# PILOTING CAA: ALL ABOARD

**Gavin Sim and Phil Holifield**

# Piloting CAA: All Aboard

Gavin Sim
Department of Computing
University of Central Lancashire
Preston PR1 2HE
Tel: 01772 895162
grsim@uclan.ac.uk

Phil Holifield
Faculty of Design and
Technology
University of Central Lancashire
Preston PR1 2HE
Tel: 01772 893168
pholifield@uclan.ac.uk

## Abstract

This paper examines the initial findings from a pilot project to determine the appropriateness of CAA software for university wide deployment. The various stakeholders within the institution have been identified and their opinions of CAA have been ascertained. Based on the findings a framework for incorporating CAA into modules is presented along with an analysis of the current barriers to adoption. The results suggest that there is overall support for the adoption of CAA but without support measures in staff development the uptake may be reduced.

## Introduction

Over the past five years the teaching and learning strategy at the University of Central Lancashire has incorporated a strategic framework for e-learning. This has emphasized the development of course content without fully utilising or embedding the capabilities of CAA. The first summative CAA test within the institution was conducted in the summer of 2003 using the assessment tool in WebCT. However, the question styles and features of WebCT are limited compared to more specialized CAA software. Both TRIADS and Questionmark were evaluated to ascertain their suitability for a pilot study into CAA within the university. Questionmark was adopted for the primary reason that staff could input the questions themselves, whereas it was felt that TRIADS would need a dedicated CAA officer.

Prior to project commencement, the technological infrastructure within the institution was analysed which raised a number of issues. These included institution-wide scalability, connectivity and bandwidth. This resulted in the acquisition of a dedicated server to host Questionmark, Internet Information Server (ISS) and SQL Sever for the deployment of CAA. Connectivity is a particular issue as the university has a number of remote franchise colleges. These colleges may be operating with different platforms, browsers and integration between systems may be a potential barrier to adoption. The

ability to integrate the CAA database with the student record system is also a crucial issue for success. These issues will be addressed at a later stage in the project depending on the suitability of the software.

The software is now being piloted within the Department of Computing to evaluate its appropriateness for widespread adoption. The key challenges to date centre around encouraging staff uptake, staff development, stakeholder acceptance and its perceived ability to test the range of cognitive skills identified in Bloom's Taxonomy. The project outcomes are to examine its suitability from a technological, practical and pedagogical perspective and this paper presents the initial findings.


**Methodology**

A questionnaire was designed and distributed to all teaching staff within the Department of Computing, n=34 with a response rate of 68%. The questionnaire comprised a number of dichotomous questions and open-ended questions to ascertain the views of staff members in relation to CAA and the level of support and training they would require to use CAA.

A framework has been developed for identifying particular areas of modules suited to CAA (Sim et al. 2003) and this was presented to six members of staff within the Department of Computing. The staff applied this framework to their modules and the results are discussed.

Another questionnaire was designed comprising dichotomous, Likert and open-ended questions to evaluate the students' experience of the assessment format. This included their acceptance of this technique, the question styles, language used and the usability of the software. The questionnaire was distributed to students on a first year web design module n=86 within the Department of Computing following a CAA test and the response rate was 94%.


**Staff Uptake**

According to the national survey of CAA use in higher education institutions, CAA is most widely practised within computing (Stephens and Mascia 1997). Computing encompasses a range of subjects from the technical disciplines such as networking to more subjective areas such as Human Computer Interaction. CAA may more readily lend itself to the assessment of a number of specific disciplines within computing for example, computer programming where there have been a number of research studies (Pathak and Brusilovsky 2002; Sayers and Hagan 2003). The results from the questionnaire revealed that only five members of staff had experience of using CAA and only three individuals are actively using CAA. This supports the notion that the adoption of CAA has usually resulted from the impetus of enthusiastic individuals rather than a strategic decision (O'Leary and Cook 2001). To encourage uptake CAA is being incorporated into the departments strategy, requiring all level one modules to utilise CAA for formative assessment, with summative being

optional. This has now resulted in five members of staff adopting CAA into their assessment strategy for this academic year. Although this still only accounts for 14% of the academic staff, it is a 500% increase from the 2002/03 academic year.

There is scepticism from lecturers about using Multiple Choice Questions for assessment within higher education (Maclaran and Sangster 2000). However, 91.3% of lectures in this study reported they would use CAA for formative assessment whilst this decreased to 56.5% for summative assessment. The difference could be attributed to the level the lecturer is teaching, as some of the lecturers questioned only teach at levels 3 and 4. Levels 1 to 3 relate to undergraduate programmes with level 4 being postgraduate. The lecturers were asked to specify the levels within computing at which they thought CAA was an appropriate technique for summative assessment and the results are displayed in Figure 1.
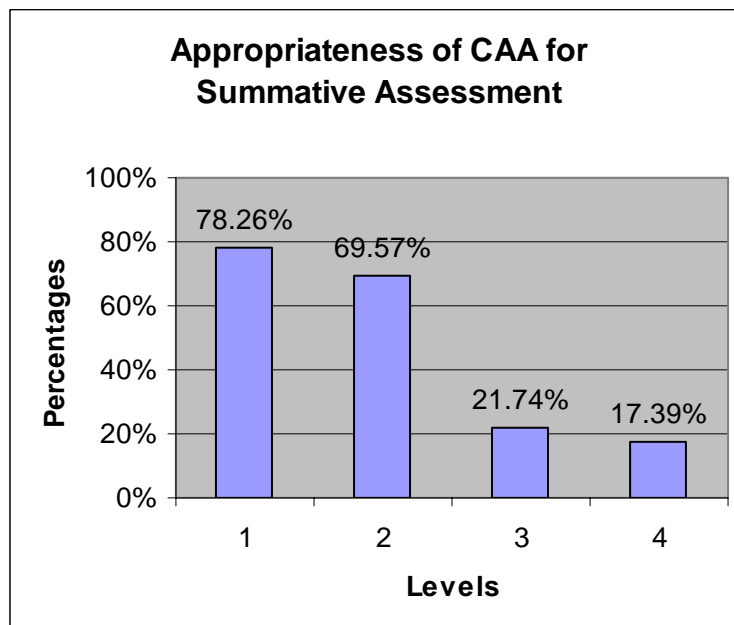


**Figure1. The appropriateness of CAA for summative at various levels**

It is clear that CAA is perceived to be an effective method for levels 1-2 but there is considerable scepticism that CAA is able to test levels 3-4 within computing. These levels are often associated with higher cognitive skills as defined in Bloom's Taxonomy (Bloom 1956). There is evidence to suggest that CAA can assess the higher cognitive skills within computing (King and Duke-Williams 2002), and as a means of addressing this scepticism a framework for breaking down modules according to Bloom's Taxonomy has been devised.


**Framework**

By analyzing the structure of a module, an understanding of how CAA can be incorporated into the assessment strategy may be achieved. Items need to be produced that cover both the content of the course and the educational

objectives (O'Hare and Mackenzie 2004). The framework identifies the learning outcomes (LO) for the module and categorises these to Bloom's Taxonomy. The syllabus area is mapped to the LO and then the syllabus is classified to Bloom's Taxonomy (Figure 2). Through this process elements at lower cognitive levels within the module can be identified and assessed using CAA. Once staff gain experience in designing questions at the lower cognitive levels they may feel more confident to experiment at the higher cognitive levels. An alternative technique can be used to test the higher cognitive skills or transferable skills required in the module, such as oral presentations or the production of a multimedia application. Paxton (2000) supports the notion that objective tests fail to examine critical or communication skills that are essential in the labour market. Therefore adopting a diverse assessment strategy may lead to a fairer assessment of the student (Race 1995).
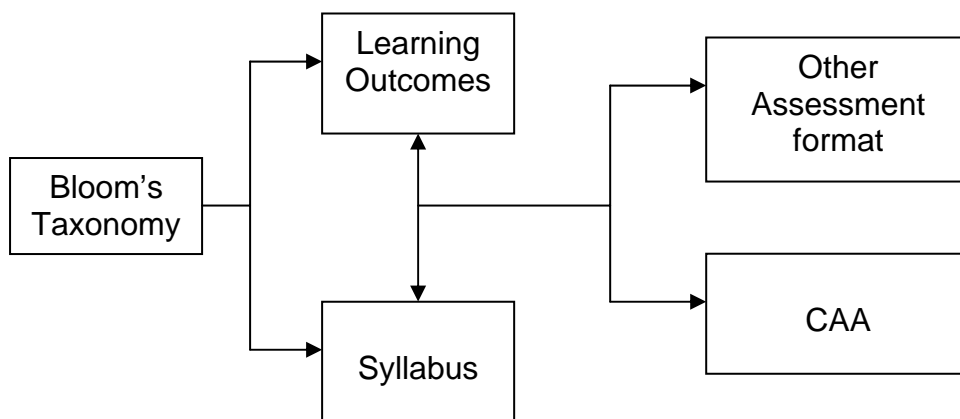


**Figure 2. Framework for classification of modules to Blooms Taxonomy**

The framework was applied to eight modules and the results are displayed in Figures 3-4. The modules analysed cover a broad range of computing subject disciplines including HCI/Multimedia and Systems.

|  | Level 1 | | | Level 2 | | | Level 3 | Level 4 |
|---|---|---|---|---|---|---|---|---|
|  | CO1652 | C01802 | C01804 | CO2751 | CO2752 | CO2601 | CO3707 | CO4707 |
| Knowledge | 1 | 1 |  |  |  |  |  |  |
| Comprehension | 3 | 1 | 1 | 1 | 1 |  |  |  |
| Application | 3 | 2 | 6 |  | 2 |  |  |  |
| Analysis |  |  |  | 1 |  | 3 | 2 |  |
| Synthesis | 1 |  |  | 1 | 1 | 2 | 1 |  |
| Evaluation |  |  |  | 1 |  | 2 | 1 | 3 |

**Figure 3. Number of Learning Outcomes at each level of Bloom's Taxonomy**

There are variations between modules in the number of LO ranging from 3 to 8. As expected, the level 1 modules mainly require students to demonstrate their ability at satisfying the lower cognitive skills. Highlighting the subjectivity of this process is the level 2 module CO2601, which requires students to demonstrate a similar ability to those found in CO3707.

The second step in the process is to identify the elements of the syllabus and their relationship to the LO. This will help to ensure that the course content is

related to the LO and prevent unrelated content being integrated into the assessment.

The syllabus is classified to Bloom's Taxonomy (Figure 4) and the results indicate that there are usually some elements of the syllabus that are at the lower cognitive levels even in a level 4 module. For example on the CO4707 module the lecturer classified nine elements of the syllabus at evaluation, however, there are still four elements at comprehension and three at application.

| | Level 1 | | | Level 2 | | | Level 3 | Level 4 |
|---|---|---|---|---|---|---|---|---|
| | CO1652 | C01802 | C01804 | CO2751 | CO2752 | CO2601 | CO3707 | CO4707 |
| Knowledge | 2 | | | | 1 | | | |
| Comprehension | 2 | | 1 | 2 | 2 | | 2 | 4 |
| Application | 2 | 3 | 2 | | 2 | | 2 | 3 |
| Analysis | 1 | | | | | 3 | 3 | |
| Synthesis | | | | | 3 | 2 | 1 | |
| Evaluation | | | | 5 | | 6 | 1 | 9 |

**Figure 4. Number of Syllabus elements at each level of Bloom's Taxonomy**

As an example the framework could be applied to an MSc Multimedia Design module. During the first semester there is a large amount of knowledge for students to digest with a steep learning curve. Regular CAA testing of this newly acquired knowledge would help both lecturers and students monitor their progress. This would provide early formative feedback to the students, which is often crucial for year long courses.

The framework shows how CAA can be incorporated into all levels within the computing discipline, however, there are still obstacles preventing lectures incorporating CAA into their assessment strategy which will be analysed.


**Staff Development**

The questionnaire ascertained the barriers to CAA adoption, and the staff development that would be required to successfully integrate CAA into the department. The lecturers were asked '*Would you be prepared to input the questions into the software yourself?*'. Of those who would use CAA 80% said that they would. However, this may not reflect the attitudes of staff in other departments who are less comfortable with using technology. The lecturers were also asked '*What additional support you would like to develop CAA?'* and 86.96% would require support using the software (see Figure 5). Therefore, although lecturers are willing to use the software it would appear staff training is necessary. Currently two research assistants within the department have developed some expertise in using the software and have implemented the assessments for staff. This has helped increase the uptake of CAA within the department but is not addressing the long term issue of staff developing their own assessments, which was one of the key goals when choosing the software.
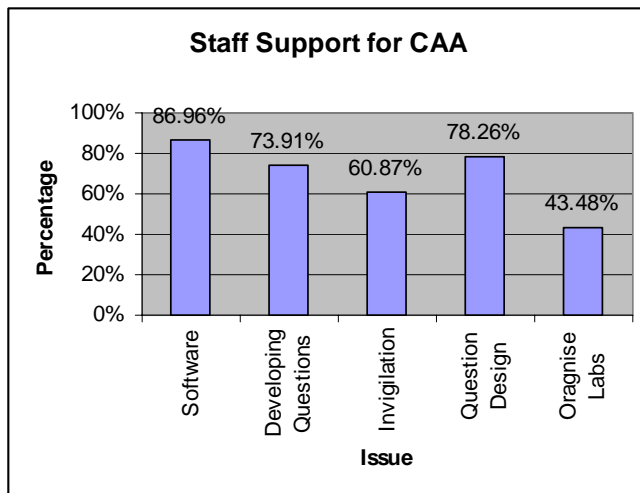
**Figure 5. Areas where staff require support**

The second area that lecturers feel they need the most support in is designing effective questions. McKenna and Bull (2000) suggest that a strong programme of staff development should help ensure competence in using the software, designing appropriate questions, invigilation and embedding CAA within the existing modules. The data gathered from the questionnaires support this synopsis. In 2003 the Learning Development Unit organised a staff development workshop entitled 'An introduction to Computer Assisted Assessment'. This proved informative for those who could attend, but without continuous development and with no allocated resources, meeting the needs of the staff is problematic and may reduce its adoption.

Although only 61% of lecturers stated they would need support with invigilation this is currently essential as there are no computer labs available that can accommodate large student cohorts.  For example, the large first year modules required 6 invigilators, 1 acting as a float, this has increased the workload of the department but with the development of new infrastructure this is envisaged to decline.

The greatest obstacle to CAA uptake by academics is lack of time (Warburton and Conole 2003). In this study lectures were asked '*By using CAA will you require more time to write questions?*' and 81% indicated more time would be necessary.  It is envisaged that a combination of staff development and experience gained through the initial pilot project should in the long term reduce development time. Also the emergence of question banks which are authored and peer reviewed by academics, such as the Electrical and Electronic Engineering Assessment Network (http://www.e3an.ac.uk), may help overcome the problem of staff development and ensure quality.

Informal focus groups have been established to initiate procedures for administering CAA within the department, discuss problems and share experiences. This includes issues such as how to accommodate students with special needs, what to do should the server fail, training and invigilation. This is important as initially lecturers were working independently and students' experiences were different from module to module.  Through this focus group support will be provided and this may encourage further staff uptake of CAA.

## Students

The students' attitude to CAA is being examined through a series of questionnaires examining their acceptance of the software and this method of assessment.

To determine their acceptance students were asked '*Would you find this format of assessment an acceptable replacement for part of your final exam?*' and they responded on a five point Likert Scale (Strong Disagree=0, Neutral = 2 and Strongly Agree =4). The mean is 2.9, standard deviation .9 and the 99% confidence interval =2.9 ± 0.26. Therefore, we can conclude that there is a reasonable level of support in this study and the majority of students would not object to CAA being incorporated into their final exam.

There is a considerable body of research acknowledging the impact computer anxiety can have in relation to CAA (Liu et al. 2001; Zakrzewski and Steven 2000). Although you would probably expect computing students to be comfortable with using a computer, this type of assessment might have caused them additional anxiety as they had no prior experience of using the software. They were asked again using a Likert scale whether they agreed to the following statement '*This format of assessment is more stressful than a paper based test*' and the mean score is .99, standard deviation .987 and the 99% confidence interval =.99 ± 0.28. Students within this study find this method less stressful than the more traditional method of a paper based test. One student commented '*I prefer completing a test in this way as it is less intimidating*'. This could be one explanation for the reduction in stress. Further analysis with different cohorts and student groups are planned but initial indications suggest the students are in favour of this method of assessment.

The students were asked '*Did you have any difficulties accessing the test?*' and 14% reported yes. It was observed that the majority of problems arose by students copying their username and password from an email and including an additional blank space at the end. Although easily rectified the software could be programmed to trim white spaces thus alleviating this problem or the authentication of the user should be achieved through the LDAP process.

Questionmark software was used to administer the tests using 'question-by-question' delivery and a standard template. Ricketts and Wilks (2002) suggest that the mode of presentation of assessment can significantly influence student performance and that appropriate screen design is perhaps the most important factor in online assessment. The students were therefore asked a series of Likert questions relating to the interface to determine its appropriateness (Figure 6) and the coefficient alpha was calculated for the scale and found to be 0.9161.

| Question | Mean | Standard Deviation |
|---|---|---|
| The test was easy to use. | 3.13 | .838 |
| It is easy to read the characters on the screen | 3.18 | .917 |
| The screen layout is clear | 3.06 | .843 |
| The screen layout is consistent | 3.15 | .823 |
| The navigation was clear. | 2.77 | .992 |
| I always know where I am in the software | 2.95 | .851 |
| The button location is consistent | 3.21 | .709 |
| The order of the navigation buttons is logical | 2.95 | .881 |
| The button names are meaningful | 3.01 | .845 |
| The on-screen navigation is easily distinguished from the questions | 3.13 | .858 |

**Figure 6. Mean scores for questions about the interface**

Overall the students responded positively regarding the layout of the interface however, the navigation would appear to be more problematic. Of the 81 students completing the questionnaire 30 provided qualitative feedback. This qualitative data revealed that the students requested the facility to go directly back to a previous question and did not like having to rely on the previous question button and this was citied 11 times. The location of the 'Proceed' button was felt to be inappropriate due to its proximity to the main navigation and this was citied 4 times. Therefore these factors will be incorporated into forthcoming tests and further analysis will be conducted to determine whether the effectiveness of the navigation has improved.

## Other Stakeholders

Information System Services (ISS) and the Management within the institution are being informed of the CAA pilot project development through a steering committee that has been established. This comprises of members of the Computing Department, Technology Department, ISS, Learning Development Unit and Faculty. The overall responsibility of the committee is to report the findings of the evaluation and determine the appropriateness of Questionmark for an institutional wide deployment. Without the overall support of the stakeholders it's unlikely that further funding will be made available for an institutional wide adoption and it may continue to be implemented by a small number of enthusiasts across the university without being fully integrated with other university systems.

## Conclusions and Further Research

It has been established that there is a certain level of scepticism from staff within the department over the appropriateness of CAA in computing modules at level 3 and 4. A framework has been devised based on Bloom's Taxonomy that may assist staff identify areas that could be suitable for CAA. Further research will be conducted using the framework to determine its effectiveness and practicality.

The department has incorporated CAA into its strategy and this has undoubtedly increased uptake. However, a number of barriers have been identified including time, staff development in using the software and question design. Developments in infrastructure, the establishment of focus groups and the continued development of procedures for CAA will hopefully overcome these problems in time.

The students in this study indicated that they would accept CAA as a replacement for their final exam and found this format less stressful than a paper based test. It is recommended that the software trims white spaces from the end of passwords or the LDAP feature is utilised to prevent students having problems accessing the test. An evaluation of the interface indicated an overall acceptance although the navigation could be made more efficient. Further research will also be conducted to evaluate the students performance in CAA compared to other assessment techniques.

A comparison will be made between the functionality, usability and capabilities of WebCT and Questionmark to determine the future strategic direction and to present to management in order to justify the deployment of further resources.

## Reference

Bloom, B. S. (1956). Taxonomy of Educational Objectives: The classification of educational goals. Handbook 1. Cognitive Domain, Longman.

E3AN (2004). Electrical and Electronic Engineering Assessment Network. http://www.e3an.ac.uk  Accessed 21/05/04.

King, T. and Duke-Williams, E.  (2002). Using Computer Aided Assessment to Test Higher Level Learning Outcomes. 5th International Computer Assisted Assessment Conference, Loughborough.

Liu, M., Papathanasiou, E. & Hao, Y.  (2001). Exploring the use of multimedia examination formats in undergraduate teaching: results from the fielding testing. Computers in Human Behaviour. 17(3): 225-248.

Maclaran, P. and Sangster, A.  (2000). Revisiting objective tests: A case study in integration at honours level. ALT-J 8(2): 58-68.

McKenna, C. and Bull, J.  (2000). Quality assurance of computer-assisted assessment: practical and strategic issues. Quality Assurance in Education 8(1): 24-31.

O'Hare, D. and Mackenzie, D.  (2004). Production of a CAA - advice for the academics. Advances in Computer Aided Assessment. D. O'Hare and D. Mackenzie, SEDA. Paper 116: 118.

O'Leary, R. and Cook, J.  (2001). Wading through treacle: CAA at the University of Bristol. 5th International Computer Assisted Assessment Conference, Loughborough.

Pathak, S. and Brusilovsky, P. (2002) Assessing Student Programming Knowledge with Web-Based Dynamic Parameterized quizzes. ED-MEDIA, Denver.

Paxton, M. (2000). A Linguistic Perspective on Multiple Choice Questioning. Assessment & Evaluation in Higher Education 25(2): 109-119.

Race, P. (1995). The art of assessing. The New Academic 4(3).

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.

Sayers, H.M. and Hagan, H.S.J. supporting and Assessing First Year Programming: The use of WebCT. Italics 3(1): 1-11.

Sim, G., Holifield, P. & McManus, B. (2003). Online assessment of multimedia: from Knowledge to Evaluation the problems of classification. CAL03, Belfast.

Stephens, D. and Mascia, J. (1997). Results of a survey into the use of computer-Assisted Assessment in Institutions of Higher Education, Loughborough University: 37.

Warburton, B. and Conole, G. (2003). CAA in UK HEIs - The state of the art? 7th International Computer Assisted Assessment Conference, Loughborough.

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.