

.....
*A forum for the
discussion of
teaching
enhancement,
provided by the UW-
Madison Teaching
Academy*
.....

The Learning Link

In This Issue

This issue of The Learning Link offers a cross-section of reflections on teaching and learning at UW-Madison. Responses and submissions are welcomed and encouraged; please see page two for contact information.

.....

Helpful Tips for Creating Reliable and Valid Classroom Tests: Writing and Scoring Essay and Short-Answer Questions

by James A. Wollack

Pages one-three

Does Investment in Higher Education Pay for States?

by Philip Trostel

Pages four-five

Web Resources for Instructional Development

Page five

Upcoming events

Page six

Helpful Tips for Creating Reliable and Valid Classroom Tests: Writing and Scoring Essay and Short-Answer Questions

In the most recent article in this series, I discussed some strategies for developing multiple-choice (MC) questions. MC items are popular for classroom tests because they are easy and efficient to score and they allow instructors to assess students with respect to many different course objectives. However, when it is desirable to assess whether students possess a rich understanding of particular material, MC is not the preferred item format. Instead, it is much more desirable to present students with a problem to solve, and to evaluate students with respect to the processes they used to solve the problem. Examples of item types measuring deep understanding include essay and short-answer questions.

Essay and short-answer items, sometimes referred to as constructed-response (CR) items, are useful when instructors are interested in learning how students arrive at an answer. In this type of a question, students decide how to approach the problem, how to set it up, what factual information or opinions to use, how much emphasis to devote to various parts, and how to specifically express their answer. Obviously, assessing writing ability is best done using an essay response format, but other examples of situations well-suited for CR items include solving math or science problems, comparing and contrasting opposing viewpoints, recalling and describing important information, developing a plan for solving a problem, and criticizing or defending an important theory, just to name a few.

Whereas the MC item is easy to score but difficult to write, most CR items are exactly the opposite. Writing good CR items still requires careful work, but the process is comparatively easy. The most difficult aspect of administering CR questions is grading them in a consistent and fair manner. The rest of this article will focus on some guidelines for writing and scoring CR questions.

(continued on page two)

Essay and Short-Answer Questions

(continued from page one)

Rules for Writing Essay and Short-Answer Items

1. Use essay and short-answer questions to measure complex objectives only. Too often, people use these questions to collect information that is easily obtained through MC items (e.g., list, define, identify, etc.). Whatever advantages are gained by asking students to produce (rather than recognize) answers for lower-level tasks are usually more than offset by the disadvantages associated with scoring such items. CR questions should be reserved for situations where either supplying the answer is essential or where MC questions are of limited value. CR items are best reserved for questions including words such as “why,” “describe,” “explain,” “compare,” “contrast,” “criticize,” “create,” “relate,” “interpret,” “analyze,” and “evaluate.”
2. The shorter the answer required for a given CR item, generally the better. More objectives can be tested in the same period of time, and factors such as verbal fluency, spelling, etc., have less of an opportunity to influence the grader.
3. Make sure questions are sharply focused on a single issue, which should be directly related to course objectives. It is difficult to write an item that identifies exactly what type of a response you expect. Do not give either the examinee or the grader too much freedom in determining what the correct answer should be.
4. Do not allow students to choose among a set of possible questions to answer. In most classroom situations, it is very difficult to compare the performances of two students who were allowed to answer different sets of questions because not all questions are equally difficult or easy to grade. Also, if students are allowed to choose, it is harder for you to control the content of the exam. Because students will likely choose to answer the items on topics which are most familiar to them, students’ performance will not truly represent how well they have mastered the entire domain of interest.

(continued on page three)

Editor's Note

In the October issue of *The Learning Link*, the article, *Helpful Tips for Creating Valid and Reliable Tests: Writing Multiple Choice Questions* was erroneously listed as co-authored. James Wollack of Testing and Evaluation Services was the sole author of that article. *The Learning Link* staff apologizes for any confusion or inconvenience this may have caused.



The Learning Link is a quarterly newsletter published by the University of Wisconsin-Madison Teaching Academy. It aims to provide a forum for dialogue on effective teaching and learning. University teaching staff, including graduate teaching assistants and undergraduate students, are invited and encouraged to submit contributions.

Editors:

John DeLamater, Sociology
(608) 262-4357
delamate@ssc.wisc.edu

Mary Jae Paul
(608) 263-7748
mpaul@bascom.wisc.edu

Designer: Mary Jae Paul



UW-MADISON TEACHING ACADEMY
133 Bascom Hall
500 Lincoln Drive
Madison, WI 53706
(608) 262-1677
[http://wiscinfo.doit.wisc.edu/
teaching-academy/](http://wiscinfo.doit.wisc.edu/teaching-academy/)

Chair:

Jay Martin, Mechanical Engineering
(608) 263-9460
martin@enr.wisc.edu

Project Assistant:

Elizabeth Apple
(608) 263-7748
EJApple@bascom.wisc.edu



Rules for Scoring Essay and Short-Answer Items

Because of their subjective nature, essay and short-answer items are difficult to grade, particularly if the score scale contains many points. The same items that are easy to grade on a 3-point scale may be very hard to grade on a 5- or 10-point scale. In general, the larger the number of points awarded for an item, the more difficult it will be to grade. Also, more complex questions will produce a wider variety of responses, also complicating the grading process. When grading CR questions, instructors should focus primarily on two goals: consistency and fairness. Consistency refers to the extent to which the same points are awarded or subtracted for comparable information across students. Two students making the same or comparable misinterpretations should receive the same deductions. Fairness refers to the extent to which the points assigned or deducted reflect the weighting of objectives in the test blueprint. If, for example, students are asked to solve a series of related problems (e.g., the answer from one problem is used to solve another problem), getting an intermediate step wrong should only result in losing points once. The second problem presumably relates to a different objective, and if the problem is solved correctly (given that the wrong initial value was used for one part of the problem), full points should be awarded.

Achieving consistency and fairness while scoring CR items is challenging. Below are a few guidelines which may help achieve these two goals.

1. Construct a detailed scoring rubric that identifies the basis for awarding or subtracting points at each phase of each item. To do this, it may be helpful to develop a model answer and think about the essential elements in producing that answer. Pay careful attention to how to score errors of omission and commission. While establishing your rubric, be cognizant of the total number of points available for the item and make sure that it is not possible to receive lower than zero points or more than the total number of points for each item.

Quarterly Quote

“It is easier to perceive error than to find truth, for the former lies on the surface and is easily seen, while the latter lies in the depth, where few are willing to search for it.”

- Johann vonGoethe

2. CR items should be graded anonymously if at all possible to reduce the subjectivity of graders. That is, graders should not be informed as to the identity of the examinees whose papers they are grading.
3. Grade all students’ responses to one question before moving on to grade the second question. This helps the grader maintain a single set of criteria for awarding points. In addition, it tends to reduce the influence of the examinee’s previous performance on other items. If multiple graders are used and it is not possible for all graders to rate all items for all students, it is better to have each grader score a particular problem or two for all students than to have each grader score all problems for only a subset of students. This strategy is effective for eliminating effects due to one person grading harder than another.
4. While grading a question, maintain a log of the types of errors observed and their corresponding deductions. It is very difficult to anticipate every error you will see, but this will allow you to maintain consistency across exams. It may be necessary to re-examine some questions that had already been graded to verify that the point deductions are consistent and fair.
5. Unless writing skill is one of the course objectives, do not take off credit for poor grammar, spelling errors, or failure to punctuate properly, unless the quality of writing clearly interferes with your ability to understand whether the student has adequately grasped the material. Never grade on the basis of penmanship.

CR items are difficult and time-consuming to grade, but with carefully planned and methodically implemented grading criteria, they can provide a richness of information not available through only MC items.

In the final article in this series, I will introduce the final step in the test development process, that of evaluating the test itself. For more information on test development, please check out Testing & Evaluation (T & E) Service’s website at <http://www.wisc.edu/exams>, call, or visit T & E Services (373 Educational Sciences Bldg., 262-5863) and ask to talk with someone about help on developing classroom assessments.

.....

James A. Wollack
Testing & Evaluation Services
UW-Madison, November 2003