

## Guidelines for Writing MCQ

**Reference:** These guidelines are taken from Ceil R. Reynolds, Ronald B. Livingston, and Victor Willson. (2009). *Measurement and assessment in Education* (2<sup>nd</sup> Edition). Pearson, Upper Saddle River, New Jersey; Columbus, Ohio.

Be flexible when applying these guidelines. Although these suggestions apply in most cases, there are exceptions.

## Tips for Writing Question Stems

1. **Have the item stem contain all information necessary to understand the problem or question.**

**Example Poor–Inadequate stem**

Absolute zero point.

- A. interval scale
- B. normal scale
- C. ordinal scale
- D. ratio scale. \*

**Example Better–Adequate stem**

Which scale of measurement incorporates a true or absolute zero point?

- A. interval scale
- B. normal scale
- C. ordinal scale
- D. ratio scale. \*

2. **It is better to write the stem in the form of a question rather than a statement.**
3. **Avoid negatively stated stem in *most* situations.** Limit the use of terms such as *except*, *least*, *never*, and *not*. Students might overlook these terms and miss the question every if they have mastered the learning objective being measured.

Occasionally, it may be necessary to state stems in the negative. For example, in some situations it is important for students to know what not to do. In these situations, you should **highlight** the negative terms by capitalizing, underlining, or printing them in bold type.

**Example Poor—Negatively stated stem**

Which state does not have a coastline on the Gulf of Mexico?

- A. Alabama
- B. Florida
- C. Tennessee \*

D. Texas

**Example Better—Negative term highlighted**

Which state does NOT have a coastline on the Gulf of Mexico?

- A. Alabama
- B. Florida
- C. Tennessee \*
- D. Texas

4. **Avoid cues in stem that inadvertently identify the correct answer.** Item stem should not contain information that gives away the answer. A cue often involves an association between the words in the stem and the correct alternative.

**Example Poor --Stem contains a cue to the correct answer**

Which type of validity study examines the ability of test scores to predict a criterion?

- A. interval study
- B. content study
- C. factorial study
- D. predictive study \*

In this example, the use of predict in the stem and predictive in the correct alternative provides a cue to the correct answer.

**Example Better—Cues avoided**

Which type of validity study involves a substantial time interval between when the test is administered and when the criterion is measured?

- A. interval study
- B. content study
- C. factorial study
- D. predictive study \*

In this example, the cue was removed. Additionally, in this second example, there is an intentional verbal association between the stem and the first distractor (i.e., interval). This association makes the first distracter more attractive, particularly to students replying on cues, who do not know the correct answer.

5. **Avoid extraneous information in the stem or concept teaching.**

**Example Poor-Extraneous information in stem**

*Radiation treatment in the head and neck can affect the microvascularization of the bone. Which arch (the maxilla or mandible) is more vulnerable to this post radiation effect?*

In this example, the first sentence is extraneous information.

6. **Avoid artificially inflating the reading level.** This does not mean to avoid scientific or technical terms necessary to state the problem, but simply to avoid the unnecessary use of complex

incidental words. (Note: This is related to test validity. You might be testing students' reading comprehension level rather than their mastery of the learning objective.)

## Tips for Writing Answer Choices

1. **Provide between 3 to 5 alternatives.** Although there is no "correct" number of alternatives, it is recommended that you use between three and five. Four are the most commonly used, but some test developers suggest using five to reduce the change of correctly guessing the answer. The use of **five** alternatives is probably the upper limit.
2. **Minimize the use of "None of the Above" and avoid using "All of the Above".** "None of the above" is criticized because it automatically forces the item into a correct-answer format, which is often limited to lower-level educational objectives and easier items. Although there are times when "none of the above" is appropriate as an alternative, it should be used sparingly.

Testing experts are more unified in their criticism of "all of the above" as an alternative. There are two primary concerns. First, students might read alternative A, see that it is correct, and mark it without even reading alternatives B, C, and D. Second, students might know that two of the alternatives are correct and therefore conclude that "all of the above" is correct. In this situation, the student's response is correct but is based on incomplete knowledge. Therefore, the recommendation is to use "none of the above" sparingly and avoid using "all of the above".

3. **Limit the use of *Always* and *Never* in the alternatives.** The use of always and never should be generally avoided because it is only in mathematics that their use is typically justified. Savvy students know this and will use this information to rule out distracters.
4. **Keep the alternatives brief.** Alternatives should be as brief as possible to make it easier for students to scan them for the correct answer.

### **Example Poor – Inadequate stem and lengthy alternatives**

Andrew Jackson \_\_\_\_\_

- A. was born in Virginia.
- B. did not fight in the American Revolution due to a childhood illness.
- C. was the 7<sup>th</sup> president of the United States \*
- D. Served three terms as president of the United States.

### **Example Better -Adequate stem and brief alternatives**

Who was the 7<sup>th</sup> president of the United States of America?

- A. Andrew Jackon \*
- B. James Monroe
- C. John Adams
- D. Martin Van Buren

5. **Ensure that all alternatives are approximately equal in length and complexity.** In addition to the stem containing cues to the correct answer, the alternatives can themselves contain cues. In

an attempt to be precise, instructors sometimes make the correct answer longer or more complex than the distracters. This is another type of cue for students.

**Example Poor –Unequal length and complexity of alternatives**

Ecology is the study of \_\_\_\_\_.

- A. genetics.
- B. organisms and their relationship to the environment. \*
- C. internal balances.
- D. evolution.

**Example Better ---Alternatives similar in length and complexity.**

Ecology is the study of\_\_\_\_\_.

- A. the genetic and molecular basis of organisms.
- B. organisms and their relationship to the environment. \*
- C. how organisms maintain their delicate internal balance.
- D. how organisms have slowly evolved over the last million years.

**6. Avoid overlapping distractors (e.g., 3 – 4 years, 4 – 5 years, 5 – 6 years, etc.)**

- 7. Make sure all alternatives are grammatically correct relative to the stem.** It is recommended to always proofread your questions before using them in the exam.

**Example Poor –Grammatical cue present**

Which individuals are credited with making the first successful flights in a heavier-than-air aircraft that was both powered and controlled?

- A. Octave Chanute
- B. Otto Lilienthal
- C. Samuel Langley
- D. Wilbur and Orville Wright \*

In this example, the phrase “individuals **are**” in the stem indicated a plural answer.

**Example Better –Grammatical cue avoided**

Which individuals are credited with making the first successful flights in a heavier-than-air aircraft that was both powered and controlled?

- A. Octave Chanute and Sir George Cayley
- B. Otto Lilienthal and Francis Herbert Wenham
- C. Samuel Langley and Alphonse Penaud
- D. Wilbur and Orville Wright \*

The grammatical cue is corrected in this example.

Another common error is inattention to the articles **a** and **an**.

**Example Poor –Grammatical cue present**

A coherent and unifying explanation for a class of phenomena is a \_\_\_\_\_.

- A. Analysis.
- B. Experiment.
- C. Observation.
- D. Theory. \*

**Example Better—Grammatical cue avoided**

A coherent and unifying explanation for a class of phenomena is a \_\_\_\_\_.

- A. Conjecture.
- B. Hypothesis.
- C. Prediction.
- D. Theory. \*

**Example Poor—Grammatical cue avoided**

A coherent and unifying explanation for a class of phenomena is a(n) \_\_\_\_\_.

- A. Analysis.
- B. Experiment.
- C. Observation.
- D. Theory. \*

8. **Arrange alternatives in an order that promotes efficient scaring.** When applicable, alternatives should be **arranged in a logical order** to promote efficient scanning. For example, numbers should be placed in ascending order, dates ordered in temporal sequence, and nouns and names alphabetized.

**Example Poor—Illogical arrangement of alternatives**

What year did the Spanish-American War occur?

- A. 1912
- B. 1890
- C. 1908
- D. 1898 \*
- E. 1902

**Example Better—Logical arrangement of alternatives**

What year did the Spanish-American War occur?

- A. 1890
- B. 1898 \*
- C. 1902
- D. 1908
- E. 1912

9. **Have all distractors appear plausible.** All distractors should appear plausible and should be **based on common student errors**. For example, what concepts, terms, and techniques are commonly confused? After you administer the test once, analyze the distractors to determine which are effective and which are not. Replace or revise the ineffective distractors.
10. **Use alternative positions in a random manner for the correct answer.** The correct answer should appear in each of the alternative positions approximately the same number of times. When there are four alternatives (A, B, C, and D), instructors tend to overuse the middle alternatives (i.e., B and C). Alert students are likely to detect this pattern and use it to answer questions of which they are unsure. Students have indicated that when faced with a question they cannot answer based on knowledge, they simply selected B or C.