Teacher- versus student-centred approaches to online assessment: experiences in a first year engineering subject

Brice Shen
The University of Melbourne, Melbourne, Australia
bshen@unimelb.edu.au

Gavin Buskes
The University of Melbourne, Melbourne, Australia
buskesgj@unimelb.edu.au

Jamie Evans
The University of Melbourne, Melbourne, Australia
jse@unimelb.edu.au

Andrew Ooi
The University of Melbourne, Melbourne, Australia
asho@unimelb.edu.au

***Abstract**: Online assessments are gaining prominence as tools that can engage students and reduce administration. They are now gaining wider acceptance as normal teaching and learning tools. With this acceptance there is a need to look at the details of such systems and the associated implications. This paper summarises the experiences of a first year engineering subject with two online assessment tools which are superficially similar, but pedagogically quite different – Maple T.A., an online quiz, and PeerWise, a collaborative multiple choice question repository. Both are question and answer systems, but one is teacher-focused, while the other is student-focused. It is found that while the online quiz is more popular, strong achievement correlations for PeerWise make it well worth pursuing, with possible extenuating circumstances impacting on its acceptance among students.*

Introduction

Online assessments are generally acknowledged to offer benefits in administration, flexibility and student convenience (Watson & Angus, 2008), and while they are often lumped together in the same basket, not all online assessments are made equal. The traditional online assessment tool is the online quiz. Typically a staff member will set up the questions, students will enter their answers and then receive a mark. Over the years, small (although sometimes quite sophisticated) variations have appeared such as the ability to randomise parameters to create individualised tests, improved usability and the addition of mathematical engine back ends, but the basic structure has largely remained the same (Allen, 2003).

This style of test has the advantage of convenience generally available to all online assessments; students (particularly in engineering) have good access to the internet (Kennedy, 2008) and have a lot of flexibility in terms of when and where they complete them, and staff can set up a test before the semester and then let it run its course. Additionally there are advantages that are largely specific to this style of assessment; it is highly controlled as the staff member picks the questions and sets the answers, it is often automatically marked by a computer which is a big advantage for large classes, the all-important feedback to students can be instantaneous (Bonham, Deardorff, & Beichner, 2003), questions and their parameters (depending on the system) can be randomised to allow students multiple attempts and to allow tests to be reused, and staff can get virtually instantaneous insight into how students are handling subject material.

The technical disadvantages of online assessments are also applicable to the online test as there can be issues ranging from browser compatibility problems for students to servers going down. Other downsides are typically that it is easy for students to cheat as tests are not usually completed in exam

conditions, there can be large setup times depending on the administration interface, the types of questions that can be entered are often limited, flexibility is generally linked to complexity which means the more flexible systems may require the learning of specific programming type languages, and answers may be required in very limited formats which may lead to support headaches. All of this means that while this style of online assessment can be useful it shouldn't be adopted lightly and setup needs to be handled carefully for the implementation to be a success for both staff and students (CSHE, n.d.). It is important to note that this style of test is very much a teacher-centred approach that lends itself towards summative assessment (Angus & Watson, 2009), although it can also be used as formative assessment particularly if an appropriate level of feedback is provided (Mestre, Hart, Rath, & Dufresne, 2002).

More recently new types of online assessments have emerged. For example, the popular learning management system Blackboard includes a range of assessment options in addition to the traditional online quiz. A few of these are: ePortfolios where students upload articles such as projects they've completed that demonstrate particular competencies, peer and self assessment where students can review and assess work done by other students anonymously and gradable discussion boards where a student's contribution to a forum is rated by a staff member.

One particularly novel type is the collaborative Multiple Choice Question (MCQ) repository. In this model of online assessment students create multiple choice questions which are placed in a repository for use by other students. While there may be learning in determining the answers to these questions, there is potentially learning in the creation of questions and answers and even in the creation of distracters, as students need to reflect on the material covered and potential misconceptions that their peers may have (P Denny, Hamer, & Luxton-Reily, 2009).

There are two key perceivable advantages of this system over the traditional quiz – the active learning that is achieved by having students create questions (Paul Denny, Hamer, Luxton-Reilly, & Purchase, 2008), and the reduced administration time as the students create the questions. Potential disadvantages are the lack of control, as students create the questions so staff can not easily control the topics covered, and the difficulty in controlling the quality of contributions, particularly for large classes. One issue that has raised concern is the validity of the answers. As answers are chosen by students there is a good probability that some will be wrong, and while this might worry some staff and students, it is in fact a good thing. When a student's answer doesn't match the suggested answer, the onus is then on him/her to determine why and in this process actively learn the material.

While both the online quiz and the MCQ repository are question and answer systems, the quiz is teacher-centred and the repository is clearly student-centred. These systems were used in a first year engineering subject in the form of Maple T.A. as an online quiz system, and PeerWise (Paul Denny, et al., 2008) as a MCQ repository. This paper will cover the experiences gained by using both of these systems with a particular focus on student experiences and achievement correlations.

Background

Engineering Systems Design 2 (ESD 2) is a first year engineering subject at The University of Melbourne, taken by engineering students from all disciplines. It is the second part of two subjects designed to introduce students to engineering, giving them exposure and basic skills in a range of disciplines. ESD 2 covers digital systems, programming, statics and dynamics. There are 36 hours of lectures and 30 hours of workshops, with tutorial and project segments in each workshop. It is run twice a year; in the second semester with around 700 students and in the summer semester with a little under 250 students. The summer semester runs over 6 weeks rather than the usual 12. Students typically only undertake one or two subjects during this time, but even so, anecdotally, they are often pressed for time. Over the summer semester 2010, assessments were comprised of 9 group assignments (2 per week), 5 weekly online quizzes (the top 4 of which were counted), a PeerWise requirement of 2 questions submitted and 5 questions answered (submission deadlines were spread across the semester) and an exam.

The lectures introduce new material which is reinforced by the tutorial segment of the workshops and the weekly online quizzes which test topic fundamentals. The project segment of the workshops and

the weekly assignments aim to develop the higher skills of analysis, design and evaluation in addition to requiring students to work in small groups.

Maple T.A.

Maple T.A.(Maplesoft, n.d.) is an online quiz system with the unique feature of being powered by the Maple software engine. This is combined with the ability to set up variables that are randomly generated within specified limits, therefore allowing for complex questions that can be asked in a vast number of variations. Maple T.A. was introduced to ESD2 in the summer semester of 2009 and has now been used in three different semesters. Currently it is hosted at The University of Melbourne.

There was a relatively steep learning curve in the set up of questions, but this was largely due to the requirement of learning the basics of the Maple language. There is also a connector for Maple T.A. with the Blackboard learning management system which reduces administration by allowing for marks to automatically be added to the Blackboard Grade Centre. This integration also improves the student experience by not requiring a second login. The amount of time required to enter questions understandably can vary significantly, where simple multiple choice questions can take a few minutes to more complex questions involving multiple calculations taking up to a couple of hours. If questions are set up carefully though, they can be used across semesters and even across subjects. Students are allowed up to three attempts per quiz, with the questions changing between attempts. The quizzes which were counted were worth 2% each. These quizzes are primarily used as formative assessment with the assessment mark acting as an incentive for students to complete the quizzes. Three attempts are permitted to allow students the opportunity to correct mistakes they've made.

PeerWise

PeerWise, while not the only collaborative MCQ repository system, is the first available for general use (Paul Denny, et al., 2008). Students create the questions, therefore guiding them to think both about the material that has been covered, what it means and possible misconceptions. The system also includes well thought out usability features to address the key issues of volume, accuracy and quality; that being how can the good questions in a large question bank be found with a reasonable level of assurance of having a correct answer. Generally speaking, a larger number of students will produce a more useful repository with less effort per member. For example, a class of 250 students can very easily create a repository of 300 questions, and if only 20% of these questions are "good" questions then there will be 60 useful questions, enough to create a good sized question bank. The issue is then sorting the "good" from the "bad" and this is handled by allowing students to rate questions as they complete them in terms of quality, a criterion that can be used for sorting.

The accuracy problem is handled by tracking user answers and allowing comments to be added after completing a question. If an answer that was suggested by the creator of a question is not the most popular, this is flagged and the creator is able to reconsider this in conjunction with the comments entered. PeerWise is hosted at the University of Auckland but available to, and used by, universities across the world. This means that there are no particular hardware or maintenance requirements. The flipside is that control over the system is very limited. There is also a feature where students can flag questions as being inappropriate which helps maintain a baseline of quality, although it is hard to stop students copying questions from other sources. Other than downloading participation and occasionally looking at flagged questions, there is very little administration. Completing all of the PeerWise assessment requirements was worth 2%. At the end of semester there were 326 questions in the repository. Of the 236 students, 159 had submitted questions, and 175 had answered questions.

The main areas investigated were student experience and correlations between achievement and final subject mark. The student experience was reviewed using the ubiquitous student questionnaire. Questions were targeted at the suitability of online assessment as being formative rather than summative. These questionnaires were administered in the final workshop and submitted anonymously. Achievement correlations were done by looking at PeerWise participation and Maple T.A. achievement level, compared to final mark achieved in the subject.

The Student Experience

The following questions were asked regarding the online assessments with answers given related to a standard Likert type scale where 1 indicates a strong negative response and 5 indicates a strong positive response. Note that the first question was actually asked for the individual modules, but has been combined, also students who said they had not participated in PeerWise have been excluded from the results.

Table 1 : Mean and Variance of student responses regarding online assessment.

Contribution to learning	**Mean**	**Variance**	**N**
How much do completing online quizzes contribute to your understanding of the subject material?	3.38	1.65	214
How helpful has creating questions for PeerWise been to your learning?	2.36	1.67	133
How helpful has answering questions in PeerWise been to your learning?	2.36	1.53	157
Suitability			
How suited do you feel online quizzes have been to your learning of the following topics?			
Digital Systems	3.59	1.45	214
Programming(MATLAB)	3.09	1.87	214
Mechanics	3.81	1.37	214
How suited do you feel PeerWise has been to your learning of the following topics?			
Digital Systems	2.35	1.66	167
Programming(MATLAB)	2.04	1.34	167
Mechanics	2.41	1.75	167

There are two clear conclusions that could be made from this data – that PeerWise was not popular, and online assessments are not equally suited to all topics.

There are a number of reasons why PeerWise may have been unpopular. The nature of PeerWise requires students to perform reflection and careful consideration and the condensed timeframe of the summer semester combined with the heavy assessment program does not lend itself to giving them the proper time to do this. The creation of questions and distracters requires synthesis and evaluation, and in terms of Bloom's taxonomy (Biggs, 2007) these are the deepest levels of understanding in the cognitive domain. If a student has not achieved the earlier levels then a push into these later levels may not be effective. In this case students would be likely to submit superficial questions, or slight variations on problems previously posed; neither of which will help the student learn. Another potential reason was that PeerWise was not properly promoted; most information about it was disseminated early in the semester with minimal follow up, a symptom of using a new system within a shortened semester. This will be changed in future semesters.

While PeerWise was not popular according the questionnaire, 36% (86 students) did more than the minimum level of participation suggesting that they saw value in the system, with 4 very enthusiastic students answering over 100 questions each, the most enthusiastic of which answered 192 questions. While question submission patterns were unremarkable, students appeared to value it as a study tool with moderate but consistent usage over the pre-exam study period.

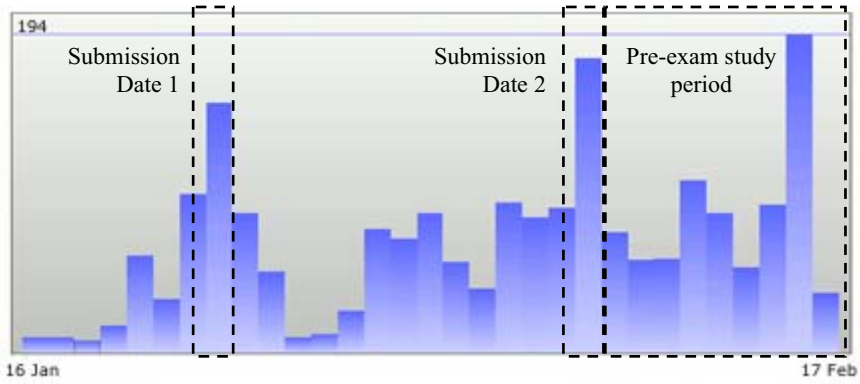

Figure 1 : Questions answered per day on PeerWise

The second conclusion made is that not all topics are equally suited to these forms of online assessment. Both mechanics and digital systems had a clear lead over programming in the online quizzes and in PeerWise, with mechanics being slightly ahead. This reflects the author's experience in creating questions. Mechanics was the easiest to set up as questions were largely mathematical in nature which suited Maple T.A. very well, and PeerWise reasonably well. Working through questions to identify shortcomings in knowledge and then reworking them to overcome these shortcomings is a traditional component of learning mathematics or physics, and a familiar process (Allen, 2003; Bonham, et al., 2003). The digital systems module requires a lot more diagrams such as circuits, Karnaugh maps and Boolean logic. While these can be done, they are more time consuming and can require compromise reducing the impact of questions. Learning the basics of programming on the other hand is best done by creating programs, with identifying "bugs" being key learning steps. The difficulty in adapting this process to questions and answers is a significant hurdle.

Achievement Correlations

Note that only the top 4 of the 5 quizzes were counted and as such are used in the average.

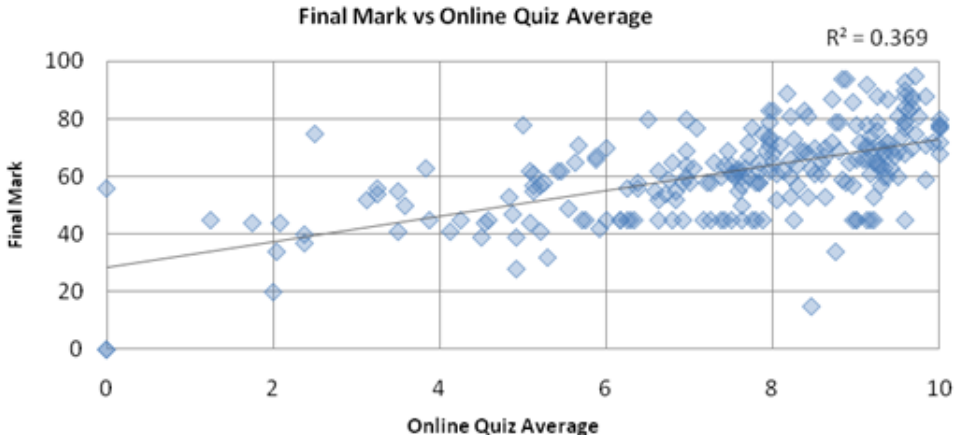

Figure 2 : Distribution of final marks against online average quiz marks

While there appears to be some correlation between the quiz mark and the final mark, statistical analysis suggests that quiz mark is in fact a very poor indicator of the final mark; the R^2 value of 0.369 being well below 1. One reason for this poor correlation is likely to be related to the high average quiz mark (7.48). While the online quiz is not reflecting deep student understanding of the material it may still be acting as a basic formative assessment.

As PeerWise is not assessed according to achievement, correlation has been based upon participation. Participation has been divided into three categories; no participation for students who have not submitted or answered any questions, some participation for students who have submitted or answered questions, but not enough to get full marks and required participation for students who have submitted at least 2 questions and answered at least 5 and therefore received full marks. The 2% awarded for completing the requirements for PeerWise is not considered to have a significant impact on the final mark.

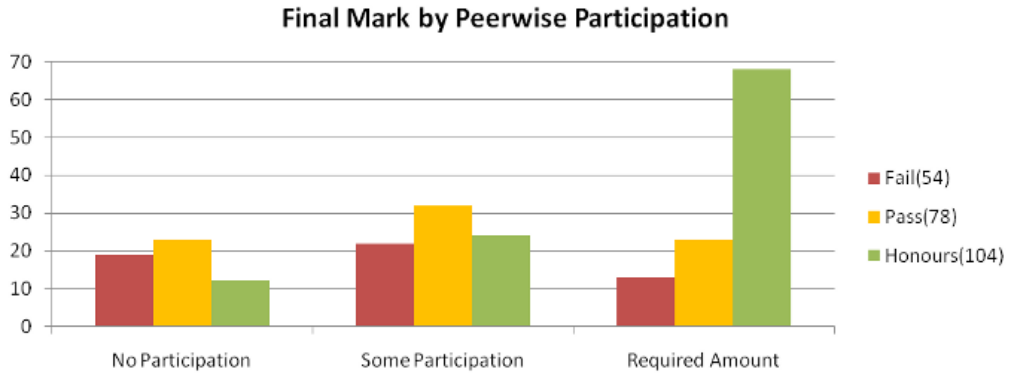

Figure 3. Final marks by Peewise Participation

While the achievement difference between students who did little or no participation and those who did at least the required participation is dramatic, it cannot be concluded that the relationship is causal. One obvious possible factor is that students who are high achieving are typically more likely to contribute to a greater degree in academic activity. Even so the strength of the correlation suggests that PeerWise deserves more consideration.

Conclusion

There are three key conclusions that can be made from the experiences gained in ESD 2 over the summer semester of 2010.

1. Not all online assessment systems are equal; whether or not an assessment tool should be used needs to be considered based on its capability and the resources available to support it.
2. Having a good online assessment tool is not enough; the way you promote and implement the tool can have a big impact on the way it is used and received.
3. A good tool is not necessarily right for all topics. Different topics may well need different tools or may not even be suitable for online assessment.

In more specific terms, the nature of the questions delivered in Maple T.A. need to be reviewed in terms of teaching value and learning objectives, particularly those used to cover MATLAB and programming. PeerWise will continue to be used, but better promotion is needed and more investigation is required into student capacity to create questions both in terms of understanding and time. Maple T.A. and PeerWise have a lot of potential, but potential which has yet to be fully realised.

References

Allen, G. D. (2003). A Survey of Online Mathematics Course Basics. *The College Mathematics Journal, 34*(4), 270-279.

Angus, S. D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology, 40*(2), 255-272.

Biggs, J., & Tang, C. . (2007). *Teaching for quality learning at university* (3rd Ed. ed.). UK/New York: McGraw Hill.

Bonham, S. W., Deardorff, D. L., & Beichner, R. J. (2003). Comparison of student performance using web and paper-based homework in college-level physics. *Journal of Research in Science Teaching, 40*(10), 1050-1071.

CSHE. (n.d.). On-line Assessment. Retrieved 16 June 2010, from http://www.cshe.unimelb.edu.au/assessinglearning/03/online.html

Denny, P., Hamer, J., Luxton-Reilly, A., & Purchase, H. (2008). *PeerWise*. Paper presented at the Proceedings of the 8th International Conference on Computing Education Research.

Denny, P., Hamer, J., & Luxton-Reily, A. (2009). *Students sharing and evaluating MCQs in a large first year Engineering course*. Paper presented at the Annual Conference for the Australasian Association for Engineering Education.

Kennedy, G., Judd, T., Churchward, A., Gray, K., and Krause, K. (2008). First year students experiences with technology: Are they really digital natives? *Australasian Journal of Educational Technology, 24*(1), 14.

Maplesoft. (n.d.). Maple T.A. Retrieved 18 June 2010, from http://www.maplesoft.com/products/mapleta/

Mestre, J., Hart, D. M., Rath, K. A., & Dufresne, R. (2002). The Effect of Web-Based Homework on Test Performance in Large Enrollment Introductory Physics Courses. *Journal of Computers in Mathematics and Science Teaching, 21*(3), 229-251.

Watson, J., & Angus, S. D. (2008). Does regular online testing enhance student learning? Evidence from a large first-year quantitative methods course. *The quantitative analysis of teaching and learning in higher education in business, economics and commerce : forum proceedings*, 139-159.