The Design and Use of Multiple Choice Questions in Testing

Duane_akroyd@ncsu.edu

Dr. Duane Akroyd, RT(R)
Professor, Department of Adult and Higher Education

Basic Guidelines Achievement Tests
- Measure clearly defined learning objectives
- Measure all instructional outcomes
- Measure representative sample of objectives
- Test should be appropriate for intended learning outcomes
- Tests should be free from measurement error

Norm or Criterion Referenced Test?
- If you want to compare a student's performance to a norm or average of performances of other students, use a norm referenced test (90 percentile).
- If you want to determine a student's level of proficiency by comparing their performance to a standard of mastery called a criterion, use a criterion referenced test (85% score).
- Most programs use...

Test Blueprint
- Constructing a test blueprint ensures that a test samples a range of content areas and cognitive levels.
- A test blueprint should be constructed before a test is administered. Why?

Developing Specific Format Tests
- The multiple choice test
  - Less subjective than an essay test but ability to test at upper levels of cognitive domain is limited (but possible)
  - The following will be a discussion of rules and techniques in writing good multiple choice questions and how to quantitatively analyze their quality

Constructing Good Multiple Choice Questions
- A multiple choice question consists of a STEM which presents a problem situation and a number of alternatives which provide possible solutions to a problem.
  - The stem may be a question, incomplete statement or the best answer form
  - Alternatives include the correct answer and several plausible wrong answers called distracters
Constructing Good Multiple Choice Questions

- Present a single clearly formulated problem in the stem

Constructing Good Multiple Choice Questions

A test blueprint
- a. indicates how a test will be used to improve learning.
- b. Provides a more balanced sampling of content
- c. Arranges the objectives in order of importance.
- d. Specifies the method of scoring for a test.

Constructing Good Multiple Choice Questions

What is the main advantage of using a test blueprint when preparing an achievement test?
- a. It reduces the amount of time required.
- b. It improves the sampling of content.
- c. It makes the construction of test items easier.
- d. It increases the objectivity of the test.

Constructing Good Multiple Choice Questions

Make sure the stem is stated clearly

Constructing Good Multiple Choice Questions

The paucity of probable but incorrect statements that can be related to a central idea pose a problem when constructing which one of the following types of test items?
- a. Short-answer.
- b. True-false.
- c. Multiple choice
- d. Essay.

Constructing Good Multiple Choice Questions

The lack of probable wrong answers will cause the greatest difficulty when constructing which one of the following types of test items?
- a. Short-answer items.
- b. True-false items.
- c. Multiple choice items.
- d. Essay items.
Constructing Good Multiple Choice Questions
Testing can contribute to the instructional program of the school in many important ways. However, the main function of testing in teaching is

The main function of testing in teaching is

Constructing Good Multiple Choice Questions
Put as much wording in the stem as possible (reduce redundancy in the distracters)

Constructing Good Multiple Choice Questions
In *objective* testing, the term *objective*
  a. Refers to the method of identifying the learning outcomes.
  b. Refers to the method of selecting the test content.
  c. Refers to the method of presenting the problem.
  d. Refers to the method of scoring the answers

Constructing Good Multiple Choice Questions
Which one of the following is a category in the taxonomy for the cognitive domain?
  a. Comprehension.
  b. (distracter needed)
  c. (distracter needed)
  d. (distracter needed)

Constructing Good Multiple Choice Questions
State the stem in a positive form when possible, if stated in the negative form underline, or bold the word *not*
Constructing Good Multiple Choice Questions

Which one of the following is not a category in the taxonomy for the cognitive domain?

- b. Application.
- c. Analysis.
- d. (answer needed)

Constructing Good Multiple Choice Questions

- Make sure the intended answer is correct or clearly the best.

Constructing Good Multiple Choice Questions

- What is the purpose of classroom testing?
  
  A better stem would be as follows;
  
  - One purpose of classroom testing is
  - The primary purpose of classroom testing is

Constructing Good Multiple Choice Questions

- Make all alternatives grammatically consistent with the stem and in parallel form.

Constructing Good Multiple Choice Questions

- Why should negative terms be avoided in the stem of a multiple-choice item?
  - a. They may be overlooked
  - b. The stem tends to be longer.
  - c. The construction of alternatives is more difficult.
  - d. The scoring is more difficult.

Constructing Good Multiple Choice Questions

- Why should negative terms be avoided in the stem of a multiple-choice item?
  - a. They may be overlooked.
  - b. They tend to increase the length of the stem.
  - c. They make the construction of alternatives more difficult.
  - d. They may increase the difficulty of the scoring.
The recall of factual information can best be measured with a multiple-choice item. Therefore, option (b) is correct.

Avoid verbal clues that may enable students to select the correct answer or eliminate an alternative.

Which one of the following would you first consult to locate research articles on achievement testing?
- Journal of Educational Psychology.
- Journal of Educational Measurement.
- Journal of Consulting Psychology.
- Review of Educational Research

Lack of attention to learning outcomes during test preparation may lower the technical quality of the items and make the construction of test items more difficult. It may also result in a greater use of essay questions and a test that is less relevant to the instructional program.
Constructing Good Multiple Choice Questions

- Make distracters plausible and attractive
  - Use common misconceptions or errors
  - Make distracters similar to the correct answer in length and complexity

Constructing Good Multiple Choice Questions

Obtaining a dependable ranking of students is of major concern when using
a. General achievement tests
b. Behavior descriptions.
c. Checklists.
d. Questionnaires.

Constructing Good Multiple Choice Questions

Obtaining a dependable ranking of students is of major concern when using
a. General achievement tests.
b. Diagnostic tests.
c. Mastery tests.
d. Pre-tests.

Constructing Good Multiple Choice Questions

One advantage of multiple-choice items over essay questions is that they
a. Provide for the measurement of more complex leaning outcomes.
b. Place greater emphasis on the recall of factual information.
c. Require less time for test preparation and scoring.
d. Provide a more extensive sampling of course content.

Advantages of Using Multiple Choice Questions

1. Versatility in measuring a variety of cognitive skills.
2. A substantial amount of course content may be evaluated in a short period of time.
3. Scoring is much more objective than other formats (essay and short answers).
4. Students may have to discriminate among options that vary in degree of correctness.
5. Degree of guessing is limited when compared to true/false questions.
6. Ability to conduct item analysis to examine student weaknesses, question ambiguity, question difficulty and the ability of an item to measure individual differences.
Suggestions for Wiring Multiple Choice (MC) Items

1. The stem of the item should clearly formulate a problem. Include as much of the item as possible, keeping the response options as short as possible. However, include only the material needed to make the problem clear and specific. Be concise—don’t add extraneous information.

2. Be sure that there is one and only one correct or clearly best answer.

3. Be sure wrong answer choices (distracters) are plausible. Eliminate unintentional grammatical clues, and keep the length and form of all the answers choices equal. Rotate the position of the correct answer from item to item randomly.

6. To increase the difficulty of a multiple choice item, increase the similarity of content among the options.

7. Use the option “none of the above” sparingly and only when the keyed answer can be classified unequivocally as right or wrong. Don’t use this option when asking for a best answer.

8. Avoid using “none of the above.” It is usually the correct answer and makes the item too easy for students with partial information.

Suggestions for Wiring Multiple Choice (MC) Items (Cont’d)

4. Use negative questions or statements only if the knowledge being tested requires it. In most cases it is more important for the student to know what a specific item of information is rather than what it is not.

5. Include from three to five options (two to four distracters plus one correct answer) to optimize testing for knowledge rather than encouraging guessing. It is not necessary to provide additional distracters for an item simply to maintain the same number of distracters for each item. This usually leads to poorly constructed distracters that add nothing to test validity or reliability.

Item Analysis of MC Tests

- Two different methods to analyze MC test items
  - Quantitative ……  A numerical method that employs student responses
    - Means / Variances
    - % that answer correct (Difficulty Index [P])
    - The ability of a question to discriminate among those who did well on the test and those who did not (Discrimination Index [D])
  - Qualitative ……  A non numerical method for evaluating tests, test objectives, content validity and technical item quality

Relating Item Analysis to Purposes of A Test

- What is the purpose of a norm referenced vs. a criterion referenced test?

- Given the purpose of norm referenced vs. criterion referenced testing, would you expect the scores for a group taking each type of test to be the same? …..Why or why not?
Quantitative Analysis of MC Questions

**Difficulty Index (P)** – the percentage of students who answered an item correctly.

- \( P \) ranges from 0 meaning no students (0%) answered the question correctly, to 1.0 meaning all students (100%) answered the question correctly.

- What \( P \) value do you want on MC questions for summative classroom tests that you do?

**Discrimination Index (D)** – A MC question's ability to discriminate between students who did well on the test and those who did poorly.

- \( D \) ranges from -1.0 to 1.0
  - A Discrimination Index of 1.0 means all students who did well on the test answered the question correctly while all those who did poorly on the test missed the question.
Quantitative Analysis of MC Questions

A negative discrimination index indicates that a greater number of students who performed poorly on the test answered the question correctly than did the high performing students.

All questions with a negative discrimination index must be reviewed for possible problems.

Recommended Book

Educational Testing and Measurement
8th edition, 2007
By Tom Kubiszyn and Gary Borich
Publisher; John Wiley & Sons