



**Kansas Assessment Program
Technical Manual
2020**

University of Kansas Achievement & Assessment Institute (AAI)

October 2020

Table of Contents

I.	Statewide System of Standards and Assessments.....	1
I.1.	State Adoption of Academic Content Standards for All Students.....	2
I.2.	Coherent and Rigorous Academic Content Standards	2
I.2.1.	Goals of Kansas Standards	3
I.2.2.	Process and Timeline.....	3
I.2.3.	Standards Review Committees.....	4
II.	Assessment System Operations	5
II.1.	Assessment Framework of the Assessed Grades	5
II.2.	Test Design and Development	6
II.2.1.	Test Blueprints	7
II.2.2.	Test Design.....	8
II.2.3.	Operational Test Construction	9
II.2.4.	Item Pool Evaluation.....	9
II.2.4.1.	Item Realignment.	9
II.2.4.2.	Oklahoma Shared Items.	9
II.3.	Item Development.....	10
II.3.1.	ELA Passage Selection and Review.....	11
II.3.2.	ELA Item Writing	12
II.3.2.1.	ELA Item Writers.....	12
II.3.2.2.	ELA Item-Writing Training.	13
II.3.2.2.1.	General Guidelines.	13
II.3.2.2.2.	Content Guidelines.	13
II.3.2.2.3.	Format Guidelines.	14
II.3.2.2.4.	Structure Guidelines.	14
II.3.2.2.5.	Stem Construction Guidelines.	14
II.3.2.2.6.	Answer-Choice Development Guidelines.	14
II.3.2.2.7.	Accessibility Guidelines.	14
II.3.2.2.8.	Bias and Sensitivity Guidelines.	15
II.3.2.3.	ELA Item-Writing Process.....	15

II.3.3. Item Review	15
II.3.3.1. ELA External Item Reviewers.	16
II.3.3.2. External Item-Review Training.....	16
II.3.3.3. Item Content-Review Process.	16
II.3.3.4. Item Bias and Sensitivity Review Process.	17
II.3.3.5. Psychometric Review.....	17
II.3.3.6. Accessibility Review.....	17
II.3.4. Field Testing.....	18
II.3.5. Field-Test Data Analysis.....	19
II.4. Test Administration.....	19
II.5. Monitoring Test Administration.....	19
II.6. Test Security.....	19
III. Technical Quality: Validity.....	20
III.1. Overall Validity, Including Validity Based on Content.....	20
III.1.1. Content Validity.....	20
IV. Technical Quality: Other.....	22
IV.1. Multiple Assessment Forms	22
IV.1.1. Cross-Year Linking Design	22
IV.2. Technical Analysis and Ongoing Maintenance	23
V. Inclusion of All Students	24
VI. Reporting.....	25
References.....	26
Appendix A. Content Framework by Grade.....	27
Appendix B. Test Blueprint by Grade	29
Appendix C. Item Statistics Flagging Criteria.....	31

Table of Tables

Table II-1. ELA Content Framework.....	5
Table II-2. Mathematics Content Framework.....	6
Table II-3. Development Timeline for KAP Assessments	6
Table II-4. Test Blueprint by Subject and Content Category for ELA and Mathematics	8
Table II-5. Fixed-Form Test Design of the KAP Assessment by Subject and Session	8
Table II-6. Number of Field-Test Items by Category, Subject, and Grade	19
Table III-1. Percentage of Items by Depth of Knowledge (DOK) Level, Subject, and Grade.....	21
Table A-1. ELA Content Framework by Grade.....	27
Table A-2. Mathematics Content Framework by Grade.....	28
Table B-1. ELA Test Blueprint by Grade and Content Category	29
Table B-2. Mathematics Test Blueprint by Grade and Content Category	30
Table C-1. Item Statistics Flagging Criteria	31

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 2019–2020* (hereafter the *Examiner's Manual*; Kansas State Department of Education [KSDE], 2020), 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 student 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 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 current technical manual provides updates, where applicable, for the 2019–2020 school year, including a description of the 2017 ELA and mathematics content standards as well as the item development and realignment activities that occurred in preparation for the planned spring 2020 administration.

Important Note on the COVID-19 Pandemic

This technical manual update provides updates made to the KAP summative assessments in preparation for the spring 2020 administration; however, the spring administration was canceled due to the COVID-19 pandemic. On March 17, 2020, after consultation with KSDE, the State Board, Kansas Association of School Boards, Kansas School Superintendents Association, United School Administrators, and Kansas National Education Association, Governor Kelly issued an executive order to close Kansas schools, thereby canceling all plans for statewide assessment administration. On March 20, 2020, the US Secretary of Education invited states to submit a waiver for reporting on 2020 state assessment results for the purposes of meeting federal accountability requirements; Kansas subsequently applied for and received the waiver. As a result of canceled testing, no results or updated analyses that rely on spring 2020 assessment results are provided in this technical manual.

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 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. The planned 2020 KAP ELA and mathematics assessments reflected the updated 2017 Kansas standards.

For science, the State Board adopted the Kansas standards in 2013. The first administration of the operational KAP science assessments aligned with the 2013 Kansas standards occurred in 2017. In 2018, Kansas science standards review committee reviewed 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. The mission of Kansas State Board of Education is 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.

1.2.1. Goals of Kansas Standards

The 2017 Kansas standards for ELA are built upon a foundation of common understandings, or practices, that provide a comprehensive view of broad goals for English language arts and literacy instruction for each student across the state. There are five foundational practices:

1. Write, speak, read, and listen appropriately in all disciplines.
2. Seek out and work to understand diverse perspectives.
3. Use knowledge gained from literacy experiences to solve problems.
4. Create multimodal versions of texts for a range of purposes and audiences.
5. Self-regulate and monitor growth in writing, speaking, reading, and listening.

The 2017 Kansas standards for mathematics were created to define what students should understand and be able to do in their study of mathematics. Mathematical understanding is the ability to justify, in a way appropriate to the student’s mathematical maturity, why a particular mathematical statement is true or where a mathematical rule comes from. The student who can explain the rule understands the mathematics, and may have a better chance to succeed at a less familiar task. Mathematical understanding and procedural skill are equally important, and both are assessable using mathematical tasks of sufficient richness.

The 2013 Kansas standards for science closely align to the Next Generation Science Standards (NGSS). The NGSS are based on the *Framework for K–12 Science Education* developed in 2012 by the National Research Council of the National Academies. However, the intent of the NGSS is to put the *Framework* into practice by coupling the practice with content, providing performance expectations while leaving curricular and instructional decisions to states and educators, and evaluating students on the degree of understanding of a full discipline core idea.

1.2.2. Process and Timeline

The Kansas ELA standards review committee met regularly to review and edit the previous 2010 standards document. Once the committee completed its task, the updated standards were presented to the State Board in October 2017. Then there was a window for public comments and review of the updated standards. The committee then presented them for adoption to the State Board in November 2017, and the State Board adopted them later that month.

The previous Kansas mathematics standards were reviewed, written, and edited by the Kansas mathematics standards-writing-and-review committee between March 2016 and May 2017. Minutes of these meetings were kept to explain the decisions that were made (KSDE, 2017). The committee presented the updated standards to the State Board in July 2017 for adoption, and then a window for public comments and review of the updated standards. In August 2017, the State Board approved the adoption (KSDE, 2017).

The 2013 Kansas science standards were reviewed in 2018 by the Kansas science standards review committee. After reviewing, the committee concluded that no updates to the 2013 Kansas science standards were needed. The detailed process and timeline of the development of the 2013 Kansas science standards can be found in the [*2017 KAP Technical Manual*](#).

1.2.3. Standards Review Committees

In an effort to ensure that educators from across the state had an opportunity to nominate either themselves or someone else to serve on the standards review committees, information about the formation of the committees was distributed to the education community through email distribution lists, meetings, and the State Board. Nominations were collected via a registration site that recorded the nominee’s demographic information, committee group of interest, years of work experience, and highest level of education. KSDE staff ensured that the standards review committees consisted of diverse genders, races, ethnicities, and teaching levels (K–12 and postsecondary) and that every state district was represented. Each committee also included an ad hoc group that consisted of representatives from various educational organizations, business communities, and KSDE, as well as legislators, parents, and other community members. Although the ad hoc group members participated in discussions during the standards review process, they did not provide an official vote on the final product that was subsequently reviewed and adopted by the State Board (KSDE, 2017).

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 but must be reasonable 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 2017 Kansas standards for ELA and mathematics are grouped by domain and cluster. Clusters are the sublevels of domains. An item is aligned to only one domain and one cluster. Tables II-1 and II-2 show the 2017 Kansas standards assessment framework for ELA and mathematics. ELA has three domains: reading for information, reading for literature, and writing. Mathematics has two domains: skills and concepts, and strategic thinking and reasoning. All domains are tested in each grade’s assessment, but not all clusters are tested in each grade’s assessment. Tables II-1 and II-2 include a comprehensive framework across grades. Appendix A includes the grade-specific assessment framework for ELA and mathematics. The framework for of science standards can be found in the [2017 KAP Technical Manual](#).

Table II-1. ELA Content Framework

Domain	Cluster
Reading—information	Key ideas and details
	Craft and structure
	Language in reading
Reading—literature	Key ideas and details
	Craft and structure
	Language in reading
Writing	Text types and purposes
	Language in writing

Table II-2. Mathematics Content Framework

Domain	Cluster
Skills and concepts	Operations and algebraic thinking Number and operations in base ten Numbers and operations with fractions Measurement and data Ratios and proportional relationships The number system Expressions and equations Algebra Functions Geometry Statistics and probability
Strategic thinking and reasoning	Problem-solving and modeling Communicating reasoning

II.2. Test Design and Development

Because the 2017 Kansas standards are only updated standards, the Achievement and Assessment Institute (AAI) worked with KSDE to determine the content to be assessed by the KAP tests for each subject area and grade level with the consideration of the continuity with the previous standards. The development leading to the 2020 KAP test administration occurred over multiple years. Table II-3 outlines the test-development timeline for ELA, mathematics, and science.

Table II-3. Development Timeline for KAP Assessments

Milestone	Date	Note
ELA		
Adoption of 2010 Kansas standards	October 2010	
First operational administration aligned to 2010 Kansas standards	Spring 2015	
Standard setting	Summer 2015	
Adoption of 2017 Kansas standards	November 2017	
2017 Kansas standards item development and realignment	2017 to 2020	
Field testing of 2017 Kansas standards items	Spring 2018 to spring 2019	Items are not included in scoring.

Milestone	Date	Note
2017 Kansas standards items included in summative assessment	Spring 2020	All operational items are aligned to 2017 Kansas standards.
Mathematics		
Adoption of 2010 Kansas standards	October 2010	
First operational administration aligned to 2010 Kansas standards	Spring 2015	
Standard setting	Summer 2015	
Adoption of 2017 Kansas standards	August 2017	
2017 Kansas standards item development and realignment	2017 to 2020	
Field testing of 2017 Kansas standards items	Spring 2018 to spring 2019	Items are not included in scoring.
2017 Kansas standards items included in the summative assessment	Spring 2020	All operational items are aligned to 2017 Kansas standards
Science		
Adoption of Kansas standards	June 2013	
Kansas standards item development	2015 to 2016	Determined annually
Census field testing	Spring 2016	Machine-scored items only
Operational testing	Spring 2017 to spring 2020	Machine-scored items only
Standard setting	Summer 2017	
Review of Kansas standards	2018	No updates to the 2013 Kansas standards

II.2.1. Test Blueprints

Table II-4 summarizes the range of the proportion of items required for each domain in the test blueprints across grade for ELA and mathematics. The ELA percentages vary slightly across grades. Mathematics has same domain proportions across grades except grade 10. Appendix B includes the grade-specific test blueprints for ELA and mathematics.

For the 2017 Kansas ELA standards, the test blueprints are the same as the previous standards. For the 2017 Kansas mathematics standards, there are slight blueprint adjustments within domain levels to reflect the updated Kansas mathematics standards for grades 3–8. The previous blueprints included 25%–35% of items measuring strategic thinking and reasoning, and the current blueprints include 15% of items in that category. The strategic thinking and reasoning items require students to use problem-solving and modeling strategies and to communicate their reasoning. These items involve analyzing complex mathematical and real-world problems, using problem-solving strategies and mathematical models to interpret and solve problems, constructing arguments to support the reasoning used, and critiquing the reasoning of others. The blueprint of science tests could be found in the [2017 KAP Technical Manual](#).

Table II-4. Test Blueprint by Subject and Content Category for ELA and Mathematics

Subject and content category	Percentage of items by category
ELA	
Reading—information	30%–36%
Reading—literature	28%–34%
Writing	36%–38%
Mathematics	
Skills and concepts	71%–85%
Strategic thinking and reasoning	15%–29%

II.2.2. Test Design

In 2019, the mathematics test design was changed from a two-stage adaptive design to a fixed-form design. In 2020, the ELA test design was also changed to a fixed-form design from a two-stage adaptive design. This fixed-form test design is the same test design used when cut scores were set. Thus, in 2020 all three subjects use a fixed-form test design. Each subject has one operational form administered in two sessions. Each session offers several blocks of items that are the same but presented in different orders to deter cheating. Students are randomly assigned to one block in each session, and each session has a designated block of items for students who need accommodations. Table II-5 shows the test design of the KAP assessment for each session by subject.

Table II-5. Fixed-Form Test Design of the KAP Assessment by Subject and Session

Subject	Grade	No. of items		
		Total	Session 1	Session 2
ELA	3–8, HS	47	22	25
Mathematics	3–8, HS	55	25	30
Science	5	35	18	17
	8, HS	40	20	20

Note. HS = high school.

II.2.3. Operational Test Construction

The test construction in 2020 followed the same procedures and guidelines as in previous years. The detailed description of test-construction procedures and guidelines can be found in the [2017 KAP Technical Manual](#). However, with ELA and mathematics changing from adaptive to fixed test design, the psychometric review process has new procedures to make sure a test includes a wide range of item difficulties, has a moderate level of average item difficulty, and provides maximum information about the proficiency-level cut scores. Then, to improve students' testing experiences and to reduce test fatigue for mathematics, the same calculator-usage criteria used in 2019 were applied for 2020 mathematics forms: all items are calculator-inactive in grades 3–5, and each test session begins with calculator-inactive items, followed by calculator-active items in grades 6–8 and 10.

II.2.4. Item Pool Evaluation

To ensure the item pools reflected the 2017 Kansas standards and provided adequate coverage of the updated standards, two activities occurred: an item realignment process and operational item sharing with another state assessment program. This section discusses these two activities.

II.2.4.1. Item Realignment. The existing item pools in ELA and mathematics were reviewed, and items were realigned to the 2017 content standards. There were minimal to moderate differences between the previous standards and the 2017 updated standards. Therefore, the purpose of the realignment activity was to determine which existing items accurately aligned to the 2017 content standards without revising the item content. AAI content staff evaluated the content of the items and analyzed the intended standard and the elicited evidence of students' knowledge, skills, and processes within the items, as described in the content standards. Items that did not align to the updated standards were removed from the item pool. Some items were realigned to a different grade according to the updated standards. The realigned items were associated to domains and clusters based on the frameworks for the 2017 standards. All items appearing on operational 2019–2020 assessments were aligned to the 2017 content standards for ELA and mathematics.

II.2.4.2. Oklahoma Shared Items. In an effort to broaden the pool of quality items from which operational forms are constructed, mathematics and science items from the Oklahoma State Department of Education (OSDE) were used for field testing in Kansas, under contractual agreement between OSDE and the KU Center for Technology Commercialization. Using specific criteria, internal reviewers decided whether each item was acceptable for field testing in Kansas. The criteria for acceptance included content, accessibility, and bias and sensitivity considerations:

- The item strongly aligns to content standards, including the level of cognitive complexity designated by the standards and item specifications.
- The item appropriately measures the intended construct, as defined by the content standards. It is grounded within grade-appropriate contexts and uses grade-appropriate vocabulary.

- The content of the item is technically correct (wording, data, and graphics).
- The item’s answer options contain only one correct answer (i.e., the key). The distractors are incorrect and not misleading, and nothing within the item cues the correct response.
- The text uses clear language and the graphics are clear. Neither text nor graphics introduces unnecessary confusion nor distraction.
- Items portray all groups accurately and fairly, demonstrating awareness of different cultures and sensitive topics and avoiding stereotypical depictions of human subject matter.

For item blocks (science only), the criteria for acceptance are:

- All of the criteria for items are met, and, where applicable, the criteria for the stimulus are also met.
- The information presented within the stimulus is needed to correctly answer the associated items.
- To correctly answer the item, students must interact with the stimulus by applying the concepts and skills targeted by the aligned content standards, instead of relying on prior knowledge alone.
- The item’s task is clear: the pathway from the stimulus to the item stem and to the answer options is the same for students of different skill levels. However, if divergent pathways do exist, the pathways are relevant to the targeted content standards.
- The order in which the item appears within the item block represents a logical position of the item within the block. The location of the item within the item block contributes to scaffolding of skills and concepts.
- Nothing in the item cues the correct response of any other item within the item block.

At least two AAI reviewers examined each item and item block to determine its acceptability for field testing. When the conclusions of the two reviewers differed, a third reviewer examined the item or item block to provide a tiebreaking conclusion.

All accepted items and item blocks were then internally reviewed and/or updated by the editing and graphics teams to ensure all content was free of grammar, punctuation, and spelling errors, as well as to ensure all content complied with KAP’s style guidelines before field testing. [Section II.3.4 Field Testing](#) presented the number of developed ELA items for field testing.

II.3. Item Development

Item development entails various efforts to ensure item quality, including ongoing research into best practices and new item types, developing and using subject-area item specifications, updating materials for item-writer training, recruiting new or additional item writers, conducting item-writer training for new item writers or a refresher training for continuing item writers, creating items, and reviewing and revising items. Item review is conducted in two phases: first, when items are created and next, after items are field tested. In the first phase, both AAI content experts and trained, external item reviewers review items. Before appearing on any assessment,

items are reviewed by content reviewers, bias and sensitivity reviewers, and KSDE staff. AAI staff use item-review feedback to revise test items as needed. Items are then prepared for field testing, according to test specifications and established guidelines for both general and accommodated presentation. After field testing, AAI content experts and psychometricians analyze the item and test data; this data analysis guides decisions about the use of items on future assessments.

During 2019–2020, AAI conducted the ELA passage selection and review and planned to conduct the psychometric review for the ELA items developed in November 2018. No new mathematics or science items were developed because the priority for those subjects was identifying and reviewing items from the Oklahoma item banks. New ELA development allowed for the creation of items written specifically to the 2017 standards. These items were developed to broaden the item pool to ensure content coverage of future forms. Mathematics and science items on 2020 forms were developed previously using the processes and guidelines described in the [2017 KAP Technical Manual](#). Therefore, the sections that follow describe the procedures for 2019–2020 ELA item development. Review protocols, however, are the same for all three subject areas.

II.3.1. ELA Passage Selection and Review

For ELA, the passage-selection process started with identifying appropriate public-domain works or commissioning passages as work-for-hire in the early summer of 2019. KAP’s passage-development team has built a strong network of both regional and national authors, allowing the team to generate high-quality, original passages capable of supporting item development.

Assessment passages included commissioned, permissioned, and public-domain readings. The ELA team used several resources, both qualitative and quantitative, to analyze text complexity and to guide grade-level placement. Passages from all sources underwent multiple rounds of internal review, including editorial, content, bias, sensitivity, and accessibility reviews. For example, AAI content experts and accessibility specialists reviewed passages for accessibility issues and content accuracy (e.g., inaccurate or outdated science information). Outside experts with knowledge of low-incidence disabilities or with specific subject-matter expertise were also consulted as needed.

Passages accepted during internal review then went to a panel of Kansas educators for external review in the summer of 2019. These panels were formed by grade band: grades 3–5, 6–8, and high school. Each panel included educators with backgrounds in teaching English learners and in special education. Reviewers received training and detailed instructions regarding how to review passages through KAP’s secure, online reviewing system. Panelists then reviewed passages at their own pace and provided feedback and placement recommendations by a given deadline.

During both internal and external review, passage reviewers used rubrics of both qualitative and quantitative measures to examine text complexity and grade-level suitability through text structure, language features, and knowledge demands. KAP used the Flesch–Kincaid and Lexile scores as quantitative measures for longer passages. However, passages of only 350–450 words

were not long enough to give an accurate Flesch–Kincaid reading, and most measures of text complexity were inadequate for the task of analyzing poetry. In those cases, KAP considered Lexile score (for prose), sentence length, and complexity to gauge passages’ initial grade placement. Qualitatively, KAP looked at each set for vocabulary, knowledge demands, topic familiarity, and interest level.

In addition, both internal and external reviewers considered several passage components.

- **Length:** Are the texts of reasonable length for students? Are the texts long enough and rich enough to support all or most of the item content established in the item and test specifications?
- **Bias or sensitivity:** Are all groups portrayed accurately and fairly? Does the passage demonstrate awareness of different cultures and sensitive topics in the state (e.g., natural disasters, politics)?
- **Overexposure:** Is the passage already commonly taught in the school or district, or is it used frequently in anthologies or lesson plans?
- **Interest level:** Will more than half of students be at least moderately interested in the passage?
- **Images:** Are there any concerns related to the accessibility or content of images? Should images be added to enhance or support the passage?
- **Prior knowledge:** Should introductions be included to provide historical context or background information?

Reviewers then recommended a grade level for each passage according to complexity and other considerations. After compiling the information and summarizing the overall data collected from the review, the passages and grade-level recommendations were shared with KSDE for approval in October 2019. Of the 39 passages sent to KSDE, 28 were approved. Based on item pool needs (e.g., complexity levels, text types, topics), some passages were selected for item development; remaining passages were held for future development.

II.3.2. ELA Item Writing

The development of new items for ELA assessments happened at an in-person item-writing event in November 2018. The recruited item writers had participated in the item-writing training before the event. During the item-writing event, the items also underwent peer review.

II.3.2.1. ELA Item Writers. Twenty-one educators from across the state were invited to participate in the item-writing event. As much as possible, these educators represented public and private schools, small and large districts, rural and urban districts, educators currently in the classroom, educators currently working at the district level, and educators working for KSDE. Among these 21 educators, 90.5% were female and 9.5% were male; 66.7% were teachers, 4.8% were test coordinators, and 28.6% were instructional coaches, reading specialists, or retired/substitute teachers; 85.7% were from public schools, and 14.3% were not working with a specific school but were instructional coaches, reading specialists, directors of curriculum and instruction, or retired/substitute teachers. AAI ELA content staff assigned educators to one of

seven teams, covering all assessed grade levels. Educators could write individually or as a team, and all educators received peer feedback during the event, allowing writers to revise their own work.

II.3.2.2. ELA Item-Writing Training. Before writing items for the KAP assessment, item writers were trained in the use of KAP subject-area item specifications in the writing and reviewing of items. All item writers received training in several topics, including

- alignment to the 2017 Kansas ELA standards
- evidence-centered design
- principles of Universal Design for Learning (UDL) and accessibility
- bias and sensitivity
- item types available for ELA

To guide the item-writing process, item writers were trained in content, format, structure, stem structure, answer-choice development, accessibility, bias and sensitivity, and traditional and nontraditional item types. Also, item-writer training taught participants about UDL concepts, including a definition of UDL and examples of test items that adhere to UDL principles. Additionally, the item-writer guidelines include many UDL principles. Besides learning fundamental principles of item writing, item writers also received training in item review so they can objectively evaluate their own and their peers' products. Key points of these guidelines, as suggested by Haladyna and Downing (1989), are presented below.

II.3.2.2.1. General Guidelines.

- Write items that have clearly correct answer choice(s), with other answer choices clearly incorrect.
- Ensure that items are clearly worded.
- Avoid the use of tricky or misleading items.
- Proofread items for correct grammar, punctuation, and spelling.
- Avoid the use of contractions.
- Use third-person perspective.
- Avoid the use of humor.

II.3.2.2.2. Content Guidelines.

- Write items to appropriate content standards.
- Ensure that multiple-choice items measure a single concept.
- Ensure that items focus on important ideas, not trivia.
- Use vocabulary that is consistent with students' grade level.
- Align items to the cognitive complexity of content standards.
- Write items to a variety of difficulty levels.

II.3.2.2.3. Format Guidelines.

- Format answer choices vertically rather than horizontally.
- Write clear instructions related to formatting.

II.3.2.2.4. Structure Guidelines.

- Avoid complex-format items.
- Write items in the form of a question.
- Avoid window-dressing of items (e.g., excessive verbiage).

II.3.2.2.5. Stem Construction Guidelines.

- Write stems positively whenever possible.
- Avoid asking for and expressing opinions in stems.
- Ensure that the central idea is in the stem.
- Place the question as close to the answer choices as possible.
- Minimize the use of qualifying words (e.g., “best,” “most likely”).

II.3.2.2.6. Answer-Choice Development Guidelines.

- Order answer choices logically.
- Create independent answer choices that do not overlap.
- Write answer choices that are roughly of the same length and parallel in structure.
- Do not offer “all of the above,” “none of the above,” or “I don’t know” as answer choices.
- Avoid cluing between the stem and answer choices.
- Avoid specific determiners, such as “always” or “never.”
- Create plausible distractors.
- Create distractors that take advantage of common errors and misconceptions.
- The order of answer keys should be roughly uniform in distribution.

II.3.2.2.7. Accessibility Guidelines.

- Consider the access needs of special populations and the ways in which accommodations affect an item’s intent.
- Use simple sentence structures.
- Minimize the use of words with multiple meanings.
- Avoid the use of slang and regional dialect.
- Avoid the use of complicated names or names that could be confused with other nouns.
- Clearly label graphics.

II.3.2.2.8. Bias and Sensitivity Guidelines.

- Avoid the use of stereotypes.
- Consider the regional and cultural nuances of words.
- Avoid the use of demeaning or offensive materials, particularly in the stimulus.
- Avoid the use of religious references, such as holidays.
- Ensure that items are not related to socioeconomic status or family attributes.
- Use artwork that reflects the diversity of the student population.

II.3.2.3. ELA Item-Writing Process. AAI’s ELA content staff supplied each grade with multiple reading passages, including informational, literary prose, and poetry texts. Writers chose the passages and created approximately 16–18 items per passage per grade to account for attrition during the review process. Writers were asked to review their chosen passage and determine which standards the content authentically addresses. They were then tasked with writing items that assessed each standard addressed by the passage at least once.

After item writers completed a full set of items, they passed the items to another item writer for peer review. The reviewer checked each item’s alignment to the standards and cognitive complexity demands and can discuss the item with the writer. The item writer then revised items as needed. The items were then passed to the content lead for further review.

II.3.3. Item Review

The item-review process involves several stages:

- internal content and editorial review
- external review (content, bias and sensitivity), using multiple panelists
- internal content-team resolution
- psychometric review
- accessibility review

The content lead performs the internal content review. After the content lead’s review, the items go through editorial review. For items needing graphics, the content specialist provides the graphic artist with instructions for rendering the stimulus and then confirms that the completed graphic meets the intended function within the item. After the editors finish editing the items, the content lead reviews the items again before external review. If substantial changes are made to an item during review, the content lead returns it to the editing team. Only after the completion of internal content and editorial reviews do the items go to external reviewers. This section describes the reviewers and review process for two external reviews: content review and bias and sensitivity review. Both external reviews occurred in February 2019 for newly developed ELA items. After external review, any items that educators in external review made recommendation for editing are edited internally. AAI might consult with KSDE about the internal editing as needed. Then, items are fielded tested and student response data is used for various analysis. The results of field-test data analysis were reviewed by both the psychometrician and the content lead during the psychometric review. The psychometric review for newly developed ELA items were

planned in summer 2020. For items passing the psychometric review and placed on the operational assessment, accessibility reviews are conducted for adherence to principles of UDL and issues in accessing the item that may be encountered by students with disabilities or students who are English learners. Psychometric and accessibility reviews are discussed in Section II.3.3.5 and Section II.3.3.6, respectively.

II.3.3.1. External Item Reviewers. AAI content experts and KSDE staff recruited item reviewers from Kansas educators for two separate types of reviews: content review and bias and sensitivity review. Prospective item reviewers completed an online survey in which they indicated their demographic information, teaching experience, professional qualifications, content expertise, experience with the standards, and endorsements or training in special education or teaching ELs. As with the item writers, the educator reviewers represent, as much as possible, public and private schools, small and large districts, rural and urban districts, educators currently in the classroom, and educators currently working at the district level. Twenty-three out of 30 educator reviewers completed the ELA external item-review process in February 2019. Among these 23 educators, 100% were female; 69.6% were teachers, 13% were test coordinators, and 17.4% were instructional coaches, reading specialists, or retired/substitute teachers; 82.6% were from public schools, 4.3% were from private schools, and 8.7% were not working with a specific school but were either instructional coaches, reading specialists, directors of curriculum and instruction, or retired/substitute teachers.

AAI staff assigned ELA content reviewers and bias and sensitivity reviewers to a specific grade according to the educators' experience. Like the ELA passage-review process, item reviews occurred through a secure, online reviewing system. After completing a web-based training session, reviewers evaluated items at their own pace and provided feedback by a given deadline.

II.3.3.2. External Item-Review Training. All item reviewers have to complete two web-based sessions of item-review training: one for the online review system and one specialized training for either bias and sensitivity training or content-review training. The training sessions include information about the KSDE–AAI partnership, test and item security, item-writing guidelines, and the item-review process. Item-review training also provide participants with practice items and AAI staff contact information.

II.3.3.3. Item Content-Review Process. Content reviewers look at every aspect of each item. They verify alignment to the content standards; judge the item's appropriateness, including its content, context, and vocabulary for the grade and subject; check the correct answer and evaluate the incorrect answers; evaluate the need for any included graphic or stimulus and comment on its utility and clarity; and identify possible concerns about accessibility. Content reviewers also attend to the alignment of items to assessment targets, checking that items adequately address part of the standard and elicit evidence for at least part of one evidence statement. In general, content reviewers check items for

- appropriate, grade-level vocabulary
- a clear, complete statement or question

- grammatically correct text
- a correct key
- accurate, relevant graphics
- well-designed answer choices that do not require background knowledge outside of the content area and that are free from clang associations. (Clang occurs when words from an item’s stem appear in one or more response options.)

Based on their analysis, reviewers could recommend that items be accepted, revised, or rejected, and they give specific reasons for their decisions (e.g., “item aligns better to this assessment target”).

II.3.3.4. Item Bias and Sensitivity Review Process. Bias and sensitivity reviewers are asked to identify barriers that might prevent students from demonstrating what they know and are able to do when those barriers are not related to the content standards. These barriers may include unfamiliar language; linguistic complexity; potentially sensitive topics; presentation of stereotypes, including emotions, regions, or occupations; accessibility for special populations; and issues with cultural or prior knowledge.

Concerns may include

- possible bias related to gender, race or ethnicity, socioeconomic factors, or other
- possible barrier related to uncommon or unfamiliar language, linguistic complexity or lack of clarity, assumed prior knowledge, cultural restrictions, accessibility, or other
- possible sensitivity concern related to stereotype, religion, socioeconomic factors, status, specific topic, or other
- other concerns

II.3.3.5. Psychometric Review. After field-test item analyses and before test construction, both psychometricians and content leads review item statistics. Items with statistical flags are used only when the item pool does not have other items for blueprint coverage; item statistical flagging criteria are included in Appendix C. When flagged items needed to be used as operational items, they undergo extra review and discussions.

II.3.3.6. Accessibility Review. After content leads and psychometricians identify items for form construction according to blueprint and psychometric specifications, an accessibility expert, along with the internal test-development team, incorporate accessibility features to ensure that the widest range of students can access the items. The accessibility work that occurs is based on knowledge in low-incidence disabilities, blind/low vision, deaf/hard of hearing, or English language learning. Every item that has not previously appeared on an accessible version of a form undergoes this review before being used on an operational form.

Accessibility features that are incorporated into items include

- verification and use of accessible color palettes
- verification and use of appropriate color contrast settings

- the presence of alternative text on images
- the ability to allow users to navigate using only a keyboard
- compatibility with commonly used assistive technology products, such as screen readers
- braille
- key word translations: translation of key science and mathematics words into Spanish
- American Sign Language (ASL) videos: uploaded videos of a translator who converts science and mathematics items into ASL
- text-to-speech

II.3.4. Field Testing

In ELA and mathematics, field testing uses the embedded-model approach; science, meanwhile, appends field-test items at the end of each session. Field-test items fall into several categories to serve different purposes. For grades 3–10 in ELA, field-test items are newly developed items passing the external review, for use in future assessments. For grades 3–10 in ELA and grades 3–8 in mathematics, there are items from predictive interim assessments¹ that will be used for future predictive interim assessments. For all grades in mathematics and science, there are items selected from the Oklahoma item bank that are field tested to obtain item statistics from the Kansas student population. For all grades in mathematics and science, there are also existing KAP items. Some of those existing KAP mathematics items included items that needed to be field tested again to ensure the field-test data reflected the condition of the resource sheet being available because a grade-specific resource sheet with conversions and formulas was available for grades 4–8 and grade 10 beginning in 2019. The rest of those existing KAP mathematics items that needed to be field tested again were realigned to a different grade after the 2017 standards were adopted. Some existing KAP science items needed to be field-tested again because of minor content changes after psychometric review of the operational field testing. The purpose of field testing Oklahoma items and existing KAP items for mathematics and science is to broaden the item pool and ensure content coverage of future forms. Table II-6 displays the number of field-test items by category, subject, and grade.

¹ Predictive interim assessments are one component of the Kansas Assessment Program and are available for grades 3–8 and 10 in ELA and mathematics. Three separate, predictive interim assessments can be administered for ELA and mathematics, for six total assessments. These assessments are available during two-week testing windows in the fall, winter, and spring of each year.

Table II-6. Number of Field-Test Items by Category, Subject, and Grade

Grade	ELA			Mathematics		Science	
	New items	Interim items	OK items	Existing items	Interim items	OK items	Existing items
3	57	8	36	22	41	—	—
4	48	16	40	19	41	—	—
5	56	8	31	26	43	67	4
6	59	16	54	6	40	—	—
7	52	24	50	10	40	—	—
8	53	24	53	27	20	41	3
10	55	16	28	72	—	—	—
11	—	—	—	—	—	67	3

Note. OK items = Oklahoma items.

II.3.5. Field-Test Data Analysis

Field-test item analyses include classical item analysis, item response theory (IRT) calibration, model-fit evaluation, and differential item functioning analysis. Items that are too easy or too difficult, that do not discriminate students' ability well, or that have large differential item functioning are flagged according to predetermined criteria. The statistics with flagging will be used in psychometric review and test construction.

Spring 2020 testing was canceled because of COVID-19, so 2020 KAP field-test items could not be administered and results are not available.

II.4. Test Administration

The planned 2020 KAP testing window was from March 16 through May 1, 2020. Because of COVID-19, KAP testing was canceled.

II.5. Monitoring Test Administration

Since spring 2020 testing was canceled, test administration could not be monitored. For information about standard procedures and protocols related to test-administration monitoring, see Section II.4.2 of the [2017 KAP Technical Manual](#).

II.6. Test Security

For information about standard procedures and protocols related to test security, see Section II.5 of the [2017 KAP Technical Manual](#).

III. Technical Quality: Validity

The next section contains relevant updates for the 2019–2020 school year. Sections III.2 through III.5 of the current manual were affected by COVID-19. For a complete description of these sections², see the original, full version of the [2017 KAP Technical Manual](#); for the most recent results associated with these sections, see the [2019 KAP Technical Manual Addendum](#).

III.1. Overall Validity, Including Validity Based on Content

Because the intended uses of the test scores are one source of evidence in a validity study, the purposes of the test should be identified before providing evidence to support test validity. The purposes of the KAP assessment, described at the beginning of this manual, include (a) measuring specific claims related to the Kansas standards in grades 3–8 and high school, (b) reporting individual student scores along with each student’s performance level, and (c) providing 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.

Evidence gathered on content validity, alignment, cognitive process, and internal structure supports the use of the KAP assessment to measure the Kansas standards content as defined in the test blueprints. Information on test reliability, fairness and accessibility, and scoring and scaling justifies the use of KAP test scores for accountability purposes and reporting students’ academic performances.

III.1.1. Content Validity

Evidence of content validity for the KAP assessments comes from the alignment between KAP items and the Kansas standards and from the congruence between the test and the test blueprint. Chapters I and II of this current technical manual and the [2017 KAP Technical Manual](#) present validity evidence related to item development, the item alignment and realignment process, and the correspondence between the test and test blueprint. As stated in the [2017 KAP Technical Manual](#), blueprints also do not specify the proportion of depth of knowledge (DOK) levels, i.e., cognitive complexity levels (Webb, 1997), required for the assessment. However, item-specification documents provide information on the expected DOK for each assessment target. Item writers use this information to write items that match the DOK expectation of each assessment target. The analyses of DOK distributions of 2020 operational items by subject and grade in Table III-1 provide the content validity evidence that the tests have good measurement characteristics across the range of examinee proficiency. Most ELA items are at level 1 and level 2; fewer items are at level 3. In mathematics, most items are at level 1 and level 2 as well, with

² These sections are: Section III.2. Validity Based on Cognitive Process, Section III.3. Validity Based on Internal Structure, Section III.4. Validity Based Relationships to Other Variables, and Section III.5. Survey Results for Validity Evidence.

even fewer level 3 items. For science, most items are at levels 2 and 3, with few or no items at level 1 across grades.

Table III-1. Percentage of Items by Depth of Knowledge (DOK) Level, Subject, and Grade

Grade	ELA				Mathematics				Science			
	Total items	DOK level, %			Total items	DOK level, %			Total items	DOK level, %		
		1	2	3		1	2	3		1	2	3
3	47	26	60	15	55	62	38	0	—	—	—	—
4	47	21	70	9	55	51	47	2	—	—	—	—
5	47	26	60	15	55	64	36	0	35	0	71	29
6	47	30	47	23	55	55	45	0	—	—	—	—
7	47	11	83	6	55	60	38	2	—	—	—	—
8	47	23	72	4	55	51	44	5	40	5	43	53
10	47	2	91	6	55	42	51	7	—	—	—	—
11	—	—	—	—	—	—	—	—	40	3	55	40

IV. Technical Quality: Other

This chapter includes two sections (i.e., Section IV.1. Multiple Assessment Forms and Section IV.2. Technical Analysis and Ongoing Maintenance) that have relevant updates for the 2019–2020 academic year. The other sections were either affected by COVID-19 (i.e., Sections IV.1, IV.3, and IV.4 in the [2017 KAP Technical Manual](#)) or did not have any planned updates for the 2019–2020 academic year (i.e., Section IV.2 in the [2017 KAP Technical Manual](#)). For a complete description of these sections³, see the original, full version of the [2017 KAP Technical Manual](#); for the most recent results, see the [2019 KAP Technical Manual Addendum](#).

IV.1. Multiple Assessment Forms

In large-scale assessment programs, different item sets are used on test forms both within and across years. Linking the scores from these different test forms puts the form scores on a common scale and ensures that all forms for a given grade level and subject area provide comparable scores. This outcome means that students will not have an unfair advantage or disadvantage simply because they took an easier or harder test form than other students did.

To maintain consistency in reporting, the same item-IRT scales used in previous years' testing were planned for use in 2020 for ELA and mathematics because the updates to the 2017 standards were minimal. The same item-IRT scales used in previous years' testing were planned for use in 2020 for science because there were no standard updates. Because all three subjects have one operational form per grade, within-year linking was not needed. Thus, only the cross-year linking was needed for 2020.

IV.1.1. Cross-Year Linking Design

To increase the number of linking items and maximize linking stability, the cross-year linking uses the preequating method. The IRT parameters of all items on the 2020 ELA, mathematics, and science tests were calibrated in previous years, and these parameters are on the same IRT scale as items in 2015 tests for ELA and mathematics and in 2017 tests for science. When the items from different years are on the same IRT scale, the student scale scores calculated from these IRT item parameters are equated and placed onto the base scale (i.e., the 2015 IRT scale for ELA and mathematics and the 2017 IRT scale for science).

³ These sections, according to their headings in the original full version of the [2017 KAP Technical Manual](#), are: Section IV.1. Reliability, Section IV.2. Fairness and Accessibility, Section IV.3. Full Performance Continuum, and Section IV.4. Scaling.

IV.2. Technical Analysis and Ongoing Maintenance

From 2018–2019 to 2019–2020, several changes occurred.

- ELA changed from a two-stage, adaptive test design to a fixed-form test design.
- Mathematics items that did not align with the 2017 standards were removed. All remaining operational items aligned to the 2017 standards. Also, a grade-specific resource sheet with conversions and formulas was available for grades 4–8 and grade 10.
- Science items that are simulation based, or did not match the NGSS, were removed without affecting blueprint coverage.

V. Inclusion of All Students

Refer to Chapter V of the [*2017 KAP Technical Manual*](#) for detailed information regarding procedures for including students with disabilities and accommodations for KAP assessments; see Sections V.1 and V.2, respectively. Section V.3 typically includes a summary of the Personal Needs and Preferences profiles submitted by teachers, which determine the availability of online test accommodations for individual students. However, because of COVID-19, the spring 2020 KAP assessments were canceled, so a report of accommodation selections is not available.

VI. Reporting

Because of COVID-19, the scheduled end-of-year KAP assessments were canceled, so score reports are not available. For detailed information about KAP assessment score reporting, see Chapter VI of the [*2017 KAP Technical Manual*](#).

References

- Haladyna, T. M., & Downing, S. M. (1989). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 2(1), 37–50.
https://doi.org/10.1207/s15324818ame0201_3
- Kan. Stat. Ann. §72-5170 (2020).
https://www.ksrevisor.org/statutes/chapters/ch72/072_051_0070.html
- Kansas State Department of Education. (2017). *2017 Kansas mathematics standards: Grades K-12*.
<https://community.ksde.org/LinkClick.aspx?fileticket=Fk5h7Uw24Kw%3d&tabid=5276&mid=15449>
- Kansas State Department of Education. (2020). *Kansas Assessment Examiner's Manual 2019–2020*. https://ksassessments.org/sites/default/files/documents/KAEM_2019-20.pdf
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Council of Chief State School Officers and National Institute for Science Education Research Monograph No. 6) (ED 414305). ERIC.
<https://files.eric.ed.gov/fulltext/ED414305.pdf>

Appendix A. Content Framework by Grade

Table A-1. ELA Content Framework by Grade

Grade	Domain	Cluster
3–10	Reading—information	Key ideas and details Craft and structure Language in reading
	Reading—literature	Key ideas and details Craft and structure Language in reading
	Writing	Text types and purposes Language in writing

Table A-2. Mathematics Content Framework by Grade

Grade	Domain	Cluster
3	Skills and concepts	Operations and algebraic thinking Numbers and operations with fractions Measurement and data Geometry
4	Skills and concepts	Operations and algebraic thinking Number and operations in base ten Numbers and operations with fractions Measurement and data
5	Skills and concepts	Number and operations in base ten Numbers and operations with fractions Measurement and data
6	Skills and concepts	Ratios and proportional relationships The number system Expressions and equations Geometry Statistics and probability
7	Skills and concepts	Ratios and proportional relationships The number system Expressions and equations Geometry Statistics and probability
8	Skills and concepts	Expressions and equations Functions Geometry
10	Skills and concepts	Algebra Functions Geometry
3–10	Strategic thinking and reasoning	Problem-solving and modeling Communicating reasoning

Appendix B. Test Blueprint by Grade

Table B-1. ELA Test Blueprint by Grade and Content Category

Grade and content category	Percentage of items by category
3	
Reading—information	34%
Reading—literature	30%
Writing	36%
4	
Reading—information	36%
Reading—literature	28%
Writing	36%
5	
Reading—information	30%
Reading—literature	34%
Writing	36%
6	
Reading—information	30%
Reading—literature	32%
Writing	38%
7	
Reading—information	32%
Reading—literature	32%
Writing	36%
8	
Reading—information	30%
Reading—literature	32%
Writing	38%
10	
Reading—information	36%
Reading—literature	28%
Writing	36%

Table B-2. Mathematics Test Blueprint by Grade and Content Category

Grade and content category	Percentage of items by category
3	
Skills and concepts	85%
Strategic thinking and reasoning	15%
4	
Skills and concepts	85%
Strategic thinking and reasoning	15%
5	
Skills and concepts	85%
Strategic thinking and reasoning	15%
6	
Skills and concepts	85%
Strategic thinking and reasoning	15%
7	
Skills and concepts	85%
Strategic thinking and reasoning	15%
8	
Skills and concepts	85%
Strategic thinking and reasoning	15%
10	
Skills and concepts	71%
Strategic thinking and reasoning	29%

Appendix C. Item Statistics Flagging Criteria

Table C-1. Item Statistics Flagging Criteria

Statistic	Criteria
Omit	Omit correlation > .1 Omit percentage > .05
Differential item functioning	Gender R ² change > 0.035 Race R ² change > 0.035 EL R ² change > 0.035
Item-total correlation	Item total correlation ≤ .249
<i>p</i> value	<i>p</i> value = 0
Item response theory— discrimination	$a < 0.3$ and $\text{abs}(b1 \dots b10) \leq 5$ $a < 0.3$ and $\text{abs}(b1 \dots b10) > 5$ $0.3 \leq a \leq 0.699$
Item response theory—difficulty	$\text{abs}(b1 \dots b10) > 3.5$
Item-total correlation of keyed response for selecting-key items	Correlation of keyed response < 0
Item-total correlation of distractors for selecting-key items	Correlation of keyed response < 0 and Correlation of distractors > 0 Correlation of distractors > .1 and Correlation of distractors > Correlation of keyed response