How To Write a Good Multiple Choice Question

INTRODUCTION

The purpose of each question is to discriminate between those who have command of a learning outcome and those who lack it. It is not to entertain. It is not an IQ test or a tricky puzzle. Our purpose is to minimize any advantage a person would have who is a “good test taker” but who has not mastered the subject being tested, and to do so by (1) requiring knowledge and (2) avoiding clues based on how the question is worded or structured.

STRUCTURE OF MULTIPLE-CHOICE TEST ITEMS

Let’s refer to each multiple-choice question as an item. Each item is composed of a stem and four responses (or answers), exactly one of which is correct. When writing an item, follow these rules of structure:

1. Capitalize the first word of the stem as you would with a sentence; follow accepted rules for capitalization;
2. If the stem is a complete sentence, end the stem with a period or question mark, as appropriate.
3. Identify each response with an upper case letter followed by a period: A., B., C., or D.
4. When using a complete sentence or complete question as the stem, begin all responses with upper-case letters;
5. Should the item stem be incomplete or require internal completion, begin all responses with lower case letters;
6. Do not end a response with a period or other punctuation mark.
7. For emphasis, use upper case letters for an entire word; e.g., MOST is capitalized for emphasis.

The examples in this guide follow these conventions, and also indicate the key with a preceding asterisk.

DEVELOPMENT OF MULTIPLE-CHOICE TEST ITEMS

To develop multiple-choice test items, the writer must:

1. Formulate a question or an incomplete sentence that clearly implies a question. This is the stem of the item.
2. Provide a good answer to the question in a few well-chosen words. This is the key of the item.
3. Produce three plausible, but incorrect, answers to the question. These are called the distracters.
THE MULTIPLE-CHOICE ITEM STEM

The function of the stem is to acquaint the examinee with the problem that is being posed. It should state or imply a specific question. In designing your questions, take particular care to focus on a specific problem. Generalities will lead to ambiguity and to confusing multiple interpretations. Below are three examples of good multiple-choice items.

Example 1: The item stem is a complete, direct statement.

At what time would a structured walk-through be MOST effective?

* A. Before program coding
  B. At post-completion audit
  C. While defining program specifications
  D. Before generation of hierarchy charts

Example 2: The item stem is an incomplete statement.

The users’ logical access to a computer application is BEST controlled by

* A. password
  B. encryption of data
  C. intrusion alarm systems
  D. complex operating systems

Example 3: The item stem is internally incomplete

A record is _____ an indexed sequential file by placing a code within a system control field of the record.

* A. added to
  B. deleted from
  C. updated in
  D. replaced in

WRITING THE ITEM RESPONSES

The item writer should adhere to certain writing practices when designing the responses. These practices are intended to keep the test objective.

1. Use parallel structure. If possible, begin each response with the same word, or part of speech. The test-taker will not be confused by multiple grammatical constructions. Neither will he/she be inclined to choose a selection simply because it is different. This is depicted in Example 4.
Example 4: Use of parallel structure in responses

The MOST common technique for handling overflow records on direct-access storage devices is

A. chaining
*B. buffering
C. overlapping
D. randomizing

By making all responses the same part of speech the item writer has ensured that the choices are clear and grammatically unbiased.

2. Avoid complex or unduly long responses. Such responses tend to confuse the test-taker. They frequently strain unnecessarily his/her power to concentrate. This particular difficulty is found in Example 5.

Bad Example 5: Avoid such complex or long responses

When we refer to the concept, “ones complement” arithmetic, what precisely do we mean?

A. All numbers are taken to less than one
B. All arithmetic operations are performed twice for validity checking
C. All numbers are subtracted from one before performing any arithmetic operations
*D. Negative numbers are formed by inverting the bits in the positive number such that all zero bits are ones and all one bits are zeroes

After wading through answers C and D, the test-taker will be strained to remember the actual question.

THE ITEM KEY: ABSOLUTE CORRECTNESS OR BEST POSSIBLE RESPONSE

The intended answer to a multiple-choice question ideally should be a thoroughly correct answer. It should admit no difference of opinion among adequately informed experts. Test designers occasionally must base their items on propositions that may not be absolutely true, but are, nonetheless, nearly 100% probable. Notice in Example questions 1 and 2 above that the item writer has stressed the words “best” and “most” in the item stem. This directs the test-taker not to look for absolute correctness, but rather the best option.

Another guideline to follow is that the stem of a multiple-choice item should ask a question that has a definite answer. Indeterminate questions may provide interesting topics for discussion, but
they do not make good items for testing achievement. Example 6 below demonstrates an indeterminate question.

**Bad Example 6: Avoid such indeterminate questions**

The practice of structured programming is an outgrowth of

A. faster computers  
B. increased computer storage  
*C. escalating programmer salaries  
D. rising costs of program maintenance

In this example, an adequate argument could be mounted for several of the options; establishing one correct answer becomes a matter of subjective choice.

**THE ITEM DISTRACTERS**

The purpose of a distracter in a multiple-choice test item is to discriminate between those who have command of a specific body of knowledge and those who lack it. To do this, a distracter must be a plausible alternative. In other words, it should be appropriate to the question or implied by the stem. One way to obtain plausible distracters is to use appropriate responses which simply do not correctly answer the stem question. Example 7 demonstrates this idea.

**Example 7: Appropriate answers that do not correctly answer the stem**

In a memory-constrained virtual storage operating system, which of the following types of interrupt is the MOST common?

A. Machine check  
B. Page interrupt  
*C. Timer interrupt  
D. Program interrupt

In this example, all of the choices are truly types of interrupts. However, since the question specifies the most common form, only one response completely answers the question: the timer interrupt.

Other sources of plausible distracters are familiar expressions and phrases. Because they have been used in common parlance, they are frequently attractive to examinees with merely superficial knowledge. See Example 8 below.

**Example 8: Using familiar expressions as distracters**
A group of characters that is read or written with each physical read or write operation on a storage medium is called a

A. file
B. field
C. logical record
*D. physical record

It is highly likely that the test-taker would have heard the terms, “file”, “field”, and “logical record”. But a person with a low-level understanding of the topic might well have not understood them. Therefore, these familiar terms provide excellent distracters at the elementary level of discrimination.

“ALL OF THE ABOVE” AND “NONE OF THE ABOVE” AS RESPONSES

A very poor device for adapting multiple-choice items to questions that seem to require several correct answers is to add as a final alternative the response, “all of the above”, or “answers A and C only”, or “none of the above”. A correct answer should not be wrong simply because there are other correct answers. In addition, the experienced test-taker who may have an incomplete knowledge of the subject will have an improved chance of guessing the correct answer. Further, when the exam is administered, the responses are displayed in random order for each examinee, making an order-dependent key unintelligible. Example 9 demonstrates this type of question.

**Bad Example 9: Avoid “all of the above” and “none of the above”**

During a preliminary study to determine the number of invoices prepared per month, you should

A. review a production log
B. ask a direct question
*C. do both of the above
D. do neither of the above

If the test-taker knows that option A is correct but is uncertain about B, he/she can eliminate options D and B, greatly improving his/her chances. This disturbs the statistical performance of the exam.