

Top 10 Best Practices for Developing Multiple-Choice (MC) Items

1	Alignment	2
2	Depth-of-Knowledge (DOK) levels	3
3	Grade-level appropriateness	4
4	Bias and sensitivity issues	5
5	Stems	6
6	Answer options	7
7	Internal and external clueing	9
8	Formatting and style	10
9	Distractor rationales	11
10	Art	12
	Appendix: Depth-of-Knowledge Resources	13

Developed by:

Assessment Solutions & Design team members, with input from Learning Architecture and support from PLU

Do you have questions or suggestions?

Please contact Laura Dershewitz or Angela Senicz.

Do you want some practice with these ideas?

We all learn best by doing! Please contact PLU if you'd like to set up a team workshop to practice and apply these best practices with members of the Assessment Solutions & Design team.

1

Align the item to the required standard, skill, and/or learning objective.

Item alignment, or connection to the standard, skill, or learning objective, is crucial.

Why?

Even a flawlessly constructed item will not provide useful data if it fails to measure what is intended.

Example

Standard: CCSS.RI.6.5: Analyze how a particular sentence, paragraph, chapter, or section fits into the overall structure of a text and contributes to the development of the ideas.

✗ Does Not Align Tightly

What is the MAIN structure of the passage?

- A. sequence of events*
- B. order of importance
- C. problem and solution
- D. comparison and contrast

✓ Aligns Tightly

What is the purpose of paragraph 7?

- A. to show how Hawaii’s volcanic activity has helped scientists better understand volcanoes *
- B. to show how Hawaii’s volcanic eruptions affect those who live on the islands
- C. to explain Kilauea’s importance in forming parts of the Big Island
- D. to explain Kilauea’s unusual appearance to visitors

The standard asks about the role of a particular sentence, paragraph, chapter, or section of a text—not about the overall structure of a text. Therefore, the item does not align tightly to the standard.

This item aligns tightly to the standard because it asks about the role of a particular paragraph.

If a standard, skill, or learning objective has multiple parts, the item need not assess every part at once. It can focus on just one component.

Why?

- MC items that assess multiple concepts can be wordy and hard to follow.
- They can also make it difficult to pinpoint the concept(s) that students are struggling with.

Refer to the item above.

Note that the item does not attempt to focus on both **structure** (how the paragraph “fits into the overall structure”) and **content** (how the paragraph “contributes to the development of the ideas”). It focuses on just one part of the standard (content).

2

Make sure the item has the required depth-of-knowledge (DOK) level.*

The **DOK level** describes the complexity of the cognitive processes demanded by an item. It refers to the type of thinking involved:

- DOK 1: recall and reproduction
- DOK 2: skills and concepts
- DOK 3: strategic thinking
- DOK 4: extended thinking

Keep in mind:

- DOK and difficulty are not the same. An item with a low DOK level is not necessarily easy. For example, this item is DOK 1 but might be difficult to answer: Who was James Polk’s vice president?
- DOK should reflect reasoning skills; it should not be influenced by other factors, such as complex wording.
- Multiple-choice items can assess higher-order thinking skills, not just recall (such as the second example shown).

*Learn more about DOK, including additional examples and verb charts, in the appendix, which begins on p. 13.

Example

<p>Standard: CCSS.RL.8.2: Determine a theme or central idea of a text and analyze its development over the course of the text, including its relationship to the characters, setting, and plot; provide an objective summary of the text.</p>	
<p>✓ DOK 1 or 2</p> <p>What is the BEST summary of paragraph 2?</p> <p>A. Ingrid helps her dad with yard work.</p> <p>B. Ingrid grows tomatoes for her experiment.*</p> <p>C. Ingrid asks her dad about growing tomatoes.</p> <p>D. Ingrid decides what she will do for the science fair.</p>	<p>✓ DOK 2 or 3</p> <p>In “Diego’s Bistro,” what lesson does Diego learn?</p> <p>A. He should trust his boss’s advice.</p> <p>B. He should share his opinions more often.</p> <p>C. It is important to keep trying after a failure.</p> <p>D. It is important to recognize other people’s skills.*</p>
<p>This item focuses on the “summary” part of the standard. It would likely be considered a DOK 1 because it involves basic comprehension—though a summary item can be a DOK 2 if it requires students to make decisions (e.g., which details are important enough to include).</p>	<p>This item focuses on the “theme” part of the standard. It could be considered a DOK 2 or 3 because it involves analysis and abstract thought. (Determining an item’s DOK level sometimes involves subjectivity.)</p>

3

Make sure the item is appropriate for the grade level.

The **vocabulary** and **syntax** should not be too difficult for the targeted audience.

Why?

If they are, then the item measures students' ability to comprehend the item, not their understanding of the standard, skill, or learning objective.

Keep in mind:

- Individual words used in an item should be on grade level or below. HMH typically uses *EDL* to make this determination. If a secondary reference is needed, HMH typically uses *Children's Writer's Word Book*.
- Exceptions may be made for discipline-specific words defined in a passage; targeted vocabulary words; tiered academic vocabulary; language that appears in the standard, skill, or learning objective; etc.

Example

Standard: RL.3.1 (grade 3) Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text as the basis for the answers.

X Not appropriate for the grade level

Which word BEST tells what kind of passage this is?

- A. fiction
- B. folktale*
- C. fantasy
- D. biography

This item includes multiple words that are above grade level in *EDL*.

If students get it wrong, we won't know if it's because they struggled with the content or with reading and understanding the item.

✓ Appropriate for the grade level

What kind of story is "Snow Cookies"?

- A. a story like real life
- B. a story that has magic
- C. a story about an adventure
- D. a story with talking animals*

All of the words in this item are grade 3 or below in *EDL*.

The **item content** should be grade-level appropriate as well.

Keep in mind:

- Item content should be comprehensible to the targeted audience. It should not rely on background knowledge that test takers wouldn't have.
- Items should not aim to teach students something new. They should adhere to the content and grade-level limits of the standard being assessed.

4

Make sure the item is free of bias/sensitivity issues.

All items must be fair to all students.

Bias: Items should not include language or content that favors some students over others. Students of the same ability level should, on average, be able to answer an item equally well regardless of their membership in certain gender, racial, ethnic, cultural, disability, linguistic, socioeconomic, and/or other groups.

Sensitivity: Topics and situations that might be perfectly acceptable in other contexts can adversely affect students taking a test on their own, with no opportunity for discussion or adult guidance. Items should therefore avoid topics that students might reasonably find controversial or upsetting. Some examples include alcohol, tobacco, or drug use; serious social issues such as poverty or sexual abuse; references to recent national tragedies; etc.

Example

X Item has sensitive content.

Winemakers, also called vintners, have been studying wine grapes for centuries. Many factors influence the flavor of wine, but by far the greatest influence is the type of grape used. Through much trial and error, vintners have learned which grapes grow best under which conditions.

A major factor affecting the grape is the temperature of the growing season. Some grapes thrive in hot climates, while others grow well only in cooler climates.

Study the chart showing ideal growing season temperatures for different varieties. Using this information, determine the best varieties to be grown in Healdsburg, California, where the average temperature between April and October is 67 degrees Fahrenheit.

[chart]

- A. Merlot, Malbec, or Zinfandel*
- B. Pinot Gris, Riesling, or Pinot Noir
- C. Riesling, Sauvignon Blanc, or Syrah
- D. Pinot Noir, Chardonnay, or Tempranillo

This science item was submitted by a vendor. It aligns to the assigned standard and DOK level, but the context is inappropriate because of the focus on alcoholic beverages.

5

Make sure the stem follows best practices.

The **stem** is the part of an item that states the question to be answered or the task to be completed. Follow these guidelines for **multiple-choice item stems**:

- Write clearly and simply, avoiding trick questions that require students to figure out a particular angle or hidden agenda.
- Be concise.

Example

X Not concise

Flowers need sunlight and water. How do sunlight and water help the flowers?

✓ Concise

Why do flowers need sunlight and water?

- Ask positive questions rather than using the word *not* or other negative words. (There may be exceptions, but generally, positive questions are easier to follow.)

X Negative

Which of the following is NOT a right guaranteed to U.S. citizens?

✓ Positive

Which of the following is a right guaranteed to U.S. citizens?

- Avoid absolutes. Try to use *most likely* instead of *always*. *Always* is an absolute; very few answers to an item will always be correct. (There may be exceptions in some cases, such as for geometry and proportional reasoning skills.)
- Ask only one question.
- Include accurate and necessary content information. This might include stimuli such as art, a passage quotation, a dictionary definition, etc.

For passage-based items

- Require students to read and understand the passage rather than rely on background knowledge. (There may be exceptions for subject areas other than reading.)
- Ask questions worth asking (focus on significant ideas, not trivial details).

For items with context & real-world scenarios

- Use content that is accurate and realistic.
- Choose content that will not be easily outdated.
- Be careful not to introduce potential bias/sensitivity concerns.
- Vary these items within standard or item sets.
- Appeal to many different student interests (art, music, reading, sports, nature, etc.).
- Be grade-level appropriate.
- Provide enough information for students to answer the question, but not so much that it is overwhelming or difficult to comprehend.
- Avoid including extraneous information.

6

Make sure the answer options follow best practices.

Most multiple-choice items have four answer options: three distractors (wrong answers) and one key (correct answer). Lower grade levels sometimes include only three options.

Follow these guidelines for **ALL answer options**:

- Be sure answer options are parallel in terms of verb usage, sentence or phrase structure, length, etc.

- Avoid outliers. One answer option should not stand out in comparison to the others.

Example (middle grades)

X Has an outlier

Based on the speech, what important victory for women paved the way for future women?

- A. A businessman believed women could be as good at bookkeeping as men.
- B. A seamstress of the city made good wages for her clothing.
- C. Men recognized a woman as an equal coworker.
- D. Elizabeth Blackwell went to medical school.*

Option D is an outlier because it's the only one with a proper noun.

✓Does not have outliers

According to the speech, what important victory for one woman helped pave the way for other women?

- A. Elizabeth Blackwell went to medical school.*
- B. Sojourner Truth did the same work as a man.
- C. Mrs. Gifford was able to support herself financially.
- D. Sarah Owen spoke about women's working abilities.

All four options include a proper noun and have a parallel structure. Another approach would be to have two options with proper nouns and two without. This would create balance across the four choices.

- Avoid use of *all of the above*, *none of the above*, or other absolutes. Students gravitate toward these options regardless of their accuracy.
- Avoid polar opposites. (There may be exceptions to this rule, as with items on synonyms or antonyms.)
- Arrange answer options in a logical order. For example, they might be stacked by length, in ascending or descending order (for numbers), in alphabetical order, or in the order they appear in a passage or piece of art.

Continues on next page.

6

Make sure the answer options follow best practices.

Continued

Follow these guidelines for **distractors (wrong answers)**:

- Use distractors that represent errors or misconceptions that students commonly make.

Example

X One of the distractors does not represent a common error or misconception.

What is 8×4 ?

- A. 12
- B. 24
- C. 30
- D. 32*

Unlike A ($8 + 4$) and B (8×3), option C does not represent an error that students would commonly make. Students know that only factors multiplied by 5 or 10 can have products that end in 0.

- Be sure distractors are plausible but clearly incorrect.

X The distractors are implausible.

Which of these is a complete sentence?

- A. he
- B. to college
- C. while he was
- D. He went to college while he was working full time.*

Options A-C are implausible because they are very short and obviously don't express a complete thought. A more plausible distractor might be "while he was working full time."

- Avoid distractors that are overly negative. These options are rarely correct, so they narrow the field of "real" choices.

X One of the distractors is overly negative.

Which word BEST describes the mood of the speaker in this poem?

- A. sad
- B. furious
- C. playful*
- D. thoughtful

It is unlikely that a speaker on an assessment would be "furious," so option B would be easy for students to eliminate without much thought.

7

Avoid internal and external clueing.

Internal Clueing: The item stem (including any art) should not clue the correct answer.

- If an important word, phrase, or number appears in the stem, it should appear in zero, two, or all four answer options.
- Internal clueing can also result from a lack of parallelism and balance, as described in section 6. For example, the key might stand out because it's far longer than the other options.

Example

X Has internal clueing

How does Michael get home?
 A. He takes the bus.
 B. He walks with a friend.
 C. He rides his bike home.*
 D. He runs for three blocks.

The word *home* appears in both the stem and the key.

✓ Does not have internal clueing

How does Michael get home?
 A. He rides his bike.*
 B. He takes the bus.
 C. He walks with a friend.
 D. He runs for three blocks.

External Clueing: One item in a set should not clue the answer to another.

- Avoid providing information that may be used to answer other items.
- In most cases, items should function independently. The answer to one item should not be required as a condition for answering subsequent items. (There may be exceptions for performance tasks, multi-part items, etc.)

X Items have external clueing.

Which word describes how the girl feels about having a party?

- A. calm
- B. upset
- C. excited*
- D. worried

What is the story MOSTLY about?

- A. a girl who has a big family
- B. a boy who wants to be a good friend
- C. how a boy makes a present for his friend
- D. why a girl is excited about having a party*

The stem and key of the first item clue the key of the second item. Also, the key of the second item clues the key of the first item.

8

Follow project guidelines for formatting and style.

Consult the **project guidelines** about details such as the following:

- Capitalizing emphasis words (MOST LIKELY, BEST, etc.) in item stems

Example

What is the MOST LIKELY reason the author starts the passage by mentioning GPS?

- Using open vs. closed item stems

Open: Native Americans used sunflowers to

Closed: How did Native Americans use sunflowers?

- Where to place stimulus material, such as art or passage quotations

Look at the picture.



Which sentence from the passage does the picture show?

- A. Then Rod sends a note back.
- B. Mom gives a note to Dad.
- C. Gram gets two notes!*

Note that the picture appears after the directive and before the question. Note, too, that this item has only three answer options. This is common for grades 2 and below, though it can vary from one project to the next. Guidelines about parallelism and outliers still apply.

- Avoiding initial caps and periods for answer options that aren't sentences
- Lining up numbers used as answer options (right-align, align on decimal points, etc.)
- Applying MathType/LaTeX rules correctly and consistently
- Capitalization/labeling in art
- Questions about grammar, usage, punctuation, and spelling

9

Include effective distractor rationales when projects allow for them.

- The purpose of a rationale is to explain why the distractor is incorrect. It can describe the mistake, misconception, misinterpretation, or incorrect inference used to arrive at the incorrect answer.
- Even if your project does not require rationales, it is best practice to develop them as you develop an item to ensure quality distractors.
- Each distractor should have a unique rationale.
- Avoid rationales such as "This answer is incorrect" and "This answer is wrong." They don't help students understand why the answer choice is incorrect.
- Rationales should be clear, concise, and written as complete sentences.
- The language and format should be consistent across sets of rationales.
- If a unique rationale is NOT required for the correct answer, then the rationale often appears as follows: Correct answer (capital C; no period).
- Requirements for rationales can change from one project to the next. Be sure to consult project-specific guidelines.

Example

Item (grade 3):

What do sharks do instead of sleeping?

- A. come up often for air
- B. rest by floating in the water*
- C. bury themselves in the mud
- D. rest on the bottom of the ocean

X Ineffective Rationales

- A. Sharks do not come up often for air.
- B. Correct answer
- C. This answer is incorrect.
- D. This option is wrong.

Rationale A: This may be true, but it is similar to saying "This answer is incorrect." It does not explain why.

Rationales C and D: These rationales point out that the answer choice is incorrect, but they do not help students understand why.

✓ Effective Rationales

- A. Dolphins, not sharks, come up often for air.
- B. Correct answer
- C. Fish and frogs, not sharks, bury themselves in the mud.
- D. Some fish rest on the bottom of the ocean. Sharks do not.

10

Make sure item art follows guidelines for content and style.

Art should

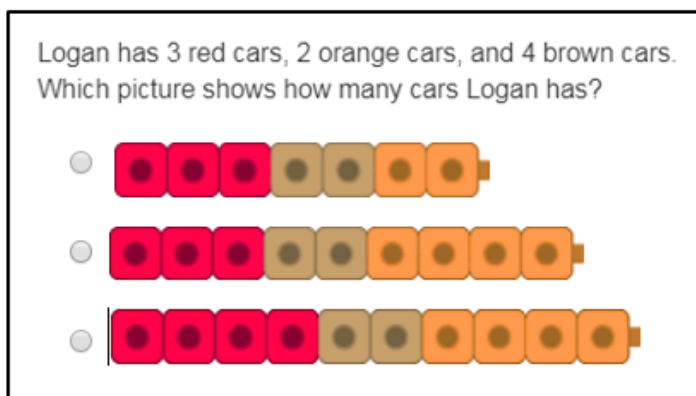
- be clear and easy to read
- be used only when relevant to the skill, standard, or learning objective
- be accurate
- be inclusive of all student populations
- use complementary colors according to the colorblind palette
- be properly labeled (unless labeling is part of the expected student response)
- be consistently formatted across items, forms, and grades

Art should NOT

- be visually challenging to students
- create scrolling issues that can be caused by the number, type, and/or orientation of certain pieces of art
- introduce ambiguity about how to arrive at a correct answer
- require a correct response that is dependent on color

Example

X Art does not follow guidelines.



- This item is both color dependent and visually challenging, which can pose issues for students who may be colorblind or have a visual impairment.
- The distinction between orange, brown, and red can be difficult for younger students and should be improved by using contrasting colors.
- Images of cars would have been more helpful and provided a tighter connection to the item's context.
- The item does not have a correct answer.

Appendix: Depth-of-Knowledge Resources

The Four DOK Levels

Depth of Knowledge 1: Recall and Reproduction

- Items in this category call for the student to remember a fact, term, or simple process. The student does not need to figure out how to attack the question.
- Verbs that typically signal this level of item include: identify, recognize, use, calculate, measure.
- Examples: Choose the correct definition for a term; identify the diagram that correctly shows a process; match a condition with a typical cause or effect.

Depth of Knowledge 2: Skills and Concepts

- These items require some decisions about how to approach the question. More complex content knowledge is involved.
- Typical verbs are: classify, organize, estimate, make observations, collect and display data, compare data. More than one step may be involved.
- Examples: Describe or explain the relationships among facts, terms, or variables; explain examples and nonexamples of science concepts; organize, represent, and interpret data.

Depth of Knowledge 3: Strategic Thinking

- These items require higher-level thinking than levels 1 and 2. The student must reason, plan, cite evidence.
- Examples: Draw a conclusion after examining data; apply a formula in an unfamiliar context; formulate a testable hypothesis based on observations and identify variables to be used in an experiment.

Depth of Knowledge 4: Extended Thinking

- Assessments in this category require high-level thinking, such as connecting ideas from disparate content areas.
- Example: Conduct an investigation from formulating a problem through analyzing data and forming conclusions.

Sample Question Stems by DOK Level

<p>DOK 1</p> <ul style="list-style-type: none"> • Can you list _____? • When did _____ happen? • Who was _____? • How can you recognize _____? • What is _____? • How can you find the meaning of _____? • Can you recall _____? • Can you select _____? • How would you write _____? • What might you include on a list about _____? • Who discovered _____? • What is the formula for _____? • Can you identify _____? • How would you describe _____? 	<p>DOK 2</p> <ul style="list-style-type: none"> • Can you explain how _____ affected _____? • How would you apply what you learned to develop _____? • How would you compare _____? • Contrast _____? • How would you classify _____? • How are _____ alike? Different? • How would you classify the type of _____? • What can you say about _____? • How would you summarize _____? • How would you summarize _____? • What steps are needed to edit _____? • When would you use an outline to _____? • How would you estimate _____? • How could you organize _____? • What would you use to classify _____? • What do you notice about _____?
<p>DOK 3</p> <ul style="list-style-type: none"> • How is _____ related to _____? • What conclusions can you draw _____? • How would you adapt _____ to create a different _____? • How would you test _____? • Can you predict the outcome if _____? • What is the best answer? Why? • What conclusion can be drawn from these three texts? • What is your interpretation of this text? Support your rationale. • How would you describe the sequence of _____? • What facts would you select to support _____? • Can you elaborate on the reason _____? • What would happen if _____? • Can you formulate a theory for _____? • How would you test _____? • Can you elaborate on the reason _____? 	<p>DOK 4</p> <ul style="list-style-type: none"> • Write a thesis, drawing conclusions from multiple sources. • Design and conduct an experiment. Gather information to develop alternative explanations for the results of an experiment. • Write a research paper on a topic. • Apply information from one text to another text to develop a persuasive argument. • What information can you gather to support your idea about _____? • DOK 4 would most likely be the writing of a research paper or applying information from one text to another text to develop a persuasive argument. • DOK 4 requires time for extended thinking.

From Depth of Knowledge – Descriptors, Examples and Question Stems for Increasing Depth of Knowledge in the Classroom Developed by Dr. Norman Webb and Flip Chart developed by Myra Collins

Sample Activities by DOK Level

Level One Activities	Level Two Activities	Level Three Activities	Level Four Activities
<p>Recall elements and details of story structure, such as sequence of events, character, plot and setting.</p> <p>Conduct basic mathematical calculations.</p> <p>Label locations on a map.</p> <p>Represent in words or diagrams a scientific concept or relationship.</p> <p>Perform routine procedures like measuring length or using punctuation marks correctly.</p> <p>Describe the features of a place or people.</p>	<p>Identify and summarize the major events in a narrative.</p> <p>Use context cues to identify the meaning of unfamiliar words.</p> <p>Solve routine multiple-step problems.</p> <p>Describe the cause/effect of a particular event.</p> <p>Identify patterns in events or behavior.</p> <p>Formulate a routine problem given data and conditions.</p> <p>Organize, represent and interpret data.</p>	<p>Support ideas with details and examples.</p> <p>Use voice appropriate to the purpose and audience.</p> <p>Identify research questions and design investigations for a scientific problem.</p> <p>Develop a scientific model for a complex situation.</p> <p>Determine the author's purpose and describe how it affects the interpretation of a reading selection.</p> <p>Apply a concept in other contexts.</p>	<p>Conduct a project that requires specifying a problem, designing and conducting an experiment, analyzing its data, and reporting results/solutions.</p> <p>Apply mathematical model to illuminate a problem or situation.</p> <p>Analyze and synthesize information from multiple sources.</p> <p>Describe and illustrate how common themes are found across texts from different cultures.</p> <p>Design a mathematical model to inform and solve a practical or abstract situation.</p>

Webb, Norman L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <<http://www.wcer.wisc.edu/WAT/index.aspx>>.