

# Multiple-Choice Tests: Getting Beyond Remembering

David DiBattista  
Brock University

If there is one characteristic that all teachers seem to share, it is that we want our students to be successful learners. To help them reach their potential as students, we devote our time and energy to preparing engaging classroom activities, selecting appropriate reading materials, and creating interesting and challenging assignments. We also encourage them to attend class, keep up on the course readings, and complete their assignments on time and to the best of their ability. Of course, part of our responsibility as teachers is to assess the extent to which the students in our courses have successfully learned the facts, concepts, procedures, and skills that have been taught. To achieve this goal, we may use a variety of techniques, such as classroom tests, in-class presentations, group projects, and term papers. When these assessments are carefully planned and carried out, we are able to assign course grades that fairly reflect our students' learning.

## MULTIPLE-CHOICE TESTING: A CRITICISM

In most college and university courses, classroom testing is the primary form of assessment, and in larger courses, which may have hundreds or even thousands of students, multiple-choice (MC) items often make up the lion's share of the classroom tests. Experts in assessment generally agree that when MC items are used properly, they can be a highly effective means of assessing student achievement (Haladyna, 2004; McDonald, 2007). Nevertheless, MC testing is often looked down on, and many teachers use MC items on their tests only with great reluctance, believing them to be a necessary evil that must be tolerated in large classes.

Perhaps the most serious criticism of MC testing is that it puts the emphasis on being able to remember information, and that it neither assesses nor encourages the use of higher-level thinking (Mitchell, 1992; Walsh & Seldomridge, 2006), such as understanding, applying, analyzing, and evaluating information. Unfortunately, this criticism seems to have a substantial basis in reality. In a recent analysis of classroom tests used in a university setting, Krista Lewis and I found that well over half of the MC items called upon students to remember information that they had encountered during the course and did not require them to engage in any form of higher-level thinking (DiBattista & Lewis, 2008). Furthermore, in some test banks that I have examined, more than three-quarters of the items require students to do nothing more than remember information!

The predominance of memory-based items on MC tests has some important and unfortunate consequences. Because students so frequently encounter memory-based MC items on the tests that

they take, it is perhaps not surprising that they tend to use a surface-learning approach rather than a deep-learning approach when studying for a MC test—that is, they concentrate on remembering material that they have encountered in the course, and they devote relatively little attention to understanding, analyzing, and interpreting course material (Scouller, 1998). In short, students seem to learn through experience that the best way to prepare for a MC test is to focus on remembering information, and they tailor their studying to achieve this goal. In contrast, students perceive essay tests to involve higher-level thinking, and they therefore tend to use a deep-learning approach when studying for essay tests (Scouller, 1998). Because factors that promote surface learning rather than deep learning generally lead to lower quality learning outcomes (Trigwell & Prosser, 1991), the extensive use of memory-based MC items and the surface-learning study strategy that it engenders in students are clearly undesirable.

Having recognized the seriousness of this problem, Nelson Education Ltd. has recently taken a significant step to address it. Specifically, as part of the Nelson Education Testing Advantage Program, they have made a commitment to ensure that their MC test banks contain a substantial number of items that assess higher-level thinking. To accomplish this goal, they have instituted an extensive training program for test bank authors to provide them with the knowledge and skills that they need to create well-structured, effective MC items that assess higher-level thinking rather than simple remembering. Of course, memory-based items are still included in the Nelson test banks, but they are far less prominent than is normally the case in publisher-supplied test banks.

As you look through your Nelson test bank to decide which MC items to include on your tests, you can examine items to determine whether they are memory-based or require your students to engage in higher-level thinking. By making your selections appropriately, you can construct tests that contain a mix of items that reflects your personal instructional goals. For a great many teachers, this will mean being able to construct MC tests that are more strongly weighted toward assessing higher-level thinking, with memory-based items taking more of a back seat than is usually the case.

Let's now look more closely at how memory-based MC items differ from items that call upon higher-level thinking. When we have done this, we will consider some ways that you can help your students to be well prepared to take MC tests that put the emphasis on higher-level thinking rather than remembering.

### GETTING BEYOND REMEMBERING

Question 1, shown in the box below, serves to illustrate a typical memory-based MC item. To answer this question correctly, students do not really have to know anything at all either about the details of classical conditioning or about how it works. By doing nothing more than remembering the definition for unconditioned stimulus, they can earn full marks for this item. Indeed, students

who cram the night before the test might very well memorize the definition for unconditioned stimulus, answer Question 1 correctly, and then promptly forget the information immediately afterward!

**Question 1.** In classical conditioning, what name is given to a stimulus that elicits a particular response even in the absence of any prior training?

- A. conditioned stimulus
- B. activational stimulus
- C. unconditioned stimulus\*
- D. discriminative stimulus

Of course, not all memory-based items involve definitions. The following questions illustrate how MC items can be used to test students' memory for facts, concepts, and procedures.<sup>1</sup>

- Who wrote *On Death and Dying*, in which the five stages of grief are described?
- For how long must the relevant symptoms be present if a person is to be diagnosed with a major depressive episode?
- In an experiment conducted by Bloggs (2004), participants were given one of four different treatments. Which treatment group showed the most improvement?
- Which of the following is a major component of the trichromatic theory of colour vision?
- If you have nine test scores, how would you compute the median?
- Which of the following procedures is used for making *post hoc* comparisons?

It is likely that memory-based MC items such as these are so commonly found in test banks because they can be created quickly and easily, focusing as they so often do on individual pieces of information. However, higher education involves much more than just remembering information, and it is therefore important that a substantial number of the MC items on tests be used to assess higher-level thinking—that is, that they test something more than students' ability to remember. As it happens, MC items can be used very effectively to assess higher-level thinking (Jacobs & Chase, 1992; McDonald, 2007), although the creation of such items certainly requires more skill than does the creation of memory-based items.

The distinction between memory-based items and those calling on higher-level thinking is nicely captured in the following quotation from Anderson and Krathwohl's work on the Revised Bloom's Taxonomy: **"If assessment tasks are to tap higher-order cognitive processes, they must require that students cannot answer them correctly by relying on memory alone"** (2001, p. 71, emphasis in the original). Let's now look at Question 2, which provides an example of a MC item

---

<sup>1</sup>I have included here only the question-portion of these MC items. In the interests of saving space, the several options that would normally be part of a MC item have been omitted.

that meets this critical criterion. Like Question 1, Question 2 deals with classical conditioning, and you will notice that both items have the same four options and the same correct answer. As we saw earlier, Question 1 is a memory-based item—that is, students can answer it correctly by relying on memory alone. In contrast, Question 2 requires students to do much more than merely remember information. Rather, they must have some understanding of the processes and relationships involved in classical conditioning, and furthermore they must be able to classify the role of Drug X in a research setting that they are seeing for the first time. Thus, students who have merely memorized the definition for unconditioned stimulus may well have difficulty with Question 2, whereas students who have truly learned about classical conditioning and understand how it works will likely be able to answer this question with little difficulty.

**Question 2.** Right after a rat smells menthol, it is always given Drug X, which reliably induces substantial water intake. Eventually, the rat drinks water whenever it smells menthol, even when it is not injected with Drug X. In this situation, what is the role of Drug X?

- A. conditioned stimulus
- B. activational stimulus
- C. unconditioned stimulus\*
- D. discriminative stimulus

Of course, in presenting Question 2 as an example of an item tapping higher-level thinking, I have assumed that students are seeing this question for the first time as they write the test. If they were to be familiarized with the situation presented in Question 2 and with the correct answer *before* taking the test, Question 2 would be reduced to a memory-based item, and its value for assessing higher-level thinking would disappear completely.

The following questions further illustrate how MC items can be used to test students' knowledge of course material while engaging them in higher-level thinking. None of these questions could be answered simply by relying on memory; rather, each of them calls on students to do more than simply remember information that they have read or heard in the course.

- Which of the following would be an example of positive feedback?
- The mean of a mound-like distribution of scores is 65 and the variance is 100. You can be pretty sure that about two-thirds of the scores lie between what two values?
- In what way are a neuron and a battery similar to each other?
- Which of the following best explains why the  $z$ -test for independent samples would be used less often than the corresponding  $t$ -test?
- A psychologist wants to compare the effectiveness of two training methods that are commonly used to teach people to perform a certain task. He obtains a group of 40 naive

volunteers and randomly assigns each person to one of the two training methods. He sets alpha at 0.05 and determines that the power of his statistical test is 0.40. Which of the following, if any, is a valid criticism of this research study?

- A researcher is planning an experiment in which each participant will receive both Treatment A and Treatment B. He is now thinking about the order in which the two treatments will be given to participants. Which of the following procedures would be most advisable?

It bears repeating at this point that an underlying assumption here is that students would be seeing the above questions for the first time as they write the test. If students have not seen these questions prior to the test, then they would not be able to answer them merely by relying on their memory, but would instead have to engage in some form of higher-level thinking while taking the test. On the other hand, if students have seen these questions previously, they might answer them correctly when taking the test merely by remembering the correct response, in which case these would essentially be memory-based questions.

Up to now, we have focused our attention on the distinction between memory-based and higher-level MC items. However, it is also important to keep in mind that different higher-level items may call on different cognitive abilities. One approach that I have found particularly useful in classifying the cognitive processes that students must demonstrate when taking a test is that of Anderson and Krathwohl (2001), who have presented a revision of the well-known Bloom's taxonomy of educational objectives (Bloom, 1956). An underlying premise of the Revised Bloom's Taxonomy is that learning situations call upon students to use a variety of cognitive processes in order to deal effectively with several different types of knowledge. Thus, the taxonomy is constructed along two dimensions—a cognitive process dimension and a knowledge dimension. Key information about both of these dimensions is presented in Table 1, but our focus here will be only on the cognitive process dimension.

As Table 1 indicates, the cognitive process dimension has six categories—*Remember*, *Understand*, *Apply*, *Analyze*, *Evaluate*, and *Create*—which are arranged along a continuum reflecting increasing cognitive complexity. As we have seen, a great many MC items are memory-based and would therefore fall in the *Remember* category at the bottom end of the cognitive process dimension. Other MC items call on higher-level thinking and will fall into the *Understand*, *Apply*, *Analyze*, or *Evaluate* categories. For example, let's think again about Question 2, which we considered earlier. To answer this question, students must examine the research scenario that is presented and determine that Drug X serves as the unconditioned stimulus. This item therefore falls in the *Understand*

category because it involves *classifying* information. Question 3, 4 and 5, which are shown on the following page, provide examples of items that would be in the *Apply*, *Analyze*, and *Evaluate* categories, respectively.

**Table 1. Key Terms in the Revised Bloom's Taxonomy**

**COGNITIVE PROCESS DIMENSION**

1. REMEMBER: Retrieving relevant knowledge from long-term memory
  - Recognizing; Recalling
2. UNDERSTAND: Determining the meaning of instructional messages, including oral, written, and graphic communications
  - Interpreting; Exemplifying; Classifying; Summarizing; Inferring; Comparing; Explaining
3. APPLY: Carrying out or using a procedure in a given situation
  - Executing; Implementing
4. ANALYZE: Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose
  - Differentiating; Organizing; Attributing
5. EVALUATE: Making judgments based on criteria and standards
  - Checking; Critiquing
6. CREATE: Putting elements together to form a novel, coherent whole or make an original product
  - Generating; Planning; Producing

**KNOWLEDGE DIMENSION**

- A. FACTUAL KNOWLEDGE: The basic elements that students must know to be acquainted with a discipline or solve problems in it
  - Knowledge of terminology
  - Knowledge of specific details and elements
- B. CONCEPTUAL KNOWLEDGE: The interrelationships among the basic elements within a larger structure that enable them to function together
  - Knowledge of classifications and categories
  - Knowledge of principles and generalizations
  - Knowledge of theories, models, and structures
- C. PROCEDURAL KNOWLEDGE: How to do something; methods of inquiry, and criteria for using skills, algorithms, techniques, and methods
  - Knowledge of subject-specific skills and algorithm
  - Knowledge of subject-specific techniques and methods
  - Knowledge of criteria for determining when to use appropriate procedures
- D. METACOGNITIVE KNOWLEDGE: Knowledge of cognition in general as well as awareness and knowledge of one's own cognition
  - Strategic knowledge
  - Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
  - Self-knowledge

Adapted from Anderson & Krathwohl (2001)

**Question 3.** Working with an ordinal data scale, Jeff obtained the following five scores: 0, 0, 2, 5, 8. What is the value of the most appropriate measure of central tendency for this set of scores?

- A. 0
- B. 2\*
- C. 3
- D. 5

**Question 4.** A rat has been trained to press a bar to receive small pieces of food. The food dispenser has now been disconnected, but the rat is still hungry. Which of these will have the greatest influence on how long the rat will continue to press the bar?

- A. the reinforcement schedule used during training\*
- B. the total amount of food the rat ate during training
- C. the number of times the rat pressed the bar during training
- D. the amount of energy the rat expended to press the bar during training

**Question 5.** You have carried out a 3×2 ANOVA for independent groups. There were 60 participants, with 10 participants randomly assigned to each cell. You have now analyzed the data and are double-checking your work. Which of the following would immediately let you know that you have made an error?

- A. You found the total degrees of freedom to be 60.\*
- B. You found the mean square for the error term to be 6.25.
- C. You found the *F*-statistic for the interaction effect to be 2.34.
- D. You found degrees of freedom for the interaction effect to be 2.

At the upper end of the cognitive process dimension, we find the category *Create*. MC items involve having students *select* a response from a list of options that are provided rather than to *generate* a response spontaneously. Because MC items provide no opportunity for students to engage in any kind of creative act, they cannot be used to assess creativity. This inability to assess students' creativity is an inherent limitation of MC testing, and other techniques, such as essays and projects, must be used for this purpose.

As a teacher, you may find it quite helpful to familiarize yourself with the Revised Bloom's Taxonomy. With a bit of practice, you will be able to determine the cognitive level of the MC items on your tests, and this will allow you to more precisely match the content of your tests to your personal instructional goals. In addition, you will find that familiarity with the Revised Bloom's Taxonomy will pay dividends in other aspects of your teaching, such as in planning your courses, setting learning objectives, and preparing all types of assignments and tests.

### PREPARING STUDENTS FOR TESTS

I mentioned earlier that students often use a surface-learning approach when studying for MC tests. This strategy ordinarily pays off for them because MC tests that they normally encounter

consist primarily of memory-based items. However, a MC test that you put together using items from a Nelson Education test bank may have a substantially greater emphasis on higher-level cognitive processing than your students are accustomed to. Consequently, if your students were to use the usual focus-on-remembering study strategy to prepare for your test, they might not do well because of the poor fit between their study strategy and the nature of your test. It is therefore important to provide your students with proper guidance about how to study for the tests that they will take in your course. Now I am not talking here about so-called “teaching to the test,” which involves giving special attention during class time to the specific questions that will later be asked on a test. Teaching to the test greatly diminishes the value of the tests that we give, making them virtually meaningless as valid measures of achievement. For this reason, teaching to the test should always be avoided. In contrast, it is always appropriate to give students proper guidance about how to study for tests and to give them opportunities to develop their course-related skills and abilities.

In the courses that I teach, the overwhelming majority of the MC items involve higher-level thinking. For this reason, I have thought quite a lot about how to prepare my students for the MC items on my tests. The suggestions that I will provide below are based both on my own experience and on a broad reading of the literature on teaching and learning. I think you may find them helpful in your own teaching and hope that you will give them serious consideration.

In my experience, I have found that a good first step in helping students to prepare for my tests is to discuss with them early in the term the difference between memory-based MC items and those that require higher-level thinking. In doing this, I keep in mind that most of my students have never given much, if any, thought to the distinction between memory-based and higher-level MC items. Even for the brightest and most thoughtful students, the most important aspect of a MC item is usually whether it is easy or difficult, and they are not likely to pay any attention at all to the nature of the cognitive processes that are involved in answering it. I therefore begin the discussion by showing my class several memory-based items, explaining why these items would not require them to do anything more than remember. I make sure to mention that most of the MC items that they have previously encountered on tests have almost certainly been of the memory-based variety, and that they therefore may have found in the past that using a focus-on-remembering study strategy has worked reasonably well for them. Then by way of contrast, I show them several items that involve higher-level thinking and explain why these items call on them to do much more than simply remember information that they have read or heard in the course. I emphasize that the tests they will be taking in my course will contain a great many items of this sort, and I also stress that they will be wise to use a study-for-understanding (i.e., deep learning) approach rather than a focus-on-remembering strategy when preparing for my tests.



I have found that it is not sufficient to tell my students just once about the somewhat unusual nature of my MC tests and the importance of their using a study-for-understanding approach when preparing for them. To increase the likelihood that they will actually *use* a study-for-understanding approach when preparing for tests, I reinforce the message during class time at various points during the term. One way I do this is to have MC items serve as the basis of a classroom assessment technique.<sup>2</sup> For example, after having covered a particular topic in class, I will sometimes present students with several MC items that they should now be ready to answer, with the majority of them typically being higher-level items. I let students use a think-pair-share approach to work through the items. Thus, I first give the students a couple of minutes during which they can work on the questions on their own, and then I allow them to work together in pairs for a couple of minutes, discussing why they agree or disagree on the answer to each item. Finally, I call on students to share their answers and their reasoning with the rest of the class, and if necessary I provide further explanation and clarification to reinforce students' learning.

This use of MC items during class time gives my students a valuable opportunity to check their learning, and by hearing what they have to say about the questions, I am able to judge how well the class has grasped the information that has been presented. In addition, while going over the items with the class, I often make a special point of asking them to examine each item to determine whether it is a memory-based or a higher-level item. I then remind the class that most of the items that they will encounter on later tests will engage them in higher-level thinking and that they must therefore study accordingly. Thus, I use this classroom assessment technique not only to provide a check on learning but also to give students practice with the type of questions that will later be on their tests. Furthermore, the reminder that most of the items on tests will assess higher-level thinking reinforces the message that their test preparation should be based on a study-for-understanding approach rather than a focus-on-remembering approach.

The use of electronic student response systems, which seem to be increasing in popularity year by year, is entirely compatible with the strategy that I have described above. With these systems, each student in the classroom has a remote control device—often referred to as a “clicker”—and by pressing the appropriate button, students can enter their answers to MC items into the system. After students have finished responding to an item, the teacher can display a summary of the responses, showing how many students answered the question correctly and how many selected each of the incorrect options. Clickers, which are available from Nelson Education, provide an enhanced level

---

<sup>2</sup>A primary purpose of classroom assessment techniques is to monitor how well students are learning during the time in between their formal assessments. See Angelo and Cross (1993) for an excellent compendium of classroom assessment techniques.

of feedback to both the students and the teacher and when used wisely they can be a worthwhile instructional tool (Reay, et al., 2007).

Another good way to prepare students for the classroom tests you will give them is to have them take online MC quizzes during the term. People often think of quizzes and tests as being useful only to assess achievement, but research has shown that frequent testing can actually promote learning (Bangert-Drowns, et al. 1991; Leeming, 2002). This may be because when tested frequently, students study more, space their studying out across the term rather than concentrating it before major exams, and have more opportunity to obtain feedback on their performance.

In a full-year course that I teach, my students take six 10-item online MC quizzes that count for two marks each. Students access the quizzes on Sakai, which is the learning management system currently used at my university. Each quiz is available online for 10 days, and students may take the quiz at any time that is convenient for them. Once logged in, they have 15 minutes to complete the quiz, which is typically plenty of time for a 10-item quiz. Moreover, students can take each quiz as many times as they want; however, because the items appearing on a particular quiz have been randomly sampled from a pool of about 80-100 items, they do not get the exact same set of questions every time they take the quiz.

The quizzes include a substantial number of higher-level MC items of the sort that students will later encounter on their classroom tests. Because I think of these quizzes as being primarily for promoting learning rather than for assessing achievement, I strongly urge students to work in groups of two or three when taking them. Thus, I suggest that one student in a group—the quiz-taker—should log into the system and that all group members should then discuss each item before the quiz-taker enters the response. Group members can then take turns being the quiz-taker, with each working toward eventually achieving a perfect score of 10.

Although the online quizzes in my course count for very few marks, they play several important roles. Not only do they give students experience with the type of higher-level MC items that they will encounter on tests, but they also encourage students to study and review course material more frequently than they otherwise might. In addition, they provide continuing opportunities for collaborative learning and thereby promote cooperation among students.

The strategies that I have described here encourage students to use a study-for-understanding approach rather than a focus-on-remembering approach as they prepare for their MC tests. In addition, they help students to be better prepared for tests by giving them opportunities to practice on the type of higher-level MC items that they will actually encounter on their tests. Over the years, these strategies have proven to be quite successful in my classrooms, and I encourage you to consider including some or all of them in your teaching.

## References

- Anderson, L. W., & Krathwohl, D., Eds. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's Taxonomy of educational objectives*. New York: Longman.
- Angelo, T. A., & Cross, K. P. (1993). *Classroom Assessment Techniques*, 2<sup>nd</sup> edition. San Francisco, CA; Jossey-Bass.
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85, 89–99.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York : McKay.
- DiBattista, D., & Lewis, K. (2008). *What cognitive processes are engaged by the multiple-choice questions on university tests?* Unpublished manuscript.
- Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items*, 3<sup>rd</sup> edition. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Jacobs, L. C., & Chase, C. I. (1992). *Developing and Using Tests Effectively*. San Francisco: Jossey-Bass.
- Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, 29, 210–212.
- McDonald, M. E. (2007). *The Nurse Educator's Guide to Assessing Learning Outcomes*, 2<sup>nd</sup> edition. Sudbury, MA: Jones and Bartlett.
- Mitchell, R. (1992). *Testing for Learning: How New Approaches to Evaluation Can Improve American Schools*. New York: Free Press.
- Reay, N. W., Li, P., & Bao, L. Testing a new voting machine question methodology. *American Journal of Physics*, 76, 171-178.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: multiple choice examination vs. essay assignment. *Higher Education*, 35, 453-472.
- Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes. *Higher Education*, 22, 251-266.
- Walsh, C. M., & Seldomridge, L. A. (2006). Critical thinking: Back to square two. *Nursing Education*, 45, 212-219.