

Writing and Evaluating Effective Multiple Choice Tests

When we gather the results of multiple-choice exams, we may notice questions many students answered the same way, but incorrectly. How do we interpret this? Is there a good time to throw out a question? This week's teaching tip focuses on some resources to help us create and use better multiple-choice exams, but can also apply to other forms of assessment.

Name: _____

Date: _____

Multiple-choice exams are often part of the assessment repertoire of many faculty because they are easy to grade. But writing good multiple-choice tests is hard to do for a few reasons:

- ☐ A **Most of us have had no training whatsoever in creating these kinds of assessments.**
Asking questions and getting answers is not hard. Asking good questions that get students to think the way you intend—that is hard.
- ☐ B **Students have developed really good test-taking skills.**
Many students know to eliminate nonsense answers to narrow choices with the best odds. But we want our tests to make students think more carefully about content!
- ☐ C **It is easy to forget what we are measuring when we use multiple-choice tests.**
Rather than just looking at norm-referencing rules for good questions, quantitatively evaluate how the class as a whole approached questions.
- ☐ D **all of the above**

For resources on how to write effective multiple-choice questions and guides on how to measure their accuracy in assessing learning, see the full teaching tip at

oakland.edu/teachingtips

Multiple-Choice Tests: When do you throw out a question?

In the last couple of weeks, I have had multiple faculty approach me asking about their multiple-choice tests that they have given in their classes and specifically, asking when to get rid of a question based upon student responses. This week's teaching tip focuses on some resources to help us create and use better multiple-choice exams but the information included applies to all types of assessment.

Multiple-choice exams are often part of the assessment repertoire of many faculty because they are easy to grade. But writing good multiple-choice tests is hard to do. I think there are a couple reasons that make this so:

1. Most of us have had no training whatsoever in creating these kinds of assessments.

When I was in grad school, we had a joint doctoral program between Exercise Science and Science Education. My Exercise Science department head gathered all of the doctoral students together to ask us what we thought the value of the education side was. Among the only people to speak up, I asked my department head how he knew if he was asking good multiple-choice questions. He responded that he kept asking the same ones for three years and threw out those where students couldn't answer correctly. He said it wasn't hard. He was right. Asking questions and getting answers is not hard. Asking good questions that get students to think the way you intend, now that is hard. Needless to say, I finished my Ph.D. in Science Education.

There are many, many resources about MC tests out there from some very quick and applied papers (i.e., http://www.theideacenter.org/sites/default/files/Idea_Paper_16.pdf) to full books and research articles (i.e.,

<http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=81790701-e732-4a68-9e0c-993437437ef1%40sessionmgr111&vid=4&hid=122>).

2. Students have developed really good test-taking skills.

As a native New Yorker, I grew up taking Regents exams – tests at the end of the year in science, math, foreign language, English, social studies, etc. In four years of high school, we took 11 or 12 of these tests and we bought these books teaching us how to take and pass the tests. Our students today have likely taken many more tests than I or you would have and may have even been privy to the prep courses that prepare people for the SAT, ACT, GRE, LSAT, MCAT or any of the plethora of multiple-choice laden tests. They know the drill. Read the choices. Eliminate the choices that make no sense with the others. You can probably narrow down the choices to two. This is not what we envision when we give a test! We want students to think! So, we need to eliminate the ability for students to do well on test taking skills alone. The BYU guide for writing MC questions has been around a long time but I think it is still one of the best guides out there for how to construct good questions:

<http://testing.byu.edu/info/handbooks/betteritems.pdf>

3. It is easy to forget what we are measuring when we use multiple-choice tests.

I have been approached in the last couple weeks by faculty telling me that they heard that they should throw out MC test questions if 50% of the students get the question wrong (I will explain in the next paragraph where this comes from). Another faculty told me that the value was 65% (I believe this is slightly confused with accepted value for how reliable a question is – a way of analyzing your tests). Now, these numbers are not incorrect but they need the proper context around them.

For instance, if you are using a MC test to identify the top performers in your class (this is also known as norm-referenced testing), then it may be proper to write a test where 5% of the items are answered correctly by 90% of the students (to boost confidence), 5% of the test items are answered correctly by 10% of the students, and the remainder of the items are answered correctly by an average of 50% of the students (Davis, 2009). This is where I believe the 50% number comes from.

Certainly, there are many ways to quantitatively evaluate your tests but it is important to recognize that it is not the only way.

If you are using a MC test to measure if students are using information, skills, and competencies (like critical thinking) that you want all students to have acquired, you are testing for something different – how well the test questions represent the things you want them to do. In this case, when students perform poorly on test questions, there are multiple possibilities: was the test item unclear or poorly written? Was the content of the question too challenging? Were the students insufficiently prepared? Looking at the choices that students made in a bar graph format will give you some insight as to how students were thinking when they answered. Here, if a good number of your students chose the same answer, whether it was the right answer or a wrong one, it would be indicative that the thinking students used was similar and that the question posed was a good question at measuring that way of thinking. It is your call as to whether that was the kind of thinking you desired to have them do.

There are many resources on campus and online to assist you in these questions and the quest to write better multiple choice tests.

Resources:

Clegg, V.L. and Cashin, W.E. (1986) "Improving Multiple Choice Tests." Idea Paper. No. 16. Found online at: http://www.theideacenter.org/sites/default/files/Idea_Paper_16.pdf

Davis, B.G. (2009). Teaching Tools. 2nd Edition. San Francisco: Jossey-Bass. Available in the CETL.

Jacobs, L.C. and Chase, C.I. (1992). Developing and Using Tests Effectively: A Guide For Faculty. San Francisco: Jossey-Bass. AVAILABLE IN THE LIBRARY AT: <http://lumen.regis.edu/search~S3/?searchtype=t&searcharg=Developing+and+Using+Tests+Effectively%3A+A+Guide+For+Faculty&searchscope=3&SORT=D&extended=0&searchlimits=&searchorigarg=ttps+for+improving>

Kehoe, J. (1995). "Writing Multiple-Choice Test Items." Practical Assessment, Research and Evaluation. 4 (9). Full text available through the library: <http://pareonline.net/getvn.asp?v=4&n=9>

Lowman, J. (1995). Mastering the Techniques of Teaching. San Francisco: Jossey-Bass. Available in the CETL.

Sehcrest, L., Kihlstrom, J.F., and Bootzon, R. (1999). How to Develop Multiple-Choice Tests. IN B. Perlman, L.I. McCann, S.H. McFadden (Eds.), Lessons Learned: Practical Advice for the Teaching of Psychology. Washington, D.C.: American Psychological Society.

Wergin, J.F. (1988). "Basic Issues and Principles in Classroom Assessment." In J.H. McMillan (Ed.), Assessing Students' Learning. New Directions for Teaching and Learning, No. 34. San Francisco: Jossey-Bass. Available through Prospector.

Submitted by:

Ken Sagendorf, Ph.D.

Director, Center for Excellence in Teaching and Learning (CETL)

Professor

Regis University

www.regis.edu