

University of Tasmania Open Access Repository

Cover sheet

Title

A COMPARATIVE ANALYSIS OF ONLINE AND PAPER-BASED ASSESSMENT METHODS: A UNIVERSITY CASE STUDY

Author

Campton, P

Bibliographic citation

Campton, P (2004). A COMPARATIVE ANALYSIS OF ONLINE AND PAPER-BASED ASSESSMENT METHODS: A UNIVERSITY CASE STUDY. University Of Tasmania. Conference contribution.
https://figshare.utas.edu.au/articles/conference_contribution/A_COMPARATIVE_ANALYSIS_OF_ONLINE_AND_BASED_ASSESSMENT_METHODS_A_UNIVERSITY_CASE_STUDY/23212691

Is published in:

Copyright information

This version of work is made accessible in the repository with the permission of the copyright holder/s under the following,

Licence.

If you believe that this work infringes copyright, please email details to: oa.repository@utas.edu.au

Downloaded from [University of Tasmania Open Access Repository](#)

Please do not remove this coversheet as it contains citation and copyright information.

University of Tasmania Open Access Repository

Library and Cultural Collections

University of Tasmania

Private Bag 3

Hobart, TAS 7005 Australia

E oa.repository@utas.edu.au

CRICOS Provider Code 00586B | ABN 30 764 374 782

utas.edu.au

A COMPARATIVE ANALYSIS OF ONLINE AND PAPER-BASED ASSESSMENT METHODS: A UNIVERSITY CASE STUDY.

Paul Campton
School of Information Systems
University of Tasmania, AUSTRALIA
Paul.Campton@utas.edu.au

Abstract

There has been a continuing trend for educational institutions to focus on e-learning, this has been accompanied by a shift towards online assessment, either of a formative or summative nature. This paper is a critical analysis of online assessment methods introduced into a first year unit in the Faculty of Commerce at the University of Tasmania. It reflects on the effectiveness on the change from a paper-based assessment method to an online system over a period of two years as well as on the advantages and disadvantages of the change for students, lecturers and tutors. The performance of students in undertaking the unit in Semester 2, 2002 and Semester 2, 2003 was measured using a common assessment test; in 2002 using a paper-based system and then in 2003 for an online system. Analysis was then performed on a cohort within a semester. The results show there is no significant difference in performance between the groups, however there existed a number of administrative benefits associated with the online assessment method.

Keywords

E-learning, assessment, computer assisted assessment (CAA), multiple choice test, summative learning.

Introduction

The importance of learning to a society should require educators involved in e-learning to treat learning with a similar importance to health and adhere to Hippocrates' direction "as to diseases, make a habit of two things - to help, or at least, do no harm" (Jones, 1923, p.165). Learning and assessment have been drastically changed by the widespread adoption of technology. This is particularly evident in the USA (McKeena and Bull, 2000). There exist well-documented technical reasons for moving from traditional paper-based assessment to online assessment; including scalability and modularisation (Huffaker, 2003). This paper will seek to build on some of the earlier research undertaken (Lloyd, Martin and McCaffery, 1996) by considering recent technological, economic and pedagogical developments. Lloyd, Martin and McCaffery (1996) found that computer-based assessment resulted in timesavings for staff and that there existed general acceptance amongst students. It is with Hippocrates' direction in mind that this research to measure the effectiveness of moving from a paper-based system to an online system of assessment was undertaken. This paper will examine the issues involved in online assessment and will specifically address the question whether there is any difference in performance between students who were assessed online compared to those who were assessed using paper-based methods. The specific form of assessment was a multiple-choice test designed for a first year Information Systems unit, BSA101, in the Commerce Faculty at the University of Tasmania. This paper will not seek to research the value of formative assessment and its role in learning and assessment (Zakrzewski and Bull, 1998; Clariana, 1993)

Background

E-learning is changing the nature of educational practices and to some degree has been driven by economic and technological factors (Zhang and Nunamaker, 2003). E-learning is fundamental to an information based economy and assessment of a students' e-learning using technology is fundamental to the e-learning process. There has been an increase in the utilisation of technology in the assessment of large student groups in international and Australian universities (McKeena and Bull, 2000). There exist a number of factors associated with the successful integration of technology into the learning process. These include strong support (Zakrzewski and Bull, 1998) and sound educational practice (McKenna and Bull, 2000). This change has been made possible in part by the introduction of Learning (Content) Management Systems (LMS) such as WebCT®. This particular LMS was introduced to the University of Tasmania and has been centrally supported since 2001. It includes assessment tools such as multiple-choice quizzes, which can be used in either a formative or summative manner (Peat and Franklin, 2003).

Prior to 2003 and in Semester 1, 2003 Information Systems students in the Commerce Faculty at the University of Tasmania were assessed using a paper-based multiple-choice test. Given the size of the group multiple sessions were required, this required the lecturer to be assisted by a number of tutors. Questions on the paper were re-ordered to create multiple versions to reduce the likelihood of students successful copying from one another during the test. Identification was required to ensure only students enrolled in the unit sat the test. At the completion of the test the completed papers were organised into alphabetical order for marking and data entry. A number of marking sheets had to be created to handle the different versions of the test. Results were then entered into a spreadsheet. Sampling and simple analysis was undertaken to measure the level to which students answered questions correctly.

There were a number of problems surrounding this paper-based system. These included the creation and copying of multiple versions of the test, the cost of supervision and marking, the collection and securing of papers and answer sheets, accuracy of marking and data entry, and the capacity to easily analyse results to check for questions which may require modification. All students had access to a much larger database of multiple-choice questions, which they could utilize on a weekly basis for formative assessment and assistance. Such feedback is an important part of the learning process and one that students typically employ (Peat and Franklin, 2003).

Methodology and Data Collection

In Semester 2, 2003 the students enrolled in BSA101, Business Information Systems, were invited to volunteer to sit the multiple-choice test online rather than by the traditional paper-based method. The specific learning management system platform used by the Faculty at that time was WebCT[®] v3.8, this had a number of online assessment tools available to designer and instructor. The multiple choice assessment test, which was initially in MS Word[®] format needed to be converted into a text file with a specific styling. This could then be imported into WebCT[®] by means of a package called Respondus[®]; this was a relatively straightforward process and took only 15 minutes. Later versions of WebCT Vista[®] should allow this process to be undertaken directly.

It is important to establish the reliability and validity of the instrument being employed (Neuman, 2003). The paper-based test was piloted on a small group (n=6) who had either lectured or tutored the unit prior to being released, thus addressing the issue of face validity. There had been previous testing of the paper-based method on a group of students (n=472) in semester 1, 2003 with the same instrument. Given there existed no significant difference between the paper-based results; 11.25 in Semester 1 and 11.87 in Semester 2, it would support the notion that the paper-based test was a valid instrument. The assessment method and multiple choice test questions remained consistent over the two semesters for the paper-based method. The online test was piloted on smaller group (n=2) as the content was identical and WebCT[®] was a robust and reliable platform. Whilst randomized groups would have been the preferred option for achieving a result that could be generalised, it was not possible due to the ethical issues relating to voluntary participation. There existed certain practical limitations in comparing the online and paper-based assessments and these will be discussed in the limitations.

An empirical approach was used to gather data and analyse the data, using the management tools available in WebCT[®]. The scope of the research was limited to BSA101. From the student group 48 students (Group A) chose to do the test online, whilst 70 students (Group B) sat the test using the paper-based system. Security and contingency plans were put in place; these included having test papers as a backup, photo identification and individual release of the online test by the tutors. Students who selected to undertake the online test were required to log into a special university user account, this was specifically created for assessment purposes. This account severely restricted the types of software students could use and the students' ability to communicate with others. Students were then required to log into their individual WebCT[®] accounts to access the test. The test, or quiz as it is known in WebCT[®], allowed students to start the test when they felt ready rather than when they were instructed to do so by the tutor or lecturer. The test had a time period of 50 minutes but most on-line students completed it well within this time and were free to leave the computer laboratory where the test was conducted. The computers were configured in a multi-island layout within the laboratory rather than long benches facing in the same direction. Students were able to change their answers repeatedly until the 50 minutes time limit expired and submission was required. Students using the online test were free not only to revisit questions but free to leave the room when they were satisfied that they had completed the test. A number of online limitations existed, specifically the ability to receive feedback on questions during and after the test. Whilst feedback is critical, particularly in formative assessment, there existed the imperative of ensuring students did not communicate answers to those students who were undertaking or intending to undertake the test.

In order to address the research question of whether there is any difference in performance between students who were assessed online compared to those who were assessed using paper-based methods the following hypothesis will be tested. Hypothesis H_{0c} Proposes there is no difference between the groups in terms of their multiple-choice test marks and the method of assessment.

Results

Assessment results between the two groups for Semester 2, 2003

Group A = Assessed by means of an online test

Group B = Assessed by means of a paper-based test

All other assessment tasks were completed using identical methods.

Table 1

| Group | Mean of Multiple Choice Test marks (/20) | Variance of Multiple Choice Test marks | Mean of other online assessment task mark (/20) | Mean of other assessment tasks marks (/60) | Total marks (/100) |
|-------|---|--|---|--|--------------------|
| A | 12.11 | 6.88 | 14.51 | 44.25 | 70.87 |
| B | 11.87 | 5.64 | 13.57 | 41.72 | 67.16 |

The assumptions on which this t -test is performed are:

- students appear in only one group;
- the groups are unrelated and
- the groups should have homogeneity of variance. (Foster, 2002)

Now $H_0: \mu_1 - \mu_2 = 0$, $\alpha = 0.05$

We assume a normally distributed population and that the samples are independent.

Thus with $\alpha = 0.05$ ($v = n_1 + n_2 - 2 = 70 + 48 - 2 = 116$) and the t -table value of 1.98, the two-tailed (non-directional) decision rule is: Reject H_0 if $t^* > (t_{0.05/2, 116} = 1.98)$ or $t^* < (-t_{0.05/2, 116} = -1.98)$.

The calculated value for t^* is 0.64 hence we cannot reject H_0 .

Hence the null hypothesis was accepted. There is no significant difference at the 0.05 level between the test results for those who sat the test online compared to those who sat using the paper-based method.

Paper-based Multiple Choice Test Results over 2 semesters

Table 2

| Semester | Multiple Choice Test mark (/20) |
|------------------|---------------------------------|
| Semester 1, 2003 | 11.25 |
| Semester 2, 2003 | 11.87 |

Discussion

Essentially, there was no significant difference between the test marks of students who were assessed online using the multiple-choice test and those who were assessed using the paper-based system. It is clear from the results, students who completed the test online were not disadvantaged in any significant manner from those who completed the test using a paper-based system (Neuman & Baydoun, 1998)

Those students who sat the test online were allowed to start when they were ready. The ability for individual students to begin, when they were ready, could have reduced the anxiety associated with a mass testing of students. This was a more individual approach and gave the students greater control than the mass paper-based testing where students started and finished according to the schedule set by the examiner. Such a student directed approach is more educationally sound and is consistent with pedagogical developments. The student's focus was on the individual test and the questions posed on the screen rather than a mass of students facing in the same direction. Students who sat the paper-based test had less freedom to leave because of the physical layout of the testing room.

The issue of whether the students who volunteered had greater computer literacy or if there existed any other contributing factors needs to be considered. Whilst the self-selection method for the groups was not as statistically sound as random selection, the former selection method was undertaken for the benefit of the

students rather than for research reasons. It would appear from a brief analysis of the performance of students in the other assessments tasks, (Table1) that very little difference existed between the groups in terms of their general performance, skills and understanding.

It must be noted that students who did not complete all assessment tasks were excluded from the analysis. It was of interest that those who completed the test online had a higher success rate in completing assessment and overall had a slightly higher total mark.

The administrative benefits were considerable. These included a reduction in marking time and associated costs, along with greater accuracy. The capacity to easily analyse results and to check for questions, which may require modification, was available. The tutor had the ability to control access to the test by selectively releasing it to the appropriate students.

Limitations

The selection of the groups remains problematic. Randomisation of the groups, whilst preferred, was not possible to the voluntary participation. The selection method may have hidden some anxiety in the paper-based group. Students studying BSA101 require significant levels of computer literacy and had already completed an online assessment task (Table 1) before undertaking the multiple-choice test. Students are required to access lecture and tutorial material online for what is essentially a computer-based unit. The level of anxiety in the paper-based group, which may have been masked by not having to undertake the test online, needs to be mitigated by the nature of the unit and the students' previous online assessment experience (Butler, 2003). Any difference may be due as much to individual variation (Clariana & Wallace, 2002) as by anxiety of the paper-based group. However, it is accepted as a limitation and one that can be overcome with further research. This would then give an ability to generalise the findings.

Conclusion

Online assessment does not disadvantage students. There are clear economic advantages associated with online assessment, these include with savings in time and costs. The ability to analyze results quickly (Zakrzewski and Bull, 1998) means that educators can better use valuable resources when improving and updating multiple-choice tests. The ability to perform in-depth analysis and quality assurance of questions is a distinct advantage for educators and a necessary function in the design of online assessment. Multiple-choice tests have a place in assessment but they need to be undertaken with the same amount of care as any other assessment method. The use of formative online tests should also reduce some of the anxiety associated with using technology for the first time for assessment. A mock on-line test is also an advantage in such circumstances. Pedagogical issues need to be considered first if we intend to reap some of the technological and economic benefits of online assessment. Sound online multiple-choice tests require an extensive investment, particularly in the initial phase, as well as some ongoing maintenance and revision. There must exist a clear and common understanding, by educators and students, of the knowledge and skills being tested if multiple-choice tests are to be effective. The assessment of higher order skills, such as synthesis and analysis, in multiple-choice tests has been considered problematic (Peat, 2000). We now have a situation emerging where the technology can be used to create a dynamic, media rich and interactive assessment environment rather than one that simply reflects a static paper-based environment. (McKenna and Bull, 2000; Daziel and Gazzard, 1999) This may be an environment in which higher order knowledge and skills can be assessed. The technological capacity to increase transmission speeds for multi-media materials should dramatically improve e-learning and online assessment (Zhang and Nunamaker, 2003).

In conclusion, multiple choice tests or quizzes have a place as an assessment tool provided educators understand the manner in which they can be successfully employed as well as considering the pedagogical issues such as learning styles and cognitive levels (Peat, 2000). There is little evidence to suggest students are disadvantaged by the use of online assessment involving multiple-choice tests, except when sound educational practice, management and support are missing. We need to have a clear idea of why and how such tests are undertaken. Indeed we may be at the emergence of a new era of enriched online assessment provided we adhere to the some basic principles of assessment.

REFERENCES

- Butler, D.L. (2003). The Impact of Computer-Based Testing on Student Attitudes and Behavior. *The Technology Source*. Available online at <http://ts.mivu.org/default.asp?show=article&id=1013>
- Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer based instruction. *Journal of Computer Based Instruction*, 20(3), 67-74.
- Clariana, R. & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology* 33 (5), 593-602.
- Dalziel, J., & Gazzard, S. (1999). Next generation computer-assisted assessment software: the design and implementation of WebMCQ. *Proceedings of the 3rd Annual CAA Conference*, Loughborough, 61-71.
- Foster J. (2002) *Data Analysis: Using SPSS for Windows*. Great Britain: Sage.
- Huffaker, D. (2003). Reconnecting the classroom: E-learning pedagogy in US public high schools. *Australian Journal of Educational Technology*, 19(3), 356-370.
- Jones, W. H. S. (1923). *Hippocrates*. Cambridge, Ma.: Harvard University Press.
- Lloyd, D., Martin, J. G. & McCaffery, K. (1996). The introduction of computer-based testing on an engineering technology course. *Assessment and Evaluation in Higher Education*, 21(1), 83-91.
- McKenna, C. & Bull, J. (2000). Quality assurance of computer-based assessment: practical and strategic issues. *Quality Assurance in Education*, 8(1), 24.
- Neuman, G., & Baydoun, R. (1998). Computerization of pencil and paper tests: When are they equivalent? *Applied Psychological Measurement*, 22(1), 71-83.
- Neuman, L. (2003). *Social Research Methods*, United States: Allyn & Bacon.
- Peat, M. (2000). Online assessment: The use of web based self assessment materials to support self directed learning. *Proceedings of the 9th Annual Teaching Learning Forum*, Perth: Curtin University of Technology. <http://cea.curtin.edu.au/tlf/tlf2000/peat.html>
- Peat, M. & Franklin, S. (2003) Has student learning been improved by the use of online and offline formative assessment opportunities? *Australian Journal of Educational Technology*, 19(1), 87-99.
- Zakrzewski, S., & Joanna Bull, J. (1998). The mass implementation and evaluation of computer-based assessments. *Assessment and Evaluation in Higher Education*, 23(2), 141 –152.
- Zhang, D. & Nunamaker, J. (2003). Powering e-learning in the new millennium: an overview of e-learning and enabling technologies. *Information Systems Frontiers*, 5(2), 207-218.

Copyright © 2004 Paul Campton

The author assigns to ACEC and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The author also grant a non-exclusive licence to ACEC to publish this document in full on the World Wide Web (prime sites and mirrors) and in printed form within the ACEC 2004 conference proceedings. Any other usage is prohibited without the express permission of the author.