IDEA PAPER No. 16

CENTER FOR FACULTY EVALUATION & DEVELOPMENT
Division of Continuing Education
Kansas State University

September 1986 KANSAS STATE UNIVERSITY

Improving Multiple-Choice Tests

by

Victoria L. Clegg William E. Cashin
Kansas State University

" . . . the tendency in course examinations is to pose the question 'How much do you remember of what has been covered?' rather than 'What can you do with what you have learned?' "
Dressel (1976, p. 208)

The classroom test is one of the most important aspects of the teaching-learning process, and designing the classroom test is one of the most demanding responsibilities facing college and university instructors. Unfortunately, most of us have had little, if any, preparation in the craft of writing tests; consequently, the process is not only difficult, it is also frustrating and often ineffective.

Writing test questions will always be demanding, even for experienced instructors, but it will be less frustrating for those who know the techniques for writing specific types of items and have some guidelines for general test construction.

The multiple-choice item has been chosen as the focus of this paper for three reasons. First, multiple-choice items can be written to evaluate higher levels of learning, such as integrating material from several sources, critically evaluating data, contrasting and comparing information. Second, multiple-choice items can be very useful for diagnostic purposes, for helping students see their strengths and weaknesses. Third, multiple-choice items are often used in college and university classes; therefore, it is especially important that instructors write them well. Although these strengths are shared by some other item types, the multiple-choice item is a powerful teaching-learning tool **if the instructor has designed the item properly.**

What Is a Multiple-Choice Item?

The multiple-choice item requires that students **select** the correct answer to a question from an array of alternative responses that are written by the instructor. All multiple-choice items have the same three elements: (1) an **item stem** that presents the problem, (2) the **correct or keyed option**, and (3) several **distractor** options, incorrect alternatives that are likely to be plausible to the student who has not completely mastered the learning being tested. Several variations of the standard multiple-choice item have been used in classroom tests. Some of these will be described later. Typically, multiple-choice items present the problem in one of two formats: the **complete question**, e.g., What is the most frequently used type of test item in college-level examinations?, or the **incomplete statement**, e.g., The most frequent type of test item used in college-level examinations is. . . . The students are directed to select either the correct answer or the best answer from the list of options provided. In the **correct answer** form, the answer is correct beyond question or doubt while the others are definitely incorrect. In the **best answer** version, more than one option may be appropriate in varying degrees; however, it is essential that the keyed or "best" response be the one that competent experts would agree upon.

It may appear to be fairly simple to construct items in the multiple-choice format. Actually, the formatting is simple; it is constructing a meaningful and worthwhile item that is so difficult and time-consuming. "An ingenious and talented item writer can construct multiple-choice items that require not only the recall of knowledge but also the use of skills of comprehension, interpretation, application, analysis, or synthesis to arrive at the keyed answer." (Thorndike & Hagen, 1969, p. 103). How many of us who teach at colleges and universities would describe ourselves as "ingenious and talented" while we struggle to write effective multiple-choice items? Wilbert J. McKeachie (1986, p. 91) has said that " . . . the greater your experience in their construction, the longer it takes per [multiple-choice] item to construct a reasonably fair, accurate, and inclusive question." In other words, as you get better, things may seem worse! We cannot promise you ingenuity and talent. We do hope to help you become a more competent and successful writer of multiple-choice items by sharing some of the guidelines that measurement experts and experienced instructors have recommended.

Many college teachers believe the myth that the multiple-choice question is only a superficial exercise—a multiple-guess—requiring little thought and less understanding from the student. It is true that many multiple-choice items are superficial, but that is the result of poor test craftsmanship and not an inherent limitation of the item type. A well designed multiple-choice item can test high levels of student learning, including all six levels of Bloom's (1956) taxonomy of cognitive objectives.

LEVELS OF COGNITIVE LEARNING

Evaluation
Synthesis
Analysis
Application
Comprehension
Knowledge

Knowledge:	simple recognition or recall of material
Comprehension:	restating or reorganizing material to show understanding
Application:	problem-solving or applying ideas in new situations
Analysis:	separating ideas into component parts, examining relationships
Synthesis:	combining ideas into a statement or product new to the learner
Evaluation:	judging by using self-produced criteria or established standards

Some writers prefer fewer levels, e.g., understanding (combining knowledge and comprehension), application, and higher-order cognitive objectives (combining analysis, synthesis, and evaluation). (See Gronlund, 1985b, for a further treatment of levels of educational objectives.)

Strengths of Multiple-Choice Tests

Multiple-choice items are often described as the most versatile of all item types, suitable to a wide range of instructional goals.

1. Multiple-choice items can be used to test **all levels of learning**, from knowledge to evaluation.

2. Multiple-choice items can assess the **ability to integrate information** from several sources.

3. Multiple-choice items are very **useful for diagnosing student difficulties** if the incorrect options are written to reveal common errors.

4. Multiple-choice items provide an **excellent basis for post-test discussion**, especially if the discussion includes why the distractors are wrong as well as why the correct answers are right.

Multiple-choice items also share many of the strengths of other selected response items, i.e., true-false, matching, etc.

5. Multiple-choice items can provide a **more comprehensive sample of subject material** because more questions can be asked.

6. Multiple-choice items adapt to a **wide range of content and difficulty levels**.

7 Multiple-choice items require relatively **less student time to answer**.

8. Multiple-choice items can be **easily and accurately scored** by a person or machine.

Limitations of Multiple-Choice Tests

Of course, multiple-choice items also have disadvantages.

1. Multiple-choice items are **open to misinterpretation** by students who read more into questions than was intended.

2. Multiple-choice items may **appear too picky** to students, especially when the options are well-constructed.

3. Multiple-choice items, when written to assess higher levels of learning, require significant intellectual effort both in reading and in answering, **causing some students to be anxious**.

In addition, multiple-choice items share the limitations of other selected response items.

4. Multiple-choice items **deny demonstration of knowledge beyond the range of options provided**.

5. Multiple-choice items are **difficult to phrase** so that all students will have the same interpretation.

6. Multiple-choice items **take time and skill** to construct effectively.

7. Multiple-choice items are so easily constructed to assess basic factual knowledge that instructors **often fail to test higher levels of thinking**.

8. Multiple-choice items are ill-suited **to assess affective or attitudinal learning** because they are easily "faked."

9. Multiple-choice items **encourage guessing**—after all, one option is correct.

Recommendations

When Should Multiple-Choice Items Be Used?

Knowing the strengths and limitations of multiple-choice items can help instructors make better decisions about whether or not to use these items in particular testing situations. Use multiple-choice items for the following instructional goals:

1. When you wish to test the **breadth of student learning**. Multiple-choice items offer the opportunity to sample a greater breadth of learning than do questions that require a lot of student writing. Because they take considerably less time to answer, many more questions can be asked and so more content tested.

2. When you want to test **a variety of levels of learning**. Multiple-choice items are extraordinarily flexible in that they can be used to assess the full range of Bloom's taxonomy (1956). Do not discount multiple-choice when you want to evaluate abilities to think critically and solve problems effectively.

3. When you have **many students who will be taking the test**, then multiple-choice tests are very efficient. If the class is very small in size, it usually is **not** worth the time it will take to construct an effective set of multiple-choice items. Carefully consider whether other item types will serve your testing purposes.

4. When you have **time to construct the test items**. Remember that effective multiple-choice items, which assess more than basic factual knowledge, require a great deal of time and effort to construct. If you do not have the time, another type of test will be a wiser choice.

5. **When time is limited for scoring**, then selected-response items are often the better choice. While it may have taken an hour to construct a multiple-choice item, it will take less than a second to score it.

6. When it is **not** important to determine how well the student can **formulate a correct or acceptable answer**. The answers are definitely provided in multiple-choice items. Even if the question requires critical thinking skills, it may be possible for a student to get the answer right because of clues in the options or by guessing. When it is important for students to formulate their own answers, multiple-choice will **not** do.

Required Preconditions

Before considering specific suggestions for writing multiple-choice items, there are a combination of abilities that, according to Alexander G. Wesman (1971), are necessary to write successful test items.

7. You must **have a thorough mastery of the subject matter** being tested. You must not only understand the implications of the facts and principles of a particular field, but you must also be aware of common fallacies and misconceptions.

8. You must **develop and use a set of educational objectives** to clearly guide your efforts to help students learn. Unless you have carefully considered what you want students to learn, you will not be able to evaluate their progress with any accuracy. This means that you must develop a test plan or table of specifications to guide your item writing. For the vast majority of tests, a two-dimensional table is sufficient. On one dimension, list the areas and subunits of the content you wish to test. On the second dimension, list the various levels of learning you wish to test, for example, understanding, application, and higher-order cognitive objectives. You must also decide what proportion of the test you want to devote to each area of content and each level of learning. Finally, as you write the test items, you should keep a tally of how many items fall into each cell of your total plan to insure that your test actually covers the learning as you originally intended. (See example.)

LEVELS OF LEARNING

Topics	Understanding	Application	Higher-Order
A	5%	10%	10%
B	5%	20%	10%
C	10%	20%	10%

According to this table of specifications, approximately 40 percent of the instruction time was spent on topics "B" and "C" at the application level and 20 percent at the higher-order level. The test should reflect that proportion. (See Gronlund, 1985a or Mehrens and Lehman, 1984, for further treatment of tables of specification.)

9. **Know the students** who will be taking the test in order to appropriately adjust the complexity and difficulty of the items. Sophomores in Introduction 101 may look the same semester after semester, but there are likely to be many differences in the educational backgrounds and intellectual abilities of the groups. Design your test so that the students can demonstrate their learning.

10. You must be a **master of written communication**, able to communicate with precision and simplicity **and** you must use language that the students understand.

Constructing Multiple-Choice Items

The following recommendations for constructing multiple-choice items reflect the collective experience and wisdom of many authors. These recommendations are written in chronological order. Several works are listed in the References and Further Readings section for those of you who wish to read more extensively.

11. **Spread the work across time.** It is unwise to wait until the night before an exam is scheduled to construct the test items. It is impossible to construct effective multiple-choice items in such a limited time. Not only do you need time to construct the items, you need an opportunity to review and revise. If you write a question or two after each class or on a weekly basis, the collection is more likely to be representative of your instruction.

12. **Use note cards for writing the items.** This makes it much easier to file according to your test plan, rearrange, rewrite, and discard items. Better yet, if you have access to a personal computer, use it.

13. Really concentrate on writing items to **evaluate higher levels of thinking.** Avoid the pitfall of writing items that test only memorization of basic factual knowledge. Many instructors (especially those who are writing the test questions just before the test) fall into this trap and pull their students in with them.

14. **Write the stem first.** The stem should present a single, definite problem as a question or incomplete statement. The problem should be one of significance in the course.

15. **Concentrate on evaluating student ability to understand, apply, analyze, synthesize, and evaluate.** It is difficult to write questions that evaluate these higher cognitive levels; but if critical thinking is what you want students to do, you will have to test for it. Students have a tendency to study "what will be on the test" and to study **only** what will be on the test.

16. **State the problem concisely, but completely.** What the student is to answer must be obvious, and the student should be able to discern the problem **without** reading all of the options. A direct question usually does this more clearly than an incomplete statement. There are times, however, when the question is just too convoluted or confusing for easy interpretation; then the incomplete statement may be preferable, or perhaps an item type other than multiple-choice is more appropriate.

17. **Write the stem to include all the information essential to determining the problem,** but omitting irrelevant material that merely serves as padding, unless the student's determination of what is relevant is part of what you want to test.

18. **Avoid unnecessary repetition in the options** by including as much of the item as possible in the stem. This is especially important when using the incomplete statement format. Forcing students to reread a phrase several times wastes time they could put to better use when taking a test.

19. **State the problem or ask the question in a positive form.** The use of negatives can be confusing to even the most intelligent reader, and anxious students often completely miss little words like "not." On those rare occasions when you decide that you must use negatives, use **boldface,** underlining, or CAPITAL letters. Do **not** use double negatives, e.g., negatives in both the stem and the options.

20. **Write the correct or best response after writing the stem.** Be certain that the best response is indeed **best**, that is, would be acknowledged as best by authorities in the field. State this response **as briefly as possible**, and without ambiguities so that all knowledgeable students will read it with the same interpretation. Having colleagues or former students critique your questions for clarity before using them on a test can help to avoid such difficulties.

21. **Avoid making the correct option longer than the distractors.** Test-wise students are very aware of this fault and use this clue to choose the correct answer without knowing the correct answer. The emphasis on the keyed response being absolutely correct sometimes leads to wordiness, and instructors tend to spend much less time developing the distractors which then tend to be shorter. Write the correct response and the distractors, and then compare the lengths. If correct answers are consistently longer (or shorter) as you write multiple-choice items, edit as necessary.

22. **Write the distractors after writing the correct option.** The effectiveness of multiple-choice items can be undermined by the sloppy preparation of the incorrect options. Designing distractors is actually quite challenging because these options must be wrong, yet be plausible enough to attract the attention of students who do not know the material as well as they should.

23. **Make all distractors plausible responses.** Avoid writing poor alternatives just for the sake of having more options; they simply become throwaway options. The criterion is whether or not the distractors test a discrimination that is important; if not, do not use it. Once in a while, a ridiculous option can relieve some of the tension that pervades a testing situation, but only once in a great while.

24. **Be sure that the distractors use words that ought to be familiar to the students.** Using highly technical language or the vocabulary of experts, terms that have not been used in class, forces students to choose correct answers without knowing the meaning of one or more of the options. If the students were not expected to learn the terms, do not include them in the options.

25. **Write distractors that are distinct from each other.** If all the distractors are too much alike, the test-wise student will use this clue to eliminate the group of look-alikes in favor of the dissimilar, correct response. Similar distractors may also indicate that the question should not be presented in the multiple-choice format. Avoid alternatives that overlap or include each other. This error is likely to be distracting to students who read carefully and know the material well, which can result in the more knowledgeable student being penalized by the instructor's lack of item-writing skills.

26. **Critique for general errors in style and format.** Delete any irrelevant clues that could lead a student to select the correct answer or eliminate one or more of the wrong options without knowing the material. Measuring the test-wiseness of the students is not the intent of the test.

27. **Be careful in using specific determiners,** such as "all," "never," "always," or other all-inclusive terms that are more likely to be found in incorrect options. Similarly, qualifiers such as "usually," "sometimes," and "maybe" are more likely to be found in the keyed item. However, sometimes the content permits using absolute specific determiners correctly, and so can keep the test-wise student "honest," e.g., "The president of the United States must always be at least 35 years old" is correct.

28. **Avoid grammatical inconsistencies** between the stem and the options. These are very useful clues for the student who is competent in syntax.

29. **Use "none of the above" as an option with caution.** Some faculty believe that the option, "none of the above," should never be used in a multiple-choice item. This belief is correct for a "best answer" type item. (Nor should options like "all of the above" or "both A and B" be used in "best answer" items.) However, for "correct answer" items, where there definitely is a correct answer, the option, "none of the above," may serve a useful purpose, especially for items requiring mathematical calculations, or perhaps correct spelling or grammar in a language. Using "none of the above" can prevent correct answers

because of guessing, or save students from spending an inordinate amount of time on a problem they cannot solve. To be effective, the option must occasionally be the keyed response; otherwise, the students will see it simply as a throwaway option.

30. **Check once more** to be certain that the correct options are not consistently longer than the alternatives.

31. **Arrange options in a logical order**, if one exists. Numerical answers should be placed in numerical order and dates put in chronological order. Sometimes alphabetizing the options is appropriate.

Organizing the Layout of the Entire Test

Once the individual multiple-choice items are written, you must decide how to organize the groups of items on the test. If you are using several types of items on your exam, be sure to group all of the multiple-choice items together, etc.

32. **List options on separate lines**, arranged in a vertical column to clearly distinguish each option from the others. Printing the responses in tandem or arranging them across the page may save paper, but the result is difficult to read. You should not be testing reading skills.

33. **Use capital letters for the response options** if the student is to write the letter to indicate the selected answer. The handwritten, lower case letters "a and d" and "c and e" can be difficult to distinguish when scoring.

34. **Check to see that the correct answers are distributed randomly among the possible option positions.** If you have had a tendency to choose one position over others, for example, "B," it may become apparent to the test-wise student who seeks out such clues. If necessary, it is easy to rearrange the order of the options to correct this problem.

Interpretive Exercises

Many teacher-made multiple-choice tests pose a series of separate, unrelated questions. In contrast to this, the **interpretive exercise** format presents a series of multiple-choice terms **based on a common stimulus**. The stimulus can be written material, also tables, graphs, maps, pictures, audio- or videotapes, etc. Interpretive exercise items can then be written to assess a wide range of student abilities, for example: to recognize generalizations, assumptions, or interferences; to apply principles; or to interpret data or experimental findings. To achieve this, however, the material must be novel or new to the students, not something previously covered in class or found in the textbook.

In addition to the general advantages of multiple-choice items in testing higher level and complex materials, interpretive exercises minimize the influence of irrelevant information because they confine the data to be interpreted to the material presented. This makes such exercises more difficult to construct; and for written material (the most common form), places heavy demands upon reading skills. Nevertheless, we believe the advantages of interpretive exercises warrant their increased use in college-level tests. (Gronlund, 1985a, has an entire chapter on the interpretive exercise which we strongly recommend for your consideration.)

Center for Faculty Evaluation
and Development
Kansas State University
1615 Anderson Avenue
Manhattan, KS 66502-4073

Conclusions

We have focused this paper on multiple-choice items because we are convinced that they permit testing higher levels of learning which are appropriate to college but which often are not tested by teacher-made tests (including essay as well as selected response tests). We are not suggesting that other forms of selected response items, e.g., true-false and matching, are inappropriate, but we have omitted them because of space limitation. Several standard textbooks in the References and Further Readings section give detailed suggestions for designing such items. Nor are we suggesting that multiple-choice items should, or can, replace essay tests (a subsequent IDEA Paper will be devoted to essay tests). What we are suggesting is that many teacher-made multiple-choice tests can be significantly improved. We hope that this paper will be of some help to readers in achieving that improvement.

References and Further Readings

Those references below which are followed by an asterisk are standard texts on educational measurement. Each has one or more chapters on multiple-choice and other selection items, as well as chapters on other aspects of testing and grading.

Bloom, B. S., et al. (1956). **Taxonomy of educational objectives: Cognitive domain**. New York: David McKay.

Campbell, J. R. (1972). **In touch with students**. Columbia: Kelly Press.

Clegg, V. L., & Owens, R. E. (1984). **Tips for writing tests**. Manhattan, KS: Graduate Services and Publications.

Dressel, P. L. (1976). **Handbook of academic evaluation**. San Francisco: Jossey-Bass.

Ebel, R. L., & Frisbie, D. A. (1986). **Essentials of educational measurement** (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.*

Eble, K. E. (1976). **The craft of teaching: A guide to mastering the professor's art**. San Francisco: Jossey-Bass.

Educational Testing Service. (1973). **Multiple choice questions: A closer look**. Princeton, NJ: Author.

Fuhrmann, B. S. & Grasha, A. F. (1983). **A practical handbook for college teachers**. Boston: Little, Brown.

Gronlund, N. E. (1985a). **Measurement and evaluation in teaching** (5th ed.). New York: Macmillan.*

Gronlund, N. E. (1985b). **Stating objectives for classroom instruction** (3rd ed.). New York: Macmillan.

Lowman, J. (1984). **Mastering the techniques of teaching**. San Francisco: Jossey-Bass.

McKeachie, W. J. (1986). **Teaching tips: A guidebook for the beginning teacher** (8th ed.). Lexington, MA: D. C. Heath.

Mehrens, W. A., & Lehmann, I. J. (1984). **Measurement and evaluation in education and psychology** (3rd ed.). New York: Holt, Rinehart and Winston.*

Milton, O. (1982). **Will that be on the final?** Springfield, IL: Charles C. Thomas.

Milton, O., & Associates (1978). **On college teaching: A guide to contemporary practices**. San Francisco: Jossey-Bass.

Roid, G. H., & Haladyna, T. M. (1982). **A technology for test-item writing**. New York: Academic Press.

Thorndike, R. L., & Hagen, E. (1969). **Measurement and evaluation in psychology and education**. New York: Wiley & Sons.

Wesman, A. G. (1971). Writing the test item. In R. L. Thorndike (Ed.) **Educational measurement**. Washington, D. C.: American Council on Education.