Using Multiple Choice Tests via Blackboard for Accounting Students

This paper describes the use of a Multiple Choice (MC) test, administered via Blackboard, as part of the assessment diet for a level 4 module in Business Accounting Systems. It starts with a brief overview of relevant literature on the topic of MC tests and then progresses to describe the process of administering the test to a group of 32 students.

The use of MC tests in Accounting Studies

The use of MC tests is seen as attractive to lecturing staff as it provides an assessment that can be quicker to mark compared to constructed response (CR) tests where students may submit a wide variety of answers. This benefit is enhanced where the test is administered via a VLE such as Blackboard, where marking (and feedback) is done automatically and promptly. However, we should be wary of using ease of operation for academic staff (and of course for administrative staff, due to fewer paper responses) as the sole criterion for the use of MC tests. Clearly, we need to ask whether MC tests are fit for purpose – do they provide similar information on student achievement as CR tests?

Comparing MC tests with CR tests

If we accept CR as the benchmark, then similar student performance on a MC test would suggest that in terms of grading, MC testing has benchmark reliability. Frakes and Lathen (1985) examined the student performance under the MC and CR type assessments. They found that there was no significant difference of student performance between the two assessments types. Bible et al (2007) also found a correlation of student performance, although they classify it as being "a somewhat strong relationship". Both of these studies were in the subject area of financial accounting. Bible et al (2007) make some comments of the relevance of subject matter, citing studies which show a stronger correlation than they found for accounting. A further point that they raise is that accounting programmes as a whole may have heterogeneous subject content. For example, taxation may appear to be a computational subject, but it may also involve an understanding of abstruse legal concepts. In other words, results of a study from one aspect of an accounting programme may not be relevant to all aspects.

A meta – analysis carried out by Rodriguez (2003) of 67 empirical studies comparing performance on MC tests with that of CR tests highlighted the inconsistency in the results of these studies. In this paper, the issue of "stem equivalency" is raised; how closely are the questions contained in the two types of tests related. The implication is that student performance may be different because different topics are being assessed – it is not the method of testing that causes a difference but the difference in question content.

In addition to the studies cited above, there are many more published studies, each of which may bring a different facet of understanding (or confusion!) to the academic trying to discern an appropriate method of assessment. It seems to the present authors that there are obvious differences in MC and CR tests, more so as we ascend Bloom's taxonomy. At a basic level of accounting expertise, a MC question such as:-

"Purchases are £100,000. Opening inventory was £5,000 and closing inventory was £6,000. What is the cost of sales?" (select one of four numerical options) is markedly similar to a CR question framed in the same way, except of course there are no options to select.

This could give the MC candidate an advantage as a guessing strategy would give a 25% chance of success, but there again a CR candidate could well obtain credit for a wrong numerical answer but a partially correct articulation of the calculation. However, it is hard to see how a CR question demanding some analysis and synthesis, such as "consider the following financial statements and recommend an appropriate financial strategy "could be administered via a MC test. This of course could simply reflect current incompetence in designing MC tests; but it seems fairer from the student perspective to use MC tests (for summative assessment) in level 4 modules only; perhaps in conjunction with CR tests at higher levels.

Administering a MC test to 32 Level 4 Business Accounting Students via Blackboard

Design

The assessment consisted of 25 questions, each of them written by a lecturer teaching on the module. Consideration was given to using questions imported from publishers' websites (e.g McGraw Hill EzyTest) but many of these had already been made available to students as formative assessments. There is also the question of style, accounting technique and vocabulary used in imported tests that may differ from those employed by the staff teaching on the module.

Available question types are as follows.

- Calculated: Students must apply a mathematical formula and enter the answer.
- Combination multiple choice: Students select a combination of correct answers from a list.
- Fill in the blank: Students fill in words or phrases left blank in the question.
- **Jumbled sentence: Students fill in blanks within a question by selecting a word or phrase from a drop-down list.**
- **Matching: Students match terms in one column with terms in another column.**
- **Multiple choice: Students select either one or multiple correct answers from a list**.
- Paragraph: Students answer with multiple sentences.
- Short answer: Students answer with a word or phrase.
- True false: Students select whether the question is true or false.

The question types used in this assessment are in bold, above. A key reason for selecting these question types is that it avoids the issue of student responses being marked as incorrect because of spelling errors or students not following the prescribed format for an answer. For example, Blackboard would treat £15,000 as a different answer to £15000 or 15000 pounds. To make the administration as simple as possible it was decided to avoid these question types and hence this problem.

Feedback can be given for each question, explaining how an answer was derived or why a selected option is incorrect.

Parameters

Various parameters can be set. These include the following.

Availability of the test - e.g. from 10 am March 17th to 11 .30 am March 17th. This can be changed at any time, so in practical terms the assessment could be made available from when all the students are in the examination room. They have an hour from starting the test and the test "times out" when the time has run out for each student. Having an **availability** of more than an hour allows for the fact that students will not all start the test at exactly the same time; but they all get an hour in which to complete it. The test was taken in one of the PC Labs in the Charles Hasting building.

Question delivery – students can be allowed to revisit answered questions, or be presented with all of the questions at once. In this assessment, students were presented with questions sequentially and allowed to revisit them.

Number of attempts – students can be restricted to one attempt or they can have unlimited attempts – In this assessment they were restricted to one attempt, but one could envisage multiple attempts if the assessment was formative in nature.

Student score – this can made available as soon as each student has finished the assessment or until the assessment is no longer available, the option in this case. The components of the student score can also be managed, to include or exclude items such as feedback.

Results and Reports

A variety of reports can be obtained from Blackboard. Students can view their completed assessment question by question, showing their answer, the correct answer and feedback.

Reports for lecturers can be at student level and/or question level, permitting an examination of questions that seemed to be particularly badly answered (or the converse of course). Reports are downloadable to Excel permitting further analysis. A statistic calculated by Blackboard is the discrimination index – discussed by Baldwin (1984). This measures the correlation between responses for each question and overall student performance of the assessment. A high positive correlation suggests that the answer was selected by high performing students, and is desirable for correct answers. A high negative correlation suggests the answer was selected by low performing students, and is desirable for incorrect answers. A correct answer with a low positive discrimination index suggests that the question was problematical for the better performing students and that it was answered correctly by students regardless of their overall performance. Therefore, is it too hard, too easy or badly expressed?

Results are automatically posted to the Blackboard grading book for subsequent input into ETM – in this case the grading book was exported to Excel and "vlookup" used to generate ETM scores and grades.

Student Feedback

The week after the assessment we surveyed students to elicit their responses. The questions and responses are as follows.

Was the time allowed for the test	n
Too long	
About right	27
Too short	

Did you find the software was

Easy to use	26
Difficult to use	1

Did you find the feedback provided

Satisfactory	21
Not satisfactory	4
I haven't looked at it	2

Would you have preferred a written examination

Yes	7
No	20

Did the test reflect what we had studied on the module

Closely	17
Moderately	10
Not really	2

Your Gender

Male	17
Female	10
Not answered	

These results suggest a high degree of student satisfaction with the process. Free texts comments were also elicited, prompting 11 responses, which were all favourable bar one. The most common response was an appreciation of the immediate post test results and feedback.

Where do we go from here?

Our conclusion from this exercise is that MC tests administered via Blackboard are a feasible way to administer MC tests. Student acceptance is favourable. From our view as lecturers, we estimate constructing the assessment took about 5 lecturer hours. A paper MC test would probably have taken 3 lecturer hours to write but marking for 32 students would have added another 3 hours. So there is a small saving of time, but there are two comments to make here. Firstly, we anticipate that future assessments in Blackboard will not take so long to construct as we have more experience in its use and Blackboard enables the re-use of questions from prior assessments. The second point is one of scale; in the context of 300 students, administration via Blackboard saves the marking overhead associated with paper based tests.

Test administered via Blackboard also produce a variety of reports and statistics that can be used to examine the appropriateness of the test, something which is possible with paper based tests but time consuming.

Of course, the comments above accept that MC tests are a valid method of assessment compared to CR tests. The discussion at the start of this paper leads us to realise that we need to give more thought to this issue. The studies we have looked at so far give differing views and suggest areas that need careful attention in constructing MC tests. However, provided we can satisfy ourselves on this issue, we feel that MC tests administered via Blackboard may provide an efficient method of assessment.

References

Baldwin, Bruce A. (1984) The role of difficulty and discrimination in constructing multiple-choice examinations: With guidelines for practical application. *Journal of Accounting Education,* 2 (1), 19-28.

Bible, Lynn, Simkin, Mark G. & Kuechler, William L. (2008) Using Multiple-choice Tests to Evaluate Students' Understanding of Accounting. *Accounting Education,* 17, 55-68.

Frakes, Albert H. & Lathen, William C. (1989) A comparison of multiple-choice and problem examinations in introductory financial accounting. *Journal of Accounting Education,* 3 (1), 81-89.

Rodriguez, M. C. (2003), Construct Equivalence of Multiple-Choice and Constructed-Response Items: A Random Effects Synthesis of Correlations. *Journal of Educational Measurement*, 40: 163–184.