Evaluating online multiple choice quizzes as formative assessment tools in an Engineering fluid mechanics subject

David Hastie
University of Wollongong, Wollongong, Australia
dhastie@uow.edu.au

Thomas Goldfinch
University of Wollongong, Wollongong, Australia
tomgold@uow.edu.au

Abstract: *In an effort to improve learning outcomes in a second year Engineering Fluid Mechanics subject, the first author has developed and piloted a series of online multiple choice quizzes (MCQ) as formative assessment tools. This paper describes the development and implementation of these quizzes and the evaluation of their effect on students' results in a formal written mid-session exam and also final exam. Students' perception of the quizzes is also captured and discussed through the use of a brief, in class survey.*

The effect of these MCQs was investigated for two main cohorts of students, those who voluntarily attempted the MCQs and those that did not participate at all. Student participation has been compared against mid-session and final examination results as well as Weighted Average Marks (WAM) (or Grade Point Averages (GPA)) and demographic information. It has been found that there was a noticeable increase in exam marks for students who attempted the MCQs. In addition, feedback from students who participated in the quizzes has been largely positive, with several requesting more MCQs. While other factors that may potentially impact on exam results are also being considered, this evaluation has indicated that there may be benefit in rolling out additional MCQs in the future, covering a greater proportion of the subject content.

INTRODUCTION

Introductory Engineering Fluid Mechanics courses are often considered 'gateway subjects', with average grades and pass rates generally lower than in other courses delivered in parallel. With such a large cohort of students (approximately 260), inevitably a wide range of academic aptitudes exist. Finding an ideal, or at least manageable, way to deal with this can be quite daunting. The students at the lower end of the academic spectrum, who are vulnerable to failing the subject, need to be targeted for additional help, but this can prove to be difficult in a large lecture environment. These same students can quite often be too shy to approach the academic(s) in charge or their tutors to ask for assistance, preferring to suffer in silence and further hindering their progress. By catering to the students at the lower end of the academic spectrum, there is little challenge for those students who are highly capable and there is the risk that these top end students may not see the continued benefit in attending lectures. The flip-side of this argument can also hold true, that is, if the more capable students are targeted, the students at the lower end of the academic spectrum are then potentially alienated and may cease to attend lectures, to their further detriment.

The utilisation of e-Learning at the University of Wollongong is commonplace, much like many other universities nationally and around the world. Web-based course content can range from the very basic (a repository for lecture notes) to the advanced (interactive sites including online assessments, submissions, discussion forums etc.).

The first author has just completed his second year teaching in the core Engineering Fluid Mechanics subject at the University of Wollongong and this year has investigated two approaches in which to engage the students within the on-line environment. One method was to provide self-assessed multiple choice quizzes to supplement the recommended weekly tutorial problems (the focus of this paper) and the second was to produce a series of narrated video solutions to tutorial problems (not being discussed further). Both approaches were developed without specific knowledge of what

There are numerous papers in the literature which focus on the implementation of on-line multiple choice quizzes, many with the purpose of using them as formative assessment (Aziz, 2003; Cox and Clark, 1998; Hayes, 1999; Henly, 2003; Lee-Sammons and Wollen, 1989; Maki and Maki, 2001; Rolfe and McPherson, 1995; White and Hammer, 2000; Velan et al. 2008) and others as summative assessment (Honey and Marshall, 2003; Nelson, 1998)

By having a facility whereby students could self-assess their progress via formative tasks, it was hoped that a large proportion of the student cohort would take advantage of the resources. It was also hoped that struggling students would realise they were in this position and make every attempt to help themselves. This has also been documented by Hayes (1999). Velan et al. (2008) have shown that students who attempted online formative assessments as part of their course benefited in the final exam, although there was not much benefit seen in repeating the formative assessments multiple times.

One foreseeable hurdle is how to motivate the students to attempt these voluntary assessments. Although in no way applicable to all students, there is a general attitude that if an activity is not assessable, why should they spend time on it? This has also been documented by Gudimetla (2006). In an attempt to circumvent this, the students were told that the problems encountered in the MCQs would be similar to those seen in the mid-session and final exams. A similar strategy was taken by Swan (2002).

SUBJECT STRUCTURE

In recent times the Engineering Fluid Mechanics subject has been taken by approximately 260 students each year, all second year mechanical, materials, civil, environmental and mining engineering students as well as third year mechatronics engineering students (who only take the first 9 weeks of the subject). Weekly there are 2 hours of lectures and 2 hours of tutorials with additional lab time inter-dispersed throughout the session. There are a total of seven 30 minute assessable tutorial quizzes spread across the 13 weeks as well as a one hour written mid-session exam in week 6 and an on-line laboratory quiz in week 11. There is also a final exam at the completion of the subject. The first 6 weeks of lectures is taught by the first author then another lecturer teaches the following 6 weeks. The last week of the course is a revision week where both lecturers are present.

PILOT PROGRAM

The initial concept was to generate a number of short voluntary multiple choice quizzes in the on-line environment (each with a maximum of 10 questions) for the students to use for self-assessment to determine if they understood the application of the lecture material. Students could re-take each quiz as many times as they liked, with the aid of any material they wished to call upon. The majority of the questions were variations of fluid mechanics problems sourced from numerous textbooks and involved some level of problem solving. It was envisaged that these MCQs would assist in student preparation for the mid-session and final exam, where some of the exam questions could be similar to those presented in the online quizzes.

The quizzes were developed for one chapter of the textbook (Cengel and Cimbala, 2009) which spanned two weeks of lectures. The main sections within this chapter were developed into 5 MCQs; pressure and manometry, hydrostatics, buoyancy, rigid-body motion – acceleration and rigid-body motion – rotation. Students had the option to pick and choose which MCQs they attempted based on their perceived areas of weakness. In saying this, it was still hoped that students would attempt all

MCQ, even if it were only for revision purposes. Each quiz contained the same bank of questions with four randomised answers.

DATA COLLECTION

The method of data collection involved extracting all MCQ data from the on-line environment and cross referencing it with the results from the mid-session exam, including the marks for individual questions. Once this was completed, the student results for the MCQs were reset and released for preparation for the final exam. Data was then once again extracted from the on-line environment and cross referenced with the individual question marks and total mark from the final exam.

Additional student demographic data was also gathered to allow a multitude of statistical analyses to be performed if deemed appropriate. This demographic information included;

- gender,
- domestic or international student,
- school leaver / mature age,
- weighted average mark (WAM) (similar to grade point average (GPA)),
- first attempt or repeat student,
- degree major.

Once all data was compiled into a spreadsheet, the data was de-identified and sent to the second author for statistical analysis.

RESULTS

Preliminary statistical analyses have been performed focusing on only student Weighted Average Mark at this time. SPSS has been used to generate Pearson correlations for WAM against; all attempts and scores in the MCQs leading up to the mid-session exam, question 2 and question 3 and total of the mid-session exam, all attempts and scores in the MCQs leading up to the final exam and all individual questions and total mark of the final exam. The data contained in the output was flagged if the correlation is significant at the 0.01 level (2-tailed) or the correlation is significant at the 0.05 level (2-tailed). To fully explain the reasoning behind linking each MCQ to specific mid-session and final exam questions, the theme of each question is explained in Table 1. The result of these correlations is presented in Table 2 and shows that there were a substantial number of variables which show at least a 0.05 significance level.

Table 1: Composition of the mid-session and final exams

Mid-session Exam
Question 1: Viscosity #
Question 2: Monometer
Question 3: Hydrostatics
Question 4: Fluid Kinematics #

Final Exam
Question 1: Hydrostatics
Question 2: Buoyancy
Question 3: Rigid-body motion - rotation
Question 4: Viscosity #
Question 5: Rigid-body motion - acceleration
Question 6: Momentum analysis #
Question 7: Dimensional analysis #
Question 8: Mechanical Energy and flow in pipes #

no corresponding MCQ was available

Table 2: Correlations of WAM versus MCQs, mid-session exam and final exam

	WAM		
	Pearson Correlation	Sig. (2-tailed)	N
Attempts manometer	-0.06	0.4796	151
Score manometer	.219**	0.0068	151
Attempts hydrostatic	0.16	0.0807	121
Score hydrostatic	.238**	0.0084	121
Attempts buoyancy	0.07	0.4831	113

	WAM		
	Pearson Correlation	Sig. (2-tailed)	N
Attempts manometer	0.02	0.7931	127
Score manometer	.452**	0.0000	127
Attempts hydrostatics	-0.01	0.8667	129
Score hydrostatics	.301**	0.0005	129
Attempts buoyancy	-0.04	0.7198	98

Score buoyancy	.293**	0.0016	113
Attempts acceleration	0.06	0.5878	93
Score acceleration	.266*	0.0101	93
Attempts rotation	0.20	0.0704	86
Score rotation	.288**	0.0071	86
Attempts total	.274**	0.0000	249
Score total	.348**	0.0000	249
Mid-session Q1	.315**	0.0000	241
Mid-session Q2	.316**	0.0000	241
Mid-session Q3	.435**	0.0000	241
Mid-session Q4	.341**	0.0000	241
Mid-session Total	.511**	0.0000	241

Score buoyancy	.339**	0.0006	98
Attempts rotation	-0.01	0.9192	88
Score rotation	0.13	0.2119	88
Attempts acceleration	-0.02	0.8398	93
Score acceleration	0.04	0.6712	93
Attempts total	.175**	0.0056	249
Score total	.299**	0.0000	249
Final Exam Q1	.456**	0.0000	241
Final Exam Q2	.266**	0.0000	241
Final Exam Q3	.401**	0.0000	241
Final Exam Q4	.499**	0.0000	241
Final Exam Q5	.329**	0.0000	241
Final Exam Q6	.410**	0.0000	241
Final Exam Q7	.540**	0.0000	241
Final Exam Q8	.444**	0.0000	215
Final Exam Total	.680**	0.0000	241

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Additionally, t-tests were performed to determine the means of students who attempted the MCQs and those that did not attempt the MCQs. The results are presented in Table 3 to Table 5. In all cases, it was shown that there was a benefit to the students if they attempted the MCQs.

Table 3: Variation in mean score when comparing total attempts at all MCQs versus student WAM and the results of the mid-session exam

	All Attempts at MCQ	N	Mean	Benefit
WAM	≥1	161	65.75/100	5.4%
	0	88	60.40/100	
Mid-session total	≥1	169	29.26/40	12.4%
	0	82	24.29/40	
Mid-session Question 1	≥1	169	7.45/10	3.4%
	0	82	7.10/10	
Mid-session Question 2	≥1	169	7.34/10	17.3%
	0	82	5.61/10	
Mid-session Question 3	≥1	169	6.93/10	12.1%
	0	82	5.72/10	
Mid-session Question 4	≥1	169	7.52/10	16.7%
	0	82	5.85/10	

Table 4: Variation in mean score when comparing attempts at specific MCQs versus the results of the mid-session exam

	Attempts Manometer MCQ	N	Mean	Benefit
Mid-session Question 2	≥1	159	7.44/10	18.1%
	0	92	5.63/10	
Mid-session total	≥1	159	29.67/40	13.9%
	0	92	24.12/40	

	Attempts Hydrostatic MCQ	N	Mean	Benefit
Mid-session Question 3	≥1	126	7.12/10	11.8%
	0	125	5.94/10	
Mid-session total	≥1	126	30.22/40	13.0%
	0	125	25.03/40	

Table 5: Variation in mean score when comparing attempts at specific MCQs versus the results of the final exam

	Attempts Hydrostatic MCQ	N	Mean	Benefit
Final Exam Question 1	≥1	126	13.73/18	11.8%
	0	125	11.61/18	
Final Exam total	≥1	126	61.56/100	11.4%
	0	125	50.17/100	

	Attempts Buoyancy MCQ	N	Mean	Benefit
Final Exam Question 2	≥1	119	4.81/8	7.0%
	0	130	4.25/8	
Final Exam total	≥1	119	62.3/100	12.3%
	0	130	49.99/100	

	Attempts Rotation MCQ	N	Mean	Benefit
Final Exam Question 3	≥1	89	5.96/8	8.6%
	0	162	5.27/8	
Final Exam total	≥1	89	64.2/100	12.9%
	0	162	51.32/100	

	Attempts Acceleration MCQ	N	Mean	Benefit
Final Exam Question 5	≥1	99	4.47/8	18.4%
	0	152	3.00/8	
Final Exam total	≥1	99	63.23/100	12.1%
	0	152	51.1/100	

DISCUSSION

From the correlations in Table 2, it can be seen that student WAMs (GPA) have no correlation to the number of attempts at any of the MCQs. Table 3 shows a minor positive relationship between student WAMs and whether or not they tried any MCQs. This indicated that utilisation of the quizzes was not strongly biased towards higher or lower performing students. Table 2 does, on the other hand, show that there is a positive relationship between student WAMs and their performance in the MCQs and the exams. Naturally, you would expect that the higher achieving students will perform better in assessments, whether compulsory or not.

Table 3 shows that students who attempted one or more of the quizzes enjoyed some benefit when it came to the mid-session exam. It could be argued that this may be a case of the more capable or motivated students undertaking the MCQs, who will inevitably perform better in the mid-session exam anyway. However, the mean differences in all but one of the mid session exam questions are much higher that the mean difference in WAM for students that attempted the MCQs. This would suggest that there is some direct benefit from the MCQs.

This point can be investigated further in Table 4. If there was direct benefit from attempting the MCQs, one would expect that attempts at the MCQs would show a greater benefit in mid-session exam questions that were directly related to the MCQs than in unrelated questions. Table 4 shows the benefits apparent in the two related MCQ/mid-session exam pairs. At 18.1% and 11.8% respectively, these percentage benefits are no substantially different to the general benefits of 17.3% and 12.1%, as shown in Table 3. A similar situation can be seen in the final exam questions in Table 5.

Left with the question ‘did the MCQs benefit students?’, the answer is difficult. Certainly, the MCQs didn’t do any harm, but this alone does not justify the effort put into developing them! Whether there is any measurable benefit to students will require more advanced analysis of the current data set. It would also be useful to perform an in-depth analysis of individual test items, and an evaluation of students perceptions of the MCQs. These are all modifications that can be made for 2011.

An anonymous survey was sought from the students the week after the mid-session exam to attempt to gauge which quizzes were more useful and also give the students the opportunity to provide written feedback on the MCQs. Some of the student feedback included; wanting quizzes for every week of the subject, many students wanted more precise feedback for wrong answers, some wanted randomised quizzes so they were different each time and some even wanted the worked solutions posted. These comments will all be considered and potentially implemented in 2011.

CONCLUSION

A series of on-line formative MCQs were setup up as a pilot project to investigate whether their use would aid the students of an Introductory Engineering Fluid Mechanics in achieving higher marks in their mid-session and final exam. There was a substantial initial outlay of time required to develop and implement the quizzes but after this period, little time was required for their continued up-keep. The results presented above are preliminary and will need further statistical analysis on individual test items. The initial feedback from the students was that they found the MCQs a positive activity to undertake provided they received sufficient on-line feedback when they obtained a wrong answer. Student comments will be used to enhance the current MCQs for the delivery of the subject next year, especially in the area of feedback, and in light of the positive reaction quizzes will start to be expanded into other key theory areas.

REFERENCES

Aziz, S.M. (2003) Online Quizzes for Enhancing Student Learning in a First Year Engineering Course, *International Conference on Engineering Education*, 21 – 25 July, Valencia, Spain

Cengel, Y.A. and Cimbala, J.M. (2009) Fluid Mechanics: Fundamentals and Applications.

Cox, K. and Clark, D. (1998) The use of formative quizzes for deep learning. *Computers and Education, 30*(3/4), 157 – 167.

Gudimetla, P. and Iyer, R. Mahalinga. (2006) The role for E-learning in Engineering education: creative quality support structures to complement traditional learning. *In Proceedings 17th Conference of the Australasian Association for Engineering Education*, Auckland, New Zealand.

Hayes, S. (1999) *Assessment online*. Learning and Teaching Services, Swinburne University of Technology. Accessed at http://www.swinburne.edu.au/ads/howto/files/Assessment%20Online%20Paper.pdf on 12 Jan 2010.

Henly, D.C. (2003) Use of web-based formative assessment to support student learning in a metabolism/nutrition unit. *European Journal of Dental Education, 7*(3), 116 – 122.

Honey, M. and Marshall, D. (2003) The impact of on-line multi-choice questions on undergraduate student nurses' learning. *In Proceedings of the 20th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE)*, Adelaide, Australia.

Lee-Sammons, W.H. and Wollen, K.A. (1989) Computerized practice tests and effects on in-class exams. *Behavior Research methods, Instruments and Computers, 21*(2), 189 – 194.

Maki, W.S. and Maki, R.H. (2001) Mastery quizzes on the web: results from a web-based introductory psychology course. *Behavior Research methods, Instruments and Computers, 22*(2), 212 – 216.

Nelson, G.E. (1998) *On-line evaluation: multiple choice, discussion questions, essay, and authentic projects*. 3rd Annual Teaching in the Community Colleges (TCC) Online Conference, Accessed at http://www.eric.ed.gov/PDFS/ED430659.pdf on 12 Jan 2010.

Rolfe, I, and McPherson, J. (1995) Formative assessment: how am I doing? *The Lancet, 345*(8953), 837 – 839.

Swan, G. (2002) *Do these problems please!* Teaching and Learning Forum, Edith Cowin University, Western Australia. Accessed at http://otl.curtin.edu.au/tlf/tlf2002/swan.html on 19 Jan 2010.

Velan, G.M., Jones, P., McNeil, H.P. and Kumar, R.K. (2008) *Integrated online formative assessments in the biomedical sciences for medical students: benefits for learning*. BMC Medical Education, 8(52), 11. Accessed at http://www.biomedcentral.com/content/pdf/1472-6920-8-52.pdf on 15 Jan 2010.

White, R.J. and Hammer, C.A. (2000) Quiz-o-matic: a free web-based tool for construction of self-scoring on-line quizzes. *Behavior Research methods, Instruments and Computers, 32*(2), 250 – 253.

Copyright statement