

Writing Good Multiple Choice Test Questions

by Cynthia J. Brame, CFT Assistant Director

<https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>

Cite this guide: Brame, C., (2013) Writing good multiple choice test questions. Retrieved [today's date] from <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>.

- [Constructing an Effective Stem](#)
- [Constructing Effective Alternatives](#)
- [Additional Guidelines for Multiple Choice Questions](#)
- [Considerations for Writing Multiple Choice Items that Test Higher-order Thinking](#)
- [Additional Resources](#)

Multiple choice test questions, also known as items, can be an effective and efficient way to assess learning outcomes. Multiple choice test items have several potential advantages:

Versatility: Multiple choice test items can be written to assess various levels of learning outcomes, from basic recall to application, analysis, and evaluation. Because students are choosing from a set of potential answers, however, there are obvious limits on what can be tested with multiple choice items. For example, they are not an effective way to test students' ability to organize thoughts or articulate explanations or creative ideas.

Reliability: Reliability is defined as the degree to which a test consistently measures a learning outcome. Multiple choice test items are less susceptible to guessing than true/false questions, making them a more reliable means of assessment. The reliability is enhanced when the number of MC items focused on a single learning objective is increased. In addition, the objective scoring associated with multiple choice test items frees them from problems with scorer inconsistency that can plague scoring of essay questions.

Validity: Validity is the degree to which a test measures the learning outcomes it purports to measure. Because students can typically answer a multiple choice item much more quickly than an essay question, tests based on multiple choice items can typically focus on a relatively broad representation of course material, thus increasing the validity of the assessment.

The key to taking advantage of these strengths, however, is construction of good multiple choice items.

A multiple choice item consists of a problem, known as the stem, and a list of suggested solutions, known as alternatives. The alternatives consist of one correct or best alternative, which is the answer, and incorrect or inferior alternatives, known as distractors.

Susie Science has discovered a mutant form of insulin that lacks a signal peptide. What will be the final cellular destination of the mutant insulin? } Stem

Answer

Distractor

Distractor

Distractor

Distractor

Distractor

A. Cytosol

B. Endoplasmic reticulum

C. Extracellular space

D. Golgi apparatus

E. Peroxisome

F. Plasma membrane

} Alternatives

Constructing an Effective Stem

1. **The stem should be meaningful by itself** and should present a definite problem. A stem that presents a definite problem allows a focus on the learning outcome. A stem that does not present a clear problem, however, may test students' ability to draw inferences from vague descriptions rather serving as a more direct test of students' achievement of the learning outcome.

STEM IS NOT MEANINGFUL

Which of the following is a true statement?

- A. Mitochondrial genomes are relatively constant in content (i.e., types of genes present).
- B. Mitochondrial genomes are relatively constant in organization.
- C. Mitochondrial genomes are relatively constant in size.

BETTER STEM

What characteristic is relatively constant in mitochondrial genomes across species?

- A. Content (i.e., types of genes)
- B. Organization
- C. Size

2. **The stem should not contain irrelevant material**, which can decrease the reliability and the validity of the test scores (Haldyna and Downing 1989).

IRRELEVANT MATERIAL

Mitochondria evolved from free-living bacteria that could carry our oxidative phosphorylation. For this reason, they have circular genomes that reproduce independently of the nuclear genome. What characteristic is relatively constant in mitochondrial genomes across species?

- A. Content (i.e., types of genes)
- B. Organization
- C. Size

3. **The stem should be negatively stated only when significant learning outcomes require it.** Students often have difficulty understanding items with negative phrasing (Rodriguez 1997). If a significant learning outcome requires negative phrasing, such as identification of dangerous laboratory or clinical practices, the negative element should be emphasized with italics or capitalization.

NEGATIVE PHRASING

Which of the following is not true about mitochondria?

- A. They contain DNA.
- B. They make some of their own proteins.
- C. They are static.

NEGATIVE PHRASING

All of the following are true about mitochondria except

- A. They contain DNA.
- B. They make some of their own proteins.
- C. They are static.

BETTER USE OF NEGATIVE PHRASING

A water-type extinguisher is suitable for putting out a fire caused by burning all of the following except

- A. alcohol
- B. cotton
- C. paper
- D. wood

Source: Cheung, Derek and Bucat, Robert. How can we construct good multiple-choice items? Presented at the Science and Technology Education Conference, Hong Kong, June 20-21, 2002.

4. **The stem should be a question or a partial sentence.** A question stem is preferable because it allows the student to focus on answering the question rather than holding the partial sentence in working memory and sequentially completing it with each alternative (Statman 1988). The cognitive load is increased when

the stem is constructed with an initial or interior blank, so this construction should be avoided.

INTERIOR BLANK

In addition to the nucleus, _____ are organelles that contain DNA.

- A. Golgi bodies
- B. Mitochondria and chloroplasts
- C. Ribosomes

BETTER

In addition to the nucleus, which organelles contain DNA?

- A. Golgi bodies
- B. Mitochondria and chloroplasts
- C. Ribosomes

Constructing Effective Alternatives

1. All alternatives should be plausible. The function of the incorrect alternatives is to serve as distractors, which should be selected by students who did not achieve the learning outcome but ignored by students who did achieve the learning outcome. Alternatives that are implausible don't serve as functional distractors and thus should not be used. Common student errors provide the best source of distractors.

IMPLAUSIBLE ALTERNATIVES

Who gathered the data that helped reveal the structure of DNA?

- A. Francis Crick
- B. George Washington
- C. James Watson
- D. Rosalind Franklin
- E. Snoopy

Note: B and E are not functional distractors and so do not contribute to the item.

2. **Alternatives should be stated clearly and concisely.** Items that are excessively wordy assess students' reading ability rather than their attainment of the learning objective

WORDY ALTERNATIVES

The term hypothesis, as used in research, is defined as

- A. A conception or proposition formed by speculation or deduction or by abstraction and generalization from facts, explaining or relating an observed set of facts, given probability by experimental evidence or by factual or conceptual analysis but not conclusively established or accepted.
- B. A statement of an order or relation of phenomena that so far as is known is invariable under the given conditions, formulated on the basis of conclusive evidence or tests and universally accepted, that has been tested and proven to conform to facts.
- C. A proposition tentatively assumed in order to draw out its logical or empirical consequences and so test its accord with facts that are known or may be determined, of such a nature as to be either proved or disproved by comparison with observed facts.

Source: Steven J. Burton, Richard R. Sudweeks, Paul F. Merri, and Bud Wood. *How to Prepare Better Multiple Choice Test Items: Guidelines for University Faculty*. 1991.

3. **Alternatives should be mutually exclusive.** Alternatives with overlapping content may be considered “trick” items by test-takers, excessive use of which can erode trust and respect for the testing process.

OVERLAPPING ALTERNATIVES

How many chromosomes are found in a typical human cell?

- A. 12
- B. 18
- C. 32
- D. 46
- E. 54

NOTE: The alternatives are overlapping because a cell that contains 18 chromosomes also contains 12; a cell that contains 32 also contains 18 and 12; etc.

4. **Alternatives should be homogenous in content.** Alternatives that are heterogeneous in content can provide cues to student about the correct answer.

HETEROGENEOUS CONTENT

Which of the following is one of the ways that the membranes of winter wheat are able to remain fluid when it is extremely cold?

- A. by increasing the percentage of cholesterol molecules in the membrane
- B. by decreasing the percentage of unsaturated phospholipids in the membrane
- C. by decreasing the percentage of short-chain fatty acids in the phospholipids of the membrane
- D. by increasing the percentage of saturated phospholipids in the membrane

NOTE: Because option A focuses on a different type of molecule (cholesterol) than items B, C, and D, it can cue savvy test-takers to the correct answer.

5. **Alternatives should be free from clues about which response is correct.** Sophisticated test-takers are alert to inadvertent clues to the correct answer, such differences in grammar, length, formatting, and language choice in the alternatives. It's therefore important that alternatives

- have grammar consistent with the stem.
- are parallel in form.
- are similar in length.
- use similar language (e.g., all unlike textbook language *or* all like textbook language).

6. **The alternatives “all of the above” and “none of the above” should not be used.** When “all of the above” is used as an answer, test-takers who can identify more than one alternative as correct can select the correct answer even if unsure about other alternative(s). When “none of the above” is used as an alternative, test-takers who can eliminate a single option can thereby eliminate a second option. In either case, students can use partial knowledge to arrive at a correct answer.

7. **The alternatives should be presented in a logical order** (e.g., alphabetical or numerical) to avoid a bias toward certain positions.

LOGIC INTERNAL TO THE QUESTION

Consider the reaction



In this reaction, which molecule is being oxidized?

- A. Fructose C. O_2
- B. CO_2 D. H_2O

ALPHABETICAL ORDER

Glucose diffuses slowly through artificial phospholipid bilayers. The cells lining the small intestine, however, rapidly move large quantities of glucose from the glucose-rich food into their glucose-poor cytoplasm. Using this information, which transport mechanism is most probably functioning in the intestinal cells?

- A. Active transport pumps
- B. Exocytosis
- C. Facilitated diffusion
- D. Phagocytosis
- E. Simple diffusion

8. **The number of alternatives can vary among items as long as all alternatives are plausible.** Plausible alternatives serve as functional distractors, which are those chosen by students that have not achieved the objective but ignored by students that have achieved the objective. There is little difference in difficulty, discrimination, and test score reliability among items containing two, three, and four distractors.

Additional Guidelines

1. **Avoid complex multiple choice items**, in which some or all of the alternatives consist of different combinations of options. As with “all of the above” answers, a sophisticated test-taker can use partial knowledge to achieve a correct answer.

COMPLEX MULTIPLE CHOICE

Who received a Nobel Prize for discovering the structure of DNA?

- A. Francis Crick
- B. James Watson
- C. Rosalind Franklin
- D. A and B
- E. B and C
- F. A and C

2. **Keep the specific content of items independent of one another.** Savvy test-takers can use information in one question to answer another question, reducing the validity of the test.

Considerations for Writing Multiple Choice Items that Test Higher-order Thinking

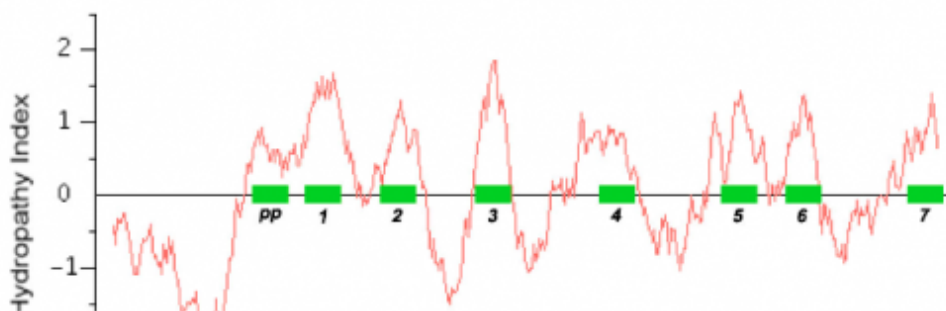
When writing multiple choice items to test higher-order thinking, design questions that focus on higher levels of cognition as defined

by Bloom's taxonomy. A stem that presents a problem that requires application of course principles, analysis of a problem, or evaluation of alternatives is focused on higher-order thinking and thus tests students' ability to do such thinking. In constructing multiple choice items to test higher order thinking, it can also be helpful to design problems that require multilogical thinking, where multilogical thinking is defined as "thinking that requires knowledge of more than one fact to logically and systematically apply concepts to a ... problem" (Morrison and Free, 2001, page 20). Finally, designing alternatives that require a high level of discrimination can also contribute to multiple choice items that test higher-order thinking.

TESTING HIGHER ORDER THINKING

A hydropathy plot for the protein AmGr2 is shown below. This protein has a signal peptide at its amino-terminus, and its hydrophobicity is shown in the hydropathy plot. Regions of the protein with a hydropathy index greater than zero are considered hydrophobic; regions with a hydropathy index less than zero are considered hydrophilic. Based on this information, what kind of protein do you think AmGr2 is?

- A. Integral membrane protein with a single transmembrane domain
- B. Cytosolic protein
- C. Multi-pass integral membrane protein with multiple transmembrane domains
- D. Lipid-anchored membrane protein
- E. Peripheral membrane protein



NOTE: This item tests higher order thinking skills because it requires that test-takers be able to *interpret* the plot and *apply* biochemical principles to choose the best answer.

Figure Source: Kent, Lauren B. and Robertson, Hugh M. Evolution of sugar receptors in insects. *BMC Evolutionary Biology* 9: 41, 2009

TESTING HIGHER ORDER THINKING

The nurse is making a home visit to a 75-y-o male client who has had Parkinson's disease for the past five years. Which finding has the greatest implication for this patient's care?

- A. The client's wife tells the nurse that the grandchildren have not been to visit for over a month.
- B. The nurse notes that there are numerous throw rugs throughout the client's home.
- C. The client has a towel wrapped around his neck that the wife uses to wipe her husband's face.
- D. The client is sitting in an armchair, and the nurse notes that he is gripping the arms of the chair.

NOTE: This item tests higher order thinking skills because it requires that test-takers be able to **apply** multilogical thinking and a high level of discrimination to choose the best answer.

Source: Morrison, Susan and Free, Kathleen Walsh. Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education* 40: 17-24, 2001.

Additional Resources

- Burton, Steven J., Sudweeks, Richard R., Merrill, Paul F., and Wood, Bud. How to Prepare Better Multiple Choice Test Items: Guidelines for University Faculty, 1991.
- Cheung, Derek and Bucat, Robert. How can we construct good multiple-choice items? Presented at the Science and Technology Education Conference, Hong Kong, June 20-21, 2002.
- Haladyna, Thomas M. Developing and validating multiple-choice test items, 2nd edition. Lawrence Erlbaum Associates, 1999.
- Haladyna, Thomas M. and Downing, S. M.. Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 2(1), 51-78, 1989.
- Morrison, Susan and Free, Kathleen. Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education* 40: 17-24, 2001.



This teaching guide is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).