A COMPARISON OF BLACKBOARD CAA AND AN INNOVATIVE SELF ASSESSMENT TOOL FOR FORMATIVE ASSESSMENT

Graham Farrell, Vivienne Farrell and Ying K Leung

A Comparison of Blackboard CAA and an Innovative Self Assessment Tool for Formative Assessment

Graham Farrell: Swinburne University Australia
Vivienne Farrell: Swinburne University Australia
Ying K Leung: Hong Kong I.V.E.

Abstract

This comparative study analyses the responses of a cohort of students using both the Blackboard Multiple Choice Computer Aided Assessment (CAA) package and another Multiple Choice Question package that utilizes Confidence Measurement (MCQCM) for revision. The Blackboard test was the simple Multiple Choice Question (MCQ) format of a stem followed by four simple text options; the other utilized the traditional MCQ format with multiple possible answers. The MCQCM evolved from a series of studies in the IT discipline where the system was used over sequential semesters to elicit feedback to improve its functionality and usability. A total of 74 students completed a questionnaire as part of the standard subject evaluation for the participating subjects. The questions were designed to evaluate their opinion of the testing procedure identifying their preferences and concerns. The results were analyzed producing some encouraging observations. It was observed that there was a strong overall acceptance of the innovative self assessment tool MCQCM as a valuable contributor to the learning experience. In addition, the MCQCM was rated equally by the participants to the Blackboard CAA in influencing their direction of learning and identifying the areas of concern. The results show that MCQCM has promise and recommends further study.

Introduction

Assessment plays a critical role in the educational process as both a means of grading and supplying valuable feedback to the student. Technology plays an integral part of the delivery of education increasing the availability of effective assessment tools for the student and encouraging self-assessment at all stages of the learning experience. It is accepted wisdom that testing for the purpose of feedback should be a “routine part of the ongoing classroom activity rather than an interruption”. (See the Principles and Standards for School Mathematics (2000)).

Traditionally formative and summative methods of assessment are reliant on the instructor to supply the feedback. With the ever-increasing demand on the instructor's time, unintentional delays in supplying feedback to the student can occur to the disadvantage of both the student and the instructor. Often the most valuable feedback occurs at the final stages of the learning path, giving no real assistance to the student's learning progress and only limited feedback of the student's progress to the instructor. Feedback must be instantaneous, which is an inherent characteristic of online assessment tools due to the nature of the encompassing technology. The formative assessment feedback cycle should occur early in the learning experience and often throughout the duration of the subject.

This paper peruses an innovative approach to a web based self assessment tool based on the traditional Multiple Choice Questions format that also facilitates the registration of the students level of confidence with their response to each of the options available. The generated students score is designed to reflect their understanding of the reviewed topic.

Doebbert (1999) emphasizes the need for the student to develop skills in managing and controlling his/her learning with the utilization of technology assisting in the process as they travel the educational path. In addition it is also important that systems that appear to provide a multitude of benefits to the students and instructors should be pursued with vigor.

Literature Review

Black and William (1998) refer to "Assessment" as being a group of activities undertaken by both teachers and students in assessing themselves, generating feedback and grades.

Instructors generally have an appreciation of the different choices of assessment and often choose the correct type for a particular purpose (Assessment Tools, 2003).

Multiple Choice Questions (MCQs) continue to be a popular assessment option for instructors today. Bacon's (2003) research demonstrates that MCQs, when constructed properly, have the ability to test a broad range of fields at various levels effectively and efficiently. MCQs thrive in the hypermedia environment, leading to numerous online MCQ testing packages available to instructors. An appealing feature of these computerized MCQ's is their ability to randomly generate questions from a predefined test bank, automatically give answers, suggest direction for the learner, offer adaptive learning strategies, automatically record grades and monitor the students' progress, all contributing to the students' management of their learning path.

MCQs offer a good method of assessing the students knowledge in areas deemed as right or wrong, but what about the levels of uncertainy, “shades of gray” or fuzzier areas? Diamond and Forrestor (1973) define knowledge as asking the question “What do you know?” followed by the meta-question “How sure are you of the answer to the question about what you know?”

Davidoff (1995) recognizes the need for a more thorough approach in assessing students in the medical area stating that much of the medical knowledge is “incomplete, ambiguous and conflicting” and that the standard MCQ testing method does not facilitate or reflect the students’ level of knowledge or their confidence in their level of knowledge. In most cases MCQs recognize and reward those areas of knowledge that is either right or wrong, encouraging guessing and often contributing to overconfidence. Davidoff (1995) considers miss-calibration of confidence is as equally concerning as lack of knowledge.

The scoring of MCQs has long been a point of discussion resulting in numerous scoring models designed and implemented to eliminate the "guessing” component. Brown and Shuford (1973) produced an MCQ calibrated scoring system designed to “encouraged honesty”, utilizing a form of confidence measurement. Pollard and Clarke (1989, 1986, and 1993) offered alternative scoring techniques where random guessing is eliminated by introducing a reward and penalty structure in the scoring system.

Pollard’s (1989) grading system is designed to reward the student who demonstrates confidence in their choices. Pollard’s (1989) system relies on the combination of boxes ticked with many variations that have to be considered.

Klinger (1997) utilized a system of scoring permitting the student to choose a position on a triangle where the apexes represented the three optional answers. He proportionally divided the lines joining the apexes into segments, which permitted the student to nominate their level of confidence. The student was then scored accordingly. In addition this assessment tool rewarded the student for choosing to declare that they were completely unsure of their answer.

This is also the case with other designed systems such as Paul's (1994) interactive response system, which requires the student to choose an option demonstrating a level of confidence in their answer.

In summary, the previous research has provided a good foundation for the development of formative assessment tools to the benefit of all concerned. The research of Pollard (1989) demonstrates the advantages of developing a scoring system that reflects the students understanding of the subject while Paul (1994), Brown and Shufford (1993) and Klinger (1997) have all demonstrated variations on systems designed to capture a numerical representation of the students confidence in their level of understanding.

The Multiple Choice Questions with Confidence Measurement Tool

This final presentation of the Multiple Choice Question with Confidence Measurement (MCQCM) has evolved after much discussions and investigative pilot programs to its present form (Farrell,Leung 2002). The initial investigations used a primitive design of the system for a group of students studying C++ eliciting feedback to improve the design (Farrell,Farrell,Leung 2001). Further studies continued to evaluate the system for cohorts of Information Technology students, producing this final operational state (Farrell,Leung 2003). In essence it is based on the traditional Multiple Choice Questions format as outlined by Kehoe (1995) and Frary (1995), consisting of a stem with four options. The MCQCM has been developed and refined to engage the user while exercising sound navigational properties for delivery across the Internet. In additional the system captures and records the scores of the students as they participate in the exercise.

The system permits more than one correct answer. This encourages the student to consider all options separately and not to identify what they consider to be the single correct answer and ignore the rest. The advantages of this approach compared to the traditional MCQ format are as follows:

- It permits the instructor to word the options to closely examine the areas of study, eliminating the need to use easily recognizable distracters.
- It forces the student to consider all options carefully, increasing their exposure to associated areas within the topic.
- The resulting score reflects the true confidence of the student in their knowledge of the subject.

The consequential feedback to the students is a simple reflection of their understanding of the concept being considered in each question.

The scoring technique is simple, placing the user in direct control of the results of their actions. The MCQCM requires the participant to commit to a “level” then to a “strength”. This is in keeping with Bandera’s (1983) work on self efficacy and Betz and Hacket’s (2002) application of same. The student first registers the level, in this case identifying the option as either being True or False. After that has occurred they are required to declare their confidence in their choice as a percentage. Registering 100% states complete confidence of their choice, while a low percentage represents very little confidence.

This is done for every option in each question.

Once the question has been completed the scoring is calculated by first identifying if they have chosen the “level” correctly (True or False) and then allocating a mark for each option from -10 to 10, depending on their registered level of confidence.

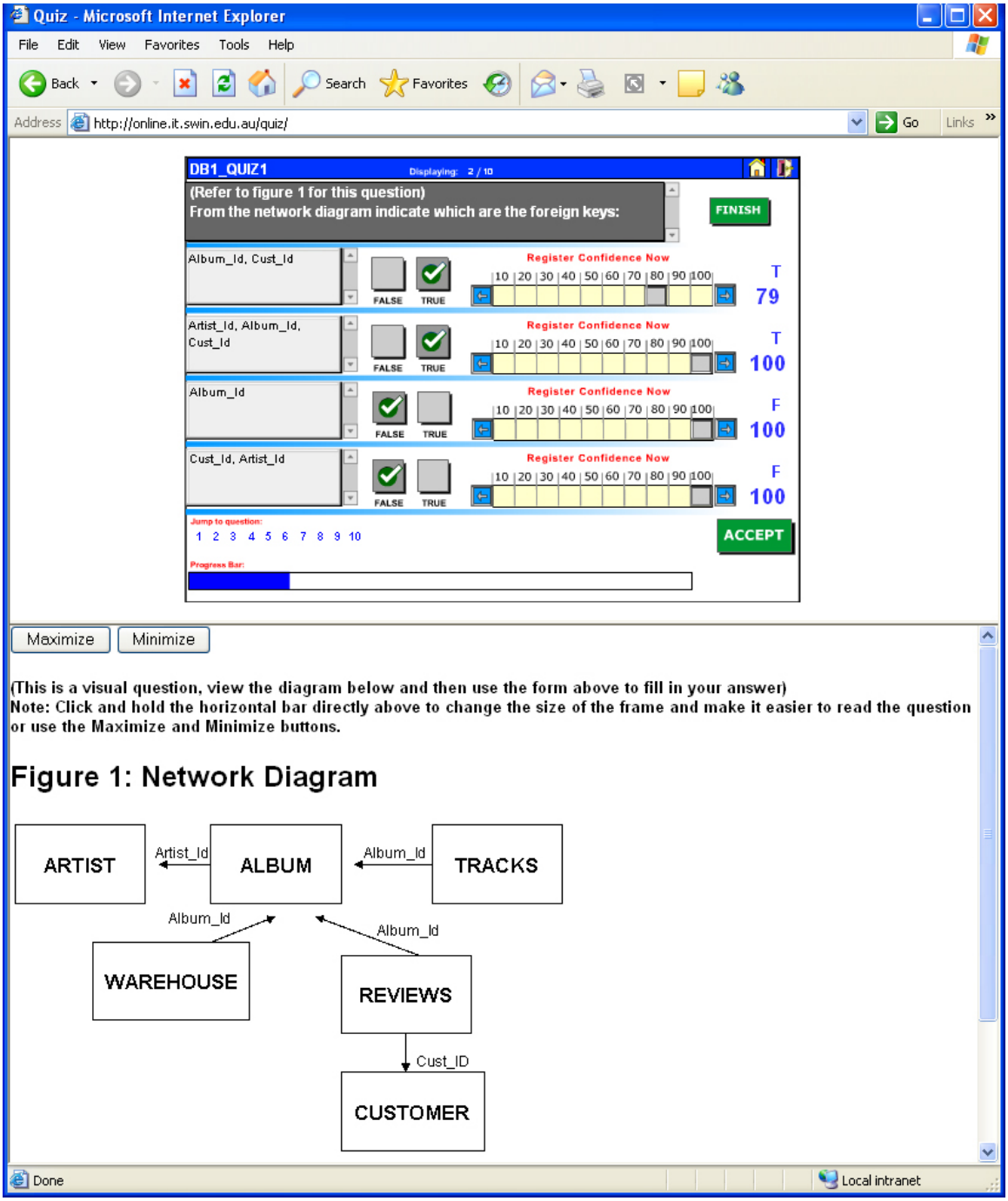

Figure 1. MCQCM display permitting the student to select the correct answers and state their level of confidence. In this case the question required the student to view a diagram directly related to the question

This tool is designed to reflect the student's level of understanding. The students are placed in the position where they must clearly state their level of confidence, knowing that they would be proportionally penalized for an incorrect choice and proportionally rewarded for a correct choice. The tool also permits the use of

questions that have multiple correct answers, which requires them to identify all possible correct combinations and scores accordingly.

At the end of the testing session the students are shown a diagrammatic representation of their knowledge for each of the questions. The diagram is colour coded highlighting the greatest areas of concern in red, the areas worth revising, areas of less concern in blue and the areas of no concern in green. See Fig 2 for the results display

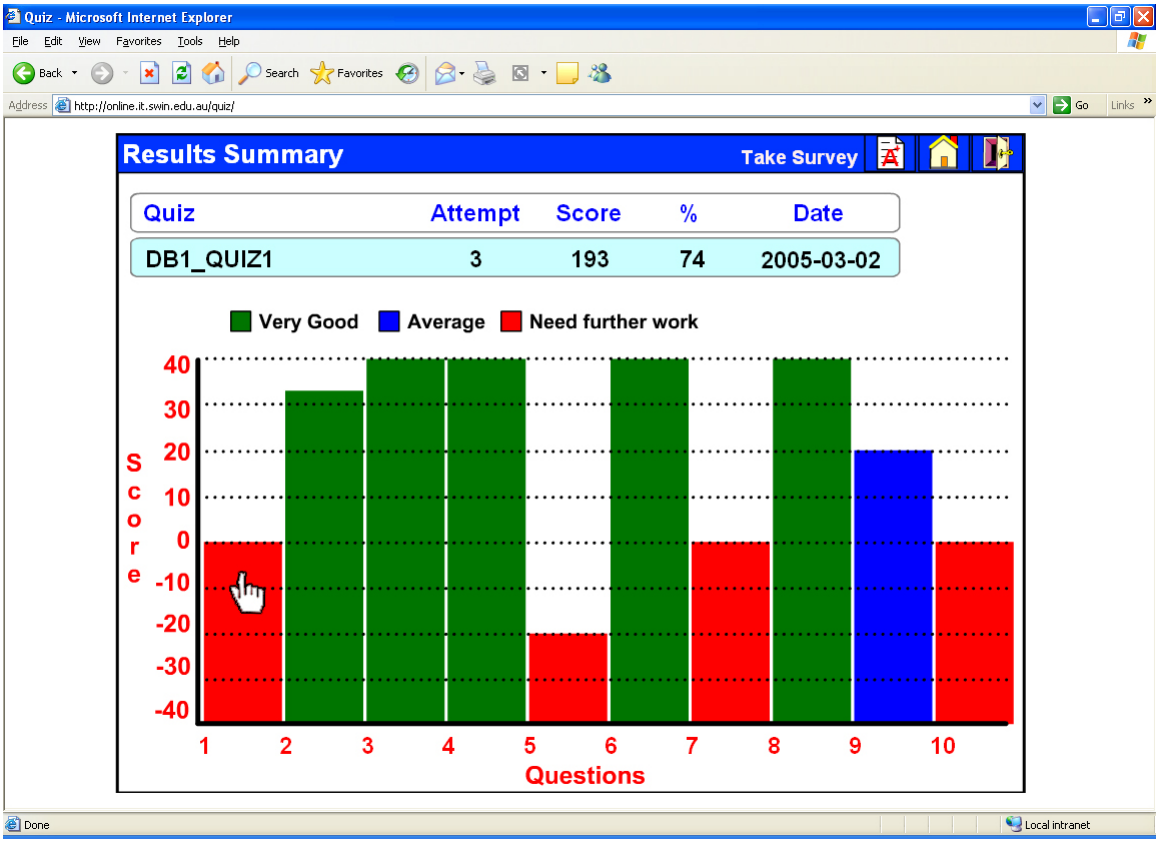

Figure 2. The Feedback screen showing the students score for each question, colour coded to represent the level of concern. The student can link directly to the question, view the correct answers and their response in order to gain assistance

This permits the students to hone in on the particular areas of concern by using the hot spot link and see the correct answers and most importantly their incorrect choices.

Method

The two subjects that this report focuses on are Database 1 (DB1) and Advanced Web Technologies (AWT). Throughout the semester all of the students were offered a series of CAAs as revision exercises. The first type consisted of the

traditional BB multiple choice questions (MCQ) self assessment tests consisting of a stem followed by 4 options, one being correct. The other was an innovation based on the traditional MCQ format with the requirement for the participant to register their level of confidence for each option. We refer to this tool as the Multiple Choice Questions with Confidence Measurement (MCQCM) as described above.

At the end of the semester the students were requested to complete a questionnaire on various aspects of the subject as part of the standard subject review process. Included was a series of questions that focused specifically on the Blackboard CAA and the MCQCM revision tests that they completed. The total number of students surveyed was 74, comprising of 41 DB1 and 33 AWT students. The data was collected and analyzed.

4.1: Demographics:

Subject/s	*Post Grads*	*PG's >25 yrs*	*Undergrads*	*UG's >25 yrs*
Database 1	30%	75%	70%	33%
Advanced Web Technologies	0%	0%	100%	66%

4.2: Test Availability for DB1 and AWT:

Both cohorts of students were offered the self assessment test online, permitting them to complete them at their convenience either in the labs, at home or any other location of their choice where they had internet access.

4.3: Surveys Structure:

The subject evaluation survey contained 8 questions, 5 specifically designed to gauge the usefulness and effectiveness of the tests, and 3 to compare the two testing methods. In addition the participants were also requested to comment on both the good and bad aspects of the tool.

The conveners of the subjects consider the response rate was medium to high for both cohorts , although exact figures are not available.

Results and Preliminary Analysis

There were no apparent differences between the two subject groups or the post graduate and undergraduate students. The preferences were consistent across all cohorts.

To assist the reader the summaries have been grouped together under each of the survey questions below.

The first five questions refer specifically to the MCQCM tool. The remaining three questions are specifically designed to compare the BB CAA to the MCQCM.

Q: 1 How would you rate the MCQCM testing method as part of your learning process?	**No help**	**Some assistance**	**Extremely helpful**
	9%	75%	16%
Q: 2 How often would you use MCQCM if available at any time?	**Never**	**Sometimes**	**Regularly**
	5%	50%	45%
Q: 3 To what level would the MCQCM influence your direction and path of your learning?	**Not at all**	**Sometimes**	**Substantially**
	13%	66%	21%
Q: 4 When viewing the MCQCM results display how clear were the scores?	**Unclear**	**Clear**	**Extremely Clear**
	13%	60%	27%
Q: 5 When looking at the MCQCM display how clearly could you identify the problem areas?	**Unclear**	**Clear**	**Extremely Clear**
	19%	52%%	29%
Q:6 How would you rate the MCQCM feedback to the BB feedback?	**Same**	**Better**	**Much Better**
	17%	63%	20%
Q: 7 Which of the two, BB or MCQCM, was the best in directing you with your revision?	**BB**	**Neither**	**MCQCM**
	25%	25%	50%
Q: 8 Which of the two, BB or MCQCM, better informed you of your understanding of the topics?	**BB**	**Neither**	**MCQCM**
	33%	33%	33%

Conclusion and Discussion

The results of this study were encouraging. The extensive use of CAAs as a formative assessment tool is widely spread and in general well received by the students

The discussion that follows summarizes the main observations and attempts to explain some of these in context.

It was observed that a significant number of the students consider the MCQCM as a good self assessment tool. Requests were made from students and instructors that it be made available on a weekly basis linked in with the lectures. In addition they commented on the ability of the system to display complex diagrams beyond the scope of many traditional MCQs.

The greater majority of the students confirmed that they would use the MCQCM tool during their studies with some further requesting it to be available for other subjects. The majority stated that the results would influence their journey along the learning path identifying the areas of concern.

The more interesting observations occurred when we compare the two offered methods of self assessment made available. Even though the greater majority

appreciated the MCQCM feedback over the BB MCQ, the preference for each was equal with regard to the direction of revision and the understanding of the topics. It was encouraging to observe that the MCQCM rated evenly against the long standing, well established standard MCQ format. The students have had previous exposure to the Blackboard CAA which could add to their “comfort” level when using the new MCQCM.

In conclusion, it is very encouraging to find the greater number of students appreciate the MCQCM tool as a valuable self assessment exercise. In particular, the confirmation that for most students the feedback was considered better than the traditional MCQ format and that it rated equally as good as the traditional MCQ in directing the students in their learning and informing them of their understanding of the content. As with all new technology it can take some time before the full uptake is realized. Given the time that students have been exposed to BB MCQ’s it is quite feasible that long term exposure to MCQCM could raise the comfort level of an even larger contingency of students over time.

As part of the ongoing commitment to this research we intend on utilizing the MCQCM as part of their formal summative assessment. Some students requested that we do so as they could see great benefit in being exposed to this type of self assessment during the semester and then having it as part of their final exam.

References

Ashburn, Robert (1938). An experiment in essay-type question. *Journal of Experimental Education 7 (1): 1-3*

Assessment tools for Assessment, Evaluation and Curriculum Redesign workshop: *month 7 http://www.thirteen.org/edonline/concept2class/month7/index_sub2.html (Last accessed Aug 2003)*

Bacon, Donald R (2003): Assessing Learning Outcomes: A Comparison of Multiple-Choice and Short-Answer Questions in a Marketing Context: Journal of Marketing Education. Vol 24. No 22. Sage publications [not yet published]

Black. P and William D (1998): Inside the Black Box: Raising Standards Through Classroom Assessment. Phi Delta Kappan October 1998· Volume 80· Number 2 P 139-149 http://www.pdkintl.org/kappan/kbla9810.htm

Black. P and William D (March 1998): Assessment and Classroom. Learning *Assessment in Education, Vol 5 March* P. 7-74

Bloom, Benjamin S (1956): Taxonomy of educational objectives, hand book 1: Cognitive domain. *New York: Longman Green.*

Farrell, G Farrell V and Leung Y (2001), Online Software Test for Efficient and Effective Assessment Using Multiple Choice Questions: An Evaluation. *Conference Proceedings American Educational Research Association*

Farrell, G and Leung Y (2003), Designing an Online Self-Assessment Tool Utilizing Confidence Measurement. *Conference Proceedings IFIP 8.4 WG*

Farrell, G and Leung Y (2002), Inproving the design of an Online Self-Assessment Tool Utilizing Confidence Measurement. *Conference Proceedings Web Based Learning*

Lambert W.T. Schuwirth and C.P.M. van der Vlueten (1996): Quality Control: Assessment and Examinations: *http://www.oeghd.or.at/zeitschrift/1996h1-2/06_art.html (Last accessed Aug 2003)*

Principle and Standards for School Mathematics (2000): *National Council of Teachers of Mathematics - Standards 2000 Project Chpt 2*

Wilson, R. B. and Case, S. M.: Extended Matching Questions: An Alternative to Multiple-choice or Free-response Questions: Journal of Veterinary Medical Education. Volume 20:3. *http://www.utpjournals.com/jour.ihtml?lp=jvme/jvme203/ExtendedMatchingQuestions.html* (*Last accessed* Aug 2003)