

Using Constructed-Response and Multiple-Choice Questions in Undergraduate Examinations.

By E. John McElvaney

This paper looks at the use of constructed-response and multiple-choice questions in first year management examinations involving 2364 students over a three semester period in 2008-9. It also compares student's performance in research and analytical assignments with their results in Multiple-Choice tests. The results show that students who perform poorly in non-multiple-choice tests are slightly advantaged by multiple-choice tests. However, students who display high distinction scores in other pieces of assessment do not receive any advantage from multiple-choice testing. Interestingly, the research also highlighted that female students perform better in their articulation of management techniques overall, but do not gain the same comparative advantage from multiple choice testing. The analysis of the assessment methods was extremely useful in predicting the level of failures for the unit in the subsequent year. Finding the right balance of assessment tasks in first year university business courses is extremely difficult, using multiple-choice tests can be useful tool as part of the assessment mix in management units.

Key words: Gender, constructed response examinations, multiple choice tests, assessment.

1. Introduction

The Management Unit is a first year core unit for many Bachelor of Commerce and Bachelor of Business degrees. It is also a popular elective subject for students from other faculties and partner institutions, because no matter what occupation the student chooses, the majority of university graduates eventually manage people and resources. (McElvaney 2008; Deakin 2009). The Management unit in many universities has a high dropout rate. (McElvaney 2008) This attrition rate of so many university students exacts a high price; in dollars, in missed opportunities, and in human lives. (Erickson, Peters & Strommer 2006) In the management unit in this study, 13% dropped out of the unit prior to taking examinations. The failure rate prior to 2008 was 20% making a total of drop outs and failures in excess of 30% of enrolments. Many of the Australian universities benchmark against the leaders in the Course Experience Questionnaires (CEQ) of graduating students in Australia (Ramsden 1991). Deakin University is no exception and

E. John McElvaney. School of Marketing and Management. Deakin University. Email: john.mcelvaney@deakin.edu.au

McElvaney

uses the Student Evaluation of Teachers and Units (SETU) as their internal method of measuring the effectiveness of teaching in units on a cohort by cohort basis. (Deakin University 2008) These 1st year units are likely to have much lower unit evaluation ratings from students than 2nd or 3rd year undergraduate units. (Hagel & Azmat 2008) .This could partly be because the larger compulsory core units of up to 2000 students employ many part time teachers. Good teachers inspire students to become the leaders in tomorrow's world. However, poor teachers, like poor managers, can alienate students, resulting in them leaving. Some management lecturers have extensive industry experience and provide a much needed balance of the theory with the practice. (Clinebell. & Clinebell 2008; Bennis & O'Toole 2005) However, because large management units are used as teacher training grounds some lecturers and tutors are new to the discipline and do not have extensive academic or industry experience.

To achieve consistency across such a diverse range of teachers and teaching locations there is a tendency to use Fredrick Taylor's management theories; where there is one best way to do the job. (Waddell et al. 2007).Ironically, when surveying the students about their likes and dislikes about the management unit, 80% of the students was unhappy about the assessment and they requested the inclusion of Tayloristic type Multiple-Choice tests in the assessment mix. This lead to a trial using their preferred assessment method of multiple choice testing. The findings are contained in this paper.

2. Literature Review

The use of multiple-choice tests is common in universities and they are generally well accepted by students and teachers (Wallace & Williams, 2003; Simkin & Kuechler 2005; Erickson, B.L. Peters, C.B. & Strommer, D.W. 2006). Multiple-choice tests are very useful when examining large classes; they can be marked very efficiently and can be used to examine almost all topics of a unit. (Betts et al. 2009; Simkin & Kuechler 2005; Wallace & Williams, 2003).

First year students have problems coming to terms with requirements of standard university research, analytical and applied assessment methods. Their preference for multiple-choice tests is to make the transition from high school to university learning in ways that they are familiar and gradual. (Erickson, B.L. Peters, C.B. & Strommer, D.W. 2006; Betts et al 2009). Teaching assessment methods are vitally important to accurately and fairly evaluate students' knowledge. The assessment methods are a useful tool in assisting the planning and organisation of the teaching of the unit. Biggs (1999 p49) says, as part of the planning stage we should be asking "what should our students be able to do intellectually, physically or emotionally as a result of the learning in that unit". Multiple-choice tests can measure if academic objectives of the unit have been achieved. They provide information on the effectiveness of teaching and pedagogy methods. (Hautau et al. 2006).

McElvaney

They test students' understanding and are important in the allocation of grades. (Bible et al. 2008 p56)

This study also looks at the issue of gender differences in the study of management and the issues with Multiple-choice tests. Studies have highlighted a possible advantage to male students relative to female students on MC tests (Bolger and Kelleghan,1990; Lumsden and Scott 1995; Bridgeman and Lewis 1994). Other studies have shown no significant difference between males and females when both are evaluated using a MC test rather than a CR test (Chan and Kennedy, 2002; Walstad and Becker, 1994). Although Simkin and Kuechler (2005) believe that the use of randomly selecting the order of question presentation for individual students to effectively avoid cheating, in some cases where Multiple-choice tests have been done using internet testing methods such as 'Blackboard' there is anecdotal evidence that cheating did occur. Consequently, Multiple-choice tests were incorporated into the strictly supervised end of trimester examination. The customized test questions were originally designed by the prescribed textbook authors Waddell, D. Devine, J. Jones, G. & George J. (2007)

3. Methodology and Research Design

- The unit objectives and assessment methods were examined. A survey of past students was undertaken to determine their likes and dislikes regarding the unit. A 10% in 2008 and 30% in 2009 multiple-choice component was included in the end of semester examination.
- A 40 MC question per chapter for the text book by Waddell et al. (2007) was developed and installed on the university's internet system. Students were required to do the MC tests weekly as part of their study. Tracking of their results could be done by their teachers.
- The methodology was similar to that used by Lynn Bible, Mark G Simkin & William L. Kuechler of University of Nevada, USA (2008) to assess MC testing on accounting students.
- The study was conducted over three trimesters to investigate the relationship between the two types of performance measures and gender differences. The sample consisted of 2364 students enrolled in the management unit at a Victorian based university in Australia. Each student was required to complete two parts of the same examination: one part consisting of Constructed Response (CR) questions and one part consisting of Multiple-choice (MC) questions.
- 44% of the students were female and 56% were male. Although the sample differed from trimester to trimester, the two types of questions on each test measured students' understanding of the same course material.
- The assessment in the management unit consists of 40% progressive assessment, and an examination worth 60%. The progressive assessment was divided into two tasks. Minor assessment: Individual research assignment (10 marks) Major assessment: Group assignment of no more than 3 people (30 marks)

McElvaney

- The examinations that provided data for this study were administered as the normal final examinations for units in the Faculty of Business and Law.
- The final examination covers the entire semester's work and took place during a three-hour independently administered session and was closed book.
- The two measures of students' understanding in this study were the scores on separate sections of the same examination. The trimester 1&2 / 2008 examinations were worth a total of 60 marks. They consisted of twenty MC questions, each worth 0.5 marks out of a total of 10 examination marks, two problem-solving/essay questions section worth 20 marks and six short essay theory questions worth 5 marks each for 30 marks.
- The 2009 examinations had sixty MC questions, each worth 0.5 marks out of a total of 30 examination marks and six short essay theory questions worth 5 marks each for 30 marks.
- In the sample tests, each MC question referred to a separate aspect of management theory and had four possible answers, labelled A to D.
- Students answered MC questions by blackening a square on a Scantron scoring sheet.
- Each CR question required a student to explain the theory behind and use practical examples in answering the questions.
- Because of the potential for inconsistency in the CR marking process, especially as ten individual markers were employed, three markers were used per examination paper and this should have resulted in a fair spread of marks.
- The cover page of the examination contained test instructions; Information about how much each question was worth and the maximum amount of time available for the examination. In practice, most students began by answering the MC questions.
- Results of the 2008 assessment were analysed and used to determine the level of multiple-choice assessment to be included in the trimester 1/ 2009 examination to achieve a predicted examination fail rate of no more than 11%. An analysis of assessment scores between different genders was carried out using the results from semester 1&2 /2008 & trimester 1 /2009.

4. Discussion of Findings

The first major question that needed to be answered was “does Multiple-Choice Testing correctly assess the students knowledge of the subject matter?” The analysis indicates that Multiple-Choice Testing correctly assesses the students' knowledge with a direct co-relation with the scores students achieved in their four methods of assessment. Early empirical tests have led some scholars to conclude that MC tests and CR tests do in fact measure the same thing (Saunders and Walstad, 1998; Traub, 1993; Wainer and Thissen, 1993; Bridgeman, 1991). See Table 1 below:

McElvaney

Table 1 Average results	Exam MCQs	Exam Exam CR	Exam Total	Guess Factor	Guess Marks
T 1-2008 577 students	68.8%	58.2%	60%	10.6%	1.06
T 2-2008 1072 students	69.7%	55.8%	58.1%	13.9%	1.39
T 1-2009 715 students	66.2%	59.5%	63%	6.7%	2.01
T 1-2-3 aver 2364 students	68.2%	57.8%	60.4%	10.4%	1.75

Some prior studies of examination effectiveness have adjusted the scores of MC tests in an attempt to correct for the effects of guessing. The findings in Table 1 do not dispute the MC tests by Zimmerman and Williams (2003) who stated that 'guessing contributes to error variance and diminishes the reliability of tests.'

However, the guess factor is only significant if MC is the major method of assessment: for instance the figures show that a 10% MC test will only give the average student one extra mark and a 30% MC test will only give the average student two extra marks in the whole 100% of the unit. It could be argued that to neutralize the guess factor, 1.25 points could be taken off for each wrong answer. The author is of the opinion that a guess as defined by Webster Dictionary to be "an intuitive judgement from data or for reasons that establish a strong probability". Accordingly, the author believes that making adjustments to the data to accommodate the guess factor is not warranted in this study. See Table 2 below for the data on students with different grades and how they performed of MC & CR tests. Multiple Choice tests really assisted the students that fail the unit. However, they failed anyway, so not doing MC tests or penalising wrong answers would be counter-productive to the higher performing students.

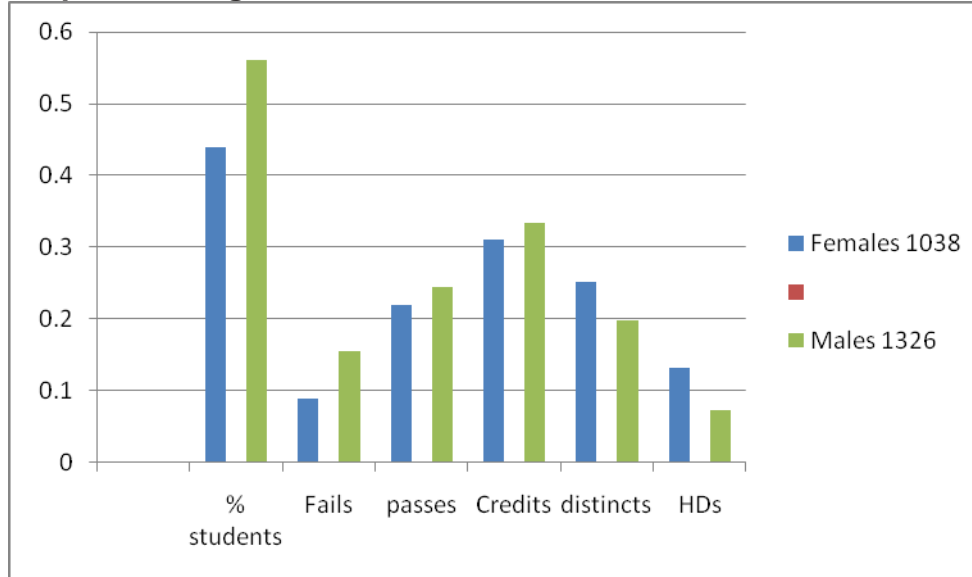
Table 2 Grade results Ave over 3 semesters	Exam MCQs	Exam Exam CR	Exam Total	Guess Factor	Guess Marks
Fails	55.1%	27.1%	34.4%	28%	4.66
Passes	62.5%	46.7%	56.7%	15.8%	2.63
Credits	68.9%	57.4%	60.7%	11.5%	1.91
Distinctions	76.6%	68.4%	70.8%	8.2%	1.36
High Distinctions	83.4%	81.2%	82.2	2.3	0.38

Interestingly, although fewer females enrol to study management, their results indicate that they are better students than males. The graph below highlights

McElvaney

the differences in performances of female and male students. It clearly shows that on average females outperform male students especially in terms of distinction and high distinction grades.

Graph 1. Management Unit: Gender differences in overall results



Regardless of the above it has been reported that males perform better than female students or actually get more from MC tests because it would appear that male students take better guesses. (Bolger and Kelleghan,1990). This research does not indicate that view to be incorrect. The results below in Table 3 below indicate that females actually do perform better at MC test, but do not have the same advantage as they do in CR tests.

Table 3. Gender Averages	Exam MCQs	Exam Exam CR	Exam Total	Guess Factor	Guess Marks
Females	70%	60%	62%	10%	1.42
Males	67%	56%	59%	11%	1.53
Variances	3%	4%	3%	-1%	-0.11

This could mean that males are slightly better at guessing, but the differences are not significant enough to be conclusive and tend to concur with Kennedy (2002) that there is no meaningful differences in the guess factor between genders. When we examine the variations between females and male students in different assessment tasks we can see that the main advantage for females in assessment tasks lies in the individual research assignments and the group analytical assignments as shown in Table 4 below.

McElvaney

Table 4

Gender Averages	Assign 1	Assign 2	Exam	Exam	Total
All assessments	Research	Analytical	MCQs	CR	Unit
Females	65.3%	60.4%	70.2%	60%	63.8%
Males	59.7%	56.3%	67%	56.%3	60%
Variances	5.5%	4.1%	3.2%	3.6%	3.8%

5. Conclusion

The findings highlight that Multiple-Choice Tests do have a place in the assessment mix of management units at university level. It appears that the students are willing to put in the time to study for Multiple-Choice Tests and they find the instant feedback is extremely useful, assisting them get better marks than they receive in other types of assessment. This study only focussed on 2364 management students over a three trimester period and more research into assessment in other business units is needed. There is a guess factor, but unless Multiple-Choice Tests are the major component of the unit assessment then it is not fair to apply an additional penalty. The Multiple-Choice Test analysis was accurately used to predict future failure rates within subsequent tests and the 30% Multiple-Choice Test were a major factor in obtaining a failure rate just over 10% of those sitting the exam (dropouts were not included in the study).

It appears that although female students out perform in every type of assessment in the management unit, they don't out perform male students by the same margin in multiple-choice tests. However, the results clearly show that although less females enrol to study management, their results indicate that they are better students than male students and are much more likely to pass the management unit with a distinction or high distinction grade. If this translates into better managers in the future then maybe this generation of female management students will be the ones to break through the 'Glass Ceiling' that seems to prevent the promotion of females in today's corporations.

References

Bennis, W.G. & O'Toole, J. 2005. *How Business schools lost their way*. Harvard Business Review. 83(5)96-104

Bible, Lynn, Simkin, Mark G. and Kuechler, William L. 2008. 'Using Multiple-choice Tests to Evaluate Students' Understanding of Accounting', Accounting Education, 17:1, S55 — S68

McElvaney

- Biggs, John. 1999 *Teaching for Quality Learning at University*. SRHE and Open University Press. Buckingham, England
- Bolger, N. and Kelleghan, T. 1990 Method of measurement and gender differences in scholastic achievement, *Journal of Educational Measurement*, 27(2), pp. 165–174.
- Bridgeman, B., and C. Lewis. 1994. The relationship of essay and multiple-choice scores with grades in college courses. *Journal of Educational Measurement* 31: 37–50.
- Bush, M. 2001. A multiple choice test that rewards partial knowledge. *Journal of Further and Higher Education* 25: 157–163.
- Chan, N. and Kennedy, P. E. 2002 Are multiple-choice exams easier for economics students? A comparison of multiple choice and equivalent constructed response examination questions, *Southern Economic Journal*, 68(4), pp. 957–971.
- Clinebell, S.K. & Clinebell J.M. 2008 *The Tension in Business Education between Academic Rigor and Real-World Relevance: The Role of Executive Professors*. Academy of Management, Learning and Education. Volume 7 Number 1 March 2008 P99-107
- Deakin University, Institute of Teaching and Learning. Retrieved 10/8/2009 (<http://www.deakin.edu.au/itl/pd/tl-modules/index.php>).
- Erickson, B.L. Peters, C.B. & Strommer, D.W. 2006 *Teaching First-Year College Students*. John Wiley & Sons Inc. San Francisco. Ca. USA.
- Greene, B. 1997 Verbal abilities, gender, and the introductory economics course: a new look at an old assumption, *Journal of Economic Education*, 28(1), pp. 13–30
- Hagel, P & Azmat, F. 2008 *HRM and Management Review Report* School of Management and Marketing, Faculty of Business and Law, Deakin University
- Hautau, B. Turner, H.C. Carroll, E. Jaspers, K. Parker, M. Krohn, K. Williams, R.L. 2006, *Differential Daily Writing Contingencies and Performance on Major Multiple-Choice Exams*. *Journal of Behavioral Education* Volume 15: November 2006 P259-276.
- Kennedy, P. E. and Walstad, W. B. 1997 Combining multiple-choice and constructed-response test scores: An economist's view, *Applied Measurement in Education*, 10(4), pp. 359–375.
- Lumsden, K. G. and Scott, A. 1995 Economic performance on multiple choice and essay examinations: a large-scale study of accounting students, *Accounting Education: an international journal*, 4(2), pp. 153–167.

McElvaney

McElvaney, E. J. (2009) Do teachers of management "Practice what they preach"?, *International review of business research papers*, pp. 174-183.

Raisman, Neal. 2007 online seminar "*Boosting Enrollment and Retention through Customer Service*," accessed March 2009
<http://www.magnapubs.com/pressreleases/111207.html>

Ramsden, P. 1991. A Performance Indicator of Teaching Quality in Higher Education: the Course Experience Questionnaire. *Studies in Higher Education*, 16(2), 129-150.

Saunders, P. and Walstad, W. B. (Ed.) 1998 Research on teaching college economics. In Teaching undergraduate economics: a handbook for instructors, pp. 141–166 (Boston: Irwin/McGraw Hill).

Simkin, M. G. and Kuechler, W. L. 2005 Multiple-choice tests and student understanding: what is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), pp. 73–97.

Traub, R. E. 1993 On the equivalence of the traits assessed by multiple-choice and constructed-response tests. In R. E. Bennett and W. C. Ward (Eds) *Construction versus choice in cognitive measurement*, pp. 29–44.

Waddell, D. Devine, J. Jones, G. & George J. 2007 *Contemporary Management*. McGraw-Hill Irwin. North Ryde, Sydney.

Wainer, H. and Thissen, D. 1993 Combining multiple-choice and constructed-response test scores: Toward a Marxist theory of test construction, *Applied Measurement in Education*, 6(2), pp. 103–118.

Wallace, M., & Williams, R. L. 2003. Multiple-choice exams: Explanations for student choices. *Teaching of Psychology*, 30, 136–139.

Walstad, W. B. and Becker, W. E. 1994 Achievement differences on multiple choice and essay tests in economics, *American Economic Review*, 84(2), pp. 193–196.