Van Blerkom Chapter 8: Multiple-Choice Items

 Multiple-choice items typically contain two components, the **stem**, which describes the problem or sets the question, and the **options or alternatives**, which are the set choices one must choose among. Of the options, usually only one is correct and the others, which are labeled **distractors or foils**, are incorrect.

1. Advantages and Limitations of Multiple-Choice Items

1a. Advantages of Multiple-Choice Items

* Multiple-choice items enable one to more adequately sample content relative to essay items. However, short-answer and alternative-choice items will usually allow for slightly better sampling than multiple-choice items, although multiple-choice items typically allow for measurement of higher-level cognitive activity.
* Because of the structured options included with multiple-choice items, one may better frame the problem to be addressed than is often possible with short-answer or essay items. The drawback to this framing, of course, is that it allows for the possibility of guessing the correct answer.
* Multiple-choice items can be efficiently and quickly scored.
* Multiple-choice items can also be objectively scored, much more so than essay or short-answer type items, and as a result, multiple-choice items enable one to achieve higher levels of reliability than is possible with essay items.

1b. Limitations of Multiple-Choice Items

* Since multiple-choice items provide the answer among a set of options, multiple-choice items are susceptible to guessing. This is problematic because it decreases both reliability and validity of scores obtained from the test. However, as the number of items and options increases, the probability of obtaining a high score from guessing alone is very small. Short-answer items are typically more reliable than multiple-choice items since short-answer items eliminate the possibility of guessing. However, essay items are less reliable than multiple-choice since less sampling, thus fewer items, is possible with the essay format.
* Unlike essay items, multiple-choice items typically do not allow for direct observation of higher-level cognitive skills. However, multiple-choice items, unlike short-answer items, allow one to measure, in an indirect fashion, higher-level cognitive skills.
* Constructing multiple-choice items is a time-consuming process. Usually, multiple-choice items require more time to construct than all other formats.

2. Qualities Desired in Multiple-Choice Items

1. **Does this item measure a specified skill**? If not, then it will not contribute to content validity of the test. Following a table of specifications or a list of performance objectives will help ensure the item measures a pre-specified skill.
2. **Will most content experts agree on the correct option**? Without such agreement, it is unlikely that the item will increase reliability or validity of scores since the item will be ambiguous.
3. **Is the item written at a reading level below that of students for which it is designed**? If not, performance on the item will be confounded with reading ability.
4. **Does the stem clearly present the problem**? A student should **not** have to read options to clarify or understand the problem or question being presented in the stem.
5. To measure something beyond recall of information (i.e., knowledge or verbal information), **the question or problem presented in the stem must be new, novel to the student**.
6. **The stem should contain only relevant, concise information**.
7. **Adjective and adverbs that reverse or alter the meaning of the stem should be highlighted or underlined** or, better, the item should be revised to remove the suspect word. The word **not** should typically be avoided in items as it often creates double negative situations that can be confusing and difficult to answer. Further, **not** should never be included in options.
8. **Options should avoid repetitive words or phrases**. Such repetitive information should be included in the stem, if possible. Note also that by including repetitive wording in the stem, it is more likely that the stem will be clearer and more easily understood.
9. **Grammar found in each option should be consistent with the stem** otherwise correct or incorrect options will be easily spotted by students.
10. **The options should be consistent in content**, otherwise correct or incorrect options will be more easily spotted by students. For example, if asking a question about sampling procedures, do not include in the option something on validity (unless this is commonly confused with the correct response).
11. **Typically, one should not use an option such as “all of the above” or “none of the above.”** All of the above options typically make items easier to answer by effectively reducing the number of options included with the item. None of the above options, however, can be used if one is sure to include items that do and do not use none of the above as the correct response.
12. **Arrange options alphabetically unless another order is more traditional**. For example, Bloom’s taxonomy should would best be arranged in the order typically learned (i.e., knowledge, comprehension, application, etc.), but race, sex, numbers, etc. should be arranged alphabetically or numerically.
13. **Include only plausible distractors**. Sometimes item developers will wish to include a certain number of options, however, this should be avoided if one is simply including items to have a certain number.

3. Variations of Multiple-Choice Items

3a. Options Representing Ranges of Values

 For mathematics tests, one may wish to have options that include not the exact correct answer, but a range in which the correct answer may fall. Such a format potentially eliminates the possibility that students learned that they have made a mistake in calculation when their answer does not agree with one of the answers given. The list of qualities listed above for typical multiple-choice items also applies with options that range in values.

3b. Matching Items

 With this technique, one lists a series of options parallel in content and asks that students match the options with a series of stems. Note that this method is more efficient than typical multiple-choice items when one has several stems that all share common options because it reduces required space. The qualities above also apply to this format.

3c. Ranking of Options

 With this format, students are requested to sort options into an appropriate order, such as ranking steps one follows in an experiment or ranking steps in calculating a correlation coefficient. The qualities above also apply to this format with little alteration.

3d. Interpretive Test Exercises

 With this format, a student is asked to read a story, paragraph, etc., or to view data of some sort and then ask a series of questions that reference the material. While one may use typical multiple-choice items to assess indirectly higher cognitive skills, this technique can be used for determining whether students can apply and analyze information. As with the other formats, the qualities above apply to this format as well.

4. Optimal Number of Choices

 When writing multiple-choice items, the determining factor for the number of options to include is the number of plausible distractors that exist. Research has shown that a total of no more than three options – one correct response and two distractors – is usually enough to generate reliable information about student performance. Also, if using fewer options per item enables one to include more items per test, then fewer options should be used to include better sampling of content and skills.

Self-Test: Multiple-Choice Items

For items 1 through 4, use the following options to indicate which item format is being described.

1. Brief Essay
2. Completion or Short-Answer
3. Multiple-Choice

1. This format is **least** able to measure complex cognitive skills.

2. This format most easily presents a concise problem to the student.

3. With which format would the **most** time be required to construct a 1-hour examination?

4. With which format would the **least** amount of time be required to score a 1-hour examination?

5. Would a 30-item multiple-choice or short-answer test likely have the higher reliability?

a. multiple-choice

b. short-answer

Items 6 through 12 propose characteristics one might see in multiple-choice items. Indicate whether each characteristic is

1. Desirable
2. Undesirable

6. If the correct answer to the item is a matter of opinion, the item indicates whose opinion is being sought.

7. The reading level is set at about what an average student can handle.

8. In items measuring higher cognitive skills, a student should have to read each of the options before fully understanding what is being asked.

9. The stem should include extraneous information when the item is measuring higher cognitive skills.

10. A word that by itself changes the meaning of the stem should be underlined.

11. “None of the above” should be used as an option to keep students from assuming the correct answer is listed.

12. To better sample students’ skills, the options to each item should include highly diverse content.

13. If a test includes 12 items, how many options per item are likely to result in the most reliable scores?

a. 2

b. 3

c. 4

d .5

14. If problems presented by each multiple-choice item can be read and answered quickly, which of these two tests would likely have the more reliable scores?

a. A test with 15 items, each with 3 options.

b. A test with 9 items, each with 5 options.