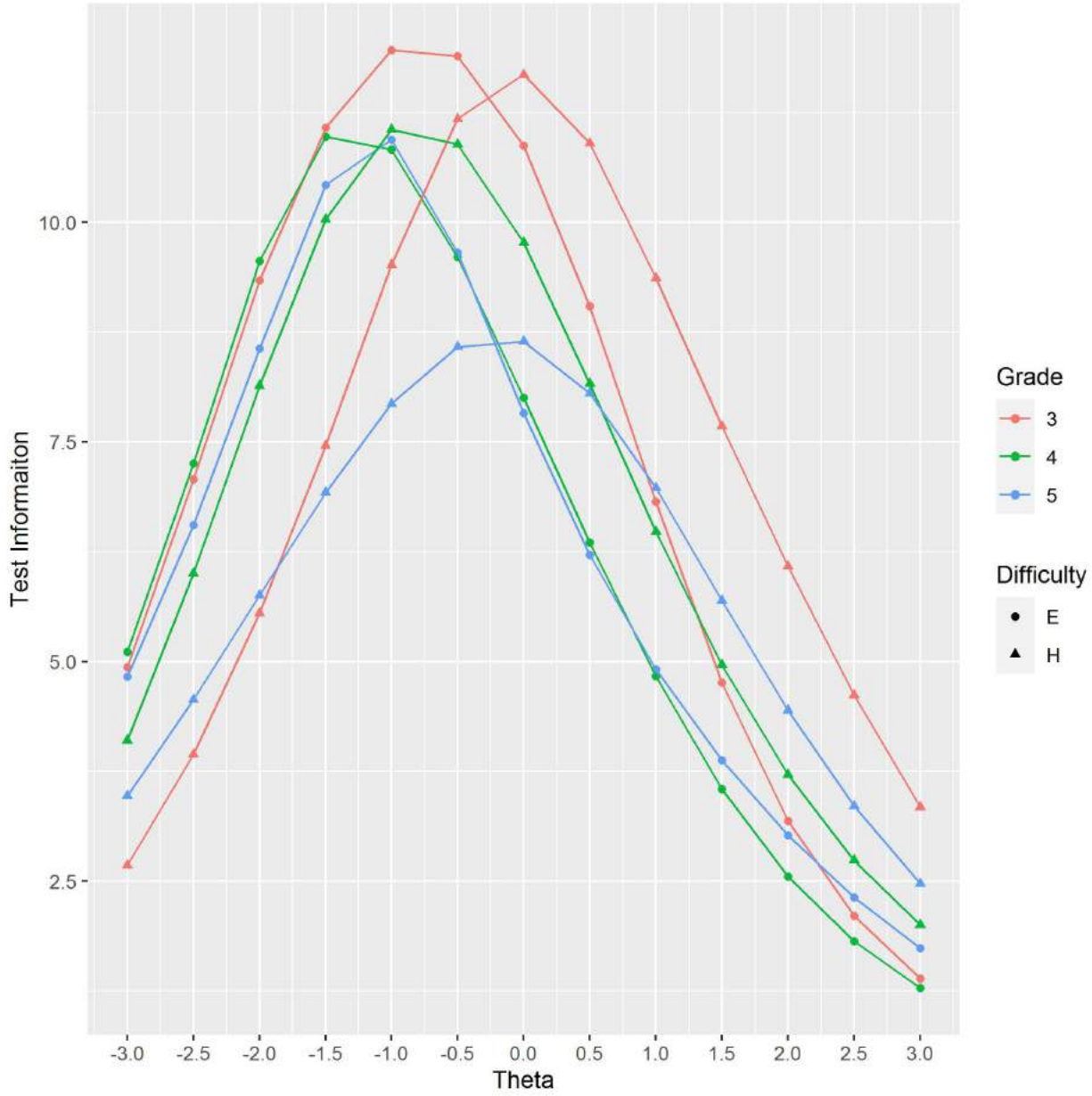
Grade	Path	Section 1	Section 2	*N*	%	Reliability
3	1	Medium	Easy	18,620	55	.93
	2		Hard	15,411	45	.92
4	1	Medium	Easy	7,672	22	.92
	2		Hard	26,437	78	.90
5	1	Medium	Easy	17,680	50	.92
	2		Hard	17,466	50	.90
6	1	Medium	Easy	7,176	20	.92
	2		Hard	27,882	80	.91
7	1	Medium	Easy	19,055	54	.93
	2		Hard	16,225	46	.88
8	1	Medium	Easy	18,159	51	.92
	2		Hard	17,195	49	.88
10	1	Medium	Easy	15,517	48	.94
	2		Hard	17,038	52	.90

IV.1.2. Test Information

KAP tests use IRT models to estimate students' latent ability (theta), which is then transformed to a scaled score. Using IRT models, test information functions (TIF) can be estimated for each theta value across the whole performance continuum. A TIF is computed as the sum of item information function of all operational items in a grade for each test. The TIF is used to estimate the amount of information the test provides at each theta; it is conceptually parallel to the reliability coefficient in classical test theory. Figure IV-1 through Figure IV-4 present the TIFs for theta values ranging from -3 to 3 in increments of 0.5 for each grade in ELA, mathematics, and science. For the mathematics and science tests where there is one fixed form per grade, there is only one TIF per grade. For the ELA tests, which represent two-stage, adaptive delivery, there are two TIFs per grade because there are two sets for total test forms, as described in Section II.2 Test Design and Development. There is one form where Session 1 is medium and Session 2 is easy and a second form where Session 1 is medium and Session 2 is hard for each grade.

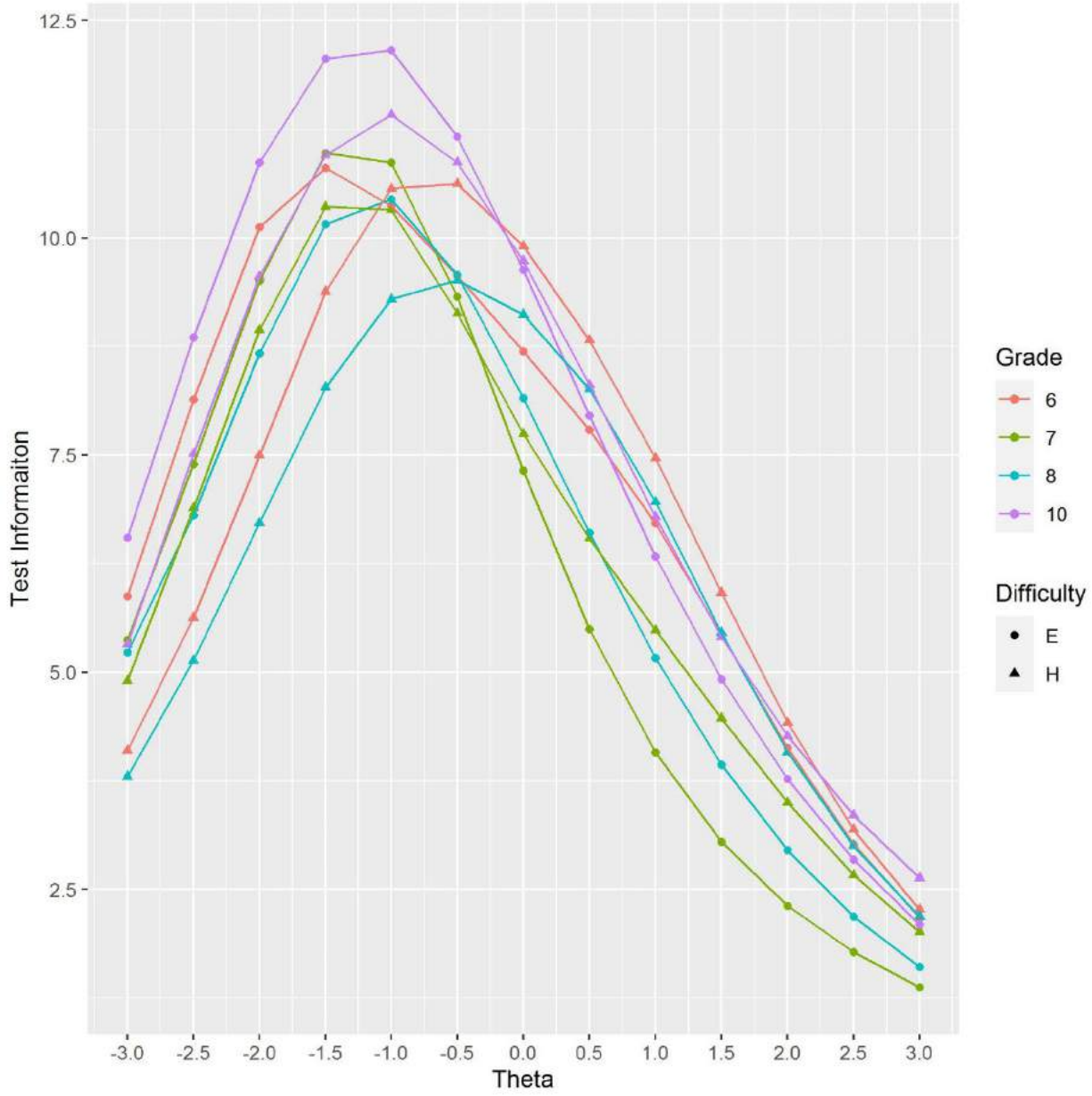
Typically, TIF values are high at the center of the theta distribution and gradually decrease toward the two ends of the theta scale, where thetas are very low or very high and result in a bell-shaped pattern. For ELA, grades 3–6 and 8 had TIFs reaching the maximum value at different theta values for easy and hard test forms. The easy form had a smaller theta value with a maximum TIF than the hard form. For ELA grades 7 and 10, the TIF of the easy form was larger than the TIF of the hard form at the low theta range and smaller at the high theta range. Mathematics had TIFs reaching the maximum value at theta values around 0, which is close to a level 3 cut. The TIF of the grade 11 science test was larger than the TIFs of grades 5 and 8, which is consistent with the reliability results, where reliability of grade 11 science was higher than the reliabilities of grades 5 and 8.

Figure IV-1. Test Information Function for English Language Arts Grades 3, 4, and 5



Note. E = medium and easy sections; H = medium and hard sections.

Figure IV-2. Test Information Function for English Language Arts Grades 6, 7, 8, and 10



Note. E = easy; H = hard.

Figure IV-3. Test Information Function for Mathematics

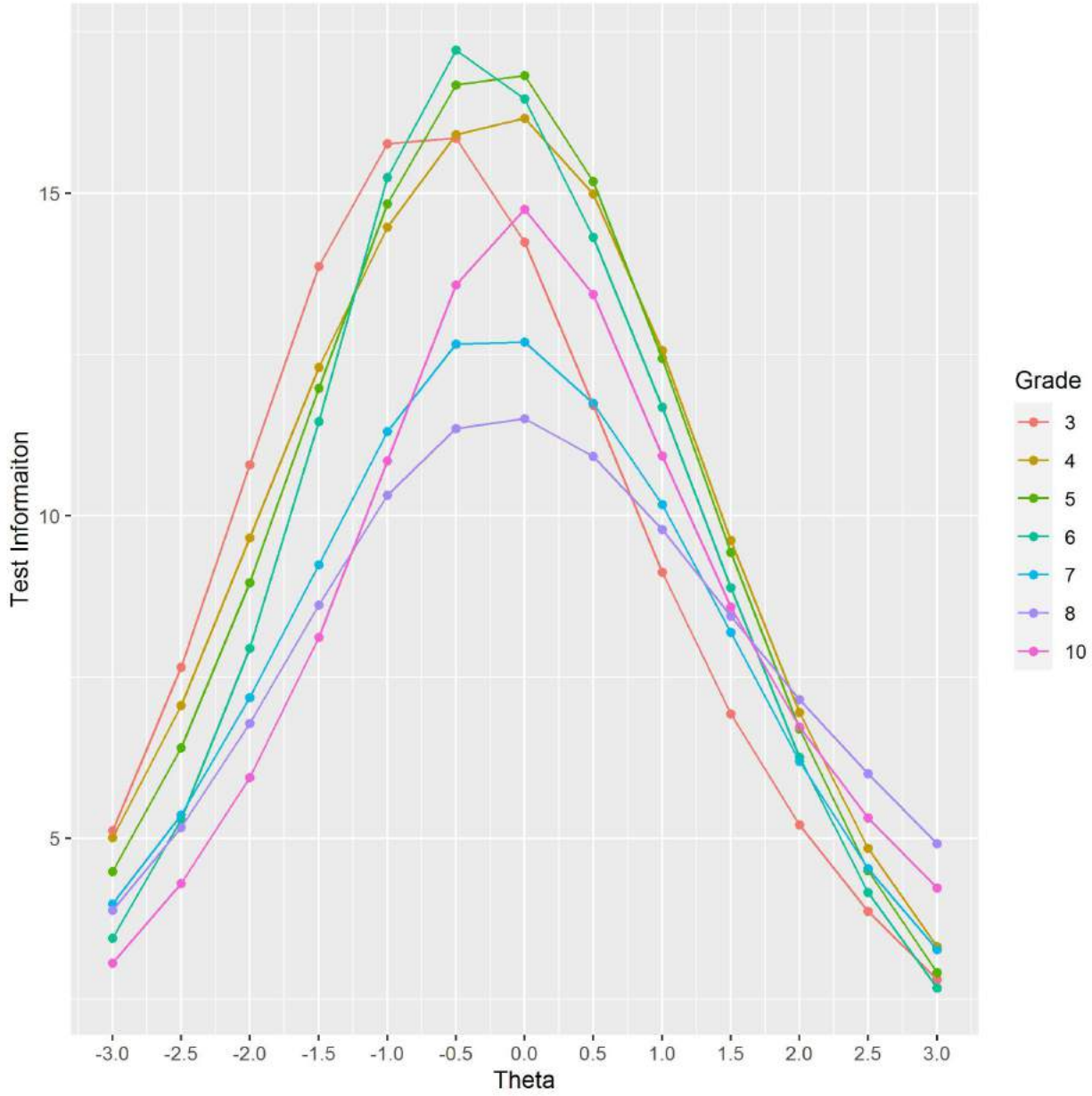
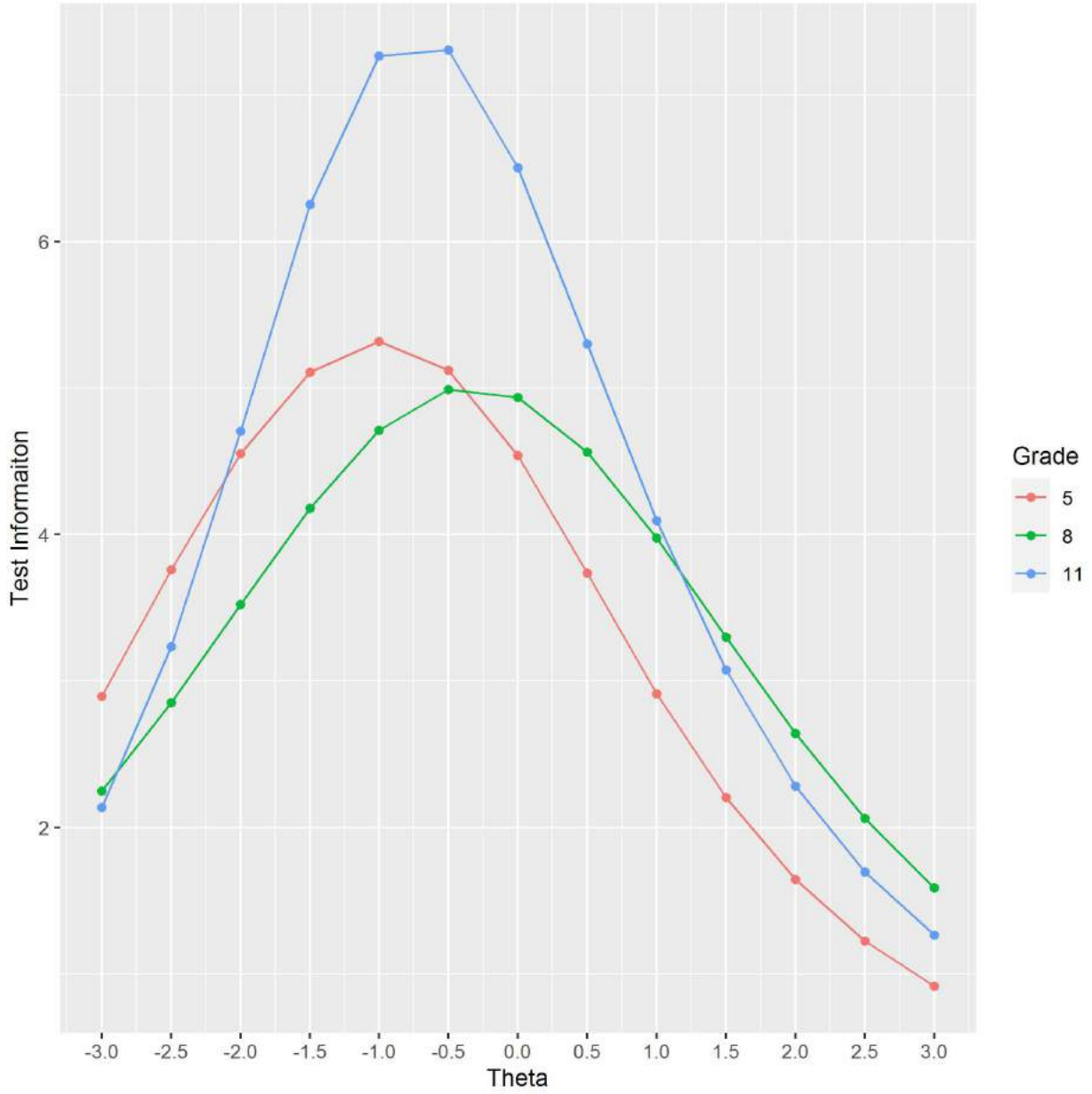


Figure IV-4. Test Information Function for Science



In IRT, a standard error is also estimated for each value of theta, called the *conditional standard error of measurement* (CSEM). CSEMs are computed through their inverse relationship with TIFs. Graphical representations of CSEM curves for theta values ranging from -3 to 3 in increments of 0.5 for each grade in ELA, mathematics, and science can be found in Appendix A. For the mathematics and science tests where there is one fixed form per grade, there is only one CSEM curve per grade because there is only one test form per grade. For the two-stage adaptive ELA tests, there are two TIFs per grade because there is either one easy form or one hard test form of Session 2 paired with the medium form of Session 1. Typical, CSEM values are low at the center of the theta distribution and gradually increase toward the two ends of the scale, whereas thetas

become very low or very high and result in a U-shaped pattern. For ELA, grades 3–6 and 8 had CSEMs reaching the minimum value at different theta values for the easy and hard test forms. The easy form had a smaller theta value with the minimum CSEM than the hard form. For ELA grades 7 and 10, the CSEM of the easy form was smaller than the CSEM of the hard form at the low theta range and larger at the high theta range. Mathematics had CSEMs reaching the minimum value at a theta value around 0, which is close to a level 3 cut. The CSEM of the grade 11 science test was smaller than the CSEMs of grades 5 and 8, which is consistent with the reliability results, where reliability of grade 11 science was higher than the reliabilities of grades 5 and 8.

IV.1.3. Classification Consistency and Accuracy

Classification consistency and accuracy indicate the decree of how accurately students are classified into performance levels. Performance level classification consistency and accuracy is of great interest for testing programs that serve accountability purpose. According to Livingston and Lewis (1995), *classification consistency* refers to "the agreement between the classifications based on two nonoverlapping, equally difficult forms of the test" (p. 180), and *classification accuracy* refers to "the extent to which the actual classifications of test takers on the basis of their single-form scores agree with those that would be made on the basis of their true scores, if their true scores could somehow be known" (p. 180). The detailed calculation of both indices can be found in the *2017 KAP Technical Manual*. Both classification consistency and accuracy indices range from 0 to 1, with 0 representing classifications that are not consistent or accurate and 1 representing perfectly consistent or accurate classifications.

The results for overall classification across all four performance levels as well as for the dichotomies created by the three cut scores are presented in Table IV-6. For the overall KAP classification, the classification consistency indices range from .39 to .65 and the classification accuracy indices range from .66 to .82 across all grades and subjects. For different KAP performance level classifications, the level 3 cut (i.e., 1, 2 vs. 3, 4) classification is the most important because the level 3 cut is the proficient cut. The classification consistency indices range from .53 to .82 and the classification accuracy indices range from .87 to .99 across all cuts, grades, and subjects. For all subjects and grades except grade 8 science, the level 3 cut classification consistency index was higher than the other two cuts' classification consistency indices. For the same grade, classification consistency and accuracy for the science tests were lower than for the other two subject tests because science tests are shorter.

Table IV-6. Classification Consistency and Accuracy

Subject and grade	Cut-score category							
	Overall		1 vs. 2, 3, 4		1, 2 vs. 3, 4		1, 2, 3 vs. 4	
	C	A	C	A	C	A	C	A
ELA								
3	.62	.80	.75	.92	.78	.92	.75	.95
4	.59	.80	.66	.93	.74	.91	.70	.96
5	.57	.77	.71	.92	.74	.91	.72	.94
6	.59	.79	.73	.91	.73	.91	.61	.97
7	.57	.78	.71	.90	.73	.91	.66	.97
8	.59	.81	.69	.91	.71	.92	.61	.98
10	.62	.82	.73	.92	.74	.93	.65	.98
Mathematics								
3	.60	.79	.64	.92	.79	.92	.79	.96
4	.65	.83	.66	.92	.81	.94	.78	.97
5	.62	.81	.68	.90	.81	.94	.80	.97
6	.61	.81	.66	.89	.81	.94	.80	.98
7	.58	.81	.57	.89	.77	.95	.70	.99
8	.61	.82	.71	.90	.78	.94	.73	.98
10	.60	.82	.68	.89	.82	.96	.80	.98
Science								
5	.39	.66	.53	.88	.64	.87	.61	.93
8	.45	.72	.66	.87	.64	.90	.57	.96
11	.44	.70	.57	.86	.69	.90	.66	.95

Note. ELA = English language arts; C = consistency; A = accuracy.

IV.1.4. Subscore Reliability

In addition to the total test score, the scores of subsets of ELA, mathematics, and science items are reported for students, and called subscores. The number of items in each subscore varies, and some items contribute to multiple subscores. The minimum number of items reported for a subscore is six. ELA and science have the same subscores across grades, but mathematics has different subscores across grades. Detailed information about the subscores in each subject can be found in the *2017 KAP Technical Manual*. These subscores are reported in three categories: below proficiency, meets proficiency, and exceeds proficiency. The detailed scoring procedure and rules for determining subscore categories can be found in the *2017 KAP Technical Manual*.

Two analyses were conducted to determine the reliability of subscores. First, the subscore marginal reliabilities were computed. Then, the classification consistency and accuracy of the subscore categories' overall classification was examined. Appendix B includes the marginal reliability, classification consistency, and classification accuracy for different subscores for each subject and grade. In summary, the averages of reliability, consistency indices, and accuracy indices are about .60, .35, and .70 respectively for all three subjects, which indicates the

subscores provide reasonable, reliable results. There is some variability in the reliability, classification consistency, and classification accuracy across each subscore by subject and form. The subscore reliabilities across the three subjects ranged from .41 to .77. The classification consistency indices across the three subjects ranged from .10 to .52. The classification accuracy indices across the three subjects ranged from .49 to .89. Reliability, classification consistency, and classification accuracy are all affected by the number of items measuring each subscore. Some subscores are measured by only six items, whereas other are measure by 30 items. The reliability, classification consistency, and classification accuracy of subscores with fewer items are expected to be low.

IV.2. Fairness and Accessibility

During the development and administration of the KAP assessment, ensuring the accessibility to all students and fairness across student groups were considered in every step. Universal design was used as a guide during the development of items, test formats, and the online test delivery interface to ensure accessibility to all students. Detailed descriptions of applying universal design in the development and administration of the KAP assessment can be found in the *2017 KAP Technical Manual*. All operational items have passed the bias-and-sensitivity review to mitigate the likelihood of content bias toward any one student group.

IV.2.1. Fairness

According to the *Standards for Educational and Psychological Testing*, "the central idea of fairness in testing is to identify and remove construct-irrelevant barriers to maximal performance for any examinee" (American Psychological Association [APA] et al., 2014, p. 74). This identifies fairness as an issue related to the validity of test-score inferences. Evidence in support of any assertion about the fairness of an assessment comes from several stages, such as item and test-development stage before test administration and differential item functioning (DIF) analyses after tests were administered to student population. Detailed fairness evidence on item and test development can be found in the *2017 KAP Technical Manual*. DIF results can be found in Section III.3.1. Differential item functioning indicates only two items were identified with moderate DIF, and further internal review did not find any content bias of these two items. DIF analysis examines whether an item shows any statistical difference between two groups of students after controlling for student proficiency. The few items with DIF contribute to the evidence in support of fairness during item writing and reviewing processes.

IV.2.2. Accessibility

According to the *Standards for Educational and Psychological Testing*, "accessibility is the degree to which the items or tasks on a test enable as many test takers as possible to demonstrate their standing on the target construct without being impeded by characteristics of the item that are irrelevant to the construct being measured" (APA et al., 2014, p. 215). Evidence in support of accessibility of an assessment consist of inclusion and accommodations and the implementation of universal design in items as well as test development and administration. Detailed accessibility evidence of KAP can be found in the *2017 KAP Technical Manual*.

IV.3. Full Performance Continuum

The KAP assessment was developed with the goal that assessment of each subject area and grade level would provide a reasonably precise estimation of student proficiency across the full performance continuum (i.e., from low-performing to high-performing students). The evidence on TIFs and CSEMs from Section IV.1.2. Test Information indicates KAP tests can accurately estimate ability across the full theta scale, especially at the middle of the scale.

This goal is also met by using items that cover different cognitive complexity levels and a wide range of difficulties. The KAP items' cognitive complexity levels were measured by depth of knowledge (DOK) framework (Webb, 1997). The Kansas Standards specify the maximum DOK for each cluster. When test items are written to each cluster, the items also have to reflect the expected DOK level as implied by the content to be measured. This expectation is emphasized throughout item writing and during both internal and external item reviews. Consequently, when the items selected for a test meet the blueprint, those items also meet the underlying DOK requirements. Because the 2019 test forms were also used in 2021, the summary of DOK levels, classical test theory item statistics, and IRT item statistics can be found in the *2019 Technical Manual Addenda*.

IV.4. Scoring and Scaling

This section introduces the procedures of scoring individual items, scoring the test as a whole, and scaling. The test results and performance level distribution of 2021 KAP testing are included. Also, the KAP performance trend for five years is presented. Finally, the quality-control procedures used to ensure the accuracy of scoring and scaling are described.

IV.4.1. Scoring

Item and test scoring in the 2021 administration remained the same as in previous years. The detailed description about item and test scoring can be found in the *2017 KAP Technical Manual*.

IV.4.2. Scaling

Scaling is the procedure of transforming thetas or raw scores to a reporting scale. The purpose of scaling is to facilitate the use and interpretation of test scores. The same scaling procedure and KAP reporting scale were used in 2021 as in previous years. Detailed information about scaling procedure, scale-transformation constants, and scale properties can be found in the *2017 KAP Technical Manual*.

IV.4.3. Operational Test Results

In this section, the results of the 2021 administration of the KAP are presented. Descriptive statistics representing the number of students tested by various sub-groups are presented. Then, the 2021 scale-score summary is included for all students and by sub-group. After that, the 2021 performance level distribution for each subject by grade is presented. Finally, the 2021 scale-score summary and proficiency rates are compared with those of previous years. Participation rates are included prominently in this report because it critical to take variability in participation into account when interpreting KAP performance within and across years.

IV.4.3.1. Participation Data

In 2021, the KAP operational test was administered in ELA, mathematics, and science in grades 3 through 8 and high school. At the high school level, students completed ELA and mathematics assessments in grade 10 and science assessments in grade 11. As described in Section I.3. Required Assessments and Intended Population, Kansas is committed to including all students in the KAP assessment.

Table IV-8 shows the number of enrolled students and tested students as well as participation rate by subject and grade. The definitions for all the indicators are as follows:

- Enrolled students are students assigned to take a KAP test.
- Tested students are students receiving a score report. Students received a score report when they were not exempt (exemption rules are described in Section I.3. Required Assessments and Intended Population), finished at least five items in each of the two test sections, and had logged out of the testing platform for the first section. This reporting rule has been used since 2015.
- The *participation rate* is calculated as the number of tested students divided by the number of enrolled students.

As can be seen in Table IV-7, for each subject and grade, more than 30,000 students were tested. Across all subjects and grades, the participation rates ranged from 86% to 96%. Elementary and middle school grades had a greater than 90% participation rate, especially at elementary grades (about 96%). High school grades had a lower participation rate, with 88% for ELA, 89% for mathematics, and 86% for science. Across all subjects and grades, the average participation rate was 93%. The participation rate in 2020–2021 was lower than participation rates in previous administrations (e.g., average participation rate in 2018–2019 was 98|%) because of the impact of COVID-19.

Table IV-7. Number of Enrolled and Tested Students With Participation Rate by Subject and Grade

Grade	English language arts			Mathematics			Science		
	Enrolled (*N*)	Tested (*N*)	PR (%)	Enrolled (*N*)	Tested (*N*)	PR (%)	Enrolled (*N*)	Tested (*N*)	PR (%)
3	35,440	34,031	96	35,455	34,044	96	-	-	-
4	35,547	34,109	96	35,557	34,131	96	-	-	-
5	36,735	35,146	96	36,743	35,125	96	36,756	34,817	95
6	37,225	35,058	94	37,224	34,971	94	-	-	-
7	38,145	35,280	92	38,142	35,204	92	-	-	-
8	38,275	35,354	92	38,286	35,269	92	38,301	34,957	91
10	36,811	32,555	88	36,813	32,622	89	-	-	-
11	-	-	-	-	-	-	35,527	30,646	86

Note. PR = participation rate (enrolled / tested).

To aid in an examination of the academic impact of COVID-19 on different student groups and regions, the participation rates by student group[3] and by School Board of Education (SBOE) districts are in Table IV-8 through Table IV-10. The 286 school districts in Kansas are distributed among 10 SBOE districts. Some school districts appear in multiple SBOE districts, when school district boundaries reach into more than one SBOE district. This document lists the school districts included in each SBOE district. Comparing the participation rates of students by gender, ethnicity, race, EL status, and disability status, within each subject and grade we see the following results:

- no difference in participation rates between gender groups
- higher participation rates for White students than for Black students
    - the difference in participation rate between White and Black students was greater than 10% (e.g., in grade-10 ELA, in grade-7 mathematics)
- a slightly higher participation rate for non-Hispanic students than for Hispanic students in middle and high schools
- a slightly higher participation rate for ELs than for non-ELs in elementary schools
- a slightly higher participation rate for non-ELs than for ELs in middle and high schools
- a slightly higher participation rate for students without disabilities than for students with disabilities

The comparison of participation rates of different SBOE districts within each subject and grade showed the following results:

- a very high ($\geq$ 95%) participation rate in districts 5, 6, and 9, with some participation rates as high as 99% in elementary and middle schools
- the lowest elementary school participation rates were in district 4
- the lowest middle school participation rates were in districts 1 and 4
- the lowest participation rates in high school were in district 8

Districts 5 and 6 are in the western part of Kansas, and district 9 is in the southeastern part of the state; all three districts are rural. Districts 1 and 4 include the Kansas City, Topeka, and Lawrence school districts. District 8 includes the Wichita school district. The detailed demographic distribution of SBOE districts can be found in Appendix C.

Table IV-8. English Language Arts Participation Rate by Demographic Characteristic and State Board of Education District

Characteristic	Grade						
	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	10 (%)
Gender							
Female	96	96	96	94	92	92	88
Male	96	96	95	94	93	93	89

[3] Economic disadvantaged status is not shared with ATLAS to protect the privacy of students, so this student group is not included in the comparison.

Characteristic	Grade						
	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	10 (%)
Race							
American Indian	97	96	96	94	93	91	89
Asian	94	95	95	90	91	91	86
Black	92	90	90	87	83	84	76
NHPI	94	96	95	89	91	88	84
Other	94	95	93	91	89	88	83
White	97	97	96	95	94	94	90
Hispanic							
No	96	96	96	95	93	93	90
Yes	96	96	96	93	91	91	84
Student with Disability							
No	96	96	96	94	93	93	89
Yes	95	95	94	93	91	90	85
English Learner							
No	96	96	96	94	93	93	89
Yes	97	97	96	93	92	91	81
District							
1	94	94	94	91	85	85	80
2	96	97	96	95	91	90	88
3	95	96	96	95	92	91	88
4	91	92	91	88	85	85	80
5	97	97	97	97	97	96	96
6	97	99	97	97	97	97	95
7	95	95	94	92	91	91	85
8	94	93	93	89	88	88	79
9	99	99	99	99	98	98	96
10	95	94	94	92	90	91	84

Note. NHPI = Native Hawaiian and Pacific Islander.

Table IV-9. Mathematics Participation Rate by Demographic Characteristic and State Board of Education District

Characteristic	Grade						
	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	10 (%)
Gender							
Female	96	96	96	94	92	91	88
Male	96	96	95	94	93	93	89
Race							
American Indian	97	96	96	94	93	91	89
Asian	94	95	95	90	91	91	86
Black	92	90	90	86	82	83	76
NHPI	94	96	95	87	91	88	83
Other	93	95	93	91	89	88	84
White	97	97	96	95	93	93	90
Hispanic							
No	96	96	96	94	93	93	90
Yes	96	96	96	92	91	90	84
Student with Disability							
No	96	96	96	94	93	92	89
Yes	95	95	94	92	91	90	85
English Learner							
No	96	96	95	94	92	92	89
Yes	97	97	97	93	92	91	82
District							
1	95	94	94	90	85	84	80
2	96	97	96	95	91	90	88
3	95	96	96	94	92	91	89
4	91	92	91	88	85	85	80
5	97	97	97	97	97	96	96
6	97	99	97	97	97	97	95
7	95	95	94	92	91	91	85
8	94	93	93	89	88	88	79
9	99	99	99	99	98	98	96
10	95	94	94	92	90	91	85

Note. NHPI = Native Hawaiian and Pacific Islander.

Table IV-10. Science Participation Rate by Demographic Characteristic and State Board of Education District

Characteristic	Grade		
	5 (%)	8 (%)	11 (%)
Gender			
Female	95	91	85
Male	95	92	88
Race			
American Indian	96	91	85
Asian	92	89	82
Black	88	82	70
Native Hawaiian and Pacific Islander	95	85	83
Other	92	87	81
White	96	93	88
Hispanic			
No	95	92	88
Yes	95	89	81
Student with Disability			
No	95	92	87
Yes	93	89	82
English Learner			
No	95	91	87
Yes	95	89	79
District			
1	94	84	78
2	91	86	84
3	91	88	85
4	91	85	77
5	97	96	95
6	97	96	94
7	94	91	81
8	92	87	74
9	99	98	95
10	94	90	81

For all tested students, Tables IV-11 through IV-13 show the percentage of students in each student group by grade for ELA, mathematics, and science. The student groups include gender, race, ethnicity, disability status, and EL status. The percentages of students in each student group were very similar across grades. There were about equal percentages of male and female students. The largest percent tested by race group was White and the largest percent tested by ethnicity group was non-Hispanic. There are more students without disability tested than students with disability tested and there are more non-ELs tested than ELs.

Table IV-11. English Language Arts Percentage of Tested Students by Demographic Characteristic and Grade

Characteristic	Grade						
	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	10 (%)
Gender							
Male	51.2	50.6	51.0	50.9	51.1	51.5	51.3
Female	48.8	49.4	49.0	49.1	48.9	48.5	48.7
Race							
AI	1.9	1.9	2.1	2.2	2.5	2.5	3.3
Asian	3.0	3.0	3.0	2.8	2.8	2.9	2.8
Black	6.8	6.8	6.8	6.8	6.6	6.5	6.2
NHPI	0.4	0.3	0.3	0.3	0.3	0.2	0.3
Other	7.0	7.1	6.8	6.7	6.5	6.6	6.3
White	81.0	80.8	81.0	81.2	81.3	81.3	81.0
Hispanic							
Yes	20.6	20.7	20.7	21.0	20.9	20.3	19.5
No	79.4	79.3	79.3	79.0	79.1	79.7	80.5
SWD							
Yes	15.5	15.6	14.8	13.9	12.9	12.1	11.1
No	84.5	84.4	85.2	86.1	87.1	87.9	88.9
EL							
Yes	13.7	13.6	12.1	11.7	9.9	8.6	8.1
No	86.3	86.4	87.9	88.3	90.1	91.4	91.9

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-12. Mathematics Percentage of Tested Students by Demographic Characteristic and Grade

Subgroup	Grade						
	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	10 (%)
Gender							
Male	51.2	50.7	51.0	50.9	51.1	51.5	51.3
Female	48.8	49.3	49.0	49.1	48.9	48.5	48.7
Race							
AI	1.9	1.9	2.1	2.3	2.5	2.5	3.3
Asian	3.0	3.0	3.0	2.8	2.8	2.9	2.8
Black	6.8	6.8	6.8	6.7	6.6	6.4	6.2
NHPI	0.4	0.3	0.3	0.3	0.3	0.2	0.3
Other	7.0	7.1	6.8	6.7	6.6	6.6	6.3
White	81.0	80.8	81.0	81.2	81.3	81.3	81.1
Hispanic							
Yes	20.6	20.7	20.8	21.0	20.8	20.3	19.5
No	79.4	79.3	79.2	79.0	79.2	79.7	80.5
SWD							
Yes	15.5	15.6	14.8	13.9	12.9	12.1	11.1
No	84.5	84.4	85.2	86.1	87.1	87.9	88.9
EL							
Yes	13.8	13.7	12.2	11.7	9.9	8.7	8.1
No	86.2	86.3	87.8	88.3	90.1	91.3	91.9

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-13. Science Percentage of Tested Students by Demographic Characteristic and Grade

Characteristic	Grade		
	5 (%)	8 (%)	11 (%)
Gender			
Male	51.1	51.5	52.0
Female	48.9	48.5	48.0
Race			
American Indian	2.1	2.5	3.4
Asian	2.9	2.9	3.1
Black	6.7	6.3	5.5
Native Hawaiian and Pacific Islander	0.3	0.2	0.3
Other	6.8	6.6	6.1
White	81.1	81.4	81.5
Hispanic			
Yes	20.8	20.2	18.6
No	79.2	79.8	81.4
Student with Disability			
Yes	14.9	12.1	10.3
No	85.1	87.9	89.7
English Learner			
Yes	12.2	8.7	8.0
No	87.8	91.3	92.0

IV.4.3.2. Operational Test Results

Summaries of scaled scores by grade are presented in Tables IV-14 through IV-16 for ELA, mathematics, and science. As noted previously, it is critical to take variability in participation into account when interpreting KAP performance within and across years.

The minimum and maximum scale score for each grade and subject were 220 and 380, respectively. As can be seen in Tables IV-14 to IV-16, the median scale scores were close to 300 in lower grades (i.e., grades 3–6 in ELA, grades 3–4 in mathematics, and grade 5 in science) and about 280 in the other higher grades. The standard deviation of scale scores are very similar across grades within one subject. Science tends to have higher standard deviations of scale scores than ELA and mathematics.

Table IV-14. Scale-Score Descriptive Statistics for English Language Arts

Grade	*M*	*SD*	Min.	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	Max.
3	293.1	29.8	220	254	270	291	313	335	380
4	297.8	28.4	220	261	277	296	318	334	380
5	294.7	29.3	220	257	273	293	316	335	380
6	288.3	28.9	220	250	267	289	309	326	380
7	287.0	30.2	220	250	265	284	306	325	380
8	281.5	28.0	220	246	261	280	299	318	380
10	282.6	29.4	220	245	261	282	301	323	380

Note. P_{10}, P_{25}, P_{50}, P_{75}, and P_{90} = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

Table IV-15. Scale-Score Descriptive Statistics for Mathematics

Grade	*M*	*SD*	Min.	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	Max.
3	299.8	28.6	220	265	279	298	318	339	380
4	289.6	28.7	220	255	266	286	307	329	380
5	287.9	27.3	220	259	266	283	305	327	380
6	287.1	26.8	223	258	268	281	303	325	380
7	284.6	26.8	220	253	265	280	299	320	380
8	281.3	27.4	220	251	261	277	296	319	380
10	282.6	26.2	220	257	266	276	294	319	380

Note. P_{10}, P_{25}, P_{50}, P_{75}, and P_{90} = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

Table IV-16. Scale-Score Descriptive Statistics for Science

Grade	*M*	*SD*	Min.	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	Max.
5	298.0	32.4	220	258	276	299	321	344	380
8	283.2	30.3	220	247	261	281	301	325	380
11	290.7	28.8	220	258	269	286	310	332	380

Note. P_{10}, P_{25}, P_{50}, P_{75}, and P_{90} = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

The percentage of students achieving each performance level (level 1 through level 4) and the proficiency rate (percentage at level 3 and level 4) are provided by subject and grade in Table IV-17 and Figures IV-5 through IV-7. The proficiency rates ranged from 21% to 48% across all subjects and grades. All three subjects tended to have lower proficiency rates in higher grade levels. Here is a summary of the results by subject:

- ELA:
  - Level 1 percentages ranged from 18% to 36%;
  - Level 2 percentages ranged from 27% to 46%;
  - Level 3 percentages ranged from 21% to 38%;
  - Level 4 percentages ranged from 4% to 15%.

  - The level 1 and level 2 percentages tended to increase, and the level 3 and level 4 percentages tended to decrease as grades increased.
- Mathematics:
  - Level 1 percentages ranged from 22% to 46%;
  - Level 2 percentages ranged from 30% to 48%;
  - Level 3 percentages ranged from 15% to 31%;
  - Level 4 percentages ranged from 3% to 17%.
  - The level 1 percentage tended to increase, and the level 3 and level 4 percentages tended to decrease as grades increased.
  - Level 2 percentage tended to be stable across grades.
- Science:
  - Level 1 percentages ranged from 28% to 43%,
  - Level 2 percentages ranged from 28% to 29%,
  - Level 3 percentages ranged from 20% to 29%,
  - Level 4 percentages ranged from 8% to 15%.
  - The level 1 percentage tended to increase, and the level 3 and level 4 percentages tended to decrease as grades increased.
  - Level 2 percentage tended to be stable across grades.

Table IV-17. Percentage of Students Achieving at Each Performance Level for English Language Arts (ELA) and Mathematics

Grade	ELA PL (%)					Mathematics PL (%)					Science PL (%)				
	1	2	3	4	P	1	2	3	4	P	1	2	3	4	P
3	31	29	26	14	40	22	30	31	17	48					
4	18	35	38	10	47	22	45	24	9	34					
5	27	30	29	15	43	34	37	19	10	29	28	28	29	15	44
6	36	27	32	4	37	37	36	20	8	28					
7	36	32	24	8	32	27	48	21	3	25					
8	29	46	21	4	25	45	34	17	5	22	43	29	20	8	28
10	34	38	23	5	28	46	33	15	6	21					
11											37	28	24	11	35

Note. PL = performance level; P = proficiency (combination of performance levels 3 and 4). Column percentages may not total 100 because of rounding.

Figure IV-5. Performance-Level Distribution for English Language Arts

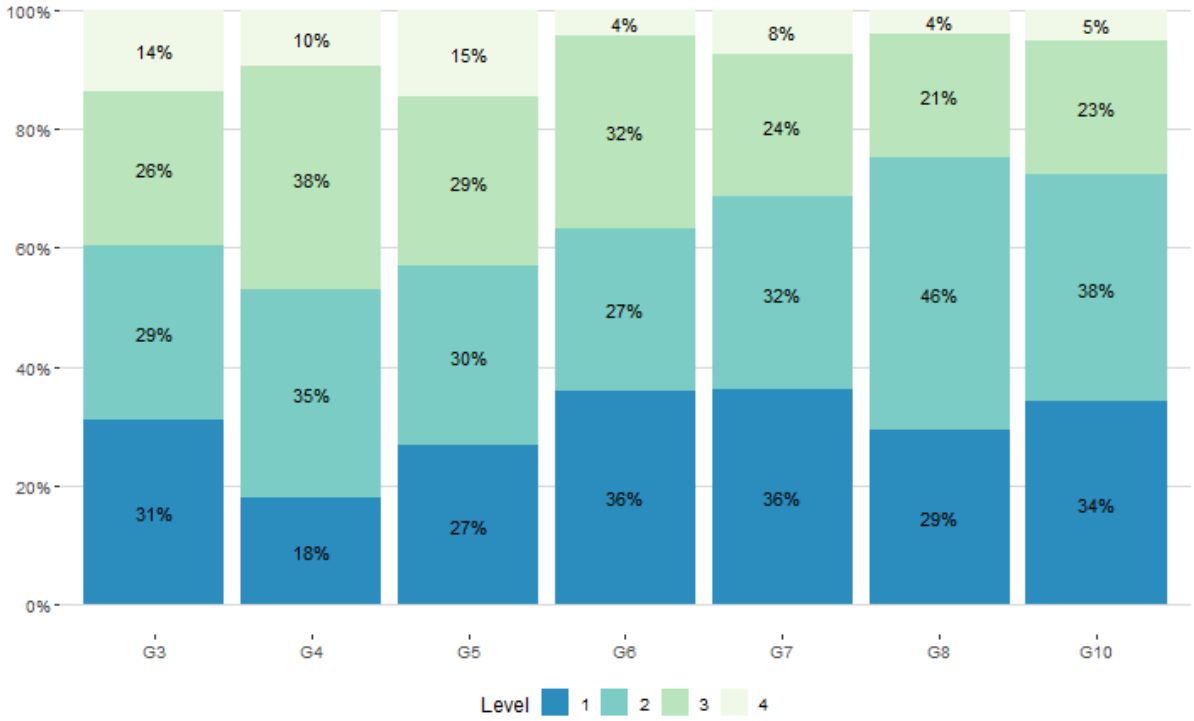


Figure IV-6. Performance-Level Distribution for Mathematics

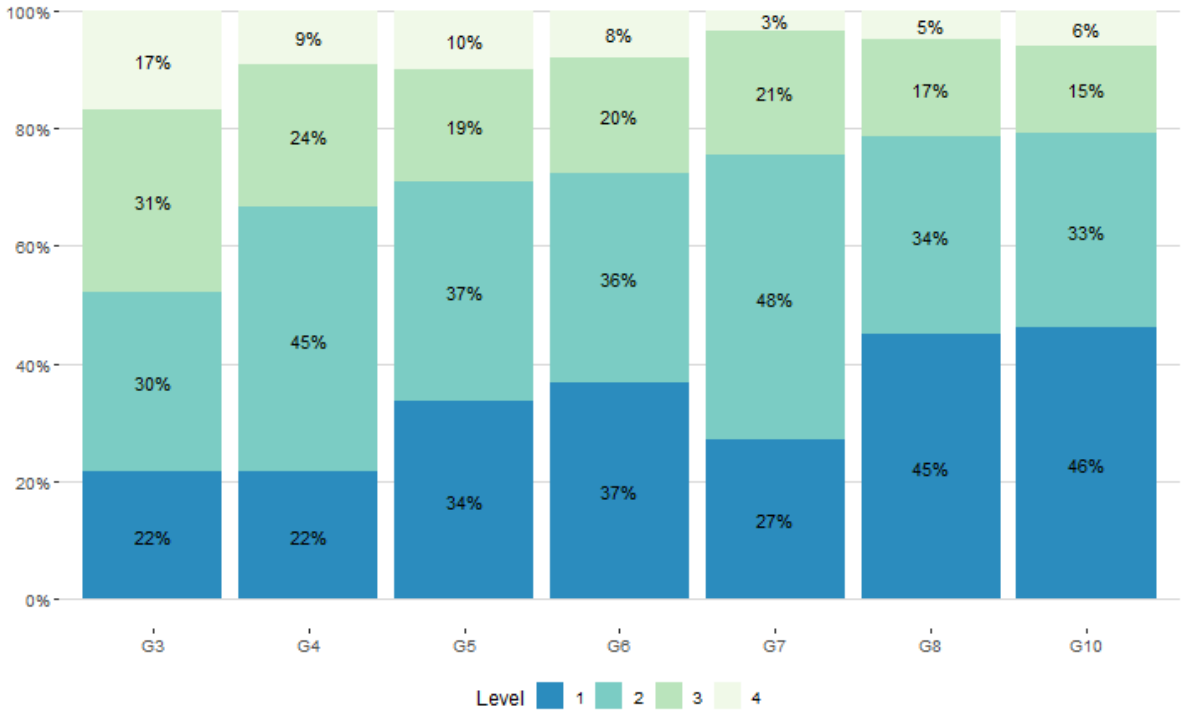
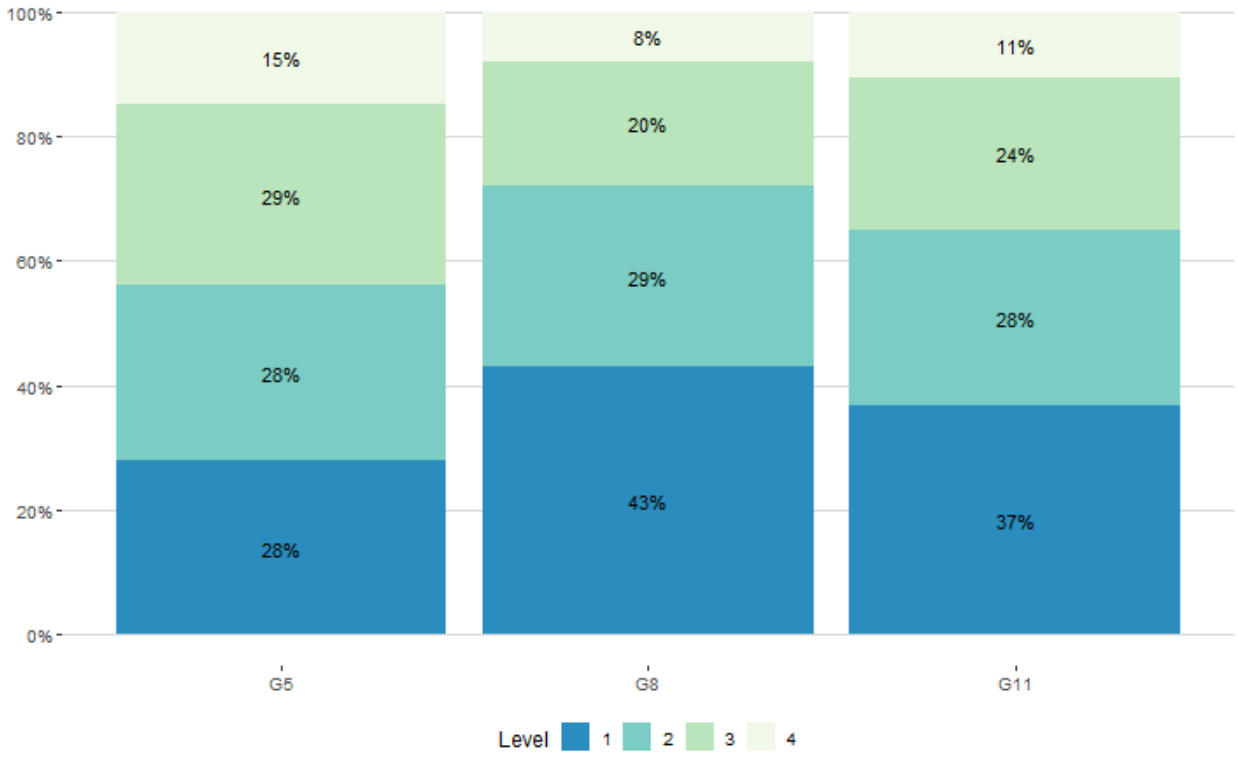


Figure IV-7. Performance-Level Distribution for Science



Tables IV-18 through IV-20 summarize the mean and standard deviation of the scale scores by demographic student groups[4]. For all subjects and grades, the mean scale score was above 280, and the standard deviation was around 30. The comparison of scale-score mean and the standard deviation of different student groups within each subject and grade indicates that female students scored higher in ELA and male students scored slightly higher in mathematics and science. Male students had higher standard deviations, Asian students had the highest means and standard deviations, and Black students had the lowest means and standard deviations. Non-Hispanic students had higher means and standard deviations than Hispanic students, non-ELs had higher means and standard deviations than ELs, and students without disabilities had higher means and standard deviations than students with disabilities.

[4] Economic disadvantaged status is not shared with ATLAS to protect the privacy of students, so this student group is not included in the comparison.

Table IV-18. English Language Arts Mean and Standard Deviation of Scale Scores by Grade and Student Subgroup

Subgroup	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 10	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Gender														
Male	291.2	29.9	295.5	28.3	292.1	29.0	286.2	28.9	284.1	30.7	278.2	27.7	278.4	29.6
Female	295.1	29.6	300.1	28.3	297.4	29.4	290.6	28.6	289.9	29.5	285.0	27.9	287.0	28.7
Race														
AI	280.0	25.4	286.7	25.4	282.3	26.7	274.8	25.6	273.2	26.4	268.5	24.9	270.4	25.8
Asian	299.3	30.6	306.3	31.2	304.9	30.7	299.1	30.4	297.8	31.7	294.8	30.0	293.3	32.3
Black	275.3	25.6	281.6	25.4	278.7	26.8	271.7	26.1	271.6	26.5	265.7	25.2	266.8	25.8
NHPI	281.0	25.3	287.5	28.0	284.0	26.8	275.4	27.8	281.6	32.9	272.1	27.5	271.7	25.2
Other	290.2	28.1	294.6	28.6	290.9	28.7	284.8	28.0	284.5	29.1	279.3	28.2	280.8	29.2
White	295.0	29.7	299.4	28.0	296.4	29.0	290.1	28.6	288.5	30.1	282.9	27.6	284.1	29.2
Hispanic														
Yes	280.4	26.2	286.4	25.4	283.2	26.3	276.0	26.2	275.2	26.7	271.0	25.0	271.4	26.4
No	296.4	29.8	300.8	28.4	297.7	29.3	291.6	28.7	290.1	30.4	284.2	28.1	285.3	29.5
SWD														
Yes	275.1	26.1	278.3	26.0	272.5	25.6	264.9	25.3	261.5	24.5	256.8	23.5	256.6	23.5
No	296.4	29.3	301.4	27.4	298.6	28.2	292.1	27.6	290.7	29.2	284.9	26.9	285.8	28.5
EL														
Yes	275.9	24.4	282.0	24.5	276.2	23.7	268.0	23.2	264.0	22.0	259.0	20.4	256.8	20.0
No	295.8	29.7	300.3	28.2	297.3	29.1	291.0	28.5	289.5	30.0	283.7	27.7	284.8	29.1

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-19. Mathematics Mean and Standard Deviation of Scale Scores by Grade and Student Subgroup

Subgroup	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 10	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Gender														
Male	301.7	29.9	292.0	30.1	289.9	28.9	288.4	27.9	286.2	28.1	282.1	28.9	283.0	27.4
Female	298.0	27.1	287.2	27.1	285.7	25.4	285.7	25.5	283.0	25.4	280.5	25.8	282.2	24.9
Race														
AI	287.6	23.6	279.5	23.8	278.3	22.2	276.6	19.4	274.2	22.3	270.0	22.0	272.4	18.9
Asian	308.5	31.1	301.6	33.6	301.2	33.7	301.1	33.7	299.7	34.2	300.1	35.3	302.4	36.0
Black	280.2	23.2	270.0	21.3	272.1	20.3	270.9	19.2	269.3	21.2	265.4	20.4	269.3	18.1
NHPI	287.9	24.5	279.1	23.1	280.5	23.9	275.3	23.8	277.2	25.3	277.4	26.7	272.3	19.7
Other	295.5	26.4	285.1	27.8	282.6	25.7	282.4	24.2	280.7	25.1	278.2	27.2	278.9	25.0
White	301.9	28.4	291.5	28.5	289.4	27.2	288.7	26.8	286.0	26.6	282.5	27.1	283.7	26.1
Hispanic														
Yes	287.4	24.2	277.6	23.9	276.8	22.0	276.0	20.8	273.7	21.6	270.7	22.2	272.8	19.9
No	303.1	28.8	292.8	29.0	290.8	27.8	290.1	27.4	287.5	27.3	284.0	28.0	285.0	27.0
SWD														
Yes	283.7	25.5	273.0	24.8	271.1	21.6	269.5	20.5	264.0	20.2	261.2	20.3	264.7	16.5
No	302.8	28.1	292.7	28.4	290.8	27.1	289.9	26.6	287.7	26.3	284.1	27.1	284.9	26.3
EL														
Yes	285.3	23.9	274.9	23.5	272.6	19.6	271.2	17.8	266.6	18.2	263.0	17.4	265.7	14.6
No	302.2	28.6	292.0	28.8	290.0	27.5	289.2	27.1	286.6	26.9	283.0	27.6	284.1	26.5

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; SWD = student with disability; EL = English learner.

Table IV-20. Science Mean and Standard Deviation of Scale Scores by Grade and Student Group

Subgroup	Grade 5		Grade 8		Grade 10	
	M	*SD*	*M*	*SD*	*M*	*SD*
Gender						
Male	299.6	33.7	285.4	32.1	292.2	30.7
Female	296.3	30.8	280.8	28.1	289.1	26.6
Race						
American Indian	286.2	29.9	269.4	25.7	276.3	23.4
Asian	305.3	33.3	293.4	31.8	299.2	31.9
Black	278.9	27.9	263.5	24.0	272.0	21.6
NHPI	283.2	26.2	273.1	25.2	280.1	21.5
Others	293.5	31.4	279.7	28.8	287.8	27.9
White	300.1	32.2	285.1	30.2	292.5	28.8
Hispanic						
Yes	285.5	28.7	270.0	26.0	279.0	24.3
No	301.3	32.5	286.5	30.4	293.4	29.1
Student with disability						
Yes	279.6	31.0	263.2	26.5	269.4	22.8
No	301.2	31.5	285.9	29.8	293.1	28.5
English learner						
Yes	278.6	26.0	259.5	21.2	267.4	18.5
No	300.7	32.3	285.4	30.1	292.7	28.7

Note. NHPI = Native Hawaiian and Pacific Islander.

IV.4.3.3. Performance Trend

ELA, mathematics, and science scale-score trends for 2018–2021 are presented in Tables IV-21, IV-22, and IV-23. The tables present the scale-score mean, standard deviation, and *N* count across administration years by grade. For all three subjects, the tested sample sizes were stable between 2018 and 2019. There was a decrease in tested sample size in 2021, and the decrease was higher in higher grades. For ELA grades 3, 4, 5, and 6, the mean scale scores were similar between 2018 and 2019, and there was a slight decrease in 2021. For ELA grades 7, 9, and 10, there was a slight decrease in mean scale scores from 2018 to 2021. The average difference in mean scale score was about 1 scale-score point across grades between years. For mathematics, the mean scale scores increased from 2018 to 2019 and decreased from 2019 to 2021.The average difference in mean scale score was about 4 scale-score points across grades between 2019 and 2021. For science grades 5 and 8, the mean scale decreased slightly from 2018 to 2021. For science grade 11, the mean scale scores decreased from 2018 to 2019 and increased from 2019 to 2021. The science grade 11 mean scale scores of 2018 and 2021 were very close to each other. For all three subjects, the standard deviations were similar across years.

Table IV-21. Longitudinal Scale-Score Trend for English Language Arts

Grade	2018			2019			2021		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
3	294.0	29.1	37,579	294.8	29.1	37,098	293.1	29.8	34,031
4	299.3	27.8	38,440	299.1	28.0	37,698	297.8	28.4	34,109
5	295.8	29.7	38,374	295.9	29.4	38,372	294.7	29.3	35,146
6	290.1	29.2	37,447	290.4	28.7	38,281	288.3	28.9	35,058
7	289.0	31.2	36,754	288.3	31.1	37,424	287.0	30.2	35,280
8	283.0	28.5	36,832	282.3	28.5	36,779	281.5	28.0	35,354
10	284.0	29.8	35,651	283.4	29.8	36,318	282.6	29.4	32,555

Table IV-22. Longitudinal Scale-Score Trend for Mathematics

	2018			2019			2021		
Grade	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
3	302.6	28.0	37,641	303.2	27.9	37,184	299.8	28.6	34,044
4	292.7	28.3	38,493	292.9	28.6	37,771	289.6	28.7	34,131
5	290.5	27.5	38,413	290.9	27.1	38,413	287.9	27.3	35,125
6	290.6	27.0	37,487	291.6	27.4	38,329	287.1	26.8	34,971
7	287.5	27.6	36,784	288.2	28.1	37,456	284.6	26.8	35,204
8	283.9	29.4	36,870	286.2	28.8	36,785	281.3	27.4	35,269
10	285.2	28.6	35,658	286.1	27.9	36,287	282.6	26.2	32,622

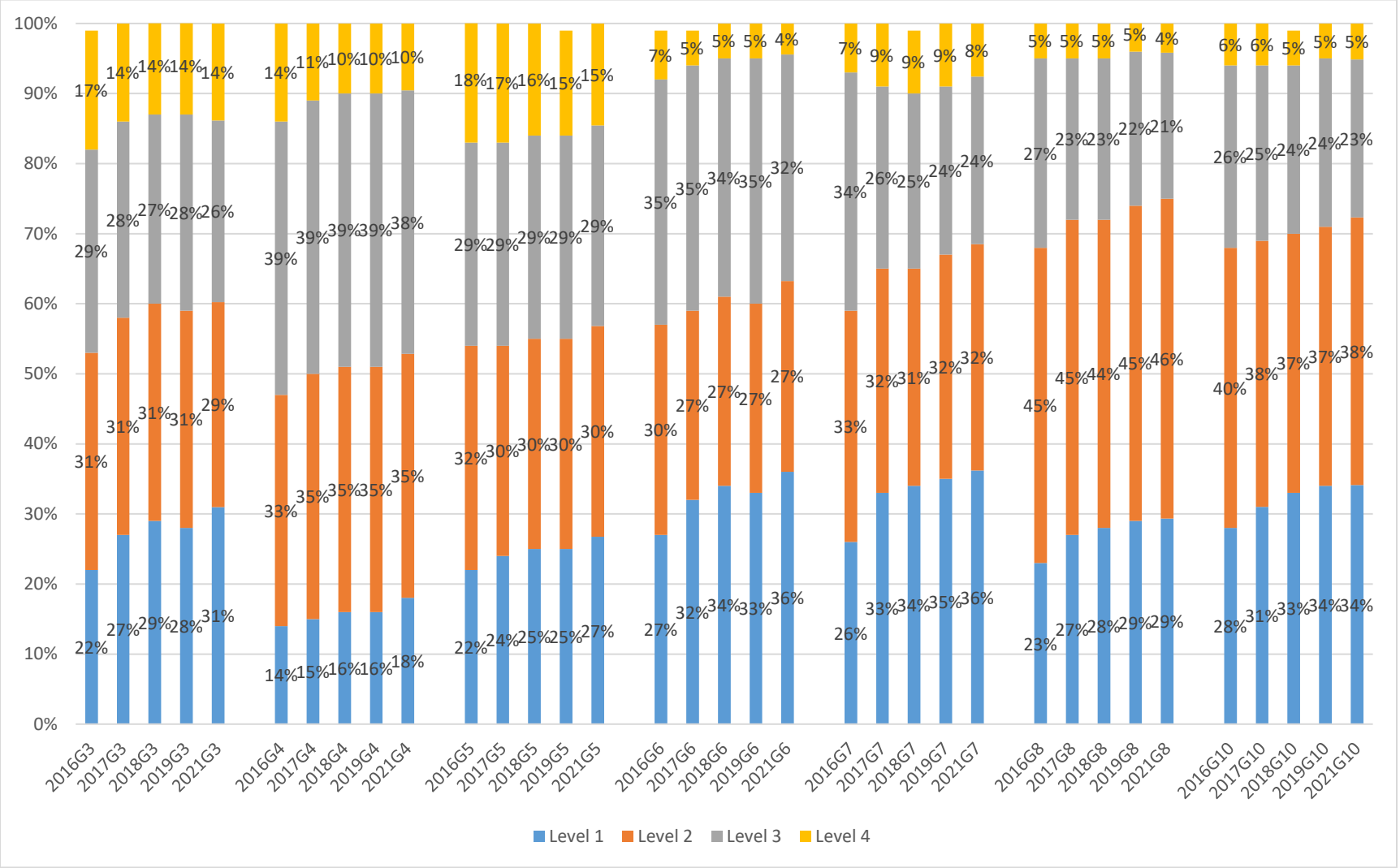
Table IV-23. Longitudinal Scale-Score Trend for Science

	2017			2018			2019		
Grade	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
5	299.0	30.6	38,458	298.7	30.8	38,442	298.0	32.4	34,817
8	287.5	29.8	36,934	286.9	30.3	36,863	283.2	30.3	34,957
11	291.2	29.4	34,314	289.3	29.8	34,081	290.7	28.8	30,646

The performance level distribution trends across years are presented in Figures IV-8, IV-10, and IV-12 for ELA, mathematics, and science. The proficiency rate distribution trends across years are presented in Figures IV-9, IV-11, and IV-13 for ELA, mathematics, and science. For ELA, there was a slight decrease in proficiency rates from 2016 through 2021, but the level 4 percentages were very similar across years. For mathematics, proficiency rates in most elementary grades decreased from 2016 through 2021, with decreasing rates increasing in 2021. Proficiency rates in middle school and high school grades increased in 2019 but decreased in 2021. Mathematics level 4 percentages were very similar across years. For grade 5 science, performance level distributions were similar between 2017 and 2018 and between 2019 and

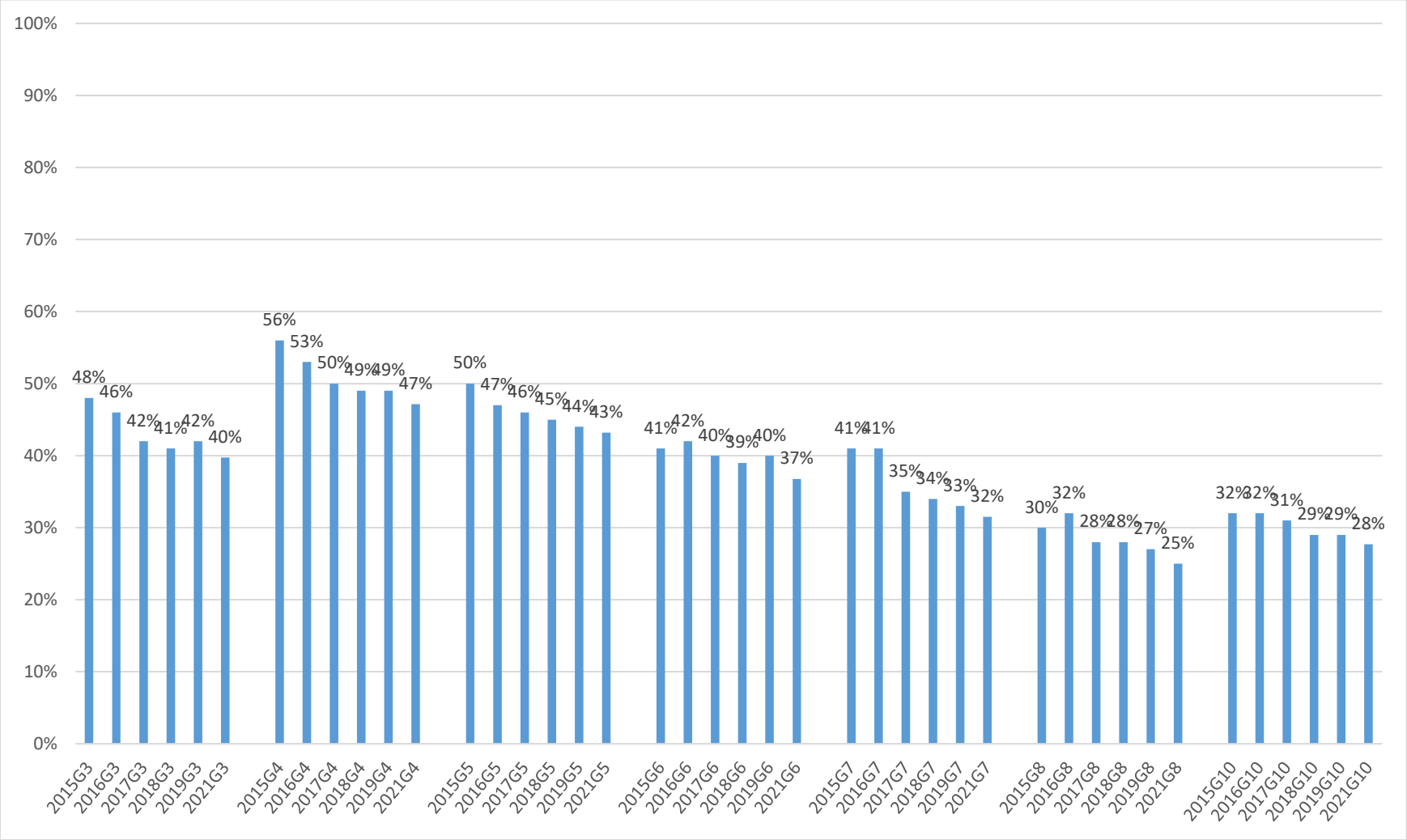
2021, with 2021 and 2019 having slightly lower proficiency rates. Also, grade 5 science had a large level 1 percentage in 2021. For grade 8 science, there was a slight decrease in proficiency rates from 2017–2021. For grade 10 science, 2017, 2018, and 2021 had very similar performance level distribution, and 2019 had a decrease in proficiency rate.

Figure IV-8. Performance-Level Distribution Trend for English Language Arts



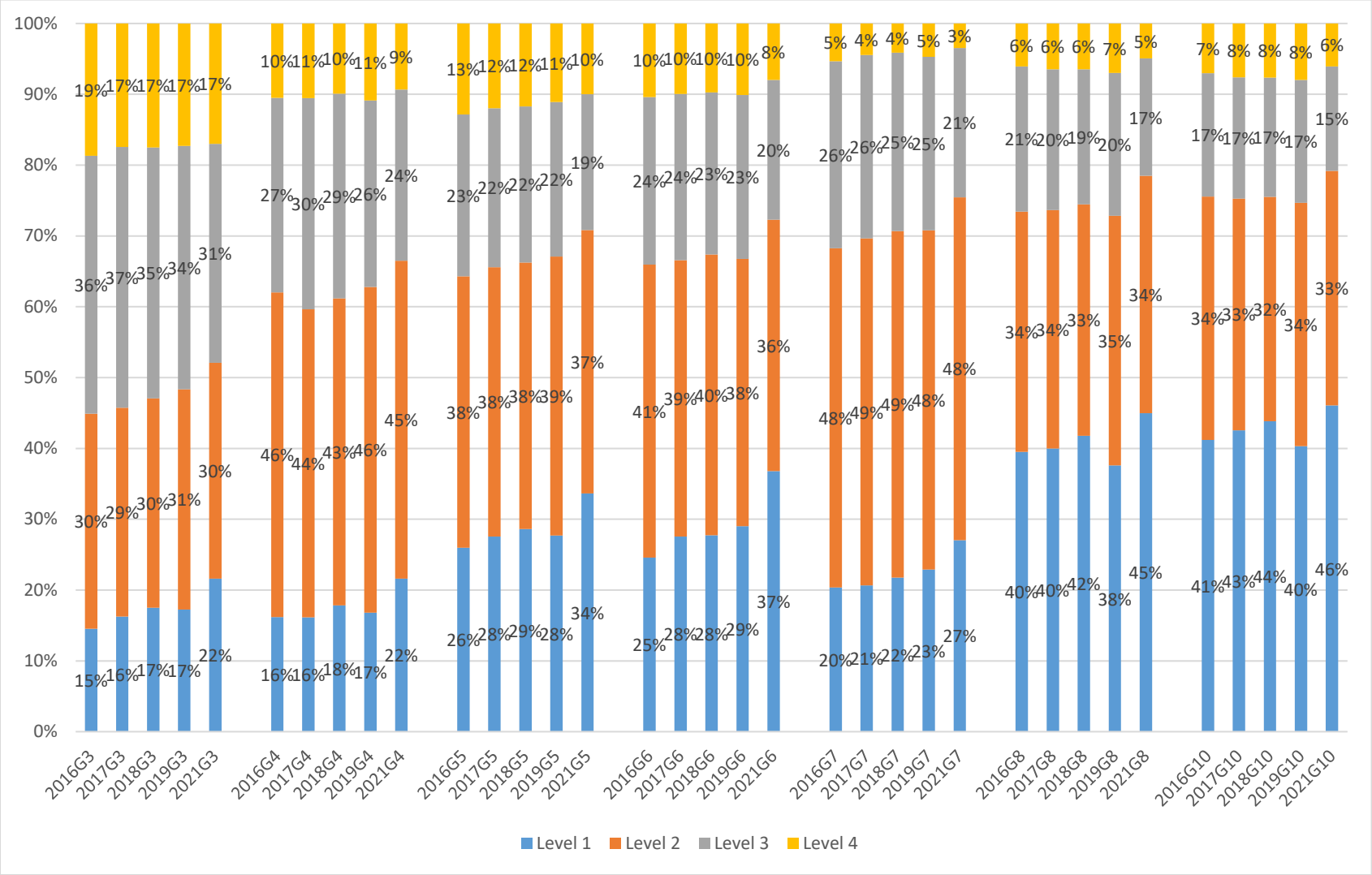
Note. G = grade. Labels of column percentages may not total 100% because of rounding.

Figure IV-9. Proficiency-Rate Trend for English Language Arts



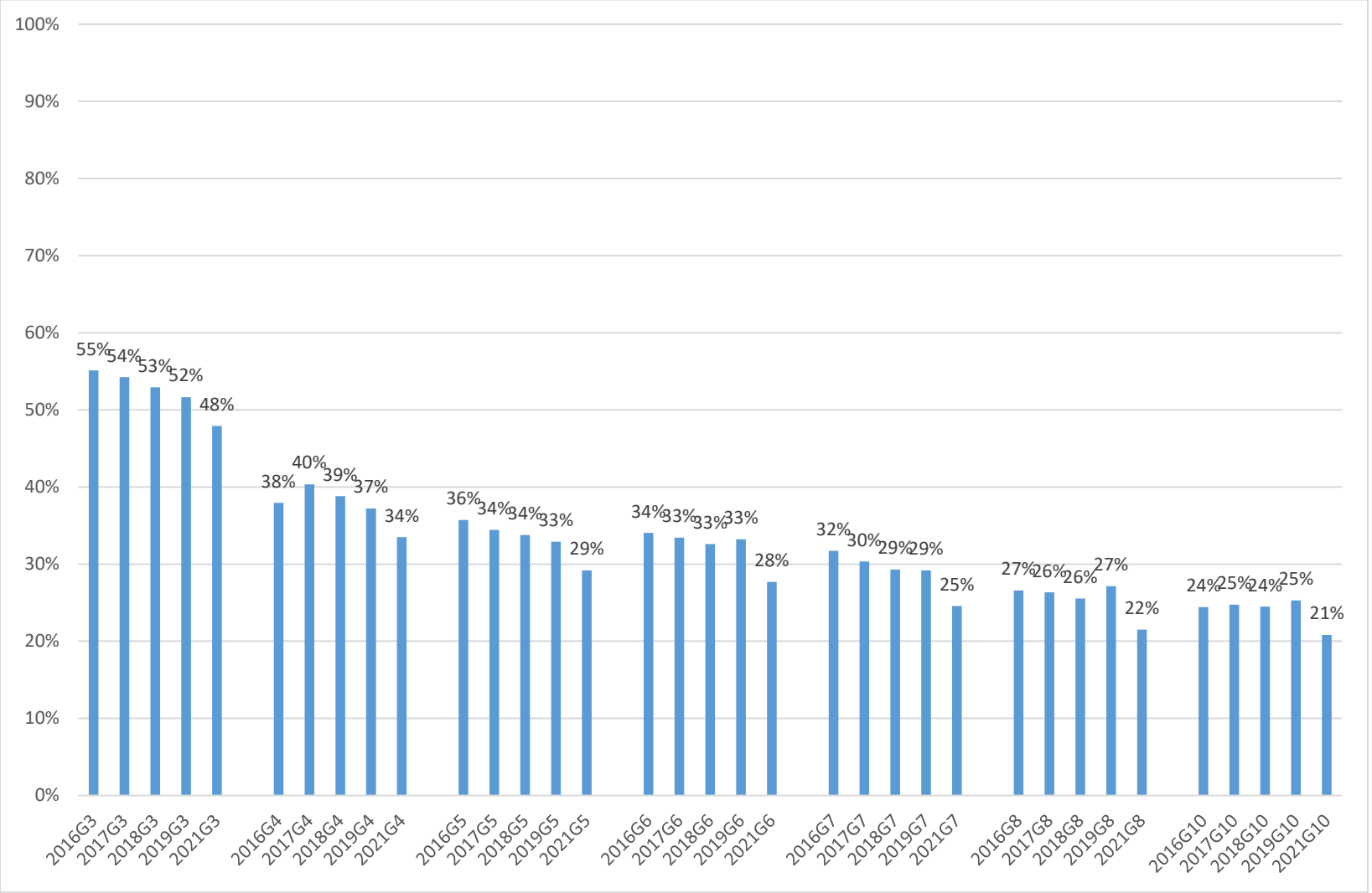
Note. G = grade.

Figure IV-10. Performance-Level Distribution Trend for Mathematics



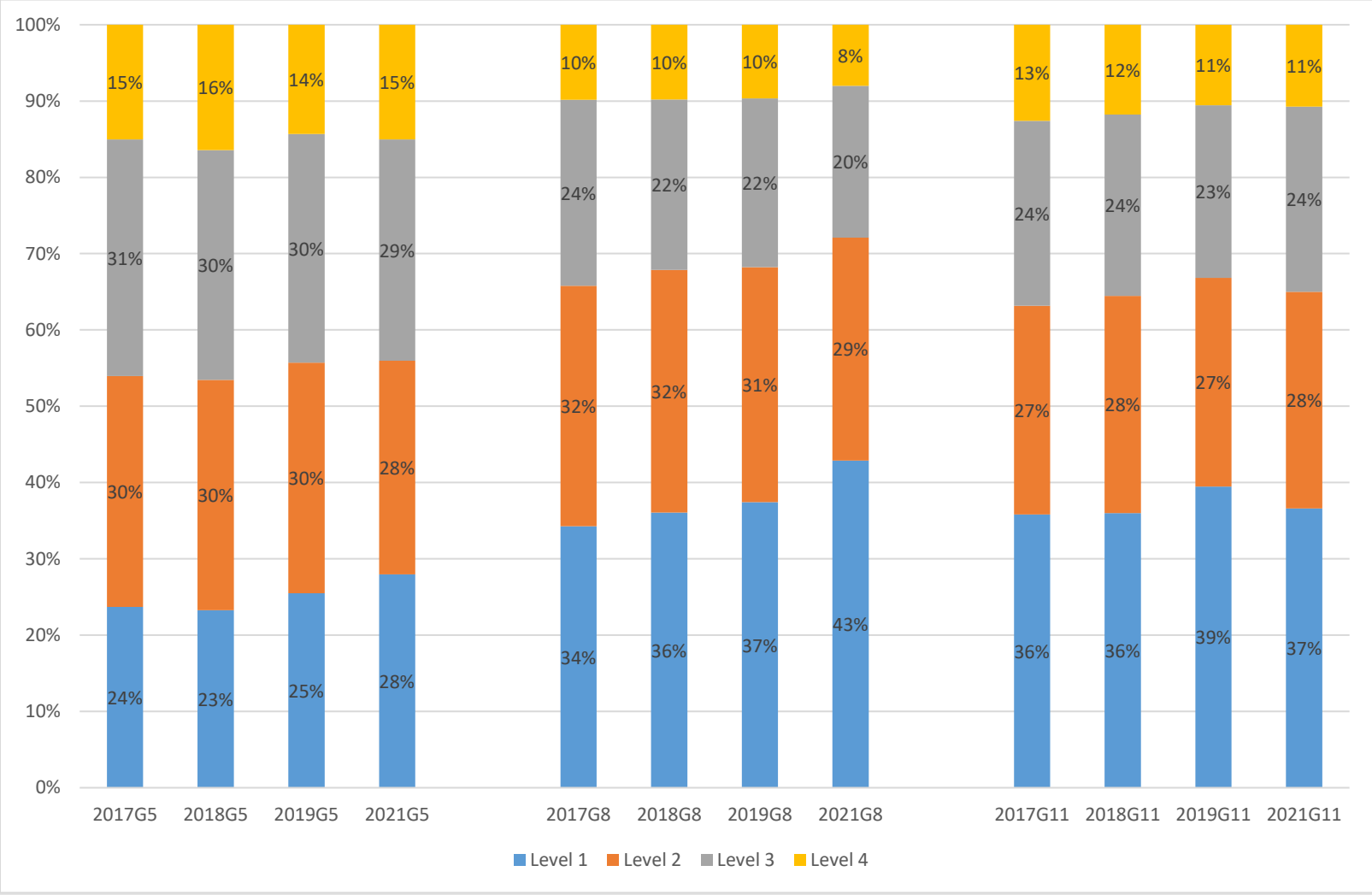
Note. G = grade. Column percentages may not total 100% because of rounding.

Figure IV-11. Proficiency-Rate Trend for Mathematics



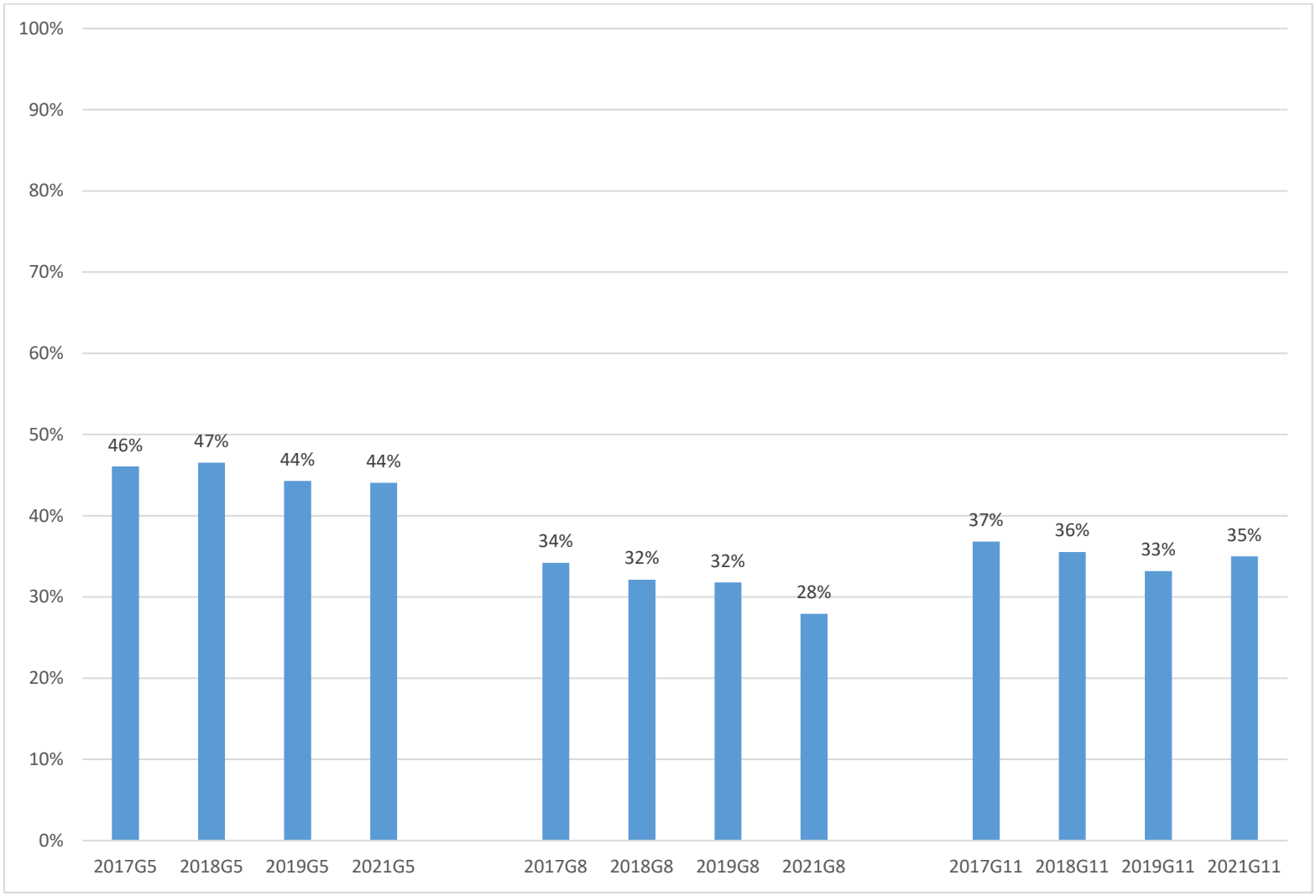
Note. G = grade.

Figure IV-12. Performance-Level Distribution Trend for Science



Note. G = grade. Column percentages may not total 100% because of rounding.

Figure IV-13. Proficiency-Rate Trend for Science



Note. G = grade.

The decrease in participation numbers may be one factor leading to the changes in performance trend in 2021. The research report on COVID-19 (Wang et al., in press) includes a detailed analysis of student groups and participation and performance trends in different regions. This report also examines the students who tested in both 2019 and 2021 and students who tested only in 2019, not in 2021.

IV.4.3.4. Quality-Control Checks

The scoring and reporting process of KAP test results has multiple quality-control steps. First, student-response data were checked at least three times during the testing window for scoring errors or duplicates. Second, classical item analysis was conducted during the testing window with approximately 20% of the overall test volume. The calculated classical item statistics from this year's data were compared with the classical item statistics from previous years' data stored. The purpose of this step was to monitor the classical item statistics trend and ensure items were functioning as expected. Third, items' IRT statistics were recalibrated, and classical statistics were calculated using this year's data after the window closed. Both newly calculated IRT and classical item statistics were compared with the statistics stored from previous years. This analysis and comparison helped us evaluate item drift. Fourth, the reasonableness and accuracy of the scoring tables were examined through predetermined criteria. Fifth, the cut scores used to classify students were checked independently by at least two people to ensure they were consistent with the cut scores approved by the SBOE. Sixth, the summary statistics of testing results were calculated and compared with those of previous years to ensure the performance trend was reasonable. Finally, the psychometric and technology teams independently calculated each individual's total score, scale score, performance levels, subscore scores, and subscore performance levels. The results from the two teams' independent calculation were compared to identify any differences or calculation errors. Students' score reports were generated only after the scoring results from both teams were identical. The purpose of all quality-control steps was to ensure the scoring results provided on students' reports were complete and accurate.

IV.5. Multiple Assessment Form

Mathematics and science tests only used one operational form per grade. ELA tests used the two-stage adaptive test design and had two assessment forms, easy and hard, per grade along with the medium difficulty in the first session of the form. All items on the ELA forms have been equated and placed on the same IRT scale. Detailed information on ELA form equating can be found in the *2017 KAP Technical Manual*. Moreover, the standardized administration procedure ensures comparability across ELA forms. Detailed descriptions of this administration procedure are in Section II.4. Test Administration of this manual.

IV.6. Technical Analysis and Ongoing Maintenance

This technical manual includes a series of technical analyses that were updated according to this year's testing data. These analyses include DIF analysis, relationships among different assessment, reliability analyses, classification consistency and accuracy analyses, test result summary, and trend analysis. All of the updated technical analyses provide evidence to indicate KAP is a reliable and valid assessment.

In addition to achievement data collected through KAP assessments, contextual data about learning environment and opportunity for learning are important for understanding the impact of the pandemic on student learning. Several surveys were administered throughout this year to collect contextual data. Research questions addressing the impact of the pandemic on student learning and achievement are reported separately in a research report on COVID-19 (Wang et al., in press). Analyses in the COVID research report were used to speculate about the impact of the COVID-19 pandemic on learning and achievement. The findings of the analyses were summarized, and recommendations based on these findings were made in the research report.

For future maintenance, the current grade-10 mathematics blueprint aligns to 2010 Kansas Standards. In 2022, there will be operational field testing for grade-10 mathematics to construct a new form that aligns to the new blueprint for 2017 Kansas Standards. Standard setting for grade-10 mathematics will be conducted in summer 2022 for this new form.

V. Inclusion of All Students

Kansas State Department of Education (KSDE) complies with the Individuals with Disabilities Education Improvement Act and the Elementary and Secondary Education Act, both of which require that all students, including students with disabilities, participate in assessments used for accountability purposes. One of the main applications of these two acts is to develop different accommodations to address each student's unique needs. Detailed information about how both acts are applied in the Kansas Assessment Program (KAP) can be found in the *2017 KAP Technical Manual*.

This chapter presents information about KAP's inclusion of all students and accommodation usage. Much of this information is also available in other KSDE documents (e.g., Tools and Accommodations for the Kansas Assessment Program and the *Kansas Assessment Examiner's Manual 2020–2021*). This chapter closes with a report on the frequency of use of specific accommodations.

To keep the test-administration conditions stable across years, especially through the COVID-19 pandemic, KSDE decided to allow only in-person testing in 2021; no remote testing was available. KSDE made efforts across the state to ensure that students returned to their schools for testing, no matter which learning mode (i.e., remote, in-person, or hybrid learning mode) the students were in. For students who were unable to return to schools for testing, special exemption rules were implemented. A student who was unable to attend in-person testing had to be approved by KSDE to be exempt from testing. KSDE also asked each district with students who were exempted from testing to provide its own evaluation of student learning that was equivalent to KAP. Moreover, the 2021 KAP testing window was postponed by 2 weeks. Districts also made efforts to maintain social distancing when bringing remote students in for testing. For more information about efforts districts made to maintain social distancing, please refer to Section II.4. Test Administration. With all of these efforts, the average participation rate was 93% for KAP across all subjects and grades.

V.1. Procedures for Including Students with Disabilities

KSDE is committed to including all students in the KAP assessments. The inclusion of students with disabilities is achieved by providing clear guidelines for educators to register their students with different needs. The *Kansas Assessment Examiner's Manual 2020–2021* describes step-by-step registration procedures for students who need accommodations.

V2. Accommodations

Assessment accommodations are practices and procedures that provide equitable access during instruction and assessments for students with special needs. These accommodations may not alter the assessment's validity, score interpretation, reliability, or security. They are designed to reduce or eliminate the effects of a student's disability; however, they do not alter learning expectations. The same rules for using accommodation on the KAP were used across years. The detailed rules for using accommodations on the KAP assessments and all available KAP accommodations can be found in the *2017 KAP Technical Manual*.

V.3. Frequency of Accommodation Use

Educators register their students with different needs through a Personal Needs Profile (PNP) in Kite® Educator Portal. The PNPs submitted by educators determine the availability of online test accommodations for individual students. The summary of PNP accommodation requests shown in Table V-1 indicates the number of students for whom each accommodation is requested. This table summarizes the PNPs by grade; note that some students may receive multiple accommodations. The table shows that "Text to Speech: Items" is the most commonly used accommodation option. This accommodation makes an audio recording of the test item available.

Table V-1. Frequency of Accommodation Requests by Grade

Accommodation	Grade							
	3	4	5	6	7	8	10	11
Auditory background	46	175	208	210	199	198	114	95
Background color	6	10	11	14	9	22	21	16
Braille	2	0	3	5	5	4	8	5
Color overlay	4	20	22	23	18	24	33	20
Foreground color	6	10	11	14	9	22	21	16
Invert color choice	0	1	1	1	3	9	4	5
Keyword-translation display	195	261	293	328	330	371	339	291
Magnification	23	26	54	49	72	53	67	52
Masking	5	6	6	6	11	13	13	2
Onscreen keyboard	4	4	11	4	2	3	7	12
American Sign Language	7	17	20	13	18	11	11	7
TTS: Items	3,345	3,793	3,932	3,775	3,550	3,298	2,797	2,611
TTS: Items and passages	900	957	900	725	652	548	286	0
Total	4,543	5,280	5,472	5,167	4,878	4,576	3,721	3,132

Note: TTS = Text to Speech.

VI. Academic Achievement Standards and Reporting

This chapter describes any updates related to achievement standards and reporting. The same achievement standards used since 2015 for ELA and mathematics and since 2017 for science were used in 2021. The format of score reports remains unchanged, with some caveat language related to the COVID-19 pandemic added to address the impact of the pandemic.

VI.1. State Adoption of Academic Achievement Standards for All Students

The state adopted the same policy academic achievement standards across grades and subjects in 2015. The detailed description of the policy academic achievement standards for the Kansas Assessment Program (KAP) can be found in the *2015 KAP Technical Manual* and *2017 KAP Technical Manual*.

VI.2. Achievement Standard Setting

The same achievement standards and cut scores used since 2015 for ELA and mathematics and since 2017 for science were used in 2021. The ELA and mathematics standard setting occurred in 2015. The procedures and outcomes can be found in the *2015 KAP Technical Manual*. The science standard setting occurred in 2017. The procedures and outcomes can be found in the *2017 KAP Technical Manual*.

VI.3. Challenging and Aligned Academic Achievement Standards

The KAP's content area and grade-specific academic achievement standards were developed to align with the state content standards—i.e., Kansas standards. The *2015 KAP Technical Manual* described the process of developing those content area and grade-specific performance-level description of ELA and mathematics. The *2017 KAP Technical Manual* described the process of developing those content area and grade-specific performance level description of science.

VI.4. Reporting

For each tested subject, the KAP assessment provides separate score reports to students, schools, and districts. The example of a KAP student score report and a KAP school/district report is in Appendix D. These reports include students' overall and subscore performances. A detailed description of KAP score reports can be found in *2017 KAP Technical Manual*. Moreover, language was added to all score reports for the 2021 administration to remind students, parents, and educators that learning conditions and student performance may have been affected by COVID-19. This caveat states

> "Please note a single test score does not provide a complete or precise measure of student achievement. When interpreting KAP results, please take into consideration other measures of student achievement. Also, consider how the conditions for learning, which may have been disrupted by the pandemic, may influence performance" (Appendix D).

To help educators and parents interpret KAP results, the KAP Educator Guide and the KAP Parent Guide are also published on the KAP website so that educators and parents can access them easily. Both guides include a letter from Dr. Randy Watson, Kansas Commissioner of Education; an overview of test purposes, content, and format; descriptions of the KAP scoring

process; suggestions for how to use test scores and improve KAP scores; and an explanation of different information presented on the score reports.

The KAP testing window ended on May 18, 2021. One week later, KAP ELA, mathematics, and science score reports were available for KSDE review. After KSDE approved the score reports, these reports were made available to districts and then to the parents.

References

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Cohen, J. (1988). *Statistical power analysis for the behavioral* sciences (2nd ed.). Lawrence Erlbaum Associates.

Forte, E., Nebelsick-Gullett, L., Deters, L., Buchanan, E., Herrera, B., Morris, J., & Phlegar, J. (2016). *Kansas Assessment Program alignment evaluation report 2015–2016*. edCount LLC.

Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, *21*(4), 347–360. https://doi.org/10.1111/j.1745-3984.1984.tb01039.x

Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error and power rate using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, *14*(4), 329–349. https://doi.org/10.1207/S15324818AME1404_2

Kan. Stat. Ann. §72-5170 (2020). https://www.ksrevisor.org/statutes/chapters/ch72/072_051_0070.html

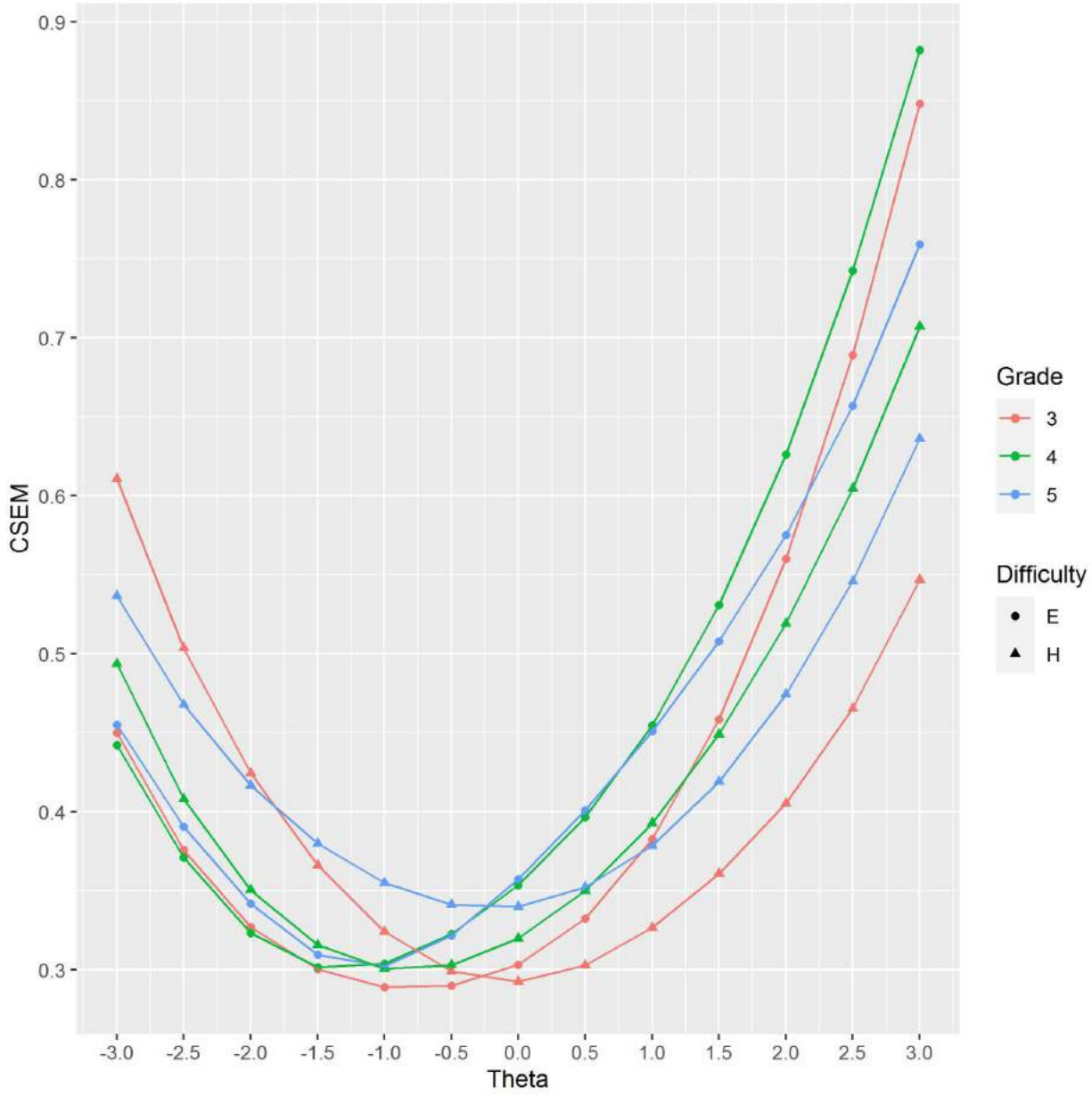
Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, *32*(2), 179–197. https://www.jstor.org/stable/1435147

Wang, W., Kobrin, J., & Tiemann, G. (in press). *Kansas Assessment Program COVID-19 effect on learning environment, opportunity to learn, and assessment 2020-2021 technical report*. University of Kansas, Center for Accessible Teaching, Learning, and Assessment Systems (ATLAS).

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 8). ERIC; Council of Chief State School Officers. https://files.eric.ed.gov/fulltext/ED414305.pdf

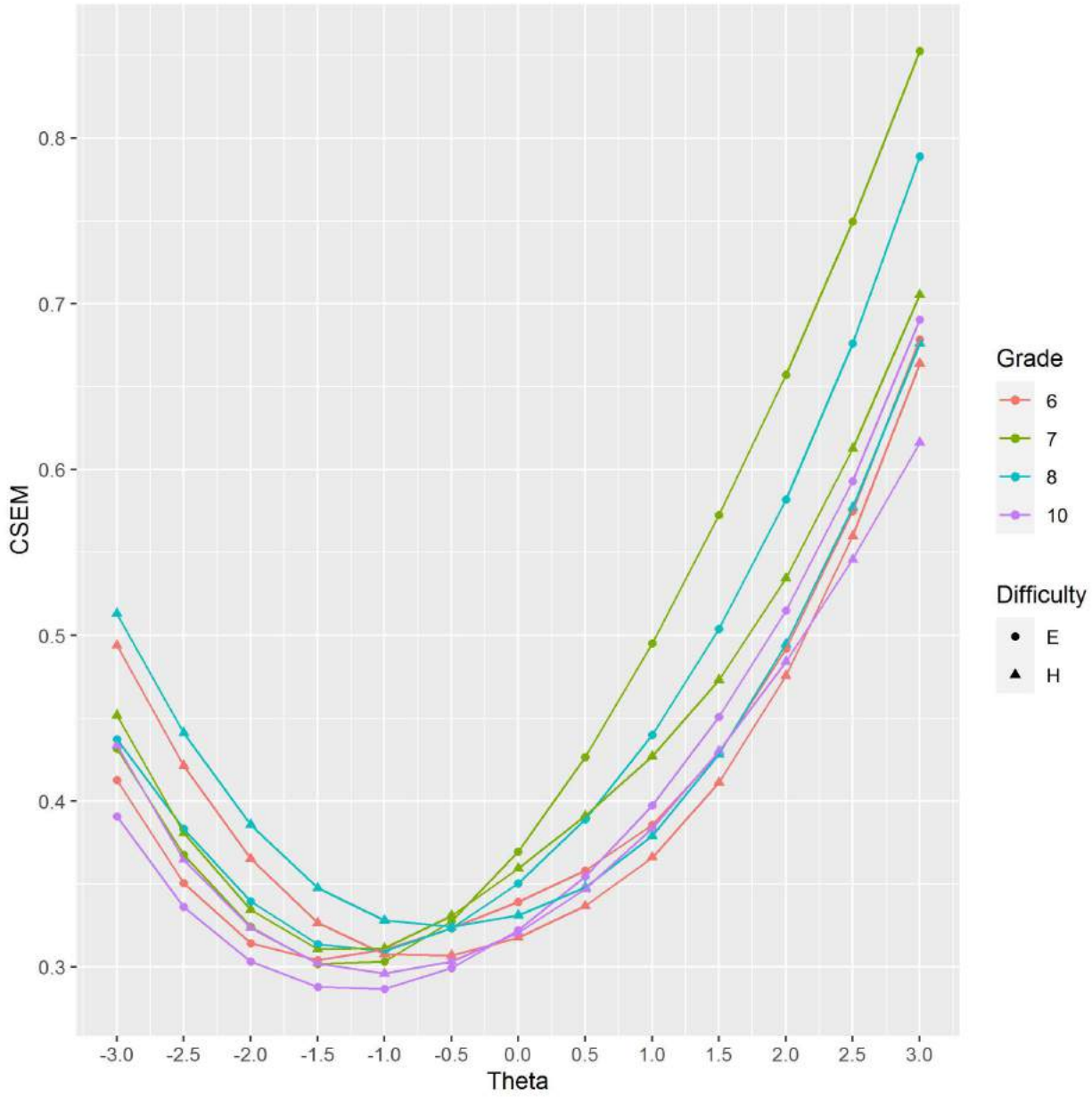
Appendix A: Conditional Standard Error of Measurement

Figure A-1. Conditional Standard Error of Measurement (CSEM) for English Language Arts in Grades 3, 4, and 5



Note. E = easy; H = hard.

Figure A-2. Conditional Standard Error of Measurement (CSEM) for English Language Arts in Grades 6, 7, 8, and 10



Note. E = easy; H = hard.

Figure A-3. Conditional Standard Error of Measurement (CSEM) for Mathematics

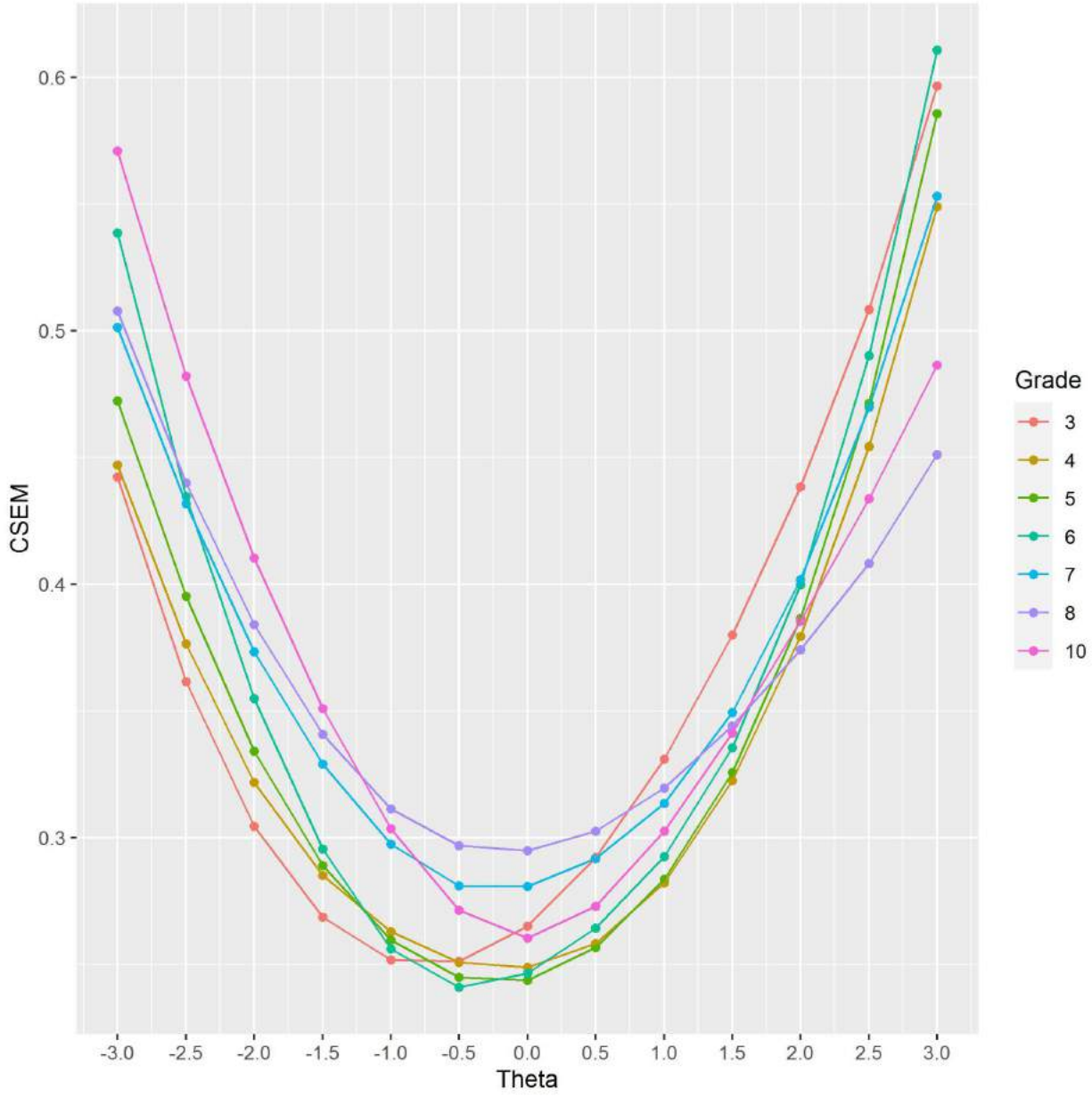
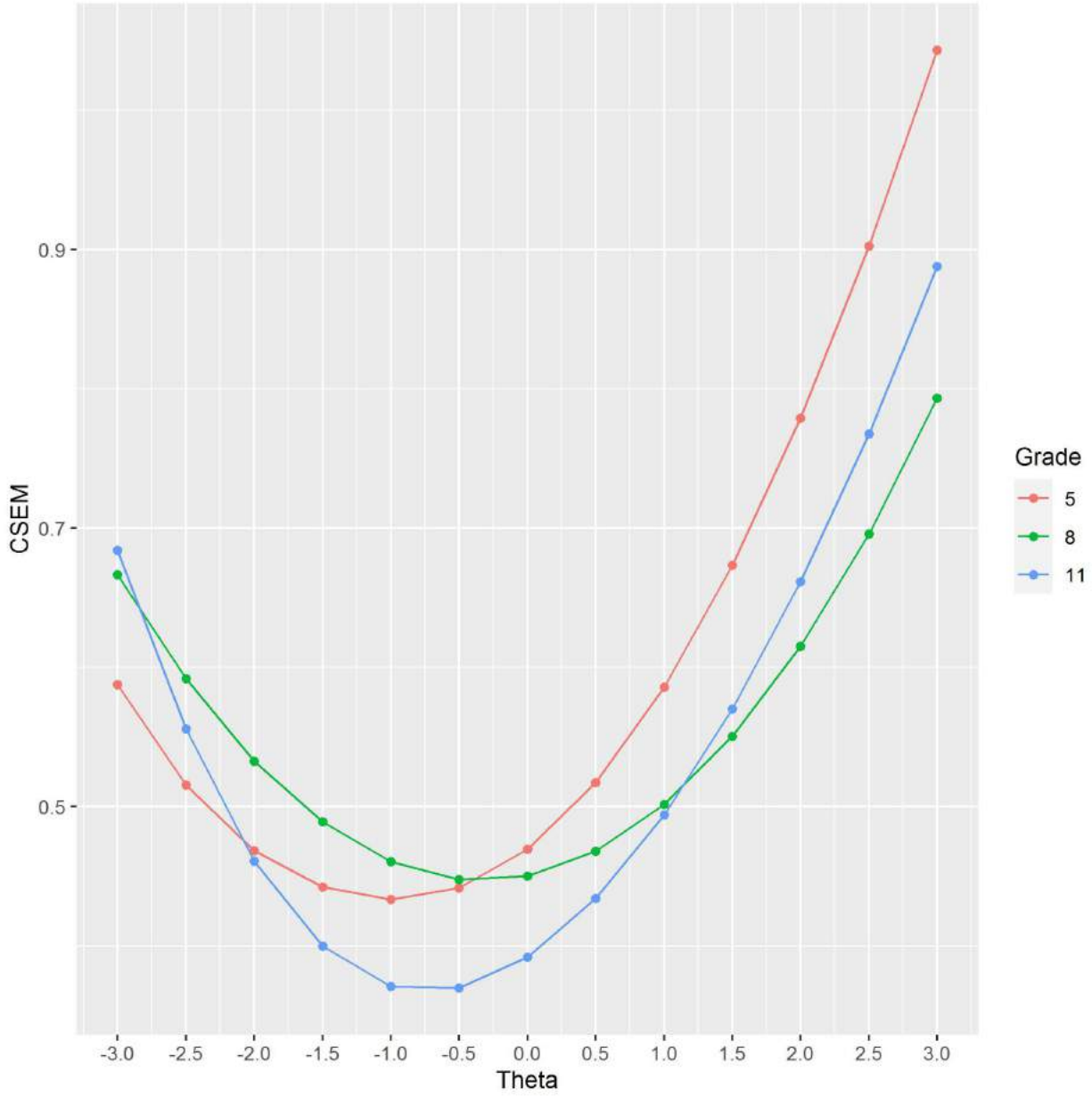


Figure A-4. Conditional Standard Error of Measurement (CSEM) for Science



Appendix B: Subscore Reliability

Table B.1. English Language Arts Grade 3-5 Subscore, Reliability, Classification Consistency, and Accuracy by Grade

Grade	Subscore name	Reliability	Consistency	Accuracy
3	Overall Reading	.73	.45	.77
3	Reading: Literary Texts	.70	.45	.77
3	Reading: Informational Texts	.64	.37	.73
3	Reading: Making and Supporting Conclusions	.63	.33	.71
3	Reading: Main Idea	.61	.32	.69
3	Overall Writing	.70	.37	.71
3	Writing: Revising	.57	.30	.66
3	Writing: Vocabulary and Language Use	.60	.29	.63
3	Writing: Editing	.64	.31	.65
4	Overall Reading	.69	.40	.73
4	Reading: Literary Texts	.62	.30	.67
4	Reading: Informational Texts	.63	.34	.70
4	Reading: Making and Supporting Conclusions	.64	.35	.69
4	Reading: Main Idea	.52	.22	.62
4	Overall Writing	.66	.34	.66
4	Writing: Revising	.54	.29	.63
4	Writing: Vocabulary and Language Use	.55	.29	.68
4	Writing: Editing	.61	.37	.70
5	Overall Reading	.70	.40	.74
5	Reading: Literary Texts	.66	.34	.70
5	Reading: Informational Texts	.61	.33	.69
5	Reading: Making and Supporting Conclusions	.61	.33	.68
5	Reading: Main Idea	.62	.36	.71
5	Overall Writing	.66	.37	.69
5	Writing: Revising	.51	.27	.64
5	Writing: Vocabulary and Language Use	.58	.32	.64
5	Writing: Editing	.60	.34	.67

Table B.2. English Language Arts Grade 6-10 Subscore, Reliability, Classification Consistency, and Accuracy by Grade

Grade	Subscore name	Reliability	Consistency	Accuracy
6	Overall Reading	.74	.39	.76
6	Reading: Literary Texts	.67	.39	.75
6	Reading: Informational Texts	.68	.39	.76
6	Reading: Making and Supporting Conclusions	.65	.37	.73
6	Reading: Main Idea	.62	.35	.72
6	Overall Writing	.59	.30	.68
6	Writing: Revising	.49	.22	.62
6	Writing: Vocabulary and Language Use	.57	.30	.63
6	Writing: Editing	.54	.32	.71
7	Overall Reading	.74	.45	.79
7	Reading: Literary Texts	.70	.43	.79
7	Reading: Informational Texts	.69	.43	.77
7	Reading: Making and Supporting Conclusions	.70	.47	.80
7	Reading: Main Idea	.60	.34	.76
7	Overall Writing	.61	.35	.71
7	Writing: Revising	.53	.25	.69
7	Writing: Vocabulary and Language Use	.51	.29	.71
7	Writing: Editing	.58	.30	.70
8	Overall Reading	.67	.36	.83
8	Reading: Literary Texts	.62	.36	.84
8	Reading: Informational Texts	.60	.30	.76
8	Reading: Making and Supporting Conclusions	.58	.33	.78
8	Reading: Main Idea	.56	.33	.78
8	Overall Writing	.62	.32	.72
8	Writing: Revising	.53	.31	.70
8	Writing: Vocabulary and Language Use	.59	.38	.73
8	Writing: Editing	.51	.32	.75
10	Overall Reading	.73	.42	.83
10	Reading: Literary Texts	.68	.39	.78
10	Reading: Informational Texts	.69	.45	.81
10	Reading: Making and Supporting Conclusions	.66	.43	.78
10	Reading: Main Idea	.63	.38	.80
10	Overall Writing	.65	.37	.73
10	Writing: Revising	.59	.32	.69
10	Writing: Vocabulary and Language Use	.56	.33	.70
10	Writing: Editing	.58	.37	.71

Table B.3. Mathematics Grade 3-5 Subscore, Reliability, Classification Consistency, and Accuracy by Grade

Grade	Subscore name	Reliability	Consistency	Accuracy
3	Overall Concepts and Procedures	.75	.47	.76
3	Concepts and Procedures: Operations and Algebraic Thinking	.68	.38	.70
3	Concepts and Procedures: Measurement and Data	.62	.32	.68
3	Problem Solving	.59	.32	.64
3	Communicating Reasoning	.55	.29	.64
3	Modeling and Data Analysis	.51	.24	.61
4	Overall Concepts and Procedures	.77	.50	.81
4	Concepts and Procedures: Operations and Algebraic Thinking	.70	.36	.74
4	Concepts and Procedures: Number and Operations in Base Ten	.61	.40	.79
4	Concepts and Procedures: Number and Operations with Fractions	.67	.42	.75
4	Concepts and Procedures: Measurement and Data	.54	.27	.70
4	Problem Solving	.49	.22	.62
4	Communicating Reasoning	.62	.40	.78
4	Modeling and Data Analysis	.57	.23	.65
5	Overall Concepts and Procedures	.75	.48	.82
5	Concepts and Procedures: Number and Operations in Base Ten	.67	.40	.77
5	Concepts and Procedures: Number and Operations with Fractions	.61	.33	.76
5	Concepts and Procedures: Measurement and Data	.63	.40	.75
5	Problem Solving	.55	.32	.73
5	Communicating Reasoning	.55	.30	.75
5	Modeling and Data Analysis	.54	.30	.71

Table B.4. Mathematics Grade 6-10 Subscore, Reliability, Classification Consistency, and Accuracy by Grade

Grade	Subscore name	Reliability	Consistency	Accuracy
6	Overall Concepts and Procedures	.74	.52	.84
6	Concepts and Procedures: The Number System	.66	.38	.79
6	Concepts and Procedures: Expressions and Equations	.63	.38	.82
6	Problem Solving	.48	.29	.69
6	Communicating Reasoning	.48	.10	.51
6	Modeling and Data Analysis	.50	.26	.71
7	Overall Concepts and Procedures	.70	.45	.84
7	Concepts and Procedures: Expressions and Equations	.56	.38	.81
7	Concepts and Procedures: Statistics and Probability	.58	.36	.79
7	Problem Solving	.48	.20	.64
7	Communicating Reasoning	.46	.24	.67
7	Modeling and Data Analysis	.59	.32	.73
8	Overall Concepts and Procedures	.71	.50	.88
8	Concepts and Procedures: Geometry	.66	.44	.83
8	Concepts and Procedures: Expressions and Equations	.61	.41	.83
8	Problem Solving	.58	.28	.63
8	Communicating Reasoning	.48	.22	.71
8	Modeling and Data Analysis	.59	.27	.63
10	Overall Concepts and Procedures	.72	.51	.88
10	Concepts and Procedures: Geometry	.53	.37	.79
10	Concepts and Procedures: Algebra	.69	.52	.89
10	Communicating Reasoning	.49	.26	.66
10	Modeling and Data Analysis	.41	.11	.49

Table B.5. Science Subscore, Reliability, Classification Consistency, and Accuracy by Grade

Grade	Subscore name	Reliability	Consistency	Accuracy
5	Physical and Chemical Sciences	.57	.34	.69
5	Life Sciences	.59	.29	.64
5	Earth and Space Sciences	.62	.33	.68
8	Physical and Chemical Sciences	.59	.32	.76
8	Life Sciences	.58	.29	.70
8	Earth and Space Sciences	.58	.32	.72
11	Physical and Chemical Sciences	.52	.31	.70
11	Life Sciences	.63	.35	.72
11	Earth and Space Sciences	.57	.31	.71

Appendix C: School Board of Education District Demographic Distribution

Table B.1. Number of Students Enrolled and Their Demographic Distribution by State Board of Education District

District	*N*	%													
		Gender		Race						Hispanic		SWD		EL	
		Female	Male	AI	Asian	Black	NHPI	Other	White	No	Yes	No	Yes	No	Yes
1	50,791	49	51	2	4	13	0	7	73	74	26	88	12	83	17
2	52,803	49	51	1	6	7	0	6	79	84	16	90	10	90	10
3	61,443	49	51	1	6	6	0	6	81	86	14	89	11	92	8
4	31,972	49	51	2	1	7	0	10	79	85	15	84	16	94	6
5	32,620	49	51	7	1	2	0	4	86	59	41	87	13	75	25
6	32,060	49	51	2	1	5	1	7	84	89	11	84	16	97	3
7	66,662	49	51	2	3	12	0	8	75	76	24	85	15	89	11
8	39,680	49	51	3	5	16	0	9	67	71	29	86	14	84	16
9	36,389	49	51	2	1	2	0	7	88	92	8	84	16	97	3
10	63,584	49	51	2	3	11	0	8	76	78	22	85	15	89	11

Note. AI = American Indian; NHPI = Native Hawaiian and Pacific Islander; EL = English learner; SWD = student with disability.

Appendix D: Sample KAP Reports

Figure C.1. Sample KAP Student Report

STUDENT REPORT: Last1150155, First1150155
GRADE: 5 English Language Arts / STATE ID: 1150155
SCHOOL: Argonia Elem
DISTRICT: Argonia Public Schools / #D0359

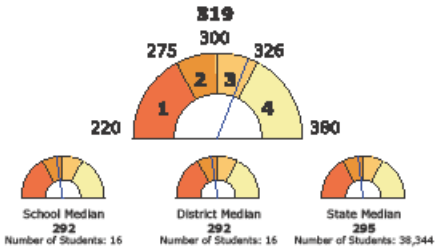
2020-2021 KAP

The KAP assessments measure students' understanding of the Kansas College and Career Ready Standards at each grade. The English language arts assessment asks students to read and answer questions about literary passages, informational texts, and writing samples. Students demonstrate their knowledge and skills related to reading and writing by selecting the right answer and sorting, matching, labeling, and ordering information.

English Language Arts Score: Level 3

In 2019 your student performed at Level 3.

A single test score does not provide a complete or precise measure of student achievement. When interpreting KAP results, please take into consideration other measures of student achievement. Also, consider how the conditions for learning, which may have been disrupted by the pandemic, may influence performance.



Students who score at this level can typically
- read and understand moderately complex grade-level texts
- summarize themes
- identify implied or clear details to support an idea
- determine meanings of more difficult words and complex figurative language
- identify literary elements and text structures and their impact on meaning
- determine point of view or purpose
- revise or edit a text to use academic language and correct grammar, punctuation, and spelling
- organize a text using sequence and logic
- determine if information is relevant
- use strategies to elaborate on ideas and structure texts

Standard error of measurement for this report:
Student — 10.4 | School — 8.1 | District — 8.1 | State — 0.2
The standard error indicates how much a student's score might vary if the student took many equivalent versions of the test (tests with different items but covering the same knowledge and skills).

Performance Level Descriptions

Level 1: A student at Level 1 shows a limited ability to understand and use the English language arts skills and knowledge needed for college and career readiness.

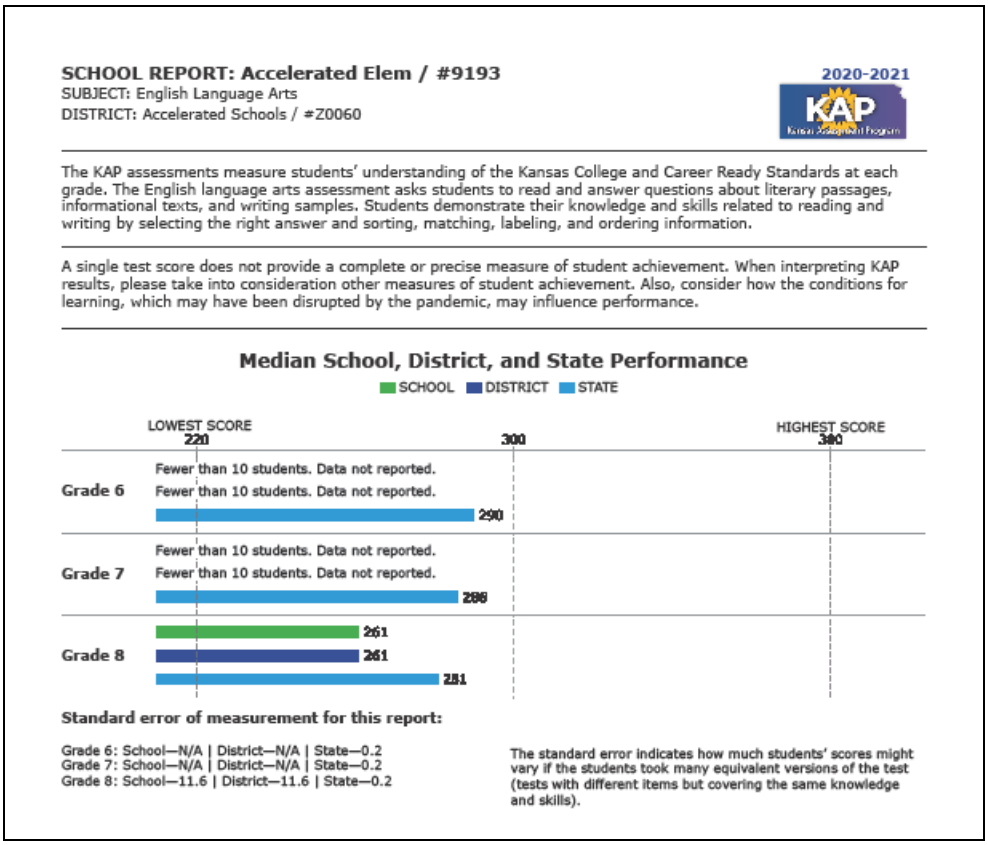
Level 2: A student at Level 2 shows a basic ability to understand and use the English language arts skills and knowledge needed for college and career readiness.

Level 3: A student at Level 3 shows an effective ability to understand and use the English language arts skills and knowledge needed for college and career readiness.

Level 4: A student at Level 4 shows an excellent ability to understand and use the English language arts skills and knowledge needed for college and career readiness.

For more details about how your student performed on specific types of test questions, see the back of this report. →

Figure C.2. Sample KAP School Report



SCHOOL REPORT: Accelerated Elem / #9193
SUBJECT: English Language Arts
DISTRICT: Accelerated Schools / #Z0060

2020-2021

The KAP assessments measure students' understanding of the Kansas College and Career Ready Standards at each grade. The English language arts assessment asks students to read and answer questions about literary passages, informational texts, and writing samples. Students demonstrate their knowledge and skills related to reading and writing by selecting the right answer and sorting, matching, labeling, and ordering information.

A single test score does not provide a complete or precise measure of student achievement. When interpreting KAP results, please take into consideration other measures of student achievement. Also, consider how the conditions for learning, which may have been disrupted by the pandemic, may influence performance.

Median School, District, and State Performance

SCHOOL DISTRICT STATE

LOWEST SCORE 220 | 300 | HIGHEST SCORE 380

Grade 6
Fewer than 10 students. Data not reported.
Fewer than 10 students. Data not reported.
290

Grade 7
Fewer than 10 students. Data not reported.
Fewer than 10 students. Data not reported.
286

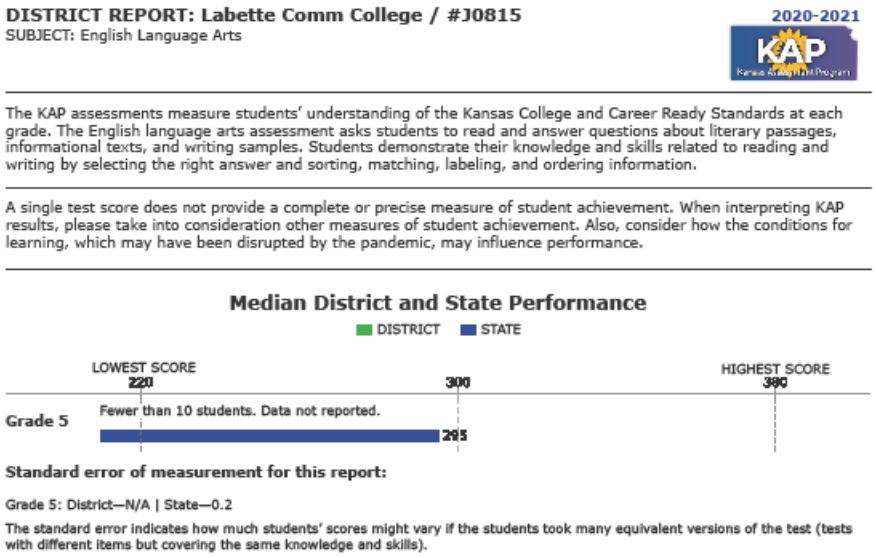
Grade 8
261
261
251

Standard error of measurement for this report:

Grade 6: School—N/A | District—N/A | State—0.2
Grade 7: School—N/A | District—N/A | State—0.2
Grade 8: School—11.6 | District—11.6 | State—0.2

The standard error indicates how much students' scores might vary if the students took many equivalent versions of the test (tests with different items but covering the same knowledge and skills).

Figure C.3. Sample KAP District Report



DISTRICT REPORT: Labette Comm College / #J0815
SUBJECT: English Language Arts

2020-2021

The KAP assessments measure students' understanding of the Kansas College and Career Ready Standards at each grade. The English language arts assessment asks students to read and answer questions about literary passages, informational texts, and writing samples. Students demonstrate their knowledge and skills related to reading and writing by selecting the right answer and sorting, matching, labeling, and ordering information.

A single test score does not provide a complete or precise measure of student achievement. When interpreting KAP results, please take into consideration other measures of student achievement. Also, consider how the conditions for learning, which may have been disrupted by the pandemic, may influence performance.

Median District and State Performance

DISTRICT STATE

LOWEST SCORE 220 | 300 | HIGHEST SCORE 380

Grade 5 — Fewer than 10 students. Data not reported.

Standard error of measurement for this report:

Grade 5: District—N/A | State—0.2

The standard error indicates how much students' scores might vary if the students took many equivalent versions of the test (tests with different items but covering the same knowledge and skills).