

INTERFACES FOR ONLINE ASSESSMENT: FRIEND OR FOE?

Gavin Sim

Department of Computing
University of Central Lancashire
Preston
grsim@uclan.ac.uk

Matthew Horton

Department of Computing
University of Central Lancashire
Preston
MPLHorton@uclan.ac.uk

Stephanie Strong

CELT
Lancaster University
Lancaster
s.strong@lancs.ac.uk

ABSTRACT

This paper examines the interface for online assessment within WebCT. Students in four undergraduate and one postgraduate module participated in the study. Post test questionnaires were distributed to students to ascertain their experience of online assessment focusing on the usability of the software. The quantitative data indicated that the users had little difficulty in achieving their tasks. However, the qualitative data revealed a number of issues with the interface regarding navigation and feedback messages. The conclusions make recommendations on how the interface could be improved and further research will lead to the production of a standard interface for test delivery within Questionmark.

Keywords

WebCT, Interface, Online Assessment, Usability, Navigation

1. INTRODUCTION

There is an increased use of Computer Assisted Assessment (CAA) within higher education, often incorporated into learning management systems (LMS). The perceived benefits to the lecturer include automated marking and question analysis. The benefits for the students usually focus on the opportunity for instant feedback, however one concern is the impact that the user interface could have on their performance. There are numerous variables that could have an influence on students' performance when tests are delivered via a

computer, including monitor resolution, the way in which the text is displayed on the screen and whether the questions require scrolling. These concerns, along with many others, have been acknowledged in the British Standard 7988 for the use of information technology in the delivery of assessment[2]. The guidelines focus on navigational and usability issues, however some of the recommendations are valueless, with statements such as the navigation should be 'simple and clear'.

The goal of the user is to complete the assessment this can be broken down into several tasks. The users' tasks when participating in an online assessment are to start the test, answer the questions, navigate between pages and end the test. When completing a coursework there is usually implied skills not associated with the subject that are necessary, for example Information Technology when constructing an essay. Therefore if the online assessment required a high level of I.T. skills it would be inappropriate if you were testing the students' ability to understand Shakespeare, this would risk construct validity. Ensuring construct validity means ensuring that the assessment content is closely related to the overall learning objectives [6]. The interface has an extremely important role as any usability problems may constitute a threat to construct validity [5].

Within a test environment where students are often anxious and placed under stressful conditions limited mental resources need to be allocated to understanding the interface. Learnability is an issue as mental resources should not be wasted on understanding the software, it needs to be intuitive. Therefore the interface should not distract the user from achieving their objective. The software needs to be robust and recoverability is a concern as students should be able to undo entries and recover data if the computer crashes.

There are different interface options within a CAA environment that are often predetermined by the software manufacturer in the form of templates. The majority of lecturers will not be experienced in evaluating the usability of an interface and will therefore not question the suitability of these default templates

offered. This paper explores the HCI issues surrounding online assessment within WebCT to inform the development of a standard interface within QuestionMark.

2. METHODOLOGY

This investigation examined the experiences of students using the assessment tool within WebCT. This was an exploratory study to determine whether there are any issues regarding the software and no initial hypotheses were drawn. The test questions were all presented using single question delivery eliminating the need to scroll as this is a factor that can have an influence on test performance [9]. The participants were drawn from four undergraduate and one postgraduate module.

Group	A	B	C	D	E
Cohort	101	119	23	75	82
Number of Responses (N)	67	31	20	55	48

Figure 1. Sample Groups

A first and second year module were selected from the Department of Technology (Group D and E) and the Department of Computing (Group A and B). The postgraduate course was from the Department of Computing (Group C) (Figure 1).

Post-test questionnaires containing a mixture of Likert (Strongly Disagree=0, Disagree=1, Neutral=2, Agree=3 and Strongly Agree=4) and open-ended questions were distributed to the students to ascertain their experience of CAA, focusing on usability issues.

3. RESULTS

The first task of the user is to access the test. Users are required to enter a user name and password and this is essential for authenticating the user and recording the results [1]. The students were asked *Did you have any difficulty accessing the test?* The results are displayed in figure 2.

Group	A	B	C	D	E
No	69.7%	71%	65%	100%	97.8%
Yes	30.3%	29%	35%	0%	2.2%

Figure 2. Did you have any difficulty accessing the test.

Overall the majority of students had little trouble accessing the test however there are two groups who had less difficulty.

Another of the users' tasks is to navigate between the pages. Regarding the navigation the students were asked

The navigation was clear, the mean scores are displayed in Figure 3.

Group	A	B	C	D	E
Mean	3.02	3.26	3.00	3.29	3.07

Figure 3. The Navigation was clear

Again the overall results suggest that the users had little difficulty in navigating the test.

To aid the students achieve their overall tasks the software should be intuitive. Students may have different levels of prior experience with using the software and learnability could be an issue [4]. Students were asked whether they agreed with the statement *The software was easy to use* and the mean scores are displayed in Figure 4.

Group	A	B	C	D	E
Mean	3.18	3.35	3.00	3.31	3.04

Figure 4. The software was easy to use

The overall results suggest that there was no difficulty in using the software, however students prior exposure to the software was not analysed.

4. DISCUSSION

4.1 Logging on

Following consultation with staff it was discovered that the three modules with a high proportion of users having difficulty accessing the test had no prior experience of using the assessment tool within WebCT. However, the other two modules had access to a practise test and this could provide some explanation for the difference between the groups. It is recommended that students have access to practise tests to understand how the process works [3] and alleviate anxiety caused by the introduction of new assessment methods [10]. The problem does not appear to be critical as it is not persistent but having exposure to the test environment appears to greatly reduce the overall problem. An alternative method which could help alleviate this problem for novice users is to extract their user name and password from their network login. This would help reduce problems associated with forgotten passwords however, this may reduce the overall security if students computers are left unattended and passwords are shared. Even under the present system security can be breached as one student commented "I have just completed the quiz on WebCT. However it has given me two scores, and states I have done the quiz twice at different times. Could it be possible that someone else has logged on to WebCT in my name? If so what should I do?"

4.2 During the test

After the user has successfully logged in, the test is generated in a small pop up window and the entire interface cannot be seen (see Figure 5). This could result in another problem if the user resizes the window they could accidentally exit the test by pressing the close button within the browser. Other applications and browsers such as Opera 7 prompt the user to confirm that they want to exit the application. However, even with Opera 7 this feature is not available when using a pop up window. Many users commented on browser incompatibility and this is supported in other studies [7]. It would therefore be recommended that the interface is displayed in a browser such as Opera 7 and the confirm exit feature is enabled.

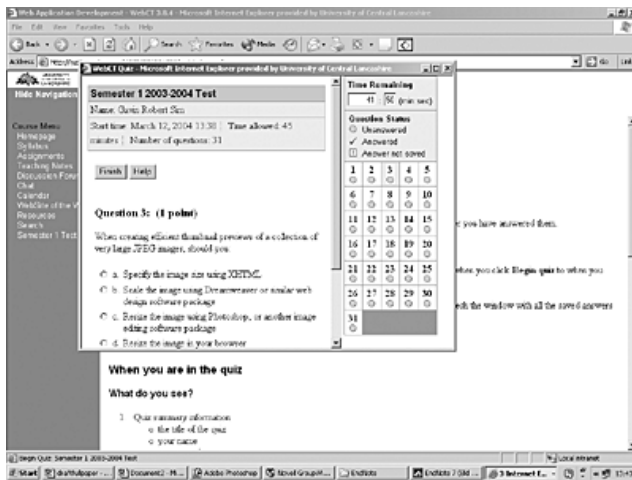


Figure 5. First screen presented to the user after logging on.

When navigating the test environment a number of users stipulated that if you did not save you would lose your answer when you left the screen. Upon further investigation it was found that if you answer a question and proceed to the next question without saving you receive the following message *warning: the current question has not been saved since the last edit. Proceed to the new question?* It does not mention that if you proceed your previous answer would be erased. Relating to this issue several users did not like having to save their answer after every question. There is no obvious reason why this process could not be automated when the user navigates between questions.

One user commented on the fact that when you were on the last question the next question button remained on the screen. When this was pressed it resulted in the user going back to the first question but the user presumed it would finish the test and thus caused confusion. It is recommended that the button is removed from the last question to help overcome this confusion. There should also be some feedback informing the user they have reached the end of the test.

Another user commented about the inability to change your answer once they had pressed save. This is one of the options available for test delivery and is required for adaptive testing. However adaptive testing was not used in this instance and the user should have been able to alter their answer. When returning to a question already answered there is no indication that it is feasible to alter the previous entry. Under test conditions users may be reluctant to experiment with the software, again supporting the need for practise tests.

Some computer assisted assessment software allows the user to flag questions but there is no option in WebCT for users to do this. This is useful when the user is unsure whether they have answered the question correctly and they wish to return to it. Instead, the user has to navigate through the questions to find the question again. This flagging feature, whilst useful, could be problematic and lead to additional navigational problems if a large number of questions are presented.

One user commented that the digital clock was distracting. This has also been found in another study into computer assisted testing [8]. However there is no evidence to suggest that this has an impact on the users' performance. The feature is essential as it notifies users how long they have left to complete the examination but the clock could be altered so the seconds are removed, this is possible within WebCT, thus eliminating the continuous distraction.

4.3 Exiting the test

One student commented that they managed to accidentally log out after just one question and could not restart the test. On the face of it, this could be deemed a critical problem with the interface. However, further investigation identified there is only one way this could occur. The user would have to press the *Finish* button which would generate the following message *Some questions have not been answered: Do you want to proceed?* The later part of the message could be deemed ambiguous as to whether it means proceed with the test or proceed to end the test. If you proceed another message appears which states *Submit quiz for grading?* If you press cancel at either of these two points it takes you back to the test interface where you can continue. This is the only way a user could not regain access to the test. The language used in the feedback messages is of particular importance within a university setting as the users first language may not necessarily be English.

If a user exceeds the time limit for the test a message appears stating *Your time is up. Please submit the quiz now.* This message is inconsistent with the action that is required by the user as they must select the *OK* button and they are then returned to the test interface rather than automatically submitting the test. They then need to press the finish button creating an unnecessary step for the user, and they can also continue to navigate around the test.

The quantitative data in this study did not highlight any issues relating to the interface within WebCT, however this could be attributed to the questions being very generalised and not probing for specific details. The qualitative data revealed a number of issues that need to be further investigated.

5. CONCLUSIONS

This paper examined the assessment interface within WebCT to determine its usability and whether there are any factors which could result in a breach of construct validity. Based on this study there are a number of recommendations for delivery of online assessment.

It is evident from the study that users would benefit from earlier exposure to the interface prior to commencing any summative assessment. This may alleviate the problems of accessing the test and help familiarise themselves with navigation. When a user answers a question their entry should be automatically saved and no data should be lost when navigating between screens. Feedback messages within the test need to be explicit and matched to the users required action. This could prevent users from accidentally logging out of the software and improve learnability. The next question button should be removed when the user is on the last question. Although only one user reported this problem it will remove the confusion over its function. The interface should not appear in a pop up window that users have to resize to see all the questions and a browser should be used that offers the facility to indicate that the window will close.

6. FURTHER RESEARCH

Further research needs to be conducted to determine the impact these factors have on test performance and whether they could result in a breach of construct validity. The results from this investigation will be used to inform the development of a standard interface for test delivery using QuestionMark. This software comes with a number of templates and initial findings indicate that these may be inappropriate and therefore a new interface will be designed and evaluated.

7. REFERENCES

1. Bonham, S.W., Beicher, R.J., Titus, A., and Martin, L. (2000). The efficacy of a World-wide Web mediated formative assessment. *Journal of Research on Computing in Education*. 33. 28-45.
2. BS7988. (2002). *Code of practice for the use of information technology (IT) in the delivery of assessment*.
3. Daly, C. and Waldron, J. (2002). *Introductory programming, problem solving and computer assisted assessment*. Sixth International Computer Assisted Assessment Conference. Loughborough.
4. Dix, A., Finlay, J., Abowd, G.D., and Beale, R. (2004). *Human-Computer Interaction*. Pearson Education Limited.
5. Fulcher, G. (2003). Interface design in computer-based language testing. *Language Testing*. 20(4). 384-408.
6. McAlpine, M. (2002) *Principles of Assessment*. CAA Centre.
7. Pain, D. and Le Heron, J. (2003). WebCT and Online Assessment: The best thing since SOAP? *Educational Technology & Society*. 6(2). 62-71.
8. Peterson, B.K. and Reider, B.P. (2002). Perceptions of computer-based testing: a focus on the CFM examination. *Journal of Accounting Education*. 20(4). 265-284.
9. Ricketts, C. and Wilks, S.J. (2002). Improving Student Performance Through Computer-based Assessment: insights from recent research. *Assessment & Evaluation in Higher Education*., 27(5). 475-479.
10. Zakrzewski, S. and Steven, C. (2000) A model for computer-based assessment: the catherine wheel principle. *Assessment & Evaluation in Higher Education*. 25(2). 201-215.