

Trends

Belverd E. Needles, Jr.

Evaluating Multiple–Choice Questions in Electronic Test Banks

Today, most first-year accounting teachers rely on the electronic test banks provided by publishers rather than making up their own examinations. How do we know how good these questions are? Most of us do not ask. But it is an important question. My recommendation is to apply the criteria below to a sample of questions from the publishers test bank. The procedure is simple:

- Choose randomly five questions from each test you plan to use.
- Rate each test question “1” if you can answer the evaluation question below “yes” for it; “0” if you answer “no.”
- Since there are five test questions and ten evaluation questions, a perfect score indicating a “good” test is 50; lowers scores indicate a lower quality test.

Score

1. **Does each question measure significant concepts and principles, avoiding items that cover trivia?**

The primary purpose of test items is to measure achievement of the learning objectives of the course. Good multiple-choice items deal with important ideas that have been covered in the course. Trivia, as used here, does not mean “easy”. If an important concept in a course is easy, there is no reason not to include test questions related to it.

2. **Does each question measure significant concepts and principles, avoiding items that cover trivia?**

The primary purpose of test items is to measure achievement of the learning objectives of the

course. Good multiple-choice items deal with important ideas that have been covered in the course. Trivia, as used here, does not mean “easy”. If an important concept in a course is easy, there is no reason not to include test questions related to it.

3. **Does each question present a single problem in the stem, and include a verb in the stem?** _____

If multiple problems are presented in a single stem, it is difficult to know what is misunderstood by students who miss the question. If an instructor wishes to test students comprehension of three concepts, it is better to use three separate questions than to use one that includes all three concepts. The task set forth should be so clear that the student can understand it without reading alternatives.

4. **Does each question include one correct or clearly best answer and three or four plausible distractors?** _____

Distractors should be limited to those that really are plausible. Research indicates that three choices are as effective as four; however, never exceed five choices. Reading distractors, even implausible ones, takes students’ time. The distractors must be incorrect, but they must be plausible enough to attract the students who do not know the materials very well.

5. **Is the question (stem) stated in clear and concise language, including all, but only relevant material?** _____

Simple, clear language; free of unnecessarily complex wording and sentence structure should be used. Attempt to keep the stem short by not loading the question with irrelevant materials. Stems can often be more clearly written if two sentences are used. One sentence should contain background material and the other should ask the question.

6. **Are clues to the right answer avoided?** _____

Common clues include: grammatical form, verbal association, length of choice-options, and use of specific determiners. The distractors should be similar in length and parallelism in structure of distractors versus the correct choice. The more homogeneous the options, the more discrimination required on the part of students and more difficult the item.

7. **Are positively stated stems used where possible; otherwise, is the student's attention called to the negative?** _____

Negative working, such as “which of the following is **not** ...” tends to confuse students. Place the student in the right frame of mind before they read the alternatives. When negatives are used be sure to underline or boldface the negative.

8. **Are questions that use “all of the above” and “none of the above” avoided?** _____

These responses are often used when there is difficulty in creating a sufficient number of distractors. “None of the above” should **never** be used with negatively stated stems because of confusion caused by the double negative.

9. **Are the position of correct answers varied in a random manner?** _____

The correct answer should appear in each position about an equal number of times, however, the placement should not follow a pattern that may become apparent to students.

10. Are options arranged in a logical order?

Are list words in alphabetical order; numbers in ascending or descending order; and dates chronologically? If no logical order is possible, are answers randomized to avoid patterns in correct responses. With accounting exams it is often difficult to list numerical choices in a logical order without violating criteria.

11. Are the questions independent of each other?

No answer to one MC question should be dependent on the answer to another question.

This procedure does not take long and can be quite revealing. By choosing the questions randomly, it gives a fair picture of the overall test bank. I have been surprised at the number of inadequate questions I have found. Sometime the problem can be fixed by editing the questions, but sometime they should be discarded and other questions chosen. When giving two versions of the same test to a class, do not inadvertently create “sequencing bias.”

The order in which items are arranged on a test can also affect the mean score. If two or more versions of a test will be given to a group of students whose grades are going to be compared, randomize all versions of the tests. If some students receive a version on which items are systematically ordered while other students receive a version that is randomly a ranged, one group may be at a disadvantage.