



Test Specifications

New Mexico Measures of Student Success and Achievement (NM-MSSA)

English Language Arts



Purpose

- Part of a Balanced Assessment System
- Claims/Score Interpretation and Use Statements

Test Specifications

- Reporting Categories and Test Blueprint
- Cognitive Complexity
- Test Design
- Fairness
- Additional NM-MSSA Considerations

Stimulus Specifications

- Specifications for Reading Stimuli
- Specifications for Language Usage Stimuli
- Specifications for Writing Prompt Stimuli

Item Specifications

- Alignment
- Item Types
- Sample Items

Purpose

Part of a Balanced Assessment System

The New Mexico Measures of Student Success and Achievement (NM-MSSA) is New Mexico's statewide summative assessment for Mathematics and English Language Arts, administered at the end of grades 3-8. As the NM-MSSA is a single measure at the end of a grade band, interpretations and uses of NM-MSSA scores should be supplemented with additional measures, including information from classroom summative and formative assessments in mathematics and English Language Arts as well as interim assessments.

Formative assessment may include the use of the Cognia Formative Assessments, which is a collection of formative assessment materials for grades 3–8 being provided by Cognia during the term of the contract with the state to administer the NM-MSSA. The materials are aligned to the CCSS and therefore to the New Mexico *Common Core State Standards*.

Claims/Score Interpretation and Use Statements

The NM-MSSA is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the New Mexico *Commons Core State Standards*.

In addition to overall ELA scale score, student performance on Reading and Language & Writing and seven subdomains is reported as noted in the reporting categories subsection in the following pages.

See the Score Interpretation and Use (SIU) documents for the complete set of NM-MSSA score interpretation and use statements. These documents can be accessed at https://newmexico.onlinehelp.cognia.org/resources-nm-mssa/.

Test Specifications – Reporting Categories and Test Blueprint

Reporting Categories

The reporting categories for NM-MSSA are based on the clusters of standards found in the New Mexico *Common Core State Standards* for both content and practices as noted below.

Reporting Categories-Reading

Grades 3-8

- Reading
- Text Type
 - Literary Text
 - Informational Text
- Reading Strategy
 - Comprehension
 - Analysis and Interpretation

Reporting Categories-Writing & Language

Grades 3-8

- Writing & Language
- Writing Analysis & Language Conventions
- Production of Writing (rubric score)
- Use of Conventions (rubric score)

Percentages for the distribution of operational (core) test points for each of the reporting categories reflect the distribution in the standards, so as not to over- or underrepresent content. The internal test blueprint has specifications for inclusion on the core forms.

Core Test Blueprints

Specifications for the full test blueprints for the construction of the core forms reflect the reporting category specifications, as well as percentage requirements for each cluster. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational (core) test points for each should be maintained from year to year.

Note that percentages in reading for (a) text type and (b) reading strategy are calculated independently. An individual item may contribute to multiple parts of the blueprint.

All NM-MSSA ELA items are based on passages, either literary or informational (e.g., expository, argument, opinion, memoir, autobiography, etc.). The configuration of texts on the ELA assessment seeks to balance national high-quality assessment guidance (e.g., NAEP, CCSSO, etc.) as well as considerations around test length. The ELA test design incorporates as much of a 50/50 split of literary and informational texts as possible in the elementary grades while still maintaining a limited summative test footprint. Beginning at grade 6, there is a shift in emphasis to informational texts at the upper grade band.

Operational Test Blueprint

	_		Grade	e 3-5	Grade	e 6-8
	Enį	glish Language Arts	Ideal # of Core Pts	Ideal % of Core Pts	Ideal # of Core Pts	Ideal % of Core Pts
	Toyt Typo	Literary Text	15	65%	8	35%
	Text Type	Informational Text	8	35%	15	65%
	Dooding Stratogy	Comprehension	12-14	52-60%	9-12	39-52%
Reading Strategy		Analysis and Interpretation	9-11	39-48%	13-16	56-70%
eac	Cluster	Key Ideas and Details	9-11	39-48%	7-11	30-48%
<u> </u>		Craft and Structure	7-9	30-39%	6-10	26-43%
		Integration of Knowledge and Ideas	4-6	17-26%	4-6	17-26%
		Total	23*	100%*	23*	100%*
	Language & Writing	Writing Analysis	9-11	42-52%	9-11	42-52%
s & age	Passage Sets	English Language Conventions	3-5	14-23%	3-5	14-23%
Writing & Language	iting	Production of Writing	4	19%	4	19%
Wr	Writing Prompt*	Use of Conventions	3	14%	3	14%
		Total	21	100%	21	100%

^{*}All items align to a text type (Literary, Informational), reading strategy (Comprehension, Analysis and Interpretation) and a cluster (Key Ideas and Details, Craft and Structure, and Integration of Knowledge and

Test Specifications – Cognitive Complexity

Percentages for the distribution of operational (core) test points across the cognitive complexity levels (DOK classification) are noted in the table below.

Depth of Knowledge Distribution - Reading

			0		0			
DOK	Grade							
DOK	3	4	5	6	7	8		
Level 1	0-20%	0-20%	0-20%	0-20%	0-20%	0-20%		
Level 2	50-70%	50-70%	50-70%	50-70%	50-70%	50-70%		
Level 3	20-40%	20-40%	20-40%	20-40%	20-40%	20-40%		
Total	100%	100%	100%	100%	100%	100%		

Depth of Knowledge Distribution – Writing & Language

DOK	Grade							
DOK	3	4	5	6	7	8		
Level 1	15-35%	15-35%	15-35%	15-35%	15-35%	15-35%		
Level 2	40-60%	40-60%	40-60%	40-60%	40-60%	40-60%		
Level 3	15-35%	15-35%	15-35%	15-35%	15-35%	15-35%		
Total	100%	100%	100%	100%	100%	100%		

Test Specifications – Test Design

Assessable Standards

The NM-MSSA assesses the New Mexico *Common Core State Standards*. The following standards are eligible to be included in the NM-MSSA.

- Grade 3-5 test: Reading: Literature (RL.1-7, RL.9), Reading: Informational Text (RI.1-RI.9), Writing, and Language standards.
- Grade 6-8 test: Reading: Literature (RL.1-7, RL.9), Reading: Informational Text (RI.1-RI.9), Writing, and Language standards.

Test Design - Student Experience

The *Student Testing Experience* tables on the following pages provide a summary of the number of items and points by item type, usage (i.e., operational items or field test items), and estimated testing time for each grade's NM-MSSA assessment.

As shown in the test design tables, the types of items on the NM-MSSA are 1-point machine-scored items (MS-1), 2-point machine-scored items (MS-2), and 7-point writing prompts (WP). Additional item type descriptions and sample items can be found in the item specifications section on page 16.

Each NM-MSSA test is administered in two sessions. The test form contains both core operational items and matrix field test items. The core operational items are items administered to all students taking that core form, and that count toward a student score. Matrix field test items are items administered to subsets of students to "try out" performance (with different students receiving different field test items), and therefore do not count toward student score.

Student Testing Experience (Full Form)

ELA Grades 3-8	Passage-Based Items			Total	Total Points		
(Spring 2021)	Passage Sets	MS-1	MS-2	WP	Items	Min	Max
Core Operational Items	6	32	6	0	38	44	44
Matrix Operational Items	0	0	0	0	0	0	0
Matrix Field Test Items	2	5	1	1	7	14	14
Total Student Experience	8	37	7	1	45	58	58
				Estimated	Test Time (min)	150 (60/90)

ELA Grades 3-8	F	Passage-Based Items				Total I	Points
(Spring 2022 and beyond)	Passage Sets	MS-1	MS-2	WP	Total Items	Min	Max
Core Operational Items	6	27	5	0	32	37	37
Matrix Operational Items	0	0	0	1	1	7	7
Matrix Field Test Items	2	10	2	0	12	14	14
Total Student Experience	8	37	7	1	45	58	58
				Estimated	Test Time (min)	150 (60/90)

Practice Test

A full-length practice test mirroring the operational test design is available beginning in the 2021-2022 school year. The practice tests and supporting materials can be accessed at https://newmexico.onlinehelp.cognia.org/practice-tests-nm-mssa/.

Test Specifications – Fairness

Fairness is defined as the extent to which the test scores are valid for different groups of test takers. Consideration of universal design, bias, and sensitivity guidelines support the construction of fair, valid assessments.

Universal Design for Assessments

The concept of Universal Design for Assessments focuses on developing content and assessments that reach the widest population of students possible. Stimuli and items on the NM-MSSA are designed to simply and clearly present tasks and to provide maximum readability, comprehensibility, and legibility. The seven elements of Universal Design for Assessments are based on the original UDL guiding principles:

Universal Design for Assessments

Principle	Explanation
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field-testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct-irrelevant cognitive, sensory, emotional, and physical barriers are removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	Test design facilitates the use of needed accommodations (e.g., all items can be translated to braille).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words kept to a minimum) for readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, tables, figures, and illustrations, and to response formats.

The concept of Bias is defined as the presence of some characteristic of an item that results in differential performance for two individuals of the same ability but from different ethnic, sex, cultural, or religious groups.

Bias can occur whenever content offends or disadvantages a student or group of students due to gender, race, regional background, socioeconomic status, or any other such classification.



Test developers take care to craft content in a way that does not misrepresent specific groups or rest on assumptions made about specific groups, that in turn could negatively impact how students interpret content.

- Stimulus and item content on the NM-MSSA must not present stereotypes or unfair representations of gender, race, ethnicity, disability, culture, or religion.
- Stimulus and item content on the NM-MSSA should not depend on overly-experiential information such as knowledge of technology, consumer goods, pop culture, geographic locations, or sports and extracurricular activities. While these topics are not completely excluded from use, care must be taken to ensure that the items are presented in a way that does not require a level of knowledge that would not be held by all students.

Sensitivity

Sensitivity refers to the presence of content that is contrary to the acceptable norms of the students, educators, parents, or other members of the community that may interact with the assessment. Sensitive subject matter can impact student performance or attitudes toward testing, and hence, their test scores.

Consideration of bias and sensitivity issues is very important when developing content for an assessment. Test developers must ensure that stimuli and items are free of content that will negatively affect a student's performance not because of what the student knows and can do but because the content evokes an emotional response from that student (or is in some other way distracting to the student).

Subjects/contexts that are likely to prompt emotional distress on the part of students cannot be used on the NM-MSSA (e.g., war, violence, human death or debilitating disease, animal-based medical research). Careful judgment should be applied to standards that cover topics that may be considered controversial by some groups (e.g., evolution examples, population dynamics including death/extinction, environmental impact). Those standards represent content knowledge to be assessed, but the assessment must be done in a sensitive, unbiased way.

Stimulus Specifications

Reading

All NM-MSSA reading items are based on permissioned passages, either literary or informational (e.g., expository, argument, opinion, memoire, autobiography, etc.). The configuration of authentic texts on the assessment seeks to balance national high-quality assessment guidance (e.g., NAEP, CCSSO, etc.) as well as considerations around test length. For grades 3–5, item sets are based on single literary passages, paired literary passages, and paired informational passages. For grades 6-8, item sets are based on paired literary passages. single informational passages, and paired informational passages. In an effort to reflect a 50/50 split of literary and informational texts in the elementary grades, the assessment at grades 3-5 does contain a balance of paired literary and paired informational texts. Beginning at grade 6, there is a shift in emphasis to informational texts at the upper grade band.

Literary passages, should include experiences with which students are familiar or provide enough context that a student unfamiliar with the experience is not disadvantaged. Passages should be excellent models of exemplary writing in including such literary elements as character development, a well-crafted plot, a text structure that supports the meaning of the passage, and the development of a point of view. Passages should include text structures and literary devices (figurative language, irony, etc.) appropriate to the grade-level being assessed.

Informational passages must include current and accurate information. In addition, care must be taken to avoid topics and details that could quickly become dated (news stories, technologies, discoveries, etc.). Text content should be consistent with the current best thinking in the various fields represented on the test and be neither overly speculative nor highly tentative, given the likelihood of such texts becoming dated over the long course of test development and administration.

Informational passages must be able to stand on their own and should not require any outside knowledge (other than common knowledge) to understand the topic or author's position. Informative/explanatory texts and arguments should be accurate, well-reasoned, and logically organized, reflecting a variety of logical text structures including, but not limited to, compare/contrast, cause/effect, order of importance, sequence/steps in a process, problem/solution, description and explanation, question and answer, and cyclical structures. It should also be noted that domain-specific vocabulary must be easily understood via context clues, minimal footnotes, and/or authorial explanation.

Quantitative measures are certainly helpful in situating a passage within a grade band; they will also help to establish a passage as appropriate (or not) for a particular grade. However, quantitative complexity evaluations must be coupled with thorough qualitative review in order to make an informed grade assignment. The qualitative measures of text complexity are evaluated using the qualitative dimensions of text complexity found in Appendix A of the CCSS, as well as the CCSSO Text Complexity Qualitative Measures Rubrics for Informational and Literary Texts.

The quantitative measures used for NM-MSSA reading passages are shown below. It should be noted that passages may sometimes exceed these specifications if the passage content is deemed appropriate in light of qualitative complexity measures.

Word Count (within the prescribed ranges)

Grade	Passage Type	Word Count
3	Single	300 – 800
3	Paired	600 – 1,000
4	Single	300 – 800
4	Paired	600 – 1,000
5	Single	300 – 800
,	Paired	600 – 1,000
6	Single	500 – 1,000
	Paired	800 – 1,200
7	Single	500 – 1,000
,	Paired	800 – 1,200
8	Single	500 – 1,000
0	Paired	800 – 1,200

Readabilities (Lexile and TextEvaluator also within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Grade Level	TextEvaluator SM
3	310 – 590
4	405 – 655
5	480 – 720
6	550 – 790
7	615 – 860
8	685 – 940

Language Usage

All NM-MSSA language usage passages will be commissioned texts which contain embedded errors. A commissioned passage is an original text written for a specific purpose (e.g., editing task). Passages should demonstrate accurate and engaging expository writing, effective argumentation, and vivid, intelligible narrative writing. The essential elements of each of these genres are derived from the applicable Common Core State Standards in Writing. (e.g., per the standards, argumentation in G7 and G8 should develop a claim in part through engagement with a counterclaim; narratives in all grades should be developed through dialogue, thoughts, and description; etc.)

Expository passages are expected to meet the highest standards of factual accuracy, syntactical and grammatical proficiency, reader engagement, and originality. Generally, expository passages will represent one of three subject areas: Social Studies/History; Science/Social Science/Technical Subjects; and, to a lesser extent, the Humanities. Although written with the general reader in mind, passages nevertheless strive to present compelling information that responds to relevant issues in each field—a new interpretation of an event or phenomenon; an examination of an overlooked (or misunderstood) movement, moment, or figure; an introduction to foundational knowledge in any of the three disciplines, etc.

Passages should assume no content background or expertise on the part of readers, but writers should invest their work with the precision and novelty that rewards attentive reading. This principle extends to technical or discipline-specific language, which should appear where necessary and natural, and never gratuitously or without supporting context. The NM-MSSA writing and language assessment is not intended to evaluate students' reading comprehension abilities. If students struggle simply to read a passage, they will surely be unable to demonstrate a meaningful understanding of how to edit or revise it.

Argument/Opinion passages maintain the same high standards of accuracy, syntax, engagement, and originality expected of expository passages. They also demonstrate cogent argumentation. That is, the writer must establish a position; provide claims, supported by evidence, that develop that position; introduce and rebut a counterclaim (in grade 7 and 8); and, throughout, use rhetorical techniques (persuasive transitions, rhetorical questions, appeals to reason or personal experience, etc.) to advance the position.

Argument/Opinion passages will tend to be informed by issues in the social sciences or current events. Successfully realized topics are not inherently controversial, and the writer's position is

not diffident or universalist—that is, the argument is primarily positive in tone and modest in scope. Nonetheless, the position is clear.

Narrative passages succinctly and lucidly describe a fictional event. Although these passages will feature many or all of the hallmarks of the narrative form—plot/conflict, climax/epiphany, conclusion, dialogue, characters' thoughts, action, description—they must be coherent in spite of their brevity. They avoid long stretches of dialogue; flashbacks and other jarring time-shifts; florid, clichéd, or self-consciously "literary" language; and more than three speaking characters.

Determining the complexity of a text requires a multifaceted approach. In addition to evaluating text based on quantitative data (readability statistics), the writers must also rely on qualitative measures to help determine whether a passage being developed is appropriate for the target audience. The qualitative measures of text complexity are evaluated using the qualitative dimensions of text complexity found in Appendix A of the CCSS, as well as the CCSSO Text Complexity Qualitative Measures Rubrics for Informational and Literary Texts.

The quantitative measures used for NM-MSSA language usage passages are shown below. It should be noted that passages may sometimes exceed these specifications if the passage content is deemed appropriate in light of qualitative complexity measures.

Word Count (within the prescribed ranges)

Grade	Word Count
3	250-400
4	275–425
5	300-450
6	325-550
7	350–625
8	350-625

Readabilities (Lexile and TextEvaluator also within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Grade Level	TextEvaluator SM
3	310 – 590
4	405 – 655
5	480 – 720
6	550 – 790
7	615 – 860
8	685 – 940

Writing Prompts

All NM-MSSA writing prompts will be partnered with one to three brief text stimuli. These may be intact (whole) passages or excerpts (only a part) of a more extended text. Some possible text types include:

- Story
- Memoir
- **Biography**
- Poem
- Article
- Review (book, music, performance, etc.)
- **Editorial**
- Website
- Letter
- Journal (diary, log) entry
- **Instructions**
- Advertisement
- **Brochure**
- Memo

- · Script
- · Transcript
- Display text (e.g., to accompany an artifact in a museum)

The number of text stimuli will vary depending on the purpose for writing. Narrative prompts will be associated with 1-2 text stimuli, while informative/explanatory and opinion/argument prompts will be associated with 2-3 text stimuli. The passages may be either permissioned or commissioned.

Word Count (maximum **total** word count 800-1000)

Grade	Word Count
3	800
4	850
5	900
6	950
7-8	1,000

Readability (Lexile within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Item Specifications

Alignment

The items on the NM-MSSA are aligned to the New Mexico *Common Core State Standards*. Each item is aligned to one of the content standards in NMCCSS.

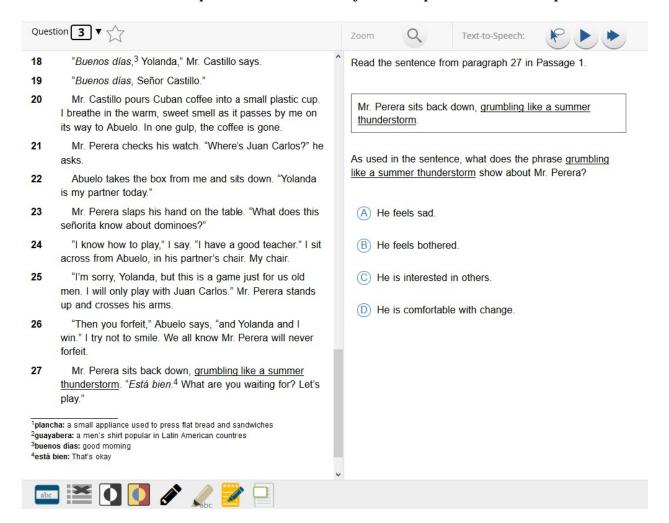
Item Types

The types of items on the NM-MSSA are 1-point machine-scored items (MS-1), 2-point machine-scored items (MS-2), and 7-point writing prompt (WP):

- MS-1 items are worth 1 point. These items may be multiple choice or multiple select.
- MS-2 items are worth 2 points. In ELA, these items are also known as evidence-based selected response items (EBSR). These items have two parts (Part a and Part b) for students to answer, and 0, 1, or 2 points total can be earned across Part a and Part b. Each part provides choices from which to select.
- Writing prompts (WPs) are worth 7 points. These items require students to write an extended response to a single prompt. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The WPs items are evaluated using a "Production of Writing" rubric on a scale from 1–4 and a "Use of Conventions" rubric on a scale from 1–3.

Sample Items

MS-1 items are worth 1 point. These items may be multiple choice or multiple select.





Text-to-Speech:

Read the passage. Then answer the questions that follow.

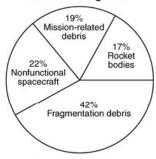
Space Junk

Satellites and space stations aren't the only objects orbiting Earth. According to the National Aeronautics and Space Administration (NASA), there are more than 500,000 pieces of debris in orbit. This "space junk" poses a threat to satellites, spacecraft, and the International Space Station. Even a small piece of space junk the size of a marble can damage them. This could put astronauts' lives at risk.

Where Does Space Junk Come From?

Some space junk comes from natural objects, such as space rocks. It can also come from human-made items, such as parts of satellites. Human-made junk that circles Earth is more accurately known as orbital debris. This debris includes old satellites, rocket launch stages, or fragments from space equipment that was destroyed upon launch.

Debris Orbiting Earth



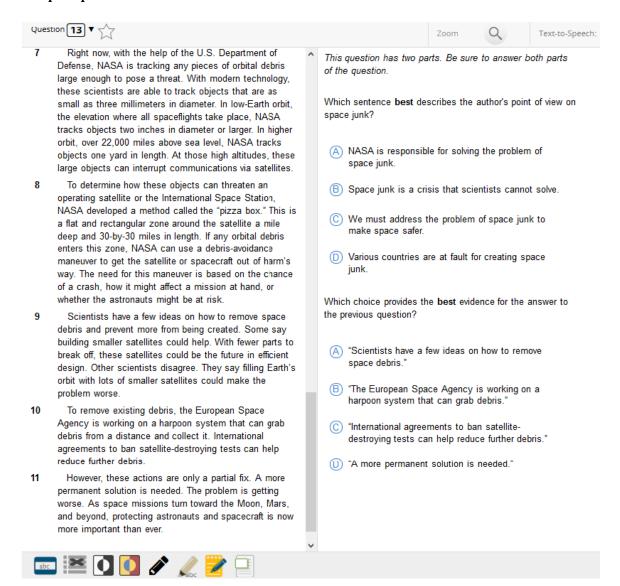
3 As more satellites and spacecraft are launched into orbit, the amount of orbital debris will increase. NASA

💌 🚺 🚺 🖋 🙎 🔀 🖳

How does the diagram support the information in the passage?

- A It explains which type of debris is the most dangerous.
- B It suggests how long debris remains in orbit around
- It represents the composition of human-made debris orbiting Earth.
- It demonstrates which types of human-made debris are the largest.

• MS-2 items are worth 2 points. In ELA, these items are also known as evidence-based selected response items (EBSR). These items have two parts (Part a and Part b) for students to answer, and 0, 1, or 2 points total can be earned across Part a and Part b. Each part provides choices from which to select.





Text-to-Speech:

Read the passage. Then answer the questions that follow.

The History of Skateboarding

- 1. Have you ever ridden a skateboard? 2. A skateboard has a deck with wheels on the bottom for riding and doing tricks.
- 3. About 100 years ago, people used skateboards, and so they were very different from the ones people use now. 4. They were made from wooden boxes and roller skate wheels. 5. Over the years, skateboards have changed. 6. They have become safer and easier to ride.
- 7. In the 1950s, skateboards were made like those made today, with a deck and four wheels. 8. At that time, the wheels were made of clay. 9. They were slippery like ice and unsteady. 10. They did not grip well to sidewalks and other places people skated. 11. In the 1970s, companies changed the materials they used. 12. Now the wheels were made of a material that was more like rubber. 13. More people became interested and took up skateboarding as a hobby. 14. In the 1990s, skateboarding became very popular. 15. People thought up new tricks to show off on skateboards.
- 16. People's love for skateboarding keeps growing today. 17. Many towns and cities have built special parks for skateboarders. 18. There is one more important thing that people who are interested in skateboarding should remember. 19. Skateboards are built better now than before, but people can still fall. 20. For things to wear, they should always wear helmets, knee pads, and gloves. 21. Then, they will be ready to learn and have fun.

"The History of Skateboarding" @ 2020 by Cognia.

This question has two parts. Be sure to answer both parts of the question.

Read this sentence.

20. For things to wear, they should always wear helmets, knee pads, and gloves.

How should the underlined part of sentence 20 be changed to best join the writer's ideas in the last paragraph?

- A For safety
- (B) To follow the rules
- On older skateboards
- (D) For extra power on skateboards

Which words from the passage best support the correct answer above?

- (A) "They were made from wooden boxes"
- (B) "More people became interested"
- C "towns and cities have built special parks"
- (D) "but people can still fall"



















Writing prompts (WPs) are worth 7 points. These items require students to write an
extended response to a single prompt. These items are hand-scored, with scorers using a
multi-trait rubric and scoring notes to evaluate responses. The WPs items are evaluated
using a "Production of Writing" rubric on a scale from 1–4 and a "Use of Conventions"
rubric on a scale from 1-3.

