Strategies for engaging students in computer based assessment – Stage 1, Taking stock

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is a conference paper.

Metadata Record: https://dspace.lboro.ac.uk/2134/1928

Publisher: © Loughborough University

Please cite the published version.
STRATEGIES FOR ENGAGING STUDENTS IN COMPUTER BASED ASSESSMENT – STAGE 1, TAKING STOCK

Joanne Smailes
Strategies for Engaging Students in Computer Based Assessment – Stage 1, Taking Stock

Joanne Smailes
Newcastle Business School
Northumbria University
Northumberland Road
Newcastle–upon-Tyne
NE1 8ST
(0191) 227 4966
Joanne.Smailles@unn.ac.uk

Keywords
Formative Assessment. Summative Assessment. Learning Approaches. Blackboard VLE.

Abstract
This paper details the first stage of an action research project within quantitative methods teaching. This project investigates student response to formative and summative computer-based assessment within the Blackboard virtual learning environment (VLE). Furthermore, initial investigation is made into the relationship between this form of assessment and student approaches to deep and surface learning. Overall, students reacted positively to these forms of assessment. Proposals to enhance participation in the formative elements were sought from the student body. Suggestions made included improving off-site access to the VLE, supplementing instantaneous feedback by including direct references to the quantitative text supporting the module, expanding resources to additionally include asynchronous feedback. However, collected evidence suggests students cannot strictly extricate formative from summative purpose of assessment. Preliminary observations into learning approaches and computer-based assessment revealed some interesting results. These were particularly pertinent within the overseas student body thus opening up further interesting lines of enquiry for future exploration. Therefore it is concluded that even though a move towards using the VLE assessment features for formative purposes only will be made, regular appraisal of strategies will be required.

Introduction
Nationally, whilst student numbers entering higher education has increased -19% to 33% over the period 1990/91 to 2000/01. (Source: Higher Education Statistics Agency HESA). Staffing at Newcastle Business School (NBS) has reduced from 141 full time equivalents in 1994/95 to 113.2 in 2000/01 (Source: HESA). Additionally, at NBS, contact time has also been reduced to typically 10% of the overall student workload allocated to a module.
As a lecturer in quantitative methods, providing alternatives to face-to-face provision of for formative assessment has been a priority. In line with many other enthusiasts (Weller, 2002; Peat and Franklin, 2002; O'Hare, 2001; McKenna, 2001; Thelwall, 1999) interest in computer-aided assessment (CAA) was based on improving student learning experience and seeking pedagogically sound forms of assessment to deal with the ever expanding diverse student numbers and increasing demands on staff time. Therefore, along with many other HE institutions (Griffiths et al, 1998; Stephens, 1994) CAA was seen as the most practical and strategic solution (Pollock et al, 2000; Harvey and Mogy, 1999; McCormack and Jones, 1998).

Background
At Northumbria University, projects involving Question Mark, the web and Optical Marking Reading. (Sambell et al, 1999) have a well-established history. However, as identified by Stephens et al (1998) to raise the profile of CAA, institutional strategy is required. At time when I began to investigate the use of CAA no such strategy was in place. As a result, organisation, implementation and administration of CAA were placed on the academic. For most this was not too onerous, but in the case of business modelling where module numbers are approximately 1000 there were major implications for staff time. However, in 2001, Elearning strategies led to the introduction of the Blackboard Virtual Learning Environment (VLE). Its integrated assessment features rekindled a personal interest in the use of CAA for large numbers. Firstly, CAA was used purely formatively, in the following year, summatively, both with limited success. In this academic year, CAA has acted as a combined summative and formative exercise.

Inclusion of summative elements is not a personal ideal. Therefore an action research project to move towards a purely formative use of CAA and investigate whether this form of formative assessment can motivate students to take a deep approach to their learning is being pursued. This paper will detail the first stage of this action research by firstly, taking stock of personal observations and student views over the last three years and secondly using the 2002/03 cohort, initially investigate students’ approaches to learning.

Background - Formative Assessment
There has been a resurgence of interest in assessment within Higher education. Formative assessment, in particular, is highlighted by several authors (Holroyd, 2000; Hager and Butler, 1996; Gipps, 1994; Wolf et al, 1991) as being worthy of increased attention. McDowell (2003) stresses that within an e-learning environment where there is less staff-student contact “effective formative assessment is likely to be particularly important”. Peat and Franklin (2002) also note that students now often expect and receive fast and direct feedback.

Some doubt the effectiveness of CAA in either formative or summative assessment, particularly in relation to building conceptual understanding and higher thinking skills (Laurillard, 1993; Jones, 1990). In a survey of academics McKenna (2001) also noted that most criticism of CAA arose when considering limitations in question design. Biggs (1999) argues that question
types most associated with CAA cannot be used to authentically assess skills where “an active demonstration of the knowledge” is required.

Quantitative methods is generally very factually based with emphasis placed on application and understanding, rather than mathematical ability. By ensuring that CAA is used to appropriately focus on “the type of learning a question is designed to assess” not the ease with which the computer can evaluate the question (Gibbs and Peck, 1995), along with Carnesin et al (online), it is believed that well written objective test questions can effectively probe understanding.

Besides, in a formative form, the factual and application based nature of quantitative methods can also be turned into a great benefit, feedback can be used to provide hints in commonly misunderstood areas in order for problems to be addressed. This view is concurred by McDowell (2003) and Brown et al (1999) were rapid feedback can be used to help readdress their current achievements against overall module learning outcomes.

**Objective testing and Approaches to Learning**

A surface learning approach is widely associated with recall and reproduction, whereas deep approaches to learning focus on meaning and understanding (Atherton, 2002). Distinctions between deep and surface learning are a well-researched area much of which makes use of the Study Process Questionnaire (Biggs, 1987b) for categorisation. Later work concentrates on refining scales etc. (Waugh, 1999; Entwistle and Ramsden, 1983) thus according to Atherton (2002) “explicating the symptoms of each approach” Atherton (2002) and McDowell (2003) report that surface learning is associated with an uphill struggle or “getting by” characterised as “fighting against boredom” as opposed to deep approaches, which are allied with the development of a genuine understanding characterised by a stimulating challenge.

There is some evidence that objective testing and hence CAA is associated with surface approaches to learning (McDowell, 2003; Sculler, 1998). However, McDowell (2003) states that lecturers can be encouraged to adopt practices within Blackboard, which could motivate students to take a deep approach to learning. This includes considering

1. Building space into the curriculum for testing.
2. Considering carefully the formative/ summative strategy.
3. Careful design and delivery of tests.
4. Bearing in mind the nature and quality of feedback.
5. Embedding of objective testing in other learning activities.

Suggestions 1, 3 and (to a certain extent) 4 are already considered to be present in the tests examined here.
Development of CAA in QM
CAA within Blackboard: Years 1 and 2

Currently, Northumbria University is only able to fully support Blackboard’s in built assessment features. Although, other alternatives, which could be incorporated, for example Question Mark Perception and Hot Potatoes are being investigated. The assessment features of Blackboard are not the most advanced or flexible available. For example, it is well established that a major advantage if CAA is the ability to take a bank of questions and present students with a random selection (Beilby, online; Kelly et al, 1996; Thewall, 1998). Even though Blackboard’s on-line assessment manager has a question pool, there is no option to present questions randomly.

Blackboard was first introduced, in pilot form, at Northumbria University in February 2001. At this point, a number of “formative self-tests” as defined by Zakrewski and Bull (1998) were designed and written for quantitative methods revision. Anecdotal evidence suggested these were well received by the students, but use was relatively low (around 25%) and tended to be concentrated around examination periods. At this stage no formal analysis was carried out.

The following year, 2001/02 CAA within Blackboard was used for one of the summative elements of the module. Given its summative nature, to make the test as secure as possible, as recommended by Whittington (1999) and Thewall (1999). The following steps were taken

- the availability feature within Blackboard, ensured questions could not be viewed prior to the test;
- users were restricted to just one attempt;
- a time limit was set.

Additionally, to counteract any potential technical difficulties (Whittington, 1999), contingency plans in the form of a paper grid to double record answers along with full replicate paper tests were produced.

This proved to be a wise decision. The week in which the summative exercise took place, there were major problems with the internal network. A one-off case of bad timing! However, subsequent feedback (Smailes, 2002) was generally positive and did serve to highlight some pertinent issues:

- Slow network response and network failure blocked out students as having taken the test. This meant restricting students to one test attempt was counter productive. Approximately 15% of students had to make use of the contingency paper version.
- The test timer was ineffective – students could continue regardless. All “timed-out” attempts had to be effectively marked manually, increasing administration time.
It was concluded that CAA within Blackboard was still worth pursuing, but with an adapted assessment strategy.

**Reviewing Year 3**

In 2002/03 a greater number of short tests with feedback were developed. Students were allowed to take these at their own pace with multiple attempts, but to counteract cheating to a small degree, during a restricted time frame of around a week. A random selection of results from the shorter tests was to be selected for summative purposes. Also in Year 3, a more formal approach was taken to evaluating the use of CAA.

**Feedback Instruments**

Formally, this has taken on two forms.

1. A three-part questionnaire, the first section collecting demographic and background information. The second consisting of five-point Likert scale attitudinal statements with the third section determining a students’ approaches to learning – detailed later.

2. Group Feedback Analysis (GFA), a staged process where
   
   a. Individuals are asked to write down anything regarding the Blackboard quizzes.
   b. Students are placed in small groups, individual comments are assembled into common themes.
   c. From these themes statements are formulated and presented to a group as whole.
   d. Individual scores 1 (definitely not important) to 7 (incredibly important) are sought, results collated and median summary score calculated.
   e. Finally, further open discussion is conducted.

The questionnaire was distributed to 110 students, 68 were returned giving a response rate of approximately 62%. Based on this sample, there was an even split between the genders with 57% male and 43% female. Around one-third of students were overseas - mainly from China. Only one student reported a disability – dyslexia. Group feedback analysis was conducted on three occasions involving a total of 18 students.
Discussion of Results
Concurrent with many other CAA studies (Thomas et al, 2002; Peat and Franklin, 2002; O'Hare, 2001; Thelwall, 2000; Charman and Elmes, 1998), in general the quizzes were well received. Students were extremely positive about quizzes' coverage of content and module learning outcomes as well as the variety and types of question used. Overall, three clear themes for further consideration emerged.

- Technology.
- Feedback.
- Separating formative from summative assessment.

Each is addressed in turn.

Technology
Due to the strategies of allowing multiple attempts in a restricted time frame no technical problems, as experienced previously, occurred.

Previous studies (Ricketts and Wilks, 2002; Weller, 2002 Thomas et al, 2002) reported technological difficulties such as the need to scroll between questions and difficulties in reading text. Author recommendations were implemented into test design and as a result, these factors were not found to be an issue.

Two-thirds of respondents reported having access to a PC with Internet capability at their residence. However, questionnaire responses indicated the tests were usually accessed from the teaching or casual access laboratories on campus. At the GFA stage, greater accessibility to quizzes from home was highlighted as an issue of importance. It was also observed that a number of students accessed the quizzes at weekends (albeit around deadlines). Therefore the desire for greater accessibility to quizzes needs to be considered in future iterations.

Feedback
In line with many studies (Thomas et al, 2002; Peat and Franklin, 2002; O'Hare, 2001; Thelwall, 2000; Charman and Elmes, 1998), instantaneous feedback was cited as one of the greatest strengths of the assessments. Additionally, during GFA, quality and speed of the feedback was consistently listed as a vital component and further suggestions for improvement were made.

At the construction stage, careful thought had been given to the form of feedback provided. Charman (1999) suggests that feedback is as “positive as possible” and kept “simple and friendly”. McDowell (2003) reports that “feedback needs to focus on task” and Black and William (1998) expressing concerns that non-specific feedback, even praise, is unconstructive. In this case, intentionally no feedback was given for a correct answer. Feedback for
incorrect answers was non critical and given in the form of hints and tips and/or a quick review of correct technique(s).

Within GFA suggestions of improvement were made including:

- “Page references and further hints would be useful as feedback.” The module under which the tests operate is directly linked to a “set text” the suggestion was clarified as pointing the user to the pages within the textbook where there are similar worked examples. This also helped clarify the usefulness of providing hints.

- “Feedback to include detailed working similar to seminar questions” As part of the contact sessions students were provided with fully worked answers to any questions. Given the nature of the subject graphics are often required. Current technical restrictions of Blackboard make this difficult to provide instantaneously.

- “Review the test questions most of the group got wrong with the group as a whole.” The nature of the request means this cannot be provided in an immediate form but as with the suggestion above, these could be implemented as an asynchronous form of feedback.

Unfortunately, a restriction of the Blackboard assessment features means feedback cannot be given without the correct answer being supplied. As test designer, this is not ideal, 60% stated they followed up on the feedback provided - but was this simply typing in the correct answer? During GFA a number of students expressed the same concern – highlighting a debate between the formative and summative aspects of the quizzes.

**Separating formative from summative**

As recommended by McDowell (2003) no measures were taken to reduce collusion and multiple attempts at the quiz were allowed. Yet a number of comments such as “feedback makes it too easy to cheat” and “tests should only be able to be taken once to stop cheating” were made during GFA discussions. Notably, overseas students appeared to be showing the greater concern. To illustrate, in one group (mostly overseas) the following statement was formulated and scored highly (median=6).

“Strict time limits should be in place on quizzes under exam conditions”

Conversely, in another group (mostly home) a similar statement was considered definitely not important (median=1).

“Individual tests should be given stricter time limits i.e. 30 mins”

Additionally, all students appeared to focus on grading. A large majority (80%) of respondents cited the provision of marks as an essential element of the quizzes. It should be noted that, for simplicity, grades for the quizzes were presented in percentages. This was not intended to reflect equal
weighting across (or even within) the individual quizzes. However, a number of students requested that an average of all the tests should be used as the summative mark - completely removing any formative provision. In GFA, students also expressed a desire to have near misses and working (e.g. rounding errors) credited in some way.

Within formative assessment, the allocation of marks is a contentious issue. Taras (2002) reports that students found it easier to “assimilate feedback and make judgements” without a grade. Sadler (1989) states that as grades represent a “one-way cipher” for students they may be counterproductive for formative purposes. Black and William (1998) suggest marks may undermine formative assessment. It is further noted that a series of bad marks may lead to a downward spiral for student learning and confidence. Does the student desire expressed here contradict much of the theory or are they simply unable to separate out formative from summative assessment? Nevertheless, students in this case were very keen to receive grades.

**Learning approaches – Initial Observations**

Studies such as those conducted by Scouller (1998) and Biggs (1999) present evidence suggesting that a surface approach to learning is more likely to be employed when multiple-choice questions are used as an assessment method. However, in this area, little is yet written which, considers the full variety of question types available within CAA. Further personal research will be conducted in this area but within the first action research cycle and hence this paper, only preliminary observation and comment is made.

These observations were made by utilising the third section of the questionnaire (Appendix 1). This was a personal adaptation of the Biggs (1987a,b) study process questionnaire combined with the Entwistle and Ramsden (1983) Approaches to Studying Inventory. It consisted of 27 statements, 9 depicting deep surface approaches, 9 surface learning approaches with a further 9 depicting a third form known as the achieving approach (Atherton, 2002). Respondents were asked to rate these statements under two circumstances, general study behaviour and behaviour when studying the quantitative methods module. Response format was based on an interval scale recommended by Waugh (1999):

3 – True of me all or nearly all the time
2 – Usually, though not always true of me
1 – Sometimes, but mostly not true of me
0 – Almost never or never true of me

When considering general study behaviours 35% of respondents were assessed as taking an achievement approach, 30% a surface approach and 20% a deep approach to learning. In he remainder the responses were inconclusive.
Given the initial literature review conducted, there was some anticipation that students would exhibit a surface or at least a different learning approach towards the quizzes. Contrary to expectations, no significant difference between approaches was found. In fact only 10% of the sample showed any change in their learning approach. Surprisingly, 75% of those appeared to take a “deeper” approach to learning when participating in the quizzes. It is possible that this could be a reflection of some of the recommendations made by McDowell (2003) such as building space into the curriculum for testing; careful design and delivery of tests; nature and quality of feedback, being already in place.

Of additional interest, as reported by Biggs (1999), there appears to be an inherent assumption in Western universities that international students, especially those from the Pacific Rim, exhibit surface approaches to their learning. Biggs (1999) argues that even with slight adjustments to accommodate cultural differences some evidence to the contrary exists (Biggs, 1987a; Volet and Renshaw, 1996). Much of the “accommodation” modifications recommended (e.g. provide as much visual back up as possible) naturally exists within CAA. In this study, it was discovered that no significant difference in the proportions of “surface” and “achievement” learning styles across “home” and “overseas” students existed. Yet, for those students exhibiting a deep approach there were a higher proportion of overseas students. Could these initial observations be evidence that CAA can promote deeper learning in overseas students?

**Discussion and Conclusion**

As expected, at the first stage of an action research project, many issues are simply highlighted rather than resolved. However, this study has provided some clear messages and points for action.

Many of the technological fears expressed by students in early studies have now been overcome (Ricketts and Wilks, 2002; Weller, 2002; Thomas et al, 2002). As Student PC ownership increases, more demands are made for greater and faster remote access to computer-based materials (including CAA). Institutional strategies may now have to include the provision of advice on how access outside of the university can be improved.

It is clear that, feedback in any form plays a vital role in student engagement with CAA. For the next action research cycle action will be taken to

- improve the instantaneous feedback within the quizzes by including page references from the set text.
- additional asynchronous feedback will be provided within Blackboard in the form of full solutions and regular reviews of group performance.

On the other hand, there is still some headway to be made on shifting student emphasis away from summative assessment. Observations in this area leave personal doubts as to whether students will engage with CAA in a purely formative form. In accordance with advice given by McDowell (2003)
engagement may be improved by making further advances to relate and embed CAA exercises in other forms of learning activity (for example a case study). Therefore it is proposed that for a trial period next year, with improvements in place, quizzes will operate on a purely formative footing. However, contrary to advice (Taras, 2002; Black and William, 1998; Sadler, 1989), grades will be given, but on a simplified basis. (i.e. each question will be awarded 1, 2 or 3 marks according to difficulty). Engagement will be monitored; dialogue of formative purpose and requests for improvement to such will be retained with the group. However, if necessary this strategy will be re-visited and summative components re-introduced.

Finally, the interesting initial observations on learning approaches has opened up fascinating new lines of enquiry. Can CAA in its new forms offset some of the previous thinking surrounding surface learning approaches? Does CAA help overcome international diversity of the student group? These investigations will be incorporated into the second action research cycle.

References

Anon (online) Higher Education Statistics Agency <www.hesa.ac.uk> (Jan 2003)


Carencson, J., Delpierre, G. and Masters, K. (online) Designing and Managing multiple choice questions. <www.le.ac.uk/castle/resources/mcqman/mcqman01.html> (February 2002)


Holroyd, C. (2000) *Are assessors professional? Student Assessment and the professionalism of Academics Active Learning in Higher Education* 1 (1) 28-44


Appendix

Section C: Approaches to Studying

This section is based of two well-established research questionnaires into processes and approaches of students to study (Study Process Questionnaire and Approaches to Studying Inventory). It has a similar purpose to that of the learning styles questionnaire used in skills module at the beginning of semester 2. Please answer the questions truthfully. If you wish to find out more about this research and its purposes, please do not hesitate to get in touch.

Please rate the 27 statements according to the following response format and place a number corresponding to your general study behaviour and your study behaviour on the Data Analysis (BM118) module in the appropriate boxes next to each statement.

3 True of me all or nearly all the time
2 Usually, though not always true of me
1 Sometimes but mostly not true of me
0 Almost never or never true of me

Example: If you “keep neat and well organised notes” for most of your modules put a 3 in the **general behaviour** box. However, If in Data Analysis you find you don’t manage to do this put a 1 in the **studying BM118** box.

<table>
<thead>
<tr>
<th>Item wording</th>
<th>General Behaviour</th>
<th>Studying BM118</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 My main motivation for this present course/programme is due to employability, rather than out of interest in the subject</td>
<td></td>
<td>Not applicable</td>
</tr>
<tr>
<td>2 I find at times studying gives me a feeling of deep personal satisfaction</td>
<td></td>
<td>Not applicable</td>
</tr>
<tr>
<td>3 I want top grades in all modules so that I will be awarded a high degree classification</td>
<td></td>
<td>Not applicable</td>
</tr>
<tr>
<td>4 When studying, I often find I reminded of material I already know and see this in a new light</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 I keep neat and well organised notes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 I work hard when I’m studying and generally manage to keep my mind on what I’m doing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 I make sure that I find conditions for studying which let me get on with my work easily</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 I make a point of doing all of the suggested reading and extra study exercises</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 I try and do all directed learning tasks as soon as I am able</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 I usually set out to understand for myself the meaning of what we have to learn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item wording</td>
<td>General Behaviour</td>
<td>Studying BM118</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>After a contact session, I re-read my notes to make sure they are clear and that I understand them</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I learn some things by rote, going over and over them until I know them by heart.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>When studying, I think of real life situations where the material may be useful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I concentrate on what lecturer’s say and believe this is more important rather than relying on my own judgement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have a strong desire to do well in all my studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tend to prefer the factual content of topics rather than theoretical components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel virtually any topic is interesting once I get into it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to relate to new material to what I already know on that topic/subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to work and review material regularly throughout the semester</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Although I can remember facts and details, I often can’t see any overall picture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I only study seriously material given out in classes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tend to become more deeply absorbed in my work, the more I do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to relate what I’ve learned from one subject to that in another</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I generally restrict study to set tasks rather than do anything extra on my own</td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I’m working on a new topic, I try and see in my own mind of the ideas fit together</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often I find myself reading things without really understanding them</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe shouldn’t have to spend significant amounts of time on material that will not be assessed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

405