applied infoKit

Effective Use of VLEs: e–Assessment

JISC infoNet

www.jiscinfonet.ac.uk

e–Assessment

Introduction

The idea of having our assessments computerised is obviously attractive and e–assessment is emerging as a major driver for e–learning for both students and staff. e–Assessment (sometimes known as Online Assessment, Computer–based Assessment or Computer Assisted Assessment – CAA) certainly has advantages. However, it also has some disadvantages and associated problems.

In this section we shall look at advantages and disadvantages, the practical considerations that need to be taken into account and some background considerations. For a start, particularly if you are new to this area, it may be useful to look at the definitions on the JISC and QCA (Qualifications &Curriculum Authority) e–assessment glossary.

What Do We Mean by Assessment?

Assessment is one of the most significant areas of an educational system. It defines what students take to be important, how they spend much of their academic time and in may ways how they value themselves. Rowntree (1987: 1) says of assessment 'if we wish to discover the truth about an educational system, we must look to its assessment procedures.' In addition, assessment is important because students cannot avoid it, as Boud (1995: 35) says:

'Students can, with difficulty, escape from the effects of poor teaching, they cannot (by definition if they want to graduate) escape the effects of poor assessment'.

This underlines the importance of getting our assessment practices right for our students.

Rowntree (1987: 1) states that assessment procedures offer answers to the following questions:–

'What student qualities and achievements are actively valued and rewarded by the system? How are its purposes and intentions realised? To what extent are the hopes and ideals, aims and objectives professed by the system ever truly perceived, valued and striven for by those who make their way within it?'

Allied to this is the fact that assessment has two main purposes within further and higher education:

- The first reason is to assist learning. When looking at this area we must always strive to make the assessment relevant to the overall goals of the unit and to make our assessment part of the learning process.
- The second is to determine the effectiveness of the education system. Only with this can we as educators improve the education of our students. However we must be able to determine not only the overall learning but which areas are not effective and need modification.

As tutors we assess for a variety of reasons:

- To pass or fail a student.
- To grade or rank a student.
- To select for future courses.
- To predict success in future courses.
- To provide a profile of what a student has learnt.
- To diagnose students' strengths and weaknesses.
- To provide feedback to students to improve their learning.
- To help students to develop their skills of self–assessment.

- To motivate students to provide feedback to teachers.
- To evaluate a course's strengths and weaknesses.

We must then question what we are assessing in the first place. A number of assessment points must be considered among which are:

1. What do we want to assess?

    Basic knowledge, skills, higher cognitive skills.
2. For what purpose?

    Diagnostic, formative, summative.
3. In which mode?

Norm–referenced, criterion referenced.

We need to be specific; for our students we need to ask the questions why, what and how, and relate these to the objectives of our courses and the learning outcomes we devise for our students

We must ask these questions to make sure that our assessment matches our educational purposes. As a tutor we then should find the most appropriate assessment method for the set assignment or to assess the desired learning. When considering the assessment tasks we set we need to consider the strengths and weaknesses.

We should also appreciate that students expect to receive much of their information, whether educational or social, from online resources and so we should be moving to assessing them by ongruent means. As student numbers have increased, and staff contact hours have in many cases decreased, students have asked for supplementary support. An example of this is given by Clarke 'Students had requested additional ways in which to learn and judge their progress during periods of low contact time with their tutors – especially in the lead up to examination' (Clarke *et al.*, 2004:250)

'From the INQUIRE evaluation it is strongly indicated that reinforcing the content of the lectures through formative assessment can act to cement students' understanding of key concepts and ideas' (Clarke et al., 2004: 259).

Methods of Assessment

We have already dealt with the different reasons for assessing students and the purposes for which students are assessed. Here we will look at the different types of assessment, whether assessment is diagnostic, formative or summative, and what different methods may be employed to assess students, whether these are in–course or end of course. Students can be assessed by their tutors, placement or professional mentors, their peers or by themselves. However, as Ramsden (1992) says, it will be rare to find one assessment method which will satisfy the assessment of all the intended learning outcomes for a course, so it will be necessary to consider a range of assessment methods for our students. Weavers (2003:13) concurs:

'diversity decreases the dependency on the traditional formal examination, a method that does not suit the learning styles of many students'

Diagnostic Assessment

As it suggests, diagnostic assessment is used to diagnose the level of learning that has been achieved by our students, and is generally used at the beginning of course units for staff to determine the level at which they should be aiming their teaching, or to suggest to staff (or students

themselves) the level of support that may be required. Staff may use diagnostic assessment at the end of a lecture, or a series of lectures, to see if students have comprehended the information conveyed, and students appear to like this, as it is a way for them to keep a track on their learning. However, diagnostic assessment does not provide a tool to enhance student learning unless it has an element of feedback within it, unless it becomes formative.

Formative Assessment

Assessment that is formative occurs during a course, and provides feedback to students to help them improve their performance. The feedback need not necessarily be derived from only the tutor, but can be from students' peers or external agents such as clinical tutors or placement supervisors. It is important that the feedback should be given in relation to the criteria against which the work is being assessed. Involving students in peer assessment aids students in understanding and using the assessment criteria (Bradford, 2003). Indeed, 'Giving feedback on another student's work, or being required to determine and defend one's own, not only increases a student's sense of responsibility and control over the subject matter, it often reveals the extent of one's misunderstandings more vividly than any other method' (Ramsden, 1992: 195–6).

Summative Assessment

Assessment that is summative may or may not include feedback. The main difference between this form of assessment and that which is purely formative is that grades are awarded. The grade will indicate performance against the standards set for the assessment task, and can either be part of in–course assessment, or assessment at the end of a course or module.

Boud (2000:160) says that assessment activities 'have to encompass formative assessment for learning and summative for certification'. We should move away from providing merely summative assessments of our students' learning, especially when these occur at the end of units of study, because students will not be able to use these to improve in their learning. Summative and formative are not types of assessment but rather purposes to which assessment are put.

A case study from Oxford Brookes University shows how the use of formative assessment can play a part in reducing the failure rate and increasing the performace of students.

Characteristics of a 'Good' Assessment Programme

Who is assessment for? There are many stakeholders in assessment. Students are obviously stakeholders, because they submit work for feedback and grades. Teachers and tutors wish to monitor student progress and can receive feedback about the quality of the students' learning experiences from their assessment performance. Employers want to know how well students have done, as do professional bodies. In addition, the wider community wants to know that education is providing value for money and they gauge this through the assessment results. Institutions are subject to Quality Assurance and assessment is regarded as a major component in this.

The key factor in determining whether an assessment programme is good depends on whether the assessment tasks are relevant to the aims and intended learning outcomes for the course, not forgetting the attitudes and skills that are to be tested. Brown *et al.*, (1997) declare some common weaknesses in assessment systems, providing a checklist against which assessment programmes can be verified.

1. Overload of students and staff
2. Too many assignments with the same deadline set in the department/school.
3. Insufficient time for students to complete the assignments in the time available.
4. Insufficient time for staff to mark the assignments before the next semester.
5. Inadequate or superficial feedback provided to students.

6. Wide variations in assessment demands of different modules.
7. Wide variations in marking across modules.
8. Wide variations in marking within a module.
9. Wide variations in marking by demonstrators.
10. Fuzzy or non–existent criteria.
11. Undue precision and specificity of marking schemes or criteria.
12. Students do not know what is expected of them.
13. Students do not know what is a good or bad assignment/project.
14. Assessment viewed by some departments/schools as an extra rather than a recognised use of staff time.
15. Project supervision seen as an extra or the real time involved is not recognised.

Students gauge the requirements of a course from the assessment that they are expected to do to attain that course:

'Assessment sends messages about the standard and amount of work required, and what aspects of the syllabus are most important. Too much assessed work leads to superficial approaches; clear indications of priorities in what has to be learned, provide fertile ground for deep approaches' (Ramsden, 1992:187–8).

Providing a range of assessment methods for students can make administration of assessment programmes more difficult, as it can prove difficult to combine marks from a number of different tasks, however, a range of methods is important to students, as this can accommodate the range of students learning preferences (Weavers, 2003).

Ideally any assessment should be valid, reliable, practicable, and fair and useful to the student. It should demonstrate whether and to what level students have met the intended learning outcome(s) of the course or programme.

Validity: The degree to which you are able to measure what you think you are measuring, and may require assessment within real life, and variable settings (Ashcroft and Palacio, 1996).

Reliability: the degree to which the scores of every individual are consistent over repeated applications of a measurement procedure and hence are dependable, and repeatable; the degree to which scores are free of errors of measurement.

Assessment should have a clear purpose, it should be clear what is being assessed and how the judgements are reached. It should enable the learner to review their progress in the light of the assessment criteria, and then plan for their further learning. It should also be subject to quality assurance procedures and allow the assessor to review teaching effectiveness. An outcome will be clear records of attainment which will be useful to all the stakeholders in assessment, as these will be required for the awarding of degrees or the achievement of professional certification.

The assessment should be based upon certain standards and these should be well articulated to form a reliable basis for determination of whether or not a student has achieved the learning outcomes of the educational programme. Simple ways of improving reliability in assessment include:

- the use of more than one method to assess achievement (Rovai, 2000, Sim *et al.*, 2004)
- setting more, albeit smaller assessment items (Gipps, 2003, Weavers, 2003)
- and the use of more than one assessor

When creating effective assessments a tutor should look into the nature of effective assessment. Brown and Knight (1994) suggest some requirements for creating effective assessment methodologies:

- give full feedback, related to the criteria that has been established for the assessment in the first place,
- identify with the student points for development,
- generate and share criteria which blend departmental requirements with student priorities,
- be supportive, foster intrinsic motivation, preferably by being interesting and enthusiastic,
- grade rapidly – effective feedback is swift feedback.

All forms of assessment are problematic to some degree. Continuously assessed work such as essays, projects and reports can be open to claims of plagiarism (Carroll, 2002 and http://www.brookes.ac.uk/services/ocsd/4_resource/plagiarism.htm) and can be very time consuming to mark. Exams can encourage surface learning, guess work, and probably don't lead to effective long–term learning (Ashcroft and Palacio, 1996) as they generally do not include any elements of feedback, and are not returned to by students.

What Do We Mean by e–Assessment?

e–Assessment is often seen as providing a partial solution to providing assessment for increasing numbers of students and declining staff to student ratios (Sim *et al.*, 2004). In addition, students may experience cognitive conflict because they are generally expected to word process essays and engage in online tasks but use pens in examination halls (Brown *et al,* 1997) such that we are training them in one system and testing them in another. Gipps (2003: 26) reasons that:

'If teaching and its associated resources become electronic, then assessment too will need to take that route, to ensure alignment between the modes of teaching and assessment.'

When e–assessment is mentioned people often have certain assumptions, that:

1. The assessment will be objective.
2. The assessment will be limited in the type of question to multiple choice questions (MCQs).
3. The assessment can only test recall or low level learning outcomes.

We will address each of these in turn.

Assumption 1: Objective assessment

- There is no such thing. Subjective judgment is always involved – when an educator creates a test they do so with their internal biases on the type and nature of material. When the limit of the assessment and the type and nature of the 'correct' answers are preset, the educator introduces their own judgement and bias into the system from the start.
- However the extent of bias can be reduced because in e–assessment, the judgments made are only made based upon the original criteria and not on 'human introduced error' (e.g. marking at 2 a.m.) so that a second level of error is not introduced. In addition, levels of correctness can be programmed into the system that can score partially correct marks in a more consistent manner.

Assumption 2: Limited Types of assessment tools

Communication and information technologies have been incorporated into assessment in further and higher education in a number of ways. For example:

- the analysis of student postings to discussion boards (Rovai, 2000; MacDonald and Twining, 2002)as discussed in the 'Computer Mediated Conferencing' section of this infoKit
- automated essay marking (Christie 1998,1999) and two commercial packages http://www.intelligentassessment.com/ and http://www.ets.org/research/erater.html)

- the delivery of exam papers (<u>Pass–IT project</u> and <u>Technologies for Online Interoperable Assessment (TOIA)</u>
- use of modified essay questions (MEQs) particularly in Medicine
- in helping to provide feedback to students on written work (e.g. <u>Electronic Feedback System</u> developed by Phil Denton from Liverpool John Moores University)
- use of Personal Response Systems to assess students' undertaning during a class(e.g. <u>JISC funded PRS project at Strathclyde University</u> by D.Nichol and J.Boyle.

Students can be asked to interact with simulations and submit the results and analysis of these results for assessment. Students can create webpages, these can be assessed by their peers or be validated electronically. Computers can create sounds which students have to replicate, for example in music, or in languages, and the computer can gauge the amount of similarity.

Some of these assessment activities can be offline assessment of online delivery and could include the assessment of a presentation on the web or powerpoint for example, as well as the assessment of online skills, including how students use particular software packages such as databases and spreadsheets.

However, most would accept that e–assessment comes into its own with objective testing (examples can be found from the CAA conferences in 2004 and 2005 available from <u>http://www.caaconference.com</u>), where computers do the marking.

While many of the available systems, particularly those which come as 'quiz' facilities with VLEs or are available free of charge are indeed limited, others include a large variety of question types. [see types of questions and choosing a software package]

Assumption 3: Computerised assessment can only test low level learning outcomes

Even straightforward multiple–choice questions can, if carefully constructed, test higher order skills. The issue here isn't so much that of technology as creativity. For instance if you are looking at the application of knowledge (a higher order skill according to Blooms taxonomy <u>http://www.le.ac.uk/cc/rjm1/etutor/resources/learningtheories/bloom.html</u>) it is possible to create an ordering question, which might appear easy to a student who has comprehended the topic whereas one who has not will struggle.

What Do We Need to Consider?

Advantages of e–Assessment

Some of the advantages of e–assessment that you might want to consider are:

- immediate feedback to students,
- allows rehearsal and revision,
- immediate feedback to staff,
- allows evaluation of a course's strengths and weaknesses,
- can be linked to other computer–based or online materials.

These are characteristics of good assessment technique and have links to a strong well evaluated pedagogy, as well as providing support for both staff and students – and of course, online assessment has all the other advantages of remote access and choice of time and place of assessment (although the latter may be limited for summative assessments that require security).

When looking to use e–assessment we can find that grading swiftly is one of its strongest points. Test feedback can be on a question by question basis and with the use of a 'knowledge tracking system' students can follow their progression and self determine their weaknesses (and strengths).

This aspect of tracking progression combined with careful nurturing of student expectations can assist in developing students as autonomous learners.

Will e–assessment actually save time for the average lecturer? If you look at it in the short term then almost certainly not. Look at this as a longer term investment. While in the first year you won't save any time; in the second and third years of that material's life span you can save considerable development and support time. Another issue is that as well as the question setting, feedback also has to be created before students sit the test, something that tutors will have to factor into their busy schedules.

Sim *et al* (2004: 217) 'The emotional and subjectivity issues that are evident in human centred marking may be removed via automatic marking offered by CAA software'.

Concerns and Issues Associated with e–Assessment

Some of the concerns associated with e–assessment that you might want to consider are:

Time Required

One of the claims most often made for e–assessment is that it saves time. This is perfectly true at the point of delivery – it is possible to process the results for a summative assessment for a class of, for example around 700 students within a couple of hours of the last one logging off, including error checking and results analysis. This has to be balanced against the time, and skill, needed to create the assessment in the first place. This may not be so important for formative assessments which can be discussed with students later (and where failings may actually be of educational interest) but it is obviously vital that an end–of–course assessment should be reliable. The time and expertise for this cannot be underestimated, nor the need for 'shredding and vetting' by colleagues. There are times when an open–ended exercise (whether we call it an essay, project or report) may be more suitable for your purposes. There is of course no reason why this cannot be delivered online, with students uploading written materials into virtual learning environments to be assessed off–line.

Misleading Clues

There is a danger that by picking out particular areas (either deliberately or inadvertently), the quizzes could send misleading clues to students about what is and isn't important. This is exacerbated by the students' tendency to be very strategic and exam–focused when considering how best to spend their study time.' Clarke *et al* 2004: 253.

Equity and Diversity

Equity and diversity – when computers are involved in the assessment process, there are equity issues for different student groups relating to language status and gender and issues around computer anxiety and exam equivalence. Brosnan (1999: 48–49) suggests that: 'computer anxiety can lead to simplification of conceptual reasoning, heightened polarisation and extremity of judgement and pre–emption of attention, working memory and processing resources. Individuals high in computer anxiety will therefore under–perform in computer–based versions of assessment'. Brosnan (1999) asserts that even those who are using computers effectively will still exhibit computer anxiety and he contends that female students exhibit higher levels of anxiety, and so poorer levels of performance. Ricketts and Wilks (2002) suggest that student performance in tests should be monitored to ensure fairness and consistency when there are any changes in delivery, whether this is a change to CAA or changes in the way that the CAA is presented.

Issues of Equivalence

The issues of equivalence between different forms of assessment are highlighted by Clariana and Wallace (2002) who assert that you cannot necessarily expect that equivalent measures of student learning will be produced from computer–based and paper–based tests, even if you use the same questions. They assume that the 'test mode effect' will diminish when students become as familiar with the medium of the computer as they are with paper, for assessment, and that computer familiarity might be an issue for some groups of students. McDonald (2002) concurs, expressing the belief that inconsistent findings relating to student scores in computer–based and paper–based tests often result from different levels of exposure to changing technologies. It is probably fair to observe generally that students perform differently under different conditions of assessment, and that innovations in CAA simply introduce a new range of variants on this construct theme.

Effect of Using Full Range of Marks

Sim *et al* (2004: 217) 'CAA, like mathematics and some science subjects, also tends to use the full range of marks therefore the trend towards a higher proportion of First Class Degrees may occur in other subject domains adopting this technique in the future.'

Attracts Greater Scrutiny

While problems with objective testing can occur whether the tests are offered on paper or online, it is the online testing that tends to attract greater scrutiny. Don Mackenzie in Brown et al. (1997: 217) contends that CAA has produced quality and efficiency gains in assessment, but for many there have been marginally lower pass rates than for essay–type assessments. He suggests that this is because there is a larger spread of marks (typically a standard deviation of 15 per cent with a mean of 50 per cent).

Design of Questions

Problems in the use of computers for multiple choice questions could derive from the design of the questions and the skills of the designer (Mackenzie, 2003), rather than from the software or the use of the computer per se, or it could be that some tutors may be reluctant to relinquish traditional modes of assessment (Mackenzie, 2003).

Disparity

Research by Clariana and Wallace (2002) has shown that the use of CAA has a positive impact on the test scores of high attaining pupils, when compared to those from paper–based tests, because they assert that higher–attaining students more quickly adapt to new assessment approaches. Noyes *et al.* (2004) suggest that lower–performing individuals will be disadvantaged when CAA is used because they assert that a greater workload and additional effort is required to complete a computer–based test.

Change in Working Practices

The savings in time that might be produced by the automated marking in CAA are instead shifted to the design and construction of the assessment activity, (including the level and amount of feedback to be given). Brown et al. (1997) see this as a profound change in working practices for academics. There is also the issue of defining requisite technical skills for students undertaking CAA such as, who should be involved in that training, and when should it take place, especially in the context of overloaded curricula (Weller, 2002). Macdonald and Twining (2002) found that their students only became competent in the use of a particular piece of software while they were completing an assignment that required its use.

Plagiarism

Plagiarism is a concern for many thinking of using CAA (Weller, 2002), but Rovai (2000) and Carroll (2002) suggest that assessment design is the key to deterring plagiarism. O'Hare and Mackenzie (2004) assert that there is a level of imagination and rigour required for the design of assessment online compared to that for more traditional forms of assessment. Weller (2002) suggests that the use of portfolios can help to counter plagiarism, as this places less reliance on single assessment items. The JISC funded Plagiarism Advisory Service gives advice and guidance on plagiarism prevention.

Off−Campus Assessment

Computer software for CAA allows for questions to be presented to students in different orders, with distracters in different orders, and if sufficient questions have been compiled of sufficient integrity then they can sit different tests. All of this allows for students to sit in adjoining desks in computer laboratories that will at other times be used for learning activities. This is fairly straightforward for on−campus students, but could be more problematic for students taking courses at a distance. However, Rovai (2000) suggests that this difficulty can be overcome by using 'proctored testing' where academics arrange for students to sit online assessments under test conditions in alternative venues.

Reasons for Using e−Assessment

As e−assessment is part of the e−learning strategy of the DfES, it should be part of our College or University Learning and Teaching Strategy. The DfES also recognises the importance of aligning assessment to the needs of pedagogy and subjects in terms of e−assessment methods for specific subjects.

e−Assessment can act as a catalyst for rethinking the whole curriculum, as well as current assessment systems (Ridgway &McCuster, 2003). The Qualifications and Curriculum Authority (QCA)in its 'The Basic &Key Skills e−assessment Experience' report is quite clear that e−assessment must not simply invent new technologies which recycle our current, perhaps ineffective, practices.

Bull and McKenna (2004: page 3) suggest a number of reasons that academics may wish to use CAA:

1. To increase the frequency of assessment, thereby:
    ◆ motivating students to learn,
    ◆ encouraging students to practice skills.
2. To broaden the range of knowledge assessed.
3. To increase feedback to students and lecturers.
4. To extend the range of assessment methods.
5. To increase objectivity and consistency.
6. To decrease marking loads.
7. To aid administrative efficiency.

Nichol and Macfarlane−Dick (2005; 2004) identified from the research literature seven principles of good feedback practice that could support learner self−regulation – active control by students of some aspects of their own learning. and Nichol and Milligan (in press) have taken this further to show how e−assessment can support these seven principles by providing:

• timely feedback,
• opportunities for re−assessment and continuous formative assessment to encourage students' self−esteem,

- statistics to help tutors evaluate the effectiveness of the assessment – questions answerred very poorly can be re–examined in case poorly specified,
- timely information for tutors to be able to help shape teaching.

They also discuss how each of the principles might be implemented in blended learning contexts.

e–Assessment can be used diagnostically at any point in an academic course. It can be used to show how much students already know when they arrive, or can be used as a means of determining how effective the teaching is during a course, and can be related to content and skills. Drew *et al* (2002) provide an example of a system of 'Skill Check Questionnaires' which are being used at Sheffield Hallam University providing diagnostic testing for key skills. The computer–based system was private to students, which encouraged them to be honest, which then directed the students to the specific support that they required.

Instant and tailored feedback can be provided for students from e–assessment, providing students with more, and more timely feedback than can be provided by tutors. Boud (1995) emphasises the importance of timely feedback for supporting student learning.

Brosnan (1999) has provided evidence of computer anxiety in students affecting exam performance, and Ricketts and Wilks (2002) concur with some students feeling disadvantaged by online examinations. However, Ricketts and Wilks (2002) have also found that some dyslexic students prefer online examinations, finding them less stressful than other examination types.

e–Assessment provides the potential for students to be provided with more complex scenarios than can be provided in paper–based tests including computer simulations, images and sounds, with which students can interact (James, *et al.*, 2002). Interactive assessment activities are time–intensive to produce in much the same way as interactive learning activities are, and make additional demands on institutional resources and support.

Summative e–assessment requires a larger degree of institutional support than formative or diagnostic assessment. The latter can generally be managed by individual academics or departments, whereas there are additional considerations for summative assessment, which are considered within the sections on policy considerations and choosing the software.

Policy Considerations: What is your institution's position?

You will need institutional support to determine the optimum system for your institution, because otherwise you will have issues related to technical and adminstrative support and upgrading, and staff and students may have to learn to use a number of different assessment engines and user interfaces. You will need to check on what your institution allows you to do in terms of policy implications and institutional impact.

Sim *et al* (2004: 222) '*The perceived benefits of CAA of freeing lecturers' time can be illusive if no institutional strategy or support is offered, successful implementation may be left to chance and CAA may be developed in an anarchic fashion. In order to utilize the features within software packages staff training and development is necessary and this may not be feasible without institutional support*'.

If you are utilising, or proposing to utilise, e–assessment across major sectors your institution will need to complete an end–to–end analysis of the systems you will use and the experience that the students will receive. E–assessment is now a proven technology, and therefore staff and student users expect secure, reliable and flexible systems that are tailored to their learning and assessment requirements. For those managing these systems, these expectations have to be delivered with the added constraints that systems should be cost effective and sustainable.

Components of e–Assessment

e–Assessment requires the provision of assessment software, server &network infrastructure, student PCs, administrative staff support and questions. To create questions and assessments across disciplines a staff education programme is normally required to ensure the staff have an accurate knowledge of the capabilities and limitations of online assessment. Publishing written policy and procedures ensure that all the assessment stakeholders – students, teaching staff, administrative staff, specialist support staff, external examiners and academic &quality assurance managers, can be clear about the process and can have confidence in the operation of e–assessment.

Policy and Procedures, Roles and Responsibilities

Designing, building and operating an institutional–wide e–assessment system is not small undertaking. It is likely to involve specialists in e–Learning, educational systems, information systems, academic departments, student IT support, quality assurance specialists and academic management. For people from these groups to work together, agreeing roles and responsibilities within policy and procedure documentation, ensures they can each apply their professional strengths and contribute to a successful institutional e–assessment system.

The University of Dundee Policy and Procedures for Computer–Aided Assessment have been in use since October 2002, although they are regularly revised. These procedures are available for other to utilise as appropriate (with due acknowledgement), but it is recognised that they are also limited in that they are written for the specific software, hardware and staffing configurations at Dundee. Another key resource is the Bristish Standard BS7988:2000 – Code of practice for the use of information technology (IT) in the delivery of assessments. Specific areas of importance are highlighted below.

IT suites and the presentation of invigilated assessments

The physical environment for taking assessments needs to be familiar to students and spacious. Screens should not overlook each other, and frequently a back–to–back arrangement is space efficient and secure. Questions may be presented to students in a random order and also even in a structured assessment, students will move at a different pace through the assessment. Some institutions favour individual screened corrals, but these carry the additional risk of participants being shielded from invigilators observation. The University of Dundee, found the use of screens unnecessary. Colchester Institute report in Case Studies of MLEs in Further Education that 'plans for the introduction of e–assessment are posing real challenges to the institution in estate terms. The computer suites are designed in 'daisy–pod' style to facilitate interaction, not as the face–on single units required by the examination bodies'. This example demonstrates that flexibility of learning space is an important consideration.

Invigilation of examinations should be carried out following the institutions conventional examination procedures.

Open Assessments and Personal Authentication

A common topic for debate with e–assessment is how a tutor can be sure that the student completing the assessment is the correct student on their own. The simple truth is that unless the assessment is delivered under invigilated conditions using secure systems, then you cannot be sure. However, many conventional assessments are taken in an open environment – in course essays for instance, and for low stakes assessments, it is often appropriate to adopt an open strategy requiring the completion of a variety of small assessments. Subsequent invigilated assessments provided a check on the open assessments.

Advanced mechanisms for personal identification are being tested (biometric keyboard use patterns, iris scans etc) and some have proposed the use of webcams to monitor participants. Dundee found the development of such approaches unnecessary.

Reliability of assessments

All institutions will aim for the highest level of reliability for online assessments. This is crucial, especially in the early days of running a large online assessment programme. If the staff and students experience an unreliable system, and then experience it again, they will loose confidence. At the University of Dundee they take particular care when a department is running a high–stakes summative e–assessment for the first time. Learning technologists and IT specialists visit the assessment for the first 15 minutes to ensure the assessment proceeds smoothly.

No system can be 100% reliable. Backup or redundant procedures should be in place to cope with the unexpected. At Dundee the backup procedures involve an entirely independent e–assessment system that is available for deployment within 15 minutes to a group scheduled to receive an assessment. Some institutions have paper copies of assessments available, but often this is not possible because of the expanding use of multimedia, innovative and flexible questions in assessments that cannot be replicated on paper.

VLE or Specialist e–Assessment systems

Software for the delivery of e–assessments is often contained within commercial VLEs or specialist e–assessment providers. An institution's VLE has the advantage that it is supported, built on a secure servers and familiar to staff and students. Specialist e–assessment systems require additional servers and support and additional tuition for students and staff, but normally offer more flexible and powerful systems.

Experiences of different institutions are varied. A major concern in the use of VLEs for summative e–assessment is that they are large, complex software and often simultaneously accessed by many individuals and groups of students from many different locations. VLE software (in 2005) is still relatively young and some systems are not as reliable or robust as may be reasonably expected. Successful use of VLEs for summative e–assessment delivery can involve timing assessments to the evening or weekends when normal usage is reduced, and specialists are on hand to monitor the systems. At the University of Dundee, the use of specialist e–assessment software (Questionmark Perception) has proved robust, flexible and cost effective. Undoubtably more complex in system design, we find that the ability to run assessments at any time using robust and redundant systems justifies the additional setup and staff development costs.

Load testing on servers

Developers of e–assessment systems are aware that their systems will be utilised with large numbers of students in a complex networked environment. Designing and deploying systems to meet institutional needs is a complex and specialist task. Load testing is frequently difficult as mimicking the activities of (say) 200 students completing an interactive assessment over a busy network cannot be easily replicated. One important tip is that even the most highly specified system, composed of a load balanced array of the latest servers, may struggle to meet the demand of simultaneous access by large numbers of students. It is good practice to make instructors and students aware of this limitation. At Dundee, assessments are started (and therefore finished) within a window of 2 to 5 minutes that effectively spreads the server load and all participants receive fast responses from the server.

Emerging e–Assessment Techniques

Alongside what is now traditional e–assessment used for the delivery of online tests and examinations, are the developing areas of e–assessment which include online submission and marking, plagiarism detection, ePortfolio assessment and assessment of contributions to asynchronous and synchronous discussions. Developing policies in these fields will involve consultation with academic staff as they evaluate these assessment mechanisms and policies should also take in to account accepted practice from published leaders in these fields.

A convention that Dundee aims to develop is that the online submission of student work is acknowledged immediately after the deadline, and that care is taken to ensure that the assessment outline are clear and the mechanism and location for delivery well understood.

A case study of e–assessement at the University of Dundee is available.

Practical Considerations

Practical considerations include having sufficient and sufficiently robust information technology systems available for students who are required to engage in e–assessment, as well as ensuring that students have sufficient access to the facilities. For security reasons it is important that summative assessments are not placed on part of the IT infrastructure where students can find them, and that results are posted securely. However, security is less of an issue when the assessment is purely formative.

Sim et al (2004) provide information about the British standard code of practice governing the use of information technology in the delivery of assessments (BS7988(2002) to which institutions should adhere. This code of practice for the use of information technology in the delivery of assessment recommends that students take a break after 1.5 hours. This has implications for invigilation of examinations.

Choosing a Software Package

The main considerations here will be in terms of cost efficacy, whether proprietary software has to be bought or whether software is available free or bundled with an institutional VLE/MLE. The amount of time required for developing the assessment items, including the programming that might be required will be an important consideration, as will the level of IT support required from the institution and locally within departments. the considerations, cost efficiency, time for development. A key issue will revolve around whether the assessment is to be formative, which can be conducted locally and probably requiring little in the way of backing, or whether it is summative requiring a large degree of institutional support to maintain quality assurance issues.

Security

Many versions of computer software for e–assessment allow for questions to be presented to students in a different order, and can provide the key and distracters in varying. If sufficient questions have been compiled of sufficient integrity then they can sit different tests, and for mathematically underpinned assessments, the programmes also provide the means to create, at little cost, variants of the same questions. All of this allows for students to sit in adjoining desks in computer laboratories that will at other times be used for other learning activities. This is fairly straightforward for on–campus students, but could be more problematic for students taking courses at a distance. However, Rovai (2000) suggests that this difficulty can be overcome by using 'proctored testing' where academics arrange for students to sit online assessments under test conditions in alternative venues. However, Sim et al (2004: 224) believe that:

'If security measures are in place there is no evidence to suggest that the integrity of the examination is more compromised by delivery over the Internet than by paper.'

Accessibility

When creating e–assessments that are to be as inclusive as possible, there are a variety of things to bear in mind:

1. ***What pedagogical issues are there to consider?***

   1. Design for All (where this can be achieved while not reducing the overall effectiveness of the assessment for other students)

      ◊ Selection of mode of assessment and subsequent assessment design with accessibility in mind
      ◊ Should enrich the experience for all users, not increase accessibility for some by reducing experience for others
      ◊ Is not 'the lowest common denominator'
      ◊ Is not bland and anti–technology
      ◊ Basic principles are easy to apply
      ◊ e–Assessment (as an alternative to paper–based assessments) can benefit those with mobility difficulties, illnesses, mental health issues and so on

   2. Think about alternatives of equivalence

      ◊ Not a case of when all else fails, thinking about possible alternatives will improve the whole assessment process
      ◊ Encourages lateral thinking about choice of mode of assessment and content
      ◊ Not 'only' for students with disabilities – all students have preferences, strengths and weaknesses
      ◊ Can be costly and time consuming, particularly in the short term, but can reap dividends later in terms of marking time and administration year after year
      ◊ Make use of student input to the design – nobody knows better what is achievable, what will aid the learning process and what will best improve the assessment experience

   3. Validity

      ◊ Alternatives and modifications need to be authorised by the validating body – the exam board, university board or whatever. The degree of difficulty with this varies tremendously between boards, but most will be interested in facilitating the assessment process for the widest possible cohort of students

2. ***What are the common mistakes?***
   Often we inadvertently assess things which are not a part of our stated outcomes, but instead are a function of the students' physical or social capabilities – we need to determine our policy on each of these prior to setting the assessment, and to communicate this policy to the students in time for them to make representations if they have issues with any of them:

   ♦ Speed of writing or typing are not normally an outcome of an assessment (except keyboard skills tests etc) – imagine a student with a very slow typing speed in an essay–style exam – what would be the best course of action to allow them to exhibit their knowledge? Sometimes extra time would be appropriate, sometimes the use of another input method such as voice or a scribe, and sometimes a completely different but equivalent mode of assessment
   ♦ We should determine prior to each assessment whether or not we are adding or deducting marks due to spelling and grammar. If this is the case, we need to

consider how we will view the work of dyslexic or blind students (for example, a blind student using voice input software will not know whether the programme has inserted the word 'court' or 'caught' – should we penalise a student for this?

♦ Are we assessing manual dexterity? Should a student be marked down because they cannot pour liquid from one flask into another, despite knowing everything about the chemical reaction involved?

♦ Are we assessing visual acuity? Auditory perception? Short–term memory? Physical endurance?

3. ***Where can I get further assistance on this subject?***

♦ <u>The Teachability project</u> is a resource booklet describing a step–by–step audit process towards making a course or module more inclusive for disabled students, including of course the consideration of assessment.

♦ <u>The Web Content Accessibility Guidelines (WCAG)</u> published by the Web Accessibility Initiative (WAI) are a checklist for making web pages accessible (a full debate about the usefulness of 'pure' accessibility as against usability can also be found on the WAI website). Although they are oriented towards web pages they are extremely pertinent to any online materials and therefore to e–assessment.

4. ***Is there a simple 'checklist' of considerations I should make?***
Not a definitive one, as every student and cohort of students will have different needs and every assessment and institution will necessarily be different. However, there are some considerations which may be considered to affect a greater proportion of students and therefore may merit the most immediate attention (NB This is not an excuse to forget about any other needs your students may have!)

♦ Navigation – can your assessment be completed without using a mouse? Can check boxes be checked using the keyboard alone? Are drop–down menus accessible using only the keyboard?

♦ Keyboard dexterity – do users need to be able to press more than one key simultaneously? If text is entered incorrectly or by accident, can the user backtrack and correct their actions? Are there alternatives for users who cannot access drag–and–drop features (NB there is often no need to remove drag–and–drop exercises because some students cannot access them, it is often far more beneficial to create alternatives)

♦ Audio content – are alternatives, transcripts or subtitles available?

♦ Language – is the assessment, particularly the instructions and navigation, clear and precise? Have you used 'trick' multiple–choice questions which use similarly spelled words as spoilers, which may confuse dyslexic students?

♦ Have you examined your images and graphics at high magnification to check they are still legible? Images of text in particular become pixelated at increased magnification, meaning they become illegible to vision impaired students.

♦ Does the layout make sense to screen reader users? Screen readers progress linearly through a page only if it has been coded to allow this, otherwise they dart about and can be very confusing. Tables in particular can be very difficult for screen reader users to cope with if badly designed. Read aloud your table row by row and see if the question is still achievable without having the table available visually. Do images and graphics have text tags or alternatives for students that cannot see them?

♦ Be careful with use of colour – do not use colour alone to convey meaning. Be sparing with bright colours – think of magnifier users who may view your pretty little lime green heading sixteen times larger than you imagined!

♦ Is it possible to give the student control over text size, font, and colour? If not, have you selected options that will be available to as many students as possible (sans–serif font, minimum 12 point font size, 1.5 line spacing etc)

♦ If you are not sure whether you have succeeded in making your assessment more inclusive and accessible, test it with students – but if you are seeking disabled

students to test your assessment with, please approach them via the student support staff, as they invariably get pulled in a number of directions at once!

Creating and Using e–Assessment

Being aware of the institutional and practical considerations is of course important, but the key issue to be resolved is that of creating the assessment tasks. As has already been mentioned, creating reliable and valid questions is a skilled task, and a lot more work is required from staff before the assessment is presented to students, especially when feedback is to be included for all of the incorrect answers in an objective test.

In common with all assessment, for e–assessment the assessment tasks should be aligned with the intended learning outcomes, so that students are able to display the extent to which they have met the outcomes for the course. Using the verbs from Bloom's taxonomy can help in the design of questions, using these to define the leavel of learning that are being tested.

Not all questions need to be written from scratch, and in the first instance it may prove useful to use tried and tested questions from other sources. Banks of objective questions are often associated with textbooks in further and higher education, and these can be used with permission from the academic publishers. Some of the subject centres of the Higher Education Academy (http://www.heacademy.ac.uk/474.htm) are developing question banks, using the expertise of academic colleagues in departments across the UK. Examples of question banks are provided in the links section.

Of course departments can choose to write questions themselves. Zakrzewski and Steven (2003) suggest that staff should create their own question banks and that they should generate an extra 10% of questions each year to allow for some questions to be removed or used less frequently. This is especially important as curricula change and develop. Gipps (2003) suggest that for cost effectiveness, staff should create question banks because: 'the true costs involved mean that CAA is only really feasible for items that can be re–used' Gipps (2003: 27). Questions should be created collaboratively and tested before being made available to students to alleviate the possibility of writing overly simple questions or questions with subtle nuances that students may inadvertently pick up (Clarke et al., 2004)

The COLA project (Sclater and MacDonald, 2004) developed a Word template to help staff with item creation for multiple choice tests, including metadata to describe the item, developing the methodology that had been used successfully by the e3an project.

McAlpine (2004) has outlined a methodology for determining whether tests that are currently administered on paper could be administered online. Initial results show that the methodology has potential for discriminating between tests that can be easily administered online, and assessment items for which more research is required in terms of the use of "emerging technologies".

Assessment Questions

Leaving aside the other uses of IT in assessment, and concentrating instead on assessments for which the computer marks the student responses, there are a number of types of objective questions which can be used. There are three main types which include students making a choice (single or multiple) inputting text (words, numbers or formulae) or manipulating information on the screen (marking relevant areas, moving items or drawing graphs or diagrams).

Many tools exist, some of which are available in the resources section.

There are a number of question types for objective test type items, examples of which can be found for many subjects within the question banks listed in the links section. This list is taken from the

CAA Centre.

Multiple choice questions (MCQs) are the traditional 'choose one from a list' of possible answers.

True/False questions require a student to assess whether a statement is true or not.

Assertion–Reason questions combine elements of MCQ and true–false questions.

Multiple response questions (MRQs) are similar to MCQs, but involve the selection of more than one answer from a list.

Graphical hotspot questions involve selecting an area(s) of the screen, by moving a marker to the required position. Advanced types of hotspot questions include labelling and building questions.

Text/Numerical questions involve the input of text or numbers at the keyboard.

Matching questions involve linking items in one list to items in a second list.

Sore finger questions have been used in language teaching and computer programming, where one word, code or phrase is out of keeping with the rest of a passage (for example a word given in the wrong tense). It could be presented as a 'hot spot' or text input type of question.

Ranking questions require the student to relate items in a column to one another and can be used to test the knowledge of sequences, order of events, level of gradation.

Sequencing questions require the student to position text or graphic objects in a given sequence. These are particularly good for testing methodology.

Field simulation questions offer simulations of real problems or exercises.

Other question types require students to identify and/or manipulate images. Students may be asked to plot a graph, complete a matrix, draw a line or build up an image using parts provided.

For descriptions of each question type and hints on how to improve your skills in developing these types of questions visit the CAA Blueprint, especially appendix A.

There are some particular terms associated with creating objective tests questions that you will need to be aware of:

- Item – the units making an objective test. These may or may not be phrased directly as questions
- Stem – introductory question or incomplete statement with which an item begins
- Response, option – suggested answers to or completion of, the stem
- Key – the correct response
- Distracter – incorrect response (s)

Tips for writing questions (thanks to Bob Matthew of The University of Glasgow for these)

Stems should:

- be as clear and concise as possible
- not contain ambiguities
- contain most of the wording
- not be negative if possible

Reponses should:

- be equally likely
- approximately the same length
- internally consistent
- in relative order

The key should be:

- in varied position
- correct

Avoid:

- double negatives
- 'none' or 'all' of the above
- clues

For further information about writing objective test questions the CAA Centre has produced a helpful document 'Designing effective objective test questions: an introductory workshop' which can be found at: CAA Centre.

Analysis and Quality Assurance

McKenna and Bull (2000) have provided some important quality assurance recommendations:

- Integrate the scheduling of computer–based tests into the timetabling for end–of–module examinations.
- Ensure the proper moderation of CAA examinations, as for traditional examinations.
- Consider appointing an additional external examiner with expertise in the construction and presentation of CAA.
- Incorporate feedback mechanisms which guide academic staff in the improvement of tests and systems.
- Ensure that staff have been offered and have attended the relevant staff development sessions.
- Develop a procedure which defines and checks that question banks have been supplemented with a percentage of new questions each year.
- Verify that piloting procedures and question analysis (to ensure reliability and validity) have been undertaken.
- Establish an upper limit on the amount of CAA examination per module. (For example, in order to encourage lecturers to offer a balanced assessment profile to students, the use of CAA might be capped at 40 per cent of the total module mark.)
- Agree standards (in terms of screen design, instructions within test, function of buttons) to guarantee consistency in presentation of tests thereby minimising student confusion.
- Integrate a programme of evaluation covering all aspects of the system

Following e–assessment it is important to determine whether the questions have been effective at discriminating between students and determining whether or not the questions have assessed what they were professed to be assessing. There are two measures that should be calculated for objective test questions, and these are often provided automatically by e–assessment software, the facility value and discrimination. The facility value is the fraction of students making a correct response to an item. Discrimination is how well an item discriminates between able and less able students, as measured by performance on the whole test.

Analysis of these statistics is important because alongside the examination results they can help academics determine which questions should be retained, and which should be altered or removed (Zakrzewski and Steven, 2003).

Feedback on the conduct of the test should be sought from the students and peers as well as from the external examiner, and the student performances should be evaluated to provide feedback on the teaching that has been occurring in the course. Johnstone (2003) suggests that a useful indicator is to look at changes in facility values. If the values are lower, then it probably relates to student ability, if the values are the same except for a particular topic then it's more likely to be related to the teaching.

Zakrzewski and Steven (2003) and McKenna and Bull (2000) suggest that effective and robust quality assurance procedures are required to successfully implement e−assessment for summative purposes.

Interoperability and Reusability

Many tutors have reservations about the time required to implement e−assessment and of its educational effectiveness. To address these concerns there has been considerable focus on the development and implementation of question banks (Bull and Dalziel, 2003). A question bank is a collection of questions that can be reused and shared across a learning community (usually within a particular subject discipline). Question banks allow users to contribute their own questions as well as reusing questions created by others tutors. This approach is useful for two main reasons. Firstly, there is concern over the effectiveness of CAA. This is because many of the questions are multiple choice and it can be difficult to draft MCQs that test students' understanding of particular concepts (as opposed to questions that jog memory recall) (Bull and McKenna, 1999). This requires both creativity and hard work before a question can be used effectively to measure students' learning (Haladyna, 1997). Therefore, reusing questions that have specifically been drafted to test students' understanding can help with this problem. Secondly, by sharing and reusing questions, tutors can benefit from economies of scale (Littlejohn, 2003).

The problem is how easy is it for you to reuse a question within your own VLE? Will this involve completely rewriting and restructuring the question? Until recently, the answer to this question has been 'yes': the reuse of question tests has simply not been worth the time invested in making a question reusable. Now, however, interoperability standards have been developed and implemented to allow questions to be easily transported from one VLE or e−assessment system to another without losing their structure. One such standard is IMS Question and Test Interoperability (QTI).

QTI is a framework for transferring assessment questions, tests and results from one VLE to another (Sclater and Cross, 2003). QTI has four main areas: the assessment tasks; the results after a test has been taken; the ordering of questions and tasks; and the processing of test scores (taking into account weighting factors, etc). This is useful for three main reasons. Firstly, it allows tutors to develop online assessments using a range of question types. Secondly, tutors can share assessment information across different e−assessment systems. Thirdly students' results can be sent to centralised, institutional student record systems.

A number of these test banks have already been set up. These include:

- COLA – a bank of items and assessments for Scottish Further Education (www.coleg.org.uk/ requires membership) and
- Electronics and Electrical Engineering Assessment Network (e3an) – a bank of peer−reviewed questions being reused across UK Higher Education
- Universities Medical Assessment Partnership (UMAP) – being developed between a consortium of Medical Schools in England

COLA Project Sclater and MacDonald (2004: 208) "The assessments had to work in a range of VLEs and it was expected that the VLEs would accept these item types if they were marked up using the IMS Questions and Test Interoperability v1.2 (QTI) specification (IMS, 2002)

COLA Project Sclater and MacDonald (2004: 211) "It was necessary to develop a program to convert the items and assessments from the word templates to the QTI format. This task was carried out by the JISC–funded Technologies for Online Interoperable Assessment (TOIA) project which had the necessary expertise in QTI in collaboration with an expert group representing the four main VLEs"

The Student Perspective

Students react favourably to e–assessment. Newark and Sherwood College report that 'Staff have noticed that students are reacting increasingly positively to online, on–demand testing. The DfES key skills provision is popular with staff and students alike.'(Case Studies of Managed Learning Environments in Further Education, July 2005)

Prior Practice

Issues of prior practice are as important in CAA as they are with any assessment activities, with students being given the chance to have practice in, and learn from, assessment activities for which there will be a summative grade. Brosnan (1999) suggests that students who are confident in the use of computers as well as having had the prior experience of the test will perform better than other students. Zakrzewski and Steven (2003) stress the importance of providing student familiarity with tests prior to them taking tests for real, and students should be encouraged to become familiar with the machines and type of test questions for which they will be assessed (Hay and Bull, 2002). It is important that students are assessed on the subject matter and not on their ability to 'press buttons in the right order' (Thomas and Milligan, 2003).

Feeding Back to Students

Feedback to students is an issue of quality assurance and quality enhancement:

'Institutions should ensure that appropriate feedback is provided to students on assessed working a way that promotes learning and facilitates improvement' (QAA Code of Practice for the assurance of academic quality in higher education, section 6 May 2000)

The importance of feedback for student learning cannot be overstated (Gipps, 2003). Improving formative feedback has been shown to raise standards in assessment, a conclusion based on a review of over 250 papers from several countries by Black and Wiliam (1998). They have also shown that the giving of marks has a negative effect, as students ignore feedback comments when marks or grades are given. Clarke *et al.*, (2004) have shown that formative assessment can reinforce the content of lectures and can 'act to cement students' understanding of key concepts and ideas' (Clarke *et al.*, 2004: 259)

There is a lot of advice given about feedback, that it should be timely to be effective, that it should provide constructive information to help with learning, that it should be related to assessment criteria that are clearly understood by the students, and that it should make explicit to students what is required for high–quality work (Black et al 2002; Cowan, 2003; Sadler 1998).

Comments on student work are only useful as feedback if students can use them to help them improve in similar further work, and Black et al (2002: 10) say that: 'to be effective, feedback should cause thinking to take place'.

Feedback should be provided in a timely manner, the longer the gap between the assessment performance and the feedback on that assessment, the more that students are likely to treat the feedback as summative, as they have already moved on to new knowledge and new learning experiences.

The information provided to the student must be of use to them: 'feedback functioned formatively only if the information fed back to the learner was used by the learner in improving performance' (Black et al 2002: 14). Feedback should focus on what needs to be done, providing a motivation that improvement is possible, rather than focussing on ability, which can cause damage to self–esteem of low attainers.

For students to be able to make use of the feedback, they have to be able to understand and apply the assessment criteria to their work. Once they can do this, they should be able to start making assessments of their own performance and begin to manage their own learning. Sadler (1989) argued that assessment criteria do not in themselves help in judging performance, but that students have to be helped to interpret the criteria for any piece of assessed work. Involving students in peer assessment, where they actively engage in using the criteria is one way of helping students to understand them, and then apply them in their own work.

Analysis of the feedback that staff give to their students can reveal more about the nature of the assessment task. Black *et al.*, (2002: 9) describe this, when talking about work with teachers in schools: 'They found that some tasks were useful in revealing pupils' understandings and misunderstandings, but that others focussed mainly on conveying information'. From this analysis, the teachers decided to modify some of the activities, remove some and find others which assessed the outcomes that were intended for their pupils. In the same way staff in higher and further education can analyse their feedback comments to evaluate the assessment tasks that they construct for their students. In addition, they can use an analysis of their feedback comments to provide information about the teaching that has been happening, and the evaluation may show areas of misunderstandings across the student body that require further attention in teaching situations. An example from Liverpool John Moores University is provided in the section overview of e–assessment.

The SENLEF project (Student Enhanced Learning through Effective Feedback) has produced a very useful publication, 'Enhancing student learning though effective formative feedback' outlining why feedback is important and how staff and students can use feedback. The associated case studies provide many ideas that staff can use in improving the formative feedback to their students.

e–Assessment gives staff the capacity to provide feedback to students on their learning in a timely manner. By associating feedback with objective test questions students can readily appreciate where they have misunderstandings, and anecdotal evidence suggests that, when allowed to sit a test multiple times, they will check each of the incorrect answers to find out why they are incorrect and so enhance their learning even further.

Resources and Links

Resource Links

- Castle project
- Question mark
- SCAAN
- CAA Centre
- Medweb
- UK Centre for Legal Education
- Triads
- Bibliography on Computer Aided Assessment and Distance Learning
- IASSESS
- WebMCQ
- hot potatoes
- CETIS
- CAA Centre on objective testing

- Centre for Interactive Assessment Development at the University of Derby
- University of Dundee information about summative assessment
- Information about Computer Aided Assessment at the University of Edinburgh
- Formative assessment in Science Teaching (FAST)
- Electronic feedback system at Liverpool John Moores University
- Links to Sheffield Hallam University web site with case studies
- Links to IMS
- Online assessment and feedback (OLAAF)
- Scottish Centre for Research into Online Learning and Assessment (SCROLLA)
- TechDis/LTSN Forum for Computer–based Assessment and Accessibility
- Information at Robert Gordon University
- Information at The University of Edinburgh
- Information on assessment of English
- Project on Assessment in Scotland – using Information Technology – (PASS–IT) based at Heriot–Watt University
- Technologies for Online Interoperable Assessment (TOIA) an advanced online assessment management system available free of charge to all UK further and higher education institutions and based at the University of Strathclyde
- FERL Focus area on e–assessment

Some Question Banks [all accessed 22nd March 2004]

- Applied social surveys
- Business Education
- Economics
- Engineering
- Foreign languages
- History
- Medicine and Business Studies; Chemistry, Pharmaceutical and Biological Sciences; Mathematics
- Social Policy. Contact bob.rotheram@ntu.ac.uk (National Teaching Fellow)
- Veterinary Science
- Universities Medical Assessment Partnership

References

Angus S., McDonald (2002), The impact of individual differences on the equivalence of computer–based and paper–and–pencil educational assessments, Computer &Education, 39, 299–312.

Ashcroft, K. and Palacio, D. (1996) Researching into Assessment and Evaluation. London: Kogan Page.

Black, P. and Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education. 5 (1), 7–71.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003) Working inside the black box. Kings College London, Department of Education and Professional Studies. ISBN 1871984394 [available from http://www.nfer–nelson.co.uk/catalogue/catalogue_detail.asp?catid=101for a small charge]

Bloom, B.S., Englehart, M. D., Furst, E. J., Hill, W. H. and Krathwohl, D. R., (1956) Taxonomy of Educational Objectives: Cognitive Domain. New York: McKay.

Boud, D. (1995a). Assessment and learning: contradictory or complementary? In P. Knight (Ed.)

Assessment for Learning in Higher Education. London: Kogan Page, 35–48.

Bradford M., A view from the top. Exchange, 4, Spring 2003: 8.

Brosnan, M. (1999). Computer anxiety in students: Should computer–based assessment be used at all? In Brown, S., Race, P. &Bull, J. (1999) (Eds), Computer–assisted assessment in higher education. London: Kogan–Page.

Brown, G., Bull, J. and Pendlebury, M. (1997) Assessing student learning in higher education. London: Routledge.

Brown S and Knight P 'Assessing Learning in Higher Education' Kogan Page 1994.

Bull, J. and McKenna, C. (2004) Blueprint for Computer–assisted Assessment. London: RoutledgeFalmer.

Bull, J., and Dalziel, J. (2003), Testing Question Banks, Chapter 14, pp171–181, Reusing Online Resources: A Sustainable Approach to eLearning, Kogan Page, London, ISBN 0780749–439507

Clarke, Sophie. Lindsay, Katharine. McKenna, Chris. New, Steve – INQUIRE: a case study in evaluating the potential of online MCQ tests in a discursive subject – Volume 12, Number 3/September 2004 of ALT–J.

Carroll, J. (2002) A Handbook for Deterring Plagiarism in Higher Education. Oxford: Oxford Centre for Staff and Learning Development

Cowan, J., (2003) Assessment for Learning, giving timely and effective feedback. Exchange, Issue 4, 21–22.

Christie, James. Automated essay marking – for both style and content. In Myles Danson and Robert Sherratt, editors, Proceedings of the 3rd Computer Assisted Assessment Conference, pages 39–50. The Flexible Learning Initiative, Loughborough University, 1999. ISBN 0–9533210–3–7.

Christie, James. Computer assisted assessment of essays. In Proceedings of the 2nd Computer Assisted Assessment Conference, pages 85–89. The Flexible Learning Initiative, Loughborough University, 1998. ISBN 0–9533210–1–0.

Clariana, R. B. &Wallace, P.E. (2002). Paper–based versus computer–based assessment: Key factors associated with the test mode effect. British Journal of Educational Technology, 33 (5), 595–904.

Clarke, Sophie. Lindsay, Katharine. McKenna, Chris. New, Steve – INQUIRE: a case study in evaluating the potential of online MCQ tests in a discursive subject – Volume 12, Number 3/September 2004 of ALT–J

Davies, P. (2004) Don't write just mark: the validity of assessing student ability via their computerized peer–marking of an essay rather than their creation of an essay. ALT–J, Research in Learning Technology. 12 (3) 261–277.

Gipps, C. (2003) Should universities adopt ICT–based assessment Exchange Spring 2003. 26–27.

Haladyna, T. M (1997) Writing Test Items to Evaluate Higher Order Thinking, Allyn and Bacon, Massachusetts.

Hay, Iain and Bull, Joanne. Passing Online Exams. Journal of Geography in Higher Education, Volume 26, Number 2, 1 July 2002, 239 – 244.

James, R., McInnis, C. and Devlin, M. (2002) Assessing Learning in Australian Universities. Canberra: Australian Universities Teaching Committee. http://www.cshe.unimelb.edu.au/assessinglearning

Johnstone, A. (2003). Effective Practice in Objective Assessment. LTSN Physical Sciences Centre

Littlejohn, A. H., (2003) Reusing Online Resources: A Sustainable Approach to eLearning, Kogan Page, London, ISBN 0780749–439507

Macdonald &Twining (2002) 'Assessing activity based learning for a networked course', British Journal of Educational Technology 33 (5) 605–620

Mackenzie, D. (2003) Assessment for E–Learning: What are the Features of an Ideal E–Assessment System? CAA Conference Proceedings 2003. 185–194. Available from the conference archive section of the CAA Conference website.

McAlpine, M. (2004) Easing the transition from paper to screen: an evaluatory framework for CAA migration. ALT–J, Research in Learning Technology. 12 (3) 231–248.

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

Nichol, D.J. &Macfarlane–Dick, D. (2004). Rethinking formative assessment in HE: a theoretical model and seven principles of good feedback practice. Higher Education Academy. Available from http://www.heacademy.ac.uk/assessment/ASS051D_SENLEF_model.doc

Nichol, D.J. &Macfarlane–Dick, D. (2005). Formative assessment and self–regulated learning: A model and seven principles of good feedback practice. Accepted for publication by Studies in Higher Education (2005).

Nichol, D.J. &Milligan, C. (in press). Rethinkng technology–supported assessment in terms of the seven principles of good feedback practice. In K.Clegg and C. Bryan (eds), Innovating in Assessment, RoutledgeFarmer.

Noyes, J. M., Garland, K. J. &Robbins, E. (2004). Paper–based versus computer–based assessment – is workload another test mode effect? British Journal of Educational Technology, 35(1), 111–113.

O'Hare, D. &Mackenzie, D. M. 2004 Advances in Computer Aided Assessment. Staff and Educational Development Association Ltd, Birmingham, UK.

Ramsden, P., (1992) Learning to Teach in Higher Education. London: Routledge

Ricketts, C., and Wilks, S. J. (2002) Improving student performance through computer–based assessment: insights from recent research. Assessment and Evaluation in Higher Education. 27 (5) 475–479.

Ridgway, J. and McCuster, S. (2003) Using computers to assess new educational goals. Assessment in Education: Principles, Policy and Practice, 10(3), 309–328.

Rovai, A. P. (2000) 'Online and traditional assessments: What is the difference?' The Internet and Higher Education 3, 3, 141–151

Rowntree, D. (1987) Assessing Students – how shall we know them? London, Harper and Row

Sadler, D. R. (1998) Formative assessment: revisiting the territory. Assessment in Education. 5 (1), 77–84.

Sadler, D. R. (1989) Formative assessment and the design of instructional systems. Instructional Science. 18 119–144.

Sclater, N. and Cross, R. (2003) What is Question and Test interoperability? CETIS Briefing Paper

Sclater, N. and MacDonald, M. (2004) Putting interoperability to the test: building a large reusable assessment item bank. ALT–J, Research in Learning Technology. 12 (3) 205–213.

Sim, G., Holifield, P. and Brown, M. (2004) Implementation of computer assisted assessment: lessons from the literature. ALT–J, Research in Learning Technology. 12 (3) 216–229.

SQA (2003) Guidelines on online assessment for further education. Available online at: http://www.sqa.org.uk

Thomas, R. C. and Milligan, C. (2003). Online Assessment of Practical Experiments. Proceedings of 7th International Computer Assisted Assessment Conference, Loughborough.

Weavers, C. (2003) What do students want from assessment? Exchange, Issue 4 Spring 2003:12–13 http://www.exchange.ac.uk/files/eissue4.pdf

Weller, Martin – Assessment Issues on a Web–based Course – Assessment &Evaluation in Higher Education, Mar 2002, Vol. 27 Issue 2, p109, 8p

Zakrzewski, S. and Steven, C. (2003) Computer–based assessment: quality assurance issues, the hub of the wheel. Assessment and Evaluation in Higher Education. 28 (6) 609–623.

Disclaimer