STUDENT VIEWS OF FORMATIVE AND SUMMATIVE CAA

D.O’Hare

Student views of Formative and summative CAA

D.O'Hare
Centre for Interactive assessment development
University of Derby
Kedleston Rd
Derby
DE22 1GB
D.ohare@derby.ac.uk
Tel 01332 591741
Fax 01332 622747

Abstract

Over the past two years students taking two biology modules at the University of Derby have been assessed using computer assessments with TRIADs (Tripartite Interactive Assessment Delivery System) in both their formal end of module examinations and for scored formative assessments. We were keen to establish the student views of the use of computer assessment and thus over this period in addition to the overall evaluation of the modules the students were also given the opportunity to evaluate these assessments. In the first instance an open ended approach was taken, and students were given the opportunity to anonymously write comments on the computer examinations. The results of this were encouraging in that only a minority of students (~5%) made non-positive comments on CAA with the majority of students being very positive on their CAA experiences. In addition a range of useful comments in relation to the application of CAA were provided by students, pertaining to comparability with traditional examinations and student learning strategy these are also discussed.

However, the results of this initial evaluation though interesting were largely qualitative and left us with a number of areas which we did not have any useful information. We also wished to gain some information on the student perceptions on the validity of the assessment in terms of its level of difficulty and its relation to the content of the course. Thus a questionaire was constructed which measured student views of the structure, clarity, difficulty of the assessment (as well as their overall opinion). This was then delivered to students at the end of a scored formative assessment on Mendelian genetics and their formal module examination both produced using TRIADs. Thus the opinion of students on computer based formative exercises and computer based terminal examinations could be compared. Statistical analysis of the data from the questionnaires has revealed that students thought both types of assessments to be well structured, fair, clear and well matched to the content of the course. However, the examination although scoring at a similar level in other categories, was seen as more difficult by students (which was evident from the grades).

In addition to these quantitative measures, students were also offered the opportunity to make any additional comments regarding the assessment. This allowed specific problem questions to be identified, which provided extra information for the post test DIF analysis.

The author suggests that use of assessment questionnaires of this type can aid both the tutor and the student in the evaluation of the learning process.
http://www.derby.ac.uk/ciad/

Keywords

Evaluation, Student views, Student CAA Questionnaire, TRIADs

Introduction

Over the past decade there has been a large increase in the use of computer based assessment in HE in the UK (Stephens & Mascia 1997). However, their has been little published to date on student views of computer based assessment, particularly that based on more complex interactions offered by the TRIADs system (Mackenzie, 1997). Given some of the published work on the prevalence of computer anxiety among students (Tseng *et al*., 1997, Brosnan, 1999), the use of computers for assessment has been laid open to question. In addition, in some instances the economic imperative to take-up CAA has perhaps left the educational imperative behind. Thus there is a potential concern that the validity of CAA has yet to be established, it is certainly true that there are “rich veins of unanswered questions” in this area (Perkin, 1999). This comes along with a general recognition within HE that assessment is no longer separate to the learning process, but impacts on all stages of the learning process (Brown & Knight 1994). Given the history of CAA in the division of biology at Derby (see below), we were interested to observe the impacts of the introduction of CAA on the learning process and to further investigate the validity of such assessments.

The introduction of computer based assessment within the division of biological sciences at the University of Derby occurred in 1993 in year I modules. The uptake was based upon the TRIADs system, which had been developed by D. Mackenzie in the division of Earth sciences (Mackenzie, 1997). Uptake was primarily to save staff time as a result of increasing student numbers during this period. However, the advantages of the TRIADS system allowing staff to ask more complex questions (than simple multiple choice) was also a contributory factor in the uptake of the system. Whilst student evaluation of the system was taken up as part of the TRIADS project, no formal evaluation of the assessments occurred in the department apart from the standard module evaluations. Thus, it was decided to undertake a formal of the computer assessments in the division of biological sciences.

We were also excited by the possibilities of the TRIADs system in allowing us to easily integrate evaluation into the assessment process. This allowed us new opportunities to involve students in the assessment process, perhaps offering them the potential to become more active in the validation and evaluation of assessments.

The aims of the study were thus to gain student feedback of the use of CAA and to investigate the potential for using student feedback in the validation of assessments.

Methods

Two evaluation processes were followed, covering two years (1999-2001) in two year I biology modules, genetics and evolution and human biology. The evaluation processes were as follows; in 1999 both modules used an open ended comment system – students were simply provided with a piece of blank paper and asked to pass any comments that they had on CAA (terminal examination). This qualitative data was then analysed in an attempt to quantify the nature of the comments made and to look across the whole dataset to examine common responses.

In the second evaluation a more formal approach was followed, a questionnaire was designed to provide a more, quantitative evaluation of the assessments. This was delivered in 2000/2001 covering both a scored formative (coursework) assessment and a terminal examination in the core biology module genetics and evolution. The students after completing the assessment (coursework/examinations) were automatically delivered a computer based questionnaire (programmed in TRIADs) asking a series of fixed response questions. The questionnaire was limited to 8 items, in order to keep it brief, to avoid gaining superficial responses from students (Harvey, 1998). Details of the questions provided to the students are given in table 1.1. In addition students were offered the opportunity to make any other comments on the assessment in a text box provided and the input was recorded

Question	0	25	50	75	100
How would you rate the assessment questions?	Very poor	poor	Ok	Good	Excellent
How clear was the wording of the assessment questions	Very clear	Unclear	OK	Clear	Very clear
The assessment mainly tested trivial knowledge	Agree strongly	Agree	Neither agree nor disagree	Disagree	Disagree strongly
How well did the assessment questions relate to the course content and objectives	Very poorly	poorly	OK	Well	Very well
How difficult did you think the assessment was	Very easy	easy	OK	Difficult	Very difficult
The assessment was a fair test of your knowledge	Agree strongly	agree	Neither agree nor disagree	Disagree	Disagree strongly
The computer assessment was more challenging than a traditional paper based exam	Agree strongly	agree	Neither agree nor disagree	Disagree	Disagree strongly
How well organised and structured did you find the assessment	Very badly	badly	Ok	Well	Very well

Table 1.1 Details of the questions in the questionnaire.

Each answer was then scored in 5 categories (from left to right) 0-100 answer had a numeric value (see table), which allowed statistical analysis of the data. The data from the questionnaires was summarised in terms of the frequency of respondent's selec ting particular answers. In addition the means and SD were calculated. Analysis was performed using MS Excel, and statistica from statsoft.

Results

Qualitative evaluation exercise

A total of 134 papers were collected from a possible 200 (a return rate of 67%), blank responses were ignored from the totals for the basis of calculations. Details of the results, including the main areas of comment, are summarised in table 1.2.

	Human biology (mixed exam part written & part computer)	Genetics & evolution (Terminal examination)
Clear & well set out	20	13
Dislike negative marking	7	-
Faster than traditional	5	-
Q's harder and more complex	3	13
Help remember	2	-
Q's over broad range	6	-
Less stressful than written	3	1
Better than traditional	13	2
Interesting	4	5
Relaxed atmosphere	8	3

Table 1.2 – An attempt to categorise the responses made by students in the open-ended evaluation of CAA

In addition to comments that could be categorised (as above), students also took the opportunity to make a range of other comments. These will be discussed later.

Quantitative Questionnaire

Question	Coursework	Exam	Interpretation
How would you rate the assessment questions?	**75** 66.25 (14.39)	**50** 61.74 (14.68)	Good – OK
How clear was the wording of the assessment questions	**50** 59.06 (21.51)	**50** 58.71 (19.36)	OK-clear
The assessment mainly tested trivial knowledge	**50** 57.50 (22.98)	**75** 69.31 (24.72)	Disagree in exam, unsure in coursework
How well did the assessment	**75**	**75**	Well

questions relate to the course content and objectives	71.87 (17.95)	70.07 (19.72)	
How difficult did you think the assessment was	**50** 60.93 (15.32)	**75** 74.24 (17.52)	OK – difficult in coursework. Difficult in the exam
The assessment was a fair test of your knowledge	**25** 31.87 (16.38)	**25** 34.09 (20.39)	Agree
The computer assessment was more challenging than a traditional paper based exam	**50** 44.68 (25.68)	**50** 47.43 (22.41)	Unsure
How well organised and structured did you find the assessment	**75** 70.31 (18.69)	**75** 68.93 (18.11)	Well – OK

Table 1.3 – Summary of results from quantitative questionnaire. Median values are given in bold together with means (SD).

A total of 66 student completed the exam questionnaire (return rate of 88%) 75 students completed the coursework questionnaire (return rate of 100%).

In an attempt to validate the questionnaire, item analysis was performed using classical test measures. Including inter item correlation and reliability estimates. However, due to the nature of the measurement scales employed (items scoring in opposite directions) validation of such measures were not possible.

Discussion

The results of each of the evaluations will be discussed in turn

Qualitative data

Despite the useful level of comments and quotations on CAA, it was quite difficult to convert the student responses into quantitative data that could give some hard data on student opinions of CAA as an assessment tool. The usefulness of the process, however, was indicated by the level of useful comments made and by the areas covered. These included; the clarity of the questions, the time available, the structure and layout of the system, how the examinations compared to traditional assessments and the nature of the questions. The comments made were largely positive only a very small number (7, 5.2%) were not positive in terms of their overall impression of CAA. Among the minority of negative comments included use of a computer being tiring, requiring concentration over long periods of time. However, many of the other comments made were often contradictory (e.g. over available time), which is perhaps unsurprising given the nature of this evaluation.

To summarise the views of a diverse group, the students as a whole seemed to like CAA thinking it to be clearly set out, with students who expressed an opinion preferring the CAA traditional exams. Only a very small minority of students made negative comments about the use of CAA. The result was the illustration that a more robust tool was required to gauge overall student opinion.

CAA and student learning strategy

Many of the comments made by the students were extremely useful to the tutor in that they offered an insight into the impact of CAA on the teaching and learning process. Several students observed that the computer based exams were “more relaxing” than traditional examinations with computer labs having a better atmosphere than examination rooms, such observations have been made in other studies (Sambell *et al*,. 1999). In addition one or two students commentated that CAA questions were better in that they actually “helped you in answering” the question or “helped re-cap memory” and that “the answer is there”. Such comments would seem to support the work of Johnstone & Abusaidi (2000) who suggest that fixed response questions (especially multiple-choice (which made up a large proportion of the exam in Human biology) are testing re-call and may include a whole number of ‘clues’. Concerns have been expressed elsewhere about the potential for such CAA to promote surface learning (Twomey 1996). However, it must be noted that the genetics and evolution paper, which included no multiple-choice items received no such comments.

In addition, the genetics & evolution students seemed to have a view on the relative coverage of CAA and more traditional examinations. In general students appeared to state that CAA offered a much wider coverage of the syllabus than traditional examinations and that the approach of being able to focus on certain areas and as one student put it “..waffle a pass” was no longer appropriate. This prompted one student to comment that revision for a CAA exam was more difficult than for a traditional paper – in that it “….challenged your knowledge – more a case of your range of knowledge than what you had set out to revise”. Indicating that the old practice of picking 3 or 4 areas to revise in the hope of writing an essay on one of them was no longer appropriate. Thus, CAA may increase the breadth of coverage of student learning, however whether this is at the cost of depth of knowledge remains to be seen. Several students did comment that the depth of knowledge required for the types of question in the assessments was great and that the written paper was therefore easier.

The nature of the questions also drew comment, in that some students did find the questions more difficult than traditional assessments, (though this was not a negative comment), the level of detail required in the questions and the diverse nature of the questions was also noted as an interesting challenge.

One interesting observation was that several of the students saw fit to mention that the CAA exams were better than traditional examinations, in that “the range of different

question styles kept your interest to the end". The inclusion of high quality graphics was cited by several students as an attractive and exciting aspect of the assessments, an obvious difference to their experience of paper based assessments. Perhaps such factors contribute to the motivational value of CAA that has been note elsewhere (Mulligan, 1999). Unfortunately, it was not possible to examine if the interesting and exciting nature of CAA led to an increased performance among the students stimulated by their experience.

The use of negative marking in CAA examinations was commented upon by a small but significant minority of students, the majority of these students stated that they considered negative marking to be unfair, although one or two students did comment positively on its use. This suggests that perhaps the students have a poor understanding of the role of negative marking, as has been noted in other institutions (Ryle, 1996).

Computer anxiety

Tseng *et al.* (1997) noted a significant number of undergraduates suffering from computer anxiety (up to 30%) in computer assessments. Whilst one or two students mentioned that they were nervous prior to the examination they observed that the experience was in fact good on the whole. We feel that the practice of preparing students for CAA by exposing them to assessment type material to use in the their own time helped in reducing computer anxiety. This approach has been shown to be beneficial elsewhere (Sambell *et al*., 1999). Comments made by students also bore this out "..doing a similar thing earlier in the course made it comfortable to work with" .

Armed with this set of interesting evaluations, we wished to construct a short questionnaire designed to test the overall student opinion on CAA in a number of key areas including the structure, clarity and difficulty of the assessments. In addition we wished to gain some information on the students' relative views on formative/summative CAA. The questionnaire was applied to that same group of students during both a formative coursework assessment and a terminal examination.

Analysis of the quantitative questionnaire data

A number of general observations can be made regarding the students' opinions on CAA as a result of the analysis of the questionnaire data. Firstly the students thought the assessments were constructed from questions that they rated highly, these were ordered into well organised and structured assessments.

There are however, one or two interesting points to come out of the analysis in relation to some of the questions. In the exam 12% of students thought the questions were unclear – a concerning point, but only 3% of the students rated the questions as poor. There appears to be a contradiction here in that it may only be a minority of questions

that the students rate as unclear and this lack of clarity is not sufficient for large numbers of students to have a negative impression of the questions as a whole. This figure of 12% was also noted on the student responses to the questionnaire for the coursework. Which consisted of an entirely different set of questions, therefore there may be a rump of students who consistently find questions unclear, or report problems that they encounter with a question (perhaps above their level of ability) as a problem with the clarity of the question. An alternative explanation may be that this minority of students failed to utilise the available preparatory assessment materials by avoiding the time-tabled sessions on offer. It is worth noting here that the opportunity for open ended comment allowed problematic questions to be identified and re-examined by the tutors, which together with the post test DIF analysis allowed the validation of such items.

Validity of the assessment

One major factor in deciding to run the questionnaires was to investigate the potential for using student evaluation as a means of increasing the validity of the assessment. Thus, the questionnaire contained questions both on the fairness of the assessment and the relationship of the assessment to the course content and objectives (content validity), to establish the student views in these areas. Whilst tutors can often agree on the measurement offered by particular items (in terms of Blooms taxonomy, Bloom *et al*., 1956), the concordance with views of the students has been demonstrated to be low (Cox, 1976). The use of open ended comments also allowed students the opportunity to make any comments on issues of concern, this provided a good deal of useful feedback on items in the test and led extra weight to the DIF analysis on deciding on strong items in the test. Such an approach it is hoped, will allow us to determine the concordance of the measurement value of future assessments by establishing student views on the validity of the assessments that they sit. The inclusion of difficulty measures in the questionnaire was also related to this aim the intention was to examine the students perception of difficulty and relate this to the actual data analysis. It became clear when analysing the results of the coursework, that students have a poor perception of difficulty, the comparatively easy coursework was rated as difficult by 44% of the students (56% rating it OK or easy). The much more challenging exam (as rated by the tutor) was rated difficult/very difficult by 74% with only 25% rating it as OK. One might reflect that the students showed reluctance to label the coursework as easy, perhaps fearing a loss of easily gained marks !

Despite this perception of the level of difficulty both assessments were seen as a fair test of ability, 60% agree (7% agree strongly) in the exam and 72% agree (4% agree strongly) in the course work. Thus a good majority of students in each case agree that the assessments were a fair test of their ability.

One area in which we were keen to establish their views, was the comparison of CAA to traditional assessments. Whilst staff had established their own views on CAA in terms of its criterion validity and superior discrimination value, we had little knowledge of

student views in this area. The questionnaire asked the students whether they agreed with the statement that the assessment was more challenging than a traditional assessment. Unfortunately as can be seen from figure 1.1 the students appeared reluctant to commit themselves to either agreeing or disagreeing with the statement. Further investigation of this area is required.

1.1 – The number of respondents in each category when asked to agree with the statement that the assessment was more challenging than a traditional assessment.

One interesting outcomes of the evaluation process was that the views of the students on the use of CAA for coursework and examinations did not differ significantly. This was surprising, given the differing modes that the assessments ran and one anticipated a much readier acceptance of CAA for coursework than examinations.

Benefits of the evaluation processes

Whilst there are obvious benefits from using a quantitative approach to the evaluation of CAA. It is noted that a greater amount of detailed information on student perceptions was gained (albeit rather nebulous) from the open ended approach to evaluation, much of which was extremely valuable to the tutor. However, in order to produce data covering the views of the whole cohort a quantitative approach was required. This generated the required coverage, but lost a good deal of the useful info gained from the first evaluation. Unfortunately students appear to be reluctant to enter large amounts of text into the computer as opposed to on paper. Therefore, to maximise the benefit of the evaluation process, this author advocates a mixed approach to gain the maximum from any future evaluation of the implementation of CAA.

One must also bear in mind however, that we are assuming in this study that students have answered honestly to both forms of evaluation. It is also worth noting that we may be asking students to comment on areas that they have little knowledge or appreciation, there perception of what is a fair or ‘valid’ assessment may be valueless as they have no conception of what a ‘fair’ or ‘valid’ assessment would be.

CAA researchers are very concerned about the validity, reliability and fairness of CAA and this is rightly justified, but it must be remembered that in the use of traditional university examinations such considerations are often taken erroneously for granted. Therefore, one might also like to reflect on the possible outcome of a similar evaluation process conducted on more traditional assessments used in year I teaching at University level.

Conclusions

The conclusions reached as a result of this evaluation is that students could potentially provide a useful input into the validation of assessment. In addition this evaluation has

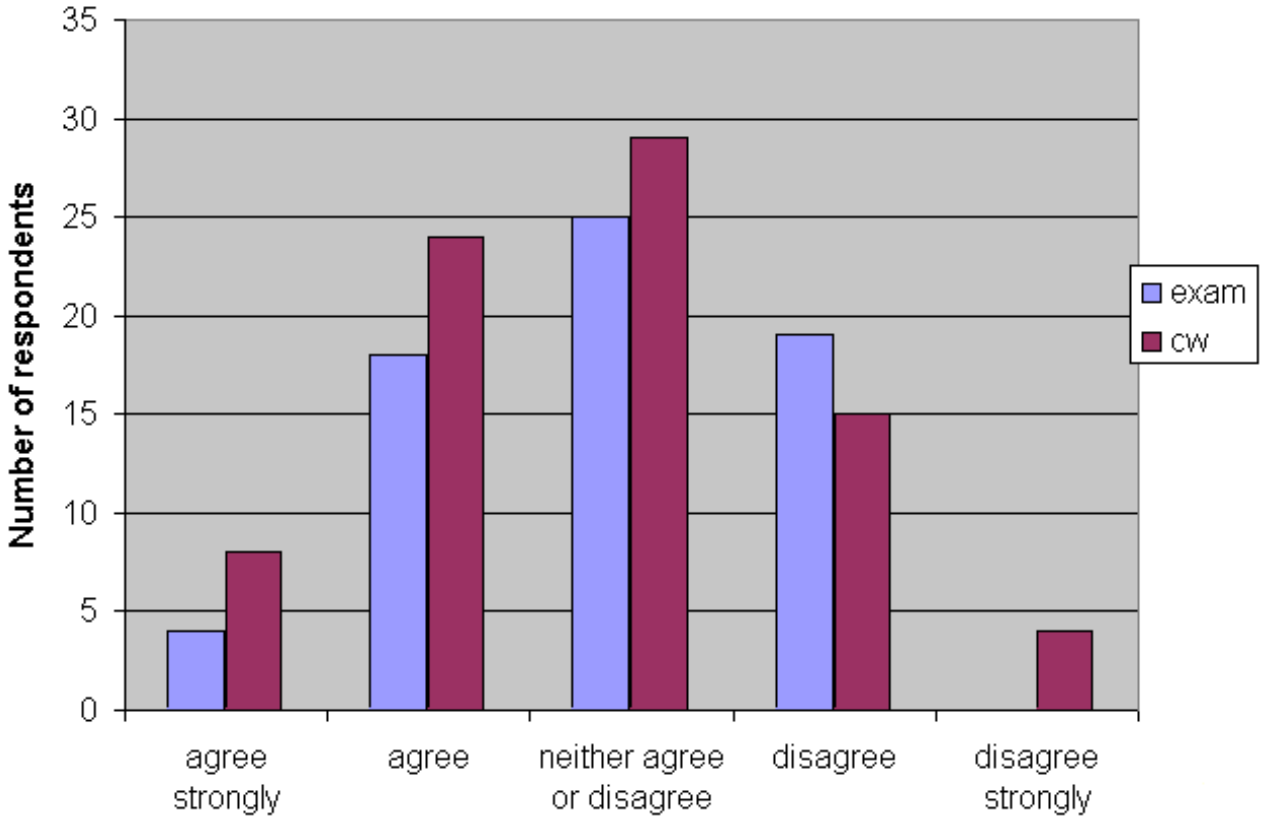
provided some useful information to staff on the best methods for the introduction of CAA i.e. the use of formative CAA based learning support materials can help in reducing student anxiety and can aid in acclimating students to CAA

The use of questionnaires of this type of questionnaire for computer based assessment could provide a useful means of increasing the knowledge available to staff on assessments that are delivered in particular allowing students to comment on particular items can provide extra information in the DIF analysis of assessments. Given the increase in the role of the market in HE the validation of assessments by this and other means may become increasingly important.

A useful by-product of such evaluations is the greater attention paid by both students and tutors alike to the assessment process.

A useful balance needs to be struck between the validity of the evaluation process and the information that can be gained as both systems utilised in this study have their own merits

Figure 1.1 - More challenging than traditional



References

Bloom, B.S. (Ed.) (1956) Taxonomy of educational objectives:The classification of educational goals: Handbook I, cognitive domain. New York ; Toronto: Longmans, Green

Brosnan, (1999) Computer anxiety in students: Should CBA be used at all. In CAA in HE (Eds) Brown, Race & Bull Kogan page

Brown, S & Knight,P (1994) Assessing students in higher education. Kogan page

Cox, K.R (1976) How did you guess? Or what do multiple choice questions measure ? Med J Aus 1 884-886

Harvey, J. (Ed) (1998) The Evaluation cookbook LTDI

Johnstone, A & Ambusaidi, A. (2000) Fixed response: What are we testing Chemistry Education: Research and Practice in Europe 1 323-328 URL :

Mackenzie, D.M. (1999) Recent developments in the tripartite interactive assessment delivery system (TRIADS). Proceedings of the 3rd Annual CAA Conference http://www.lboro.ac.uk/service/fli/flicaa/conf99/pdf/contents.pdf

Mackenzie, D.M. (1997) Fully interactive assessments on the web – A shock to the system. Proceedings of the 1st Annual CAA Conference

Mulligan, B. (1999) Pilot study on the impact of frequent computerized assessment on student work rates. Proceedings of the 3rd Annual CAA Conference http://www.lboro.ac.uk/service/fli/flicaa/conf99/pdf/contents.pdf

Perkin, M. (1999) Validating formative & summative assessment. In CAA in HE (Eds) Brown, Race & Bull Kogan page

Ryle, A.P. (1996) Objective tests: In defence of negative marking Life Science Educational Computing 7 no1

Sambell, K. Sambell, A. & Sexton, G. (1999) Student perceptions of the learning benefits of computer assisted assessment : a case study in electronic engineering in Computer assisted-assessment in higher education Ed Brown, Race & Bull Kogan page

Stephens D & Mascia, J (1997) Results of a survey into the use of CAA in HE in the UK January 1997

Tseng, H Macleod, H and Wright, P (1997) Computer anxiety and measurement of mood change Computers in Human Behaviour 13 305-316

Twomey, E (1996) Is there a role for computer based assessment. Computers in University Education University of Liverpool