Van Blerkom Chapter 6: Completion and Short-Answer Test Items

 This chapter deals with standard, systematic methods for developing completion and short-answer test items. Employing the methods outlined below will assist with establishing both content validity and reliability.

Note that completion and short-answer items are, in most cases, equivalent formats, with only appearance the being the distinction.

A completion item requests that students fill-in a blank space with the most appropriate response.

Example:

Abraham Lincoln was born in \_\_\_\_\_\_\_\_ (bad example, as we will see later)

A short-answer item requires that students provide a written response to a question rather than simply a fill-in the blank type response.

Example:

Abraham Lincoln was born in which year?

Given this limited distinction, the term short-answer item will be used to refer to both completion and short-answer.

1. Advantages and Limitations of Short-Answer Items

 The discussion of limitations and advantages of short-answer items is normative; that is, the discussion must be relative to other test item formats such as essay and multiple-choice. Keep this in mind as you read the advantages and limitations.

1a. Advantages of Short-Answer Items

* Are easy to construct since they most always measure low-level cognitive capabilities such as knowledge and comprehension, or verbal information and discriminations.
* Require students to supply answers. This, of course, reduces the affect of guessing found with formats like multiple-choice and true-false. As a result, decreases in reliability due to guessing is less problematic.
* Many short-answer items can be incorporate into a test because they typically require less time to complete. This will allow one of cover content in more detail, thus increasing the odds of having a more reliable and valid test.

1b. Limitations of Short-Answer Items

* Usually limited to recall or comprehension of content, thus, higher-level skills (rules, synthesis, analysis) are typically not found with short-answer items. However, for some content areas short-answer items can reflect higher-level scores (e.g., What is the variance for 1, 3, 5, and 7? \_\_\_\_\_ ).
* Short-answer items are more likely to introduce scoring problems since short-answer items, at times, are ambiguous. Recall the above example, Lincoln was born in \_\_\_\_\_\_\_\_\_. Is this question asking for a country, state, town, house, hospital, year, or something else? This problem can be alleviated by following the item development procedure listed below.
* Require more time to score than do multiple-choice and true-false formats, but requires less time to score than essay formats.
* Computerized scoring not yet feasible for shot-answer items, but very common for multiple-choice and true-false items.

2. Qualities Desired in Short-Answer Items

 The following criteria are useful for evaluating short-answer items.

1. Does item measure specific skill. This refers to the content-validity of the item. If item developed using table of specifications or performance objectives, then the item should match a specific skill.
2. Does item require an answer longer than a short sentence? If yes, then perhaps a better format would be short- or brief-response essay. For short-answer items, the goal is to keep responses as brief as possible.
3. Are a limited number of homogenous – highly related – responses likely? If not, then the item is ambiguous, like the Lincoln example above. Starting with the desired response and then developing the short-answer items helps alleviate this potential problem.
4. Will content experts agree as to the specified correct response? If not, the item will not adequately measure achievement. Topics that are based upon opinion are typically not appropriate for short-answer format since opinion justification is often necessary and desired.
5. Is reading skill level appropriate for students who will be taking this test? If the reading skill level is beyond the capabilities of the students responding to the test item, no matter what item format, then both reliability and validity will be sacrificed. Reading skill level is recommended to be below that of the students taking the test.
6. Avoid using sentences found in instructional sources (i.e., books, notes, etc.) as the shot-answer item. Using key sentences and omitting key terms encourages memorization and can be problematic should the key sentence be content specific.
7. Does the physical format of items and layout of test facilitate scoring efficiency? Since one benefit of short-answer items is that they are easier and quicker to score than essay, one should attempt to maximize this benefit by arranging items and responses in such a way to facilitate ease of reading for grader and scoring. The following example shows one method of aligning items and responses to assist in efficient scoring.

|  |  |
| --- | --- |
| Response Area | Item |
| 1. \_\_\_\_\_\_\_\_\_\_\_ | 1. A. Lincoln was born in which state? |
| 2. \_\_\_\_\_\_\_\_\_\_\_ | 2. A. Lincoln was born in which year? |

1. For items that require numerical responses, specify the unit of measurement desired, i.e., indicate the desired unit of measurement such as inches, meters, pounds, ounces, etc.
2. The blank area must represent the key word in the test item if a completion item.
3. For completion items, blanks should be placed at or near the end of the item rather than at the beginning or in the middle. This makes it less likely that students will have to re-read each item, thus enabling one of include more items on the test.
4. Use a consistent length for each blank with completion items. If blanks vary across items, students may either be alerted and recognize that a longer word or response is necessary, or they may be confused thinking that a longer or shorter response is necessary.
5. If possible, include only one blank per completion item. Excessive use of blanks may destroy the meaning of the item and make it very difficult to elicit a correct response.

Self-Test

 For items 1 through 5, indicate (yes or no) whether each statement describes an **advantage** of the completion and short-answer format.

1. Easier to construct than multiple-choice items.

2. For a given number of items, provides more reliable test scores than the multiple-choice format.

3. Compared to the essay format, can measure more skills with a 30-minute test.

4. Can measure students’ ability to synthesize ideas.

5. Can be scored as quickly as multiple-choice items.

 Items 6 through 11 propose characteristics one might see in completion and short-answer items. Indicate whether each characteristic is

A. Desirable

B. Undesirable

6. Short-answer items should be designed so students’ responses are one to three sentences in length.

7. Ideally, all knowledgeable students will use the same words to answer a short-answer item.

8. The best completion items use an important sentence from a source such as a class, lecture, replacing the most significant word with a blank.

9. If a short-answer item asks students to measure a book using a metric system, the item should indicate whether the answer is to be in centimeters or millimeters.

10. to facilitate scoring of completion items, blanks should be inserted near the beginning of sentences.

11. The length of blanks within completion items should correspond to the space needed to write the answer.

 For each of the following, indicate, using the qualities presented above, which of the qualities for short-answer items is violated, if any.

|  |  |
| --- | --- |
| 12. An animal that eats the flesh of other animals is | \_\_\_\_\_\_\_\_\_\_\_\_ |
| 13. If oranges weigh 5.67 oz. each, how much will a dozen oranges weigh? | \_\_\_\_\_\_\_\_\_\_\_\_ |
| 14. What is the name of the part of speech that connect words, clauses, and sentences? | \_\_\_\_\_\_\_\_\_\_\_\_ |
| 15. \_\_\_\_\_\_\_ animals that are born \_\_\_\_\_\_\_ and \_\_\_\_\_\_\_ their young are called \_\_\_\_\_\_\_. | \_\_\_\_\_\_\_\_\_\_\_\_ |
| 16. What causes rain clouds? | \_\_\_\_\_\_\_\_\_\_\_\_ |