

# An Inquiry into the Effectiveness of Student Generated MCQs as a Method of Assessment to Improve Teaching and Learning

Damien Hutchinson, Jason Wells

School of Information Technology, Deakin University, Melbourne, Australia  
Email: [drh@deakin.edu.au](mailto:drh@deakin.edu.au)

Received June 3<sup>rd</sup>, 2013; revised July 3<sup>rd</sup>, 2013; accepted July 10<sup>th</sup>, 2013

Copyright © 2013 Damien Hutchinson, Jason Wells. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In anticipation of helping students mature from passive to more active learners while engaging with the issues and concepts surrounding computer security, a student generated Multiple Choice Question (MCQ) learning strategy was designed and deployed as a replacement for an assessment task that was previously based on students providing solutions to a series of short-answer questions. To determine whether there was any educational value in students generating their own MCQs students were required to design MCQs. Prior to undertaking this assessment activity each participant completed a pre-test which consisted of 45 MCQs based on the topics of the assessment. Following the assessment activity the participants completed a post-test which consisted of the same MCQs as the pre-test. The pre and post test results as well as the post test and assessment activity results were tested for statistical significance. The results indicated that having students generate their own MCQs as a method of assessment did not have a negative effect on the learning experience. By providing a framework to the students based on the literature to support their engagement with the learning material, we believe the creation of well-structured MCQs resulted in a more advanced understanding of the relationships between the concepts of the learning material as compared with plainly answering a series of short-answer questions from a textbook. Further study is required to determine to what degree this learning strategy encouraged a deeper approach to learning.

*Keywords:* Assessment; Multiple-Choice Questions; Student Generated Content; Computer Security; Information Technology

## Introduction

Assessment refers to the processes used to evaluate student achievement of the learning objectives of a unit or course (Phye, 1997: p. 525). “Of all the activities associated with teaching and learning, assessment have the potential to have the most influence in directing students’ energies and in determining their approach to learning in a unit of study” (Elton & Johnston, 2002). In order to assess understanding and measure learning, students are typically given an assignment that may vary from problem-solving exercises to project work and essay writing. One of the critical challenges the teacher faces as part of the assessment process is how to encourage a deeper understanding of the key areas of the curriculum (Palmer & Devitt, 2006). Teachers prefer students to undertake their subject with the intention of engaging with it and understanding it (a deep approach) rather than with the intention of merely reproducing material without having a good understanding of it (a surface approach) (Elton & Johnston, 2002).

During our teaching experience as an instructor and lecturer over the past decade, it has become apparent that few students engage with the learning material on a progressive basis. To develop quality assessment tasks that prepare students for professional practice it is imperative to measure the effectiveness in terms of student learning that these tasks impart. The student cohort at the centre of this study was enrolled in a unit called

Introduction to Computer Security (ITCS). This is a first year unit offered on and off campus at Deakin University in Victoria, Australia. ITCS is one of eight units offered as part of the IT security major leading to a Bachelor of IT (IT Security) by the School of Information Technology, within the Faculty of Science, Engineering and Built Environment. The unit may also be taken as an elective. In ITCS students are required to learn the principles of computer security for the protection of company assets and information systems of an organization.

The current assessment tasks for ITCS require the students to read the textbook and related material then provide solutions to short answer questions. These solutions are subsequently marked and a grade is returned to each student with feedback and a suggested solution. The solutions provided by the majority of students were largely reproduced from the textbook. This made it extremely difficult to determine if the students had actually learned the material or just taken a surface approach with the intention to complete the task requirements by reproducing material (Ramsden, 1988: p. 19).

Applying the advice of Brown, Rust & Gibbs (1994); “if you want to change student learning then change the method of assessment”, having students create their own MCQs was implemented as the new method of assessment. The basis for this decision was that student-questioning as a learning strategy plays a crucial role in students’ active processing of given materials (Wong, 1985); for students to be active learners and independ-

ent thinkers, they must generate questions that help shape, focus, and guide their cognition during the learning process” (Singer, 1978). Furthermore, several authors consider the student generated question as an enabler for participants’ cognitive elaboration and as an effective alternative for achieving meaningful learning by reinforcing higher-order thinking skills (Wong, 1985; English, 1998; King, 1995; Leung & Wu, 1999; Silver, 1994).

This paper delivers the findings of a study into the use of an alternative assessment approach involving students creating MCQs. Previous studies where students created MCQ’s were found to be successful (Yu & Liu, 2004b; Palmer & Devitt, 2006). Studies indicate that students could potentially be motivated to study a subject in greater depth if provided with the opportunity to write their own problem-solving exercises (Sircar & Tandon, 1999).

By having students create MCQs it is anticipated that students would be involved in a deep approach to learning by having to both engage with and understand the ITCS principles rather than just reproduce them. It is expected that this approach will add value to teaching by encouraging the student to develop a deeper approach to learning by providing a new and different assessment method that motivates greater interaction with the unit material (Toohey, 1999: p. 167). It is also expected that this approach will provide formative feedback quickly to the students as well as provide quantitative data (Censeo, 2007) and insights into those areas on the unit material that require more emphasis or improvement. A repository of MCQ’s will also be created by the students that can be provided back to them as an additional learning resource. The aim of the study was to measure the effect of students constructing their own MCQs on learning. The hypothesis to be tested was that when given a topic to research and write on, student understanding would be enhanced if they were asked to construct MCQs based on the material under study.

## Background

This section provides an overview of the literature surrounding the use of MCQs for assessment. The premise for using the MCQ as a method of assessment is outlined together with the associated advantages and disadvantages in relation to teaching ITCS. Second details of how the MCQ can be used to achieve active learning is described within the context of two other studies that used the construction of student generated MCQs as the method for learning.

The 21st century has brought new challenges for teaching and learning in higher education. In particular assessment needs to be innovative in practice, responsive to individual needs and relevant to real life. Both Parry (2004) and Parker (2004) emphasize the need for a learning paradigm that uses more formative, technology-mediated assessment. Moving in this direction, **Table 1** provides an analysis of the relationship between key elements of assessment and the use of MCQs for assessment. With the intention of using student generated MCQs and software to support the assessment process, the third column of **Table 1** presents how this relationship is applied to teaching computer security by highlighting the advantages and disadvantages of using MCQs as an assessment method.

## Using Student Generated MCQs to Achieve Active Learning

The literature reviewed indicates some work has been under-

taken in the area of student generated questions for learning and assessment purposes. This section presents two examples of studies that applied the concept of student generated MCQs for the purpose of achieving active learning.

Palmer and Devitt (2006) conducted a study about constructing MCQs as a method for learning. Their study was performed to measure the influence of student-based construction of MCQs as a stimulus for the learning and understanding of topics in clinical surgery (Palmer & Devitt, 2006). The hypothesis of this study was that in order to create high quality MCQs, deeper understanding of the material involved in the question would be created (Palmer & Devitt, 2006). If the students can understand the method of construction of a good MCQ, “it will teach the difference between mere knowledge acquisition and how that knowledge can be used in terms of comprehension, application, analysis and evaluation” (Palmer & Devitt, 2006).

At the completion of the study it was found that the MCQs created were of a high standard and the students looked more favorably at the exercise, although many found that it did not replace traditional teaching methods (Palmer & Devitt, 2006). In terms of assessment the students “produced a large bank of potentially viable MCQs which are perfectly suitable for both formative and summative assessment purposes” (Palmer & Devitt, 2006). The following outlines some of the other significant findings of this study (Palmer & Devitt, 2006):

- one quarter of the questions created were evaluated to be capable of testing higher order cognitive skills;
- the majority of the MCQs created tested knowledge and comprehension;
- both teachers and students need to be educated about the benefits of untried learning methods for them to be more receptive to new learning initiatives;
- an unanticipated weakness of the learning strategy in this study (which is often viewed as a strength of an MCQ assessment) was the inability to assess a wide range of material; “the students may have gained a deep understanding of their particular area of study, but they obtained only a superficial understanding of other areas, thus showing no substantial net increase in their overall understanding”.

The authors suggest that this problem may be rectified if students are required to create a larger number of MCQs. However we would suggest that this would only create a larger bottleneck in terms of marking. Instead we would recommend an alternate approach which would involve peer review of the created MCQs. This would be a topic for future research.

Yu & Liu (2004a) conducted their study into the perceived potential value of student MCQ construction in the introductory physics laboratory. The study was performed to determine whether a more active learning atmosphere in physics labs would be cultivated by having students construct MCQs pertaining to interacting physical phenomena during the learning process (Yu & Liu, 2004a).

The hypothesis of this study was that the creation of MCQs would stimulate students to be more attentive to their tasks and be more reflective on their own thinking and learning (Yu & Liu, 2004a). Similar to the study performed by Palmer and Devitt (2006), in the process of constructing the MCQs, students would be more likely to be intellectually active in order to generate a question-stem, provide the correct answer to the posted question and ponder three plausible distracters that can essentially distinguish the students who have learned the concepts from those who have not (Yu & Liu, 2004a). In terms of active

**Table 1.**

Relationship between key elements of assessment and the use of MCQs for assessment in relation to our teaching.

| Key Elements of Assessment   | Assessment Using MCQs  | Teaching Computer Security   |  |
|--|--|--|--|
|  |  | Advantages   | Disadvantages  |
| <p>3 contemporary assessment issues in Australian higher education (Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• potential of online assessment;</li> <li>• assessment design for large classes;</li> <li>• academic honesty;</li> </ul>  | <p>Immune to the influence of bluffing (not guessing) (Burton, Sudweeks, Merrill, &amp; Wood, 1991).</p>   | <p>(Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• test knowledge quickly with large groups;</li> <li>• students own work;</li> </ul> |  |
| <p>Formative assessment (improving):</p> <ul style="list-style-type: none"> <li>• tasks that form developmental or ongoing teaching/learning process;</li> <li>• provide ongoing feedback to the student to improve their learning &amp; decision making (Lublin, 2000; Phye, 1997: p. 514).</li> </ul>  | <p>Excellent opportunity to offer efficient feedback:</p> <ul style="list-style-type: none"> <li>• Comments usually limited to indicate score;</li> <li>• Need explanation for wrong answer and marking key &amp; explanation for correct answer (Ramsden, 2003: p. 188);</li> <li>• Total score—how the test taker compares to others.</li> </ul>   | <p>(Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• prompt feedback on student performance.</li> </ul>                                 | <p>(Phye, 1997, pp. 16-17; Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• Difficult and time consuming to design and develop e.g. 1 hour per question (ACVIM, 1997); Feedback often withheld to save expense of constantly designing new tests (Lublin, 2000).</li> </ul> |
| <p>Summative assessment (documenting):</p> <ul style="list-style-type: none"> <li>• tasks that often occur at the end of a unit;</li> <li>• used primarily to indicate how much student has learned e.g. assign grades (Lublin, 2000).</li> </ul>  | <p>Excellent opportunity to offer efficient feedback:</p> <ul style="list-style-type: none"> <li>• Section score—strengths and weaknesses can be understood; Summative MCQ tests work best when question bank built up over time: <ul style="list-style-type: none"> <li>◆ from questions subjected to item analysis;</li> <li>◆ used to indicate its degree of difficulty.</li> </ul> </li> </ul>   |  | <p>Students' tend to revert to rote-learning mode when they know they will be tested by MCQs (Lublin, 2000).</p>   |
| <p>Reliability of assessment—extent to which results of assessment can be trusted (Miller, 1987: p. 73; Toohey, 1999: p. 180; Lublin, 2000).</p>   | <p>(Burton et al., 1991):</p> <ul style="list-style-type: none"> <li>• Reliable; no argument about the right answer;</li> <li>• Objectively scored;</li> <li>• Not affected by scorer inconsistencies (like essay questions). (Phye, 1997: pp. 16-17; Lublin, 2000):</li> <li>• Marking is reliable, cheap, &amp; mechanically scored;</li> <li>• Can be reused/item banked.</li> </ul>  | <p>High reliability &amp; motivating regular study (Toohey, 1999: p. 173).</p>   |  |
| <p>Validity of assessment—extent to which the assessment reflects the learning objectives of the unit (Miller 1987: p. 73; Toohey, 1999: p. 180; Lublin, 2000).</p>  | <p>Can be invalid (Burton et al., 1991). Group of people should be involved in MCQ preparation:</p> <ul style="list-style-type: none"> <li>• Generating the questions and trialling the questions.</li> </ul>  |  | <p>(Phye, 1997: pp. 16-17; Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• surface approach to learning i.e. swotting for recall. At mercy of one person's way of communicating the discipline.</li> </ul>   |
| <p>Learning objectives</p> <ul style="list-style-type: none"> <li>• Good assessment starts with author determining objectives &amp; use of results (Phye, 1997: p. 186; Lublin, 2000; Censeo, 2007).</li> </ul> <p>Student understanding of assessment process facilitated by (Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• explaining assessment approaches, how approaches relate to the unit's objectives &amp; criteria by which students will be assessed.</li> <li>◆ e.g. problem-solving abilities; vs. recall of information only.</li> </ul> | <p>MCQ Design:</p> <ul style="list-style-type: none"> <li>• A standard multiple-choice test item consists of two basic parts: <ul style="list-style-type: none"> <li>◆ 1. a problem (stem); question or an incomplete statement</li> <li>◆ 2. list of suggested solutions (<i>alternatives</i>); contains one correct or best alternative (<i>answer</i>) and a number of incorrect or inferior alternatives (<i>distractors</i>) (Burton et al., 1991).</li> </ul> </li> <li>• Select and write down the general content area;</li> <li>• Then write down the testing point. According to (Burton et al., 1991; ACVIM, 1997; Phye, 1997: p. 521; Lublin, 2000; Censeo, 2007) criteria for writing a "good" MCQ include keeping the: <ul style="list-style-type: none"> <li>• alternatives homogeneous in content;</li> <li>• alternatives free from clues as to which response is correct;</li> <li>• grammar of each alternative consistent with the stem;</li> <li>• alternatives parallel in form;</li> <li>• alternatives similar in length.</li> </ul> </li> </ul> |  | <p>(Burton et al., 1991): Low (knowledge of terms, facts, methods, and principles) vs. high-level (comprehension, application, and analysis) objectives.</p>   |

**Continued**

|   |  |  |  |
|---|--|--|--|
| <p>Assessment Objectives:</p> <ul style="list-style-type: none"> <li>• must specify observable, preferably measurable, changes in the learner's behaviour at the end of the course (Miller, 1987: p. 10);</li> <li>• help teachers plan more efficiently for students' learning experiences (Miller, 1987: p. 14).</li> </ul> <p>Category of educational objective for higher education:</p> <ul style="list-style-type: none"> <li>• Cognitive—objectives are those dealing with the acquisition of knowledge; recall, processing, stating information (Bloom, 1956; cited in Miller, 1987: p. 51).</li> </ul> | <p>Suggestions for recommended actions to take based on test results e.g. self-study plan (Censeo, 2007).</p> <p>Students appreciate feedback that assists them to improve future efforts (James, McInnes, &amp; Devlin, 2002).</p> <p>Promotes the development of deep approaches to learning (Ramsden, 2003). Designed for the cognitive levels they can test (ACVIM, 1997);</p> <ul style="list-style-type: none"> <li>• e.g. Teaching IT Security; securing a small business information system.</li> </ul> <p>MCQ: Knowledge (+); Comprehension (+); Application (+); Analysis (?); Synthesis (-); Evaluation (?). Key: (+) = usually; (-) = sometimes; (?) = rarely (Bloom, 1956, cited in Miller, 1987: p. 51).</p> | <p>Students' tend to revert to rote-learning mode when they know they will be tested by MCQs (Lublin, 2000).</p> <p>(Burton et al., 1991):</p> <ul style="list-style-type: none"> <li>• Comprehend principles;</li> <li>• Apply principles to new situations &amp; solve problems.</li> </ul> <p>Allow broad coverage, higher order thinking (Lublin, 2000; Censeo, 2007; Phye, 1997).</p> | <p>(Burton et al., 1991; Lublin, 2000):</p> <ul style="list-style-type: none"> <li>• unique thinking processes;</li> <li>• progressive problem solving; providing examples.</li> </ul> |
|---|--|--|--|

learning they found that MCQ construction helped make students monitor consciously and actively their own learning' and concluded that MCQ construction "is an instructional strategy with great potential" (Yu & Liu, 2004a). The following outlines some of the other significant findings of this study (Yu & Liu, 2004a):

- by identifying important principles or concepts that were difficult to comprehend students engaged in cognitive strategies skills indicative of active, self-regulative learners, including rehearsing, organizing and continuously monitoring their state of cognition;
- it made students "think and reflect more on physics-related questions and phenomena", and "discuss more frequently and intensely with group members".
- "it seemed to transform the classroom into a more reflective and inquisitive learning atmosphere that are more active and interactive".

Further studies conducted by Yu & Liu into active learning through student generated questions in physics experimentation classrooms continued to support the positive potential of students creating their own MCQs. Their findings were that "MCQ generation activity qualitatively changed students learning behavior, and helped students become more active learners" (Yu & Liu, 2004b).

From these studies it is evident that using student generated MCQs as an alternative form of assessment offers potential for enabling active learning. However as indicated there are several issues that must be considered for success of this learning method to be realized.

### Methodology

The aim of this study was to show that a student's knowledge on a given topic will increase after they create MCQs relating to that topic, and that by following predefined criteria and guidelines, students would create high quality MCQs which can be then redelivered in an online environment as either assessment or self-review tests for revision. This section presents a description of the methodology undertaken for this study.

### Method for Collecting Evidence

The method for collecting evidence involved performing an

experiment. The aim of the experiment was to compare the results of a pre-test set of MCQs with a post-test set of MCQs completed by the participants. Part of the experiment involved analyzing the quality of students' formal assessment and using this as evidence of improved learning. The following is a description of the procedure undertaken.

### Procedure for the Experiment

Students enrolled in ITSC were invited to take part in a three part experiment to determine if creating MCQs has an educational value. One activity undertaken as part of the experiment was an assessment task. However the students had the choice to opt in or out with regard to being part of the experiment. In this case the students' results were excluded from the sample group and in no way had an adverse affect on their course result.

Part 1: A paper based MCQ quiz (pre-test) with a mixed level of difficulty associated with each question was provided to participating students to individually complete at the start of their tutorial class in the second week of semester. The hypothesis made here was that they have not learned or only have limited knowledge of the material to be covered at this stage. Due to the theory based nature of ITCS the question and learning style employed was more knowledge based rather than problem solving. The pre-test required each participant to recall knowledge on a given topic and apply that knowledge to find the correct answer. The pre-test was collected and marked by a third party. For the purpose of feedback only a result was provided to the sample group at this stage.

Part 2: All enrolled students were required to complete the assessment task of individually creating 15 MCQs on a given topic relating to their studies. The students were given four weeks to complete the task. The students were given a brief description of the topics which were framed as short answer questions as an impetus to create the MCQs. They were required to research the topic using the prescribed textbook to create high quality MCQs. "As most successful student generated question instructional studies appeared to involve either direct instruction or explicit written instruction on question generation" (Wong, 1985; cited in Yu & Liu, 2004a), two methods of training were provided before having students construct MCQs on their own. First participants were provided

with a set of criteria for creating good quality MCQs (Burton et al.). Second a complete example consisting of a sample MCQ and its elements was provided as a guide to what was expected for the assessment.

One of the key challenges with using MCQs is the bottleneck associated with creation, formatting and managing of the test bank and conversion for use in an online learning system e.g. Blackboard (Blackboard, 2007). With this in mind, a software application called “Test Monkey” was used as a tool to enable students to enter their MCQs and manage the test bank. This is represented in **Figure 1**.

MCQs are well suited for use in a software solution because answers are provided for the student to choose from, thus the computer system is not required to attempt to interpret a student’s answer, it instead matches the student’s response to the correct answer, giving a final result. All students were provided access to and were required to use this software to enter their questions. The questions the students created were marked based on the criteria for creating “good” MCQs as displayed in **Figure 2**.

The students were required to access the software from a Web site. The intention was to enable uploading of the test bank into an online learning system for future use.

Part 3: A second MCQ quiz (post-test) comprising the same questions as the pre-test was provided to the sample group to complete 6 weeks later following completion of the assessment task. The reasoning behind using the same quiz MCQs was to later enable the direct comparison of where the participants’ knowledge increased the most. Again the paper based post-test was completed by the participants in their tutorial classes. The post-test was collected and marked by a third party. For the purpose of providing meaningful and quality feedback the participants were provided with: the correct answers to the MCQs and the relationship between the Test Monkey rating criteria, the MCQ creation guidelines and the assignment general requirements.

The results from the pre-test and the post test were compared in an attempt to show that allowing students to create assessment content is beneficial from a learning perspective. Since the students were not provided with the correct answers for the pre-test, the assumption was that learning would take place which can be measured because the sample group would have to engage with the learning activity on a deeper level to produce “good” MCQs.

A summary of the method of inquiry is provided below.

Experiment:

- Sample group—participants enrolled in ITCS.

Part 1: Pre-Test: Students completed a MCQ Quiz on the



**Figure 1.** Test monkey software application.

**Rate the question**

|  |  |
|--|--|
| Question grammar appropriate and free of spelling errors                     | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2   |
| Question clear and unambiguous   | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| Question is original, well thought out and of high quality                   | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| Correct answer is well thought out and of high quality                       | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| 3 * Distracters plausible and unobvious                                      | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| Answers grammar appropriate and free of spelling errors                      | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2   |
| Question and Answers are related   | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| Reference / Feedback accurate  | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3   |
| Reference / Feedback is informative and clear                                | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 |
| Overall type rating (recall = 1, critical thinking = 2, problem solving = 3) | <input checked="" type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3   |

[Submit Answer](#) [Return to Question List](#)

**Figure 2.** MCQ rating criteria in test monkey.

following topics:

- Organizational Security and Human Factors;
- Cryptography and Public Key Infrastructure;
- Standards and Protocols and Network Security;
- Infrastructure and Remote Access Security.

Part 2: Students researched, created and submitted a set of MCQs on the same topics:

- The MCQs were graded by the teacher based on defined criteria.

Part 3: Post-Test:

- The sample group completed a post-test using the same questions as the pre-test.

Analysis of results:

- Results from the pre and post tests were compared to determine whether student generated assessment does encourage learning.

Analysis of student learning:

- Determine if learning has occurred;
- Compare scores from pre and post tests to measure improvement;
- Analyze if questions of high quality from part 2 produce high results in post-test from part 3.

In this experiment any improvement in learning was measured by comparing the results of the pre-test and post-test MCQ quiz before and after the creation of the MCQs in the assessment task. As indicated the same ITCS topics were used in order to measure the difference between what was known before and following the creation of MCQs.

**Results and Discussion**

The focus of the analysis was on the improvement of a student’s marks after creating MCQs, to justify the educational benefits which MCQ creation provides. It was anticipated to see students improve their test scores after creating MCQs and that

there would be a correlation between the performance on the post-test and the mark awarded for the assessment task. This section delivers the findings and discussion of the results reflecting on what the findings mean in relation to the research question.

A series of tables and charts were used to display the results from the experiment. The results and basis for discussion have been taken from these displays.

A sample of 30 students accepted the invitation to partake in the experiment. The students first completed the pre-test and the results were collated. The pre-test determined the student's knowledge of the material before they commenced the assessment task of creating MCQs. As indicated by the percentages and grades from the students' pre-test results it was clear that the student's knowledge was not of a high standard. This, however was expected because the students had not been exposed to the topics covered in the pre-test. The pre-test was designed in this manner so it could be determined if creating MCQs would enhance the student's knowledge on the given topics.

The analysis conducted showed the average mark obtained on the pre-test was 23.27 (51.71%) out of a possible 45 marks. This means on average students only just passed the pre-test by a margin of .77 or 1.71%. The maximum mark achieved was 32 (71.11%) and the minimum was 14 (31.11%). The median score was recorded at 23.5 sitting one mark above the pass mark.

An analysis of the students grades was also displayed and it is interesting to observe that 10 students or 33.33% failed the pre-test (which is rather high when a normal bell shape curve is considered), a further 14 students achieved a pass grade and only 5 of the students were able to obtain credits. Finally only 1 distinction was awarded and no high distinctions were allocated.

A range of statistics in relation to the pre-test results, including the mean, mode and standard deviation were all recorded around the pass mark of 22.5, with the results being 23.3, 23.9 and 4.59 respectively. The negative skew of the data confirmed that the students understanding of the topics tested in the pre-test was low to non-existent. From these findings it was expected with some confidence that any knowledge obtained to improve the students' results in the post-test would be obtained in the assessment task.

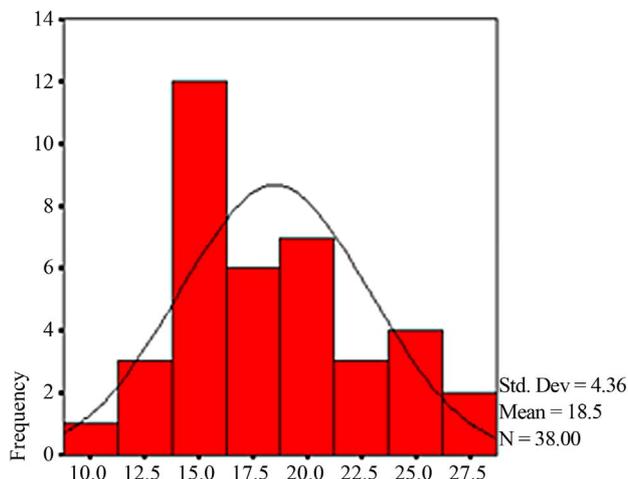
The histogram presented in **Figure 3** shows the frequency distribution of the results obtained in the pre-test.

In summary the results of the pre-test were low, as expected. The data collected had a normal distribution so statistical significant testing was determined to be appropriate.

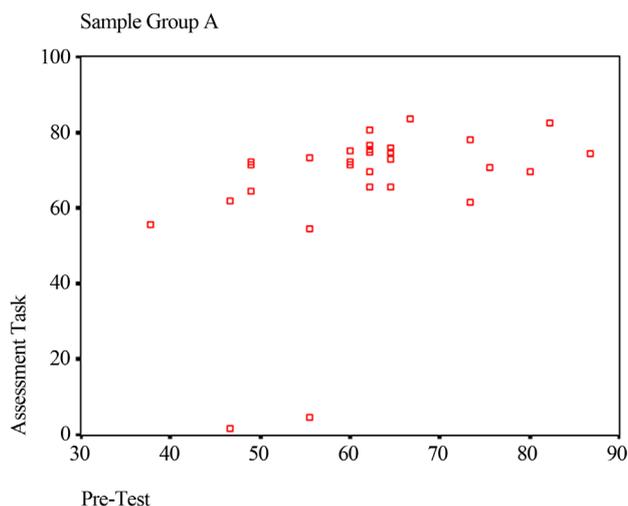
The next step in the experiment was to conduct the assessment task which represented the learning process. Students were required to create MCQs and those questions were rated by the teacher to determine if they were of a high quality. The criteria against which the MCQs were rated are displayed in **Figure 2**.

To show the level of correlation between a student achieving highly in the assessment task and highly on the post-test, the raw scores the students achieved on the assessment task as well as the percentages were calculated. As displayed in **Figure 4** the level of correlation is high with two outliers.

The final step in the experiment was having the students take the post-test to determine if by creating MCQs they were able to increase their level of knowledge on the tested topics. An increase in post-test score would indicate that the assessment



**Figure 3.** Histogram of pre-test results.



**Figure 4.** Correlation of assessment task vs. pre-test.

task did have a positive affect on the students learning and would contribute to validating the research question.

The results showed that the post-test score increased from the pre-test after the students undertook the process of creating MCQs. The highest improvement was a 33.33% increase; with this test score increasing 15 marks from a 21 (Fail) to 35 (High Distinction) marks. This was the highest improvement. The average improvement was an increase of 4.57 marks representing 4 more questions being answered correctly on the post-test than the pre-test. In general all students were able to achieve a higher post-test score. In fact the fail rate in the post-test fell by 13.33% down to a more reasonable 20% when a normal bell shaped curve is considered.

Further comparison between the results of the pre and post tests indicated that the average post-test score increased by 4.56 to 27.83 out of 45, moving the average result to a low credit, an increase of a whole grade. The maximum mark on the pre-test achieved after the completion of the assessment task was 39, just 6 marks off 100%, a difference of 7 marks from the pre-test. The minimum mark was raised to 17, a difference of 3; the median showed an increase of 4.5 moving to 28 and four less

people failed the post-test.

One of the more significant findings was the change in grades. Three participants achieved high distinctions and there was a major increase in students moving from a pass to a credit grade. In relation to the research question this shows that it was possible to significantly raise the level of knowledge that students were able to recall after they undertook the creation of MCQs.

The statistics calculated for the post-test, including the mean, mode and standard deviation all increased to, 27.83, 28 (18.4% of the sample achieving this score) and 28 respectively. The data shifted from a negative to a positive skew and overall the results of the post-test were higher than what was previously achieved on the pre-test.

The histogram displayed in **Figure 5** clearly indicates an increase in the post-test scores in comparison to the pre-test scores. The highest achieved results are now located around the 27.5 mark, compared with previously identified 25 mark. As displayed the frequency has increased for higher marks and the graph is plotting higher test scores, starting from 17.5 compared with 15 from the pre-test.

The t-test performed was a Paired Sample t-test and was the most important part of the results in terms of validating the research question. The paired samples t-test compares the means of the two variables and computes the difference for each case, testing to see if the average difference is significantly different from zero. This test determined whether the assessment task had a statistically significant affect on students test results.

First this test needs to have paired data which means that the pre-test and post-test scores were taken from the same individual and second the data must be normally distributed which was previously established.

The null hypothesis tested was that there is no significant difference between the means of the pre and post test; the alternate hypothesis which was hoped to be achieved was that there is a significant difference between the means of the pre and post test, meaning that the students did gain knowledge from the creation of MCQs.

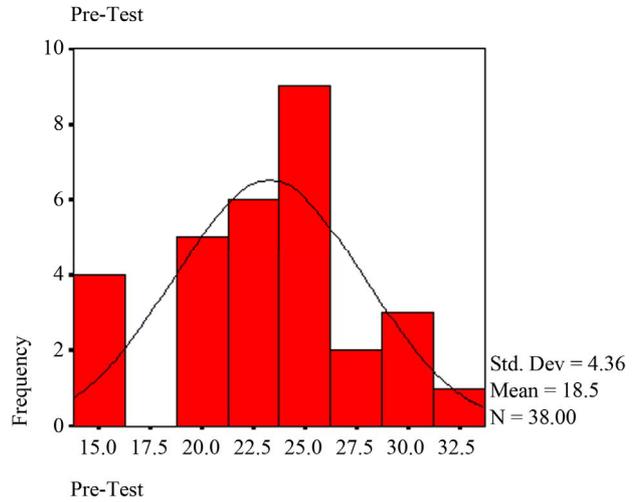
The Paired Sample Statistics showed the statistics for both sample groups and were used to calculate the significance. The Paired Sample Correlations showed that the data sets had a positive correlation of .692 meaning that students who did well on the pre-test generally did well on the post-test, confirming that the data is consistent and that statistically students did not perform worse on the post-test.

The final test was the Paired Samples Test and this indicated if the differences in the results were statistically significant. The analysis indicated an increase in the mean of 4.5667 and a large improvement in the lower and upper bounds of the 95% confidence interval, recording  $-5.9730$  and  $-3.1603$  respectively. The T value was  $-6.641$  and the degrees of freedom equaled 29.

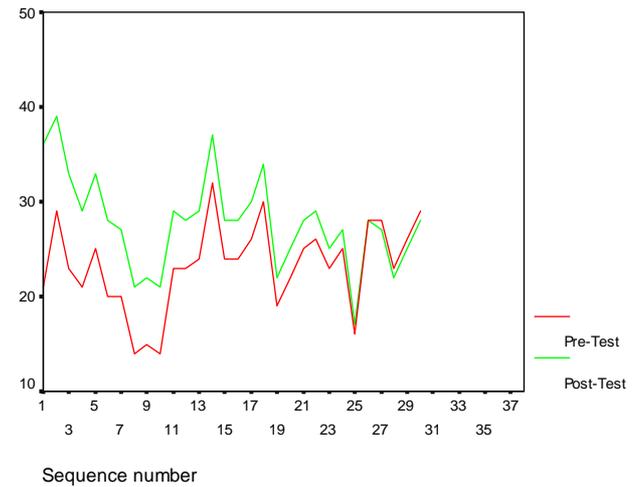
The final figure that this test presents is the significance of the 2-Tailed-t-test. The significance value for the paired sample (pre & post test) record was .000 which means that the null hypothesis can be rejected. It can be concluded that the difference between the pre and post test was a significant difference, validating the educational value of students undertaking the process of creating MCQs.

**Figure 6** displays a Sequence Diagram mapping the results of the pre and post test against each other. From this it is clear that the majority of post-test results were higher than the pre-test.

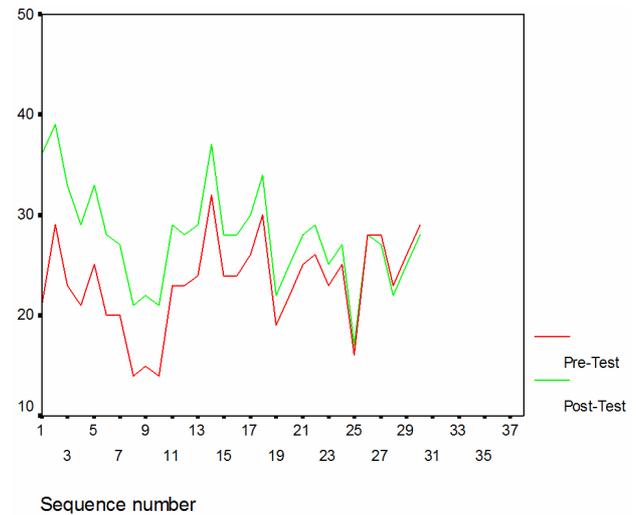
**Figure 7** displays a plot of the assessment task results against



**Figure 5.**  
Histogram of post-test.



**Figure 6.**  
Sequence diagram pre-test and post-test.



**Figure 7.**  
Assessment task vs. post-test results.

the post test results. It is clear to see that the majority of assessment results were higher than the post-test results. This provides further justification that the creation of MCQs resulted in effective learning.

All the analysis presented in this section demonstrated that there were statistical significant differences in the pre and post test results and that having students create MCQs in ITCS as part of their assessment does have a valid learning benefit.

### Conclusion

This inquiry set out to implement a new learning strategy to facilitate a more active and inquisitive learning atmosphere than what was currently being experienced. It was contemplated by incorporating an assessment requiring students to construct MCQs which could have a positive impact on students' approach to learning. By quantitatively analyzing the results from the pre and post tests and comparing the results of the post test to the assessment task, the desired outcome of improving student knowledge and understanding in relation to the concepts and practices of computer security was realized. The significance of the method was proven in raising the level of knowledge gained by these students.

Using "Test Monkey" enabled a unique style of feedback allowing students to review the expert rating for each of their MCQs as marked by the teacher. This process provided each student with a customized view of their results, and importantly not just a mark. It is our belief that this provided opportunity for self-reflection to enhance their learning experience and motivated students to study and keep up to date by encouraging a more interactive approach to learning.

The results suggested that students may have been involved in a deeper approach to learning and the development of cognitive skills through a more intimate interaction with the unit material. However the degree of deeper learning that this inquiry achieved is not conclusive at this stage. This is understandable as it is "not a realistic expectation for students to produce MCQs testing higher order cognitive skills at their first attempt". Also "exercises of this nature are not likely to be greeted with much enthusiasm as they involve learning methods unfamiliar to many students" (Palmer & Devitt, 2006). Further analysis of the data is required, for instance to determine and rank the overall quality of the MCQs created.

From a teaching perspective the experience was certainly valuable on several levels. It provided the opportunity to try something new. It exposed what students did not know and enabled informed adjustments to the teaching to target those areas earlier on in the semester. It has provided a resource for future use and most importantly feedback about ITCS for reflection and improvement.

The major dilemma experienced was the marking of approximately 2000 MCQs all at one time. This created a bottleneck in the assessment process. To overcome this dilemma the assessment model should be adjusted so that students would create MCQs on a weekly basis; thus the assessment would be spread over the duration of the semester. This would be expected to encourage a more progressive and active approach to learning allowing students to receive feedback on a regular basis to help them improve their learning strategy. This again would provide direction for teaching; more towards those specific areas on what our students don't know or understand.

In terms of future work it is planned to perform additional

inquiries across multiple disciplines to assess and validate the effectiveness of this assessment approach on a wider scale. Also it is intended to introduce another feature into the software to allow a peer review process of the MCQs. In this way students would be able to learn from each other which is thought to add additional value to the learning experience.

### REFERENCES

- ACVIM (1997). *Question-writing guidelines*. American College of Veterinary Internal Medicine.  
<http://www.acvim.org/uploadedFiles/Candidates/exam/ABIMQuestionWritingGuidelines.pdf>
- Blackboard (2007). *Blackboard homepage, Blackboard Inc.*  
<http://www.blackboard.com/us/index.Bb>
- Brown, S., Rust, C., & Gibbs, G. (1994) *Strategies for diversifying assessment in higher education*. Oxford: Oxford Centre for Staff Development.
- Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1991). *How to prepare better multiple-choice test items: Guidelines for university faculty*. Brigham Young University Testing Services and the Department of Instructional Science.  
<http://testing.byu.edu/info/handbooks/betteritems.pdf>
- Censo (2007). *Guidelines for writing effective tests: A practical "short course" for test authors*. Censo Corporation.  
<http://www.censocorp.com/downloads/whitepapers/guidelines-for-writing-effective-tests.asp>
- Elton, L., & Johnston, B. (2002) "*Setting the scene*", *assessment in universities: A critical review of research*. London: Generic Centre Learning and Teaching Support Network (LTSN).  
<http://eprints.soton.ac.uk/59244/1/59244.pdf>
- English, L. D. (1998). Children's problem posing within formal and in formal context. *Journal for Research in Mathematics Education*, 29, 83-106. doi:10.2307/749719
- James, R., McInnes, C., & Devlin, M. (2002). *Assessing learning in Australian universities*. Canberra: Australian Universities Teaching Committee.
- King, A. (1995). Inquiring minds really do want to know: Using questioning to teach critical thinking. *Teaching of Psychology*, 22, 13-17. doi:10.1207/s15328023top2201\_5
- Leung, S. S., & Wu, R. X. (1999). *Problem posing with middle grades mathematics: Two real classroom examples. Mathematics teaching in the middle school*. Reston, VA: National Council of Teachers of Mathematics.
- Yu, F., & Liu, Y. (2004a). Perceived potential value of student multiple-choice question-construction in the introductory physics laboratory. *Proceedings of International Conference on Engineering Education ICEE-2004*, Gainesville.
- Yu, F., & Liu, Y. (2004b). Active learning through student generated question in physics experimentation classrooms. *Proceedings of International Conference on Engineering Education ICEE-2004*, Gainesville.
- Lublin, J. (2000). *Guidelines for good practice in assessment*. University of Western Sydney, Campbelltown: Centre for Enhancement of Learning and Teaching.
- Miller, A. H. (1987). *Course design for university lecturers*. New York: Nichols Publishing Company.
- Palmer, E., & Devitt, P. (2006). *Constructing multiple choice questions as a method for learning*. Singapore: Annals Academy of Medicine.  
<http://www.annals.edu.sg/PDF/35VolNo9Sep2006/V35N9p604.pdf>
- Parker, L. (2004). Intersecting sets: Carrick, assessment and evaluation. *Key Note Address Presented at the Evaluations and Assessment Conference*. Melbourne: RMIT.
- Parry, S. (2004). Student assessment. *Key Note Address Presented at the Evaluations and Assessment Conference*. Melbourne: RMIT.
- Phye, G. D. (1997). *Handbook of academic learning—Construction of knowledge*. California: Academic Press.
- Ramsden, P. (1988). *Improving learning—New perspectives*. London: Kogan Page.

- Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). London & New York: Routledge Falmer.
- Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics*, 14, 19-28.
- Singer, H. (1978). Active comprehension: From answering to asking questions. *Reading Teacher*, 31, 901-908.
- Sircar, S. S., & Tandon, O. P. (1999). Involving students in question writing: A unique feedback with fringe benefits. *Advances in Physiology Education*, 22, 84-92.
- Toohey, S. (1999). *Designing courses for higher education*. Buckingham, UK: SRHE and Open University Press.
- Wong, B. Y. L. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, 55, 227-268.  
[doi:10.3102/00346543055002227](https://doi.org/10.3102/00346543055002227)