



Keeping Up-to-date: Regular Online Multiple-choice Quizzes and Post-graduate Students' Performance

Harry TSE¹, Liam Kane²

¹School of Finance and Economics, University of Technology Sydney, Australia, ²Faculty of Arts and Social Sciences, University of New South Wales

Abstract

This article argues that, while not without their limitations, multiple choice quizzes (MCQs) have a place in innovative pedagogical practice. It evaluates how the introduction of a series of online MCQs affected students' performance in a post-graduate introductory economics subject in 2016. The authors also explore students' engagement with and attitudes toward MCQs through surveys and focus groups. Drawing on an econometric analysis, the authors demonstrate that regular practice multiple choice questions that provide high-quality feedback can have a significant positive effect on students' performance. This analysis also revealed that some demographic characteristics – namely, ability, mathematics skills, and region of origin – had a strong effect on students' outcomes. While focus groups indicated that MCQs helped students to self-regulate their learning, there was some negative feedback, especially about perceived prevalence of cheating. The paper concludes by providing suggestions for how some of the limitations of MCQs can be overcome.

Keywords: Multiple Choice Quizzes; Electronic Assessment; Post-graduate; Economic Education; Business Education; Students' Performance; Demographic Characteristics

1. Introduction

Encouraging students to revise subject content, participate in class discussions, and arrive ready to take on new content are not easy. Some educators have provided creative ways to remedy this issue, although they potentially involve substantial changes to course structure making their implementation challenging (Bentley et al., 2009). This paper evaluates a more modest intervention to increase students' motivation, encourage them to learn independently, and ultimately, perform well in a post-graduate economics subject at the University of Technology Sydney (UTS). In 2016, the midsemester examination was replaced with a series of graded online multiple-choice quizzes (MCQs) that, it was hoped, provided an incentive for learners to continually revise. Furthermore, a series of non-graded practice MCQs that provided automatic high-quality feedback was also created to help students revise. Drawing on an econometric analysis, this paper argues that MCQs – a form of assessment often overlooked in the literature about creative pedagogical practices – can positively influence students' overall performance. It also explores students' perceptions of MCQs and highlights the challenges educators can expect to face when implementing such a change of assessment.

2. Literature Review

MCQs are not commonly associated with innovative teaching practices. At worst, MCQs are associated with an outdated “chalk and talk” model of economics teaching, which is reasonable since widespread overreliance on this form of assessment was given as evidence for pedagogical stagnation in the field of

economics education in the USA in the late-1990s (Becker and Watts, 2001). However, because many universities are simultaneously reducing teaching resources and increasing class sizes – meaning that instructors must do more with less – online MCQs are widely used by academics (Boitshwarleo et al., 2017). Thus, it is prudent to determine how MCQs might be used to enhance student learning.

Indeed, some scholars have recently provided instructive models for using MCQs to good effect.

MCQs are typically used to measure the lower learning objectives of Benjamin Bloom's well-known taxonomy, that is, asking students to recall concepts as well as comprehend and apply them (Simkin and Kuechler, 2005; Reback and Astra, 2012; Douglas et al., 2012; Boitshwarleo et al., 2017). It may also be the case that students themselves believe that constructed responses (short answer questions and essays, for example) provide a better opportunity to demonstrate the depth of their knowledge than MCQs (Mingo et al., 2017). However, some scholars have argued that MCQs can measure elements of students' in-depth knowledge. Buckles and Siegfried (2006) have described how to craft multiple-choice questions that gauge economics students' analytical skills, namely if students can identify economics concepts and reason through logical steps without explicit guidance. Given that the design of publishers' test-bank MCQs may not be sophisticated enough to reward students who take an in-depth approach to study, which is positively correlated with good performance (Yonker, 2011), instructors would do well to craft their own more complex MCQs; and furthermore, they should do so with close reference to a coherent set of learning outcomes such as Hanson's "expected proficiencies" (Hanson, 2011).

Perhaps, the most useful aspect of online MCQs is that they can provide automatic feedback. Nicol (2007) has convincingly argued that well-designed and creative MCQs can provide feedback that, in terms of quality, correspond to seven principles of good feedback practice, to develop student "self-regulation," or in other words, independent learning: (1) Clarifying what good performance is; (2) facilitating self-assessment; (3) delivering high-quality feedback information; (4) encouraging teacher and peer dialogue; (5) encouraging positive motivation and self-esteem; (6) providing opportunities to close the gap between current and desired performance; and (7) using feedback to improve teaching (Nicol and Macfarlane-Dick, 2006). Some studies have provided empirical evidence to suggest high-quality MCQ feedback helped students self-assess their level of knowledge, gain self-confidence, and develop a dialog with instructors (Honey and Marshall, 2003; Douglas et al., 2012).

There are some limitations of MCQs. Scholars have strongly suggested that MCQs must be used in conjunction with constructed response assessments (Nicol, 2007; Boitshwarleo et al., 2017). It is also easier for students to cheat on online MCQs, especially in unproctored tests. Not only does cheating give instructors a false picture of students' understanding of a subject's content, but this behavior may negatively affect students' performance in later courses (Arnold, 2016).

While this article addresses safeguards against cheating, its main contributions to this literature are twofold. First, it examines MCQs within the context of economic education, which is underrepresented in the literature. Second, and more importantly, this study focuses specifically on post-graduate students and MCQs; and as far as the present authors are aware, it is the first study to do so.

This study sought to answer the following questions: (1) What effect did the introduction of the both graded and practice MCQs have on students' performance in "Economics for Management (EFM)" and (2) how did students perceive graded and practice MCQs.

3. The Subject and the Intervention

This article examines an intervention in a post-graduate subject – "EFM" – at UTS in semesters 1 and 2 of 2016. EFM covers the foundations of macroeconomics and microeconomics. While EFM is a principals-level economics subject, students enrolled usually have a broad range of academic backgrounds, for many, this is their first university-level economics subject. As each seminar in EFM builds on the concepts delivered in the previous one, it is vital that students stay up-to-date. While performance in this subject was generally good, the course coordinator and tutors noticed that many students did not participate in class discussions and there was concern that they did not fully grasp the content introduced in seminars.

This prompted a change in assessment structure in 2016. Hitherto, the assessment structure of EFM was as follows: Midsemester examination covering class material from lectures 1–5 (40%), a written assessment (20%), and final examination covering class material from lectures 6–10 (40%). The most significant change was the replacement of the midsemester examination with five in-class online MCQs held fortnightly throughout the semester. Each quiz is worth 5% of the final grade and the best 4 out of 5 were counted toward students' final marks. Students were given 10 min in class to complete these MCQ quizzes on the university's online teaching platform – UTSONline – using their personal electronic devices. The graded MCQs were an incentive to engage with pre-class activities, the most important of which was weekly practice (non-graded) MCQs. These MCQs too were accessed through UTSONline, so that students could attempt them at their own leisure.

One of the present authors designed all the multiple-choice questions for EFM. While this was a laborious task, creating original multiple-choice questions had numerous advantages over using a publisher's test bank. First, creating a bank of questions saved money. It is common for textbook publishers to include online tests as part of a free package of teaching resources for new customers but later hide these features behind a paywall. More importantly, creating a test bank provided the opportunity to create questions that targeted desired learning outcomes. Although it is possible to create sophisticated questions that test students' analytical skills (Buckles and Siegfried, 2006), both the practice and graded MCQs were designed to measure students' ability to recall, comprehend, and apply economics concepts introduced in EFM, in other words, measure the lower learning objectives of Bloom's taxonomy.

Given the importance of feedback to students' learning discussed in the literature, all MCQs provided automatic quality feedback. Incorrect responses prompted feedback that explained how the correct answer is arrived at and directed students to the appropriate study resources. It was hoped that the practice MCQs would help students to identify their strongest and weakest areas and guide their revision. The course coordinator could monitor how many times students attempted both kinds of MCQs and what they scored. This feedback schema complies with the first three and final seventh of Nicol and Macfarlane-Dick's seven principles of good feedback for the development of student self-regulation: (1) Clarifying what good performance is; (2) facilitating self-assessment; (3) delivering high-quality feedback information; and (7) using feedback to improve teaching (Nicol and Macfarlane-Dick, 2006).

The following are two examples of questions and feedback from practice MCQs. The examples reflect the range of difficulty of questions in all the graded and practice MCQs. In general, the multiple-choice questions consisted of both simple and complex types. The simple questions asked students to identify concepts and complex questions prompted them to apply a concept to economic data.

3.1. Simple question

To compute the gross domestic product (GDP) you:

- Use a weighted average of a survey regarding how much people value different goods and services.
- Sum the cost of producing the final goods and services.
- Simply sum the number of final goods and services.
- Sum the market value of final goods and services.

Incorrect response feedback. GDP is the market value of all currently produced final goods and services within a country in a given period by domestic and foreign-supplied resources. For the definition of GDP, refer to Lecture 1 Slide 10.

3.2. Complex question

The labor force is 100 million and the unemployment rate is 5%. One million people quit looking for a job. What is it called when an individual leaves the labor force? What is the new unemployment rate?

- Discouraged worker, 4.04%
- Encouraged worker, 5%

- c. Discouraged worker, 3%
- d. Discouraged worker, 5.05%.

Incorrect response feedback. Discouraged workers (also known as marginally attached workers) are not included in the labor force and therefore, by definition, are not officially counted as unemployed. Officially, unemployment figures only count those who have no job and are actively looking for work. If the labor force is 100 million and the unemployment rate 5%, then the number of people unemployed is 5 million. If one million people quit looking for jobs (i.e., they are not officially included in the labor force), then the number of unemployed people is 4 million and the labor force is 99 million. The formula for the unemployment rate is number of people unemployed ÷ by the Labor Force x 100%; therefore, the new unemployment rate is equal to $4 \div 99 \times 100\% = 4.04\%$. Refer to Lecture 1, slide number 55.

4. Data Collection

Data used for quantitative analysis were drawn from three sources. First, attendance records were taken manually by seminar leaders. Attendance can be used as a proxy for students' motivation (Self, 2012; Andrietti and Velasco, 2015). In semester 1 and 2, seminar attendance was 71.51 and 75.55% respectively. Second, students' records allowed the present authors to account for a range of variables that potentially affect performance in economics. Weighted Average Marks (WAM) is a proxy "pre-class aptitude" or "ability," which is known to have strong effect on students' outcomes (Mumford and Ohland, 2011; Mallik and Lodewijks, 2010; Thiele et al., 2014; Tse and Tam, 2017). In terms of performance in EFM, there may be some variation between students at different stages of their post-graduate education, which can be measured by students' credit points. It is also logical to assume that having failed EFM is negatively associated with performance in this course and repeating students were accounted for. Gender potentially influences performance in economics (Buckless et al., 1991; Keef and Roush, 1997; Fogarty and Goldwater, 2010), and as such, dummy variables (0 - male, 1 - female) were established. Students' nationality or race may influence performance in economics (Borg and Harriet, 2002; Kevin and Hill, 2009, so for the purposes of this study, countries were divided into geographic regions used by the United Nations' Statistics Division (United Nations Statistics Division, 2011). Finally, students' mathematical ability may influence their performance in economics (Auyeung and Sands, 1993; Mallik and Basu, 2009), and this was captured by students' scores on a voluntary, non-graded basic maths test completed on UTOnline in the first week of the semester.

With ethics approval, focus groups were conducted around the end of the semester to gain insight into students' perceptions of the MCQs and EFM more generally. With students' consent, each session was recorded and transcribed. Participants were assured that their feedback would be published anonymously and that they could withdraw from the study at any time without having to provide an explanation. Eight sessions were held in semester 1 and another six in semester 2. Each focus group consisted of between three and six participants and lasted roughly 15–20 min. The focus groups were conducted by one of the present authors who was not known to students. Participants were asked the following open-ended questions:

- Did anybody find the fortnightly in-class quizzes difficult?
- How effective did you find being allowed to do the quizzes on your phones and devices?
- Did you prepare for the online quizzes?
- Did these continuous quizzes help you learn the economics concepts introduced in the course?
- Do you think there are flaws in the in-class quizzes as they are now?

In addition, surveys were also administered by tutors in the second last week of each semester. Of course, students were assured that the surveys were voluntary and that their responses would have no bearing on their grades. The survey questions sought to elucidate students' study habits and clarify

what sources they used to revise. A copy of the survey is provided in Appendix 1. A total of 299 useable surveys were returned in semesters 1 and 2 of 2016.

5. Statistical Analysis

Table 1 provides summary statistics for all the variables that may affect students' performance. A total list of the variables under consideration and their abbreviations can be found in Appendix 2. In total, students' records and surveys provided data for 525 students. Table 1 presents the descriptive statistics.

About 52% of respondents were male and about 48% were female. Around 6% of students were repeating the subject. Across all the responding students, about 17% of students were from Oceania region, about 55% from East Asia region, about 27% from South Asia region, about 7% from Southeast Asia region and about 5% from the rest of the world (Appendix 3). The means of students' total marks and WAM were very close: 65.6 and 63.6, respectively. Furthermore, total marks and WAM showed a strong correlation (Pearson coefficient = 0.79, $P < 0.01$). As expected, the standard deviation for the WAM was found to be lower than for the single subject scores (standard deviation of 11.28 against 14.72). If students attempted the weekly practice MCQs, then their total marks increased by 0.68 per week. Furthermore, the correlation between the students' attempts of the weekly practice MCQs and their total marks is positive and significant (Pearson coefficient = 0.33, $P < 0.01$). In other words, the more that students attempt their weekly practice MCQs, the higher their total marks will be.

5.1. Econometric analysis

To determine what variables affected students' performance in EFM (measured by total marks) an econometric analysis was conducted. Out of the original eleven variables under consideration, five variables emerged as significant 10% level. The original regression model was:

$$\text{TOTALMARK}_i = \beta_1 + \beta_2 \text{ATTENDANCE}_i + \beta_3 \text{COUNT_PRE-CLASS}_i + \beta_4 \text{CREDITPTS}_i + \beta_5 \text{Male}_i + \beta_6 \text{OEX}_i + \beta_7 \text{REG_1}_i + \beta_8 \text{REG_2}_i + \beta_9 \text{REG_3}_i + \beta_{10} \text{REG_4}_i + \beta_{11} \text{REPEAT}_i + \beta_{12} \text{WAM}_i + \varepsilon_i$$

The original model included eleven variables and was estimated by ordinary least square. All models used Newby–West corrected standard errors. This model, the results of which are presented in Table 2, provided a starting point for a more sophisticated statistical analysis. Given the results of previous studies discussed in the above sections, the relationships between the predictors under examination and total marks aligned with the authors' expectations. Combined, these variables explain 62% of students' total marks and are highly significant ($F\text{-stat} = 78.62$).

Table 1: Descriptive statistics

Variables	Mean	Median	Mode	SD	Skewness	Min	Max
Totalmark	65.6	67.00	75	14.72	-0.74	0	96.00
Attendance	8.2	9.00	11	2.84	-1.17	0	11.00
OEX	9.3	11.00	12	3.63	-1.44	0	13.00
Count_pre-class	6.8	9.00	10	3.50	-0.67	0	10.00
Repeat	0.06	0.00	0	0.32	5.34	0	2.00
WAM	63.67	63.70	62	11.27	-0.69	0	93.80
Creditpts	58.98	66.00	72	23.99	-0.30	0	102.00
Male	0.52	1	1	0.50	-0.08	0	1
reg_1 (Oceania)	0.16	0	0	0.37	1.80	0	1
reg_2 (E Asia)	0.54	1	1	0.49	-0.18	0	1
reg_3 (S Asia)	0.17	0	0	0.37	1.75	0	1
reg_4 (SE Asia)	0.06	0	0	0.24	3.49	0	1
reg_5 (ROW)	0.05	0	0	0.22	4.01	0	1

SD: Standard deviation, WAM: Weighted Average Marks

To identify the most significant determinates of students' performance in EFM, a number of variables examined in the original regression model were eliminated based on the results of sequential t-tests, that is, the least significant variables were eliminated one at a time. Consequently, our model evolved from the original model to the final model, whereby each variable was significant at the 10%

Table 2: Original regression model

Variable	Symbol	Coefficient	se	t-value	P value	Adj R ²	F-stat
						0.62	78.62
Constant		0.631	3.760	0.168	0.867		
Attendance	(ATTENDANCE _i)	0.150	0.166	0.903	0.367		
Weekly pre-class work attempts	(COUNT_PRE_CLASS _i)	0.423	0.144	2.940	0.003*		
Credit points accumulated	(CREDITPTS _i)	0.0141	0.020	0.695	0.488		
Male	(Male _i)	-0.605	0.835	-0.725	0.469		
Basic mathematic skill diagnostic exercise results	(OEX _i)	0.512	0.116	4.424	0.000*		
Students from Oceania region	(Reg_1 _i)	-3.41	1.436	-2.374	0.018**		
Students from East Asia	(Reg_2 _i)	0.924	1.137	0.702	0.483		
Students from South Asia	(Reg_3 _i)	-1.448	1.592	-0.910	0.363		
Students who repeat the subject	(REPEAT _i)	-0.362	0.964	-0.376	0.707		
Weight average marks for all subjects undertook	(WAM _i)	-0.894	0.0464	19.270	0.000*		

***Significance at the 10% level; **Significance at the 5% level; *Significance at the 1% level

Table 3: Final Regression Model

Variable	Symbol	Coefficient	se	t-value	P value	Adj R ²	F-stat
Constant	C	3.600	2.934	1.226	0.221	0.613	167.16
Weekly Pre-class work attempts	(COUNT_PRE_CLASS _i)	0.381	0.151	2.521	0.010***		
Basic mathematic skill diagnostic exercise results	(OEX _i)	0.452	0.122	3.708	0.000*		
Students from Oceania Region	(Reg_1 _i)	-4.957	1.087	-4.553	0.000*		
Students from South Asia	(Reg_3 _i)	-2.850	1.061	-2.685	0.008*		
Weighted average marks for all subjects undertook	(WAM _i)	0.900	0.047	19.165	0.000*		

***Significance at the 10% level; **Significance at the 5% level; *Significance at the 1% level

level ($P < 0.10$). The final model removed six of these variables. This left five variables that were significant at the 10% level. The final regression model as follows:

$$\text{TOTALMARK}_i = \beta_1 + \beta_2 \text{COUNT_PRE_CLASS}_i + \beta_3 \text{OEX}_i + \beta_4 \text{REG_1}_i + \beta_5 \text{REG_3}_i + \beta_6 \text{WAM}_i + \epsilon_i$$

The variables in Table 3 were all significant at the 10% level. This model explains 61.3% of variation in the total scores for the subject. The overall F-statistic of 167.76 indicates that the hypothesis that $R^2 = 0$ can be rejected at the 10% level. This final model was also supported by the value of the Wald Test – a joint F-test – where all the insignificant independent variables were excluded at once ($P = 0.5699$).

5.2. Coefficient diagnostics

To ensure that there was no strong correlation between the three continuous variables – WAM, OEX, and COUNT_PRE_CLASS – an analysis to check for multicollinearity was conducted. Diagram 1 shows the matrix plot for the correlations between these three variables.

The intercorrelations between the variables were weak, which suggests that there was little potential for multicollinearity. The correlation between WAM and OEX was 0.186; WAM and was 0.390; and OEX and COUNT_PRE_CLASS 0.096, respectively. Furthermore, the high sample size allows precise estimation of marginal effects.

5.3. Residual diagnostics

We ascertained that our data adhered to the linear regression model's requirement that the regression error is bell shape and normally distributed. The residuals are normal at 1% level of significance ($P = 0.0136$). Hence, our residuals do not reveal any departures from normality ($F\text{-stat} = 15.1$, $t\text{-stat} = 14.3$). This was also corroborated by a visual normal quantile plot in Diagram 2.

This model suggests that students' total marks were made up of a base mark 36. For every weekly practice MCQ that students attempted, we can expect that their total marks would increase by about 3.8 marks. For each mark higher that students achieve in the mathematics orientation exercise, we can expect that their total marks would increase by 4.5 marks. Students from Oceania were worse off by 4.9 marks; students from South Asia were worse off by 2.8 marks. For every ten marks higher in WAM, the students' total marks would increase by nine marks.

Diagram 1: Coefficient diagnostics

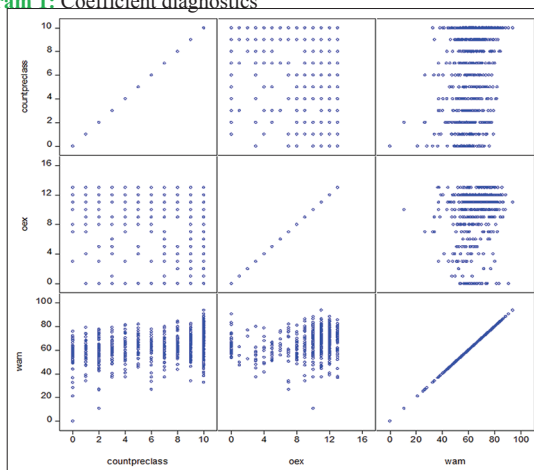
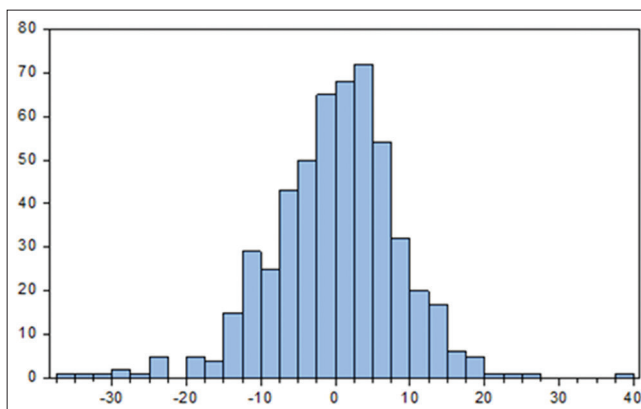


Diagram 2: Residual diagnostic normality test



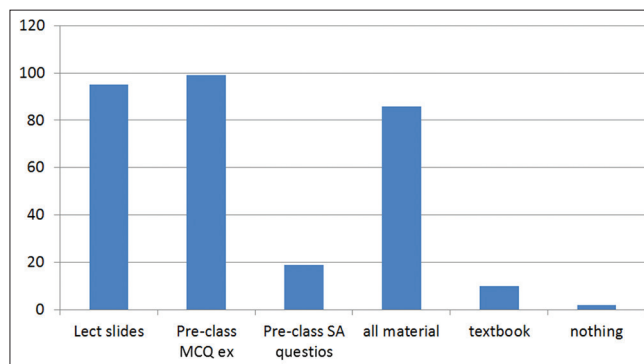
6. Qualitative Analysis

In general, focus group participants had more criticism than praise for the MCQs. Most promisingly, many participants clearly perceived the intended purpose of the fortnightly in-class MCQs, which was to prompt them to “keep up-to-date.” This was articulated particularly well by one participant, “It was not to test our IQ, it was to push us and drive us to keep practicing these things and not have a pile up with it at the end of the semester.” Similarly, several participants also stressed that they did not feel the need to “cram” for the final exam because they had studied continually throughout the semester. Many students found the MCQs useful because they highlighted their strongest and weakest areas. As one participant noted, “So if nothing else, the quizzes remind you which bits you’re probably a little bit weaker on.” Some participants also appreciated that the best four graded MCQs out of five were counted toward total marks: “They give a chance to fall over in one and pick up in another.”

While it is true that more participants criticized the MCQs, it is crucial to acknowledge that this was constructive criticism – much of this feedback shaped this assessment in following semesters. Very few participants suggested that the MCQs should be abandoned. Initially, there were some “teething issues” as both students and tutors learned how to use the technology. It seems that while the online MCQs generally worked well, some devices were not very compatible with the software on UTSONline. Nonetheless, paper examinations were available for the few students who experienced technical difficulties. Some participants believed a post-graduate degree should have difficult MCQs: “I want to be challenged in this degree; I didn’t find this challenging.” Conversely, some focus group participants found the questions too hard and did not have time to study. Others suggested that a weighting of 5% per quiz was not a strong enough incentive to study hard. Most troubling was the fact that some students cheated in the graded MCQs by sharing their answers with peers or searching for the answers on the internet. It is unlikely that a high number of students cheated in the fortnightly MCQs, but the ease with which it could be done prompted some focus group participants to question the fairness of the assessment.

The data from 299 surveys provided an insight into students’ study habits. While it is possible that participants inflated the number of hours they spent studying, there is no reason to believe that they were dishonest and the results were not alarmingly high. Most (37%) respondents did 2 h of private study for EFM; 30% did one hour; 28% did over 3 h; and 3% did no study at all. Most (32%) used the pre-class MCQs; 31% of respondents just used lecture slides for revision; 28% used a combination of all available study materials; 6% used pre-class short answer questions; just 3% used the textbook; and <1% used no study materials at all (Diagram 3).

Diagram 3: Study material



7. Discussion

The most significant finding of this study was that attempting the practice online MCQs had a positive and significant impact on students' total marks in EFM ($P < 0.1$). The more students attempted the practice MCQs the higher their overall marks would be. While this finding was very promising, the possibility for self-selection bias must be acknowledged: It is possible that the better students chose to do the voluntary pre-class MCQs. Nonetheless, including those that used a combination of all the study materials, the vast majority of survey respondents used the practice MCQs as a study tool.

Focus group participants' perceptions of the graded MCQs were mixed. It was most gratifying that some students recognized that the MCQs did not simply test them for the sake of testing, but, rather, were intended to prompt them to "keep up-to-date." Student feedback also highlighted the importance of automatic feedback, which suggests that the MCQs helped them to self-regulate their learning. Instructors also stressed the usefulness of being able to quickly check students' scores in the fortnightly MCQs and clarify students' overall weakest areas. However, during the in-class MCQs students could share the feedback (including the correct answer) with their peers. Seeing other students cheat gave some focus group participants doubts about the fairness of the assessment. Furthermore, some students expected the MCQs to be very difficult and were disappointed that they did not feel challenged.

Much of this feedback was incorporated into the EFM's assessment structure in 2017. The idea of creating more difficult questions was discussed; however, there were major concerns that difficult MCQs could potentially have an adverse effect on introductory-level students' self-confidence. Nonetheless, academics that are coordinating intermediate or specialized post-graduate economics courses could certainly consider designing MCQs that test students' critical thinking skills. Nonetheless, some safeguards against cheating were implemented in 2017. The test bank for each fortnightly MCQ was increased to 60 questions per quiz, and questions were randomly selected, which meant that each student got a different set of questions. Thus, a strong anti-cheating measure was introduced without abolishing automatic feedback, which students valued highly. In addition, instructors were encouraged to be more active in the classroom during the graded MCQs to discourage cheating.

While the practice MCQs had influenced performance in EFM, it was not the strongest determinant. Given the significance of pre-class aptitude stated in the literature, it is unsurprising that pre-class aptitude – also referred to as ability – had the most significant effect on students' performance in EFM ($P < 0.01$). Certainly, there is nothing that instructors can do to increase students' ability before they enroll their subjects: Remedying this issue would involve broad changes both in the teaching and administration of undergraduate and post-graduate level courses that are beyond the scope of this paper. There are internal and external supports for students who have a low WAM or are generally struggling in EFM. At UTS, the main support group is Higher Education Language and Presentation Support. Students enrolled in EFM also had the opportunity to attend additional "U-Pass classes" – weekly peer learning sessions run by exceptional former students.

The second strongest determinant of students' outcomes in EFM was their mathematical ability – as measured by students' performance in a mathematics orientation exercise ($P < 0.01$). While introductory economics subjects like EFM usually do not feature advanced mathematics, it is expected that students have certain foundational mathematics skills – namely, basic algebraic literacy and diagrammatic skills. Remedial mathematics classes may be organized by subject coordinators to help students with poor mathematics skills. While there is some evidence to suggest that remedial maths classes do not improve students' performance in economics subjects (Lagerlöf and Seltzer, 2009), both instructors and students would be well-advised to explore this course of action. One of the present authors has also argued that some basic mathematics revision could be “embedded” into subject content (Tse and Tam, 2017).

Furthermore, it was discovered that there was a negative correlation between students' whose country of origin was Oceania region or South Asia and total marks in EFM ($P < 0.01$). This finding deserves closer attention than has been given in this paper, and the influence of students' demographic characteristics on performance in post-graduate subjects will be the subject of another study by the present authors. However, it is possibly the case that teaching at high school and undergraduate levels explains the poor performance in EFM of students from the Oceania Region and South Asia. It is also likely that many students have English as a second language and require additional help using and reading the academic language.

8. Conclusion

This study sought to determine what effect the introduction of a series of graded and practice MCQs had on students' performance in a post-graduate economics subject. It also sought to explore students' perceptions of this form of assessment. The main finding of this study was that attempting weekly practice MCQs was positively and significantly correlated with overall performance in EFM. Surveys also found that a large number of students used the practice MCQs as a study tool. Despite these promising results, students' perceptions of the MCQs were mixed. Some students recognized that the purpose of the MCQs was to help them to self-regulate their learning. However, a number of students suggested that the questions were too easy. Ironically, other focus group participants thought the questions were too hard and found it hard to find time to study regularly. The most negative comments related to cheating, and based on this feedback, safeguards against cheating have been implemented, which has resolved the issue.

Regular graded and practice online MCQs can be usefully incorporated into post-graduate teaching. This could be an easy alternative to more complex interventions designed to enhance students' learning. Instructors would be well advised to create their own test bank and design questions with reference to a coherent set of learning objectives. MCQs should also provide high-quality automatic feedback to help students locate their strengths and weaknesses and focus their study efforts, which in turn assists them in self-regulating their learning. Further research into the role of MCQs in post-graduate education – both within and outside the field of economic education – is needed. This study also revealed that student characteristics – namely nationality, mathematical ability, and pre-course aptitude – were strong predictors of outcomes in EFM. Further research is needed to clarify the determinants of success in post-graduate economics subjects.

References

- Andrietti, V., Velasco, C. (2015), Lecture attendance, study time, and academic performance: A panel data study. *Research in Economic Education*, 46(3), 239-259.
- Arnold, I.J.M. (2016), Cheating on online formative tests: Does it pay off? *Internet and Higher Education*, 29, 98-106.
- Auyeung, P.K., Sands, J.S. (1993), An evaluation of secondary school studies as predictors of performance for accounting majors. *The Australian Educational Researcher*, 20(3), 51-61.
- Becker, W., Watts, M. (2001), Teaching methods in U.S. undergraduate economics courses. *The Journal of Economic Education*, 32(3), 269-279.
- Bentley, K.A., Brewer, P.C., Eaton, T.V. (2009), Motivating students to prepare for class and engage in discussion using the hot seat. *Journal of Accounting Education*, 27(3), 155-167.

- Boitshwarleo, B., Reedy A.K., Billany, T. (2017), Envisioning the use of online tests in assessing twenty-first century learning: A literature review. *Research and Practice in Technology Enhanced Learning*, 12(16), 2-16.
- Borg, M.O., Stranahan, H.A. (2002), Personality type and student performance in upper-level economics course: The importance of race and gender. *Journal of Economics Education*, 33(1), 3-14.
- Buckles, S., Siegfried, J. (2006), Using multiple-choice questions to evaluate in-depth learning of economics. *The Journal of Economic Education*, 37(1), 48-57.
- Buckless, F., Ravenscroft, S.P., Lipe, M. (1991), Do gender effects on accounting courses performance persist after controlling for general academic aptitude? *Issues in Accounting Education*, 6(2), 248-261.
- Douglas, M., Wilson, J., Ennis, A. (2012), Multiple-choice question tests: A convenient, flexible and effective learning tool? A case study. *Innovations in Education and Teaching International*, 49(22), 111-121.
- Fogarty, T.J., Goldwater, P.M. (2010), Beyond just desserts: The gendered nature of the connection between effort and achievement for accounting students. *Journal of Accounting Education*, 28(1), 1-12.
- Hanson, W.L. (2012), An expected proficiencies approach to the economic major. In: Hoyt, G.M., McGoldrick, K.M., editors. *International Handbook on Teaching and Learning Economics*. UK: Elgar Publishing Limited.
- Honey, M., Marshall, D. (2003), The impact of on-line multiple-choice questions on undergraduate student nurses learnings. In Geoffrey, C., Thiele, D., Scholten, I., Barker, S., Baron, J., editors. *Interact, Integrate, Impact: Proceedings of the 20th Annual Conference of the Australian Society for Computers in Learning in Tertiary Education (ASCILITE)*. Adelaide: ASCILITE.
- Keef, S.P., Roush, M.L. (1997), New Zealand evidence on the performance of accounting students: Race, gender and self-concept. *Issues in Accounting Education* 12(2), 315-329.
- Kevin, J., Hill, C. (2009), Race and the development of career interest in accounting. *Journal of Accounting Education*, 27(4), 210-222.
- Lagerlöf, J.N.M., Seltzer, A.J. (2009), The effects of remedial mathematics on the learning of economics: Evidence from a natural experiment. *The Journal of Economic Education*, 40(2), 115-137.
- Mallik, G., Basu, P. (2009), Does high school mathematics improve student learning in economics in the university? *The International Journal of Learning*, 16(4), 515-520.
- Mallik, G., Lodewijks, J. (2009), Student performance in a large first year economics subject: Which variables are significant? *Economic Papers*, 29(1), 80-86.
- Mumford, K., Ohland, M.W. (2011), Student performance in undergraduate economics courses. *Journal of Economic Education*, 42(3), 275-282.
- Mingo, M. A., Chang, H., Williams, R. L. (2017), Undergraduate students' preferences for constructed versus multiple-choice assessment of learning. *Innovations in higher education*, 43(2), 143-152.
- Nicol, D. (2007), E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 53-64.
- Nicol, D., Macfarlane-Dick, D. (2006), Formative assessment and self-regulated learning: Model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Reback, K., Astra, C. (2012), Methods of assessment in the college economics course. In Hoyt, G.M., McGoldrick, K.M., editors. *International Handbook on Teaching and Learning Economics*. UK: Elgar Publishing Limited.
- Self, S. (2012), Studying absenteeism in principles of macroeconomics: Do attendance policies make a difference? *Research in Economic Education*, 43(3), 223-234.
- Simkin, M.G., Kuechler, W.L. (2005), Multiple-choice tests and student Understanding: What is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), 73-97.
- Thiele, T., Singleton, A., Pope, D., Stanistreet, D. (2014), Predicting students' academic performance based on school and socio-demographic characteristics. *Studies in Higher Education*, 1(23). Available from: <http://www.trandfonline.com/doi/abs/10.1080/03075079.2014.974528>. [Last accessed on 2014 Nov 27].
- Tse, H., Tam, K.T. (2017), Getting the basics right: Factors shaping student performance in intermediate economics. *Economics Analysis and Policy*, 53. Available from: <https://www.sciencedirect.com/science/article/pii/S0313592615301533>. [Last accessed 2018 Apr 10].
- United Nations Statistics Division. (2011), Standard Country or Area Codes for Statistical Use (M49). Available from: <https://www.unstats.un.org/unsd/methodology/m49>. [Last accessed on 2018 Apr 07].
- Yonker, J.E. (2011), The relationship of deep and surface approaches on factual and applied test bank multiple-choice questions performance. *Assessment and Evaluation in Higher Education*, 36(6), 673-686.

Appendix

Appendix 1: Extract from the survey

Section E

9) How many hours of revision did you do to prepare for the fortnightly in-class MCQ quizzes?

- (a) 0 h
- (b) 1 h
- (c) 2 h
- (d) 3 h
- (e) 4+ h

10) What materials did you look at to prepare for the fortnightly in-class quizzes?

- (a) Lecture slides
- (b) Pre-class MCQ exercises
- (c) Pre-class short answer question
- (d) All of the above
- (e) Textbook

11) How challenging did you find the fortnightly in-class quizzes?

- (a) Very challenging
- (b) Challenging
- (c) Neutral
- (d) Simple
- (e) Too simple

12) How useful did you find the fortnightly in-class class quizzes in helping you to learn and remember the economics concepts?

- (a) Very useful
- (b) Useful
- (c) Neutral
- (d) Not useful

13) How many hours do you work per week in paid employment? _____ hours

14) How many hours of private study do you perform every week (on average) for the subject Economics for Management (excluding attending classes)? _____ hours

Appendix 2: Variables and abbreviations

Constant symbols:

Student' weekly class attendance (ATTENDANCE_{*i*})

Weekly pre-class work attempts (COUNT_PRE_CLASS_{*i*})

Credit points accumulated (CREDITPTS_{*i*})

Basic mathematic skill diagnostic exercise results (OEX_{*i*})

Students from Oceania Region (Reg_1_{*i*})

Students from East Asia (Reg_2_{*i*})

Students from Southeast Asia (Reg_3_{*i*})

Students from the rest of the world (Reg_4_{*i*})

Students who repeat the subject (REPEAT_{*i*})

Male (Male_{*i*})

Weighted average marks for all subjects undertaken (WAM_{*i*})

Appendix 3: Student region of origin

