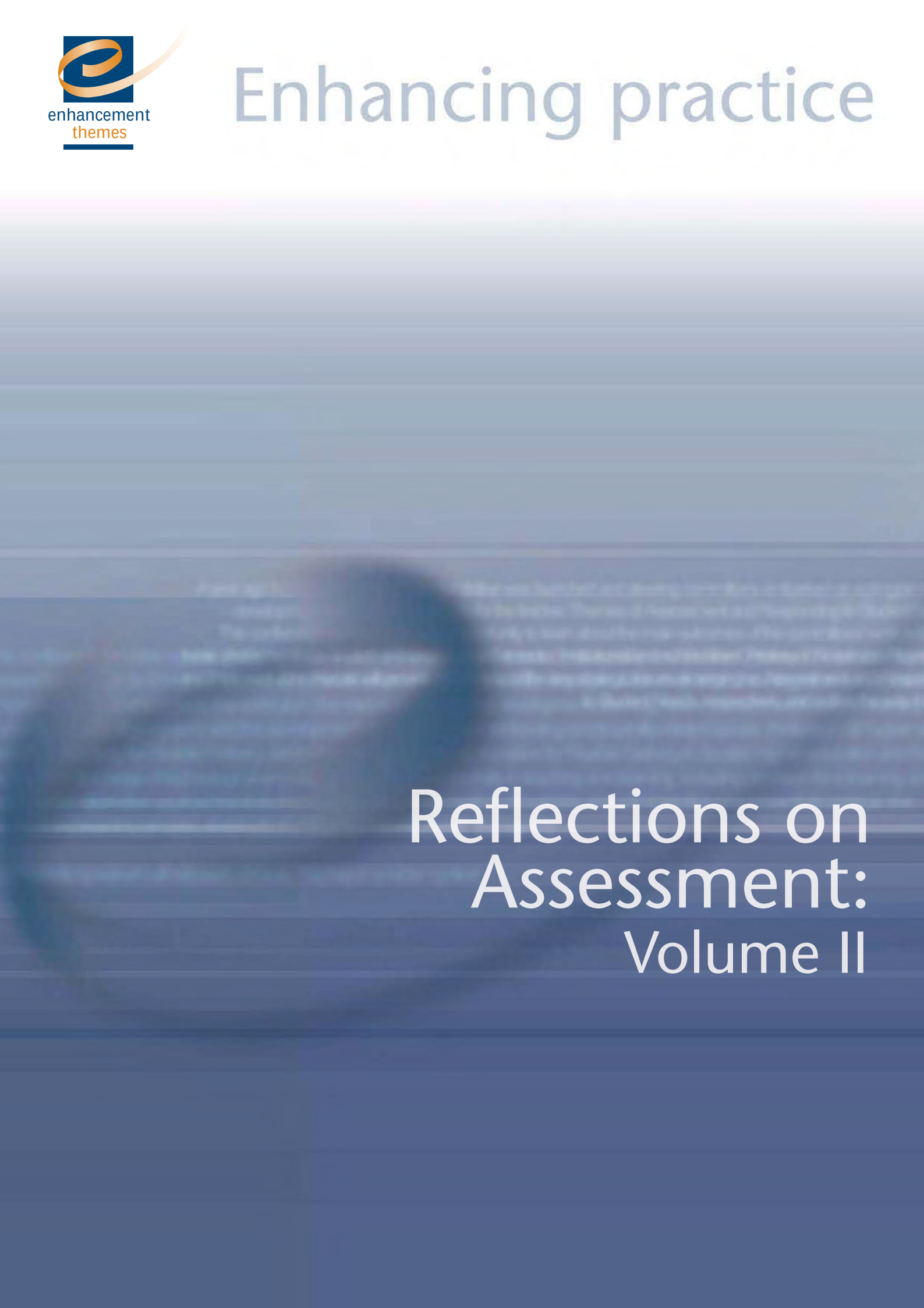
enhancement themes

Enhancing practice



Reflections on Assessment: Volume II

Preface

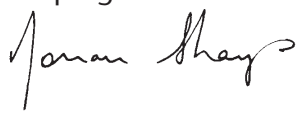
The approach to quality and standards in Scotland is enhancement-led and learner-centred. It has been developed through a partnership of the Scottish Higher Education Funding Council (SHEFC), Universities Scotland, the National Union of Students in Scotland (NUS Scotland) and the Quality Assurance Agency for Higher Education (QAA) Scotland. The enhancement themes are a key element of a five part framework which has been designed to provide an integrated approach to quality assurance and enhancement, supporting learners and staff at all levels in enhancing higher education in Scotland drawing on developing, innovative practice within the UK and internationally.

The five elements of the framework are:

- a comprehensive programme of subject level reviews undertaken by the higher education institutions themselves; guidance on internal reviews is published by SHEFC (www.shefc.ac.uk)
- enhancement-led institutional review (ELIR) run by QAA Scotland (www.qaa.ac.uk/reviews/ELIR)
- improved forms of public information about quality; guidance on the information to be published by higher education institutions is provided by SHEFC (www.shefc.ac.uk)
- a greater voice for students in institutional quality systems, supported by a national development service - student participation in quality scotland (sparqs) (www.sparqs.org.uk)
- a national programme of enhancement themes aimed at developing and sharing good practice to enhance the student learning experience, which is facilitated by QAA Scotland (www.enhancementthemes.ac.uk).

The topics for the themes are identified though consultation with the sector and implemented by steering committees whose members are drawn from the sector and the student body. The steering committees have the task of developing a programme of research and development activities, which draw from national and international good practice. Publications emerging from each theme are intended to provide important reference points for higher education institutions in the ongoing strategic enhancement of their teaching and learning provision. Full details of each theme, its Steering Committee, the range of research and development activities, and the outcomes are published on the enhancement themes website (www.enhancementthemes.ac.uk).

To further support the implementation and embedding of a quality enhancement culture within the sector, including taking forward the outcomes of the various enhancement themes, a new overarching committee has been established, chaired by Professor Kenneth Miller (Vice-Principal, University of Strathclyde). It will have the important dual role of keeping the five-year rolling plan of enhancement themes under review and ensuring that the themes are taken forward in ways that can best support institutional enhancement strategies. We very much hope that the new Committee, working with the individual topic-based themes' steering committees, will provide a powerful vehicle for the progression of the enhancement-led approach to quality and standards.



Norman Sharp
Director, QAA Scotland

Contents

Workshops 1-4 can be found in *Reflections on Assessment: Volume I*

Assessment workshop series - No 5

Assessing online

Contents

Assessing online - An overview

Dr Mary McCulloch; Dr Hamish Macleod, Centre for Teaching, Learning and Assessment; and Nora Mogey, Media and Learning Technology Service, The University of Edinburgh

This is the fifth in a series of eight enhancement workshops concerning different aspects of the assessment of student learning. The issue is about assessing online, but forefronting this should be the issue of assessing per se.

Boud (2000) asserts that assessment always has to perform a number of tasks, he calls this doing 'double duty', and maintains that assessment activities:

- 'have to encompass formative assessment for learning and summative for certification
- they have to have a focus on the immediate task and on implications for equipping students for lifelong learning in an unknown future
- and they have to attend to both the learning process and the substantive content domain' (Boud, 2000).

These assertions are no less true when computers are involved in the assessment process. Online assessment should only be 'one part of a balanced approach to assessment' (Rovai, 2000), and like other assessment should be 'valid, reliable, fair and useful to the student' (Seale, 2002). The input of technology into the assessment process should make us think again about our practices, creating opportunities for quality enhancement, although these opportunities can be missed, and computers can be deployed in support of bad practice just as easily as good practice.

The increased use of e-learning in higher education curricula has led some to suggest that the use of online assessment or computer aided assessment (CAA) should also increase, providing a closer alignment of learning outcomes and learning experiences with assessment methods (Brown et al, 1997) assert that students may experience cognitive conflict because they are required to word process essays (and indeed engage in online tasks), but use pens in examination halls. They suggest that we may not be fair to our students if we train them in one system and yet test them in another.

What do we mean by online assessment or CAA?

Computer aided assessment is frequently assumed to mean the online assessment of activities which most typically have been undertaken offline, for example, an online objective test delivered by a computer, marked automatically which is used in the assessment of a student's understanding of a particular topic or topics. Suitably constructed online tests can be used for summative, formative or diagnostic assessments.

However, CAA also encompasses other types of online assessment. Students may be asked to interact with a simulation, with the results they generate also being assessed within the system. A student who is asked to build a web page may be assessed electronically by submitting it to a validation service, with the report from the validation service being used as a measure of success. A student of linguistics or music may be asked to create a sound which exactly matches a computer generated pattern, how closely they can replicate the sound is recorded and the success of the match evaluated by the computer.

A further family of CAA activities are those where there is an offline assessment of an online activity. A common example of this would be assessing a student's contribution to a discussion forum, but it could also be the assessment of a presentation (a web page, Microsoft PowerPoint presentation, desk top published document) or the assessment of an online skill (creating a database, manipulation of a spreadsheet or a host of other specialist software packages).

In addition, CAA can be defined to include administrative and management aspects of assessments. These would include not only tasks such as collation and moderation of grades, but detailed analytical statistics such as facility values and measures of discrimination which can provide detailed information for the academic team to help them improve and tailor future tests. Straight forward electronic submission of assignments is now frequently enhanced with the electronic return of feedback and results. Students can also build and manage their own electronic portfolios evidencing their skills and achievements.

Bull and McKenna (2004) 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.

Online assessment for different assessment strategies

Diagnostic assessment

Computer aided assessment can be used to conduct diagnostic assessment of students, perhaps at the beginning of a term or semester to inform the lecturers how much the students know before they arrive, or perhaps part way through a term or semester to let the lecturers and students know how effectively the taught elements of the course are progressing. Diagnostic testing can be related to content or skills.

Drew et al (2002) show the use of diagnostic testing for key skills support where students were helped to identify their strengths and weaknesses with 'Skill check questionnaires'. These provided tailored support to students, and the students liked the computer-based system because it made the testing private and personal to them. The honesty from the students was important because this helped to guide them effectively to the support they required. This system has been used in a number of institutions, and more details can be found on the Sheffield Hallam University website provided in the list of links.

Formative assessment

Students can be provided with computerised self-assessment which they can use for formative assessment. This allows for instant and tailored feedback and often provides more, and more timely, feedback than can be provided by tutors. It can also be used to point students to other resources to support their learning. Mackenzie et al (2004) suggest that using CAA for formative assessment allows for detailed feedback and so provides 'substantial learning quality benefits' (Mackenzie et al, 2004).

Denton (2003) outlines a formative assessment system developed at Liverpool John Moores University, a Microsoft Office Marking Assistant called 'Electronic Feedback 9'. Comments are stored relating to a particular piece of assessed work, students can receive standard comments, those to the whole group, as well as individual comments. Staff can see which comments are required most frequently and can ensure some consistency in feedback across different markers. More information is given on the website in the list of links.

The importance of the formative aspects of using CAA are emphasised by Ecclestone (2002) who maintains that e-learning should not privilege summative over formative assessment as this will be detrimental to learning and motivation.

Summative assessment

Summative assessments are often high stakes assessments and so involving computers in the assessment process can be seen to include additional risks and anxieties, more of which will be pursued in following sections. Some students feel that they are disadvantaged by online examinations because these examinations are more stressful or because they hate computers (Ricketts and Wilks, 2002). This is in line with the comments of Brosnan (1999) about computer anxiety affecting performance. However, it is interesting to note that dyslexic students remarked that online examinations were advantageous to them, and some students remarked that this format was less stressful than other exams (Ricketts and Wilks, 2002).

Using computers for summative assessments can provide the potential for students to be presented with more complex scenarios, including interactive resources such as images sounds and simulations, than could be presented in paper-based tests (James et al, 2002). Developing these assessment activities increases the effort required by academics, as well as the technical support required from the institution.

Online summative assessments often require a larger degree of institutional support than diagnostic or formative assessments which can be largely within the control of individual academics. Central units will often provide knowledge and expertise, as well as managing software and servers. Two examples of the support provided can be seen at The University of Edinburgh and the University of Dundee, links to pages can be found in the list of links.

Issues surrounding CAA

Student acceptance

Although students find it difficult to read from a computer screen for long periods of time (Ricketts and Wilks, 2002), they have been shown to accept computer-based assessment because of the speed of marking and the availability of feedback (Ricketts and Wilks, 2002). In addition, Weller (2002) showed that students appreciated the greater degree of individuality that was allowed by using a web format for their assessment items, giving students a greater sense of ownership and pride in their work. However, students are affected by the interface of online tests, preferring questions that are presented one at a time, rather than having to scroll through questions (Ricketts and Wilks, 2002).

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).

Disability

The *Special Educational Needs and Disability Act 2001* makes us think more clearly about the needs of disabled students, especially as these needs may be more apparent when computers are involved. Wiles and Ball (2003) are concerned that as assessments are designed for online arenas they should be designed with all students in mind, and in particular that any variants of the assessments that are designed for the needs of disabled students have the same validity. The issue is whether there can be true equivalence between tests that are offered online to non-disabled students, versus those which are perhaps offered on paper or in other formats to disabled students. This issue and others are discussed on the TechDis assessment pages, the link for which is given in the list of links and in Wiles (2002).

Students should be encouraged to disclose any special needs that they may have in relation to examinations in sufficient time for any special arrangements to be made for them, for example a preferred screen background and text colour, or the allowance of extra time for the examination.

Equity

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) 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.

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.

Where questions come from

Many textbooks associated with higher education courses now come with banks of objective test questions on CD, available from the publisher on formal adoption. These can be used, especially formatively, being integrated into other online materials provided for the students (assuming copyright clearance). Some of the Higher Education Academy Subject Centres are developing question banks with the help and support of colleagues from institutions around the country, and a list of some of these is given at the end of this paper. Alternatively, departments can write questions themselves. Zakrzewski and Steven (2003) suggest that academic staff should create their own question banks and that they should generate an extra 10 per cent of questions each year. Creating question banks will help to ensure the growth and cost effectiveness of CAA in the long term, as Gipps (2003) contends that 'the true costs involved mean that CAA is only really feasible for items that can be re-used'.

Quality assurance

Quality assurance procedures that are effective and robust will be essential to successfully implement CAA systems, some quality assurance strategies for dealing with the risks associated with CAA are provided by Zakrzewski and Steven (2003) and McKenna and Bull (2000, summarised at the end of this paper).

In addition, it is important to engage with a statistical analysis of the questions, post test, to determine whether these questions have been successful as discriminators of student performance. Such analyses can have great power, alongside the examination results to help academics determine which questions should be retained, and which should be altered or removed (Zakrzewski and Steven, 2003). It can be easy to blame the students if there is a set of disastrous exam results one year, however, the reasons may be a combination of student intake, the quality of the testing or the teaching. For a set of questions from a bank with previous statistics, Johnstone (2003) suggests that a useful indicator is to look at changes in facility values (the proportion of the students who choose the correct answer, expressed as a fraction). 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.

Drawbacks

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) 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).

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).

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.

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.

Security

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.

Assessment challenges

Much of the assessment currently taking place online is in the form of objective testing. Some of the challenges facing academics who wish to increase the range of online assessment are how to assess activities which happen online or items which are submitted electronically.

Assessing online discussions is one of these challenges. Parry and Dunn (2000) compared two cohorts of students in online distance education courses, and stressed the importance of assessment to encourage online discussions. One of their respondents said: 'if it's not assessed you probably wouldn't have interacted so much - there's less enthusiasm if it's not assessed' (Parry and Dunn, 2000). Rovai (2000) provides a rubric for grading online discussions, with three grades. He suggests that it is possible to objectively grade the level at which students are interacting in a discussion, relating to the amount that students read, comment on, respond to and collaborate. An example of how to grade the activities of students in asking and responding to questions would be:

Grade C
'Never includes questions that stimulate discussion. Rarely responds to questions raised by others.'

Grade B
'Rarely includes questions that stimulate discussion. Sometimes responds to questions raised by others.'

Grade A
'Sometimes includes good questions that stimulate discussion. Frequently responds to questions from others.'

The full rubric can be found in Rovai (2000).

Conclusions

Yet again with assessment we are looking carefully at the computer-based methodologies in a way in which we have never scrutinised the 'conventional' approaches. We must acknowledge that many of the parameters of our traditional

approaches to assessment (the size of exam halls for example) can lead to differential effects mediated through situational anxiety. We should be aware of how our students experience the assessment activities that we require them to undertake.

We need to question whether there are core assessment processes for which equivalents need to be found in the online environment, or whether we are transposing procedures through habit? If by assessing online we are producing an electronic equivalent of traditional forms of assessment, we may find that the additional issues that arise from the use of technology may outweigh any of the benefits (Weller, 2002).

References

Brosnan M (1999) Computer anxiety in students: should computer-based assessment be used at all? in Brown S, Race P and Bull J, *Computer-Assisted Assessment in Higher Education*, pp47-54, Kogan Page, London

Brown G, Bull J and Pendlebury M (1997) *Assessing Student Learning in Higher Education*, Routledge, London

Boud D (2000) Sustainable assessment: rethinking assessment for the learning society, *Studies in Continuing Education,* 22 (2) 151-167

Bull J and McKenna C (2004) *Blueprint for Computer-Assisted Assessment*, RoutledgeFalmer, London

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

Clariana R and Wallace P (2002) Paper-based versus computer-based assessment: key factors associated with the test mode effect, *British Journal of Educational Technology,* 33 (5) 593-602

Drew S, Thorpe L and Bannister P (2002) Key skills computerised assessments. Guiding principles, *Assessment and Evaluation in Higher Education*, 27 (2) 175-186

Ecclestone K (2002) E-qualifications - a new basis for learning or another bandwagon? *Broadsheet EdExcel*, March 2002, 8-9

Gipps C (2003) Should universities adopt ICT-based assessment? *Exchange*, Issue 4, Spring 2003
www.exchange.ac.uk/issue4.asp [Accessed 22 March 2004]

Hay I and Bull J (2002) Passing online exams, *Journal of Geography in Higher Education*, 26 (2) 239-244

James R, McInnis C and Devlin M (2002) *Assessing Learning in Australian Universities,* Australian Universities Teaching Committee, Canberra
www.cshe.unimelb.edu.au/assessinglearning/ [Accessed 22 March 2004]

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

MacDonald J and Twining P (2002) Assessing activity-based learning for a networked course, *British Journal of Educational Technology*, 33 (5) 603-618

Mackenzie D (2003) Assessment for E-Learning: What are the Features of an Ideal E-Assessment System? *CAA Conference Proceedings 2003*, pp185-194 www.lboro.ac.uk/service/fli/testsite/contents.pdf

Mackenzie D, O'Hare D, Paul C, Boyle A, Edwards D, Williams D and Wilkins H (2004) Assessment for Learning: The Triads assessment of learning outcomes and the development of a pedagogically-friendly computer-based assessment system, in O'Hare D and Mackenzie D (2004) *Advances in Computer Aided Assessment*, SEDA Paper 116, Birmingham

McDonald A S (2002) The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments, *Computers and Education,* 39, 299-312

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

Noyes J, Garland K and Robbins L (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 and Mackenzie D (2004) *Advances in Computer Aided Assessment,* SEDA Paper 116, Birmingham

Parry S and Dunn L (2000) Benchmarking as a meaning approach to learning in online settings, *Studies in Continuing Education*, 22 (2) 219-233

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

Rovai A P (2000) Online and traditional assessments: what is the difference? *Internet and Higher Education,* 3, 141-151

Seale J (2002) Using CAA to support student learning, Starter Guide 4, LTSN Generic Centre/ALT York www.heacademy.ac.uk/resources.asp?process=full_record&section=generic&id=38

Thomas R and Milligan C (2003) Online Assessment of Practical Experiments, *CAA Conference Proceedings 2003*, pp421-430 www.lboro.ac.uk/service/fli/testsite/contents.pdf

Weller M (2002) Assessment issues on a web-based course, *Assessment and Evaluation in Higher Education*, 27 (2) 109-116

Wiles K (2002) Accessibility and computer-based assessment: a whole new set of issues? in L Phipps, A Sutherland and J Seale (eds) *Access All Areas: disability, technology and learning*, ALT/JISC/TechDis www.techdis.ac.uk/index.php?p=1_1_20042209080936_20040610021026

Wiles K and Ball S (2003) Constructing accessible CBA: minor works or major renovations? *CAA Conference Proceedings 2003* www.lboro.ac.uk/service/ltd/flicaa/conf2003/pdfs/contents.pdf [Accessed 22 March 2004]

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

List of links [all accessed 22 March 2004]

Computer Assisted Assessment Centre on objective testing
http://caacentre.lboro.ac.uk/resources/objective_tests/index.shtml

Centre for Interactive Assessment Development at the University of Derby
www.derby.ac.uk/ciad/dev/

University of Dundee information about summative assessment
www.dundee.ac.uk/learning/dol/caa.htm

University of Edinburgh information about summative assessment
www.elearn.malts.ed.ac.uk/

Electronic feedback system at Liverpool John Moores University
http://cwis.livjm.ac.uk/cis/download/xlfeedback/welcome.htm

Formative assessment in Science Teaching
www.open.ac.uk/science/fdtl/index.htm

Links to Sheffield Hallam University website with case studies
www.shu.ac.uk/keytokey/index.htm

Online assessment and feedback
www.bbk.ac.uk/olaaf/

Scottish Centre for Research into On-line Learning and Assessment
www.scrolla.ac.uk/about/assessment.html

TechDis/LTSN Forum for Computer-based Assessment and Accessibility
www.techdis.ac.uk/cba/forum.html

Some question banks [all accessed 22 March 2004]

Applied social surveys
www.socstats.soton.ac.uk/cass/

Business education
www.bized.ac.uk/stafsup/options/bus/bus_g_4.htm#qbank

Economics
www.economics.network.ac.uk/qnbank/

Engineering
www.e3an.ac.uk/

Foreign languages
www.well.ac.uk/languageexercises/

History
www.arts.gla.ac.uk/www/ctich/assess.htm

Medicine and business studies; chemistry, pharmaceutical and biological sciences; mathematics
www.science.ulst.ac.uk/caa/banks.html

Social policy
Contact bob.rotheram@ntu.ac.uk (National Teaching Fellow)
www.swap.ac.uk/learning/assessment4.asp

Veterinary science
www.ltsn-01.ac.uk/resources/best_practice/display_single_item?BestPracIndex=259

Some important quality assurance recommendations provided by McKenna and Bull (2000)

- 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.

Assessment as a catalyst for innovation

Professor Gráinne Conole, Research and Graduate School of Education, University of Southampton

Abstract

This paper provides a critique of the impact of technology on assessment and considers whether innovative uses of information and communications technology (ICT) might result in new e-pedagogies and practices in assessments. The paper consists of an overview of the key characteristics of ICT and associated pros and cons, along with a reflection on computer-assisted assessment (CAA) specifically and the associated key research questions and issues.

Introduction

Technologies provide potential new opportunities for teaching, learning and assessment. Conole and Dyke have developed a taxonomy of the 'affordances' of ICT (Conole and Dyke, 2004), where affordance refers to the 'perceived and actual properties of a thing, primarily those functional properties that determine just how the thing could possibly be used' (Salomon, 1993). They argue that a better understanding of the properties of different technologies might enable practitioners to make more informed judgement about how technologies can be used to support teaching, learning and assessment. Their taxonomy consists of the following affordances:

- speed of change
- diversity
- communication and collaboration
- reflection
- multimodal and non-linear
- risk, fragility and uncertainty
- immediacy
- monopolisation
- surveillance.

With technologies there are two sides to every story (Figure 1). What's fascinating about technology is the way in which whatever aspect of it you look at there are pros and cons. So yes the internet now means we have access to vast quantities of materials but there are increasingly real issues about information overload and issues about access and the quality and authenticity of materials. The wealth of new tools for communication offer opportunities for new forms of dialogue and collaboration but brings with them associated issues of the need for new skills for both students and teachers.

Figure 1 A spectrum of advantages and disadvantages of the affordances of ICT

An inevitable feature of ICT is that technologies are continually developing in scope and sophistication. A particularly important aspect of this is the prevalence of abundant and rapidly changing information mediated through extensive communication technologies (Conole and Dyke, 2004). This has a significant impact on social behaviour and practice in that our world is increasingly constituted by information rather than pre-given modes of conduct. This requires practitioners to continuously reassess their needs and also the information, which is part of these experiences, instead of relying on custom and tradition to guide action. The immediacy of access to rapidly changing information or events is a core feature of new technologies, enabling unprecedented speed of access to materials and world events as they happen. However, this speed can also raise issues about quality and lack of authority of sources. The speed of change may also militate against reflective and critical thought, fostering surface approaches to learning. The issue here is that the speed of change in a world full of conflicting and changing information presents a challenge for the educational use of the new technologies. In other words, how can ICT be used to enable practitioners and students to navigate their way through the myriad of changing information and make more informed decisions? There is also a constant tension and mismatch in terms of the skills levels of students and tutors. Students in some cases have more sophisticated skills levels in terms of the use of ICT than their tutors. However, these might be skills levels of a particular kind, for example, experience of gaming environments, whereas other e-literacy skills (such as critical evaluation of the value of online resources or experience of using office applications) may not be so well developed.

The communication and collaborative abilities of technology present another key affordance that offers the potential for learning enriched by engagement with the 'other'. New technologies have opened up the possibility of new forms of dialogue and communication. ICT offers the potential to develop new forms of online communities and new means of communicating and sharing information (Preece, 2000). However, this can lead to issues in terms of individuals being 'spread too thinly' across communities, as well as issues of lack of identity and peripheral engagement.

Asynchronous communicative tools (in particular) offer the potential for encouraging reflection and critique with users engaging in discussions over a longer time frame than is possible in face-to-face discussions. The use of CAA tools for formative assessment also has the potential to promote reflection by providing students with immediate feedback on their progress. Clearly, there is nothing inherent about ICT that nurtures reflection - the key is how it is used. ICT has the potential to enable reflection and criticality to be enhanced. There is, equally, a risk that the speed and pace of information change outlined above militates against reflection. It leaves no space for contemplation and considered judgement, and promotes a more pragmatic, reflexive immediate response to new information.

Another affordance of ICT is the potential for multimodal and non-linear approaches to navigating through information. The non-linearity of the web (epitomised by hypertext and the use of powerful search engines) leads to the potential for different routes through, and forms of, learning. ICT enables the learner to move beyond linear pathways of learning. Yet much current computer-based training material still appears to follow a linear, assembly line, mode of learning. Many e-learning packages are built

on behaviourist principles of atomised experiences that need to be completed in a specified order before the individual is positively reinforced and permitted to move on - a form of electronic page-turning. More complex, multimodal and non-linear approaches also create issues in terms of increased navigational skills and problems with student potential getting lost or confused.

E-learning clichés and reality

There is much hype around the potential use of technologies to support learning and Table 1 lists some of the main culprits.

- Use of ICT will lead to new forms of learning.
- ICTs offer the potential to be adaptive, flexible and tailored to individual needs.
- Mobile technologies means learning can take place anywhere, anytime.
- Smart technologies can be adapted to individual needs.
- ICTs offer new forms of rich multimedia representation.
- ICTs enable better and more tailored learning.
- New models of e-learning are emerging.
- Learning objects offer the potential for materials to be reused and repurposed across a variety of contexts and applications.
- The internet enables a globally connected, information-rich society.

Table 1 E-learning clichés

Although there is more than a grain of truth in all of these, understanding the nature of technologies and their impact is much more complex and multifaceted than these simple headline statements might suggest. For example, there is the notion that the use of technology may lead to new forms of learning, but what evidence do we have of this? In a review of the use of tools and their impact in practice, Conole concludes that the tools which have had the greatest impact in terms of changing practice are Microsoft Word, email and the internet (Conole, 2004a). The two main advances in terms of the use of tools over the last five years have been the use of tools for presenting and receiving information, and tools for communication. The best example of the increased use and importance of ICT is the now ubiquitous use of the computer as a work tool and the replacement of many traditional work modes of communication, such as memos, with online communication via email. There has been a commensurate growth in the use of technology to support learning, fuelled by the increased use of all-in-one software such as virtual learning environments (VLEs). The emergence of VLEs in the mid-nineties quickly gained popularity because they offered ease to use, all-in-one environments to support learning and teaching. Practitioners could easily relate to the different tools which aligned well with their existing practice in terms of preparing, delivering and assessing. There has also been a growth in the amount and variety of resources to support learning and research, through specialised gateways and portals, and niche resource providers. The types of software tools, hardware systems and online environments have also increased in variety and complexity with tools now available to support everything from research publication management to online assessment and monitoring.

There is a lot of rhetoric about how new technologies offer the potential to be tailored and adaptable to end user needs, and related to this is the idea that appropriate use of technologies will lead to better and more tailored learning provision, with new technologies offering them rich authentic multimedia learning environments more suited to the increasingly diverse student population in higher education. A hot topic at the moment is the search for underpinning models of e-learning and the belief that such models might better help us develop good practice in e-learning or mechanisms for learning from the good practice of others (Beetham, 2004; de Freitas and Mayes, 2004). Another key Holy Grail is the search for reusability and repurposing and the notion of small transferable chunks of learning which can be adapted and reused for multiple purposes (Littlejohn, 2003). Finally, there is the idea that technologies now enable us to live and learn in a globally connected society offering potential for new forms of discourse and access to new distributed communities.

So what is the reality? On the negative side there is actually a lack of widespread uptake of the use of e-learning - use of communication tools is patchy, with interactions online often forced or stilted and despite the wide scale use of VLEs now much of the use is little more than their acting as content repositories. More worrying, most of the use is not pedagogically informed and there is a dearth of underpinning theory. Despite the variety and the potential of tools to support learning, evaluation research shows that it is difficult to encourage authentic virtual learning or collaboration. Discussion board use, for example, often shows a pattern of peak use directly related to teacher intervention or responses to particular 'hot' topics. Collaborative group work needs to be carefully set up and orchestrated to achieve desired results and despite this may still end up as a rather stilted collaboration exercise not comparable with direct face-to-face equivalent group work (Jones, 1999). Integrated learning environments are still predominantly used as shells for displaying web pages and rarely get beyond basic information dissemination and administration (Thomas and Wyatt, 1999). With respect to information seeking and handling the sheer volume available to learner/researchers is increasing exponentially, unmatched by the sophistication of the searching and handling tools (Lawrence and Lee Giles, 1999). Information overload, coupled with confusion of where to look, is increasingly problematic and, despite a growth in the range of searching tools and portals, it is not evident that the right information is being dispatched to the right users in a timely and quality assured fashion (Conole, 2002).

Therefore, despite the variety of ICT tools and resources available and the recognised potential benefits of using these to support teaching and learning, practitioners lack the necessary e-learning skills to take full advantages of the potential affordances (Conole and Dyke, 2004) that these technologies offer and complain that support and training in this area is inadequate.

However, the use of ICT in education has increased significantly in recent years and we could argue that there is now a critical mass of tools and resources available to support learning and teaching. Teachers are becoming more information technology literate, for example, use of email and the internet is now ubiquitous and in general there has been a shift to a more socially situated approach to learning and teaching and some novel examples of using technologies to support this.

Focusing on CAA

If we focus in particularly on CAA, evaluation studies consistently reveal a set of specific pros and cons (Table 2). There is the attraction of using CAA in terms of potential time saving and, in particular, time saved on teacher marking, but this is coupled with the upfront investment of effort in terms of writing the questions. Furthermore, good question writing is a skilled technique and takes time to develop. The use of shared question banks across subject domains has become increasingly popular, particularly in science and engineering. Interestingly, one of the perceived benefits of CAA from the student perspective is the idea that somehow a computer-marked test is more objective! A potential draw back for some is the view that CAA is too restrictive and can't be used effectively to measure higher order thinking. A national survey of CAA revealed that on the whole it was being used to support formative assessment, the lack of widespread use of CAA for summative assessment is almost certainly associated with the perceived risks and security issues. Finally, CAA users complain that many of the CAA tools themselves are still very rudimentary and restrictive

Pros	Cons
Potential time savings	Considered restricting
New pedagogical models	Time and effort
Repurposing year on year	Difficult to measure higher order thinking
Reflection on practice	Security issues
Shared question banks	Stress!
'More objective'	Tools still rudimentary

Table 2 Pros and cons of CAA

The role of technology and how it might impact on assessment is still in its infancy and we need to develop new models for exploring this. For example, there is a wide range of software available to facilitate research. Bibliographic software facilitates powerful online literature searches and specialist statistical packages that support quantitative data analyses are well established, enabling researchers to focus less on routine calculation and more on the analysis of statistical outputs. Computer-assisted qualitative data analysis software facilitates management of data sets from large-scale projects, from coding through to sophisticated analysis and modelling. However, little research has been done on trying to understand the ways in which these tools are impacting on and changing research practice. Similarly, little is understood about the barriers and enablers to using these tools effectively. In light of this, there are, therefore, real issues in terms of what we might expect or want students to learn. There is a lot of rhetoric about shifting from an information-focused approach to learning to one more based on problem-based learning and inquiry. However, this shift is even more urgent if we consider the fact that students have automatic access to information. In addition, there is a general shift across education from assessment of products or outputs to assessing the processes of learning. Therefore, we need to consider what we actually want to be assessing and how best to do this.

Similarly, the new forms of communication and collaborations which are now possible with technologies also raise a number of issues about what we should be assessing and how to assess it - how do we measure the interactions which occur within an online discussion forum and how can we attribute this in terms of student learning? What about those who don't contribute - 'the lurkers' - are they opting out or learning differently ie vicariously by reading and reflection on the postings of others? Interactive e-journals are changing the nature of academic discourse and the relationship between authors and referees (Hey, 1997; Ingraham, 2000). Similarly, specialised gateways such as the Social Sciences Information Gateway have emerged which help to categorise resources but raise interesting questions about quality control and scope. Free academically owned publishing mechanisms such as ePrints are now becoming more respected and referenced within the research community, but how does this impact on the future of more traditional journals? Finally, copyright and plagiarism issues can be discussed in relation to use and repurposing of other researchers materials (Carroll and Appleton, 2001).

Overall, it is clear that technologies are having, and will continue to have, a major impact on all aspects of learning, teaching and assessment and therefore we need to consider carefully how these resources and tools can be harnessed effectively.

Changing roles and identities

Individual roles and identities are also changing as a consequence (Table 3). Students are becoming increasingly ICT literate. The 'gaming generation', for example, have grown up with interactive games which enable them to control and manipulate their environment (Prensky, 2001). As a consequence, students have increased expectations of higher education and increasing demands for more technology enhanced environments including access to wireless networks, information about courses online and the ability to use their own laptops onsite. However, it is also true that students recognise the value and benefit of face-to-face teaching and many do not want this significantly replaced by online resources.

Students	Practitioners
Changing	Research versus teaching
Skills base	Conflicting demands
Support needs	Increasing collaborative
Employer expectations	Need to link them
Expectations	Roles and structures
Access to resources	Increasing skills gap
Equality of experience	Shifting roles and structures
Student ownership understanding	Lack of senior management
Associated issues	
Plagiarism, copyright	
Monitoring, surveillance	

Table 3 Changing roles and identities

Academics' roles have also changed as a result of the increased availability of technologies, and academics are facing increasing and conflicting demands - there is a particular tension between increased pressure to evidence research excellence and a requirement to be a reflective practitioner engaged in innovative uses of technology to enhance learning.

Similarly, technologies are beginning to have an impact on the very nature and structure of organisations and the ways in which they are organised and run (Conole, 2004b). The advent of commercial VLEs such as WebCT and Blackboard has enabled practitioners to explore the potential of these all-in-one packages to support teaching and learning. As a consequence, many institutions have now begun to consider how these systems can be more systematically integrated with student record systems, finance systems and recruitment processes, and how these might be used as a starting point in the development of an overarching institutional managed learning environment. There has been a significant shift in higher education in the past decade in terms of the nature and potential composition of the environmental context within which UK institutions work. Previously, the environment was relatively stable, with standard and established administrative processes and stable methods and technologies for learning and teaching. The increased range and potential of new technologies over the past decade has begun to have a major impact on organisational structures, roles and identities and new forms of learning and teaching innovation. Much of this is now crystallising through managed learning environement developments which foreground these changes more visibly and by their nature demand cross-institutional engagement and impact.

Conclusion

This paper has provided a critical review of technology and its impact in education, specifically focussing on the issues this raises for assessment and in particular:

- what new forms of assessment might arise as a result of the impact of technologies?
- can technologies enable new forms of pedagogy?
- what new forms of literacy are emerging?
- how can technologies be effectively integrated?
- how might the technologies support distributed-knowledge networks?
- in what ways might the new technologies offer new models of distributed cognition?
- what is the role of policy in all of this?

These are the central questions which we need to research and address if we can achieve the desired goal of maximising the potential technologies offer to improve learning, teaching and assessment.

References

Beetham H (2004) *Review: developing e-Learning Models for the JISC Practitioner Communities*
www.jisc.ac.uk/index.cfm?name=elp_outcomes

De Freitas S and Mayes T (2004) *Review of e-learning theories, frameworks and models,* commissioned review report as part of the JISC-funded e-pedagogy desk study on e-learning models
www.jisc.ac.uk/index.cfm?name=elp_outcomes

Carroll J and Appleton J (2001) *Plagiarism - a good practice guide*, JISC report, May 2001
http://online.northumbria.ac.uk/faculties/art/information_studies/lmri/Jiscpas/site/jiscpas.asp

Conole G (2002) Systematising Learning and Research Information, *Journal of Interactive Media in Education*, Special Issue on Theory Learning for Technologies, 2002(7)
www-jime.open.ac.uk/2002/7/

Conole G (2004a) Report on the effectiveness of tools for e-learning, commissioned review report as part of the JISC-funded e-pedagogy desk study reviewing resources, tools and services for e-learning

Conole G (2004b) *Taking account of organisational context in the development of managed learning environments,* paper presented at the American Educational Research Association conference, San Diego, April 2004

Conole G and Dyke M (2004) What are the affordances of Information and Communication Technologies, *ALT-J*, 12.2, 113-124

Conole G and Oliver M (2002) Embedding Theory into Learning Technology Practice with Toolkits, *Journal of Interactive Media in Education*, Special Issue on Theory Learning for Technologies, 2002(8)
www-jime.open.ac.uk/2002/8/

Hey J (1997) E-journals for Research: the User Perspective, *Serials* 10 (1): 65-68

Ingraham B (2000) Scholarly rhetoric in digital media, *Journal of Interactive Multimedia Education*
www-jime.open.ac.uk

Jones C (1999) From the sage on the stage to what exactly? Description and the place of the moderator in co-operative and collaborative learning, *ALT-J, 7* (2) 27-36

Lawrence S and Lee Giles C (1999) Accessibility of information on the web, *Nature*, 400, 107-109

Littlejohn A (ed) (2003) *Reuse of Educational Resources for Networked Learning,* Routledge, London

Preece J (2000) Online communities: *designing usability, supporting sociability,* John Wiley, New York

Prensky M (2001) *Digital game-based learning*, McGraw-Hill, New York; London

Salomon G (ed) (1993) *Distributed cognitions - psychological and educational considerations*, Cambridge University Press, Cambridge

Thomas G and Wyatt S (1999) Shaping cyberspace - interpreting and transforming the Internet, *Research Policy,* 28 (6)

Websites

Eprints
www.eprints.org/

Journal of Interactive Multimedia Education
www-jime.open.ac.uk

Online assessment: quality production and delivery for higher education

Professor Don Mackenzie, Centre for Interactive Assessment Development, University of Derby

Abstract

This paper focuses on the practical steps that can be taken to ensure the maintenance of quality of online assessment from design through production and delivery to data retrieval, moderation and feedback into course design. The general stages in the design to delivery process are outlined and the quality assurance checkpoints are discussed. The relative merits of two models of production are examined. In the first of these, the Devolved Tutor Development model, the academic tutor is responsible for a large proportion of the process from design, through production to delivery and reporting. In the second of these, the Integrated Team Development model, the academic tutor works closely with a central support team of assessment developers who provide advice and manage the production, delivery and analytical data reporting. The likely outcome from the application of each model is reviewed in terms of quality, diversity and level of assessment. The challenge is to develop a method of working that is scaleable and economic while delivering assessments that have the rigour required by the higher education environment.

Introduction

In the discussion that follows I focus on the issues associated with the production and delivery of automatically marked, computer-delivered assessments (CDA) in their broadest sense on an institution-wide basis. The more remote the delivery, the greater the complexity of quality assurance issues.

The benefits of CDA have been discussed many times (eg Mackenzie et al, 2004) and it is tempting to overemphasise the potential time savings on offer and overlook the substantial enhancement in quality that can be achieved especially by the application of 'advanced' computer-based assessment (ACBA).

Advanced computer-based assessment goes beyond the simple multiple-choice/response item types into the area of complex question types, adaptive branching, scenarios and simulations of real life situations or problems. In formative mode, feedback may be immediate, extensive, context-sensitive and include annotation of diagrams, links to web resources or full courseware/e-learning tutorials making it an extremely powerful learning tool. Indeed, taken to its limits, the intimate embedding of both formative and subliminal summative assessment within e-learning materials could allow production of examination-less e-learning modules in some disciplines with reduced problems of authentication of end-users. Space precludes the description of examples here but the reader is referred to an interactive demonstration of a range of CDA and ACBA questions that is available from www.derby.ac.uk/ciad

Despite these potential benefits, it is important to realise that CDA is not a panacea and that there are many areas where it is not practical or appropriate to use it. Indeed it is beneficial for students to be exposed to a wide range of assessment environments during their course. The large scale computer delivery of summative assessments is extremely resource and expertise intensive, requires robust networks, specially equipped examination rooms and a raft of failsafe and disaster recovery procedures, particularly for 'high stakes' (BS 7988) examinations.

When implementing computer-based assessment (CBA) on an institutional scale there are many issues to be resolved to ensure its quality and successful delivery. Individual aspects have been discussed in detail by Bull and McKenna (2004) but here I wish to take an overview of the whole process.

Risks to quality and barriers to take-up

There are wide ranging threats to quality in the assessment creation process and a number of barriers to take-up of CBA. When converting from traditional models of assessment, these need to be appreciated at the outset and steps taken to ameliorate their effects. Not the least of these are the skill sets of academic assessment authors. Designing assessments for computer delivery is a specialist skill that combines imagination with sound pedagogical principles and knowledge of the capabilities of the assessment system. Even with simple question (item) types there are significant pitfalls in scoring strategies for the unwary (Mackenzie and O'Hare, 2002). Direct conversion of paper-based to computer-based seems a good way to get started but it

may not produce the best assessments. Some of these points are further expanded and set in national context by Boyle and O'Hare (2003).

Assuming that these issues have been overcome, the broad elements in the production and delivery cycle for a summative assessment are:

- course design and definition of learning outcomes
- question (item) and assessment design
- assessment programming
- assessment testing
- deployment/installation on network/server
- assessment delivery - BS 7988 Code of Practice
- marks processing, analysis, moderation and reporting
- evaluation of results and assessment performance
- assessment and/or question modification for future delivery where necessary
- feedback of evaluation into the course delivery or design in subsequent runs.

Successful delivery requires quality assurance checks at all stages in this cycle. The feedback loop into course design/delivery and the design of subsequent assessments is often overlooked in the dash for time saving but is particularly important and one of the principal benefits that the very detailed records of performance provided by CBA can provide.

The process for the development of formative assessments is somewhat simpler but more time consuming in the production of useful feedback. It can also be helpful for tutors if student progress can be monitored and tracked so that help can be targeted to those who need it most.

There is a common assumption that individual academic tutors will engage in a high proportion of the process outlined above when converting traditional modes of assessment to computer delivery. However, if we are attempting to use CBA to test higher order skills it may be necessary to use more sophisticated assessment tools that require steeper learning curves in order to gain the level of interactivity required. Should we really be expecting tutors to be software programmers as well as subject material and pedagogic experts? What model of development should we adopt to ensure quality and scalability on an institution-wide basis?

Models for institutional CBA development

Two end-member models for CBA development are compared in this section and in Table 1.

- Devolved Tutor Development model
  Here the design, programming and delivery of the assessment is entirely handled by individual academic tutors.

- Integrated Team Development model
  In this case the academic tutor is supported by a team of experts throughout the whole process which I shall refer to as the Assessment Support Team (ASTeam).

Some institutional procedures may fall between these end-members.

Process	**Devolved Tutor Development**	**Integrated Team Development** **Assessment Support Team**
Course design	Academic course team, including tutor with responsibility for the assessment **IQC**	Academic course team, including tutor aided by ASTeam on request **IQC**
Assessment design	Academic tutor	Academic tutor in collaboration with ASTeam **IQC**
Programming	Academic tutor programs assessment and tests program	ASTeam programs assessment and tests program **Intra department IQC**
Testing	Academic tutor + peers check content and refer back as necessary **IQC**	Academic tutor + peers check content and refer back as necessary **IQC**
Network deployment and testing	Tutor uploads to virtual learning environment (VLE) (formative only) or negotiates with Information Technology (IT) Services for deployment. Tutor tests network version	ASTeam uploads to network and checks functionality **Intra department** **IQC**
Delivery	VLE automatic or tutor negotiates with IT Services and exam administration for rooming and delivery window	ASTeam monitors delivery and provides failsafe technical support. Tutor negotiates with exam administration for rooming and delivery window. **IQC**
Marks collation	Tutor retrieves from VLE or network folder. Undertakes item analysis where able	ASTeam retrieves results and provides report with statistical analysis and question performance indicators **IQC**

Process	Devolved Tutor Development	Integrated Team Development Assessment Support Team
Moderation and evaluation	Tutor evaluates/ moderates results - peer moderation? **IQC**	Tutor evaluates/moderates results with recourse to ASTeam for support regarding performance indicators etc + peer moderation? **IQC**
Feedback into future course and assessment design	Tutor may modify future course delivery or assessment content **IQC**	Tutor may modify future course delivery or assessment content aided by ASTeam on request **IQC**

Note: **IQC** = Independent Quality Check

Table 1 Comparison of institutional CBA development models

Features of the Devolved Tutor Development model

This could be undertaken using the rudimentary functionality embedded in the institutional VLE. The advent of VLEs has been instrumental in empowering many tutors to place their course resources on line in a structured environment. However, VLEs can be restrictive and are generally unsuitable for summative assessments because any candidate who is enrolled on the course may view the assessment whether or not they are in the examination room. A dedicated assessment system is normally necessary to deliver tests outside this environment and accessed via separate logins. Some commercial systems are now sufficiently user-friendly for tutors to engage with, even on an occasional basis, to extend the functionality that is normally included within a VLE.

In the Devolved Tutor Development model the academic tutor programs and tests assessments then uploads to network/VLE for delivery. This could be simple for formative assessments in VLEs or departmental networks or complex for summative assessments on university networks where it might involve negotiation with network managers, room booking clerks and examination administrators together with a further phase of testing of the network version by the tutor.

Some kind of independent quality assurance gateway would seem appropriate here in the case of summative assessments. It is essential in this model that the assessment is subjected to independent peer review prior to deployment since ambiguities and typographical errors are often overlooked if the programmer is also the assessment author.

In this model, the tutor retrieves results files, analyses and moderates marks and reports them. If the assessment is deployed through a VLE, the marks may go unmoderated into the VLE grade book and are usually immediately available to the student. This is not a problem for formative assessments but it is another reason for

not using VLEs for summative assessments where some moderation of final scores may be necessary as a result of item and test performance analysis. More numerate tutors could calculate item analysis statistics to inform them about question performance and reliability if this functionality is available in the delivery system. However, there is a danger that this quality assurance step may be missed if the tutor is not familiar with the statistical procedures with the result that the benefits may not feedback into future iterations. The level of IT skills possessed by the academic tutor has a strong influence on the nature of the resulting assessment.

Features of the Integrated Team Development model

Development using this model is usually undertaken with a more sophisticated, dedicated assessment development system. Such a system may have a steep learning curve that makes it less suitable for individual tutors to pick up on an occasional basis. On the other hand such systems can provide the functionality to create highly interactive and very rigorous assessments that may involve the use of multimedia.

In this model tutors/course teams design assessment with the help of an ASTeam and/or instructional designers where required, paying due attention to alignment of assessment with learning methodology and outcomes. The ASTeam is familiar with the development software and scoring strategies so that errors are less likely there is a potential to deliver assessments that begin to tackle higher order skills in some disciplines.

The ASTeam programs the assessment which is then checked by the tutor and ideally independently by their peers. Once checked, the ASTeam assumes responsibility for the delivery, collation of results, item analysis and reporting to the tutor. The whole process is quality assured throughout, the screen presentation is consistent and the tutor has the benefit of the wider assessment experience of the ASTeam that may cross many disciplines.

The ASTeam can guide tutors through the results analysis so that they can make informed moderation decisions and can understand the implication for future runs of the assessment and course implementation.

Devolved Tutor Development	Integrated Team Development
Richness of assessment restricted by the IT skills of the tutor and user-friendly software functionality.	Design not limited by IT capabilities of tutor. Interactivity can be efficiently produced by expert programmers in ASTeams resulting in more sophisticated assessments. Teams can be more imaginative than individuals.
Assessment paradigm may be restricted to the discipline of the tutor.	Transfer of ideas from other assessments in other disciplines is facilitated by ASTeams.
Risk of inappropriate scoring leading to anomalous results.	ASTeams have wide experience, and can advise on scoring strategies.

Devolved Tutor Development	Integrated Team Development
Development of assessments that test higher order skills is difficult to achieve with simple systems.	ASTeams are using assessment systems with less restriction on the range of question and assessment types. Easier to design assessments to test higher order skills.
Quality and graphic design and screen presentation is likely to be variable.	The graphic design and the user interface is likely to be of a higher quality because it is produced by experts in the ASTeam. Less variability between assessments from different tutors. Easier to use for candidates.
Item analysis may not be undertaken because of lack of understanding. Marks may be unmoderated.	Item analysis fully integrated into the results reporting and evaluation. Marks more likely to be moderated on basis of item performance.
Represents a significant increase/change in tutor workload/pattern.	Less of an impact on tutor workload/pattern. Tutor can concentrate on pedagogic design of assessment rather than programming.
Danger of few independent quality checks.	More independent quality checks within the process. An expert undertakes each task therefore there is less risk of error.
May be best suited to production of simple formative assessments but lacks the rigour necessary for the development of medium or high stakes summative assessments.	Rigorous production and delivery controls make this model suitable for summative assessments at all levels and for innovative formative assessments.

Table 2 A comparison of likely outcomes of each development model

Outcomes of the application of each of model

The likely outcomes of the application of each model are summarised in Table 2.

The challenge is to develop a method of working that is scaleable and economic while delivering assessments that have the rigour required by the Higher Education environment.

Clearly, the Integrated Team Development model, with independent checking at each stage and input from a wider range of expertise, appears to have the greatest overall potential for the development of higher quality assessments, both in technical and in educational terms. This model provides the best opportunities to extend the scope of CBA into new areas of assessment and will promote a more rapid and informed take-up with higher quality output. In some disciplines, it will allow us to undertake assessments simulating real situations that would be expensive or impractical to assess

by traditional means and to test higher order skills very effectively. However, to implement this university-wide for all types of CBAs could be expensive and difficult to scale up. This begs the question as to what extent should individual academic tutors become software developers.

Moving beyond the level of simple formative quizzes might seem difficult using the Devolved Tutor Development model, but some commercial assessment systems now offer relatively user-friendly creation of question styles that can be used very effectively both formatively and summatively. Some institutions have used this strategy to great effect particularly for formative work (eg Universities of Loughborough and Dundee). It is important to stress here that it is the quality of design and underpinning pedagogy that creates a good assessment, not the complexity of the question (item) types within it. However, tutors need appropriate tools with which to work and usually the simple functionality of a VLE is not enough. Furthermore, there is need for some sort of 'quality assurance gateway' to ensure the reliable and consistent delivery of summative assessments using this model.

Conclusion

Perhaps the answer lies in compromise, with simple formative and summative assessments being created by individual tutors with user-friendly systems. The ASTeam could then focus on providing advice to course development teams where required, programs more sophisticated and high stakes assessments as well as providing the quality assurance gateway for assessments produced using the Devolved Tutor Development model. This team could also oversee the delivery of summative assessments.

References

Boyle A and O'Hare D (2003) Finding appropriate methods to assure quality computer-based assessment development in UK higher education, *Proceedings of the Seventh International Computer Assisted Assessment Conference*, pp67-84, University of Loughborough

Bull J and McKenna C (2003) *Blueprint for Computer-assisted Assessment*, Chapter 10, pp115-134, RoutledgeFalmer, London

Technical Committee IST/43, British Standards Institute (2002) BS 7988, *A Code of practice for the use of information technology (IT) for the delivery of assessments,* BSI

Mackenzie D M and O'Hare D (2002) Empirical Prediction of the Measurement Scale and Base Level 'Guess Factor' for Advanced Computer-based Assessments, *Proceedings of the Sixth International Computer Assisted Assessment Conference, Loughborough 9 and 10 July 2002,* University of Loughborough

Mackenzie D M, O'Hare D, Paul C, Boyle A, Edwards D, Williams D and Wilkins H (2004) Assessment for Learning: The TRIADS Assessment for Learning Outcomes Project, and the development of a pedagogically friendly computer-based assessment system, in O'Hare D and Mackenzie D M (2004) *Advances in Computer Aided Assessment,* SEDA Paper 116, Staff and Educational Development Association, Birmingham

Aligning e-learning with assessment and learner support

Dr Janet Macdonald, The Open University in Scotland

Abstract

The use of online media allows us to develop courses that give students access to the opinions of peers and the resources of the web. The result is that they may need to study in a far more self-directed approach than has previously been possible, or even necessary. So, instead of telling them what they need to know, we are increasingly expecting them to find it out for themselves. But how might such courses be assessed? We know that assessment plays a major formative role in driving student learning appropriately, but what implications does this have for e-learning courses? Should we reconceptualise the ways in which we assess students, or are existing methods, tried and tested in conventional teaching and learning situations appropriate? Is the assessment of course content alone still appropriate? This paper discusses some practical approaches to supporting learner development through assessment design, drawing on examples from recent research at the Open University.

Results and discussion

E-learning and lifelong learning: it's not what we learn, but how we learn

Online media offer the opportunity to develop more open and self-directed courses than have been possible in the past. So rather than 'opening their heads and pouring in the content', and then testing students to see how much remains at the end of the course, many course writers are turning their hands to courses in which students are 'learning by doing'. Such courses place emphasis on the development of skills which students will need throughout life.

Online media may be used to give students access to the wider resources of the web, or perhaps to a more limited set of resources chosen by the tutor. Such courses may also involve the use of online conferencing for discussion on course concepts. But these courses often make unfamiliar demands on the students who study them, because they need to develop **communicative and interpretive ability** in using online media. This probably involves the use of computers as an effective study tool, but it may also mean the development of the critical and analytical abilities to work as a self-directed learner, together with the communicative abilities to work with, and learn from, their peers. In fact, competence in e-learning has many parallels with competence in learning and is probably acquired in a developmental progression, alongside developing confidence and reflective learning, and a growing understanding of a discipline. In a conventional university course of study, we might expect students to be competent in these skills by the time they graduate (Perry, 1970). But we are increasingly expecting our students to undertake these rather demanding tasks at earlier stages of their undergraduate studies, and very often in their first year.

The question is, how might such skills be developed? Biggs (2003) suggests that we should be thinking in terms of 'constructive alignment', in other words making sure that our students are offered the opportunity to actively construct their understanding by undertaking tasks which are relevant to the course outcomes. Since it is well known that assessment is a very significant driver for student learning, we need to devise ways of harnessing its power to encourage students to practise and hone their skills. In fact, assessment can be used not only to check students' knowledge and understanding of the subject, it has an increasingly important role in helping them to build the skills they need to be competent e-learners and competent lifelong learners.

In this paper, we discuss the potent formative value of assessment for developing lifelong learners, by drawing on examples from three qualitative studies of two e-learning courses at the Open University (UK); these studies are reported in detail elsewhere (Macdonald, 2001; Macdonald, Heap and Mason, 2001; Macdonald J, Weller M and Mason R, 2002; Macdonald and Twining, 2002; Macdonald, 2003).

Assessment design for supporting e-learning skills

We have learnt that assignments may provide much valued practice and support in the process of learning on online courses, as well as covering course concepts. For example, in one assignment, which is drawn from a second level undergraduate technology course, we encouraged students to engage in resource-based learning. They were required to use a search engine to select a limited number of academic papers to provide material for an essay on a course issue. Not only were they

expected to write the essay, they also had to reflect on the process of selection, and describe how they chose the papers. In this way, we wished to encourage them to develop some critical and analytical skills, as well as skills in using the search engine, alongside a developing knowledge and understanding of the subject field.

- Write an essay on a course issue, drawing on four articles from the electronic resources.
- Compare the use of two different search methods used to collect material for your essay, and comment on their usefulness.
- Describe three criteria that you used to decide on the relevance and applicability of your chosen articles.

Figure 1 Developing students' investigative skills for resource-based learning

The powerful formative effective of assessment is such that it was only when our students were required to prepare for the assignment, that they started to undertake the activities we had recommended. The assignment effectively provided them with an **irresistible opportunity to learn**. It provided useful practice in the basic search tools, which (despite our advice in course materials) they had not previously practised. We believe that students need to learn these basic skills before they can be expected to undertake complex tasks. So the more complex areas of critical and analytical ability will takes time and iterative practice, and their development must surely form part of a coherent strategy to provide practice through the assessment, throughout a course of study.

We have observed a similar effect, while using assessment to encourage online participation by our students. Although some students are all too eager to contribute to online discussion, many will treat this as an optional activity unless it is integrated with the assessment. Some of this reluctance may be attributed to a lack of skills and confidence, the development of which probably needs to be built incrementally into the assessment, but it is also true that our increasingly 'time poor' students will tend to prioritise their efforts on what is assessed.

Summarise an article on a course topic.
Post your summary to online conference.
Discuss the issues arising, in the conference.
Marks for:

- summary of article
- reflection on key points of online debate
- two of your own messages, each supported by a peer's message to illustrate your ability to interact and build on contributions.

Figure 2 Encouraging online participation and development of online collaborative skills

Another way in which students can be helped to develop self direction in learning is to help them to develop an understanding of what we are expecting of them, and what we value in a good piece of work. In other words, they need to develop the ability to judge their own work. One simple approach we have used is to make available on the network the title and assessment criteria for each assignment. Students are then free to discuss their interpretation of the meaning with their tutor, and this can help them to develop their understanding, not only of the requirements of the course, but also of the terms which we use in assessment criteria.

This assignment will be judged on the following criteria.

1 Understanding of key concepts x, y & z.

2 Coherent explanation.

3 Good presentation.

Figure 3 Open assessment criteria: Developing self judgement

Developing this idea further, there is a growing body of evidence that students need help in understanding how to write appropriately for a particular discipline (see for example Lee and Street, 1998). And one critical time when students are particularly looking for help in their writing is at the start of a course, at a stage where they are grappling with a new approach to writing and the requirements of the course assessment.

Message from tutor

'Every course has its own particular approach to writing, and this course is no exception. I have put together a few examples from your scripts to illustrate some approaches to writing. You can learn from each other!

'Remember that you need to set the scene, when planning the introduction to your essay. "Why" is a good question to ask yourself when you are planning an introduction. Then "what" definitions...And "how" you are going to discuss it. These three don't have to come in a particular order. Have a look at the following extracts.'

1 'I am going to look at the situated view of learning as explained in the E211 course and give examples of personal experiences which show the connections between the theory and the practice of this view of learning....'

2 'What is a "learning situation"? Although in this analysis I use terms such as "my personal learning situation" in order to situate my argument within the context of the course, the truth is my "situation" interacting with another student in First Class is not the same as my "situation" reading a course text or authoring hypermedia....'

3 'The screenshot above, taken from a First Class conference, shows some of the participants in this year's E211. Traditional approaches to learning might see us a uniform group of learners, all ready and waiting to be filled with the same information in exactly the same way. A "situated" view of learning would, on the other hand, acknowledge us as individuals, each encountering a different learning situation as a result of our prior experiences and the unique personal context within which we are studying....'

Figure 4 The electronic scrapbook: Developing literacy

Figure 4 illustrates an electronic scrapbook of writing samples, which is drawn from assignments written by students, as a way of using student resources to illustrate a number of teaching points and a range of writing styles. Such a scrapbook might be compiled after assignment submission, as formative feedback to the group. We have tried this approach on a second level education course, using it as the basis for online discussion and dialogue on interpretations of assignment wording or criteria, while at the same time offering some relief to the marker's workload, because instead of giving feedback to individual students, it can be given to the group.

The electronic scrapbook can also be archived and used in subsequent course presentations, where it might seed a discussion as part of assignment preparation, although of course it tends to lose some of its relevance today if the scripts are not from recognisable students.

Students are also concerned to get additional help in the revision period before the final course exam.

We have experimented with sharing old exam questions with them on the network, and having given them the opportunity to attempt their own answers, and share these with the group, we produced an answer written by the lecturer, as a way of illustrating a successful approach. This has been very popular with students who are highly motivated to seek advice at this stage in the course, and it is cost effective, as we have been able to provide formative support to classes of 60-100 students at one time.

Conclusions

Beyond its acknowledged importance in judging the quality of students' work, assessment has powerful formative potential in helping students to learn what we want them to learn. We have argued here that it can be used not only for developing knowledge and understanding, but also to develop the communicative and interpretive skills which are required for e-learning, and which are a necessary part of lifelong learning.

Our illustrations of assessment design have shown how the formative aspects of assignments may form part of feedback, **after** assignment submission, being the point where the student gets feedback on their efforts, thus playing an important role in completing a reflective learning cycle (Kolb, 1984). Alternatively, some innovations may be effective in developing student thinking, and influencing their activity **before** assignment submission. And of course, the summative aspects of assessment can also have a powerful formative effect, for if students know they are to be awarded marks for particular aspects then they will assume that this must be an significant part of the course: this has important implications for the award of marks on the process of learning, if this is what we wish to develop in our students.

Of course, any attempts to develop formative assessment which supports student learning effectively must be balanced against the need for reliability and validity, and costs. We have illustrated here how we have addressed these issues in a variety of ways, for example:

- by the provision of networked formative assessment which can be used with a group of students, rather than to individual students

- by targeting particular points in the course, where support will be particularly valued by students
- by the engagement of student input and debate, so that students can learn from each other
- by offering the opportunity for students to develop self judgement.

References

Biggs J (2003) *Teaching for quality learning at university*, Society for Research into Higher Education/Open University Press, Buckingham

Kolb D (1984) *Experiential learning*, Prentice Hall

Lea M and Street B (1998) *Student writing in higher education: an academic literacies approach*

Macdonald J (2004) *Developing competent e-learners: the role of assessment, Assessment & Evaluation in Higher Education*, 29 (2) 215-226

Macdonald J, Heap N and Mason R (2001) *Have I learnt it? Evaluating skills for resource based study using electronic resources, British Journal of Educational Technology*, 32 (4) 419-434

Macdonald J (2001) *Exploiting online interactivity to enhance assignment development and feedback on an ODL course, Open Learning*, 16 (2) 179-189

Macdonald J and Twining P (2002) *Assessing activity based learning for a networked course, British Journal of Educational Technology*, 33 (5) 605-620

Macdonald J, Weller M and Mason R (2002) *Meeting the assessment demands of networked courses, International Journal on E-learning*, 1 (1) 7-16

Macdonald J (2003) *Assessing online collaborative learning: process and product, Computers and Education*, 40 (4) 377-391

Perry W G (1970) *Forms of intellectual and ethical development in the college years*, Holt Reinhart & Winston, New York

Ensuring quality and efficiency with online assessments

Dr Richard Parsons, Centre for Learning and Teaching, University of Dundee

Introduction

The University of Dundee is making moderate usage of computer aided assessment (CAA) to support learning throughout the University. These assessments are delivered online and are provided to students in both formative and summative modes. Given the importance of assessments and the costs of providing secure and stable online systems, the University has invested in central resources to provide these facilities to all departments. This approach has been important for ensuring the quality and efficiency for the underlying systems, and it has also been useful for promoting good practice within the learning activities that utilise the systems.

Support for online assessment by the central systems can be divided into six parts (Figure 1) and each is discussed below. This represents a systematic approach to supporting the enterprise, but does not include the important activities of staff developing questions and students interacting with the assessments. Specifically the success of using online assessment depends crucially on having quality questions that either stimulate student learning (for formative deployment) or provide accurate determination of student competences (for summative deployment). Creating these questions is an academic activity and this largely lies with the academic staff who may view this as an additional burden and therefore restrict the successful use of the systems.

At the University of Dundee, we have had some excellent success with innovative staff who are using online assessment to produce demonstrable increases in student performance in assessments. Another approach we are evaluating is to provide students with the opportunity to collaborate and develop revision assessments by each submitting a number of questions to a pool, which are then collated and made available in a formative fashion to the students to support their revision. Writing questions for online assessment that address the higher orders of learning - skills, competences, synthesis and analysis is possible, but it requires careful thought and commitment on behalf of the question author. We aim to increase our development and usage of these questions to support 'deeper learning' through online assessment.

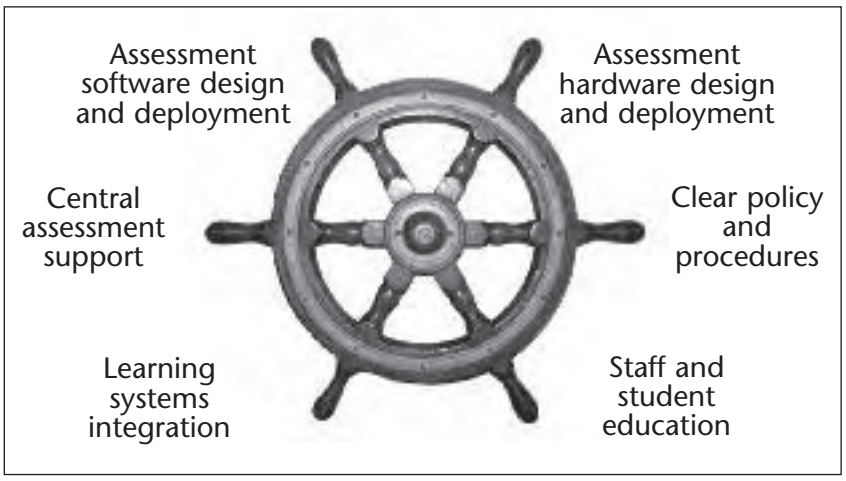


Figure 1 Components of a central online assessment system

Assessment software - design and deployment

The University of Dundee has a central virtual learning environment (VLE) that is available for all departments to use. This is presently a Blackboard version 6.0 system, and while the VLE provides some functionality for supporting online assessments, it also has some limitations when compared to specialist software tools that are devoted entirely to assessment. The assessment tools within the VLE are used by many staff to carry out small scale formative quizzes and they are also used for the end of module evaluations. For staff who wish to provide enhanced, flexible or summative assessments, the University has a full licence for QuestionMark Perception™ (currently version 3.4). Frequently, staff become familiar with the utility of online assessment by using the tools within the Blackboard VLE and then subsequently progress to the more powerful and flexible environment that Questionmark Perception™ provides.

Authoring of questions on the Blackboard system is carried out using the internet, while we generally recommend that users of the Questionmark Perception™ software use the Microsoft Windows client. This system can operate on Macintosh computers using Virtual PC to create a Microsoft Windows environment. Academic authors receive full instruction in the use of Questionmark Perception™ software and online assessment through our staff education programmes (see below). Authors can create assessments and try out these assessments on their local computers, but uploading and configuration on the VLE and CAA servers is an activity carried out only by central staff in a gatekeeper role, again see below. It is important to note that various staff around the campus have widespread experience of CAA systems and some have been using QuestionMark Perception™ for over nine years.

For the students, usage of the two systems is largely indistinguishable. Access to both Blackboard and QuestionMark Perception™ assessments is provided from the VLE login, using a seamless transfer. Again the web interface provided by the QuestionMark Perception™ software is more configurable and can be developed to provide specialist support for special assessments. The marks from the assessments taken in Blackboard or QuestionMark Perception™ are automatically and privately available to the students within their Blackboard gradebook.

Assessment hardware - design and deployment

This is a specialist area and one that will only be briefly introduced here. The server systems that support the VLE and QuestionMark Perception™ systems at the University of Dundee are typical of major information technology systems that are in place at many institutions. The University of Dundee has a policy of supporting Sun-Unix system platforms and Oracle databases for mission critical applications and these are the platforms that underpin the VLE and the QuestionMark Perception™ databases. The Blackboard VLE is a dual server system with data storage managed through a storage area network using redundant array of independent disks five. Two QuestionMark Perception™ application servers are deployed (one predominately for formative assessment delivery, and one for summative assessment delivery), and the formative server is coupled to a Unix Oracle database located on one of the VLE servers. The QuestionMark Perception™ servers run on Microsoft Windows 2000 and use the Internet Information Service Recycle tool at 5am every morning to refresh

their memory usage. They have proved stable and reliable. Development of the production systems has now demanded the creation of test servers for both the VLE and QuestionMark Perception™ systems, and these are utilised to test upgrades, new approaches and new installations, before subsequent deployment on the production systems. A diagram illustrating these server configurations is presented in Figure 2.

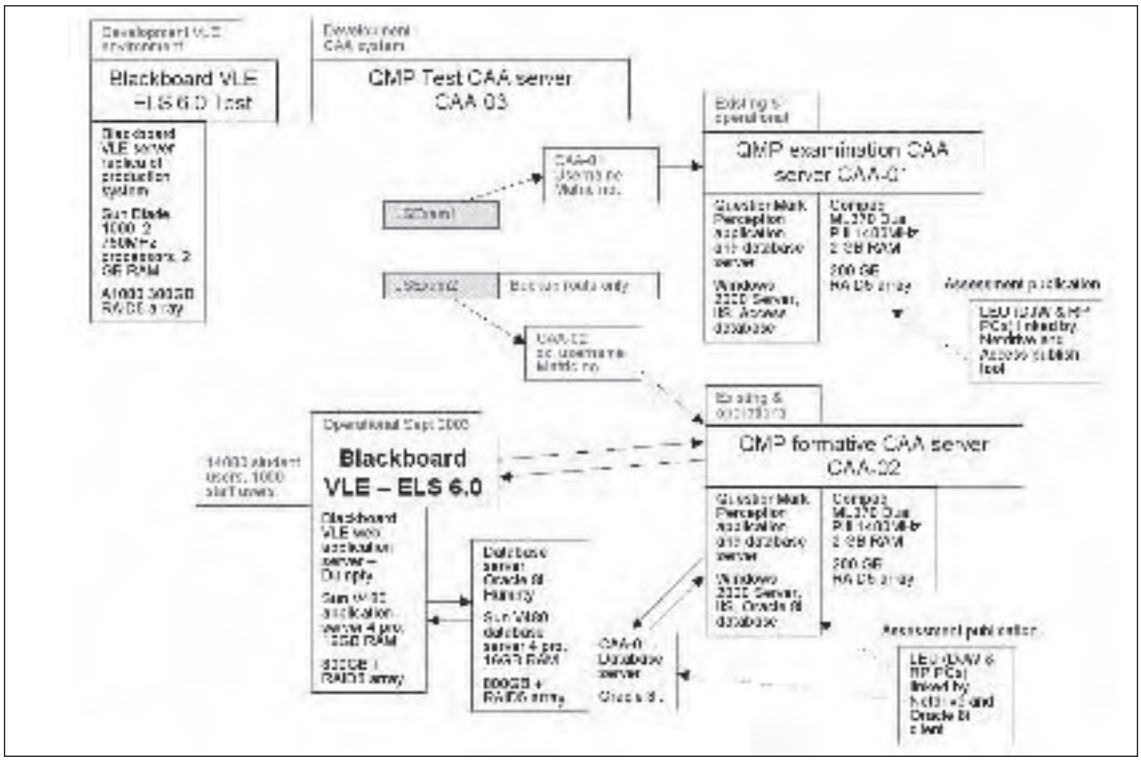


Figure 2 Central CAA server configuration at the University of Dundee (March 2004)

Clear policy and procedures

From the outset of developing a central assessment system, we created a draft policy and procedure document to ensure that a quality system was developed and understood by all. This document is designed to ensure that our systems and operations are compliant with the British Standard BS 7988:2002 - *Code of practice for the use of information technology (IT) in the delivery of assessments*. Our central principles are that assessments conducted with the CAA systems must be fair, reliable, confidential, accurate, secure, accessible and safe. The policy and procedures were originally introduced in draft format and they have now been passed officially at the University of Dundee Senate committee and now represent formal guidelines.

The policy and procedure documentation at the University of Dundee refers predominantly to summative assessments and students must have received formative assessments in a similar format to any summative assessments. While the details are largely relevant to the specific University of Dundee environment, the author is happy to pass on copies of these guidelines to assist others who may be developing their own.

Staff and student education

It is clearly crucial that both the students and staff are comfortable using the online technologies to receive their assessments and to support their learning and teaching.

Students entering the University of Dundee all receive an IT induction session of 90 minutes that includes a brief introduction to the Blackboard VLE and QuestionMark Perception™. They complete a diagnostic assessment which is delivered by QuestionMark Perception™. In addition, all students receive documentation to support their use of the VLE and CAA systems. In modules where the VLE and CAA systems are used, generally one of the lecturers will demonstrate how to access the material during a lecture and support is available from the IT Service's helpdesk. The vast majority of students are very comfortable and familiar with the IT environment and seem to have an expectation that the systems will be integrated and operate in a straightforward and reliable fashion.

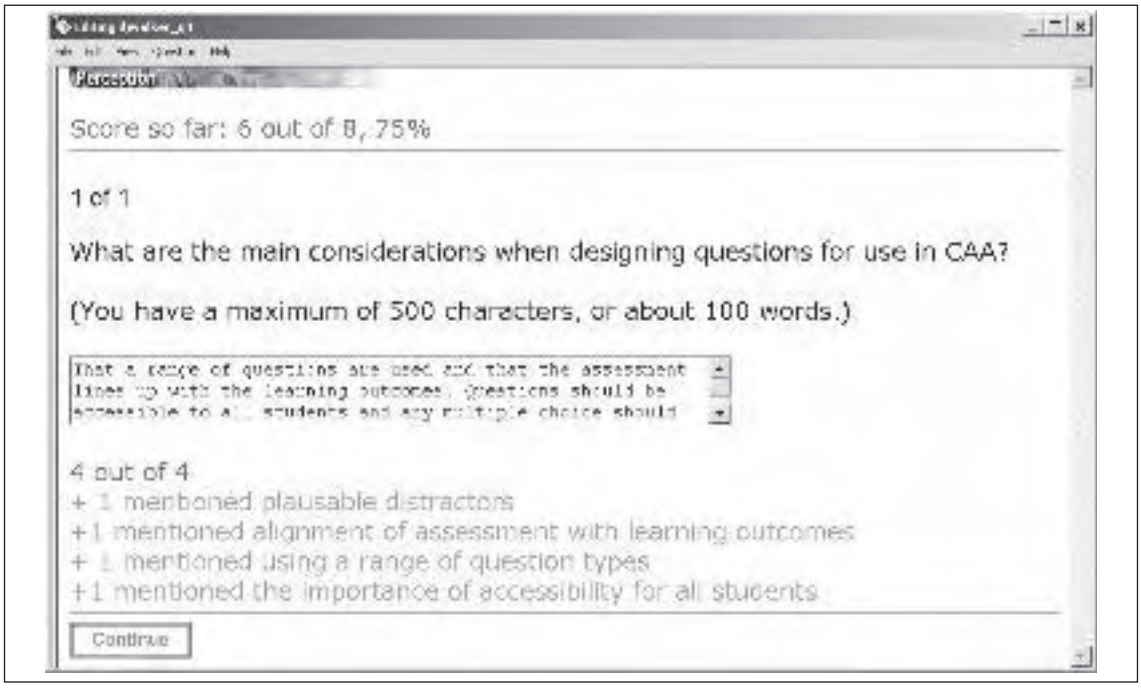


Figure 3 A text-matching question, automatically marked and provided with feedback

Staff at the University of Dundee have a more varied experience of IT and also have a more challenging task to complete as they have to construct the online material and author complex and specialist questions. Staff education is provided by the traditional face-to-face staff development activities, and these are sometimes focussed on a department where the staff have a particular interest in developing online assessment. The most successful staff education activity has been the development of an online course to support CAA and authoring within QuestionMark Perception™. This course (CAA online) runs for five weeks, has been formally accredited at 10 Scottish Credit Accumulation and Transfer points by the Faculty of Education and Social Work and is run biannually. It is a demanding course where the two matters of pedagogy of

online assessments and the technical authoring and delivery of assessments are examined. Typically those completing the course spend about five hours each week on the course material and are then well qualified to lead their students in online assessments. The course is open to staff from other institutions.

With a detailed knowledge of the benefits and limitations of online assessment, staff are free to develop innovative approaches for their teaching. An example of this is our online self and peer assessment system that has been utilised by staff in environmental sciences, medicine, education and accountancy. Using this system lecturers can set a text-based exercise for students to research and submit their answers online. The students are then presented with a set of marking criteria which they apply first to their own answer and then subsequently to two of their peers answers. The marking is completed anonymously and the lecturer can moderate and mark the work themselves if they wish. The results are emailed and texted (Figure 4) to the student author. Feedback from the students is very favourable as they find that they gain a much deeper understanding of the assessment process, particularly realising how criterion-based reference marking operates and how they can successfully predict the marking criteria. This system is available for lecturers to use within and beyond the University of Dundee at www.dundee.ac.uk/learning/ilt/

Figure 4 Online peer assessment of text using software developed at the University of Dundee

Learning systems integration and Dundee specialisations

As VLEs and CAA systems are relatively new, integrating these systems is also a new activity, but one that pays dividends in that the resultant systems can be straightforward and consistent for users. The University of Dundee was one of the first institutions to install the Blackboard 6/Questionmark Perception™ bridge and the utility of this integration has proved worthwhile. Student use of the system is particularly straightforward and they are only challenged once for username/password when they log into the VLE, and all other username and access information is passed automatically between the systems.

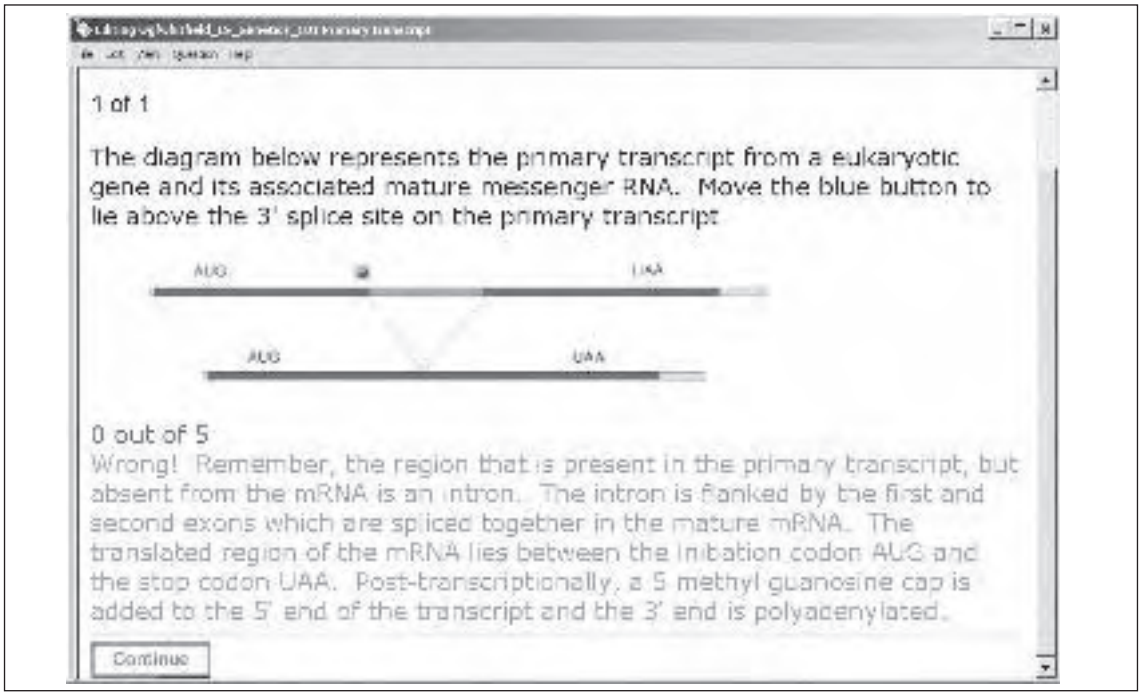


Figure 5 Drag and drop question with visual cues and feedback designed to support learning

Some development work has been carried out to create tools to summarise assessment results in Questionmark Perception™ for subsequent direct upload into the Blackboard VLE. In this way, low stakes assessment marks can be based on any of the maximum, first, mean or final mark obtained in an assessment. This has proved important as it provides for flexibility in the system, allowing lecturers to design the deployment of their assessments as appropriate for the activity being completed. For example, some assessments, where 20 questions are randomly selected from a pool of 400 questions, are open for multiple attempts over a two week period and the maximum mark achieved is taken as the final score.

We have developed templates for the creation of questions in multiple response formats that allow for partial correct scoring and for numerical questions that include random variables within the question. We are developing an automated way of transferring questions with feedback from text to IMS QTI XML that is designed to allow questions to be directly incorporated into assessments. These tools are important to make the question authors task as technically straightforward as possible. They are also useful for allowing us to provide an efficient authoring service should students wish to develop their own revision assessments, or when authors provide questions for conversion to an online format.

Central support for online assessments

From the matters presented above, it is clear that online assessment systems require substantial commitment on behalf of those who wish to use them. At the University of Dundee, a decision was taken to provide these resources centrally, and to make the software, servers, staff education and general support available to all faculties. The

CAA operations are managed by staff in the Centre for Learning and Teaching, who have an academic teaching focus and are supported by IT Services who provide technical expertise. At present a single learning technologist provides specialist support to assessment authors, although this person is part of a group supporting the learning technology systems on campus. On the server and security side, the systems are supported by a team of people who support a large number of server-based systems. No single individual is responsible, but combined, they deal with issues of hardware, networking, operating systems, databases, storage area network, web servers, firewalls and security settings.

The central learning technology group operate in a gatekeeper fashion and are the only people with access rights to upload assessments to the CAA servers. This provides two purposes.

1 It frees the teaching staff of the need to learn about the upload and setup procedures for assessments.

2 It permits a quality assurance check, where general format, scoring and presentation can be double checked before publication to students.

This role of gatekeeper may be seen as restrictive by some, but it ensures the systems are stable and secure for everyone.

Attempting to detail the costs of deploying the VLE and particularly the CAA systems is of course challenging. Total cost accounting is not appropriate, as the academic staff, support staff, network, IT suites and many other infrastructure costs are required for other purposes, as well as for the CAA systems. Direct costs associated with the CAA systems are one staff member, software licences, hardware (lifetime of three years) and other directly attributable support costs. A fair estimate of these costs is about £60,000, representing £4 per student, or 50p per assessment delivered (about 120,000 assessments are delivered each year at present). Undoubtedly the use of CAA has a positive influence on student learning and they provide an efficient mechanism of completing summative assessments across a range of disciplines.

Summary

The example of the deployment of an online assessment system at the University of Dundee described above clearly illustrates that these relatively new learning technologies have a role to play in the efficient presentation of quality assessments. They are certainly not yet appropriate for carrying out all university assessments, but they can play an important role in providing and efficient, flexible and interesting component of the assessment of student learning.

Assessing online - Post workshop report

Dr Mary McCulloch; Dr Hamish Macleod, Centre for Teaching, Learning and Assessment; and Nora Mogey, Media and Learning Technology Service, The University of Edinburgh

Reflections on the workshop

The workshop provided, both in the contributions from the invited speakers and in the discussions in the wider group, a stimulating balance between pragmatic concerns with the maintenance of quality and rigor in assessment practices, and aspirations towards innovation in teaching, learning and assessment. The title of 'Assessing online' was interpreted as widely as possible, such that we were able to concern ourselves with both the application of online techniques to the needs of assessment, and also the emerging challenge of assessing the work that our students carry out online.

The contributors

Gráinne Conole provided a critical analysis of what information and communications technologies (ICT) may have to offer the teaching and learning enterprise, attempting to chart a course between the very real and exciting potentialities and the more expansive hype. She emphasised that there were almost certainly going to be two sides to every story of technological innovation, with potential benefit being balanced against negative, perhaps unanticipated, consequences. Technological developments should not be judged merely by the extent to which they appear to support or undermine existing educational agendas, but rather their transformative potential should be embraced. New technologies change the ways in which information is handled and used, and so will, at least in part, change the nature of what education should be trying to achieve. ICT tools have radically changed the ways in which professionals relate to, and handle, their evidence base, and therefore call for different sorts of skills to be developed in our students, and for different approaches to assessing their progress and success.

Janet Macdonald focused on the increasing self-direction and autonomy that we expect to see in our students and about the consequent shift in the assessment agenda that this new set of expectations demands. Skills that we would previously have hoped to see present in our graduates (of collaboration, communication and interpretation), we now require to see deployed by our students at much earlier stages of their academic development. Acknowledging the important role of assessment in the minds of students to direct and motivate learning, there is a need to consider carefully the alignment between the assessment tasks presented and the evolving learning outcomes sought.

Don Mackenzie provided an analysis of the 'life cycle' of assessment procedures and presented two different models of how the wider deployment of computer-based approaches to assessment might be stimulated and supported at an institutional level. Rigorous assessment involves a number of different stages of activity, from the design and development of the assessment, through its secure and equitable delivery, and the gathering and moderation of results, to the feeding back of the conclusions of the assessment for developmental purposes both to students and to course designers and

teachers. This cycle calls for a range of different skills. Often these skills are assumed to be present within the individual academic, or academic course team, and the assessment activity proceeds as something of a 'cottage industry' within individual courses. This Mackenzie calls the Devolved Tutor Development model. As an alternative to this the Integrated Team Development model is offered, in which is embodied much more explicit division of labour and centrally coordinated support. Course teachers work alongside a central Assessment Support Team that can bring a wide range of domain-specific expertise to the assessment cycle. This team approach can build in many more 'independent quality checks', encourage dissemination of good assessment practices and ensure equity of student experience across courses. The centrally supported model can also help to extend the sophistication of assessment practice, addressing the oft-stated concern that computer-aided approaches can only address the lower order learning outcomes.

Colleen MacLean presented three case studies from the UHI Millennium Institute of how students' online participation is being assessed. On these courses, a small proportion of the overall course mark is set aside to encourage and reward student participation in online discussion activities. The picture which she described is one of evolving practices, highlighting a range of different approaches used to align assessment with the nature of the learning tasks presented to students. A range of different experiences were described, some of which provided support and encouragement for the practice, while others raised concerns. Difficulties included anxieties around the ownership of contributions to what is fundamentally a collaborative activity, and concerns about the validity of the measure of participation being used for assessment purposes. More positively, students appeared to appreciate the fact that engagement that took up a considerable amount to their time and effort was directly rewarded.

Richard Parsons' case study of online assessment from the University of Dundee conveyed something of the complexity of introducing these innovations, and described how his institution has deployed central resources to support staff in this area. The hardware and software resources were described, with particular reference to their usability, security and reliability, and the importance of a clearly articulated central policy was emphasised. A particularly interesting aside was about some courses engaging with students to create assessment items that could then be used by their colleagues for formative or revision purposes. Not only was this beneficial for those who used the items, but it demanded analytic engagement with the topic on the part of the question setter.

Using the technologies

The workshop took two opportunities, beyond the increasingly ubiquitous use of screen projection in support of the presentations, to make use of the technologies on which we were reflecting. The first was to bring one of our keynote speakers, Professor Gráinne Conole, into the programme for the day via a pre-recorded video presentation. Professor Conole was out of the UK at the time of the workshop and initially we feared that she would have been unable to participate. One initial thought was that she might be brought into the day by live video-conference, but it was quickly decided that this would be technically extremely difficult given the planned venue, and altogether too risky from an organisational perspective. We finally settled

for a recorded presentation of a short paper, supported by slides prepared in Microsoft PowerPoint, displayed in synchrony through two data projectors. The enduring impression of the organisers was of Gráinne's 'presence' with us on the day, and this was echoed by informal observations by members of the wider group.

Extending the notion of using the technological affordances to involve colleagues not able to be present on the day with the discussions of the workshop, we also planned to use a web-based online discussion forum to hold a post-workshop event. The hope was that discussions initiated on the day might be extended online for a brief period thereafter. A forum was set up using WebBoard, and announced to participants on the day. In brief, this venture was not as successful as the recorded video presentation. First of all, it became immediately obvious that some colleagues trying to access the discussion appeared unable to do so for technical reasons outwith the immediate control of the server site. We finally concluded that this was probably due to perfectly reasonable security measures in place at the remote sites. The message here is that we must not underestimate the technical difficulties that may be encountered by our students (and potentially our tutors also) when we engage in any but the most straightforward exercises in computer-mediated communications extending across a number of institutional networks, or commercial internet service providers. Further, for those who were technically able to gain access, the stimulus to participate was not great. Again there are lessons here. Online discussions and collaborations are unlikely to blossom without a significant degree of cultivation (Preece, 2000; Salmon, 2000), which is difficult to engineer for a one-off meeting of this sort. If this is true for a group of engaged and motivated colleagues interested in the topic in question, how much more so will it be true for our students for whom their immediate motivation cannot be assumed. If this particular innovation was unsuccessful, at least it can be said to be interestingly and definitively unsuccessful!

Issues of debate and discussion

As the one meeting in the series focusing on computer-based approaches to assessment, the Assessing online workshop inherited all of the general assessment concerns discussed in the other meetings, overlaying them with the special problems and potentials related directly to the medium. The following section summarises the issues and topics that were raised within the breakout discussion groups on the day. While the breakout groups were arranged and briefed to address the issues of assessing online at the local (course and department), institutional (strategic) and technical and pedagogic levels, in practice it was found that the discussions of all the groups eventually ranged across these levels, the prominent barriers and drivers identified frequently relating to interactions between these domains.

One general observation arose, and recurred at many points in the discussion, about the way in which technological innovation is often regarded. While it is of course necessary that we should carefully evaluate and monitor any innovations or changes in teaching and assessment practice, it often feels as if far more is being asked of the new approaches by way of evidence of their efficacy than has ever been demanded of the so called 'tried and tested' methods. In truth, the reliability and validity of conventional methodologies are often assumed based on traditions of use, rather than upon any objective evidence. We use what has always been done as the touchstone

against which to measure what might be done, without questioning the fundamental reliability of customary practices. We will return to this point as it arises in relation to specific concerns explored in the discussion.

Assessing of higher level learning

It is often assumed that computer-assisted assessment (CAA) can efficiently test factual knowledge, but that it cannot assess the analytic and synthetic abilities that one would expect to be concerned with in higher education, particularly in the senior years. It is certainly true in general terms that the assessing of higher level achievements is more difficult than is assessing the retention of factual information, and this will be true in CAA settings too. But Bull and McKenna (2004) are able to provide convincing examples of how CAA questions might address achievement over the range of Bloom's taxonomy (Bloom et al, 1956), and the examples provided by Don Mackenize (this meeting) showed the testing of analytic and evaluative ability.

It may be however, that we are not comparing like with like. It is frequently observed that the use of CAA does not save the time of the teacher, but rather moves the burden of effort from the end of the assessment process to the beginning. CAA demands that we clearly identify the learning outcome, devise a means by which the student might be seen to demonstrate the achievement of that learning outcome, then construct assessment items that can be handled within the particular CAA system in use. At that point, most of the work is done. Conventional approaches require that we formulate a 'good' question or task, then place most of our effort into scrutinising the response which this evokes in the student, and judging the quality of the outcome. One might question, in a parallel way, whether our conventional practices actually test higher order learning, or whether it is more correct to say that conventional assessment methodologies allow students to display these capabilities. A tightly worded essay question in an examination may call for analysis and evaluation on the part of the student, and the highest grades will be awarded to those who rise to that challenge. A less carefully worded question may not explicitly call for an evaluative answer, yet the able student will recognise that that is what is expected.

Returning to the earlier point about the demands we make upon the computer-based approach, there may be a tendency to find the best practices in conventional assessment being compared with the poorer practices from CAA. Good assessment is a demanding business and takes time, whether that time is invested in advance to plan good test items, or at the end to apply professional judgement to the product.

Caution and conservatism

It might seem sometimes that progress towards the wider deployment of CAA has been, and continues to be, slow. There are many stakeholders in the assessment process however, and progress needs to move slowly enough to bring them all along (Bennett, 1998). In the oft-quoted words of Boud, students can avoid bad teaching, but they cannot avoid bad assessment (Boud, 1995). Assessment of performance, and by implication ability and potential, is a high stakes business in the life of the student, and students themselves, as well as parents, politicians and potential employers, need to believe that the procedures are reliable and valid. Innovations in approaches to

teaching and learning can be embraced or avoided, but all students have to pass through the process of assessment. Thus we can all find ourselves more conservative about the matter of assessment than in any other area of our practice. This state of affairs carries with it the attendant risk of increasing the mismatch between teaching and learning practice and assessment practice.

Feedback and formative assessment

It may be that the application of CAA in formative and diagnostic situations will be where teachers begin their engagements with the medium. Formative feedback, where the stakes are lower, may provide the opportunity to try out new approaches, and to gauge their acceptability with the student group. While students are often reluctant to engage seriously with assessments that they perceive 'don't count' to the final grade awarded, the relative informality of the computer-based quiz seems to be something that students find more attractive.

Many discussants looked very positively to CAA to provide new opportunities to enhance students' experiences of feedback and a number of possibilities and potentialities are raised. Increasing the frequency of formative feedback is known to be desirable, helping students to judge their progress, reassure them or spur them on, and thereby contributing to improved retention and progression. Computer-based approaches afford opportunities for self-paced and on-demand self assessment which, when the initial investment in the creation of the item bank has been made, can prove highly sustainable. Many publishers produce significant banks of test items in association with major textbooks in certain subjects, and various subject-specific consortia (for example, based around the Higher Education Academy Subject Centres) have begun to assemble item banks, and stimulate collaboration. Like so many other areas of computer-based learning activity, the pressure of 'not invented here' does seem to apply to such item bank initiatives. Some textbook collections of items are regarded as being of questionable quality, developed largely as adjunctive resources intended to encourage lectures to adopt the book.

The inclusion of different levels of feedback was discussed. Significant value can be added to simple multiple-choice questions by giving students not only feedback about whether their chosen answer was right or wrong, but by providing elaborations on the range of possible answers indicating why a particular answer was incorrect, or suggesting possible misunderstandings which might lead one to choose a particular incorrect answer. When such systems are available, students are often found to go repeatedly thought the tests, trying out the various wrong answers to see what additional useful information they can find. It has also been observed that, even when students could claim a computer to themselves, small groups of students will collaborate around one machine, discussing their predictions and the interpretation of the feedback that the system provides. The test is then acting as a catalyst to promote discussion.

Other potential innovations in practice were suggested, such as the possibility that formative tests might be constructed, possibly by randomised computer selection, from the overall item bank from which the final summative examination will be constructed. In this case, the student motivation would be high, knowing that performance with the formative set of items would be highly likely to predict

performance in the final examination. Analogous to this would be the possibility of students submitting themselves repeatedly for re-examination until they reached an acceptable criterion level of performance before passing on to the next topic. These sorts of models of assessment prompt fears in some that the system might be subverted by students merely remembering the answers to all the items in the bank through repeated exposure. Others argue that this would be a good way to ensure that students covered the syllabus. Fears of this sort might alternatively be addressed by a regime that required students to obtain a particular grade in the formative assessment before being allowed to proceed to the summative stage, which might be of a different form, or be based on items selected from a different bank.

Close to this, and the idea of on-demand assessment, is the possibility that items of computer-based assessment might be tightly embedded alongside learning materials with the aim of cultivating a reflective approach to self-paced learning. There then emerges the possibility of intelligent learning tools that could capture something of a student's transit through a body of learning materials and resources, and thereby contribute to a summative judgement about the success of the learning achieved.

There was also discussion of the evidence about the negative consequences of the close association of the feedback provided on students' work with the grades awarded to them on the basis of that work. Research (for example, Black and Wiliam, 1998) has suggested that feedback provided disconnected from grades is more likely to be influential, and to be used to direct future study. Perhaps CAA approaches might be developed such that feedback can be uncoupled in the minds of the students from the subsequent grade awarded.

Alignment

Following the theme addressed by Janet Macdonald in her presentation, much discussion focussed on the need for there to be correspondence between the ways in which students learn, and the ways in which their achievements are assessed. As mentioned, there was a feeling that innovations in teaching methods might proceed in advance of innovations in assessment methodology, and that this was inconsistent and prejudicial. Students have the right to expect that forms of assessment should be in keeping with the skills and practices that they had developed in their courses, and that where a range of assessment methods were to be used (in itself a desirable thing) that they should have the opportunities, perhaps through participation in formative exercises, to practice with the forms of assessment through which they were ultimately going to be judged.

There is also a set of wider contexts with which our assessment procedures need to be aligned. Prensky (2001) has coined the terms digital natives and digital immigrants to distinguish between those who have grown up in an environment rich in media technologies, and those who have grown up in an earlier era, and for whom ICTs will always be encountered as something slightly alien. The rapid advances in technology over the last two decades mean that the undergraduate body participating in higher education at the moment have grown up in an environment radically different from that of their parents, and even from the younger of the academic staff who are now their teachers. Prensky argues that these changing experiences of media technologies

have led to radical changes in cognitive orientations of modern young learners, calling for corresponding changes in the ways in which their learning should be supported. This calls for new approaches to education, and corresponding developments in assessment.

Similarly, the world of work has been transformed by new technologies, and new clusters of skills will be required. The ways in which students are prepared, and the ways in which the success of that preparation is best assessed, should be undergoing corresponding changes. This would suggest that assessment procedures that require students to work online, with access to computer-mediated resources and tools, will come to have increasing face validity.

Issues of equity

It has been observed that enhancement in the experience for some may result in increased problems for certain particular groups. The classic example would be the arrival of direct manipulation 'windowing' interfaces on microcomputers. This development represented a significant improvement in usability for most people, but greatly complicated the lives of those who were dependent on systems to read and speak the contents of their screens.

The introduction of the computer per se no doubt constitutes a barrier for some, prompting Mark Brosnan (1999) to ask rhetorically 'should computer-based assessment be used at all?' It may be that people who suffer from a degree of anxiety about the use of technology will be less well served when the assessment is carried out using a computer. That is, while the computer-based approach is as valid and reliable as the conventional approach overall, there might yet remain a degree of differential non-equivalence based on particular characteristics of the individual student. It should be noted that there is no strong research evidence that this is indeed a problem, but we should continue to be aware that such test-mode effects might, in some circumstances, arise. It should be assumed, for example, that various stressors would combine to the detriment of student performance in an assessment situation. That which would be of no account when one is confident about the material in hand could prove to be an intolerable distraction when one feels less secure about the topic being assessed. This is, of course, a point of more general relevance than just to CAA. Ambient noise, as when an examination is held in a large, echoing gym hall, can prove particularly distressing for some people (Singleton, 1999).

Increasing costs

Bennett (1998) describes a generational analysis of the potential developments of CAA. We are, he argues, working with first generation approaches at the moment, which amount to little more than using the technology to automate practices already established in some pre-automated form. Like any other instance of the introduction of technology, this is an expensive phase in which resources have to be built up. These resources may be in the form of the necessary equipment to enable sufficient numbers of students to participate, the experience and skills within the body of the academic and support staff to be able to deploy these approaches, and in terms of the banks of test items that have to be created. Good test items are difficult and

time-consuming to create and, at least in the early stages of the introduction of CAA, the item bank needs to be constantly replenished to address the demands of security and equity. All of these areas will require investment, meaning that in the short term at least the introduction of CAA will not bring cost savings. Further, if, as Don Mackenzie and Richard Parsons argued, particular advantages of CAA will be best realised through a system of division of labour between central support and local academic domain knowledge, then new funding and staffing models will have to be developed and introduced. Cost savings will not be brought about by innovation per se, but rather when the innovations become embedded and routine.

Technical concerns

Many saw barriers to the increased use of CAA to locate in relatively basic issues of access to technology. Self-paced formative testing may simply be a special case of online learning with its attendant demands on access to computing equipment for students. Significant use of computers in summative assessments raises the access issue by an order of magnitude. Individual courses may be able to arrange for the setting aside of institutional computing laboratories in which to process large cohorts of students at examination times, but were these approaches to be adopted on a wide scale whole new computerised assessment centres would have to be equipped. Again, we will need to develop new ways of deploying resources such that they can be efficiently used across the course of an academic year, either desynchronising examination demands, or finding other ways to use the facilities between examination periods. Interestingly Bennett (1998) suggests that this might be addressed by new liaisons between academic and commercial organisations.

Assuming that access to the necessary resource can be organised, some feel that moving to CAA exposes us to the risk of serious disruption to examinations due to failures of individual computers, or network components. The truth is that we have been making extensive use of computers in assessment for some time, and have experienced some of the difficulties. Every time a student submits a word-processed essay or dissertation we are participating in CAA, and we know that difficulties can arise which are due to (or at least attributed to) the role of the technology. Richard Parsons' presentation addressed these concerns, emphasising the importance of backup and redundancy (both at the level of individual machines, and also of network components), and of thorough contingency planning. Clarity about the handling of particular contingencies is important not only for the staff involved, but also to reassure students so that the examination does not engender more anxiety in them. It is also important that examination regulations are also framed to take account of technical difficulties, such that serious problems can ultimately be resolved in ways that are seen to be fair and equitable. Reassuringly, Parsons' maintains that no system can logically offer 100 per cent reliability, be it computer-supported or conventional. Those who have had experience of the administration or invigilation of examinations over any period of time will know that conventional examinations are not free of crisis and failure. Here again we need to be careful not to apply more demanding standards to the computer-supported setting than we know to pertain in the conventional setting.

Examination regulations

There is often a feeling that the assessment regulations of our institutions are framed in such a way as to proscribe certain sorts of practices, and thereby to get in the way of innovations such as CAA. It is to be hoped however, that these problems may be more apparent than real. Regulations will be worded to take account of circumstances that are known about and understood, and may have nothing to say about novel or innovative assessment procedures. Where the wording of particular regulations does seem to have inadvertent implications for practices that were not anticipated when the regulations were drafted, mechanisms will exist for the regulations to be reviewed and brought into line with current practice. It is also the case that this review process can provide an important forum in which issues of clarity and equity can be debated; issues that might otherwise be overlooked in the face of the enthusiasm of innovative teachers or course teams.

Student authentication

There are concerns that when students take assessments and examinations online it will be difficult to be sure that the person taking the assessment is who they claim to be. Again, we may be exposing the computer-mediated assessment to more rigorous criticism than is routinely applied to the conventional approaches. Most of us believe that our institutions have procedures in place (like the requirement to display a student photo-ID card in an examination hall) that ensure that the person completing the assessment is the person who is given credit for the achievement. Similar procedures can readily be put in place in computer-mediated examinations, although this would only apply where students are brought to an examination centre of some kind to sit the assessment. Where some form of distance participation is envisioned, the solution may not be as simple. And yet one must concede that when a student on a distance learning course submits an assignment for grading by post or email one does not know that the work was actually carried out by that student.

Some discussion focused on the hypothesis that it might be possible for a computer-mediated assessment to monitor some aspects of the student's interactional style, and thereby to provide some checks as to the identity of the person taking the test. It might be, for example, that the patterns of keystrokes at a keyboard could provide a form of behavioural signature analogous to the information conveyed through handwriting. Were such techniques to be successfully developed the computer-mediated testing situation has the potential to become more secure than a conventionally administered test.

Conclusion

Many of the issues addressed in discussion in the workshop - the relationship between innovation and regulation for example, or the ways in which feedback can be deployed to enhance learning - are not particular to the business of CAA, but arise in consideration of assessment through any medium. Here again however, thinking about the role that technology can play in support of our practices in education, we are given the opportunity to reflect upon what is fundamental to those practices, what it is that we are trying to achieve, what can be changed and what must be preserved at all costs. If the majority of current applications of CAA are based on the use of technology

to automate existing objective testing procedures, the presentations and discussions at the workshop demonstrated the potential of CAA significantly to augment current practice, and to exert a transformative influence on higher education.

References

Bennett R E (1998) *Reinventing assessment; speculations on the future of large-scale educational testing*, Policy Information Centre, Educational Testing Service, Princeton, New Jersey

Black P and Wiliam D (1998) *Inside the Black Box: Raising Standards Through Classroom Assessment*, Department of Education and Professional Studies, King's College London

Bloom B S, Engelhert M D, Furst E J, Hill W H and Krathwohl D R (1956) *Taxonomy of Educational Objectives. Handbook 1 [etc] ... by a committee*, Longmans, Green & Co, New York

Boud D (1995) Assessment and learning: contradictory or complementary? in P Knight, *Assessment for learning in higher education*, pp35-48, Kogan Page, London, in association with the Staff and Educational Development Association

Brosnan M (1999) Computer anxiety in students: should computer-based assessment be used at all? in S Brown, P Race and J Bull, *Computer-assisted assessment in higher education*, pp47-54, Kogan Page, London

Bull J and McKenna C (2004) *Blueprint for computer-assisted assessment*, RoutledgeFalmer, London

Preece J (2000) *Online communities: designing usability, supporting sociability*, John Wiley, New York

Prensky M (2001) *Digital game-based learning*, McGraw-Hill, New York; London

Salmon G (2000) *E-moderating: the key to teaching and learning online*, Sterling, Kogan Page, London

Singleton C H (1999) *Dyslexia in higher education: policy, provision and practice* (Report of the National Working Party on Dyslexia in Higher Education), University of Hull on behalf of the higher education funding councils for England and Scotland

Web links

National Foundation for Educational Research database of links
www.nfer.ac.uk/research-areas/computer-based-assessment/computer-based-assessment-links.cfm

Scottish Centre for Research into On-line Learning and Assessment (SCROLLA)
www.scrolla.ac.uk/

Dr Joanna Bull

On 8 June 2004, the CAA community were shocked to learn of the death of Dr Joanna Bull. Jo was a young woman who has made a major contribution to the field of CAA in her short career. After completing her PhD, Jo became a member of the Higher Education Staff Development Agency team at Sheffield in the early part of her career. She took up a post at Luton University in the middle 1990s where she was integrally involved, along with colleagues, in establishing the CAA Centre there. Jo coordinated a practical demonstration of what CAA could deliver across a university campus. She wrote many articles and books with the *Blueprint for Computer-Assisted Assessment* (2004), co-authored with Colleen McKenna, as perhaps her most prominent recent work.

Jo experienced a course of treatment for cancer over the last three months of her life but typically continued to work as an independent consultant in the field with many of her colleagues unaware of the fight in which she was engaged. On 16 June 2004, a SCROLLA meeting entitled Assessment Futures anticipated a presentation from Jo on her recent work in Scotland. Over the last couple of years, Jo had been project managing a pilot programme for the delivery of the National Assessment Bank of questions into Scottish secondary schools in PASS-IT (www.pass-it.org.uk). The audience at 16 June 2004 meeting were shocked and saddened to learn of Jo's recent death. In her role as project manager she exhibited all those talents which friends and colleagues will recall with fondness: she would encourage without rancour, anticipate to eliminate problems and all with humour, thoughtfulness and a smile. CAA has lost an intelligent, inspiring and energetic worker and many have lost a good friend.

This tribute was written by Professor Cliff Beevers, Director of the Computer Aided Learning in Mathematics Project and co-director of the SCROLLA.

Assessment workshop series - No 6

Issues of validity, reliability and fairness

Contents **Page**

Issues of validity, reliability and fairness - An overview

Pamela Flanagan, Royal Scottish Academy for Music and Drama and Workshop Director

So why did we have this conversation?

'Assessment is the sharp end of learning' (Race 2001).

In the overview paper for the workshop, Pamela Flanagan posed the question as to why such a conversation on a large scale should be taking place, in view of the breadth and depth and sheer quantity of research and consequent literature on the issue of assessment, particularly within the last 25 years. So much of it focuses again and again on similar topics, particularly the need for change, the identification, addition and closer involvement of newer stakeholders in the process such as parents, peers, employers; the movement towards a more holistic view of the those being assessed; the increasing recognition of the usefulness of formative as opposed to summative assessment; changes which are or which should be happening; and advice on how to effect change etc. There are so many definitions of what assessment is or should be that one could be forgiven for feeling slightly overwhelmed. Many have contributed to the debate over the years with advice/recommendations and examples for change and progress. Sally Brown identified some common strands some years ago which might have given some grounds for optimism that the nature of assessment might be about to change:

> 'the broader concept of assessment, in terms of its purposes and closer integration within teaching and learning; the increasing range of young people's achievement and qualities towards which assessment is directed in both traditional and new contexts; the move towards assessment which emphasises description rather than comparative judgements; the changes in ideas about who should have responsibility for making the assessments; and the emerging recognition that formal certification should not be restricted to the privileged few' (Brown, 1988a),

and again

> 'Our ideas about the process of assessment itself, however, have undergone even more major reforms…Assessment is no longer seen as an end in itself; it has to earn its keep by contributing to the effectiveness of the education and training which is offered to young people. And it is that criterion on which it should be judged' (Brown 1988b).

However, more recent observations, not least from Brown herself, would appear to suggest that many of the issues confronting assessment stakeholders then are still with us now; if there has been change it has occurred slowly:

> 'Ensuring that assessment is fair, accurate and comprehensive - and yet manageable for those doing it - is a major challenge. It is a challenge which has been grappled with by many…but one which, however, often has to be tackled in relative isolation. Despite the fact that there is a considerable body of international research about assessment and related issues, we experiment largely in ignorance of the way others have affected positive change, and we have limited opportunity to learn from the lessons of others' (Brown S and Glasner A, 1999).

> 'We have moved at warp speed in developing models for classroom assessment; but…we have not followed suit in developing assessment models that accommodate learning styles' (Anderson, 2001).
>
> 'Desirable though reliability is, achieving it is another matter' (Knight, 2002).

One major issue, which is not necessarily new, is the question of who is really being assessed. Is it students, or teachers, or institutions, or government? The fear or pressure of a suspected hidden agenda might be one reason as to why the process of change has been slow to come about; constraints imposed by lack of resources may be another; pressure exerted by conflicting expectations of different groups of stakeholders yet another. Again there have been growing indications of this in the literature:

> 'There is one further point which should perhaps concern us, and that is **who** is the examination for? Are some of us more worried about ourselves than about our pupils?...Is it important to be able to measure the achievement of our students in terms readily acceptable to others…because this gives us reassurance about the value and the success of what we do?' (Paynter, 1982).

Though it has also been acknowledged that occasionally the best laid plans go awry despite one's best efforts:

> 'Although I wanted my students to understand the content I was teaching, the need for them to be able to cope with (and succeed in) the forms of assessment they would face at the end of the year eventually influenced their view of learning and their understanding of what was "important" to learn, which also inevitably affected how they learnt' (Loughran 1997),

reflecting, perhaps, the frustration of coping with a syllabus in which the teacher has not always had a role in assembling, and the perennial combat with the need to 'get through' or 'get them through' the summative assessment that is the end of course exams.

Sometimes the 'how to' is the main barrier, particularly when a whole range of diverse assessment methods is involved:

> 'The main source of the deficiency is our own ignorance about how to do the job properly. Teachers in higher education frequently assess as amateurs when the task demands grave professionalism' (Ramsden, 1992a).

It is clear from the above that there are no easy answers to the issues of who is being assessed, and the reasons for the apparent slow progress in change, to date. In Scotland, however, the higher education sector is now embracing a new enhancement-led approach to quality that aims to improve the student experience, and enhance and encourage innovation in learning and teaching. As such, the workshop offered a significant opportunity for practitioners from a variety of different backgrounds and disciplines to come together, share current good practice and develop new ideas and possible ways forward for the future to facilitate real change, in essence the basis, hopefully, for a type of 'improvement conversation' (Knight, 2001a).

The timing of the workshop was particularly apposite for the first keynote speaker, Linda Suskie (Towson University, Maryland), whose most recent book, *Assessing Student Learning: A Common Sense Guide,* has just been published. In her address, she highlighted six principles of good practice in relation to assessment which cover the

areas of usefulness, accuracy and truth, fairness, ethics, regularity/review and cost effectiveness. She also provided advice and recommendations on creating and implementing the best possible assessment culture in today's modern environment, challenging her audience to reflect on their own particular situation and 'to make the fair appropriate use of assessments ubiquitous' (Suskie, 2000).

> The main criticism of demonstrable objectivity (and the keyword here is demonstrable) is that, in the formation of the assessment model, the need to demonstrate objectivity is frequently of such overriding concern that it is applied to the detriment of all other considerations (Spruce, 1996).

This viewpoint was echoed by Professor David Lines (The Robert Gordon University) who argued that two cultures of assessment, described as being 'assessment of education' and 'assessment **for** education' respectively, have now been created, the first being largely summative, the second formative. However, these do not sit equally side by side, the former having become more and more dominant, virtually excluding the latter, with a consequent distortion of both learning and teaching, and raising questions concerning the validity, fairness and reliability of assessment instruments. Stating that it was time for a review of the purposes of assessment, Professor Lines proposed a way forward with recommendations for the reconciliation of the two cultures that aligns them with, and integrates them fully into, the learning and teaching process, in a manner that satisfies, and is fair to, all stakeholders thus creating the 'powerful learning environment'.

> And no unbiased study of the written machinery of assessment procedures could fail to conclude that we think that students are at heart plagiarists and cheats (Ramsden, 1992b).

Another issue which has become increasingly prominent is the desire to succeed almost at any cost, though here again students are not the only people this affects:

> In many subjects plagiarism says more about the quality of thinking than it does about students' moral failures. Plagiarism often shows students responding intelligently to teachers' slack assessment practices (Knight, 2001b).

On the topic of plagiarism, Jude Carroll (Oxford Centre for Staff and Learning Development, Oxford Brookes University) shared her own considerable experience of observing institutions and individual lecturers, under the headings of fair assessment, transparency, consistency, and natural justice and focusing very firmly on the responsibility of the institutions and lecturers as well as the students in this regard. The issues of communication and knowledge loomed very strongly here; how are students told what is expected of them and how far is their trespass the result of the lecturers'/institutions' failure to communicate this adequately and explicitly to them rather than automatically assuming the desire to cheat? How are lecturers told/trained to communicate this information, and, in the case of the truly dishonest student, not shy away from confronting this practice but respond with appropriate sanctions which are fair to all? What can be done to keep pace with the changing educational climate which is becoming more and more electronically based, and which now presents a different set of challenges in the detection of dishonesty, but without descending into a battle of wits between lecturer and student? Jude Carroll suggested

appropriate strategies for crediting students for their own work and applying sanctions for that which is not, which can be easily and effectively implemented by both lecturers and institutions alike.

'Good pedagogic practice tends to be inclusive practice' (McCarthy and Hurst, 2001).

Perhaps the most difficult and challenging area of assessment was highlighted by Karen Robson (University of Wales Institute, Cardiff), who in her lecture, Assessment - The final frontier, considered the whole area of special needs in which both legislation and research is still developing and where educational programmes, both mainstream and specialist are still evolving. It is challenging not least because of the sensitivities involved (not necessarily restricted to those with special needs) and difficult because of the need for a sea-change in perception, in motivation, and willingness to adapt without necessarily compromising standards. It has been acknowledged that there are now more declared instances of students with special needs than ever before, a tribute, perhaps, to the success of the research and innovative educational practices resulting from that research, which is encouraging many more to come forward and to have confidence that an educational system will serve them well. And yet again there are problems of communication and knowledge here; though many fine studies have been carried out in the last decade or so (notably the Teachability project in Scotland, spearheaded by the University of Strathclyde) the fact remains that many misconceptions and misunderstandings abound, a situation which often the wording of the legislation does little to clarify (what constitutes 'reasonable adjustment', for example). The problem of integration at higher level education has its roots in the school system, where the separateness of both mainstream and special needs is still a problem, and where many teachers still lack the knowledge or skills necessary to cope with such integration, never mind evolving a system of assessment which is fair to all. As Allan, Brown and Riddell observe:

> 'Training for staff in mainstream and special needs schools is clearly a major priority for the future if pupils are to be supported effectively in both settings' (Allan, Brown and Riddell, 1995).

The same is no less true for those in the higher education sector. Taking the recent amendment to the *Disability Discrimination Act 1995* (Part IV - *Special Educational Needs and Disability Act*) Karen Robson considered the issue that while an abundance of materials continue to be produced to assist staff in making their curriculum more accessible, much remains to be done in the area of assessment to ensure that:

a disabled students can demonstrate their knowledge

b staff have confidence both in the adjustments necessary for, and in their own ability to conduct, assessment of these students and

c the test nonetheless satisfy the requirements of reliability, validity, fairness and standard that are applicable to and inclusive of all students.

Throughout all of the deliberations on the above issues however, it was suggested that all should perhaps keep in mind their role not as assessors, but more as 'sensitive critics' as Swanwick (1988) terms it, and not lose sight of the fact that: 'Though accountability matters, learning still matters most' (Angelo, 1999).

It is perhaps concern for this last that is probably the main reason why so many delegates turned up to have this conversation.

References

Allan J, Brown S and Riddell S (1995) *Provision for Special Educational Needs: Current Trends and Future Prospects*, (Interchange) Vol 38, p6, The Scottish Office: Education and Industry Department, Research and Intelligence Unit, Edinburgh

Anderson J (2001) Tailoring Assessment to Student Learning Styles - A Model for Diverse Populations, *American Association for Higher Education,* Vol 53, no 7, 4

Angelo T A (1999) Doing Assessment As If Learning Matters Most, *American Association for Higher Education*, Vol 51, no 9, 10

Brown S (ed) (1988) *Assessment - A Changing Practice* (*Professional Issues in Education*) Vol 2, Scottish Academy Press, Edinburgh

Brown S and Glasner A (eds) (1999) *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*, Society for Research into Higher Education/Open University Press, Buckingham

Knight P T (2001a) *A Briefing on Key Concepts: Formative and Summative, Criteria and Norm-referenced Assessment,* Vol 7, Learning and Teaching Support Network, York

Knight PT (2002) *Being a Teacher in Higher Education*, Society for Research into Higher Education/Open University Press, Buckingham

Loughran J (1997) Teaching about Teaching: Principles and Practice, in Loughran J and Russell T (eds) *Teaching about Teaching: Purpose, Passion and Pedagogy in Teacher Education*, p57, Falmer Press, London

McCarthy D and Hurst A (2001) *A Briefing on Assessing Disabled Students,* Assessment series No 8, Learning and Teaching Support Network, York

Paynter J (1982) *Music in the Secondary School Curriculum*, Cambridge University Press, Cambridge

Race P (2001) *A Briefing on Self, Peer and Group Assessment*, Assessment Series No 9, Learning and Teaching Support Network, York

Ramsden P (1992a) *Learning to Teach in Higher Education*, Routledge, London

Scottish Higher Education Funding Council (2000) *Teachability: Creating an Accessible Curriculum for Students with Disabilities*, University of Strathclyde, Glasgow (email: a.simpson@mis.strath.ac.uk)

Spruce G (1996) Assessment in the arts - issues of objectivity, in Spruce G (ed) *Teaching Music*, Routledge/The Open University, London

Suskie L (2000) Fair Assessment Practices: Giving Students Equitable Opportunities to Demonstrate Learning, *AAHE Bulletin*, Vol 52, no 9, 7-9

Swanwick K (1988) *Music, Mind and Education*, Routledge, London, 150

Fair assessment, fair policing and fair punishment: building on reliability and validity

Jude Carroll, Oxford Centre for Staff and Learning Development, Oxford Brookes University

Fair assessment assumes good practice in assessment as a whole through measures described by others at this event. At the level of individual teachers, good practice means designing courses that link assessment, learning outcomes and teaching methods. Assessment tasks are described via clear, specific briefs that explain what the student must do and how it will be judged then students are taught the skills they will need to do those tasks. This kind of practice makes it less likely that students see assessment as arbitrary or even contrary to their needs as learners. Once they have submitted, if teachers spend time and effort marking what the students have produced reliably, then students may even read the feedback and see their teachers' judgments as worthy of notice. The chances that this will happen go up if the feedback is timely, legible, and includes comments linked to the assessment criteria. The literature about assessment, much of it reviewed at this event, confirm the practices that encourage students to see assessment as a valuable and integral part of their own learning.

At the programme level, fair assessment rests upon good practice as well. This includes such measures as increasing the range and variety of assessment methods and ensuring there is not too heavy an assessment burden. Good induction programmes inform students of how their programme will work, including assessment matters and early diagnostic exercises to identify those needing additional help.

At the institutional level, fair assessment requires policies and procedures that explain, police and uphold the rules in ways that treat students equitably, consistently and transparently. At this level, good practice relies on rules being clear, on efforts to enforce the rules, and on procedures that ensure that students have had ample opportunity through the measures listed in previous paragraphs to learn what their responsibilities are.

However, ensuring students are assessed fairly means moving beyond good practice to address other issues such as how consistently, speedily and transparently students are treated. This paper addresses the extra demands fairness sets for teachers, programmes and institutions. In summary, it discusses how students are helped to understand and use the rules and requirements that study at higher education level puts upon them, including those who need additional help to understand and use academic conventions and skills. It discusses how students' skills evolve over time and what helps and hinders that development. Some students will not follow the rules, either intentionally or through misunderstanding or misapplication of complex academic conventions. Fairness addresses how these students should be handled. For example, students accused of misconduct should not face undue delays before the case is resolved. Students who adhere to rules and regulations should not feel that their experience is undermined or threatened by those who do not. Those called upon to enforce rules need to feel they are being asked to act reasonably, fairly and free from over onerous burdens. These are not easy requirements but they are achievable if assessors, programmes and institutions move beyond the good practice recommendations already cited to address the extra demands of fairness.

Transparency

As already stated, unless students know what they must do and how it will be judged, they have difficulty accepting the assessment as fair. However, real transparency is difficult to achieve because spelling out in some detail what the student must do nevertheless also contains tacit and implicit requirements. For example, if an assessment brief reminds the student to 'make sure your essay includes a wide range of sources to support your argument', this appears to be explicit. But looked at more closely, the brief assumes the student is familiar with a particular discourse style (the academic essay) and many are not. Indeed, almost all students are not skilled essay writers when they start their undergraduate study and a few are not when starting postgraduate study. The instruction assumes the student sees learning as shaping and supporting an argument though many have never encountered this idea in previous study. Finally, it assumes that words like wide and range mean the same thing to both the teacher setting the task and the student reading the brief. Indeed, all three points may be mysterious to students who qualified via A levels and probably did not in text citation. It probably baffles those who recently returned from practice or other work and may never have written an essay. Those who are dyslexic may know what is required but lack the skills to provide it. And students recently arriving in the UK from Jordan or Korea or Greece may never have encountered a writing task that is done independently and certainly none that involve arguments and use of a number of texts. Fairness means helping all these students learn the **new** academic skills they need in as painless a way as possible and as quickly as possible.

Sometimes, it is difficult for academics to accept there is a UK-specific and higher education-specific set of assumptions and values to be learned. Academic culture is so much a part of how lecturers think and teach that it becomes (paradoxically) both invisible to themselves, classed as simply 'normal', and assumed to be obvious to others. In fact, they are asking students to become proficient in a complex, arbitrary and culturally specific way of writing which many have called 'the academic game'. Unless students learn the expectations, beliefs and values that their teachers espouse (as well as learning subject-specific information or skills) they will not be successful. Some students seem able to read implicit messages but many only discover what the teacher wants by submitting work that follows old ways but in this new context, that 'breaks the rules'. When this happens, students learn by paying attention to the feedback and making sense of it. This is not easy. Perhaps they loose marks with no explanation, in which case, they may not notice the cue to learning. Perhaps they receive opaque feedback ('Where are your references?'; 'Is this your own words?') and have to guess what this means. Without knowing how it **should** be done, they might be criticised ('This argument lacks structure') and try and decode the message even though UK higher education means something different by structure than other higher education systems. All students find such lessons painful and some, in short programmes or programmes with a large amount of compulsory work, may be unable to recover their good standing when they finally grasp the lesson.

Apprenticeship: yes; ignoring and hoping: no

Fairness suggests a kind of apprenticeship to allow all students time to set aside methods and beliefs that have previously served them well as students and adopt new

ones. Often, staff accept this but adopt strategies such as ignoring inappropriate activity in the early days of the programme or hinting at acceptable behaviour. For example, a teacher might:

- write 'ref?' four times in the margin of a first year essay that has no in-text citations then give it a good mark
- mark the first 2,000 words only (ie the stated limit) when the student hands in 3,500 words
- deduct five marks for poor referencing in a year two student's essay
- tell two students they colluded on work worth 60 marks so they will get 30 marks each.

Some students will take the hint but many might not notice or they might draw the wrong conclusion.

- 'Ah, references are optional because you can get good marks without them.'
- 'Here, hard work does not get more marks' (with hard work defined as more words).
- 'Ok, referencing isn't very important.'
- 'Ah, it's OK to copy as long as you do really good work so your half is a pass.'

Academic expectations and conventions are best learned through early diagnostic activities and safe practice through tasks that focus on important issues like academic writing, reading analytically and using appropriate ways to structure information. Such an approach need not be overly demanding of teachers' time if the task is small and if peer feedback or group feedback is used. The teacher also needs to design into the course ways to ensure students attend to the feedback (and without these measures many would not) such as:

- requiring students who failed a diagnostic test or did not do a diagnostic task adequately to do a retake in a set time
- not marking subsequent work unless students successfully retake early diagnostic tasks
- making 25 per cent of the student's final mark for the module dependent on showing they have used early feedback to modify their subsequent efforts.
- making 25 per cent of their module mark dependent on giving one or two fellow students correct advice about citation and attribution, using texts.

Of course, many students will find even these measures insufficient. These students would need extra help to meet these requirements; most would need to be taught skills such as following citation rules, using authoritative sources, expressing personal opinions correctly, providing analytical reviews and so forth. All would need written guidance. They would have to practice using the new skills and receive tailored feedback. The key point is that their apprenticeship must not be spent with academics passively waiting and hoping as this is not fair. Some students will pick things up, many won't, and a few cannot without extra help.

Fair treatment of plagiarism

Requirements about transparency and practice apply to a whole range of academic skills but are especially pertinent given the upsurge in punishments for student plagiarism. Parks (2003) notes that information on the incidence of plagiarism is inconsistent, making it hard to judge 'the scale and nature of the problem, the extent to which it changes through time or varies from country to country, from subject to subject or between undergraduates and postgraduates'. Despite this variation, Parks concludes plagiarism is becoming 'more common and more widespread'. Others agree and focus specifically on the serious, potentially fraudulent end of the spectrum. A large US study (McCabe, 2003) found that in 1999, 13 per cent of US students said they regularly cut and pasted from the internet without attribution and in 2003, 41 per cent said they did so. An Australian study (CAVAL, 2002) used electronic detection software to check students' use of web sources and found that nearly nine per cent of student coursework contained more than 25 per cent of unattributed web-based material. By extrapolating from such studies, by listening to the worries of hundreds of academics and by keeping in mind students' own complaints about others' cheating, I believe it is possible to make rough guesses as to what an academic might expect to find in their students' coursework. I suggest that academic anticipate that about 10 per cent of their students' coursework will not comply with rules of academic citation and attribution to such an extent that it warrants **attention beyond normal assessment**. Of course, some tasks will offer little or no chances for students to claim credit for work they did not do. Others will have much larger percentages of students who set out to trick the assessor into thinking that what they submit is work they themselves have done rather than work they have lifted, purchased, downloaded or copied. I recognise that the rough 'look out for about 10 per cent' suggestion is a hostage to fortune, perhaps opening the way for accusations of being punitive, harsh or over-zealous. It also looks unrealistic as I know of no institution where this level of activity in defence of academic conventions and citation rules can be shown. In most, it is likely that many cases, even serious ones, are overlooked because assessors missed the signs or turned a blind eye or used implicit strategies to inform the student of the rules. Fairness rests on punishing students who do not follow the rules appropriately and on ensuring that students who do follow the rules feel their efforts are valued and rewarded.

Unfairness of under-detection

Fairness is more likely in situations where detection is assumed to be a normal part of an assessor's duties, where detection skills are shared among colleagues or where reporting is encouraged and dealt with quickly as a normal part of academic life. Under-detection and under-reporting of plagiarism leads to unfairness through practices such as:

- academics ignoring plagiarism until the third year then going for severe penalties
- using stricter criteria for high-status work than for run-of-the-mill coursework
- ignoring plagiarism in some students on a programme but not others
- treating collusion differently in different disciplines
- having 80 per cent of all cases reported from five of 40 modules in a school
- students developing informal 'swap shops' for recycling last year's coursework

- two academics punishing identical misconduct differently (one choosing an informal chat and the other, zero for the module)
- all cases being passed to a head of department who does nothing for six months then opts for no action.

Inconsistencies and unfairness like these occur because policies and procedures for dealing with plagiarism in most universities and colleges are not used consistently or not used at all. The two main reasons for this occurring are the institution's inability to deal efficiently with a large number of cases and disincentives on detection due to the impact on the person who spots it. Under-detection makes sense if the alternative is devoting time, energy and attention to something that the institution does not obviously value and might not support if the student contests the charge. Fairness rests on institutions seeking to make the detection more effective, on ensuring the subsequent action is fair, consistent and transparent; and on protecting the detector from the consequences of their actions.

Effective detection of plagiarism

As already stated, ensuring students understand and use academic conventions is vital to helping the majority avoid plagiarism. However, assessment would be fairer if detection of those who cheated was less random and less accidental than is currently the case. Detection rates would rise if colleagues shared skills about what signs to look for and if more universities and colleges used detection tools more frequently and judiciously. As things stand, most instances of student misconduct are found by looking out for signs in the student's work with the most commonly used indicator being a change in the student's writing style, type of discourse or use of language. Other signs range from the blatant to the obscure. At the most extreme, staff report encountering 'smoking guns' such as urls left at the top and bottom of the page or even a Tipexed name of the original author with the new 'author' written in biro below. Sometimes the clue lies in the formatting of the document with fonts wandering for no reason, hyperlinks left in grey and paragraphs moving between double and single spacing. Bibliographies are a good source of clues through noting:

- mixed bibliographies eg half Chicago, half Harvard
- dated bibliographies with nothing more recent than 1999
- a list of obscure texts or texts not in the local library
- in-text citations not included in the reference list.

The language the student uses can sometimes trigger questions. Perhaps there are anachronisms ('…now, at the end of the millennium') or unexpected words ('ineffable') or American spelling. Academics report being uneasy if a large piece of work appears with no evidence of how it was made such as a dissertation arriving on your desk without supervision. Sometimes, they are suspicious if the level of the work suddenly changes, especially if it changes markedly for the better in what some call the 'end of year miracle' for their students who are most at risk of failure. And there are those moments when you read something that is strangely familiar, sometimes because you yourself wrote it.

Electronic detection and authentication

An increasingly common way that plagiarism can come to light is by using tools designed to reveal it. Reactive electronic detection (that is, checking an individual case after suspicions are aroused) is most frequently done using Google's Advanced Search facility, although it is possible to find things via other metasearch engines such as Yahoo if Google is not productive. Proactive screening of student work (ie to see what, if any, plagiarism it might contain before it is assessed) is probably best done using the JISC-sponsored Plagiarism Advisory Service. Their website (www.jiscpas.ac.uk) hosts access to the detection tool, iParadigms. The service is available to all higher education institutions in the UK at no cost though this will change in 2005 when a graduated tariff will be introduced. Teachers using this tool can upload all the students' scripts and receive a report of its similarity to websites and to material held on two databases. The first database holds previously submitted work and the other, a growing range of textbooks, journals and subscription resources including lecturers' own notes and publications if they wish to have them included for checking against student submissions. (Note: Students must agree to their personal information being held on the first database before this service can be used.) After a slow start, use by UK higher education institutions is rising rapidly. A third electronic tool that many institutions find useful is designed to identify collusion across a student cohort is called CopyCatch (www.copycatchgold.com/copycatchesreview.htm).

Electronic tools are a significant addition to the range of detection available and can be used fairly if good practice underpins assessment as a whole and if staff are trained to use the electronic detection tools well. However, they cannot help with plagiarism arising from ghost writing, prewritten essays bought off the internet or the use of translation programmes. These fraudulent actions may come to light using the 'eagle eye' tactics already described or may become causes for suspicion if authentication exercises bring anomalies to light. Suspicions might be raised by a significant difference between exam performance and coursework, or by interviewing the student as a normal part of assessment. Alternatively, suspicions can be followed up by interviews about the writing, vivas as to the content, or by asking to see drafts and main sources.

Strategies and suggestions such as those in the last few paragraphs increase the frequency of spotting and pursuing plagiarism cases and when used widely, lessen the random or accidental aspects of identifying cheating. They are part of the support for the majority of students who do not cheat. More academics would use them if they were aware of them, if their colleagues were doing the same, if the department was able to foster a 'no blind eyes' culture, and if assessors were confident that they would not suffer personally from doing so.

Fair assessment at institutional level

Many colleges and universities are now dealing with (or should be dealing with, if my earlier guesses are correct) a relatively high-volume of relatively minor cases plus a growing minority of increasingly severe ones. In my own institution, the rates are roughly 80/20. This double challenge - rising numbers and rising severity - needs careful thought. On the one hand, a light touch is needed to deliver fast verdicts to a high volume of cases; on the other, decisions may need to withstand robust challenge by

students, students' advocates, legal representatives and, often, the media. If the decisions are too centralised (for example, requiring all cases to be handled by the Dean or even in some cases, the Vice Chancellor), they become so slow and bureaucratic that staff will not use them. If they are too devolved, usually to the level of individual academics, this has the same result: staff won't use them. Linking of detection and dealing with the results encourages turning of blind eyes by the potential 'spotter', inconsistent treatment once the accusation is made and accepted as true, no recording of outcomes and, should the outcome be challenged, difficulty in defending the result.

An alternative is being used in many institutions, including my own, where in 2000, we appointed specialist officers (in other institutions, this role is taken by specialist small panels) based in schools and departments. These officers would have a range of other duties besides academic misconduct but all would be allocated time for the role (usually between 60 and 150 hours per year). They are called Academic Conduct Officers (ACOs) and we now have one or two in each of eight academic schools. ACOs are authorised to award penalties from a restricted list of five following a face-to-face interview with the student. Penalties range from recording the fact that a conversation occurred to reduction in marks, resubmission of a corrected piece of work for a capped pass, zero for the piece of work and zero for the module. If the ACO is asked to handle a case where, if the accusation was accepted as correct, it would warrant a more severe penalty, the ACO could pass the case to a central disciplinary panel where the process would unfold in ways described in most institutional policies. The ACO informs the student of their decision, either of a punishment or a referral, at the meeting. The student can accept if a punishment is awarded or ask for a disciplinary panel to be convened.

The strength of the ACO role is that all decisions are recorded; the bulk of cases can be handled soon after they are reported, academics can alert ACOs of cases without the fear of burdening themselves with the consequences, and ACOs encounter enough cases to develop expertise in allocating suitable and fair punishments. ACOs decisions can be monitored centrally and a search for consensus and consistency between programmes, while not easy, becomes possible. To encourage cross-university consistency, ACOs can be asked to meet regularly, discuss cases and explain the grounds for their decisions.

Institutions that have not yet addressed these issues will find help on the Plagiarism Advisory Service site (www.jiscpas.ac.uk). The Service offers guidance to institutions wishing to adapt their current procedures or to audit the way they are handling plagiarism. A growing number of institutions can now document that they are adopting similar ways of working towards ensuring fast, consistent and recorded decisions.

Natural justice

Much of the recommendations in this paper underpin the requirement that the student should be assessed in ways that comply with natural justice. They deserve to know what they are being asked to do, to be helped to understand that requirement, to understand how it will be judged and to have their efforts treated consistently by assessors. This serves the dual purpose of enhancing the student's experience of assessment and ensuring the institution can defend it actions if and when it is ever challenged to do so.

References

Caval Collaborative Solutions (Sept 2002) *Victorian Vice-Chancellors' Electronic Plagiarism Detection Pilot Project: Member Project Report*, Caval Collaborative Solutions, Melbourne

McCabe D (2003) Promoting academic integrity - a US/Canadian perspective, *Proceedings of the Educational Integrity Conference: plagiarism and other perplexities, University of South Australia, 21-22 November 2003*

Parks C (2003) *In other (people's) words: plagiarism by university student - literature and lessons,* Assessment and Evaluation in Higher Education, 28, 5

What are good assessment practices?

Linda Suskie, Towson University, Maryland

Abstract

While perfectly accurate strategies to assess student learning aren't possible, because of factors that we can't control such as a student's health, we can maximise the quality of our assessments by addressing six characteristics of 'good' assessment that we **can** control to a certain degree. Good assessments are useful; they give reasonably accurate, truthful information; they are fair to all students; they are ethical; they are systematic; and they are cost-effective. This paper gives practical suggestions for maximising and documenting the quality of our assessment strategies.

Regardless of what or how we are assessing, our assessment activities should conform to six principles of good practice.

Good assessments:

- give us **useful** information
- give us **reasonably accurate, truthful** information
- are **fair** to all students
- are **ethical** and protect the privacy and dignity of those involved
- are **systematised**
- are **cost-effective**, yielding value that justifies the time and expense we put into them.

This paper discusses each of these principles.

Create useful assessments

Perhaps the most important assessment principle is that assessments be useful. If an assessment doesn't help improve teaching and learning activities, why bother with it? In order to be useful, **assessments must correspond to your key learning goals and your curriculum.** No one strategy is right for every course or programme in every institution.

To ensure the usefulness of your assessments, periodically evaluate your assessment programme and ask yourself whether your assessments are giving you useful information. If a particular assessment is not helping you or your students, stop doing it. Similarly, if a particular survey question isn't providing information that you can use to help make decisions about your programme, stop asking it. And periodically compare your assessment tools against your learning goals to ensure that they continue to align.

Create accurate, truthful assessments

What is a good assessment? More than anything else, it is an assessment that gives us truthful information; it tells us what our students have **truly** learned. Students who have truly learned what we want them to will do well on a good assessment; students who truly have **not** learned what we want them to will not do well on it.

Unfortunately, it's not possible to determine with complete confidence exactly what our students have learned. We can't get inside their heads to find out what they truly know and what they don't. The best we can do is to look at samples of their behaviour - what they write, produce, say and perform - and from those samples try to estimate or infer what they truly know. Even under the best of circumstances, making an inference from these snapshots of behaviour is bound to be at least somewhat inaccurate because of what psychometricians call 'measurement error' - fluctuations in human performance that we can't completely control.

We can't control, for example:

- whether a student is ill on the day they complete an assignment or takes a test
- whether a student is preoccupied with an argument they've had and therefore isn't focusing sufficiently to do their best
- memory fluctuations (we all periodically 'blank out' on key names and facts)
- luck in whether a particular assignment or test question focuses on something a student knows well (we all learn some aspects of a subject better than others)
- luck in guessing on multiple-choice questions or
- mental 'set' (sometimes we have flashes of insight; sometimes we seem inexplicably in a mental rut).

While we thus can't create assessments that will give us absolutely accurate information about what students have learned, we must strive to make them sufficiently truthful that we will have confidence in our findings and can use them with assurance to make decisions about goals, curricula and teaching strategies. The following approaches will help increase the accuracy and truthfulness of assessment strategies.

- **Start with clear statements** of the most important things you want students to learn from the course or programme.
- **Teach what you are assessing.** Purposefully help students learn the skills needed to do the assessment task.
- Because each assessment technique is imperfect and has inherent strengths and weaknesses, **collect more than one kind of evidence** of what students have learned. If you are assessing learning across an entire programme, for example, rather than only give students a culminating examination, you might also look at samples of papers they've written and perhaps internship supervisors' ratings of their skills.
- Before creating an assignment, **write a rubric:** a list of the key things you want students to learn by completing the assignment and to demonstrate on the completed assignment.
- Likewise, before writing test questions, **create a test 'blueprint':** a list of the key learning goals to be assessed by the test and the number of points or questions to be devoted to each learning goal.
- **Make assignments and test questions crystal clear.** Write them so that all students will interpret them in the same way and know exactly what you want them to do.
- **Make sure that your assignments and test questions clearly relate to your key learning goals**. Each test question, for example, should clearly correspond to the learning goal you've identified for it in your test blueprint. A writing assignment intended to assess how well students organise an essay shouldn't be graded primarily on grammar and spelling.
- **Ask colleagues and students to review drafts** of your assignments, rubrics and (using former students) test questions to make sure they're clear and appear to assess what you want them to.
- **Try out surveys and similar tools** with a small group of students before using them on a larger scale. Check students' responses to make sure they are giving answers that make sense. Ask them if they found anything unclear or confusing. Ask some students to 'think out loud' as they answer a test question; their thought processes should match those you intended.
- Collect enough evidence to **get a representative sample** of what your students have learned and can do. Collect a sufficiently large sample that you will be able to use the results with confidence to make decisions about a course or programme.
- **Score student work fairly and consistently**. Before scoring begins, have a clear understanding of the characteristics of meritorious, satisfactory and inadequate papers. Then use a rubric to help score assignments, papers, projects etc consistently.
- **Use assessment and quality assurance results appropriately. Never base any important decision on only one assessment.** (Failure to adhere to this maxim is one of the major shortcomings of many high-stakes testing programmes.) Assessments shouldn't make decisions for us or dictate what we should teach; they should only advise us as we use our professional judgment to make suitable decisions.
- **Evaluate the outcomes of your assessment efforts** and revise your strategies to address any shortcomings.

How can the quality of assessment and quality assurance methods be documented?

Should you document evidence of the quality of your assessment methods? This depends on how the results may be used. An assessment used to make minor curricular modifications does not need as much evidence of its quality as one used to help determine who graduates, whether expensive modifications should be implemented, or whether a programme should be terminated, or whose findings are likely to be challenged.

Obviously, the more rigorous and extensive your evidence, the more compelling it is, but also the more time-consuming it is to collect and evaluate. Be forewarned that, no matter how extensive your efforts to document the quality of your assessment strategies, you can never **prove** that your assessments are accurate and truthful; you can only collect evidence that your assessments **appear** to be accurate and truthful. Someone who wants to dispute your findings will always be able to poke a hole in your assessment strategy.

Should you decide to document the quality of your assessment activities, here are some ways to do so.

- **Keep records of everything you've done to maximise assessment quality,** including reviews of your assessment tools by others, tryouts of your assessment strategies, rubrics used to score student work, blind scorings by your colleagues and other strategies discussed in the previous section.
- **Use other kinds of assessments to corroborate your findings.** A student whose writing sample receives a high score, for example, should also receive a high score on a published writing test and a high rating from her professor on her writing skills.
- **See if results fall in appropriate patterns.** Students at the end of a programme should generally do better on an assessment than students at the beginning, while students with high grades should generally do better on an assessment than students with low grades. Some results should predict current or future performance; scores on a pre-calculus test, for example, should predict calculus grades at least somewhat accurately. And sometimes students should perform differently by major. Physics majors, for example, may score higher on a quantitative reasoning assessment, on average, than English majors.

These are only a few of the many approaches that can be taken to appraise and document the quality of assessment measures. To learn more, ask a psychology or education staff member for information on reliability and validity.

Create fair assessments

A fair assessment is one in which students are given equitable opportunities to demonstrate what they know. This does not necessarily mean that all students should be treated exactly the same. Equitable assessment means that students are assessed using appropriate methods and procedures, which may vary from one student to the

next depending on the student's prior knowledge, cultural experience and learning style. For example:

- Marla is not a strong writer but great at visualising concepts. She will better demonstrate her understanding of a complex concept if she can draw a diagram rather than write an explanation.
- Robert's culture values collaboration and he learns more from working with others than by studying alone. He will better demonstrate his understanding if he can work with others on a group presentation rather than make a solo presentation.
- Janice is not a good test taker but very creative. She will better demonstrate her understanding if she can create a video explaining a complex concept rather than take a test.
- Jason was home-schooled in a home without a computer, so he's still insecure on computers. He will better demonstrate his understanding on a paper-and-pencil test than on a computer-based test.
- Lisa attended a school that stressed rote memorisation and drill. She will better demonstrate her knowledge of American history on a fill-in-the-blank test than in a term paper that requires critical thinking skills.
- Dan has poor test-taking skills. If question 2 stumps him, he'll likely spend the whole testing period on that question and never answer the remaining questions. He will better demonstrate his understanding by writing a term paper than by taking a multiple-choice test.

Creating custom-tailored assessments for each student is, of course, largely impractical, but we can work toward assessing students equitably by providing a variety of assessment venues. Instead of assessing students solely through multiple-choice tests or solely through writing assignments, assess them using a combination of tests, writing assignments and other projects. Students might convey the essence of a novel's protagonist, for example, through a diagram, video or oral presentation rather than through the traditional essay.

Create ethical assessments

A number of professional organisations engaged in the assessment of human performance have developed statements of ethical standards. Two pervasive themes in these statements are protecting the privacy and dignity of those being assessed and using results in a fair and appropriate manner. Virtually all these statements agree that ethical assessment programmes:

Protect the privacy of those who are assessed. Take appropriate security precautions before, during and after you conduct an assessment, and protect the confidentiality of individually identifiable information. Password-protect computer files with identifiable information and store paper records with identifiable information in locked file cabinets. If several people are reviewing samples of student work or accessing a computer file, removing information that identifies individuals may be a wise precaution.

While it's important to protect student privacy, staff must have sufficient information to be able to do their jobs and this can often involve sharing identifiable information.

Some departments, for example, periodically hold staff meetings to discuss the progress of each of the students on their programme. Staff also consult with their colleagues about their students less formally; a staff member concerned about a student's slipping performance might consult with the student's advisor for ideas on how to help the student get back on track. Staff are simply carrying out an important part of their responsibilities when they hold such conversations, and considering identifiable assessment results can make the conversations more fruitful.

Keep students informed about the nature and purpose of each assessment. Students should be informed as early in their programmes as possible, in writing, of graduation or programme completion requirements beyond successful completion of coursework, such as compiling a portfolio, completing a survey, participating in a focus group or taking a comprehensive examination. These statements should also make clear if, in order to progress or graduate, students are expected to earn a minimum score on a special assessment such as a portfolio or published test.

Minimise potential bias. Obviously we wouldn't want to use an instrument with stereotyping or offensive material. But an unbiased instrument goes farther than that; it describes activities that are equally familiar to all and uses words that have common meanings to all. An item on a quantitative skills test that asks students to analyse American football statistics wouldn't be fair to women, for example, as they're typically less familiar with the sport than men.

A good way to detect potential bias is to ask yourself, 'If someone wanted to see the exact opposite of the results that I'm hoping for, would they conduct the same assessment in the same way?' You're probably hoping, for example, that your assessments will demonstrate that your students are learning all kinds of important things. Imagine (however difficult this may be for you!) that someone is convinced that your course or programme is of very poor quality and expensive to boot and wants it eliminated. What strategies to assess student learning might you both conceivably agree on?

To ensure further that your assessments are equitable and don't favour students of a particular gender or background, ask colleagues and students of varying backgrounds to review drafts of your assignments and test questions. And engage and encourage your students; the performance of some is greatly influenced by positive contact with staff.

Give appropriate attribution to the work and ideas of others. Don't use items from someone else's test or survey in your own assessment instrument, for example, without obtaining permission from the author or copyright holder and acknowledging the contribution.

Make the following information available to anyone considering your assessment results.

- The exact wording of assignments and questions given to students.
- How the participating students were selected and any evidence that the students who participated are a representative, unbiased sample of the students you wanted to assess.
- The number of students or student works in the sample, the number actually participating, and the participation rate. (For example, 'A random sample of 50 seniors was invited to participate in exit interviews. Twenty students or 40 per cent of those invited participated'.)

- Information on the precision of the results. (For example, 'Eighty-two percent of our alumni are satisfied with their education here, with an error margin of plus or minus four per cent'.)
- A fair, objective presentation of the results, both intended and unintended, without censorship.
- Qualifiers and caveats regarding the conclusions drawn from the results. (You might, for example, want to caution your audience about a low survey participation rate, a test question that you've learned was misinterpreted by many students, or that male students are underrepresented in the group of papers you evaluated.)

Discourage others from making inappropriate interpretations or otherwise false or misleading statements about assessment or quality assurance results.

Promote the use of multiple sources of information when making any major decisions.

Create systematised assessments

Good assessments are not once-and-done affairs. Assessments should be conducted on a regular basis to see if course and programme improvements are having their desired effect and to make sure past performance levels haven't slipped.

Programme assessments should be repeated fairly frequently, not just once every five or 10 years. Less frequent assessments can take more time in the long run, as there's a good chance that no one will remember, find the documentation for or understand the rationale behind an assessment done several years ago, which means spending far more time planning and designing the new assessment - in essence, reinventing the wheel. Imagine trying to balance your cheque book once a year rather than every month (or your students cramming for a final rather than studying over an entire term), and you can see how difficult and frustrating infrequent assessments can be compared to those conducted routinely.

Keep assessment efforts cost-effective

The business world's concept of 'return on investment' applies to assessment and quality assurance activities. Assessments should yield dividends - namely more effective learning experiences for students - sufficiently worthwhile to justify our investment of time and resources. Assessment is like putting together a jigsaw puzzle when we don't have enough time to assemble the entire puzzle. We want to put together just enough pieces to get a reasonably good sense of what the completed picture would look like.

Here are some strategies for keeping assessment manageable.

- **Focus your assessments.** It's better to do a few assessments well than many poorly. Concentrate on assessing just a few key learning goals rather than every goal of your course or programme.
- **Make maximum use of existing information** before creating or purchasing new tools.
- **Focus on those assessment strategies that give the greatest dividends** for time and resources invested.

- **Limit the volume of assessment information** you collect from students. Perhaps a one-page chart will give you just as much information on students' analysis skills as a three-page essay. Perhaps a two-page abstract will give you just as much information on students' writing skills as a 20-page term paper.
- **Use rubrics** - they really speed up the process of evaluating student papers and projects.
- **Stop doing something else**. Consider dropping your mid-term examination to give you more time to assess student projects. Consider moving some of your more straightforward lectures to handouts that students read on their own, creating more class time for students to collaborate on assignments and for you to review assignments with individual students.
- **Look at samples** rather than censuses of student work. If students maintain journals in your course, for example, spot check a random sample of them each week rather than read them all. If all students in a programme complete a senior thesis, evaluate just a sample of them for writing and critical thinking skills.
- **Stagger your assessments.** Stagger the due dates for assignments so each class's assignments are turned in a few weeks apart and you're not overwhelmed with papers at any one point in the term. Similarly, stagger programme assessments across a multi-year period. A three-year assessment cycle might include an examination of student portfolios every first year, a survey of alumni every second year, and exit interviews of graduating students every third.
- **Adapt your assessment schedule to meet your evolving needs.** Suppose that focus groups show high levels of student satisfaction but senior theses show poor organisational skills. You may want to put the focus groups on a back burner, conducting them only once every three years just to make sure student satisfaction isn't slipping, and begin reviewing theses every term to monitor the effectiveness of your efforts to strengthen organisational skills.

We're not talking dissertation-quality research here; establish realistic expectations for quality. Assessment is a form of action research, a branch of research that, while disciplined and systematic, is inherently imperfect, so don't expect perfection. While it would be wonderful if every assessment project were designed to meet standards for publication in peer-reviewed research journals, realistically most staff don't have the time - or interest - to do this. Aim not for replicable, generalisable research but for results that are simply good enough to use with confidence to make decisions about teaching and learning in your course, programme or institution.

Recommended reading

American Association for Higher Education (1996) *Nine principles of good practice for assessing student learning,* AAHE, Washington
www.aahe.org/assessment/principl.htm [Accessed 2 June 2003]

Anderson J A (1988) Cognitive styles and multicultural populations, *Journal of Teacher Education*, 24 (1) 2-9

Badger E (1999) Finding one's voice: A model for more equitable assessment, in A L Nettles and M T Nettles (eds) *Measuring up: Challenges minorities face in educational assessment*, Kluwer, Boston

Campbell D T and Fiske D W (1959) Convergent and discriminant validation by the multitrait-multimethod matrix, *Psychological Bulletin*, 56 (2) 81-105

Conference on College Composition and Communication (1995) *Writing Assessment: A Position Statement,* National Council of Teachers of English, Urbana www.ncte.org/about/over/positions/category/write/107610.htm

Conference on College Composition and Communication (2001) *Guidelines for the Ethical Treatment of Students and Student Writing in Composition Studies*, National Council of Teachers of English, Urbana www.ncte.org/about/over/positions/level/coll/107670.htm

Fleming J (1998) Correlates of the SAT in minority engineering students: An exploratory study, *Journal of Higher Education*, 69, 89-108

Gonzalez V (1996) Do you believe in intelligence? Sociocultural dimensions of intelligence assessment in majority and minority students, *Educational Horizons*, 75, 45-52

Greater Expectations Project on Accreditation and Assessment (2002) *Criteria for recognizing "good practice" in assessing liberal education as collaborative and integrative,* Association of American Colleges and Universities, Washington DC

Huba M E and Freed J E (2000) Applying principles of good practice in learner-centered assessment, in *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning,* Allyn & Bacon, Needham Heights

Joint Committee on Standards for Educational Evaluation (1994) *The program evaluation standards: How to assess evaluations of educational programs,* second edition, Sage, Thousand Oaks

Joint Committee on Testing Practices (1988) *Code of fair testing practices in education,* National Council on Measurement in Education, Washington DC

Lam T C M (1995) *Fairness in performance assessment: ERIC digest,* Office of Educational Research and Improvement, Washington DC (ERIC Document Reproduction Service No ED 391 982)

Linn R L (1999) Validity standards and principles on equity in educational testing and assessment, in A L Nettles and M T Nettles (eds) *Measuring up: Challenges minorities face in educational assessment*, Kluwer, Boston

McCabe D L and Pavela G (1997) The principled pursuit of academic integrity, *AAHE Bulletin,* 50 (4) 11-12

National Council on Measurement in Education (1995) *Code of Professional responsibilities in Educational Measurement*, NCME, Washington DC www.natd.org/Code_of_Professional_Responsibilities.html [accessed 2 June 2003]

National Research Council (1993) *Leadership Statement of Nine Principles on Equity in Educational Testing and Assessment,* NRC, Washington DC www.ncrel.org/sdrs/areas/ issues/content/cntareas/math/ma1newst.htm [Accessed 2 June 2003]

Office for Protection from Research Risks (13 November 2001) *Code of Federal regulations: Title 45: Public welfare: Part 46: Protection of human subjects,* National Institutes of Health, Washington DC

Parkes J (2000) The Relationship between the Reliability and Cost of Performance assessments, *Education Policy Analysis Archives*, 8 (16) http://epaa.asu.edu/epaa/v8n16/ [Accessed 2 June 2003]

Shavelson R J and Huang L (2003) Responding responsibly to the frenzy to assess learning in higher education, *Change,* 35 (1) 10-19

Spangehl S D (1994) Latent and leveraged benefits of assessment: Expanding the value of assessment information, In Bers T H and Mittler M L (Eds) Assessment and Testing: Myths and Realities, *New Directions for Community Colleges*, Vol 88

Suskie L (2000) Fair assessment practices: Giving students equitable opportunities to demonstrate learning, *AAHE Bulletin*, 52 (9) 7-9

Thompson B and Daniel L G (1996) Seminal readings on reliability and validity: A "hit parade" bibliography, *Educational and Psychological Measurement*, 56 (5) 741-745

Walvoord B E and Anderson V J (1998) *Effective grading: A tool for learning and assessment*, Jossey-Bass, San Francisco

Note: This paper is adapted from Chapter 2 of:
Suskie L (2004) *Assessing Student Learning: A Common Sense Guide*, Anker Publishing, Bolton
Reprinted with permission of the publisher

A powerful learning environment

Professor David Lines, Centre for the Enhancement of Learning and Teaching, The Robert Gordon University

Introduction

This paper will suggest that given current knowledge of the ways students learn, we need to review the purposes of assessment. The principle of 'active learning' must be accompanied by an appropriate, aligned assessment regime, otherwise all efforts to enhance final learning outcomes will fail.

Two cultures of assessment

Though different writers use slightly different terminology, all generally agree that assessment has three purposes (inter alia, Brown et al, 1997; Yorke, 1998; Black, 1998). Firstly, assessment is designed to support and thus enhance learning. Secondly, it provides certification for progress or transfer and thirdly it is a form of accountability (quality assurance) to stakeholders.

For the student, it is the second of these that is crucial; for funding agencies, government, taxpayers and so on, it is the third. In both cases it could be argued that assessment of education is taking place. The student wants to know that, however organised, the successful passing of examinations (in whatever form) will lead either to entry into the next stage of education or into an appropriate job. The external stakeholder is also judging education; put at its crudest that judgement is based on a 'value for money' argument, though what 'value' means in this context is highly contestable.

The worry is that the assessment **of** education has become so pervasive in a world of competitive league tables, that assessment **for** education is increasingly crowded out, with the result that summative tests dominate formative ones; courses are sub divided into units, each of which require assessment in high stakes contexts that threaten a synoptic appreciation of subject knowledge, and so on. It would seem that in these two cultures of assessment there is only one winner.

Yet if we take a step back and examine how learning takes place, we can quickly see that assessment **for** education can also provide an assessment of education. As Palomba and Banta (1999) have shown, better assessment provides better information for **all** stakeholders whether they be staff, students, administrators or taxpayers. They quote Colorado State University, which has involved employers in the assessment process, and Eastern New Mexico University where every fine arts student has their portfolio or audition assessed each semester in front of a panel that includes staff, students, community representatives and other professional staff from outside the fine arts department. Such an eclectic engagement is crucial for a better understanding of what and why we assess and is needed to demonstrate to stakeholders, especially policy makers, that education, as in most walks of life, has moved on.

Yet for the 'two cultures' to move together there must be a shift in attitudes. What Hesse (1989, quoted in Wilson and Scalise, 2003) described as a 'pass the buck' approach, which is that failure is the student's fault (not intelligent enough/didn't

study hard enough etc) must be replaced by a recognition that learning requires a partnership between the teacher and the taught. For that to happen, we must recognise changing views on the way people learn.

Constructivism

In contrast to the notion of a student's brain being a 'void' into which knowledge is 'poured' (leading to the rote learning, through instruction, of 'facts' to be recalled in a final examination), the fundamental idea of constructivism is that the learner has to make sense of the data being supplied in his or her own way. Within this simple statement there is a great deal of controversy (see Light and Cox, 2001) but there is general agreement over its basic tenets. One of these is that within a teaching situation, instead of supplying 'facts' - a slippery term in itself - the teacher's role is to provide what is called the 'scaffolding' for learning. That is to say, the support upon which learning rests, rather than the learning itself.

The objective of constructivism is, as its name suggests, to help the student construct their own meaning for the information presented. The teacher can assist the process (provide the scaffold), but in the end the student has to 'own' the data and be comfortable with it. Self evidently, the only way the student can achieve this is by being an 'active' participant in the learning process.

Instinctively we learn actively, all the time. We watch our parents, our peers, even our employers and we acquire knowledge from them, but we process the information in our own terms and in our own ways, mediated by our personal experiences elsewhere. As we grow older we increasingly make an informed use of heuristics (rules of thumb), intuition and pattern recognition, but then, as experts, we go a stage further and discover shortcuts (Drefus, 1986).

The idea of constructivism builds on Schon's earlier notion of the 'reflective practitioner' (1983). The reflective practitioner or professional is someone who can step outside themself and observe their own actions in a given situation, evaluate them and consider ways in which they might be improved. Although such reflection is largely an instinctive or intuitive trait, becoming a reflective practitioner requires teaching. The role of the teacher, once reflection becomes embedded, is much more that of mentor, 'critical friend' or coach. The mentor's role is to challenge the trainee, to ensure that evaluation is properly and comprehensively carried out and then to provide suitable challenging and stimulating 'real-life situations and contexts' to anchor the learning (Direick and Dochy, 2001). This new environment recognises that for deep learning (ie that which is embedded and long-lasting) to occur, the learner has to make his own sense of the data, the situation and the context and take ownership of all the inputs. These can then be processed and returned as actions when new, similar contexts and situations are faced. Such a style of teaching and learning sits in complete contrast to the instructional approach, where knowledge is 'owned' by the teacher and is 'handed out' in what the teacher considers to be an appropriate amount in an appropriate package at an appropriate time.

It must be emphasised again that this new learning environment does not remove the teacher/mentor's role, although it certainly changes it. Similarly, the assessment

regime must also change, because if the teaching and learning environment changes, then, to be valid, the assessment regime associated with it must alter as well. Encouraging learners to self-discover (in a 'safe' environment, that is one where mistakes are accepted), to develop new strategies for new situations and so on, require an assessment regime that provides constructive feedback (or, more properly feedforward, since reflections on past performances are of little value unless they alter behaviour in the future). Such a regime cannot be predicated on pass/fail notions, nor can it occur only at the end of a course, which implies a shift away from summative towards formative assessment, or more simply, from testing to (true) assessment.

From 'teacher directed' to 'competence orientated'

Elshout-Mohr et al (2002) have described the shift outlined above as a movement from a teacher directed configuration to one that is competence orientated (see Figure 1). In between there is one described as 'standard orientated'.

Configuration	Learning arrangement	Presentation of learning outcomes	Standards and assessment criteria
Teacher directed	Teacher directed: **equal** for all students	Teacher directed: **equal** for all students	Established by the teacher: **equal** for all students.
Standard oriented	Student directed: **different** learning routes	Teacher directed: **equal** for all students	Established by a panel: equal for all students.
Competence oriented	Student directed, affected by personal preferences and opportunities: **different** learning routes	Student directed and affected by personal preferences and opportunities: **different** for students	Adapted to preferences and opportunities: **different** for students. Even the composition of the assessment panel might be **different**.

Figure 1 Equality for all students of learning arrangement, presentation of learning outcomes and assessment criteria in the three educational configurations (Elshout-Mohr et al)

In reading the chart, it is important to see the link between learning outcomes, teaching and assessment, which has been called 'constructive alignment' (Biggs, 1996) or 'the congruent curriculum' (Lines, 1999). If the assessment processes and procedures fail to match the other elements, for instance, then assessment 'backwash' comes into play (Dochy, 2001). The 'backwash effect' is an educational version of Gresham's Law - that the bad drives out the good. In other words, if the certification actually and ultimately depends on a final, pencil and paper test, then students will abandon any other learning strategy, however noble and effective, and concentrate on the test, for it alone provides the route to their ultimate goal.

In Figure 1, teacher directed is what might be described as a conventional, perhaps even old-fashioned configuration. There are apparent efficiencies in this system, in that the teacher and the tests can focus strongly on what are considered to be most important. The problem, however, is that material learnt and tested in this way has a very short half-life indeed - perhaps no longer than it takes for the student to leave the examination hall!

The second, standard orientation, describes a movement towards giving the learner more responsibility. Standards are laid down and the teacher or mentor provides coaching to help the student achieve them. It is the student who largely determines what learning strategies he or she will adopt in working towards them. The third configuration, competence orientation is one where students operate in professional environments and mentors work with them as co-learners. In this phase, individualised learning is matched by individualised assessment.

Although it is impossible to generalise, it is probable that large parts of higher education have moved or are moving away from the teacher directed configuration and are making progress towards standards orientation. The final shift is the one advocated here, though the challenges that such a change presents should not be underestimated and will require intensive training and support, as well as a wholehearted commitment by all interested parties.

Creating a powerful learning environment

From the above it is possible to identify certain characteristics of what is called a powerful learning environment and the assessment that is associated with it.

Learning, teaching and assessment are integrated and aligned. This means that careful consideration is given in advance as to what the outcomes of learning will be for successful course completion; the means of delivering appropriate skills and knowledge are put in place and appropriate methods of assessing the specified outcomes are constructed. If any one of these three components is out of alignment, the entire structure fails.

The student is an integral part of the process. This is crucial, since the focus is and must remain, the learning outcomes that the individual should achieve by the end of the course. That does not mean that the student determines those outcomes, or the assessment process by which decisions will be made on whether or not the outcomes have been achieved, but the process of acquiring the required knowledge and skills becomes much more the student's responsibility, and indeed, in close association with the mentor, the student may help to determine when or whether they have achieved them.

Both the outcomes and the process of achieving them are assessed.

The assessment process uses a variety of approaches, including real life scenarios that require decisions to be made. The scenarios should require candidates to take a variety of perspectives and also to take context into account so that the knowledge and skills applied once can be transferred into new perspectives and different contexts. Causal mechanisms should be investigated. As a guide, the interrogative words 'when?', 'where?' and 'why?' should be used rather than 'what?'.

The evidence for success is presented in a portfolio, which implies that a single score success or fail mark is no longer tenable. Birenbaum (1996) describes this as a shift from quantification to a portrayal. It also implies that the assessment cannot be time constrained (though of course there can be elements that are, depending on the learning outcomes sought).

The process of learning involves tasks that engage the student, are meaningful, challenging and 'authentic'. An assessment that is authentic is one that closely matches the desired performance and takes place in an authentic context. It should be pointed out that the assessment process is itself dynamic and impacts upon the person being assessed. Research has shown, for instance, that 'easy' questions at the start of an examination result in higher overall scores, because early success builds confidence in the candidate's mind (Goldstein, 1994). Put crudely, this can be translated as 'success breeding success'. This phenomenon can, unfortunately, also work the other way, with failure on one occasion causing, or contributing to failure in the future.

A reflective diary is maintained by the student and is central to the learning process. The diary or journal may or may not be confidential, but if it is to be used as part of the course assessment, then it must be shared, at least with the mentor.

To summarise the above, a powerful learning environment requires the adoption of a variety of assessment techniques, but with examinations emphasising higher order skills, all to be enclosed within a portfolio. The portfolio should include examination results, a reflective diary, personal observations and so on. The emphasis of the entire assessment 'package' is on the engagement of the learner in developing competencies, though straightforward knowledge and skills would not be ignored.

Pie in the sky or down to earth?

It may seem that persuading students to keep a reflective diary and then to construct a portfolio of evidence, which includes failures as well as successes, is unrealistic. Yet students in art and design courses have done this for many years, and increasing interest in personal development planning perhaps suggests a sector-wide interest in such an approach. Fortunately, technological changes can also help. The existence of e-portfolios offers the potential for accessible, easy-to-manage documents that can be cross-referenced with ease, far removed from conventional notions of weighty, impenetrable paper versions.

So, perhaps the proposal is not so unrealistic after all. Instead of emphasising the negative aspects, we should instead ask what will happen if we don't change. Arguably, this will mean a continuation of shallow learning, of education being assessed almost exclusively in 'value for money' terms and a sector containing students who don't enjoy higher education for the intellectual enhancement it provides, but who instead simply see it as a means to an end.

It really is time to move on.

References

Birenbaum M (1996) Assessment 2000: towards a pluralistic approach to assessment, in M Birenbaum and F Dochy (eds) *Alternatives in assessment of achievements, learning processes and prior knowledge,* Kluwer, Boston

Black P (1998) *Testing: Friend or Foe? Theory and Practice of Assessment and Testing,* Falmer Press, London

Brown G, Bull J and Pendlebury M (1997) *Assessing Student Learning in Higher Education*, Routledge, London

Direick S and Dochy F (2001) New Lines in Edumetrics: New Forms of Assessment lead to New Assessment Criteria, Studies in Educational Evaluation 27 (4) 307-329

Dochy F (2001) A New Assessment Era: different needs, new challenges, *Research Dialogue in Learning and Instruction* 2, 11-20

Elshout-Mohr M, Oostdam R and Overmaat M (2002) Student Assessment within the Context of Constructivist Educational Settings, S*tudies in Educational Evaluation* 28, (4) 369-390

Goldstein H (1994) Recontextualising Mental Measurement, *Educational Measurement: Issues and Practice* 13 (1) 16-19

Hesse J (1989) From Naive to Knowledgeable, *The Science Teacher*, September, 55-58

Light G and Cox R (2001) *Learning and Teaching in Higher Education*, Paul Chapman Publishing, London

Lines D (1999) *Cultural Values and the Curriculum: Economics and Business Education at Different Stages in the Maturation Process*, unpublished PhD thesis, University of London

Palomba C A and Banta T W (1999) *Assessment Essentials*, Jossey-Bass, San Francisco

Schon D A (1983) *The Reflective Practitioner: How Professionals Think in Action,* Temple Smith, London

Wilson M and Scalise K (2003) *Assessment to Improve Learning in Higher Education: The BEAR Assessment System,* under review

Yorke M (1998) Assessing capability, in J Stephenson and M Yorke (eds) *Capability and quality in higher education*, pp174-191, Kogan Page, London

Assessment - The final frontier - Just how valid, reliable and fair are assessments of disabled students?

Karen Robson, Disability Services Manager, University of Wales Institute, Cardiff

Abstract

'Any attempt to challenge the boundaries of conventional assessment is bound to provoke many new questions' Broadfoot 2002.

In September 2002, the *Disability Discrimination Act 1995* (DDA) was amended by the addition of Part IV *Special Educational Needs and Disability Act* and with it came new institutional duties relating to the treatment of disabled students and applicants. This change in the law has resulted in the production of a plethora of good practice materials in order to assist institutions and individual academic staff in the review of practices to ensure they are appropriate to the increasing number of disabled students studying in higher education. Many of these resources address learning and teaching practices to enable staff to deliver the curriculum in an accessible way. While much has been written on assessment generally, little has focussed on assessing disabled students. Resources associated with assessment of disabled students are becoming increasingly available, but this area would appear to be the one which provides some of the most challenging questions and therefore a 'blueprint' approach maybe some way off. Implicit within this, is the requirement to ensure that any change to conventional assessment is undertaken in such a way that staff are confident in utilising these methods and that they continue to fulfil the principles of assessment and in no way lower academic standards. Simultaneously, they must be valid, reliable and fair tests so that students also have confidence in this approach. Research undertaken by Sharp and Earle (2001) found evidence of a lack of consistency between higher education institutions in their approach to assessment of disabled students and identified a general lack of an explicit policy, thus alternative assessments offered were often not 'genuine' ones. The issue is therefore, how do we ensure disabled students are given the opportunity of being able to demonstrate their knowledge in an appropriate and accessible way, but at the same time ensure any tests used are genuine alternatives and meet the themes of the conference in terms of validity, reliability and fairness.

Introduction

The short case study presented at the conference focused on the assessment of disabled students, considering the context of this recent change in legislation and the implications of not identifying possible barriers appropriately, while addressing them in a valid, reliable and fair way. The concept of treating all students equally, but not necessarily treating them the same was also considered and alongside some examples of genuine alternative assessments were discussed. To end, some suggestions for inclusive practice on which to build alternative arrangements were offered as a basis for discussion within the breakout session.

As Brown and Knight (1994) comment, 'assessment is at the heart of the student experience'. Assessment plays a significant role in the learning experience. It defines what students regard as important and has a profound impact upon motivation and achievement. I think few in education would disagree with this, so why therefore does Ramsden (1992) argue that 'assessment of students is a serious and often tragic enterprise', warning us of the consequences of embarking upon it in an inappropriate way. These consequences have the potential to adversely affect disabled students even more so, though since the introduction of the DDA Part IV in September 2002, institutions should no longer treat a disabled student less favourably for reasons relating to their disability and also have a duty to make reasonable adjustments to ensure that disabled students are able to access the curriculum and affiliated services. The legislation also requires institutions to be proactive and anticipate access difficulties. These duties apply to all aspects of learning and teaching, but it is important to emphasise that there is no requirement to reduce academic standards. The introduction of the DDA has inspired a vast range of useful resources for both institutions and academics and many like the Teachability audit tool (Simpson, 2000) and Accessible Curricula (Doyle and Robson, 2002), address pedagogical issues and curriculum design and delivery. Assessment however remains more elusive and is perceived by many to be the 'hardest nut to crack'. The South West Academic Network for Disability Support (SWANDS) document (Waterfield and West, 2002) provides a really useful chapter on assessment, giving helpful examples of alternative assessment methods. What this document suggests is that developing assessment techniques for disabled students isn't that difficult and just requires a degree of lateral thinking and maybe a change in philosophy, as opposed to any major changes; I'll return to this point later.

Georgina Follett (2003) commented, 'the hardest task left is for the shift in culture to result in real change, change embraced for the value perceived, not for compliance'. While responding to our duties under the DDA is a serious matter of legal compliance, with transgressions having numerous negative implications for both institution and student, it is also worth remembering that the law merely imposes a base line and although the 'big stick' approach is often needed to encourage institutional change, it may be more helpful to see this shift as being a 'carrot' - an opportunity to review traditional practices and enhance the value perceived; for example reviewing the fairness of assessment methods for a wide range of students. With regard to disabled students in particular (though this applies to all students in different ways), barriers can exist in all forms of traditional assessment methods; from exams to practicals, group tasks to essays; for some students such methods will not enable the student to demonstrate their knowledge/ability on a par with other

students. Given the latest data available (which can be accepted as not being wholly accurate for a wide range of reasons and thus is considered unrepresentative of the true number), 4.7 per cent of students in higher education have a declared disability (HESA, 2001-02). Thus there is a real risk that traditional forms of adjustments ie extra time, are going to become unmanageable (SWANDS, 2002) for such numbers which continue to increase. That's a practical point, but to bring this back to the issue of validity, reliability and fairness, how can academics be sure that, for example, an exam with 25 per cent extra time is indeed an appropriate way of assessing a dyslexic student with short-term memory difficulties? What are we trying to test in such circumstances? The student's ability to recall facts and, as is also likely, the ability to marshal thoughts into a coherent argument. There is little evidence to support this type of adjustment; it could be argued it is almost a measure of custom and practice for want of an alternative. There is however, much evidence to suggest that 'special exam arrangements' for disabled students is expedient, bolt-on solutions to existing practices which are not working. Earle et al (1999) noted that UK higher education institutions made provision for disabled students on an ad hoc basis. In 2001, Sharp and Earle noted that there was little consistency between UK higher education institutions, no explicit assessment policy and that the alternatives offered were not genuine alternatives. Many would suggest that little has changed since and there is still a preponderance of extra time to address disabled students' needs, almost regardless of whether that reflects their learning style. If we are truly concerned about validity, reliability and fairness, our approach towards the assessment of students must be reviewed. Clearly, as Broadfoot suggests at the beginning of this paper, this is going to provoke new questions and may also be considered to be controversial, but it is a process we must pursue in academia, if we are to be confident that disabled students are being given the opportunity to demonstrate their knowledge in a fair and consistent manner.

Sharp and Earle (2001) describe different types of assessment. They suggest that a genuine alternative form of assessment utilises, 'methods which test exactly the same skills and knowledge as the original assessment'. Conversely, compensatory assessments are 'tests which do not assess the full range of skills and knowledge as the ones they replace'. Genuine alternatives are designed to minimise the impact of the student's disability on their performance. Compensatory forms of assessment could be perceived as discriminatory and counter-productive and arguably for many students this is still their experience. So what is the answer? Linda Suskie (2004) suggests that 'a fair assessment…where students are given equitable opportunities to demonstrate what they know…this does not necessarily mean all students should be treated the same'. Another way of putting this could be - same assessment, different process. This approach requires a very clear understanding of what the learning outcomes are that need to be demonstrated. Once this is clarified, it is possible to align an appropriate method of assessment. This could be for example replacing a dissertation with a project on designing an interactive CD-ROM, accompanied by a viva perhaps. This on the surface can appear quite challenging and even may be an assault on academic standards. As Broadfoot suggests, this raises a number of questions, such as is a dissertation essential to an honours degree or what are the intended learning outcomes and can they be met in an alternative way? In our respective disciplines, we need to be clear about what is essential or what may be custom and practice that we

perpetuate without really being sure why. Further, in a multimedia technological age would the underlying skills requiring for producing a CD-ROM be more appropriate or relevant to future employers? Will traditional forms of assessment move over to forms of assessment which utilise more diverse skills? Through addressing the challenges posed by reviewing traditional assessment methods for disabled students, maybe the underlying principles of assessment in general are being questioned. There has been a steady move towards continuous and more varied assessment over recent years and in fact the subject benchmark statements, produced by the Quality Assurance Agency for Higher Education, list 47 different forms of assessment methods considered valid and reliable methods worthy of incorporation into our programmes. Clearly the method utilised must be considered fit for purpose, thus what is being tested? Is it a good memory, an ability to retain information, an ability to select and interpret sources etc? This is supported by providing clear learning objectives and accessible module descriptors or programme specifications.

A common question raised when alternative assessment is proposed for disabled students is, how do we know this is genuine and fair and corresponds to the existing method utilised? One answer is to pilot alternatives testing their validity, but maybe the most straightforward is to consider offering more than one form of assessment from the outset of the course, as opposed to a knee-jerk reaction on demand, for example, offering a 500 word book review, alongside the traditional 2,000 word essay. If nothing else the former would be quicker to mark! By adopting this approach, students who find certain sorts of assessment unconducive are offered a choice by which to demonstrate their knowledge. The SWANDS resource offers a number of examples whereby students are assessed maybe using three different forms of assessment and the results compared. What is evident in most examples is that the student performs better using the alternatives, for example, the student with Asperger's syndrome who undertook a traditional timed exam (third class equivalent), an essay and a cloze (both of an upper second level). This approach however, can go one step further and address the key issues focussed upon in this conference. In searching for 'alternatives', maybe we need to look no further than to our colleagues in other academic disciplines. While producing a report may be unknown in the history department, in engineering or science it would be common practice. So maybe an alternative assessment to address a disabled student's needs would be less of a stab in the dark, or an ad hoc measure as Sharp and Earle suggest, but a measured approach grounded in years of testing on multiple students, albeit in a different discipline. The principle behind inclusive education is not only beneficial to disabled students but to mature learners, or other groups that make up the diverse student population. Thus offering perhaps four different forms of assessment per assessment opportunity reflects differing learning styles as opposed to focusing on the person's impairment. As the Social Model of Disability espouses, a person is disabled by their environment; in this case an inappropriate form of assessment. Offering validated choices, evaluated prior to need, changes the disabling environment and critically can meet the principles of assessment by being reliable, valid and fair tests; this returns to the theme of perceived value rather than purely addressing compliance issues without being confident of the merits of such an approach.

Confidence in results is clearly of primary importance to both students and staff alike. As Carroll (2004) suggests in her presentation, institutions must be more aware of the issue of plagiarism. Moving away from a timed and invilgilated exam can open up concerns about whether it is the student's work being assessed, or that of someone else. A recent article appearing in *The Guardian* (4 June 2004) written by a retiring academic raises this point and he argues forcibly to 'save the three-hour desk exam'. Concerns regarding the difficulties students face are dismissed by acknowledging the 'structural fairness of the British system'. He asserts that 'three hour finals exam produce results which tally well with classroom performance' and advocated that 'it is fair - the most level of playing fields'. As with any new developments and change, there can be a tendency to throw 'the baby out with the bath water' and so recommending the phasing out of such methods of testing is not being advocated here. However, I would question whether it is indeed the most level of playing fields - how often do we hear tutors reporting in exam boards that the student performed well in class but was let down by exams? This is clearly not consistent. What is being proposed is a mixture of testing methods, choice and also, where possible, that the methods are considered during course review or validation, embedding and mainstreaming them into the curriculum, moving away from ad hoc untested solutions, which can't be guaranteed of meeting the principles asserted in the focus of this conference. As Cottrell (no date) recommends, 'courses should make clear what alternatives in assessment are or are not feasible and have good reasons, from an educational and disability perspective, why restrictive practices cannot be modified'.

Inclusive practices

So, how can we move this agenda forward, ensuring that we as individuals and as institutions are legally compliant, utilising assessment methods that give us and our students confidence in the validity, reliability and fairness of the system and also reflect the wealth of good practice recommendations available about using assessment tasks. Firstly, lets talk about this; in our course teams, in our disciplines and across the sector; lets find solutions together working from the start point of what is it that is being tested? As Boud (1995) suggests 'there is more ignorance of significant issues in assessment than in any other aspect of higher education', so there is a staff development issue here that must be addressed and the enhancement theme explored in these workshops provides an ideal opportunity to reflect upon our practices. Secondly, vary assessment methods and offer a choice across programmes for all students, thus not singling disabled students out for special treatment. By doing this, the needs of all students are considered in an inclusive way. Use events such as Course Review and Validation to reconsider, reflect and then embed the slight change in approach and work with colleagues from other disciplines in this process in order to increase awareness and familiarity of differing assessment forms; what do they consider to be the pros and cons of their current choices? Consider formulating a policy on assessment to facilitate change; at the conference a number of delegates reported 'I couldn't do that at my institution'. Our policies should be there to enable us, not to create slaves! From the conference breakout sessions, it was evident that there was a huge amount of good practice being implemented in institutions, but many delegates themselves reported on the 'ad hocness' of this practice and also the inconsistency. This surely undermines the principle of fairness we are all seeking to

employ. The Staff-Student Partnership for Assessment Change and Evaluation project managed via the University of Plymouth is a research project incorporating a number of institutions, exploring the validity and fairness of differing forms of assessment. The project is seeking to address questions such as 'can we assess ability and not the effects of disability?'; 'can we accommodate the learning styles of a range of learners at assessment?'; 'can we reduce the 'discriminatory and exclusionary features of current policy?' (Barton, 2003). The project team are still gathering their data, however, the results look to be providing a really useful bank of information and empirical evidence, regarding one of the last areas of pedagogy that must be tackled with regard to access for disabled students. Such evidence will only seek to add to the credibility of any proposed changes to conventional assessment methods.

I called my presentation 'the final frontier'. This can make encouraging a shift in the culture of assessment appear elusive. On reflection I don't think it is, especially after hearing about the good practice that exists in higher education institutions. Nor do I think it requires us all to become experts in disability in order to address the requirements of the DDA. More that we become experts in aligning the principles of assessment with the learning outcomes and we consider opening up the 'closed shop' of assessment and provide students with choice, making the most of their learning styles and preferences. As John Sutherland, the author of the article in *The Guardian* (June 2004) reminded us, 'the three-hour, bookless, desk exam derives from medieval institutions, when libraries were few and far between and knowledge had to be stored in the student's head'. Is this all we are testing? Isn't it time for a change?

References

Barton L (2003) *Inclusive Education - Teacher Education: A basis for hope for a discourse of delusion*, Professional lecture, Institute of Education, University of London

Boud D (1999) Assessment and Learning: Contradictory or Complimentary? in Knight P (ed) *Assessment for Learning in HE*, Kogan Page, London

Broadfoot P (2002) *Assessment in Education: Principles, policy and practice*, Vol 9, No 1, Carfax Pub

Brown S and Knight P (1994) *Assessing Learners in HE*, Kogan Page Limited, London

Carroll J (2004) Fair assessment, fair policing and fair punishment: building on reliabililty and validity, in *Reflections on Assessment: Volume II*, QAA, Gloucester

Cottrell S (2001) *Developing positive learning environments for dyslexic students in HE*

Doyle C and Robson K (2002) *Accessible Curricula, Good Practice for All*, UWIC Press

Earle S, Adams M and French D (1999) A national survey of assessment practices in HE: special provisions for students with disabilities and learning disabilities, *Skill Journal* (no issue no), Skill, London

Follet G (2003) The place of assessment in quality enhancement in Scotland, *Exchange*, Issue 4, Spring 2004

QAA (2002) *Code of practice for the assurance of academic quality and standards in higher education: Section 6: Assessment of students,* QAA, Gloucester

Ramsden P (1992) *Learning to teach in HE*, Kogan Page, London

Sharp K and Earle S (2001) Assessment in HE: compensating students with disabilities and learning difficulties, *Skill Journal,* No 69, Skill, London

Simpson A, Teachability Project (2000) *Creating an Accessible Curriculum for Students with Disabilities*, University of Strathclyde, Glasgow

Suskie L (2004) *Assessing student learning. A Common Sense Guide,* Anker, Publishing, Bolton, MA

Sutherland J (2004) My final word - save the three-hour desk exam, *The Guardian*, June

Waterfield and West (2002) *SENDA Compliance in Higher Education. An audit and guidance tool for accessible practice within the framework of teaching and learning*, South West Academic Network for Disability Support

Issues of validity, reliability and fairness - Post-workshop report

Pamela Flanagan, Royal Scottish Academy of Music and Drama and Workshop Director

On 7 May 2004, 60 delegates gathered at the University of Stirling to hear presentations on issues of validity, reliability and fairness in assessment from four distinguished speakers drawn from the UK and United States, to debate and discuss issues arising from these presentations, to make recommendations for future progress and to air their views and ask questions of the workshop panel at the conclusion of proceedings. This report attempts to distil and summarise the essential themes and recommendations which emerged over the course of the day through references to the papers presented by all involved, reports of the discussions and main points made by the participants, and the views of those who were not there, but who were nonetheless the main reason for the gathering, namely the students whose opinions were voiced through a scoping survey of Scottish higher education institution students' associations conducted by the Quality Assurance Agency for Higher Education (QAA) Scotland during February/March 2004.

The first part of the conversation: the speakers' perspective

'Assessment is a big jigsaw puzzle; we don't have time to put the whole puzzle together but just enough pieces in order to get an idea of what the whole picture might look like' (Linda Suskie, Stirling Workshop, 2004).

So what themes emerged following the four presentations? The creation of a student-centred environment was clearly a priority, taking account of the students' needs and at the same time their right to privacy in what is often a highly public process. The necessity for the student to become more involved in the learning and teaching process was highlighted; as David Lines observed: 'If you make it important then it will be important to them' and again 'The student should be helped to construct his/her own meaning for the information presented'. This view was supported by the other members of the panel: 'They need time to learn what is required and how to do it' (Jude Carroll) and 'Give them the skills needed to do assessments' (Linda Suskie).

The need for clarity and communication for, and between, both staff and students was reiterated by all four speakers: clear rules and procedures, clear statements of what students are required to learn, clear assignments and questions, clear feedback and clear feedforward, clear marking and assessment criteria aligned to learning and teaching outcomes etc. The difficulties posed for the teacher/lecturer, coupled with the fear of failure on their part, was acknowledged, but reassurance was on hand, beginning with Linda Suskie who asserted that 'We can never have a perfect assessment', a point echoed by David Lines who opined that 'We're bad at using failure as a scaffold for success, and yet we learn more from the failures than we do from the successes'. People's reluctance to tackle the difficult area of plagiarism whether through uncertainty, ignorance, fear of legal consequences etc was also an issue but, as Jude Carroll observed, 'It shouldn't hurt to uphold the law'. Allied to this, consistency and transparency, both in the application of the assessment process and in the application of sanctions in the event of transgression, was considered

paramount. On the different but no less sensitive issue of disability, Karen Robson urged her listeners to think of the assessment process as being 'not just about compliance, but about incentive; we need to be prepared and we need to think ahead'.

The need for reflective practice to become a natural aspect of any teacher's skill base and for the teacher to become a reflective practitioner rather than a teacher 'telling' a student how it goes, was reiterated by David Lines and Linda Suskie respectively: 'Any assessment is telling you something as a teacher' and 'Assessment is a form of action research'. The importance of partnership was stressed, whether this is institutional, resulting in not just moral support for staff to be innovative in their assessment methods but practical assistance through staff training, or departmental, using and acknowledging the help of colleagues in evaluating, reviewing, developing and maintaining fair and reliable assessment methods not just departmentally but across an institution as a whole, or working together with students to enhance and support their learning: '[the teacher should be] a mentor and critical friend' (David Lines). All were in favour of a variety of assessment methods but cautioned against over assessing, stressing that all such methods should be relevant to the subject and firmly and clearly aligned with learning and teaching outcomes.

In the final paragraph of his paper, David Lines states that 'Instead of emphasising the negative aspects, we should instead ask what will happen if we don't change'. The present writer would feel that we should also ask: 'What is it that is stopping us from moving forward?' And as will be seen later in this report, one of the prevailing questions from the concluding minute paper session is 'How?'

The second part of the conversation: the workshop participants' perspective

'Making judgements is part of life, whether the ones made are right or wrong, valid and reliable or not' (Heywood, 2000).

Two breakout sessions were conducted at the workshop and indeed were central to its programme, providing as they did an opportunity for participants to reflect upon and debate the issues raised by the workshop speakers as well as consider set fictional scenarios and key questions posed by QAA Scotland. At the conclusion of the day, a minute paper was distributed by Linda Suskie for completion by the participants as part of the final plenary session. What follows here is a distillation of the discussions, comments and questions received.

The breakout groups - Session 1

'If you aim to make it useful you will end up with real quality' (Linda Suskie, Stirling Workshop, 2004).

First topic

The first topic for discussion (evaluating effective communication, and ancillary skills such as organisation and mechanics, through writing a dissertation) provoked some very different reactions. One group reported back with recommendations for

obtaining accurate and truthful information and maintaining cost-effectiveness, while another challenged the proposed assessment mechanism itself, feeling quite strongly that it was in itself questionable, having the potential for exclusion of those with special needs, namely dyslexia, and that if a major dissertation was required it should be fully used for evaluating a whole range of skills and learning outcomes and not just the designated ones of organisation, focus, style and mechanics. All were agreed that it was important to establish what was being assessed, why it was being assessed, what we wanted students to achieve and that the proposed method of assessment should be fit for the purpose. The following suggestions were made.

- Explore different modes of assessment, attempting to balance both the needs of the student and that of the discipline concerned.
- Develop benchmarks (which may have a bearing on the type of criteria established).
- Establish a marking scheme and criteria which is clear to both staff and students.
- Be clear on what is expected by way of original work so as to stress the seriousness of plagiarism (a theme that was revisited later in the day).
- Specify clear, unambiguous learning outcomes, linking these clearly to marking criteria/marking rubric, shared with the students, to ensure that students are clear on what is being assessed and what is being looked for in their work for various 'levels' of mark (it was suggested that students should practice using them on their own work, or in the context of anonymous peer assessment within the group, or reviewing previous years' work).
- Incorporate double blind marking for all scripts and external examination, for at least a representative sample, to ensure consistency and accuracy of marking.
- Use a feedback template (based clearly on specified marking criteria) to ensure consistency of feedback to students and to simplify marking for tutors, ensuring that different tutors address similar issues in their marking and feedback; this could possibly incorporate a database of feedback comments (derived from generic issues arising).
- Provide feedback and support throughout the course/programme of study rather than expecting students to pour all their efforts into a final summative assessment.
- To this end, therefore, consider using smaller, shorter pieces of work which could build up into the final dissertation, providing the opportunity for formative feedback to students, reducing the marking burden for the first marker (who will have already seen much of the dissertation by the time the come to mark the final version); this could also be of relevance in minimising instances of plagiarism.
- Alternatively consider using shorter assignments which stand on their own thus eliminating extensive student and staff effort and maintaining cost effectiveness.
- Encourage students to work in peer support study groups to discuss and share generic issues relating to their dissertations.
- Embed skills assessment in other assignments to maximise cost-effectiveness, thus avoiding the need to devise additional assessments.

- Give generic feedback to the group as a whole, thus reducing the amount of feedback that has to be given on a one-to-one basis, as a further aid towards making the process more cost-effective.

The issue of protecting the students' privacy, and that of their professors, was considered and the following recommendations were put forward.

- Anonymous marking should be used wherever possible (although this may not be feasible where individual focus identifies individuals eg placements etc).
- Enabling students to view their marks individually eg through an online portal (rather than positing on a board - even in an anonymised format) will allow students to decide how widely they wish to share their marks.
- Results could be given by number, not by name.
- Anonymous peers' assessment is another option (and should be possible if the criteria are explicit and detailed), with the tutor moderating/marking a sample.
- Evaluating student feedback in an anonymised fashion, removing any reference to individual tutors (by name) will protect professors' anonymity in published evaluation results; specific issues relating to individual tutors should be taken up through an alternate, developmental route (such as appraisal) - not through 'naming and shaming'.
- Providing generic, anonymised feedback to the group (on collated issues) can permit for issues to be tackled without individuals being identified explicitly.
- On the other hand, individualised feedback may be necessary (see first bullet point above).

Second topic

'Who is more aware and, therefore, in a better position to assess their personal skills, qualities and attitudes than the students themselves?' (Bowen, 1988).

With regard to the second fictional scenario as to how students' tolerance for perspectives other than their own are assessed, the question was posed as whether or not this can be tested and how. It was acknowledged that this could be quite a passionate topic and one which could be quite difficult to measure; nonetheless some suggestions were put forward, namely that:

- students be set a task to research a viewpoint completely contrary to their own, and to report on this either through written work, or discussion, which would lead eventually to the production of a balanced argument which incorporated both their own and the contrary view
- as an effective way of achieving this, opportunities for students to experience another perspective before making a judgement on it could be provided, for example, going round campus in wheelchair to experience things from disabled student's point of view
- teachers/lecturers should lead by example, presenting contentious issues in a balanced and non-biased manner
- students should interview someone with the opposite viewpoint, or adopt that viewpoint within a role-playing debate, which would enable students to reflect

better on their own attitudes and perspectives. One participant suggested that there probably was no need to do this formally, as much of the above was frequently present in any student bar!

It was felt by some that assessing how far someone's attitude had changed, and in a way which yielded unbiased results, would be extremely difficult to achieve.

Third topic

'Self-esteem is an issue for many students and there is evidence that self and peer evaluation can help promote self-esteem' (Hunter, 2004).

Concerning the third topic of the session (the use of group work in researching a particular issue, with an oral presentation expected at the end of this assignment) the importance of clarity of intent and purpose at the start of such an exercise was reiterated. Again the fitness of one of the proposed assessment methods (oral presentation) was challenged, and a number of suggestions for supporting those students for whom such a method might prove difficult were made (a student with a speech impairment was part of the fictional group). However, many felt that, unless such a method of presentation is an integral part of the course, it should not be assessed at all. The following is a summary of the suggestions made.

That:

- students should be required to undertake/submit both a group and an individual piece of assessed work and should be required to pass both components
- peer marking (appropriately annotated) should be incorporated within any group work, enabling peers to indicate the contribution (by effort) made by members of the group
- mixing groups episodically may prove useful to enable patterns of student working to emerge, particularly for indicating where a student encounters problems in more than one group scenario
- students and assessors should be involved jointly in setting criteria, thus allowing students to have a role in determining how to achieve the learning outcomes, and generating a student contract for the assessed work; this could be particularly important in the case of students with special needs, though it was also stressed that appropriate advice and information should also be sought through alternate means (special needs advisory staff, learning needs identification documents produced in response to the student's disclosure of disability etc) to ensure that teachers/lecturers make informed decisions on student support needs
- to avoid unfair balance between 'workers' and 'passengers' within the group, many modes of assessment should be employed such as the use of group diaries, an attendance record, the setting of clear and equal goals to be achieved by each student, anonymous peer review, tutor moderation, one to one presentation, video recordings etc
- a mixture of marks should be used: group mark, peer assessed mark, individual mark

- opportunities to find out about the development of the final assignment, what ongoing work/discussions have taken place, how was the work allocated and completed, should be built into the course in order to get a sense of who was/wasn't involved (it was felt that this background material could be assessed, but wouldn't necessarily have to be)
- students should be encouraged to use online discussion boards, if these exist, to talk about their assignment, share ideas etc, and give the tutor access so they can use the facility to get a sense of how the group is working together, who is having an input and who isn't
- the student with the speech impairment should be asked if they wanted to participate in the delivery of the presentation (staff should not automatically assume they can't, or wouldn't want to); if arising from this the student felt that they couldn't, then the tutor should consult with the student as to what they would see as an alternative, and endeavour to provide it
- alternative methods of presentation should be explored such as video, Microsoft PowerPoint etc.

The breakout groups - Session 2

First topic

'Working towards a no-blind-eyes culture' (Jude Carroll, Stirling Workshop, 2004).

In the afternoon session, the subject of plagiarism was considered at length. Many agreed that in a lot of cases no-one knew just how prevalent the practice was, though it was noted that the number of known cases were on the rise. This acceleration was in part being aided through the use of the internet, which in turn made the evolution of new methods of detection through JISC or Copycatch a necessity, particularly in the case of electronically submitted assessments. It was also acknowledged that other factors can play a part in inadvertent plagiarism, such as cultural issues (academic conventions and cultural attitudes vary not just from country but within the UK itself), the outcomes-focussed approach of many staff and students and the pressure to get results, language problems, students with specific learning difficulties, lack of resources and lack of support for both staff and students in dealing with the issue.

All were agreed that communication and clarity about what constitutes plagiarism was essential at the start of any course for both staff and students; it was noted that many of the former were unsure about this and so such clarity was lacking - this made things particularly difficult when staff are then obliged to deal with suspected instances of plagiarism. While the need for informing students at the start of their course (through the use of whatever mode of induction employed by different institutions) was recognised it was agreed that this should not be limited to just the start of the course, but reinforced again and again, encouraging students to reflect, understand and consider their own practice. The necessity for a consistent approach across an institution as a whole was vital; one manifestation of this could be the use of academic misconduct officers to minimise the possibility of confrontation between individual staff and students. The following actions were recommended.

- Make assessment criteria explicit and establish clear definitions of plagiarism/collusion for both staff and students.
- Put necessary support in place for students (study skills, language support etc).
- Address assessment issues early, include advice on writing skills, how to avoid plagiarism etc at induction.
- Identify problems early and tackle them - don't ignore them.
- Institute a central policy to deal with plagiarism and collusion, which is clearly defined, and clearly communicated to everyone; in addition, establish institutional support so that staff will feel confident in dealing with cases, and eradicate a culture where cases are ignored.
- Following on from this, establish consistency of approach (vital but hard to achieve).
- Eliminate unnecessary rules and regulations which can complicate the issue and lead to confusion among staff and students.

Second topic

'The hardest task left is for the shift in culture to result in real change, change embraced for the value perceived, not for compliance' (Follett, 2003, cited by Karen Robson, Stirling Workshop, 2004)

The second topic for discussion, and one that seemed to pre-occupy most of the respondents to the minute paper at the conclusion of the workshop (see below), was the creation of an assessment culture which would provide reliability, validity and fairness, but which would also be consistent, transparent and fair to the students. However, as the discussion progressed it became clear that reliability, validity and fairness in relation to the tutors/professors was just as much as issue for the workshop participants as it was for their charges. In particular, the barriers preventing the establishment of such a culture were a major consideration and suggestions towards overcoming these were not always forthcoming. Some of these barriers were identified as follows.

- The threat to formative assessment from modularisation, which raised issues related to creating a culture that encourages a student-centred approach to learning. Precisely what threat or threats were not elucidated by this particular group though there are a number that could be conceived namely:
    a the fact that breaking a subject down into modules means that continuity, consistency, and a continuous learning curve is difficult to achieve, that it becomes impossible to cover anything in depth, and that the aim becomes one of just finishing the module (summative) rather than attempting a solid basis for an educational continuum

    b the danger therefore of loss of ownership by both staff and students of an inclusive and fair learning and teaching process (through formative assessment) to those who are only interested in results and figures, and that prevention of this can only be achieved by engagement of the support of senior managers/academic staff which is not always possible.

- Fear of the loss of ownership of the learning and teaching process by staff if students were to be involved in the design of their own learning.
- Some participants felt that the influence of professional accreditation bodies could stifle innovation and free-thinking in learning and teaching methods.
- Lack of the courage to be innovative: the need for innovation was viewed as being necessary (the example of the University of Edinburgh veterinary school was cited - this particular department has started to use case-based learning in the first year therefore exposing students to work-related learning from the outset of their course).
- Finally, the need to possibly re-educate colleagues and senior managers/staff was emphasised by several if a truly student-centred approach to learning is to be adopted as this would entail a shift in culture and practice.

Some incentives were suggested.

- As a means of overcoming the fear of failure among students (the 'prat' factor as one group termed it): it was observed that students engage well in tasks that are peer-assessed as they fear and respect criticism from their peers (a point that could be applied just as much to teachers and lecturers).
- Importance of the physical environment in which teaching and learning are conducted was stressed; indeed the question was asked as to how much learning goes on outside the programme/formal teaching hours. This in turn raised estates issues and the need to ensure such opportunities for informal learning and communication existed eg a common room/coffee machine at departmental/school level for staff and students.
- It was felt that a range of assessment methods was needed even in the first year of study and the importance of the first year curriculum (and in particular the importance of articulating and establishing clear assessment criteria in the first year of study) were cited as having a bearing the improvement of retention rates.

Third topic

'We assume, in a word, that the student has a right to be fairly assessed on what benefit he or she has taken from the discipline. On the other hand, it may well be that our own approach to assessment falls short of such an ideal…'
(Teachability, 2004).

Georgina Follett's comment cited earlier in this paper, together with the above, is no less valid when applied to the third topic for discussion, that of assessing students with special learning needs and disabilities, a student population which includes not only disabled students with specific learning difficulties but also international students (particularly problematic because academic conventions and cultural attitudes towards the correct use of sources vary) and access/non-traditional entrants. Indeed, many of the suggestions made in the previous discussion outlined above can be

applied here, though specific recommendations did emerge at the end of the afternoon's debate. The following needs were identified.

- A movement towards an anticipatory culture and away from the ad hoc/ compensatory model that is still prevalent.
- Where specific adjustments are required to be made, decisions concerning these to be made in a three-way partnership between the academics, specialists and the student.
- Such adjustments to operate systematically, within an assessment framework that aligns them to learning outcomes.
- The anticipation of, and therefore the engagement of curriculum design and validation with, a diversity of learners and needs including those with impairments, people who are ill prepared for learning and international students.
- Ongoing collaboration with students and enabling mutually negotiated assessment tasks and objectives within the course criteria for all students, supported by skills development throughout the course.
- A wider engagement with issues relating to supporting students learning in the curriculum, particularly in cases where plagiarism is viewed by some students as a coping strategy.
- Academics within disciplines to develop approaches that relate to core skills.
- More research evidence about inclusive assessment practices.
- A forum where practice is shared.
- A team approach to assessment and agreement about core course objectives which in turn generates understanding and 'buy in' from all participants.
- Assessment options which are available to all and supported by relevant practise and skill development.
- The separation of intellectual engagement with a course from fitness to practise issues as a possible way forward on some currently restrictive 'vocational' courses.

Concerns were raised, however, in the course of discussions, particularly in relation to the 'Fitness to practise' criteria and competency requirements of professional bodies which were experienced by many as being problematic, specifically in the areas of health and safety issues, professional responsibilities and, perhaps most worryingly, the question of who would be sued if something went wrong.

The minute paper response

'Resistance to change is normal' (Brown, Bull and Pendlebury, 1997).

At the end of the day, participants and speakers gathered for a final brief question and answer session based on minute papers distributed by Linda Suskie which invited comments on what had been learned from the workshop and questions. Many of the comments received indicated that participants had been somewhat reassured by their debates and discussions throughout the day and by what they had learned.

- 'That what my (what I believed!) "off the wall" methods of teaching and assessment are really ok!'
- 'Never say never - think laterally.'
- 'We should feel able to be innovative.'
- 'That there are many like-minded people with great assessment ideas around locally.'
- 'The idea of assessment of education and assessment for education.'
- 'The importance of linking learning outcomes to assessment, and the fact that I know this isn't always happening in courses at my institution.'
- 'Fairness does not equal equality.'
- 'We identified the need for a sector-wide forum on approaches to disability, to promote fairness and consistency and share good practice.'

However, the questions raised reiterated the fact that many still felt confronted by barriers preventing them moving forward and were unsure as to how these should be overcome; more tellingly, as will be seen below, the major barrier appeared to be a very human element.

- 'How to engage colleagues with the issues?'
- 'How do we spread the word beyond those already interested in/committed to excellence?'
- 'How to put things learned in practice within a "set in its ways" department?'
- 'What can I do initially to start to change my practice?'
- 'How do we encourage all in universities - staff and students - to engage with making assessment and reasonable adjustments to assessment - relevant to the learning objectives and needs?'
- 'How to work on the issue of getting my institution to link learning outcomes to assessment?'
- 'How do you implement a variety of assessment and imagination while attempting to reduce over assessment?'
- 'How best to engage senior management in cultural change to encourage innovative teaching/assessment methods?'
- 'How will I gain support from senior management for taking risks (by innovating) - they are very risk averse?'

Nonetheless, several participants indicated their wish to prolong this debate beyond the conclusion of the workshop, thereby ensuring the sharing of good practice, developments of new ideas and solutions to problems continues.

The third part of the conversation: the students' opinions

'[Assessment] is not a one-way street' (David Lines, Stirling Workshop, 2004).

Students' perspectives on assessment, which were expressed through the scoping survey conducted by QAA Scotland during February/March 2004, appeared to support many of the points made by both speakers and participants at the workshop

(in the examples which follow, quotes from the overall report appear in italics, while those from individuals are in quotation marks). For example, on the value of written examinations, several responses highlighted that unseen time-constrained written examinations tended not to assess the learning that had taken place but encouraged cramming (soon forgotten) and regurgitation.

'Yes, it enables me to demonstrate about half of what I've learnt, but the other half, although expressed during tutorials, seems a bit wasted', thus apparently feeling that end of course summative assessments were not the best way of eliciting what the student had learned.

Many students wanted more feedback.

Quite strikingly, the usefulness of feedback and formative assessment - and a need for more of it - were explicitly mentioned by some respondents and hinted at by several others in justifying the other methods they would like to see in practice. Students liked to know they were 'on the right track'.

'Good feedback is quickly available, is individual, and contains clear directions for action/learning.'

'Feedback is as important, if not more so, than a mark - all assessment including final exams, I think, can be part of the learning process, as well as a "hurdle", but feedback is vital to make this possible.'

A wish for more varied forms of assessment was also a recurrent theme. More group work, peer-assessments, presentations, report writing, multiple-choice, open-book exams, take-home exams and need for feedback all featured. It was striking that a number of respondents were clearly thinking of these methods of assessment as being useful also in terms of the acquisition of skills which would be useful to them later.

Arising from this the strengths of group work and peer-assessment were also mentioned.

'Peer-assessment gives you an idea of how assessment works, range of standards, understanding of what gains/loses marks' and 'Group work is useful as you have to think on your feet to solve problems'.

Once again the issue of fairness was highlighted.

The terms 'fairness' and 'unfairness' featured in a number of responses. Some students believed that a variety of methods of assessment meant the system was fairer to students in general, who all have different strengths and weaknesses. Students have a strong sense that systems and processes should be fair and there is a sense that diversity of approaches promotes this.

Clarity and consistency were also concerns.

Inconsistencies in practice and in unclear provision of information about assessment were mentioned. Students needed such guidance to be clear and unambiguous and students believe that approaches and processes should be made clear to them and applied consistently.

Despite the fact that the opinions of the workshop participants and students were elicited on separate occasions, and by different methods, there is clear coincidence of views between the two groups and evidence of a basis for partnership in assessment - the next stage of the conversation should surely now take place between the two groups face to face.

Conclusion

'It's about going out there and trying it' (Professor Simon van Heyningen [Chair, Quality Enhancement Themes Steering Committee for Assessment], Stirling Workshop, 2004).

Earlier in this report the current writer asked 'What is it that is preventing us from moving forward?' There are a few more questions to be added to that one such as 'Is it a question of different educational generations, entrenched viewpoints and mindsets?' Arising from that, is it a fear of giving students or individual staff members too much power? Is it a fear of league tables/pressure to get results/risk taking? Is it lack of time/lack of finance and resources/lack of willingness? Are we waiting for the 'results' of changing methods of assessment themselves to be tested, to feed through and be surveyed? Or is it 'tiresome'? (Knight, 1995). And, having identified potential barriers against change, there remains the question of 'how?'

Clearly there is need for more conversations (and for that read vigorous debates!) to take place between staff and students, between staff and senior management, between staff and colleagues either within or outwith institutions, between staff and employers. It is also clear from the workshop, and indeed reports of other workshops in this series, that some of these are already underway and the vast body of evidence which will be amassed by this enhancement series can only add weight to the arguments and drive necessary to effect meaningful change. More research is required in some areas, notably those of special educational needs and disability, and electronic plagiarism, but we need more than conversations, research and literature - these are only the supports to the decisive action (urged by Simon van Heyningen at the end of the workshop) that must be taken if David Lines' predicted continuation of shallow learning is to be avoided. But perhaps (lest the conclusion of what proved to be a most productive day be considered all doom and gloom) the last word on this should go to Linda Suskie.

'[If you have] confidence in results enough to make changes in future delivery then that is quality/truthfulness enough' (Linda Suskie, Stirling Workshop, 2004).

References

Bowen J (1988) Student Self-Assessment, in Brown S (ed) *Assessment - A Changing Practice,* (*Professional Issues in Education*), Vol 2, Scottish Academy Press, Edinburgh

Brown G, Bull J and Pendlebury M (1997) *Assessing Student Learning in Higher Education*, Routledge, London

Follett G (2003) The place of assessment in quality enhancement in Scotland, *Exchange,* Issue 4, Spring 2004

Heywood J (2000) *Assessment in Higher Education - Student Learning, Teaching, Programmes and Institutions*, Jessica Kingsley Publishers, London

Hunter D (2004) *How am I doing? Valuing and Rewarding Learning in Musical Performance in Higher Education*, University of Ulster, Faculty of Arts, Centre for Creative and Performing Arts, Londonderry

Knight P T (ed) (1995) *Assessment for Learning in Higher Education*, Kogan Page, London

Scottish Higher Education Funding Council (2004) Creating accessible examinations and assessments for disabled students - booklet produced for *Teachability project: Creating an Accessible Curriculum for Students with Disabilities,* University of Strathclyde, Glasgow (email: a.simpson@mis.strath.ac.uk)

Assessment workshop series - No 7

Improving feedback to students (link between formative and summative assessment)

Contents **Page**

Rethinking formative assessment in higher education: a theoretical model and seven principles of good feedback practice

Dr David Nicol, Centre for Academic Practice, University of Strathclyde and Debra Macfarlane-Dick, Careers Service, University of Glasgow

Introduction

This briefing paper explores how higher education (HE) institutions might use assessment more effectively to promote student learning. Assessment provides a framework for sharing educational objectives with students and for charting their progress. However, it can generate feedback information that can be used by students to enhance learning and achievement. This feedback information can also help teachers re-align their teaching in response to learners' needs. When assessment serves these purposes it is called 'formative assessment'. It is argued that formative assessment should be an integral part of teaching and learning in HE and that feedback and feedforward should be systematically embedded in curriculum practices.

Formative assessment aids learning by generating feedback information that is of benefit to students and to teachers. Feedback on performance, in class or on assignments, enables students to restructure their understanding/skills and build more powerful ideas and capabilities. However, the provision of feedback information is not the sole province of the teacher. Peers often provide feedback, for example, in group work contexts, and students generate their own feedback while engaging in and producing academic work (see below). Formative assessment also provides information to teachers about where students are experiencing difficulties and where to focus their teaching efforts.

This paper summarises the research on formative assessment and feedback. It includes the following.

- A conceptual model of the formative assessment/feedback cycle.
- Seven principles of good feedback practice: these are drawn from the model and a review of the research literature.
- Some examples of good practice strategies related to each principle.

There are two central arguments within this paper:

i that formative assessment and feedback should be used to empower students as self-regulated learners and

ii that more recognition should be given to the role of feedback on learners' motivational beliefs and self-esteem.

A number of writers have argued that feedback is under-conceptualised in the theoretical literature in HE and elsewhere, and that this makes it difficult to design effective feedback practices or to evaluate their effectiveness (Yorke, 2003; Sadler, 1998). While there has been a move over the last decade to conceptualise learning from a constructivist perspective (eg Laurillard, 2002), approaches to feedback have, until recently, remained obstinately focused on simple 'transmission' perspectives. Teachers 'transmit' feedback messages to students about strengths and

weaknesses in their work assuming that these messages are easily decoded and turned into action. In contrast, in this paper, students are assumed to construct actively their own understanding of feedback messages from tutors. Moreover, these messages are assumed to be complex and difficult to decipher (Higgins, Hartley and Skelton, 2001; Ivanic, Clark and Rimmershaw, 2000).

The conceptual model and the seven principles presented in this paper are intended as tools that teachers might use to analyse and improve their own formative assessment and feedback practices.

A conceptual model

In a review article, Black and Wiliam (1998) drew together over 250 studies of formative assessment with feedback carried out since 1988 spanning all educational sectors. The studies that formed part of their meta-analysis were ecologically valid ie they were drawn from real teaching situations. Black and Wiliam's analysis of these studies showed that feedback resulted in positive benefits on learning and achievement across all content areas, knowledge and skill types and levels of education. One of the most influential papers underpinning the Black and Wiliam review, and the writings of other researchers, is that by Sadler (1989). Sadler identified three conditions necessary for students to benefit from feedback. The student must:

a possess a concept of the goal/standard or reference level being aimed for

b compare the actual (or current) level of performance with that goal or standard

c engage in appropriate action which leads to some closure of the gap.

Sadler argued that in many educational settings teachers give students feedback information on b, ie how their performance compares to the standard, but this feedback often falls short of what is actually necessary to help students close the gap. For example, such information might be difficult to understand (eg a comment such as 'this essay is not sufficiently analytical') and especially if the learning goal a has not been fully assimilated in the first place. Black and Wiliam (1998) further elaborate on this communication issue when they discuss the links between the way a feedback message is received and what students do with that message.

> '...those factors which influence the reception of a [feedback] message and the personal decision about how to respond…[include]….beliefs about the goals of learning, about one's capacity to respond, about the risks involved in responding in various ways and about what learning should be like.'

Any model of feedback must take account of the way students make sense of, and use, feedback information. More importantly, however, is Sadler's argument that for students to be able to compare actual performance with a standard, and take action to close the gap, then they **must already possess some of the same evaluative skills as their teacher.** For many writers, this observation has led to the conclusion that as well as focusing on the quality of the feedback messages teachers should focus their efforts on strengthening the skills of self-assessment in their students (Yorke, 2003; Boud, 2000).

Figure 1 presents a conceptual model of formative assessment and feedback that synthesises current thinking by key researchers into this topic (Sadler, 1983, 1989; Black and Wiliam, 1998; Yorke, 2003; Torrance and Pryor, 1998). The figure is based on a model of feedback and self-regulated learning originally published by Butler and Winne (1995). A key feature in the model that differentiates it from commonplace understandings of feedback is that the student is assumed to occupy a central and active role in all feedback processes. They are always actively involved in monitoring and regulating their own performance both in terms of their goals and in terms of the strategies being used to reach those goals.

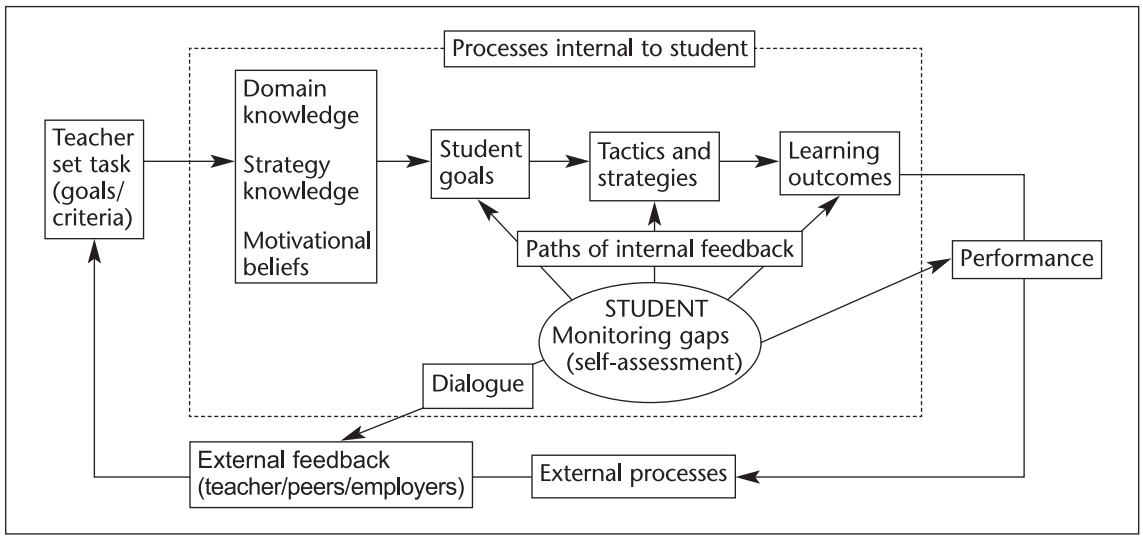


Figure 1 A model of the formative assessment and feedback

In the model, an academic task set by the teacher (in class or set as an assignment) is the starting point for the feedback cycle. Engagement with the task requires that students draw on prior knowledge and motivational beliefs and construct a personal interpretation of the requirements and properties of the task. Based on this internal conception, they formulate their own task goals (which may be different from those of the teacher) and engage in actions to achieve these goals by applying tactics and strategies that generate outcomes. Monitoring these interactions with the task and the outcomes that are being cumulatively produced, generates **internal feedback.** This feedback is derived from a comparison of current progress against internal goals or standards - gaps are identified (between progress and goals) and further actions are taken to close these gaps (Sadler, 1989). This self-generated feedback information might lead to a reinterpretation of the task or to the adjustment of internal goals or of tactics and strategies. Students might even revise their domain knowledge or beliefs which, in turn, would influence subsequent processes of self-regulation. If **external feedback** is provided, this additional information might augment, concur or conflict with the student's interpretation of the task and the path of learning (Butler and Winne, 1995).

In the model, external feedback to the student might be provided by teachers, peers or others (eg placement supervisor). However, students are always actively engaged in feedback processes. First, they generate aspects of their own feedback as they

monitor performance and identify and make sense of gaps while carrying out tasks. Second, they interpret and filter feedback information from external sources. The teacher's feedback response (based on their monitoring and assessment of student performance) must be interpreted and internalised by the student before it can influence subsequent action (Ivanic, Clark and Rimmershaw, 2000). This has important implications for feedback processes in HE. If students are always involved in monitoring and assessing their own work, then rather than just thinking of ways of enhancing the teacher's ability to deliver high quality feedback we should be devising ways of building upon this capacity for self-regulation (Yorke, 2003).

Seven principles of good feedback practice

From the conceptual model and the research literature on formative assessment it is possible to identify some broad principles of good feedback practice. A provisional list might include the following seven.

Good feedback practice

1 Facilitates the development of self-assessment (reflection) in learning.

2 Encourages teacher and peer dialogue around learning.

3 Helps clarify what good performance is (goals, criteria, expected standards).

4 Provides opportunities to close the gap between current and desired performance.

5 Delivers high quality information to students about their learning.

6 Encourages positive motivational beliefs and self-esteem.

7 Provides information to teachers that can be used to help shape the teaching.

The following sections provide the rationale for each principle in terms of the conceptual model and the associated research literature. Brief examples of how these principles might be applied are also suggested.

1 Facilitates the development of self-assessment (reflection) in learning

Over the last decade there has been an increasing interest in strategies that encourage students to take a more active role in the management of their own learning (see Nicol, 1997). Black and Wiliam (1998) make the argument that 'a student who automatically follows the diagnostic prescription of a teacher without understanding of its purpose will not learn' while Sadler (1989) argues that the purpose of formative assessment should be to equip students gradually with the evaluative skills that their teachers' possess. These writers are concerned that an overemphasis on teacher assessment might increase students' dependency on others rather than develop their ability to self-assess and self-correct.

In the conceptual model, the student or learner is always engaged in **monitoring gaps** between internally **set task and personal goals** and the **outcomes** that are being progressively produced. This monitoring is a by-product of purposeful engagement in a task. However, in order to build on this process, and the student's

capacity for self-regulation, teachers should create more formal and structured opportunities for self-monitoring and the judging of progression to goals. Self-assessment tasks are a good way of doing this, as are activities that encourage reflection on both the processes and the products of learning.

Research shows that direct involvement by students in assessing their own work and frequent opportunities to reflect on goals, strategies and outcomes are highly effective in enhancing learning and achievement (McDonald and Boud, 2003). Moreover, if the skills of self-assessment are developed progressively over the course of an undergraduate degree this would support a model of HE where students are prepared for lifelong learning (Boud, 2000).

An important aspect of self-assessment involves helping students both to identify standards/criteria that apply to their work and to make judgements about how their work relates to these standards (Boud, 1986).

Examples of structured reflection and/or self-assessment are varied and might include students:

1 requesting the kinds of feedback they would like when they hand in work

2 identifying the strengths and weaknesses in their own work in relation to criteria or standards before handing it in for teacher feedback

3 reflecting on their achievements and selecting work in order to compile a portfolio

4 setting achievement milestones for a task and reflecting back on progress and forward to the next stage of action.

Having students give feedback on each other's work (peer feedback) also helps support the development of self-assessment skills (eg Gibbs, 1999).

2 Encourages teacher and peer dialogue around learning

While research shows that teachers have a central role in helping develop student's own capacity for self-assessment in learning, external feedback from other sources, for example, tutors or peers is also crucial. Feedback from tutors and peers provides additional information that helps challenge students to reassess their knowledge and beliefs. Teacher feedback also serves as an authoritative external reference point against which students can evaluate, and self-correct their progress and their own internal goals.

In the conceptual model (Figure 1), for external feedback to be effective it must be understood and internalised by the student before it can be used productively. Yet in the research literature (Chanock, 2000; Hyland, 2000) there is a great deal of evidence that students do not understand the feedback given by tutors (eg 'this report is not logically structured') and are therefore not be able to take action to close the gap (ie they may not know what to do to make the report more 'logical in structure'). External feedback as a transmission process involving 'telling' ignores the active role the student must play in constructing meaning from feedback messages.

One way of increasing the effectiveness of external feedback and the likelihood that the information provided is understood is to conceptualise feedback more as a **dialogue** rather than as information transmission. Feedback as dialogue means that the student not only receives initial feedback information but also has the opportunity to engage the teacher in discussion about that feedback. This is shown in the conceptual model by the two-way arrows that link external processes to those internal to the student. The idea that feedback encourages dialogue, is considered good practice by many writers on assessment. For example, Freeman and Lewis (1998) argue that the teacher 'should try to stimulate a response and a continuing dialogue - whether this be on the topics that formed the basis of the assignment or aspects of students' performance or the feedback itself'. Discussions with the teacher help students to develop their understanding of expectations and standards, to check out and correct misunderstandings and to get an immediate response to difficulties.

Unfortunately, with large class sizes, it can be difficult for the teacher to engage in dialogue with students. Nonetheless, there are ways that teachers might increase feedback dialogue even in these situations. For example, by reporting feedback in class and structuring break out discussions of feedback or by using classroom technologies that collate student responses in-class and then feed the results back visually as a histogram. This feedback can act as a trigger for teacher-managed discussion (eg Nicol and Boyle, 2003).

Another source of external feedback are the students themselves. Peer dialogue is beneficial to student learning in a variety of ways. First, students who have just learned something are often better able than teachers to explain it to their classmates in a language and in a way that is accessible. Second, peer discussion exposes students to alternative perspectives on problems and to alternative tactics and strategies. Alternative perspectives enable students to revise or reject their initial hypothesis and construct new knowledge and meaning through negotiation. Thirdly, by commenting on the work of peers, students develop objectivity of judgement (about work in relation to standards) which can be transferred to the assessment of their own work (eg 'I didn't do that either'). Fourthly, peer discussion can be motivational in that it encourages students to persist and gives a yardstick to measure their own performance against (see Nicol and Boyle, 2003). Finally, it is sometimes easier for students to accept critiques of their work from peers rather than tutors.

Good examples of feedback dialogue in class include:

1 providing feedback using one-minute papers (Cross and Angelo, 1990)

2 reviewing feedback in tutorials where students are asked to read the feedback comments they have been given and discuss these with peers (they might also be asked to suggest strategies to improve performance next time)

3 asking students to find one or two examples of feedback comments that they found useful and to explain how they helped.

Other ways of using feedback dialogue in a planned way, for assignments, might involve:

1 having students give each other descriptive feedback on their work in relation to published criteria before submission and

2 group projects.

3 Helps clarify what good performance is (goals, criteria, expected standards)

Students can only achieve a learning goal if they understand that goal, assume some ownership of it, and can assess progress (Sadler, 1989; Black and Wiliam, 1998). In the model (Figure 1), understanding the goal means that there must be a reasonable degree of overlap between the task goal set by the student and the goal originally set by the teacher. However, there is considerable research evidence to suggest that there are often mismatches between tutors' and students' conceptions of goals and of assessment standards and criteria.

Hounsell (1997) has shown that tutors and students often have quite different conceptions about the goals and criteria for essays in undergraduate courses in history and psychology and that poor essay performance is correlated with the degree of mismatch. In a similar vein, Norton (1990) has shown that when students were asked to rank specific assessment criteria for an essay task they produced quite different rankings from those of their teachers. Weak and incorrect conceptions of goals not only influence what students do but also the value of feedback information. If students do not share (at least in part) their tutor's conceptions of assessment goals (criteria/standards) then the feedback information they receive is unlikely to 'connect' (Hounsell, 1997). In this case, it will be difficult for students to evaluate gaps between required and actual performance.

One way of clarifying task requirements (goals/criteria/standards) is to provide students with written documents embodying descriptive statements that externalise assessment goals and the standards that define different levels of achievement. However, many studies have shown that it is difficult to make explicit assessment criteria and standards through written documentation or through verbal descriptions in class (Rust, Price and O'Donovan, 2003). Most criteria for complex tasks are difficult to articulate; they are often 'tacit' and unarticulated in the mind of the teacher. As York (2003) notes:

'Statements of expected standards, curriculum objectives or learning outcomes are generally insufficient to convey the richness of meaning that is wrapped up in them' (York, 2003).

Hence there is a need for strategies that complement written materials and simple verbal explanations. An approach that has proved particularly powerful in clarifying goals and standards has been to provide students with 'exemplars' of performance (Orsmond, Merry and Reiling, 2002) alongside other resources. Exemplars are effective because they define an objective and valid standard against which students can compare their work.

Strategies that have proved effective in clarifying criteria, standards and goals therefore include:

1. providing better definitions of requirements using carefully constructed criteria sheets and performance level definitions
2. providing students with exemplar assignments with attached feedback
3. increasing discussion and reflection about criteria and standards in class
4. involving students in assessment exercises where they mark or comment on other students' work in relation to defined criteria and standards
5. workshops where students in collaboration with teacher devise their own assessment criteria for a piece of work.

Combinations of the above five have proved particularly effective.

4 Provides opportunities to close the gap between current and desired performance

According to Yorke (2003), two questions might be asked regarding external feedback. First, is the feedback of the best quality and second, does it lead to changes in student behaviour? Many researchers have focused on the first question but the second is equally important. External feedback provides an opportunity to close the gap in the learning process between the current learning achievements of the student and the goals set by the teacher. If feedback information is not turned into action soon after it is produced then this is a missed opportunity. As Boud notes:

> 'The only way to tell if learning results from feedback is for students to make some kind of response to complete the feedback loop (Sadler, 1989). This is one of the most often forgotten aspects of formative assessment. Unless students are able to use the feedback to produce improved work, through for example, re-doing the same assignment, neither they nor those giving the feedback will know that it has been effective' (Boud, 2000).

In the conceptual model (Figure 1), Boud's arguments about closing the gap can be viewed in two ways. First, closing the gap is about supporting students while engaged in the act of production of a piece of work. Second, it is about providing opportunities to repeat the same 'task-performance-feedback cycle' by, for example, allowing resubmission. External feedback should support both processes: it should help students to recognise the next steps in learning and how to take them both during production and for the next assignment.

Supporting the act of production requires the generation of concurrent or intrinsic feedback that students can interact with while engaged in an assessment task. This feedback would normally be built into the task (eg a group task with peer interaction is an example here) or the task might be broken down into components each associated with its own feedback. Many forms of electronic feedback can be automatically generated to support task engagement (multiple-choice, frequently asked questions). Providing feedback at sub-task level is not significantly different from other forms of feedback described in this paper.

In HE, most students have little opportunity to use directly the feedback they receive to close the gap especially in the case of planned assignments. Invariably they move on to the next assessment task soon after feedback is received. While not all work can be re-submitted, many writers argue that resubmissions should play a more prominent role in learning (Boud, 2000). In addition, the external feedback provided to students often focuses on identifying specific errors rather than providing constructive advice about how performance relates to standards and about how to make improvements in subsequent tasks; and even when corrective guidance about how to improve is given students often do not fully understand it or know how to turn it into action.

Specific strategies to help students use external feedback to close the gap are:

1 to increase the number of opportunities for resubmission

2 for teachers to model the strategies that might be used to close a performance gap in class (eg model how to structure an essay when given a new question)

3 teachers might also write down some 'action points' alongside the normal feedback they provide. This would identify for students what they should do next time to improve their performance

4 a more effective strategy might be to involve students in identifying their own action points in class based on the feedback they have just received. This would integrate the process into the teaching and learning situation and involve the students more actively in the generation and planned use of feedback.

5 Delivers high quality information to students about their learning

Another finding from the research is that a great deal of external feedback given to students is not of good quality: it may be delayed, not relevant or informative or over-whelming in quantity etc. Good quality external feedback is defined as information that helps students troubleshoot their own performance and take action to close the gap between intent and effect. In the model (Figure 1), processes internal to the student (shown by the dotted line) are strongly influenced by contextual factors in the environment over which the teacher has considerable control. The teacher sets the task, assesses performance and provides feedback. Research shows that in each of these areas there is considerable scope for improvement.

Feedback needs to be relevant to the task in hand and to student needs. Despite this, research shows that feedback information is often about strengths and weaknesses in handed-in work or about aspects of performance that are easy to identify (eg spelling mistakes), rather than about aspects that are of greater importance to academic learning but that are more abstract and difficult to define (eg strength of argument).

Students might also receive too much feedback making it difficult to decide what to act on. In the literature on essay assessment, researchers have tried to formulate guidelines regarding the quantity and tone of feedback comments. For example, Lunsford (1997) has advocated providing only three well thought out feedback comments per essay. Moreover, these comments should indicate to the student how the reader experienced the essay as it was read (ie playing back to the students how

the essay worked) rather than offer judgemental comments. Such comments help the student to understand the difference between his or her intentions and the effects. Comments should always be written in a non-authoritative tone and where possible they should offer corrective advice (both about the writing process as well as about content) instead of just information about strengths and weaknesses.

Other researchers have argued against following positive comments with lists of criticisms (eg this essay was well-structured…however…) arguing instead that descriptive information about performance in relation to defined assessment criteria is better received by students and is more likely to be acted upon.

It has become common practice in recent years to provide feedback sheets with assessment criteria as a way of informing students about task requirements and of providing consistent feedback in relation to expected goals. However, the construction of such feedback sheets does not always encourage students to engage with a task in a way desired by teachers. Sadler has argued that the use of such criteria sheets often has unwanted effects: for example, if there are a large number of criteria (12-20) they may convey a conception of an assessment task (eg essay) as a list of things to be done (ticked off) rather than as a holistic process (eg involving the production of a coherent argument supported by evidence). So as well as being responsive to student needs, teachers should also consider whether the instruments they use to deliver feedback are commensurate with the expected goals and task requirements.

Strategies that increase the quality of feedback drawn from research include:

1 making sure that feedback is provided in relation to pre-defined criteria but paying particular attention to the number of criteria

2 providing feedback soon after a submission

3 providing corrective advice not just information on strengths/weaknesses

4 limiting the amount of feedback so that it is used

5 prioritising areas for improvement

6 providing online tests so that feedback can be accessed anytime, any place and as many times as students wish

7 focusing on students with greatest difficulties.

6 Encourages positive motivational beliefs and self-esteem

How can we make assessment a positive learning experience for students? A key feature of the model of feedback (Figure 1) presented in this paper is the importance attached to motivational beliefs and self-esteem. In the model, students construct their own motivation based on their appraisal of the teaching, learning and assessment context. This influences the goals that students set (personal and academic) as well as their commitment to these goals. However, research has shown that external feedback can have a positive or negative effect on motivational beliefs and on self-esteem. It influences how students feel about themselves which, in turn, affects what and how they learn.

Many studies have shown that, contrary to expectation, frequent high stakes assessment (where marks or grades are given) can lower the motivation to learn (Harlen and Crick, 2003). Such assessments encourage students to focus on performance goals (passing the test) rather than learning goals (Elliot and Dweck, 1988). In one study, Butler (1988) demonstrated that feedback comments alone improved students' subsequent interest in learning and performance when compared with controlled situations where marks alone or feedback and marks were given. Butler argued that students paid less attention to the comments when given marks and consequently did not try to use the comments to make improvements.

Butler (1987) has also argued that grading student performance has less effect than feedback comments because it leads students to compare themselves against others (ego-involvement) rather than to focus on the difficulties in the task and on making efforts to improve (task-involvement). Feedback given as grades has also been shown to have especially negative effects on the self-esteem of low ability students (Craven et al, 1991).

Dweck (2000) has interpreted some of these findings in terms of a developmental model that differentiates students into those who believe that ability is fixed and that there is a limit to what they can achieve (the 'entity view') and those that believe that their ability is malleable and depends on the effort that is input into a task (the 'incremental view'). These views affect how students respond to learning difficulties. Those with an entity view (fixed) interpret failure as a reflection of their low ability and are likely to give up whereas those with an incremental view (malleable) interpret this as a challenge or an obstacle to be overcome.

These motivational beliefs, however, are not immutable. In part, they depend on how teachers provide feedback. Praising effort and strategic behaviours and focusing students on learning goals leads to higher achievement than praising ability or intelligence which can result in a learned-helplessness orientation. In summary, 'feedback which draws attention away from the task and towards self-esteem can have a negative effect on attitudes and performance' (Black and Wiliam, 1998).

The implication of these studies for teaching practice is that motivation and self-esteem are more likely to be enhanced when a course has many low-stakes tasks with feedback geared to providing information about progress and achievement rather than high stakes summative assessment tasks where information is only about success or failure or about how students compare with peers. Other strategies that would help encourage high levels of motivation to succeed include:

1 providing marks on written work only after students have responded to feedback comments
2 allocating time for students to rewrite selected pieces of work - this would help change students' expectations about purpose
3 automated testing with feedback
4 drafts and resubmissions.

7 Provides information to teachers that can be used to help shape the teaching

Good feedback practice is not only about providing good information to the students about learning but it is also about providing good information to teachers. As Yorke (2003) notes:

'The act of assessing has an effect on the assessor as well as the student. Assessors learn about the extent to which they [students] have developed expertise and can tailor their teaching accordingly' (Yorke, 2003).

In order to produce feedback that is relevant and informative, teachers themselves need good data about how students are progressing. They also need to be involved in reviewing and reflecting on this data and in taking action to help close the learning gap.

In the conceptual model (Figure 1) information about students is provided when the learning outcomes are translated into public performances. Teachers generate this public information about students through a variety of methods - by setting assessment tasks and in-class through questioning of students and through observation. Such information helps teachers uncover student difficulties with subject matter (eg conceptual misunderstandings) and difficulties with study methods while carrying out assessment tasks.

Frequent assessment tasks, especially diagnostic tests, can help teachers generate cumulative information about students' levels of understanding and skill so that they can adapt their teaching accordingly. This is one of the key ideas behind the work of Angelo and Cross (1990) in the US. They have shown how teachers can gain regular feedback information about student learning within large classes by using short test-feedback cycles. These strategies benefit both the student and the teacher (Steadman, 1998) and they can be adapted to any classroom situation or discipline. Moreover, implementation allows teachers and students to share, on a regular basis their conceptions about both the goals and processes of learning (Stefani and Nicol, 1997).

A variety of strategies are available to teachers to help generate and collate quality information about student learning and help them decide how to use it. For example:

1 one-minute papers where students carry out a small assessment task and hand this in anonymously at the end of a class (eg what was the main point of this lecture? what question remains outstanding for you at the end of this teaching session?)

2 having students request the feedback they would like when they make an assignment submission

3 having students identify where they are having difficulties when they hand in assessed work

4 asking students in groups to identify 'a question worth asking', based on prior study, that they would like to explore for a short time at the beginning of the next tutorial

5 quick evaluation strategies at key points in teaching.

References

Angelo T and Cross P (1990) *Classroom Assessment Techniques*, Jossey-Bass, New York

Black P and Wiliam D (1998) Assessment and classroom learning, *Assessment in Education*, 5 (1) 7-74

Boud D (1986) *Implementing Student Self-Assessment,* Higher Education Research and Development Society of Australia, Sydney

Boud D (2000) Sustainable assessment: rethinking assessment for the learning society, *Studies in Continuing Education,* 22 (2) 151-167

Butler D L and Winne P H (1995) Feedback and self-regulated learning: a theoretical synthesis, *Review of Educational Research*, 65 (3) 245-281

Butler R (1987) Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance, *Journal of Educational Psychology*, 78 (4) 210-216

Butler R (1988) Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and involvement, *British Journal of Educational Psychology*, 58, 1-14

Chanock K (2000) Comments on essays: do students understand what tutors write? *Teaching in Higher Education*, 5 (1) 95-105

Craven R G, Marsh H W and Debus R L (1991) Effects of internally focused feedback on the enhancement of academic self-concept, *Journal of Educational Psychology,* 83 (1) 17-27

Dweck C (2000) *Self-theories: Their Role in Motivation, Personality and Development,* Psychology Press, Philadelphia

Dweck C and Elliot E (1988) Goals: An approach to motivation and achievement, *Journal of Personality and Social Psychology*, 54, 5-12

Freeman R and Lewis R (1998) *Planning and Implementing Assessment*, Kogan Page, London

Gibbs G (1999) Using assessment strategically to change the way students learn, in S Brown and A Glasner (eds) *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches,* Society for Research into Higher Education/Open University Press, Buckingham

Harlen W and Crick R D (2003) Testing and motivation for learning, *Assessment in Education*, 10 (2) 169-207

Higgins R, Hartley P and Skelton A (2001) Getting the message across: the problem of communicating assessment feedback, *Teaching in Higher Education,* 6 (2) 269-274

Hounsell D (1997) Contrasting conceptions of essay-writing, in F Marton, D Hounsell and N Entwistle (eds) *The Experience of Learning*, Scottish Academic Press, Edinburgh

Hyland P (2000) Learning from feedback on assessment, in A Booth and P Hyland (eds) *The practice of university history teaching*, Manchester University Press, Manchester

Ivanic R, Clark R and Rimmershaw R (2000) What am I supposed to make of this? The messages conveyed to students by tutors' written comments, in M R Lea and B Stierer (eds) *Student Writing in Higher Education: New Contexts*, SHRE/Open University Press, Buckingham

Laurillard D (2002) *Rethinking University Teaching: a conversational framework for the effective use of learning technologies*, second edition, RoutledgeFalmer, London

Lunsford R (1997) When less is more: principles for responding in the disciplines, in M Sorcinelli and P Elbow (eds) *Writing to learn: strategies for assigning and responding to writing across the disciplines*, Jossey-Bass, San Francisco

McDonald B and Boud D (2003) The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations, *Assessment in Education*, 10 (2) 209-220

Nicol D J (1997) *Research on Learning and Higher Education Teaching*, UCoSDA Briefing Paper 45, Universities and Colleges Staff Development Agency, Sheffield

Nicol D J and Boyle J T (2003) Peer instruction versus class-wide discussion in large classes: a comparison of two interaction methods in the wired classroom, *Studies in Higher Education,* 28 (4) 457-473

Norton L S (1990) Essay writing: what really counts? *Higher Education,* 20 (4) 411-42

Orsmond P, Merry S and Reiling K (2002) The use of formative feedback when using student derived marking criteria in peer and self-assessment, *Assessment & Evaluation in Higher Education,* 27 (4) 309-323

Rust C, Price M and O'Donovan B (2003) Improving students' learning by developing their understanding of assessment criteria and processes, *Assessment and Evaluation in Higher Education*, 28 (2) 147-164

Sadler D R (1983) Evaluation and the improvement of academic learning, *Journal of Higher Education*, 54 (1) 60-79

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

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

Steadman M (1998) Using classroom assessment to change both learning and teaching, *New Directions for Teaching and Learning,* 75, 23-35

Stefani L and Nicol D (1997) From teacher to facilitator of collaborative enquiry, in S Armstrong, G Thompson and S Brown (eds) *Facing up to Radical Changes in Universities and Colleges*, Kogan Page, London

Torrance H and Pryor J (1998) *Investigating formative assessment: teaching, learning and assessment in the classroom*, Open University Press, Philadelphia

Yorke M (2003) Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice, *Higher Education,* 45 (4) 477-501

Understanding the economies of feedback: balancing supply and demand

Dr Randy Swing, Policy Center on the First Year of College, Brevard College, North Carolina

For those of us who are successful learners and self-motivated to acquire new knowledge, the idea that feedback is part of the natural process of learning appears self-evident. Our idealised picture of the eager learner may be that of a young child running to their parent asking if they like the drawing the child made in school that day. Yet the ideal clashes with the reality when the eagerness for feedback of young children is contrasted with the actual reactions of many students in college classrooms.

Much that can be learned by simply observing our students as feedback is presented to them. Envision a professor returning a graded paper or essay exam to a classroom of students. Even if the professor provides carefully constructed individual comments throughout the paper, students are often reported to simply turn to the back page to locate the grade. Why does it appear that instructor comments have so little value in comparison to grades?

In spite of the obvious, that not all students desire feedback from their teachers, the myth of the eager learner breathlessly awaiting wise feedback from professors endures. Our beliefs have shaped the professional literature. Two strands dominate the research on feedback in college settings. The first strand is research on how professors can improve the delivery of feedback to students. These efforts focus on processes for providing feedback in terms of timing, extent, delivery options etc. The second strand is on how to manage the process of providing feedback so that the professor's life is not overtaken by the time-consuming practice of grading and commenting on student work.

There is good news to be shared: the research efforts about college-level feedback have produced clear evidence about effective practices in both delivering feedback and managing the process of producing feedback. One such synthesis is The American Association for Higher Education's (AAHE) *Principles of Good Practice for Assessing Student Learning* (1996). Among the best practices listed are four that focus directly on the process of providing feedback to students.

Answers questions that people really care about.
Leads directly to improvement in learning.
Embedded in the context of learning.
Takes place repeatedly over time.

AAHE's list of good practice is but only one of a number of typologies and helpful guides on the process of providing feedback to students (Black and William, 1998; Nicol and Macfarlane-Dick, 2004).

The abundance of information about best practices in delivering feedback to students shows that the focus of higher education has been on the 'supply side' of the feedback equation. Educators know a great deal about effective feedback strategies (Pascarella and Terenzini, 1991). Clearly there are many challenges remaining in the

application of our knowledge to actual practice, but the largest gap in perfecting feedback practices is not the need for more information about how to deliver feedback but rather on how to increase student desire for feedback. To that end I propose that feedback be considered using an economic model of supply and demand.

What if feedback were a commodity which was bought and sold in the free market? How would value of feedback be determined? The answer would likely follow the basic rule of any free economy - price is a function of both supply and demand. When the supply of a commodity is greater than demand, the commodity carries a low market value. When demand is greater than supply, the commodity carries a high value. A stable economy occurs when supply and demand are about equal. If we apply that reasoning to feedback in higher education, wouldn't it be fair to report that higher educators provide more feedback to students than they ask for? Perhaps feedback is so 'cheap' in higher education because the supply side of the equation is out of balance with student demand for feedback.

Should we simply wait for students to discover the obvious high value of feedback from their professors? Can we have an impact on the demand side of the equation? Such efforts can, of course, be frivolous. Without the assistance of marketing experts and late-night television, many Americans would not know of their urgent desire for a Chia Pet®, a terracotta novelty pot for growing watercress sprouts, or a Ginsu Knife® which never needs sharpening. Other efforts to increase demand can be of critical importance to societies, such as the efforts to increase demand for AIDS testing in Africa or demand for information on the consequences of obesity in the UK and United States. The latter examples, efforts to increase desire and awareness because to do so benefit society, best illustrate the need for intentionally engaging students in wanting and valuing growth-producing feedback.

Educators might adopt two sets of lenses for viewing our feedback initiatives. First, we could think like anthropologists and study our students in their native habitats to establish their basic instincts and motivations. My past experiences suggest that such efforts would show that students respond positively to rituals and have a natural curiosity about how they fit in and compare with their peers. Second, we could think like advertising managers who create demand for products and services. Past experiences also suggest that even mature students often fail to see the connection between educational practices and the desired end results unless educators are very intentional in communicating both hopes and plans.

Below are a number of practical ideas for raising the value of feedback by increasing student demand for it. These ideas are a blend of building on the natural tendencies of students and purposefully creating adjustments in attitudes and behaviours.

1 Provide feedback that allows students to compare themselves with their peers in non-threatening ways. Traditional aged students are especially interested in learning how they fit in and measure up in comparison to their peers. Western education, until recently, has seldom provided opportunities for students to work together in the classroom in meaningful ways which include sharing feedback. Educators should use this natural curiosity to our advantage. Unfortunately, student interest in learning about themselves can easily be stifled by the scary aspect of being on the receiving end of poorly executed feedback.

This lesson was driven home to me when, long after I had graduated from college, I decided to take an introductory art class. I had undergraduate experience in art appreciation, but this was my first attempt at an art performance class. Interest in studying art was stimulated by marriage to an artist and not because of self-perceived artistic skills. My classmates had extensive prior experience in art classes as shown by their portfolio of work: I was admitted to the class as a special exception since I did not have a portfolio to present in the admissions process. I was keenly aware that my classmates, all far younger than me, had 'earned' their place in the class and so I was appropriately intimidated from the first day of class. Near the middle of the term students were required to show a piece of art to the class and open ourselves to a critique from the whole class. At that point in my life I had been a teacher for some 15 years and had presented speeches to audiences of over 2,000 people - being in front of a class was not difficult for me. Yet the position of seeking feedback from this class of students had my voice trembling and the nervous sweat stained my shirt and dripped off my forehead. Still, I recall this experience both for the emotions of the day and for the incredible feedback my peers gave me. Their feedback was powerful and positive, but the opportunity to compare my artwork with their creations allowed me to engage in self-reflection that would have otherwise been impossible.

While this example may fail to qualify as 'non-threatening', the value of the opportunity to compare my own work with that of my peers was a rewarding experience. Judging from the ease at which many of my peers undertook this critique I surmise that mature artists may find peer critique less threatening than I did in my first experience with the process.

2 Focus on formative feedback rather than summative feedback. An anthropologist studying college students would likely establish that grades and transcripts are the currency of higher education for many students. (Educators who hope that students will learn for the pure sake of learning may not like hearing this news!) Tightly connecting feedback to the grading process so that there is real opportunity to turn that feedback into a higher grade could raise the demand for feedback. The use of multiple drafts of assignments and other opportunities to improve student work before final grading is more helpful than feedback that comes only at the end of an assignment or activity. The use of comments to improve a first draft is intuitively obvious to mature learners but may not be so to less experienced learners. New students may need assistance in the process of using feedback from early drafts and will need to gain experience in seeing the positive impact of later grades.

Unfortunately most students have a less than confirming experience with assignments that use multiple drafts. Student experiences in many multiple-draft assignments follow a predictable path:

1 student receives feedback about errors in the first draft

2 student rewrites paper

3 professor points out unresolved errors from the first draft, new errors created in the edits, and finds errors that were in the first draft but were not noted.

While the chance of any future draft being error-free could seem hopeless, the more important lesson taught may be that corrections do not produce positive feedback.

Using the opportunity of a second draft to praise improvements could significantly increase student interest in receiving feedback.

3 Leverage the power of peers. Students learn a great deal from other students. Alexander Astin's review of the research literature found that 'the student's peer group is the single most potent source of influence on growth and development during the undergraduate years' (1993). Educators should formalise the peer feedback process to help students develop a habit of seeking feedback from others.

A recent study of the power of peer-based feedback confirms the veracity of this pedagogy. Less academically prepared students were paired with a stronger student as either a roommate or laboratory partner. Control groups consisted of student pairs matched on their level of academic preparation. In this set of experiments, weaker students matched with a stronger student performed at higher levels than students matched as equals. Furthermore, there was no evidence that learning outcomes were diminished for stronger students (Winston, 2003).

Intentional efforts by faculty members can increase opportunities for students to work together in ways that encourage both formal and informal feedback exchanges. Peer to peer feedback is too important to leave to serendipity.

4 Role model your own demand for and willingness to seek external feedback. Students need models of adults who actively seek feedback from others. They also need models of how adults evaluate feedback and decide how to use feedback for improvement. Sometimes the really smart move is to dismiss misguided or erroneous feedback. Understanding that effective people retain full power to use or dismiss feedback is an empowering epiphany. One way to increase demand is to teach students that asking for feedback does not mean that they give up power to make their own decisions and determine their own course of action.

5 Use rituals to establish high expectations. Fraternity and Sorority members often pass through admissions rituals which include feedback about their performance. Such rituals can create demand for feedback by providing an appropriate space and context.

One institutional example of an academic ritual used to shape expectations and behaviours is the opening day at Elon University in North Carolina. The Elon campus is known for its beautiful mature oak trees and so the campus initiated a ritual of giving each graduate an oak tree sapling on graduation day as a symbol that students take some of Elon with them wherever they go in the world. The ritual became an important graduation experience. After some years, someone realised that only the successful students received this symbolic reward. In effect, keeping this ritual a secret from new students missed the opportunity to motivate students and set high expectations. Following that realisation, the campus added to its welcoming activities a new campus convocation where new students crossed the 'graduation' stage on their first day on campus. To share the expectation that students will graduate from Elon, the campus presents each student with an acorn and explains that another acorn was planted that day to start growing the sapling that will be presented to graduating students in four years.

Elon uses this ritual to set high expectations that students will graduate and as the background for talking with students about what it means to be a successful student, including expectations for their involvement in receiving and giving feedback as part of the 'Elon Experience'. This ritual is but one example of how colleges can establish high expectations and increase the demand for feedback that could help students achieve success.

Conclusion

There will always be new staff members entering the profession and returning staff members working to enhance their level of skill in providing feedback to students. Exploration and dissemination of best feedback practice is unlikely to diminish anytime soon. In fact, the published literature shows a rich array of evidence about how best to supply feedback to students.

The challenge that continues to face higher education is to increase student demand for feedback as a key process in learning and psychosocial growth. To date, higher educators may have focused too exclusively on the supply side of the equation with too little attention to the demand side of the equation. The challenge for maximising the value of feedback as a pedagogical technique falls on institutions of higher education. It is necessary, but not sufficient, to supply feedback to students - we must also actively work to create student demand for feedback.

References

American Association for Higher Education (1996) *9 Principles of Good Practice for Assessing Student Learning*, AAHE, Washington DC www.aahe.org/assessment/principl.htm [Accessed 2 July 2004]

Black P and William D (1998) Assessment and classroom learning, *Assessment in Education*, 5 (1) 7-74

Nicol D and Macfarlane-Dick D (2004) *Rethinking formative assessment in higher education: a theoretical model and seven principles of good feedback practice*, conference presentation at Enhancing the Student Experience in Scottish Higher Education: Workshop 7, Glasgow, 4 June 2004

Pascarella E T and Terenzini P T (1991) *How College Affects Students*, Jossey-Bass, San Francisco

Winston G C (2003) $50,000 for Your Thoughts: Why Colleges Pay Wages to Their Students, *The Chronicle of Higher Education*, 11 November 2003

Formative assessment and student success

Professor Mantz Yorke, Centre for Higher Education Development, Liverpool John Moores University

Abstract

There are plenty of claims to the power of formative assessment, yet formative assessment is rarely used to fullest advantage in UK higher education. Causes of this weakness can be discerned in curricular structures and pedagogy, and their effect is magnified as the pressures increase on the sector to enhance student employability and to make 'efficiency gains'. The argument of this paper is that a realisation of the full potential of formative assessment will require substantial attention to both curricular structure and pedagogy.

Formative assessment

The basic idea of formative assessment is straightforward - to contribute to student learning through the provision of information about performance, either formally or informally. Rowntree (1987) sees it as spanning 'a spectrum [...] ranging from the very informal, almost casual, to the highly formal, perhaps even ritualistic'. This spectrum can be transmuted into a four by two matrix (Table 1).

	Formal	**Informal**
From teachers	(probably the main approach in higher education; feedback from computerised packages might be included here)	(where circumstances permit)
From peers	(eg via peer assessment activities)	(perhaps over coffee or a stronger beverage)
From others	(though possibly problematic if the 'other' is also a mentor or supervisor)	(probably the main approach in work-based learning contexts)
From self	? (only if an assessment requirement)	(where the student is acting self-critically)

Table 1 A typology of formative assessment

Most academics probably tend to think of formative assessment in terms of the top left-hand cell. Of course, formative assessment from teachers may be oral (particularly in fieldwork, studios and in practice placements) and possibly informal (top right-hand cell), and intended to encourage the student towards progressively higher levels of achievement. The key issue underlying this paper is how to encourage students to develop their capacity for self-assessment (bottom right-hand cell).

Summative assessment and a blurred distinction

Summative assessment, in contrast, is concerned with establishing the extent to which a student has achieved the outcomes specified in the curriculum design. However, the distinction between formative and summative assessment is far from sharp. Some assessments (eg in-course assignments) are deliberately designed to be simultaneously formative and summative - formative because the student is expected to learn from whatever feedback is provided, and summative because the grade awarded contributes to the overall grade at the end of the study unit. Summative assessments in relation to a curricular component (the student passes or fails a module, for example) can act formatively if the student learns from them: in the case of examinations it is rare for feedback to be provided upon performance.

Formative assessment is powerful

The value of formative assessment has been stressed by a number of authors (eg Brown and Knight, 1994; Hounsell, 2003). Both qualitative and quantitative

studies have provided convincing evidence of its effectiveness. The extensive meta-analysis of quantitative data from school and college settings[1] that was undertaken by Black and Wiliam (1998) provides as strong a confirmation of the effectiveness of formative assessment as is likely to be obtained from experimental studies in an arena in which control is always problematic.

If it is reasonable to extrapolate the findings that have emanated mainly from school settings across the breadth of higher education, then the need is to consider whether better use may be made of formative assessment in the latter milieu. This is not only a matter of pedagogy because the structures of academic curricula provide the framework (or, perhaps, the cage) within which pedagogy operates.

A weak aspect of curriculum provision

It is widely acknowledged that assessment is generally weak in comparison with other aspects of curriculum provision. Criticisms from a range of Quality Assurance Agency for Higher Education (QAA) reviews have included the slowness of feedback, and the failure of the feedback to offer adequate guidance for future work[2]. Boud (1995) coined the term 'consequential validity' to signify the importance of feedforward in formative assessment - indeed, the word 'formative' itself embodies this teleological perspective.

Even though students tend to be more generous when responding to survey questionnaires[3], it is evident that not all is well in the assessment domain.

Formative assessment is under threat

The pressures on higher education exert a continuous threat in respect of formative assessment. Among these are the following.

- Paradoxically, the governmental concern with standards of attainment and with the accountability of higher education, which privileges summative over formative assessment.
- Research and related activities, which draw attention away from teaching and learning.
- Increased student/staff ratios.
- Curricular unitisation, which has increased attention on summative assessment at the expense of formative assessment.
- The legacy of the 'scientific measurement' paradigm that was dominant in the twentieth century (Shepard, 2000), and still casts a shadow over the contemporary constructivist approaches to pedagogy.

[1] The bulk of the data had been gathered in school settings.

[2] See the range of subject overview reports published by QAA, the *Overview report on Foundation degree reviews: conducted in 2003* (QAA, 2003), *Learning from higher education in further education colleges in England* (QAA, 2004a) and *Learning from subject review 1993-2001* (QAA, 2004b).

[3] Data from a survey of Foundation Degree students suggests that they were happier with feedback than were members of QAA visiting parties. Further, the recent pilot survey of student opinion for a national survey in the UK seemed to evoke similarly generous responses (Wojtas, 2004). Of course, opinion ratings can only be properly interpreted in the light of the respondents' expectations.

These are big pressures whose effects can be mitigated to only a limited extent by the actions of individual teachers. The need is for a strategic approach on a broader front.

There are also threats from students' perceptions of the assessment regime, such as students playing the assessment game by finding out the expected response and providing it, rather than taking a risk with something more ambitious - a grade/performance trade-off that is often learned in school (Doyle, 1983).

Effective formative assessment

Formative assessment is perhaps more demanding than some teachers and students appreciate. Knight and Yorke (2003) summarise the demands in Exhibit 1. Although Exhibit 1 refers to the student's personal development, it probably underplays the importance of the personal in formative assessment - a matter that is discussed further below.

Teachers:

1 are aware of:

- the epistemology of the discipline
- stages of student intellectual and moral development
- the individual student's knowledge and stage of intellectual development
- the psychology of giving and receiving feedback

2 provide:

- tasks sufficient in number to create opportunities for giving feedback on all **key** module/programme learning outcomes
- tasks of progressively graded difficulty, appropriate to the students
- criteria against which performance(s) will be judged

3 communicate with students:

- clearly regarding the standards expected of students
- in a timely manner
- highlighting the strengths and weaknesses of presented work (and not of the students themselves)
- indicating how their work might subsequently develop.

Students:

- understand what is expected of them (with reference, inter alia, to the assessment criteria)
- elicit the meaning from formative comment
- act on the basis of their developed understandings.

Exhibit 1 Components of effective formative assessment[4].

[4] This exhibit draws upon Knight (2002), Gibbs and Simpson (2002) and Yorke (2003).

It is argued elsewhere (Knight and Yorke, 2003) that formative assessment can be considered as a complex signaling system that calls into play at various times the background characteristics of student and teacher/assessor; the assessment task; the criteria against which performance is being judged; the giving and receiving of feedback on performance; and - crucially for success - the taking of action by the student on the feedback received. The signaling system conveys various 'messages' which are subject to interpretations of varying accuracy. A successful signaling system minimises inaccuracy in interpretation[5]. When and how that signaling system is operated are important pedagogic matters that need to be addressed in the context of contemporary higher education.

Widened participation

The government's aim for English higher education is that, by 2010, half of young people[6] should have experience of higher education. In Scotland, this level of participation has already been reached; in England, the percentage of young people in higher education at the time of the 2003 White Paper was estimated to be 43 per cent (DfES, 2003, para 5.7). This contrasts with a ratio of roughly 1:8 at the beginning of the 1980s. However, the proportion of students entering higher education from socio-economic groups IIIm to V remains well below that from the other socio-economic groups, although it has risen steadily over the years (DfES, 2003). The demographics of higher education have changed in another (and not unrelated) way, in that many more 'mature students' have been recruited to higher education, with a wide range of background experiences. The student body in higher education is more diverse than it has ever been, yet, as Wagner (1995) observed nearly a decade ago, the now-massified system was still being run on lines similar to the previous elite system.

The performance indicators published by the funding councils (eg HEFCE 2003a) show that the new universities and general colleges of higher education attract a higher proportion of students from disadvantaged backgrounds than the old universities, and that their student completion rates are lower. Poor completion is linked to a variety of factors that are unrelated to institutional provision (eg poor choice of programme, financial difficulties, health problems), but also to some which appear to be quite strongly related (eg quality of the student experience; resourcing)[7]. In the latter group, dissatisfaction with tutorial support and feedback figure as influences on non-completion.

The vital importance of tutorial support was captured by a student who had entered higher education from an access course:

> 'I completed an Access course prior to attending [university] where the staff were really helpful and knew you on a 1 to 1 basis. At university this wasn't the case and…I couldn't cope with the workload with no tutorial support' (Student reading for a diploma in higher education. From Yorke, 1999).

[5]There is often an opportunity, in higher education, for students to interpret tasks in ways that were not envisaged by the task-setter - indeed, some would argue that higher education should offer an invitation to students to do more than fulfil specified tasks, for example, by 'reading around' the subject. Hence the level of interpretive inaccuracy in the signalling system may not reach the zero that could perhaps be achieved in competency-driven curricula.

[6]Taken as spanning the age-range 18 to 30.

[7]See for example Yorke (1999), Davies and Elias (2003), Yorke and Longden (2004).

Since a quarter of the 2,151 non-completing students who responded to a survey of their experiences of higher education indicated that dissatisfaction with staff support was a moderate or considerable influence in their departure (Yorke 1999), the quotation carries some weight[8]. However, it is impossible on the evidence available to make an explicit causal connection between poor feedback and non-completion.

The policy drive towards widened participation draws attention more sharply than before to the characteristics of the entering students, and hence to how their chances of success might be maximised. A central issue, well understood by those running access courses, is the need to encourage the development of self-confidence, self-esteem and self-efficacy. This brings into focus the student as a person - a perspective that was well appreciated by writers such as Rogers (1961; 1969), but is currently at some risk of being backgrounded by the instrumentalism of outcomes-led curricula.

The importance of 'the personal'

The USEM account of employability[9] provided by the Enhancing Student Employability Co-ordination Team (Knight and Yorke 2004; Yorke and Knight 2003) emphasises an aspect of student learning that has been given relatively little attention until recently - the range of personal qualities and attributes that can influence student (and graduate) achievement. The argument made in respect of USEM is that these personal variables influence performance across the board. There is a substantial body of research that testifies to the importance of the personal dimension (Table 2).

The point of including Table 2 is to demonstrate, albeit with considerable brevity, that 'the personal' is of significance in the learning process[10]. In a higher education system in which some students[11] will, for a variety of reasons, be uncertain of their capacity to succeed, formative assessment is of particular significance. The opportunity exists for staff to **support students psychologically**[12] in addition to demonstrating ways in which their academic performance might be developed. The need is to develop their self-efficacy in relation to higher education.

Contribution	Theorist(s)
Having a malleable self-theory is preferable to having a 'fixed' self-theory (eg 'my intelligence is "developable" rather than fixed for all time').	Dweck (1999)
Adopting goals focused on learning is generally preferable to concentrating on performance ('looking good' or 'not looking bad') …	Dweck (1999)
…though striving for good performance may not be deleterious, especially for the able	Pintrich (2000)
Practical intelligence plays an important part in success in life; academic intelligence is not sufficient	Sternberg (1997)

[8] Even though the particularities of individual students' circumstances were varied.

[9] **U**nderstanding; **S**kilful practices; **E**fficacy beliefs and **M**etacognition. Although USEM was developed with employability in mind, it applies to other life-situations as well. It is consistent with the concept of 'capability' advocated by Stephenson (1998), which can be summarised as 'effectiveness in the world'.

[10] Note that Marzano (1998) found size effects comparable to those of Black and Wiliam (1998) in respect of experimental studies of influence on the self-system (the 'E' of USEM) and on metacognition (the 'M'). As with Black and Wiliam's meta-analysis, the bulk of the studies were undertaken in school contexts.

[11] And not only those who might be labelled as 'widening participation students'.

[12] This should not be interpreted in terms of unwarranted praise (Dweck, 1999).

Contribution	Theorist(s)
An internal 'locus of control' is preferable, since it relates to a sense of self-agency	Rotter (1966)
Self-efficacy (the belief that one can, probabilistically 'make a difference') influences performance	Bandura (1997)
'Learned optimism', likewise	Seligman (1998)
'Emotional intelligence', also likewise	Salovey and Mayer (1990); Goleman (1996)
Emotional state is of significance in learning	Boekaerts (2003)

Table 2 Sources of support for valuing 'the personal'

Pedagogy

Coming to terms with the demand of higher education is a challenge for any first-time student, irrespective of their background. Apparently privileged students (such as those going to public schools) who have reasonably good A level scores may struggle when they are confronted with the need to take considerable initiative in their learning. There is circumstantial evidence to this effect from Naylor and Smith (2002) and HEFCE (2003b) (as regards degree classification) and from unpublished work by the Student Assessment and Classification Working Group on data relating to year one performances at a new university. In crude terms, the privileged did less well than might have been expected. One can only speculate as to the cause(s), but a plausible contributory explanation would seem to be the amount of coaching given in respect of the A level examinations which is typically not available in the higher education environment.

The standards of higher education are difficult to appreciate. Incoming students have to adjust to the norms and expectations that may have a substantial tacit dimension, and 'tuning in' can take time. The good performer at A level may not easily realise that - to parody slightly - an elegant reassembly of received opinion may attract only a modest grade in a system that looks for analytical challenge and imagination in awarding the highest grades[13]. The student entering with vocational qualifications and/or with life-experience may also find the demand initially mystifying.

The main pedagogical implication is clear. In the early stages of a programme it is important to provide early study tasks and to give feedback on them. To do so allows a dialogue to open up between teacher and student regarding expectations which, crucially, is based on real activities and not the abstractions of learning outcomes[14]. Note the plural, 'tasks'. Once is almost certainly not enough. There needs to be a sequence of tasks and associated feedback, if the student is to be helped to appreciate exactly what is expected. Bandura advises against demanding large

[13] There are probably some marked differences between subject disciplines here. The early stages of science-based programmes may focus on the development of basic understandings at the expense of the criticality expected from the start in programmes in the humanities.

[14] See Wolf (1995) on the importance of exemplification in assessment to teachers, let alone students.

cognitive jumps, on the grounds that it can be demoralising not to make good progress towards a distant goal:

> The less individuals believe in themselves, the more they need explicit, proximal, and frequent feedback of progress that provides repeated affirmations of their growing capabilities (Bandura, 1997).

One approach that responds to Bandura's point is that of the 'patchwork text' (Winter et al, 2003) which involves students in completing a series of short assignments that culminate in a reflective commentary on what has been learned, instead of requiring them to complete a single assessment at the end of the module of study.

As the student's capabilities develop, the expected cognitive jumps can be larger.

There is, however, a possible problem with formative assessment. If feedback is given on draft material, the student may respond to the feedback without necessarily having developed their understanding to an appropriate extent. The revised piece of work may gain a good grade, but it may in the longer term prove an inadequate launching pad for the next, more advanced, phase of study. Subsequent performances may turn out to be weaker than the student had come to expect.

An underlying concern in formative assessment is the development of the student's autonomy and their capacity for self-regulation. Formative assessment is aimed at the student's metacognitive development, as well as the cognitive and affective (and in appropriate cases the psychomotor). In terms of student development, the ideal shift is from formal feedback provided by teachers to students' evaluation of their own achievements (part of the development of autonomy in learning).

Knight and Yorke (2003) argue that some aspects of performance are not warrantable by institutions unless the resourcing devoted to them reaches a prohibitive level. Performances in work settings are vulnerable to the lack of training of assessors in the workplace, and the kinds of role conflict that follow from a colleague acting as both mentor and assessor. A more viable option is to make better use of formative assessment - to enable the student or graduate to make claims regarding their achievements that are grounded in evidence collated in a portfolio. The success of such an option depends on the student's capacity to reflect upon their performances and on the formative feedback that they have received. While formative assessment is typically construed as being 'low stakes' in character, in the longer term it implicitly acquires a veneer of 'high stakes' (such as when a student or graduate draws upon it in the construction of a job application, or at interview).

It is sometimes overlooked that formative assessment is part of the teaching/learning engagement. An implication of seeking to improve formative assessment is that more staff time will need to be devoted to it, and less to other aspects of teaching. This is more challenging than a mere redistribution of time, since it raises the issue of pedagogical strategy within the unit of study.

Increased student/staff ratios as UK higher education has expanded have led to growth in class size and to a diminution of contact between teachers and students. Yet the pedagogic changes have tended to adjust incrementally to accommodate to the growing pressures: the opportunity has not generally been taken to reflect on whether incremental adjustments are the optimal response to a massified system

whose resourcing is accepted by government as being currently inadequate. Using Senge's (1992) metaphor, the higher education frog continues to swim in the bucket of gradually heating water, since no single change in circumstances has been sharp enough to trigger a move to jump out.

How the formative assessment is handled is important. There is a danger that some students will interpret criticism of their work as being criticism of them as persons. Table 2 indicates the importance of dealing sensitively with the psychological aspects of formative assessment. In their longitudinal study of learning at Alverno College in the US, Mentkowski and Associates (2000) exemplify good practice in the College's approach to formative assessment, which has requires students to construe formative assessment in terms of constructive criticism and its capacity to help them evaluate their own work[15]. What Alverno College has achieved is an institutional culture focusing on **learning**, without denying the importance of performance.

Academic structures

Dealing with pedagogy at a relatively localised level will not solve all the problems relating to the improvement of formative assessment, since the academic structures within which pedagogy is conducted exert their own constraints. It is widely appreciated that the move towards modularised curricula has made the modal time-span for a curricular component the semester, in contrast with the academic year. This has militated against formative assessment, since there is considerable pressure to cover the syllabus in the time available in the semester. Much formative assessment takes significant time, and the limited time available in a semester makes it difficult to accommodate the 'turn-around time' needed for students to submit work, have it assessed, and to act upon whatever recommendations may be made by the assessor. Staff in a number of institutions have complained that summative assessments now occur more frequently than they did in the past.

Recently a number of institutions have relaxed their modular structures in order to lessen the pressure on students in the first semester. This is not a version of 'dumbing down', but the opposite - a way of helping students towards the attainment of standards.

The espoused logic can be summarised as follows (Yorke, 2001).

- Students merely need to pass the first year of full-time study in order to qualify for the honours-bearing phase of a degree. Grades higher than that of a pass are, in effect, irrelevant.
- Students who take a little time to adjust to the demands of higher education may fail their assessments at the end of the first semester.
- This may be sufficient of a discouragement for them to decide to discontinue and, if they do continue, they are likely to be burdened by 'trailing' one or more modules into the second semester.
- Since such students are, self-evidently, not the strongest at this stage, it is better to act formatively during and at the end of the first semester, and to place the only critically important (summative) assessment at the end of the second semester.

[15] Although they do not call upon some relevant psychological theorising.

The 'logic-in-use' however, has probably factored in the institutional concern with retention statistics, and it is debatable whether the educational or managerial consideration has been the dominant driving force.

The need for a strategic approach

Institutions are in a position to take strategic initiatives to enhance the provision of formative assessment. They can, for example, review their curricular structures in order to test whether they are facilitating to the optimum the development of their students. Since institutions in the UK are required to produce learning and teaching strategies and are expected to update them at intervals, there are external stimuli to undertake curricular review.

Two studies suggest that institutions in England which have addressed the first-year experience systematically are reaping the benefit in terms of student success (Action on Access 2003). In the first study (of six institutions that had bettered their benchmark expectations for retention despite having a challenging demographic profile), interviews with senior managers found inter alia that these institutions were likely to have emphasised formative assessment in the early phase of their programmes, and were committed to being supportive to students and 'friendly'. In one of the institutions, there was also a recognition that the social dimension was important in learning activities (echoing comments above regarding 'the personal'). A study of a further nine institutions with high proportions of students from disadvantaged backgrounds elicited findings that were broadly similar, even though the retention statistics had yet to reflect the policy changes that had been put into effect.

The Action on Access studies concentrated their attention on the first-year experience, in which induction (both socially and academically) are important, with formative assessment playing a key role. However, formative assessment is significant across the whole curriculum, and it would be a mistake to concentrate excessively on the first-year experience, critically important though it is. Students can benefit from a curriculum richly pervaded with formative assessment, though the nature of these benefits is likely to evolve over a student's period of engagement in higher education.

Where the institution finds it difficult to deal with formative assessment at a corporate level, there is nevertheless plenty of scope for appropriate action at the level of the department and/or programme team. Programmes are reviewed at intervals, often quinquennially, which provides the opportunity to reconsider curricular aims, content and - of particular importance here - methods and assessment. Less obviously strategic, the department or programme team can instigate a process of 'tuning' curricula (Knight and Yorke, 2004) by making appropriate adjustments to existing practices which remain consistent with the approved curricular structure. Placing oneself in the shoes of the student, and being imaginative as regards pedagogy, can lead to quite high gains for relatively little pedagogical pain.

Strategic activity requires vision and leadership, together with the managerial skills to ensure that the espousal of change does not get transmuted into actionless rhetoric. As Fullan (2001) observes, educational change is relatively easy to envisage, but difficult to implement in a socially complex environment. It is a demanding challenge to give formative assessment greater prominence in curricula, since it is likely to involve considerable cultural change as regards pedagogy. The challenge can be expected to take considerable time, and is unlikely to succeed without sustained leadership and commitment. There is no quick fix.

References

Action on Access (2003) *Student success in higher education*, Action on Access, Bradford

Bandura A (1997) *Self-efficacy: the exercise of control,* Freeman, New York

Black P and Wiliam D (1998) Assessment and classroom learning, *Assessment in Education*, 5 (1) pp7-74

Boekaerts M (2003) Towards a model that integrates motivation, affect and learning, in L Smith, C Rogers and P Tomlinson (eds) *Development and motivation: joint perspectives,* pp173-189, British Psychological Society, Leicester

Boud D (1995) Assessment and learning: contradictory or complementary? in P Knight (ed) *Assessment for learning in higher education,* pp35-48, Kogan Page, London

Davies R and Elias P (2003) *Dropping out: a study of early leavers from higher education* [*Research Report RR386*], Department for Education and Skills, London

DfES (2003) *The Future of Higher Education* (Cm 5753) The Stationery Office, Norwich

Doyle W (1983) Academic work, *Review of Educational Research,* 53 (2) pp159-199

Dweck C S (1999) *Self-theories: their role in motivation, personality and development*, Psychology Press, Philadelphia

Fullan M (2001) *The new meaning of educational change*, third edition, Teachers College Press, New York, and RoutledgeFalmer, London

Gibbs G and Simpson C (2002) *Does your assessment support student learning?* unpublished paper, Centre for Higher Education Practice, The Open University, Milton Keynes

Goleman D (1996) *Emotional intelligence*, Bloomsbury, London

HEFCE (2003a) *Performance indicators in higher education in the UK, 2000-01, 2001-02,* Higher Education Funding Council for England, Bristol www.hefce.ac.uk/learning/perfind/2003/reports/ [accessed 4 February 2004]

HEFCE (2003b) *Schooling effects on higher education achievement* [Report 03/32], Higher Education Funding Council for England, Bristol

Hounsell D (2003) Student feedback, learning and development, in M Slowey and D Watson (eds) *Higher education and the lifecourse,* pp67-78, Society for Research into Higher Education/Open University Press, Maidenhead

Knight P T (2002) *Being a teacher in higher education,* Society for Research into Higher Education/Open University Press, Buckingham

Knight P T and Yorke M (2003) *Assessment, learning and employability*, Society for Research into Higher Education/Open University Press, Maidenhead

Knight P T and Yorke M (2004) *Learning, curriculum and employability in higher education*, RoutledgeFalmer, London

Marzano R J (1998) *A theory-based meta-analysis of research on instruction,* Mid-continent Regional Educational Laboratory, Aurora

Mentkowski M and Associates (2000) *Learning that Lasts: Integrating learning development and performance in college and beyond*, Jossey-Bass, San Francisco

Naylor R and Smith J (2002) *Schooling effects on subsequent university performance: evidence for the UK university population [Warwick Economic Research Papers No 657*], Department of Economics, Warwick University, Coventry www2.warwick.ac.uk/fac/soc/economics/research/papers/twerp657.pdf [accessed 23 November 2003]

Pintrich P R (2000) The role of goal orientation in self-regulated learning, in M Boekaerts, P Pintrich and M Zeidner (eds) *Handbook of self-regulation*, pp451-502, Academic Press, New York

QAA (2001a) *Subject Overview Report: Business and Management* [QO5/2001] Quality Assurance Agency for Higher Education, Gloucester

QAA (2001b) *Subject Overview Report: Education* [QO10/2001] Quality Assurance Agency for Higher Education, Gloucester

QAA (2003) *Overview report on Foundation degree reviews conducted in 2003,* Quality Assurance Agency for Higher Education (mimeo) Gloucester

QAA (2004a) *Learning from higher education in further education colleges in England,* Quality Assurance Agency for Higher Education, Gloucester www.qaa.ac.uk/reviews/academicReview/learningfromHEinFE/HeFe.asp

QAA (2004b) *Learning from subject review 1993-2001,* Quality Assurance Agency for Higher Education, Gloucester www.qaa.ac.uk/reviews/SubjectReview/learningfromSubjectReview/subjectreview.asp

Rogers C R (1961) *On becoming a person: a therapist's view of psychotherapy,* Constable, London

Rogers C R (1969) *Freedom to learn,* Charles E Merrill, Columbus

Rotter J B (1966) Generalized expectancies for internal versus external control of reinforcement, *Psychological Monographs,* 80, pp1-28

Rowntree D (1987) *Assessing Students: How Shall We Know Them?* (revised edition) Kogan Page, London

Salovey P and Mayer J D (1990) Emotional intelligence, *Imagination, Cognition, and Personality, 9,* pp185-211

Seligman M (1998) *Learned Optimism*, Pocket Books, New York

Senge P (1992) *The Fifth Discipline: The Art & Practice of the Learning Organisation,* Century Business, London

Shepard L A (2000) The role of assessment in a learning culture, *Educational Researcher,* 29 (7) pp4-14

Stephenson J (1998) The concept of capability and its importance in higher education, in J Stephenson and M Yorke (eds) *Capability and quality in higher education*, pp1-13, Kogan Page, London

Sternberg R J (1997) *Successful intelligence: how practical and creative intelligence determine success in life*, Plume, New York

Wagner L (1995) A thirty-year perspective: from the sixties to the nineties. In T Schuller (ed) *The changing university?* pp15-24, Society for Research in Higher Education/Open University Press, Buckingham

Winter R, Parker J and Ovens P (2003) The patchwork text: a radical re-assessment of coursework assignments, *Innovations in Education and Teaching International,* 40 (2) whole of special issue

Wojtas O (2004) Students rate staff highly, *The Times Higher Education Supplement,* No 1625 (30 January) p11

Wolf A (1995) *Competence-based assessment*, Open University Press, Buckingham

Yorke M (1999) *Leaving early: undergraduate non-completion in higher education,* Falmer, London

Yorke M (2001) Turn first-semester assessments into richer learning experiences, *Innovations in Education and Teaching International,* 38 (3) pp277-278

Yorke M (2003) Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice, *Higher Education,* 43 (4) pp477-501

Yorke M and Knight P (2003) *Embedding employability into the curriculum*, Learning and Teaching Support Network, York

Yorke M and Longden B (2004) *Retention and student success in higher education*, Society for Research into Higher Education/Open University Press, Maidenhead

Improving feedback to students (link between formative and summative assessment) - Post-workshop report

Dr Robert Matthew[1], Teaching and Learning Service, University of Glasgow and Workshop Director

A one-day workshop 'Improving feedback to students' was held on 4 June 2004 at the St Andrew's Building, University of Glasgow. The workshop consisted of four keynote presentations and two breakout sessions. The keynote presentations were given by Professor Dai Hounsell, University of Edinburgh; Professor Mantz Yorke, Liverpool John Moores University; Dr Randy Swing, Policy Center on the First Year of College, Brevard College, North Carolina; and Mrs Debra Macfarlane-Dick, University of Glasgow and Dr David Nicol, University of Stratchclyde.

The keynote presentations

Dr Randy Swing gave an interesting and thought provoking talk on 'Understanding the economies of feedback: Balancing supply and demand'. He started his talk with an observation on student behaviour when receiving feedback. He then moved on to review the research literature (albeit from a strong North American perspective) with a view to developing a model for giving learner feedback which is firstly likely to enhance how it is received and secondly that the effort in producing it is somehow balanced against the demand for the feedback from the students. This application of the law of supply and demand provoked discussion in a number of the breakout groups.

Professor Dai Hounsell gave a talk entitled 'Reinventing feedback for the contemporary Scottish university'. In his talk, Professor Hounsell discussed nine assumptions which could be said to underpin current practice on feedback at the present time. For each of the assumptions, he offered grounds for the questioning the assumption and then offered advice for reinventing feedback practice. Throughout the presentation, the discussion was reinforced by what the research literature had to say on the subject. Overall, this was a paper which gave a pragmatic but research-informed way forward on the subject of feedback.

Professor Mantz Yorke presented a paper entitled 'Formative assessment and student success'. In his presentation Professor Yorke began by reviewing the blurred boundary between formative and summative assessment, then moved on to briefly review the UK experience of the Quality Assurance Agency for Higher Education reviews of 'teaching quality' and the emerging general conclusion that feedback to students is an area where legitimate criticism can be made. Particular areas of concern were the length of time for the return of feedback to students in some instances and the failure of feedback to offer adequate guidance for future work. In the next part of his talk, Professor Yorke examined what makes effective formative assessment and feedback. He used the USEM model[2] to stress the importance of the 'personal' and 'personal development' in making feedback effective. This, to me, is an interesting approach to getting students to take ownership for responding to feedback and also for developing the skills associated with self and peer assessment. Finally, Professor Yorke stressed the need for institutions to adopt a **strategic** approach to enhance the provision of formative assessment and feedback. This is certainly a viewpoint that would seem in keeping with having assessment as a specific enhancement focus in Scotland.

[1] Unable to attend on the day due to bereavement.
[2] **U**nderstanding, **S**kilful practices, **E**fficiency beliefs and **M**etacognition.

In their joint presentation, Debra Macfarlane-Dick and Dr David Nicol outlined the model they developed as part of the Learning and Teaching Support Network Generic Centre-funded project Student Enhanced Learning through Effective Feedback (now part of the Higher Education Academy). They presented a conceptual model to describe formative assessment and feedback. The model is a synthesis of the work of many researchers and is primarily based on a model of self-regulated learning and feedback developed by Butler and Binnie. Like almost all such models this one provoked debate among participants about its description of such a complex process. The presenters then went on to describe seven principles of good feedback practice linked to the model, which helped to demonstrate its usefulness. The obvious benefits of these principles were reinforced by means of several case studies, which helped to reinforce the concept that feedback is intended to help enhance student learning

Group discussions

Participants were divided into breakout groups at two points during the day. The discussions were wide-ranging and what follows is my own distillation of the record of these discussions provided by the note-takers appointed on the day for which I am extremely grateful.

From my reading of the discussions, it is clear that the breakout groups covered a wide range of issues. I have identified a number of recurring of themes and these are presented as follows (in no order of importance).

Demand

The demand for feedback is a factor which requires consideration in the development of any institutional strategy on assessment. The participants gave many examples to show the differences that disciplines make on formative feedback eg in the visual arts, students are very demanding and there are tensions between giving formative feedback, ensuring students become independent learners and staff time. It was also noted that even within a single discipline the demand for feedback can change during the named award eg in nursing the demand for formative feedback grows during the programme of study as students seek more feedback on their professional practice.

Time

The issue of time and timing was commented on by many of those attending. The staff time involved in writing high quality feedback, and the methods used to reduce this (eg checklists, pro formas or computer generated feedback) were discussed at length. The timing of feedback to learners was also discussed and the seven principles described by Macfarlane-Dick and Nicol were thought by some to offer helpful advice on the timing of the feedback ie not all feedback needs to be returned as quickly as possible.

Anonymity

The role that the increased used of anonymous marking plays in the feedback process was discussed by some groups alongside the issue of giving feedback in a way that helps with both low and high achievers without demotivating either. While

anonymous marking has been introduced with the best interest of the students in mind, it was felt by many that it was often difficult to write impersonal (to an anonymous student) and yet appropriate feedback.

Student numbers

Clearly the issue of formative feedback to large classes is one which vexes many academic colleagues. Almost everybody thought that giving 'helpful' feedback was easier in smaller classes. However, ideas of peer and self assessment, the use of checklists and pro forma and computer-based assessment and feedback were all discussed as possible ways of ameliorating the problem of large classes. Similar issues were discussed under the next heading.

Student use of feedback

The comment that was made by Randy Swing in his presentation, namely 'the professor provides carefully constructed individual comments throughout the paper yet students are often reported to simply turn to the back page to locate the grade' rung true for many. As a result, a considerable amount discussion took place around how to get students to engage with feedback. It was felt by many that self and peer assessment had a key role to play here, both in developing the skills of assessment of learners but also in developing notions of what is effective feedback. It was thought by some attendees that self and peer-assessment was also a beneficial way of generating useful formative feedback for learners in large classes.

Self and peer-assessment

While this arose indirectly in several groups, it was also explicitly discussed by other groups. It was thought by many that using self and peer-assessment, in both formative and summative situations, offered possible solutions to many of the difficulties discussed such as large classes, raising student expectations, motivating students to demand feedback and the development of learner independence and autonomy. Examples of methods used for self and peer-assessment included the use of model answers, pro formas and checklists.

Final thoughts

Perhaps unsurprisingly, at the end of the day, there were still many issues unresolved. Some of these included:

- training of students for self and peer-assessment
- training of students to use feedback effectively
- development of student abilities to recognise good and bad work
- staff time to develop new feedback methods
- getting the balance between formative and summative assessment.

Finally, feedback received from those who attended on the day suggests that this was both a stimulating and enjoyable day, which at the height of the exam season implies a successful day.

Personal note

I would like to acknowledge publicly the help and support I had from Fiona MacKinlay in the Research Office, Faculty of Education, University of Glasgow in organising this event. A great deal of this took place in my absence (due to the sudden death of my wife, Anne). Without the hard work of Fiona (particularly in dealing with the proposal from Scottish Water to cut the water supply to the building on the day of event!), this event would not have taken place. As they say, weel done cutty sark.

References

Knight P and Yorke M (2004) *Learning, curriculum & employability in higher education,* RoutledgeFalmer, London

Juwah C, Macfarlane-Dick D, Matthew B, Nicol D, Ross D and Smith B (2004) *Enhancing Student Learning through Effective Formative Feedback,* Higher Education Academy, York

Butler D L and Winnie P H (1995) Feedback and self-regulated learning: a theoretical synthesis, *Review of Education Research*, 65 (3) pp245-281

Assessment workshop series - No 8

Assessing personal transferable skills

Contents **Page**

Assessing personal transferable skills - An overview

Dr Colin Mason, SALTIRE, University of St Andrews and Workshop Director

Abstract

This short paper provides an introduction to the last of the enhancement themes workshops on assessment, Assessing personal transferable skills. It seeks to draw together the challenges of those approaches outlined in keynote and case study presentations as ways, or potential ways, of addressing the problematic nature of some of the issues for academic and support staff and students as well as institutions. This topic is underpinned by discussions that have arisen in many of the preceding workshops on assessment and develops some of the themes in practical ways. Thus, topics included here are:

- the nature and definition of 'skills'
- what is meant by personal learning and how or whether such learning is transferable
- whether learning of this sort is integrated within the curriculum (often implicit) and thus is 'acquired' through processes of engaging with teaching of disciplinary content
- whether such learning is explicitly identified, perhaps involving curriculum mapping to guide students throughout their discipline-based studies and sometimes compartmentalised (a stand alone module or learning opportunity)
- how and whether learning of this sort is assessed formally (often for credit) or informally (often as part of development)
- how best to represent the totality of a student's learning experience when this includes knowledge and understanding as well as skills that may have been assessed for credit with grades, or on a pass/fail basis, or that have simply been developed by opportunities both within and outwith the taught curriculum
- what might the role be of personal development planning (PDP) as a contributor to the learning processes involved and as a means of differently (better?) representing such learning
- and finally, how such learning is part of a wider set of attributes that contributes to the development of student 'employability' in the context of career development and planning.

The set of topics described above is addressed by different presenters in different ways, some focusing upon specific topics only, others bringing together two or more aspects, while others address holistically the whole set from a particular perspective. It is not my purpose here to provide a structured framework drawing out key messages from each presenter's contribution in the sequence of the programme. Rather, I intend to draw upon the research literature (particularly the Learning and Teaching Support Network Generic Centre commissioned series of briefing papers on Learning

and Employability and Assessment - now part of the Higher Education Academy) as well as presenters' contributions to pose a series of key questions that provoke further exploration of the theme for this workshop, especially during the breakout session discussions.

How do students develop personal transferable skills?

'Some (higher education) teachers were recently asked for their views on the teaching of personal transferable skills. Most thought that the teaching of personal transferable skills occurred incidentally when the students were encouraged to…There was conviction that personal transferable skills were developed as a means to an end. There was little evidence of actual instruction in personal transferable skills. Comments such as 'I am in a complete state of uncertainty about the problem of personal transferable skills. Please help if you can'. Such a cri de coeur points to the dilemma in which teachers have been placed in relation to the teaching of personal transferable skills.'

Now this paragraph is actually an amended version of the opening section of a classical work on spelling[1] (see original paragraph reproduced in Appendix I), but it highlights a dilemma facing higher education today. Are personal transferable skills caught or taught? The summary conclusion about whether spelling is caught or taught derives from the research literature available at that time, that a wide variety of training techniques should be introduced to supplement existing teaching approaches as well as acquired skill that arises through reading, not least a change of attitude by teachers to become more positive about spelling and thus pass on to their students a more positive attitude toward their own (good) spelling abilities. Clearly, there is a message for higher education and personal transferable skills - as for spelling, they are both caught and taught. This begs the questions, what are personal transferable skills and if they can be taught as well as caught, how can we assess whether they have been acquired or developed, and how is such learning then represented as student achievement?

What is the place of personal transferable skills in the higher education curriculum?

So, what are personal transferable skills? There is considerable confusion in the higher education literature about the terminology used to describe 'skills'. Yorke and Knight[2] have constructed a conceptual model, USEM, to better define employability (in higher education). Thus, employability is influenced by three other components (**U**nderstanding, **E**fficacy beliefs and **M**etacognition), in addition to **S**kills. The phrase 'skilful practice' has been introduced as being a more appropriate working definition of the retained term, skills, which includes key skills. Other terms used to qualify different subsets of skills are core skills, key skills, generic skills. There may be some artificial national boundaries surrounding nomenclature eg core skills, a term often used in Scotland, is perhaps interchangeable with key skills, as used in England. What constitutes particular skills in either of these definitions also varies however. After Dearing (1997), key skills included communication, information technology skills, numeracy and learning to learn skills. Nevertheless, many other skills lists have been drawn up and an example compendium of these, and which institutions are doing

[1] Peters M L (1967) *Spelling: Caught or Taught?* Routledge and Kegan Paul Ltd, London.
[2] Yorke M and Knight P T (2004) *Embedding Employability in the Curriculum*, LTSN series, Learning and Employability No 3.

what to develop them in their students, can be found in the project, **TRAN**sferable **S**kills in **EN**gineering and their **D**issemination, **TRANSEND**[3]. Generic skills are sometimes referred to as the transferable skills, ie those skills that students may require regardless of their discipline or employment context. These may include a variety of those aspects of employability defined under three headings: personal qualities, core skills, process skills as summarised in Learning and Employability No 3. In short, a definitive, agreed set of skills, including personal transferable skills, that students in higher education should be fortunate enough to be born with or expected to acquire, simply does not exist. This is not only true but also desirable. Different disciplines place emphasis on different aspects of their curricula and thus core or key skills for a geographer differ (albeit sometimes only slightly) from a chemist or maybe more so from an historian, though generic skills should be more common. We will see examples of the range of transferable skills developed by undergraduate students in chemistry at the University of St Andrews presented by Professor David Cole-Hamilton and other skills sets that are important for career planning identified by contract research staff in the presentation of Dr Colin Mason.

This may highlight another source of confusion - the use of the term 'academic skills'. Are these universal, or are these also discipline-related? I do not, anymore than others in this field, have clear answers to these questions. In one sense, I also wonder whether it really matters, but nevertheless present a classification view of higher education knowledge for debate (Appendix II). This draws on sources already mentioned but also includes views of education based upon work in the 1950s and 1960s, on learning objectives.

Finally, a further view of what constitutes effective learning in the higher education curriculum, incorporating skills, classified as either disciplinary or trans-disciplinary, is presented. This model is part of the ongoing work of a group in Scotland defining an effective learning framework as a tool to help institutions design curricula that facilitate personal and professional development planning, integrating academic, personal and career development opportunities both outwith and within the curriculum (Appendix III). Perhaps the beauty of this model is that the trans-disciplinary component does not necessarily convey the meaning that the skills involved are transferable, a moot point raised by Mantz Yorke[4]. One argument is that many of the skills such as 'arguing' in an academic context, are appropriate, but are context-dependent (unlike generic practical skills such as using email or a spreadsheet) since deployment of similar skills in employment might be problematic for a new graduate employee. Thus, direct transfer may be inappropriate, but knowing when and how to use or adapt particular skills in new situations is a higher order skill, characteristic of professional behaviour and part of the set of skills of self-awareness, self-regulation or metacognition, what Schon[5] might refer to as the 'reflective practitioner'. Peter Knight will discuss personal transferable skills in the wider context of employability in higher education in his presentation.

[3] Contact: Andrew J Merchant, Stefaan J R Simons, Ming Tham and David B F Faraday (2001) The TRANSEND Project, an FDTL project, carried out by a consortium of institutions: University of Birmingham, University College London, University of Newcastle-upon-Tyne and University of Surrey.

[4] Yorke M (2004) Employability in higher education: what it is - what it is not, LTSN series, Learning and Employability No 1.

[5] Schon D A (1983) The Reflective Practitioner - How professionals think in action, Arena, reprinted 1996.

How (or even) should transferable skills be taught?

The TRANSEND project neatly captures the nature of the debate - offering two workshops for the dissemination of good practice focusing on:

- embedded and integrated skills development
- bolt-on skills development.

There is heated debate about whether skills per se should be taught explicitly or caught, ie developed by providing learning activities within curricular designs that leave students no alternative but to engage in more than a mere cerebral way with disciplinary knowledge. Some (traditionalists often) argue that students come to university to study a discipline, English say, and that it is not even a consideration. The argument here is of a different order. Should stand-alone activities where skills are developed be provided within the curriculum, and if so how? Study skills, or learning to learn academic skills, are clearly important to enable students to gain maximum benefit from exposure to new information, ideas and conceptual frameworks presented by leading scholars in their disciplines, but rarely (until maybe their final year of studies!) do the majority of students appreciate taught courses in academic reading, speed reading, note-making, essay writing, examination technique etc that are provided either by central student services or even by academic departments. Such courses are perceived as remedial and even when they carry academic credit are unpopular. An alternate approach is to provide learning to learn opportunities within the curriculum. These require much more attention by academic departments to providing formative learning experiences - essays, laboratory reports, projects etc that are carried out by students and handed in for assessment by tutors who then provide constructive feedback that acknowledges and compliments what is good, corrects what is wrong in a sympathetic manner and points to areas for improvement and how this might be achieved. There are no marks at stake here, the student is engaged in the learning activity of the discipline but through feedback is encouraged to improve a particular academic skill (in this case, writing). Similar opportunities arise through collaborative group work for the development of inter-personal skills such as leadership, negotiation and giving feedback; oral skills through giving verbal presentations and so on. These are all examples of what Peter Knight[6] refers to as low risk, low stakes assessment (Feedback: formative assessment) compared to high risk, high stakes (Feedout: summative assessment). Such activities provide opportunities for guided practice, a pre-requisite for skills development.

So, why isn't there more of this going on? I suspect two forces are conspiring together to limit such opportunities. On the one hand, while students recognise the benefits of formative assessment and wish to have more feedback on their work, they are also aware of the need for strategic effort directed at those assessments that 'count' - summative assessments. Consequently, students do not readily engage in non-summatively assessed learning tasks - they have lots else to do: working (part or even full-time), socialising and doing summative assessments. They believe they are already over assessed. On the other hand, academics also believe that students are over assessed and it is they (academics) who have the increasing marking loads. They also

[6] Knight P T (2002) Summative Assessment in Higher Education: practices in disarray, *Studies in Higher Education*, 27, No 3, 275-286.

think that they themselves are over assessed (research assessment exercise; subject reviews, by the Quality Assurance Agency for Higher Education; peer review etc) and they too behave strategically, perhaps avoiding additional assessment that does not 'count'.

In summary, maybe it is more desirable to embed and integrate skills development in the curriculum. Professor Cole-Hamilton will present the case for this approach in the chemistry curriculum at the University of St Andrews. However, this requires serious attention to be paid to curriculum design and commitment (as well as skill) from staff to implement such programmes. An alternate approach where the students develop career development - associated skills through taking an elective credit-bearing module in Career Planning has been pioneered at Stirling University in Scotland. This approach will be presented by Graham Nicholson, now Director of the Careers Service at the University of Dundee.

How should transferable skills be assessed?

This question naturally flows from the preceding section where transferable skills are formally assessed for credit. In some cases, assessment is indirect. Here, students only engage in a process designed to develop skills as a means to an end, creating some academic discipline-based product of learning (an assessed assignment) that is judged for quality of content. More often, assessment is direct, where skills, such as giving a verbal presentation, are explicitly defined as a learning outcome and criteria for levels of performance are given quantitative credit ratings. Marks or grades (for different components eg job study, CV, application, presentation, interview, action plan) in Stirling University's Career Planning module are aggregated to produce a final credit-bearing grade. Graham Nicholson will expand on this approach in his presentation.

An alternative approach is for transferable skills to be assessed informally. This often means engaging students in assessing their own level of development in particular skills. Marks or grades - as such - are not at stake, but often, especially if students make early self-assessments, they can then gauge their own progress after engaging in learning opportunities where further skills development has been possible.
Jean Gowans of the Careers Advisory Service at the University of St Andrews, presents an informally assessed, 'certificated' but not academic credit-bearing, voluntary (but monitored) attendance programme in Career Development. Further, an example of pre and post self-assessment of transferable skills by contract research staff at the University of St Andrews, demonstrates the effectiveness of even short-course taught programmes. Finally, another informal route for acquiring transferable skills is afforded by students who volunteer as part of a network of peer support for their colleagues and engage with PDP as a means of recording their experiences. These activities are completely extra-curricular. Chris Lusk, Director of Student Support Services at the University of St Andrews, presents the way she informally assesses skills development, through provision of a reference for students wishing to pursue further related employment. This reference is informed by student-authorised consultation with PDP records. Clearly, the learning, in this case development of transferable skills, can therefore be represented in different ways, some credit-bearing, some not.

How should such learning be represented?

The UK approach to higher education, until relatively recently, has focused upon the final year (classified honours) degree certificate. At least this was the focus, perhaps until the skills revolution, instigated in the early 1990s by the Enterprise in Higher Education Initiative (in England) and subsequently by additional focus upon teaching quality audit processes that searched for examples of implementing the development of personal transferable skills within taught programmes. The honours degree, predominantly though not exclusively, reflected attainment of knowledge and understanding of various disciplines at higher level, sometimes at the frontier of current scholarship through engaging students in research projects and dissertations. With the exception of professional degree level vocational disciplines such as medicine, law, architecture, nursing, performing arts etc, skills assessment was very definitely indirect - academic skills such as analysis and creativity; generic practical skills such as word-processing; and personal skills such as oral communication through presentations were subsidiary to the 'content' which was generally what was rewarded in assignments. Such content was (relatively!) easily marked or graded and notwithstanding the difficulties involved in appropriately aggregating grades, marks and scores, and so on, an overall measure of performance could be ascertained and one of four 'class' labels (First, Upper Second, Lower Second and Third) could then be attached to a student, thus describing their 'attainment'. The students, their institutions and other educational institutions, as well as employers, had a common understanding of what the labels meant. The process of involving peers in external examination of degree programmes 'guaranteed' standards for each of these labels right across the sector. Behind closed doors no one necessarily believed this. Even within institutions, departments argue about how much easier it is for students to get a First in a different discipline and how much harder it is for them to get a First in their own discipline. For departments that award high proportions of 'Firsts' engagement with this debate may be less, especially if their perception is reinforced by discipline-dependent variations in student achievement across the higher education sector.

The current picture is even more complex. Disciplinary knowledge and understanding are still important, but skills are very much on the agenda. In some disciplines, practical skills are not only crucial but the level of attainment is defined as 'mastery', and not necessarily achieved by some arbitrary (40 per cent, 60 per cent or even 80 per cent) pass mark, but the full pass (100 per cent) mark that must be achieved by meeting all the criteria defined in a competency profile for a particular skill. Various transferable skills outlined above may be assessed and graded in such a way that may allow aggregation with other grades from modules. However, they may also present a problem similar to that provided by discipline-based practical skills that require competency (100 per cent to pass), and if not achieved, then fail. Accordingly is it fair - or even correct - to attempt to aggregate a set of grades such as 55 per cent, 70 per cent, 45 per cent, 65 per cent, Pass (100 per cent) and Pass (100 per cent)?

Other learning opportunities either within or outwith the curriculum that are only assessed informally do not afford award of grades and thus cannot be integrated at all with academic grades. Rob Ward, Director of the Centre for Recording Achievement,

will be addressing how such seeming incompatibilities of the higher education summative assessment system can be circumnavigated. The introduction of progress files, that comprise both a transcript, preferably that conforms to an agreed presentational format, and a way for students to represent their own personal, academic and career development through a personal and professional development planning process, has been proposed to offer a representational route that is both more informative and accurate.

There is an increasing interest in technology-assisted compilation of the records of such activities through electronic checklists, electronic portfolios and so on. Professor Cole-Hamilton will mention a system devised by the Royal Society of Chemistry that he and colleagues are piloting at the University of St Andrews with first year students and third and fourth year student mentors. However, such innovations involving developing more efficient, smarter and more aesthetic software systems should not detract us from the key business of assisting students (and more of them) entering higher education from increasingly diverse backgrounds to develop as flexible, adaptable people, committed to learning for life. Some characteristics of such students in higher education as lifelong learners are that they are morally, ethically, socially, culturally, politically and economically sensitive, as well as aware of their own strengths and weaknesses when working codependently as well as independently, and are able to communicate articulately and effectively about these qualities and attributes both in similar and new contexts with potential employers, by using examples and referring to experiences within the curriculum of degree programmes and through other personal experiences gained outwith the curriculum as part of ongoing commitment to career development and employability.

Appendix I

Margaret L Peters, The spelling problem, Ch 1, pp1-17, in *Spelling: Caught or Taught?* Routledge and Kegan Paul Ltd, 1967

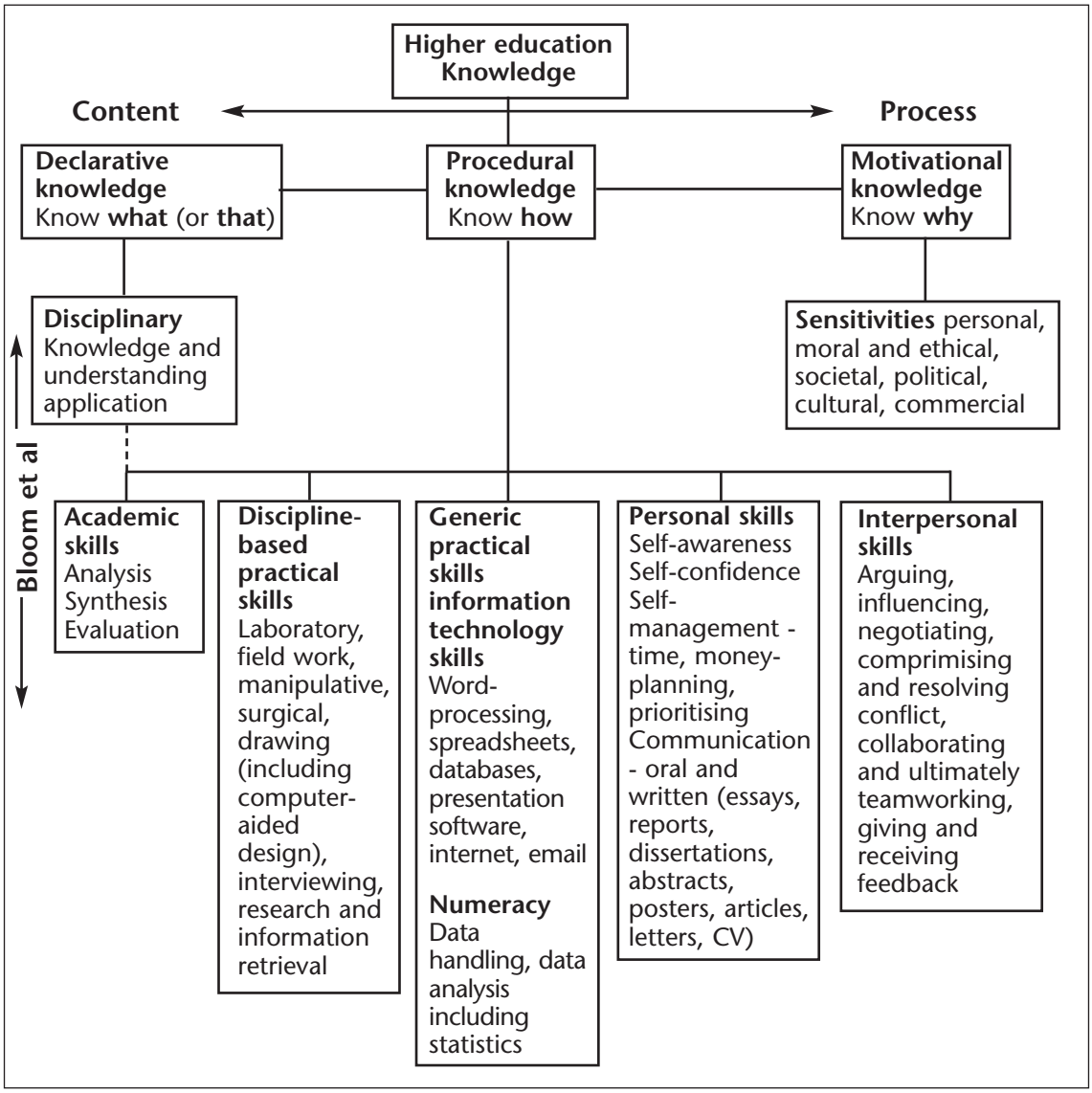
Some primary school teachers were recently asked for their views on the teaching of *spelling*. Most thought that the teaching of spelling occurred incidentally when children were encouraged to write about their interests, with the enlargement of vocabulary involved in such activities. Phrases such as 'when the need arises', 'association of ideas' appeared. There was conviction that *spelling* was a means to an end, and that it should only be the servant of creative writing. There was little evidence of actual instruction in *spelling*. Comments occasionally appeared such as, 'I am in a complete state of uncertainty about the problem of spelling. Please help if you can'. Such cri de coeur points to the dilemma in which teachers have been placed in relation to the teaching of spelling.

The learning of skills, however, is not just a matter of practice; for there are other very important factors as well which will be discussed later. Much of this monograph will be concerned with the factual question of the extent to which a skill such as spelling can be picked up or caught.

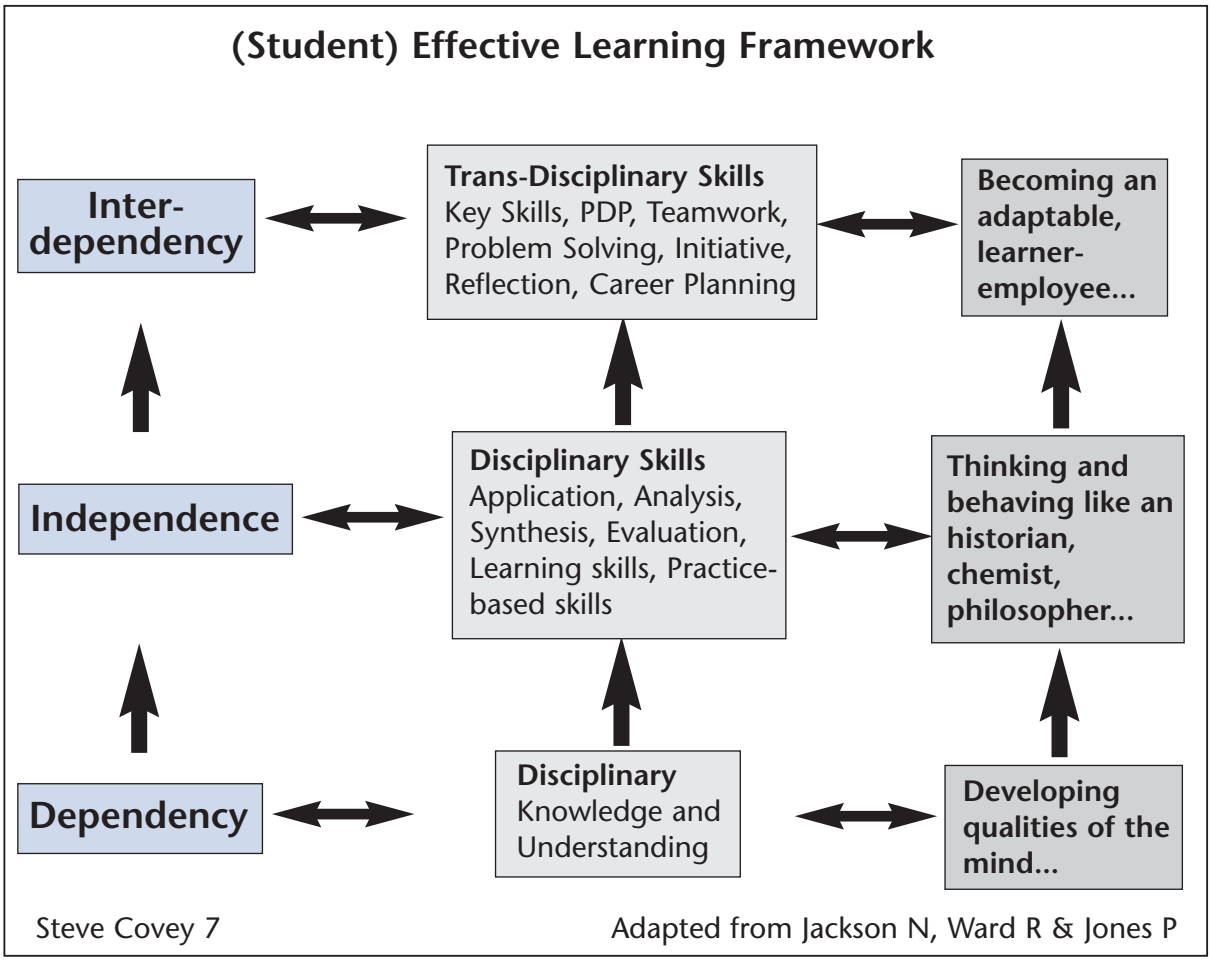
It may well be the case that there are good reasons for learning to spell which the traditional teacher never bothered to make explicit. It may also be the case that there is something to be said for lessons of a formal type in teaching this as well as other skills, which must supplement other more informal ways of handing them on.

Appendix II

Classification view of higher education



Appendix III



Employability: judging and communicating achievements

Professor Peter Knight, The Open University and Professor Mantz Yorke, Liverpool John Moores University

Part of the Learning & Employability series published by the Learning and Teaching Support Network (LTSN) and the Enhancing Students Employability Co-ordination Team. Reproduced with kind permission from the Higher Education Academy.

Introduction

This Guide is for colleagues who are responsible for programmes and who need to ensure that those programmes make a clear contribution to student employability. It is also for those who work with programme leaders, particularly for educational developers. It is probably goes into more detail than senior policy-makers want.

The main idea is that employers value achievements that we find it hard to assess in traditional ways. The argument is that we need a more differentiated, programme-level approach to assessment. This involves disrupting the assumption that assessment has to be about measurement and numbering and substituting the view that assessment is about judgement, which can take many forms.

Preview of the main points

1 Curriculum goals should be assessed because that which is assessed gets taken seriously. That which isn't, doesn't.

2 Different views of employability imply different learning, teaching, curriculum and assessment approaches. This paper concentrates upon employability as the development of complex achievements. Other definitions are not excluded, but they are not central to this paper.

3 Many of the complex achievements that teachers and employers value can only be reliably and affordably assessed on a pass/fail basis. Even then, costs may be greater than a programme can afford.

4 Students need to be 'knowing students' in order to benefit from assessment arrangements that sustain complex learning.

5 Students can be helped to develop claims to those complex achievements that cannot, or ought not to be, assessed by more traditional means.

6 Portfolios are one way of helping students to make claims to complex learning achievements.

7 Self and peer-assessment should be included in a programme assessment plan.

8 The assessment of competence, particularly of fitness to practise, is expensive. It needs to be addressed through a programme assessment plan.

1 Premises

We may plan a fine curriculum and try very hard to implement it faithfully, only to find that what students experienced was rather different. Typically any such mismatch takes the form of aims, such as promoting autonomy, creativity, critical thinking or understanding, being frustrated by assessment arrangements that somehow encourage students to play safe, to rehearse any party line they can detect and to stockpile information in preparation for examinations.

Assessment drives the understood curriculum:

- It tells students what the aims of the curriculum really are, because 'what matters' gets assessed;
- It tells them how to work, because it seems sensible to prefer ways that pay off in terms of good grades;
- It tells them when to work, because tasks that are not assessed give students implicit permission to work longer on their part-time jobs or to spend more time enjoying themselves.

This Guide suggests ways of bringing the curriculum goal of enhancing student employability closer into line with course and programme assessment practices.

2 Employability

The word 'employability', in the context of higher education, implies ability to be employed in a 'graduate job', something rather different from actually being employed. Figure 1 summarises five common descriptions of 'employability' and adds notes on the assessment implications of each.

Each definition of 'employability' has value in some circumstances but the Learning and Employability series concentrates on the fifth, which is most firmly based on research evidence about what employers value.

When employers are asked what they look for when hiring graduates, they are inclined to say that

- having a good degree is necessary but it is little more than a ticket to compete for a job;
- chances are improved when applicants have credible claims in respect of 'key' or 'transferable' skills;
- what they are really looking for - what they use to choose amongst the skilful graduates - is something more complex.

Figure 1 Seven meanings of 'employability' and some assessment implications

What is employability?	Notes	Assessment implications
Getting a (graduate) job	Employment figures are taken as a robust indicator of employability.	No particular implications, as long as HEIs' assessment practices do not impede students in getting jobs.
A consequence of 'having' key skills	The Dearing Report said all students should develop four 'key skills'. Others have been added. There is some scepticism about the whole 'key skills' enterprise (see the companion Guide, *Employability in Higher Education: what it is - what it is not*).	The search is for reliable and valid ways of certifying things such as communication, numeracy and problem-solving. Measurement theory demands repeated, high-quality judgements before achievement is warranted. Arguably, such achievements are too complex for affordable and reliable measurement.
A likely effect of having had good work experience	Work experience consistently correlates with success in the labour market.	See section 7. Reliable and valid assessment of workplace performance is expensive. Formative, conversational assessment is cheaper but cannot contribute to warrants of achievement.
A product of skilful career planning and interview technique	Employability is in part about knowing the rules of the job-seeking game. Most of the unemployed graduates interviewed in Skills *plus* project had fallen down here.	Assessments that help students to identify and then present their achievements effectively are invaluable. This says less about assessment methods and more about making the rules of the learning and employment games very clear.
A mix of cognitive and non-cognitive achievements and representations	See, in this series, the following Guides: *Employability in Higher Education* and *Embedding Employability in the Curriculum.*	Understanding can be assessed reasonably reliably, and often affordably. Skills present problems (see above), although simplified skills can be reliably assessed. Beyond that, assessment may be best treated as an aid to learning and claimsmaking: thought needs to be given to letting go of trying to certify complex achievements.

Note: The darker the shading, the more appropriate it is to give priority to formative assessment to support claimsmaking because it is harder to see how reliable, useful, affordable and valid assessments could lead to generalisations about competence and performance - to certificates and warrants.

Consider the three lists in Exhibit 1, which are fairly representative of what researchers find that employers say.

Such research underpins our description of employability as

> a set of achievements, understandings and personal attributes that make individuals more likely to gain employment and be successful in their chosen occupations.

Now, if employability is 'being attractive to employers' and if that means 'degree and skills and what is in Exhibit 1', then a conclusion is that higher education is firmly in the business of promoting complex learning - in other words, its goals are necessarily more fuzzy and extensive than many might acknowledge. Elsewhere we (Knight and Yorke, 2003b) have described employability as a blend of understanding, skilful practices, efficacy beliefs (or legitimate self confidence) and reflectiveness (or metacognition). Notice that some of these elements resist conventional, measurement-driven assessment approaches. This is a central point. Insofar as employability involves the promotion of achievements that cannot be specified completely and unambiguously, it cannot be measured, although local judgements can be made and others, such as employers, might choose to generalise from them.

Exhibit 1 What employers value in new graduates

Peter Hawkins and Jonathan Winter (1995) highlighted 'career management skills and effective learning skills': self-awareness; self-promotion; exploring and creating opportunities; action planning; networking; matching and decision-making; negotiation; political awareness; coping with uncertainty; development focus; transfer skills; self-confidence.

Lee Harvey and colleagues (1997) found that employers want graduates with knowledge; intellect; willingness to learn; self-management skills; communication skills; team-working; interpersonal skills.

John Brennan and colleagues (2001) found that UK graduates considered the top ten competencies required in current employment to be: working under pressure; oral communication skills; accuracy, attention to detail; working in a team; time management; adaptability; initiative; working independently; taking responsibility and decisions; planning co-ordinating and organising.

3 Judging achievement

Assessment is often a high-stakes business, by which we mean that it is vital that judgements have considerable public significance, as with examinations and graded coursework. When achievement is to be warranted - to be publicly certified or attested - then the judgements need to be reliable.

Some achievements, especially those connected with understanding and the more straightforward skills, can be fairly reliably assessed in much the same way that they are currently assessed. Practice might be improved by refining the assessment tasks; writing programme assessment plans to ensure that these learning outcomes are repeatedly assessed throughout the programme; developing assessment criteria which students have and understand, and which assessors use; and making resources available for double marking of all summatively-assessed work, with the exception of work which leads to clear right/wrong answers.

Measurement theory suggests that much higher education assessment practice, which attends to complex achievements, cannot deliver the certainty it feigns. This is:

- Partly a funding problem (measurement of any except the simplest things is expensive);
- Partly an orchestration matter (greater certainties would be possible if there were programme assessment arrangements that generated multiple estimates of programme learning outcomes);
- Partly a matter of pragmatics (how much assessment is it reasonable to impose on learners and teachers);
- Partly an epistemological fact (some things resist measurement).

It is therefore timely to re-appraise 'assessment'.

The problem for high-stakes assessment is that it does involve trying to generalise from performance on a couple of tasks in one course. However, in social science there is a widely-held view that generalisation requires multiple judgements, made by more than one observer, using indicators that are understood by them and by the students, and which the observers are skilled in using in a consistent manner. This is inconvenient because it is expensive and difficult to come up to these standards. If we understand 'assessment' to mean high-stakes measurement for the purposes of warranting, then there are problems with seeing how we could assess some of the notions of employability in Exhibit 1, let alone how we could afford to do it.

The less that assessment practices conform to the demands of measurement theory, the less reliable are any predictions about future performance, especially when it comes to performance in quite different contexts. This helps to explain why employers are so often disappointed by graduates. Universities and colleges often suggest that graduates have certain achievements to their credit but base their judgements on scanty evidence which is often associated with 'tame', rather artificial tasks done in academic surroundings.

Yet complex achievements are assessable, although they may resist (affordable) measurement. Apart from the scientific approach to judgement (measurement), there is a legal approach (weighing evidence and claims) and an artistic one (connoisseurship). Many interesting results of higher education cannot be well captured by the measurement approaches developed in natural science, but they can be reached by other approaches. If we adopted either of those other approaches:

1 Judgements would be based on appraisal of evidence of achievement.

2 There would be indicators to inform the identification of common features of better and worse performances.

3 Those judging would be familiar with the indicators, as would the learners who create the evidence.

4 If the stakes were high, perhaps because a licence to practise rested upon the outcome, several expert judges would review many pieces of evidence, taking particular care over boundary decisions (Pass/Fail). If the stakes were lower, as when creating feedback to help further development, assessment could take any form likely to create useful and informed suggestions. Given the relative costs of

formative and summative assessment and the difficulties high-stakes assessment has with complex achievements, the effect would be to prefer low-stakes approaches to judging many achievements.

A major objection to this position is the claim that students will not take this formative assessment seriously; nowadays they are so instrumental and preoccupied with their part-time employment that they will only do the minimum to get their upper second class degree, so the suggestion that assessment systems should make more use of low-stakes procedures looks otherworldly. The fear is that whatever is touched by low-stakes assessment alone will be ignored. Four responses are:

- Students are more likely to take low-stakes, formative assessment seriously if they understand the purposes. This is not a matter of telling them once but of saturating programme and module handbooks, as well as teachers' discourses, with messages about the importance of formative assessment.
- Low-stakes tasks are taken seriously when they are preludes to high-stakes ones. Set two low-stakes task, telling students that the third task in the sequence will be of a similar sort and will be for high-stakes purposes.
- Low-stakes tasks can be done in class as seminar activities. For example, students come to class with a one-page plan of a paper, which is then reviewed by two peers.
- Formative assessment is valuable in its own right. A recent review concluded that formative assessment improves learning; if best formative assessment practices were adopted in mathematics it would raise 'average' countries such as England and the USA into the top five. The effect size of 0.7 is '... amongst the largest ever recorded for educational interventions' (Black and Wiliam 1998: 61).

The claim is that alternatives to the measurement model of assessment are available for summative and formative purposes, although the costs of maximising reliability mean that there is much to be said for using them for low-stakes formative purposes.

How does this come together in an approach to the assessment of employability? We suggest that teams look at programme specifications and put the learning outcomes into one of three groups:

- Those that can be readily assessed for high-stakes purposes - recall of information, routine application of formulae and procedures.
- Those that, for a variety of practical, theoretical and ethical reasons, virtually defy high-stakes assessment - legitimate self-confidence, taking responsibility, willingness to learn.
- Those which can be judged in a tolerably-reliable way if sufficient time and money is invested in them - assessments of workplace competence, portfolios, performance in groups.

This is the basis of a differentiated approach to assessment.

4 Orchestrating the assessment of employability at programme level

In order to see better how employability might be assessed, we exploit this more differentiated view of assessment. Recall that some of the objections to formative assessment have just been addressed.

Biggs (2003) has been influential in his advocacy of constructively aligned curricula. The basic idea is simple and powerful: students have the best chance of learning when curriculum, learning, teaching and assessment are pointing in the same direction: when they are aligned. The 'orchestration' or 'tuning' approach described here involves making a series of small changes to a programme to enhance the contribution that assessment makes to the enhancement of student employability.

Typical moves are:

1 As we have just suggested, take the outcomes of learning identified in the programme specification and identify those that can readily be summatively assessed - knowledge and understanding goals are often assessed in this way. Then identify those that call out for formative assessment approaches. Re-examine the residue, making an economic judgement about the assessment arrangements that would be implied, and whether the programme can afford to warrant their achievement.

2 With some programmes, all of the modules that lead to the award are prescribed. Where this is not the case, the next step is to identify the combinations of modules that students most commonly take for the award: you are identifying the main pathways they take towards the award.

3 Once pathways are identified, approach the leaders of pathway modules and ask them to refer to the programme specification and identify, say, the three outcomes that get sustained attention in each course. They will no doubt say that their work touches upon many outcomes, but the aim here is to identify those that are most seriously addressed.

4 Collate the returns and, if need be, negotiate with course leaders to tune the programme for:

- Gaps - programme outcomes that are not addressed.
- Redundancies - outcomes that get too much assessment attention.
- Bunching, where all the attention to an outcome is at one level and there is no obvious educational rationale for that being the case.

5 Now ask leaders how the assessment arrangements touch the three key module outcomes.

6 Collate the returns, again looking for gaps, redundancies and bunching in terms of

- Assessment of learning outcomes - are some outcomes missed or over-addressed?
- Task variety - are essays over-used, for example?

7 Again, follow this with negotiations to achieve a better orchestration of assessment tasks and the learning achievements to which modules give priority.

8 It is essential that those teaching on the programme know what is being addressed and where, and that the material they give students explains how the module and its learning intentions relate to the programme and its learning intentions. For example, a module handbook should state that, say, three outcomes identified in the programme specification will get sustained attention, and remind students of where they can refresh their memories of what the specification says.

9 Students need to be 'knowing' students - they need to know what they are supposed to be learning, how, how their achievements will be judged and for what purposes. This usually entails rewriting programme and module handbooks and ensuring that parallel guidance purposes, whether for academic or career purposes, also carry the same messages. To do this, programme teams have produce a coherent and convincing account of their programme.

10 Students need to learn ways of representing their achievements to employers and graduate schools.

The following sections develop the last two points.

5 'Knowing' students

Students come to class with learning histories that have shaped their beliefs about the rules of the academic game, particularly beliefs about what learning is, what teachers do and what assessment is for. Many innovative teachers have found that students resist academic practices that do not conform to those expectations, partly because they do not understand the good sense behind them. The approach to assessment and employability that has been outlined here is sufficiently distinctive to need full and repeated explanations if students are to understand, follow and appreciate the new rules of the assessment game. It is necessary to explain at least three things very clearly:

1 Why there is such an emphasis on formative assessment.

2 Why students should expect to undertake peer and self-assessment. Formative assessment works well when it creates thoughtful feedback on improving performance, especially when feedback is related to assessment criteria that are known, understood and used. The practices of judgement learned through an active engagement in peer- and self-assessment contribute to student employability and are a basis for self-regulation and life-long learning.

3 That formative assessment will not work unless students and teachers take it seriously. Teachers might want to reinforce the principle by requiring students to provide evidence that they contributed criteria-related feedback to others on a specified number of occasions during the course.

These explanations should go in the course handbook and, ideally, be closely related to the course assessment plan.

Many students will resist attempts to involve them in novel practices - for example, self and peer-assessment: some because they lack confidence and dislike the uncertainty that comes from unfamiliar practices, and others because feel that they have paid a great deal to be taught and expect the tutor to do the marking and not shuck it off on to other students.

They are least likely to be upset by the idea of peer- and self-assessment if they are introduced to it in Year 1, understand the purposes and benefits, and see others taking self or peer-assessment for granted. Yet it takes persistence and a coherent curriculum to form the learning communities and cultures that embrace new approaches to assessment, teaching and learning, as Mentkowski and colleagues (2000) show. When this sustained, programme-level action is not possible, teachers may still innovate in

individual modules, while being prepared for objections from students who prefer the familiarity of established methods and are suspicious of new ones.

6 Portfolios, PDP and assessment

It is quite common in the professions to consider portfolios when judging fitness to practise, for appointment or promotion. Students or applicants will usually select from their collection of material those items that can be presented as good evidence to support their claims to the achievements that define competence or higher grade performance. Teachers in higher education are also increasingly expected to produce portfolios in support of their claims to achievement (Wright and Knight 1999) and most postgraduate programmes validated by UK Institute of Learning and Teaching in Higher Education require teaching portfolios.

David Baume (2001) has produced a briefing on portfolio assessment for the Generic Centre and its website holds a number of documents on good practice in personal development planning (PDP) and the creation of portfolios.

Portfolios are notoriously difficult to assess reliably, although Baume and Yorke (2002) describe an approach to doing so. Five sources of difficulty are:

1 The claims to achievement and the evidence used to support them tend to diversity. Greater convergence, which is necessary for reliable and efficient grading, requires indicators that helpfully describe and illustrate the assessors' expectations. The price is that these measures can curb students' creativity, limit flexibility and reduce students' feelings of having some ownership of the PDP process and the portfolios it produces.

2 Even when indicators are helpful, there will be considerable variations in the evidence presented. The variations may represent different degrees of achievement, but they will also represent different circumstances of achievement and different judgements of how best to make the claim to success.

3 There will always be differences in weighting between elements of claims to achievement. For example, there are some 30 elements to an English specification of teaching competence. It is unlikely that claims will treat them all equally. The more that assessors have to judge how to respond to such imbalances, the more elusive is reliability.

4 Portfolios tend to be long. Long documents are costly to assess. Costs multiply if grading is more complicated than pass/fail.

5 The reliable assessment of portfolios demands expert judges who are well-trained in using indicators in consistent ways. However, the more complex the assessment task - and portfolio assessment is as complex as it gets - the more elusive is reliability, and the higher the training and quality assurance costs.

The more that reliability is emphasised, the more assessment costs soar and the more students' freedom to develop their claims is curbed. In the context of assessment for employability, the suggestion is that portfolios should only have formative purposes. Exhibit 2 contains some notes for those wanting to use them summatively.

Exhibit 2 The summative assessment of portfolios
If you want to get reliable grades from portfolios:

1 Invest resources in the summative assessment so that graders can be well-trained, portfolios can be independently graded by more than one assessor, and there are resources for thoroughly resolving differences. This implies not using a lot of tutor time on other summative assessment tasks in the same course, unless the contributory judgements are just 'sufficient for progress to the next stage' or 'not ready to progress to the next stage'.
2 Ensure that there are clear indicators and plenty of examples of good practice, available to students and teachers. The tighter the brief, the easier it is to get reliability (but the harder it is for students to develop the claims they want to make in valid ways). However, the danger in tightening the brief is of creating an approach that students will see as a strait-jacket.
3 Encourage or require students to discuss their portfolio claims with each other before submission. This will clarify understanding of what is required and make reliable grading easier.
4 Reduce the number of decision points. 'Pass/fail' grading (one point) is cheaper than 'distinction/merit/pass/fail/non-redeemable fail'. Grading on four elements of a portfolio is cheaper than grading on 24. In all cases, detailed assessment attention might be concentrated upon borderline and failing portfolios so as to help the students to improve to an acceptable level.
5 Consider grading only the claims, which can be set out as a one or two thousand word preface to an annotated file of evidence. Sample the evidence for appropriateness but only look in any detail where there is cause for concern.
6 Look over the portfolios to ensure that they pass the threshold of adequacy but do not grade them. Set students whose portfolios are adequate a separate task, perhaps under exam conditions, that capitalises on the learning that the portfolios represent. Good portfolios should support better performance on this task than those that showed minimal effort. Those producing inadequate portfolios are not given their grades until their work is judged adequate.

Consult Baume and Yorke (2002) for an account of attempts to improve the reliability of portfolio assessment.

When used formatively, portfolio-making is treated as an opportunity for PDP. What follows is a summary of the ways in which portfolios are used developmentally in one social science department.

Students start with the programme specification, which explains the programme's learning intentions. They begin by adding to the standard list the outcomes of learning that they value and either can document on the basis of what they have done in school, in their part-time and vacation jobs, and through their leisure activities. They then review this new set of outcomes and do two things: identify the sorts of claims to achievement they can make in respect of each outcome on their lists; and identify areas for development and consider ways of doing something about them. The portfolio they create and develop throughout the programme has three main parts, described in the course handbook as:

> *Section 1: claimsmaking*. First, there are your claims to achievement, which will be written in continuous prose, highlighting the points that you think present you

to your best advantage. Although you will inevitably refer to your cv and say something of the courses you have done, jobs you have had and qualifications gained, this section is about making claims based on those experiences and achievements…

Section 2: associating claims with evidence: The second part, which may be best presented as a table which you create and maintain in electronic form, should list your achievements - such as practical, intellectual and key skills - say a little about each and refer readers to the evidence that fleshes out the claim…

Section 3: The evidence. The third section is likely to be a box or a more sophisticated filing system containing the evidence you want to use in support of your claims…it is important that employers - and you - are quickly able to understand which claims are supported by one or more items of evidence and why. For example, you might have put a particularly good essay in your file because it shows high academic achievement, good presentational skills, ICT skill and numeracy. In which case, make sure there is a note explaining how this item is to be read as evidence of the claims you are basing upon it.

This claimsmaking enterprise rests on the programme assessment plan in that:

- Students need to be quite skilful at reflecting on their own learning and achievements if they are to appraise their attainments and plan for future learning. If this is not encouraged by programme assessment practices, students will generally be disadvantaged.
- Students need to be familiar with the programme learning indicators, to have seen examples of their use in practice and to have a good, experience-based understanding of what they mean, as expressed by the grade indicators.
- Each of the foregoing points assumes an experience - probably a substantial experience - of peer and self-assessment.
- Students need evidence of achievement, particularly in respect of those outcomes of learning that the HEI does not warrant. This means that they need to do tasks that support development in those areas and that provide feedback both on performance and also for improvement.
- There need to be plenty of tasks with formative assessment purposes in order to support development.
- Portfolio development is integral to the curriculum. That means telling students that it is an important curriculum activity, giving them guidance on creating and maintaining a portfolio, providing tutor support and guidance, creating opportunities for learning conversations around portfolios, and aligning this part of the assessment system with the HEI's academic and personal guidance systems.

This portfolio work, if well-planned, should help students to develop claims to achievements that a department does not summatively assess, help them to review all their learning, and prompt them to identify areas and opportunities for development. However, enthusiasm for progress files, portfolios, dossiers etc. has not always been shared by students. Unless students are enrolled in a programme culture that values and supports portfolios, resistance and indifference are likely to follow.

7 Assessing competence and work-based learning

Employability is often associated with competence, particularly when an employer advertises for someone with particular achievements - for example, in operating Linux software, in teaching children with special educational needs, or in post-partum care. Although there is a history of treating competence as a stable set of distinct but generic elements, researchers such as Michael Eraut (1994) have a lot to say about the degree to which it is content-specific, situationally-variable and holistic rather than an agglomeration of separate skills. As with the notion of employability itself, definitional matters have ramifications for assessment, as Figure 2 shows.

Concept of competence	Assessment implications
1 Competence as having sufficient knowledge.	Assess knowledge. Better still, assess understanding, although this is harder to do reliably.
2 Competence is being adept at problem-solving.	Assess quality of solutions to well-defined professional problems. Better still, assess quality of suggestions about ill-defined situations (problem-working, which is more authentic than problem-solving).
3 Competence as having clinical or practical skills.	Assessment of individual skills through OSCE (objective, structured clinical exam). Better still, observation of skills in authentic settings.
4 Competence is being effective and efficient in practice.	Judgement on the basis of evidence of effective and efficient practice 'in the wild'. Will include observation, peer-appraisal, appropriate performance indicators and, perhaps, portfolio claims to achievement.
5 Competence is tantamount to effective and reflective practices	As above, plus evidence of reflection, perhaps in the form of evidence of continued learning within an area of specialism.

Figure 2 Assessment implications of varying views of competence

When the stakes are high, as they are when fitness for practice is at issue, then high levels of reliability are needed. This is not always easy to achieve because it depends on repeated assessments of the elements of competence in different settings and using a variety of authentic assessment tasks. Different assessors should be involved. They should understand and use the same criteria. Their consistency should be monitored and disputes should not be settled by splitting the difference between two assessors' marks.

However, as definitions of assessment move from the first row of Figure 2 to the fifth, complexity and cost increase as well, and curriculum designers have to think in terms of programme, not module, assessment plans. Direct summative assessment of competence, as it is defined in rows 3, 4 and 5, is expensive, which is why it is so tempting to use the simpler methods listed in rows 1 and 2, even though they only give evidence about impoverished concepts of competence. The temptation is to make the easily assessable important, regardless of whether what is important is easily assessable.

We might sidestep this problem by turning high-stakes assessments of competence into coaching or formative ones. It is cheaper and appropriate where 'fuzzy' or 'soft' achievements are concerned. However, it is seldom realistic to say that workplace learning will only be assessed formatively. Alternatively, we might have high-stakes assessments but treat them as purely local verdicts, not general warrants to competence. (A series of local assessments might, over a programme, be sufficient for some generalisation about competence.) The snag is that many HE programmes need to produce warrants in the form of statements of fitness to practise, which also limits the use they can make of the cheaper options of assessing knowledge or skills proficiency. Although professional bodies are often prepared to negotiate about the interpretation of their regulations for registration, departments can often find themselves severely constrained by the rather quaint ideas about assessment held by some professional bodies. The best advice on meeting professional bodies' requirements is likely to come from subject associations and from the relevant LTSN subject centre.

When summative data are wanted, then the costs will be high. Eight contributions to tackling the provision of summative data are as follows.

1 Establish a programme assessment plan. The savings may not be obvious because they will stem from reduced uncertainty for all concerned.

2 Invest in materials that explain to students, employers, assessors and others how achievements are being understood and assessed, what competent performances will look like, and how assessors will make their judgements. This helps all concerned by reducing misunderstanding and confusion.

3 Make sure that students know the rules of the game, have plenty of exposure to examples of competence and follow a well-designed professional programme. It is easier to assess competence when the curriculum routinely promotes it and students know what the curriculum is doing.

4 Invest in coaching and assessment training for staff. The more that teachers are agreed on what would count as evidence of competence, the more efficient and effective the system. However, 'staff' in this context includes workplace assessors and those departmental colleagues who liaise with them. Not only is it expensive to train workplace assessors and often hard to get them to agree to be trained, it is also expensive to establish quality assurance systems to make sure that comparable standards are being similarly applied across a range of settings.

5 Aim to have few assessment decision points by, for example, preferring pass/fail judgements to percentage grades. More-than-competent performance can be appreciated without having to put numbers on it.

6 Where competence is established, stop summatively assessing and concentrate on other areas of concern. Again, that is no bar to giving less formal feedback to encourage a move from competence towards excellence.

7 Provide plenty of formative tasks that lead to fewer, sharp high-stakes tasks upon which judgements of competence are based.

8 Encourage a claimsmaking approach, in which learners are themselves responsible for producing evidence of achievement.

These eight suggestions comprise a systemic approach to the assessment and promotion of competence which is a model for the assessment of employability in general. There is no avoiding the fact that the assessment of competence is expensive (unless competence is defined in Figure 2, row 1 terms). For example, it is said that a major bugbear in the assessment of competence is establishing the validity of claims to achievement coming from work experience or other activities outside of higher education. Agreed. Point 4 above alludes to the costs of making sure that judgements within a programme are sound. When it comes to appraising claims made on the basis of extra curricular activity, it is common to require students to produce portfolios to support the claims and/or to viva the student. Both are expensive.

A differentiated programme-wide approach to assessment can cope with expensive elements, such as the assessment of competence and claims based on workplace learning, because the programme team can decide to use cheap assessment approaches in some modules in order to free up the resources to invest in expensive assessment practices in others.

Further advice can be found in Gray's (2001) Generic Centre paper on the assessment of work-based learning.

8 Suggestions for action

Teachers and module team leaders could*:*

1 Reconsider the balance between formative and summative assessment purposes.

2 Consider extending the range of assessment methods.

3 Network. LTSN subject centres, professional associations and other interest groups in this country and overseas are good sources of ideas that can be borrowed and customised.

4 Hold on to the idea that many of the assessment problems you would like to solve are either not solvable or most sensibly tackled at system level. Teachers are prone to feel guilt (Hargreaves, 1994) but this is seldom appropriate here because solutions often lie outwith their power.

5 Aim to extend the range of assessment methods in use and concentrate them on directly assessing a few - three or four - achievements per module.

Programme leaders could:

1 Make programme assessment practices a priority for departmental attention over, say, the next three years.

2 Review the amount of assessment on a programme, looking at the range of methods and the balance of formative and summative assessment. It is not unusual to find considerable imbalances.

3 Look for consultancy/evaluation help on the design and management of assessment systems. There is a place for workshops on topics of interest but there is, we suggest, a massive, unmet need for consultancy support. Educational development units, LTSN subject centres, subject and professional associations and national quality enhancement agencies, such as the LTSN Generic Centre, can all help here.

4 Get some programme-wide criteria in place to help thinking about assessment. Concentrate on identifying the sorts of performance associated with, say, a lower second class degree. Don't take these indicators too seriously but treat them like a 'starter culture', a way of developing conversations about what is involved in assessing learning. They can be the beginnings of a common assessment language.

9 Questions

Question. What does the assessment of employability mean for assessment practices?

Suggestion

- Using a wide range of assessment methods.
- Orchestrating assessment arrangements so that they inform us about a wide range of achievements.
- Taking a more planned approach to assessment, for example by ensuring that module practices are better integrated with the programme specification.
- Ensuring that assessment practices dovetail with learning tasks and teaching sequences.
- In short, arranging things so that students experience coherence, rather than disarray, and breadth of engagement, rather than narrowness.

Question. What strengths are there in current assessment practices?

Suggestion

The past fifteen years have seen the use of a wider range of assessment methods, which in turn has encouraged a wider range of achievements. Admirable as these developments are, they tend not to be joined-up, so student assessment experiences can be uneven. In some subject areas practices remain quite traditional, touching a narrow range of achievements.

Another strength is that students continue to take assessment seriously and are generally motivated to do well when the stakes are high.

Question. Where do current practices fail?

Suggestion

Where the range of methods is too narrow, the range of learning intentions promoted by current methods is often too limited, the workloads on teachers can be excessive and alignment with teaching and tasks is not always good. The course-programme fit may not be as close as it could be.

Current practices do not, generally, support PDP and claimsmaking, nor do they routinely create feedback to students that is couched in terms of course and programme learning indicators.

They perpetuate the unhelpful view that assessment means measurement.

Question. How do assessment practices need to evolve in order to accommodate the employability agenda?

Suggestion

The answer is implicit in the response to the last question and detailed suggestions for course and programme leaders have just been presented.

Arguably, if assessment to support employability and other complex learning intentions, the most important thing is for there to be a change of thinking - a recognition that assessment, as described here, is not what it is often assumed to be (ie 'measurement').

Question. What is going on that might help or hinder this?

Suggestion

Hindrances for teachers include innovation fatigue; the allure of research; lack of resources; lack of educational consultancy support; wariness of the term 'employability'; high-choice modular programmes; a belief that assessment=measurement; a lack of cross-departmental thinking; weak traditions of departmental leadership; league tables; confusing of employment indicators with 'employability'; the growing 'casualisation' of academic employment; and the cost of coping.

Hindrances for students are the experience of very different assessment cultures; their frequently fragmented experience of programmes; the need to do paid work; the psychological allure of playing safe; the instrumental need to get a 2:1; and, for some from non-traditional backgrounds, the strangeness of it all.

Opportunities for all include the requirement that PDP be available to all students by 2005; the beginnings of a move away from highly modularised curricula; the range of work now being done on programmes and coherence; the professionalisation of teaching, notably in the Academy for the Advancement of Learning in Higher Education; increased rewards for good teaching; and the subject-based resources, events and guidance provided by LTSN subject centres and subject associations.

Learning more

This Guide draws heavily on the 2001 pamphlet *Skills plus: assessment and employability* (www.open.ac.uk/vqportal/Skills-Plus/publications.htm). It anticipates some ideas that will be in Knight, P. and Yorke, M. (2003a) *Assessment, Learning and Employability.* Buckingham: the Society for Research in Higher Education and the Open University Press.

The Generic Centre's Assessment Pack is a useful and accessible set of booklets on assessment matters (www.heacademy.ac.uk/resources/asp?=filter_fields&section=generic&type=some&id)

John Biggs' work on constructive alignment takes a line similar to one developed in the Skills *plus* project - Biggs, J. (2003) *Teaching for Quality Learning at University* (2nd ed). Maidenhead: Society for Research in Higher Education and Open University Press. His work has been very influential.

David Gosling and Jenny Moon (2001) give clear advice on the design and use of learning outcomes, which guide low-stakes judgements and determine high-stakes assessment - *How to Use Learning Outcomes and Assessment Criteria.* London: Southern England Consortium for Credit Accumulation and Transfer.

For a useful guide to common assessment practices, see Hounsell, D., McCulloch, M. and Scott, M. (Eds.) (1996) The ASSHE Inventory (Edinburgh: University of Edinburgh and Napier University) and at www.heacademy.ac.uk/asshe/project.asp Also Brown, G., Bull, J. and Pendlebury, M. (1997) *Assessing Student Learning in Higher Education.* London: Routledge.

Two further sources are as follows:

On assessment-as-learning: Mentkowski, M. and associates, (2000) *Learning that Lasts: integrating learning development and performance in college and beyond.* San Fancisco: Jossey-Bass.

On self-assessment: Boud, D. (1995) E*nhancing Learning through Self-assessment,* London: Kogan Page.

The references in the text are to:

Baume, D. (2001) the assessment of portfolios, accessed via www.heacademy.ac.uk/resources.asp?process=full_record&section=generic&id=6

Baume D and Yorke M (2002) The reliability of assessment by portfolio on a course to develop and accredit teachers in higher education. *Studies in Higher Education* **27** (1), pp.7-25.

Black, P. and William, D. (1998) Assessment and classroom learning. *Assessment in Education.* 5 (1), pp.7-74

Brennan, J., Johnston, B., Little, B. Shah, T and Woodley, A. (2001) *The Employment of UK Graduates: comparisons with Europe and Japan.* London: The Higher Education Funding Council for England.

Eraut, M. (1994) *Developing Professional Knowledge and Competence,* London: Falmer Press.

Gray, D. (2001) the assessment of work-based learning, accessed via www.heacademy.ac.uk/resources.asp?process=full_record&section=generic&id=11

Hargreaves, A. (1994) *Changing Teachers, Changing Times.* London: Cassell.

Harvey, L., Moon, S. and Geall, V with Bower, R., 1997, *Graduates' Work: Organisation change and students' attributes.* Birmingham, Centre for Research into Quality (CRQ) and Association of Graduate Recruiters (AGR).

Hawkins, P. and Winter, J. (1995) *Skills for Graduates in the 21st Century*. London: Association of Graduate Recruiters.

Knight, P. T. and Yorke, M. (2003b) *Learning, Curriculum and Employability in Higher Education.* London: Routledge/Falmer.

Mentkowski M & Associates (2000) *Learning that lasts: integrating learning, development, and performance in college and beyond.* San Francisco: Jossey-Bass.

Wright, W.A., and Knight, P. T. with Pomerleau, N. (1999) Portfolio people: teaching and learning dossiers and the future of higher education, *Innovative Higher Education,* 24(2), 89-102.

Personal development planning with tutor and peer student mentoring: interim report of an experiment in implementation (warts and all)

Professor David Cole-Hamilton and Dr Fiona Gray, School of Chemistry, University of St Andrews

Introduction

The School of Chemistry made the decision to introduce personal development planning (PDP) to undergraduate students in the academic year 2002-03. The aim is that by 2005, when all higher education institutions are required to implement PDP as a result of recommendations of the Dearing Review, an established system will be in place. The School of Chemistry is using the Undergraduate Skills Record (USR) produced by the Royal Society of Chemistry (RSC). It is available as hard copy or can be downloaded from the website for students to complete electronically. To implement PDP, the School of Chemistry proposed a cascaded mentoring approach using tutor and peer student mentoring and this has been developed on a pilot scale.

Royal Society of Chemistry Undergraduate Skills Record

Both printed and electronic versions (www.rsc.org/ugskills) of the RSC's USR are available. The first section is for recording personal details, qualifications, work experience, as well as awareness of departmental and university-wide facilities and services. The main section, Skills Audit, is for the students to complete throughout their course, and guides them through assessment, reflection and the planning of different skill development. The final section is a Skills Profile where the student has the opportunity to generate a summary of their skills development using examples highlighted in the Skills Audit section.

The Skills Audit is broken down into nine skills:

- planning and organisation
- study skills
- handling information
- communication skills
- working with others
- scientific/practical skills
- improving learning and performance
- information and communications technology skills, problem solving.

By answering a series of questions on each skill, the students rate themselves from the Ability Ratings given, rate themselves overall and then summarise the reasons for this scoring, including evidence to support it. After reflection, they are asked to set themselves a realistic development target for the next phase. Guidance and examples are given to help students fill the Skills Audit in appropriately.

Although the USR is designed to be of particular benefit to students working towards a degree in chemistry, the Skills Audit is general enough to be appropriate to students of most scientific disciplines.

Method

The RSC's USR was made available to first and third year students. These two groups were targeted for the following reasons.

- It would be difficult to support and mentor students from all years at the same time.
- First year students will have the opportunity to build up a complete profile over their entire university career.
- As mentor, I was already known to about half the first year class through tutoring in the previous semester.
- Third year students were keen to be involved in supporting the first year chemistry students in a number of areas and this fitted in well with their enthusiasm.
- We knew the third year students well and they would be able to give an honest evaluation of the usefulness of the forms. The benefits were clearer to them as many had either prepared for industrial placement interviews or were beginning to focus on career interviews in the near future.
- Third year students would be available the following year to support the introduction of PDP to a new first year class.

In week 8, semester 1, all participating students were given an introductory talk, received a printed copy of the USR and were made aware of the availability of the electronic version.

Proposed time course of the project

Semester 1

Week 8	Launch project
Week 9-10	Individual discussions with third year students who will fill out their USRs retrospectively
Week 11-12	Individual discussions with first year students

Semester 2

Week 5-6	Individual discussions with first year students who should have started to complete their USRs
Week 7-8	Individual discussions with third year students. Introduce the concept of mentoring
Week 9-10	Individual discussions between third year and first year students to pass on the perceived benefits of the process and assist with completing first year USRs

Staff involvement

All staff received a memo outlining details of the project, the benefits to students and the role of the teaching staff in providing feedback to students.

- They were asked to be aware that assessments would, in future be used in students' PDPs and therefore to take this into account when providing assessments of work.
- The USR provides a photocopiable page where students may summarise relevant parts of their Skills Audit and present it to an appropriate member of staff to comment on. Staff were therefore asked to familiarise themselves with the USR and what might be required.
- For feedback on practical skills, it was decided that laboratory demonstrators were best placed to give this but it was up to the member of staff in charge of the lab to brief demonstrators.

Implementation

Third year students

There were 30 students in this group. Of these, 22 completed the USR and made time to discuss and evaluate its usefulness.

The main problem encountered was the time course. For the project to develop, the students needed to complete the USR and be confident and familiar with it to be of assistance to first year students. Due to pressures of work towards the end of semester 1, only one student had completed the USR on time and the remaining members of the class asked if they could complete the USRs over the Christmas vacation. However, the reality of the situation was that only when individual appointments were made for the student to come and discuss the USR with myself did the USR get completed.

As a result of this rescheduling, students had one, rather than two individual interviews. The interviews were therefore longer than anticipated (on average 20 minutes), to allow time to discuss and evaluate the USR and develop a strategy for mentoring. These were spread out over several weeks, in order to work round students' other class commitments.

Not all students wished to or would have been appropriate for mentoring but a sizeable number were enthusiastic and willing. This group was coordinated by one student, who liaised directly with myself.

First year students

There were 64 students in this group. The focus was on students whose subject intention was chemistry because it would be easier to support them in future years. A small number, however, had other subject intentions but were keen to participate and continue to use the USR, because of the obvious benefits.

Once interviews with the third year students were underway, interviews with this group were set up. As all first students are in the labs at least once a week, it seemed appropriate to conduct interviews there. Because of knock-on effects of the third year timetable change, a new interview timetable was drawn up for this group.

Semester 2

Week 1	All students received an email reminding them to start filling in their USR and explaining that interviews would commence the following week.
Week 2	The first group of students were seen individually by myself, the benefits of keeping an USR were discussed and a brief explanation given on how best to approach it. They were then asked to fill in the USR over the following week and bring it to the lab, so that any questions and difficulties could be addressed. A reminder slip, with date, was also given.
Week 3	A second group was given an introductory interview, as in week 2. The first group had the opportunity to go through their USR with me and discuss any concerns or uncertainties.
Week 3-5	The third year mentors went into the laboratories at a different time to myself and spoke to the students individually. They emphasised how difficult it was to fill in the USR retrospectively and the advantages there were in completing it from first year.
Week 4-8	The process was repeated until all students had been seen. The interviews became spread out because of absence, forgetfulness and some unwillingness. Some persistence was required but it produced results in the end.

Student mentors

Mentors went into the laboratories in small groups and each spoke to a number of students on an individual basis. The discussion areas they were asked to focus on were:

- general information about the chemistry course, particularly with respect to skills development
- why keeping a record of their skills development has been/would have been useful
- the problems they had filling in the USR retrospectively
- how they filled in the USR, the time involved, sources of evidence for skills.

Any particular points or problems raised were fed back to myself.

Outcomes

First year students

- Some students were reluctant to fill in the USR because of bad experiences at school with the National Record of Achievement (NRA). All in this category admitted that the USR was much more straightforward and less time-consuming than the NRA.
- Many wanted to know if completing the USR was compulsory. As the answer to this was 'no', it was important to reinforce the benefits and to discuss individual situations.

- A number of students were not very confident about filling in the USR. Some just needed reassurance - checking and commenting on their completed record. Others, however, found it difficult to self-assess primarily because 'someone was going to look at it'. This raised the important issue of ownership.
- The student mentoring running in parallel seems to have been beneficial. A number of first year students commented on points raised by the third year students. At the very least, it kept the momentum up and kept USRs high profile throughout the semester.

In summary, the combined mentoring efforts appear to have been worthwhile because as the semester progressed, so did the general acceptance that the USR should be completed. Ultimately, only six students had not completed the USR by the end of semester 2.

Third year students

- Of those who participated, many saw it as an unnecessary exercise at this stage in their studies. Resistance was largely due to an existing awareness of skills development; many students kept a record of their own making and the MChem. Students had recently prepared their CVs for industrial placement applications. The majority felt it would have been much more useful in first and second years.
- A major concern was asking students to carry out mentoring work on top of an already heavy workload. Time was always an issue.
- Mentors required more guidance than was anticipated at the outset. They were comfortable with the system devised. A 'buddy' system would have taken more time and organisation and seemed more appropriate as a longer-term strategy, introduced in semester 1.

In summary, the majority of third year students did not find the introduction of an USR at their stage appropriate, but appreciated it was necessary in order to familiarise themselves with the system, before mentoring the junior class. Although many students were willing to be mentors, some found the timescale of the project too narrow and were concerned about their own workload.

Expanding the project

School of Chemistry

The pilot scheme should make the expansion of PDP within the School of Chemistry relatively straightforward. It has, however, highlighted a number of key issues that need consideration.

- The first year of a student mentoring system is the most difficult because senior students need time to familiarise themselves with all aspects of the adopted scheme. For 2003-04 onwards, there will be groups of senior students familiar with the USR and it should therefore be easier to organise a more rigorous mentoring system.
- The issue of ownership is clearly a difficult one. We have taken the view that the USR is the students' property, for them to complete for their own benefit. By doing this, we may be letting some students down by not giving proper guidance and direction. However, with a good staff/student mentoring team this may not be a problem.

- The student mentors themselves need direction, support and encouragement and to this end, staff involvement will be essential.
- The School of Chemistry offer many opportunities for skills development from first year onwards (creating posters, newspaper articles, presentations, workshops on communications and interview skills), in addition to the skills learned in formal teaching and practical classes. For the junior classes, it is worthwhile highlighting these opportunities and it takes little effort on the part of staff to give appropriate feedback on class and project work, which can be used directly as 'evidence' of skills development. It is therefore important that all members of staff are aware that students are building a personal development file and assist, as appropriate.
- The timing of the introduction of the USR is, at least in the short term, something of a problem. Due to the timing and requirements for different chemistry modules, many students who received the USR in semester 1, did not continue with chemistry in semester 2 and likewise, students doing chemistry in semester 2 did not necessarily receive the USR in semester 1. It is very doubtful if students with non-chemistry intentions will continue to use the USR, until other schools implement PDP. Students with chemistry intentions who are not required to do the semester 1 Foundation course may not receive any mentoring until semester 2.

University-wide

As highlighted above, until the whole university is involved in undergraduate PDP, there will be problems with giving support and feedback to students. However, as the whole university comes on-board, additional problems arise. There will inevitably be many different PDP record books in use throughout the university. With the modular system of courses, decisions on who issues the record books and when, who mentors students, etc will be required to implement the scheme effectively.

Conclusion

PDP has been introduced successfully to the first year chemistry students. Senior students are now familiar with the system and assisted in the mentoring of first year students. A number of these students will be available to help introduce PDP to new students in session 2003-04. However, for the scheme to continue effectively, student mentors will have to be organised and supported and a reasonably high level of staff input will still be needed, at least in the short term. In addition, the second year students will need reminding at some point during the session to fill in their records, with perhaps some mentoring, but hopefully not at the level required in the first year. Direct-entry second year students should also be identified early in semester 1 and introduced to PDP. Until PDP is introduced university-wide, there will be difficulties in implementing an effective and robust system at first year level.

Acknowledgements

Funding from the Fund for Innovation in Learning, Teaching and Assessment is gratefully acknowledged. The help of the students involved, especially those in third year has been essential to the successful implementation of the scheme.

Personal development planning with student volunteers in support work

Chris Lusk, Student Support Services, University of St Andrews

Synopsis

Student volunteers working within a university support service have used Personal Development Plans (PDPs) as an informal assessment of the experience and analysis of the skills they are developing in their work. Students are supervised by a full-time coordinator who assists them in identifying tasks with opportunities for them to develop their skills. However, the students themselves compile the PDP and it is the process of self-reflection and one to one identification of what they have achieved which is seen as the greatest advantage.

Many of the students who volunteer and train for this volunteer group have a long-term target of gaining postgraduate qualifications and training for employment in one of the caring professions. The experience they gain from working within the support network allows them to provide evidence of prior training and experience in this area of work already. The use of the PDP, plus the development of summary information incorporated into a reference from the department on graduation, has allowed past students to gain entry to the courses of their choice.

The disadvantage of the PDP process is the difficulty of selling the concept to students who remain unconvinced of its worth and are therefore reluctant to participate if it is not mandatory. It also requires a drive from staff convinced of its worth and can easily slip into non-usage if staff are stretched elsewhere in their work.

Context

The University of St Andrews has an integrated central unit for support for students, The Student Support Service. A team of 25 professional and administrative staff serve the 6,500 student population in a geographical area with limited welfare, counselling and health resources provided in the local hinterland. The service therefore aims to provide an all-round service on a range of issues and covering emotional and practical support to students themselves and pastoral care staff, 24 hours a day, seven days a week during term time.

Assisting this team of staff is a team of 30-40 trained student volunteers self-named the Supnet (support network). These students are selected (from application through voluntary training and interviews) from a substantial number of applicants and are then trained over the first year in aspects of social and/or welfare skills (eg anti-discriminatory practice, confidentiality, boundaries) and in specific areas of support work, eg eating disorders, ME, health awareness, depression etc. Some of this training is provided in-house, some of it involves sending the 'Supnetter' to national services. All training subsequent to selection is mandatory.

Supnet activities

Once trained, the Supnetter operates for the department for the remaining three years. They run a variety of activities for students, from leading teams of volunteers

for orientation, to providing first aid cover at student balls and social events. They offer one to one support for students with disabilities, students experiencing emotional crises and they organise support self help groups for students suffering from conditions such as ME, depression and eating disorders or with unexpected situations arising such as bereavement or an unplanned pregnancy.

All Supnetters attend mandatory training as described, attend group and single supervision sessions and are debriefed after each round of activities by a full-time Supnet Co-ordinator - a full time member of professional staff who acts as a liaison for referrals from other members of Student Support staff or University pastoral staff.

History

In 1995, the Supnet was established to offer some extra peer support for students. By 1997, the numbers had risen to 25, students joined intermittently and took ad hoc training as it was possible to timetable within their personal diaries and that of the staff within the unit. Commitment was intermittent with student attendance unreliable. By 1997, the powerful contribution which had been demonstrated by some committed Supnetters had demonstrated a role here which was thought to be worth developing. However, in order to encourage worth in the project a new beginning was planned.

In 1997, a clean sweep allowed the project to start from scratch again. A full time coordinator was hired with funding part-time from Lloyds TSB Trust and matched by the University central funds. Recruitment criteria were updated and strengthened with unit demands on time commitment heightened. Students were asked to come for training (voluntary at that stage) and were offered a series of interviews. The interviewers stuck rigidly to the criteria decided and turned away people who did not have the required approach or experience. Academic references and character references were required. The numbers took some time to build up but by 1998 we had 30 Supnetters appointed.

These Supnetters have demonstrated the worth of the project by their commitment since and work, on average, 10 hours per week for the unit for the three years following their recruitment. In return, they have each been offered a PDP which they can fill in themselves throughout their training and three years of working within the group. The Supnet Co-ordinator helps them maintain their PDP and when gaps in skill development are identified, opportunities to address this are created.

The PDP has offered evidence to employers or future educational institutions on the skills obtained by the student. The personal reference which refers to extracts of the PDP has proved its worth and to date all students using this system have obtained entry to their desired course of action.

The 10 PDP key skills (Mason, Collier et al)

Presentation skills

The ability to give a structured presentation to an audience utilising effectively audio/visual aids and successfully demonstrating the ability to build up a rapport with an audience.

Analytical skills

The ability to collect, collate, analyse, adapt and classify data and to be able to use your results effectively.

Creative thinking

The ability to develop strategies to solve complex problems requiring initiative, imagination and flexibility.

Teamwork

The ability to work with others effectively; to exchange ideas as well as giving and receiving feedback.

Time management

The ability to keep to schedules, to structure your own time and to prioritise your workload. The ability to complete work to a deadline.

Communication (written and verbal)

The ability to express ideas and be understood through a variety of communication media, including public speaking, talking in small groups or one to one, presentations, letter writing, reports and telephone.

Leadership

The ability to organise, motivate and lead others, to take decisions and to listen to all relevant opinions before reaching a decision. The ability to accept and handle responsibility well. The leader effectively pulls a team together to give it direction and purpose. A good leader enables the group to work through differences and become high performing, well able to do more work than a group of individuals on their own.

Interpersonal skills

The ability to listen and react to the needs of others. The ability to initiate relationships and to build a rapport with a variety of people.

Practical skills

The ability to operate machinery safely, to be computer and numerically literate as well as showing competence in managing own financial affairs.

Self reflection

Last but by no means least, the ability to reflect on your experiences and learn from them is a skill which will benefit you greatly throughout your life.

Process

The student, having undertaken an activity, will fill in a Skill Development Sheet.

Skill Development Sheet - example (Collier 2000)

Project Eating Disorder group coordination **Date** 6 April 2004

Main purpose/specific challenge ensuring safe space for participants to meet, raise concerns and offer mutual support to each other.

Main activities engaged in organising the EDA group throughout event, publicity, room booking, contacting student referrals by phone, organising speaker (contacting national network), teas/coffees, summary handouts by Counsellors, feedback forms, presenting speaker on night.

Time commitment 9 hours over 3 weeks. 3 hours on the night. 2 hours post event.

Skills involved Leadership, creative thinking, teamwork, time management, communication, practical skills.

Evaluation of success

Communication - I don't like presenting. Presentation skills need work - I don't like it because I don't think I'm good at it. I waffle - I stammer. I also prefer to phone than do face to face when dealing with stroppy hall-keepers.

Feedback said that event was a success - only small numbers. but quality was good and some of our students who were in most distress last week really said they gained from it. Time management was magnificent! The talks went off to the second!

At the end of each month, the Supnetter will go for a debriefing session with the Supnet Co-ordinator. At this session they will discuss the activities they have undertaken. They will use the PDP as a guide for identifying key skills developed and will update the Skills Development Sheet. They will also identify areas where development is required or where strategies to get round a problem have been identified.

All the development sheets are summarised each year into an annual Skills Summary itemising time commitment, areas of strength and weakness, areas requiring development.

At the end of the Supnetter's time, the three Skills Summaries will be used by the Director of the Unit to draw together a personal reference for the Supnetter for further training or employment.

The reference

The reference, at the end of the day, provides detailed information on the Supnetter's involvement in the unit. The length of commitment (including hours) is documented; training completed satisfactorily (eg counselling skills, mental health issues, drug and alcohol awareness.) Examples of activity are included (organising welfare publicity,

managing small self-help group for depression, helping run our first year orientation events and is a valued team leader in our crisis management team providing first aid at balls).

Personal recommendations noted in the PDP might be included.

> Her commitment and trustworthy approach have been commented upon by professional contacts (Warden and local GP).
>
> As part of her learning while carrying out these tasks, [Name] has completed a PDP, a copy of which she will supply to you upon request. This identifies the categories of areas, which she has given thought to developing through her time with us.
>
> By examining this PDP, you will note that [Name] has experience in giving structured presentations in public, organising team members and leading projects with creativity and motivation. She has used self-reflection on a continuous basis to analyse her motives, aims and targets, and she has had to account to us for her decisions - with much success.
>
> The timing of the dates in the PDP will emphasise her ability to keep to deadlines with serious commitment once targets are identified.

Not all is perfect - the impression is given of a realistic self-assessment of capabilities.

> The self-reflective element of the PDP has identified some areas where [Name] would wish to develop further, eg when making public presentations, her verbal and written work is excellent but she requires further experience in the use of technological visual aids such as Microsoft PowerPoint. [Name's] enthusiasm and adaptability would make the opportunity to learn the only requirement here.

Advantages

The obvious advantages are identifying skills in a focused format allowing for a concise assessment by future agencies/employers.

At the time, however, there is a great advantage in the process. These students are developing self-reflective skills which are difficult to replicate in training. They are also supported by their focused once a month debriefing session. Feedback from the students was that they found the one to one session on the PDP with their Supnet Co-ordinator a hugely satisfying experience with them identifying their gains and crediting themselves with work achieved. The confidence they gained from this was described by them as invaluable.

Difficulties

Although ultimately rewarding, this scheme is intensive in terms of staff time and commitment to manage. If a member of staff coordinating it is not 100 per cent behind the project, the momentum required to maintain it slows.

The advantages of the PDP tend to be appreciated retrospectively. At the time, and certainly in advance, it is difficult to convince students that there is anything to be

gained from this. Some students with prior school experience of Record of Achievement were reluctant to re-enter the process having found the first disillusioning.

Since 2000, the numbers of students participating in Supnet has increased and, through word of mouth, students see it as a challenging and exciting group of people to join. However, the numbers of students filling in a PDP has declined since we made it voluntary. As students develop and become more involved emotionally with the work they occasionally neglect, forget or deliberately omit to complete the recording process, favouring instead, to experience the moment and grow and develop from that alone.

At this point in time we have very few students filling in a PDP - a fact rued by the ex-Supnetters who visit and try to convince the new ones that there 'really is so much to be gained' in such a process. When it worked best at its peak, the PDP system was supported by three members of staff - as part of their roles. This, unfortunately, would allow some to see it as resource intensive and a luxury.

In defence of its operation, at its peak, some students were attracted to join Supnet on account of the opportunity to complete a PDP. This, it could be argued, has some disadvantages in selectors ensuring the recruited have an appropriate motivation. However, it was accepted within the staff team that the additional use of the PDP in reference format, with links to the original document, was a valuable tool for ensuring entry into future work or the next stage of training while at the same time developing the self analytical skills required to offer added protection that the team members were developing professional conduct in their work.

References

Collier S (2000) *Guidelines on Preparing a PDP*, University of St Andrews, St Andrews

Mason C, Collier S and Baxter C, *Personal Development Planning with Support Network Team Volunteers*, University of St Andrews
www.recording achievement.org/Case.Studies/cs_detail.asp?sid=69

Assessment of personal transferable skills - Post workshop report

Dr Colin Mason, Director of Learning and Teaching Development, University of St Andrews

Introduction

Colleagues were reminded that the workshop was designed to explore theoretical frameworks underpinning the very nature of transferable skills and how these were being assessed in different contexts in different institutions across the UK. Further, views expressed were likely to be contradictory but if the Scottish sector was to derive useful lessons in an endeavour to develop a strategic vision of innovative approaches to assessment in higher education then we should be open to such diversity and welcome challenges to our own assumptions and prejudices. The whole group were reminded that this was particularly important in working in the four breakout sessions during the afternoon, after the presentation of all keynote and case study presentations.

Keynote and case study presentations

Assessment of personal transferable skills - models and approaches

Peter Knight addressed the title of this conference head on by querying the very nature of the key words involved in the phrase (assessment of) personal transferable skills, with which he feels there is some inherent ambiguity. He explored in some detail the nature of 'skills' and, in his view, the more useful set of attributes that underlies developing students employability. The two significant challenges posed by Peter's thesis were:

- the assessment of complex, high order intellectual capacities (analysis, evaluation, synthesis) and employability skills, in which validity and consistency or reliability are sometimes trade-offs, necessitates a major shift towards predominantly formative assessment of such activities designed to foster their development in students and
- that there is a need to re-examine programme-level assessment, moving away from unconnected module-based, predominantly summative assessment practices to more integrated developmental and strategic, formative assessment practices.

These two ideas alone (among many others) provided the workshop with a provocative, engaging challenge for all participants, and indeed none more so than future speakers presenting either keynote addresses or case histories.

Graham Nicholson presented the case for an explicitly-identified credit-bearing module that facilitated the development of personal transferable skills. The case for making explicit the learning outcomes and giving the unit of study credit weighting were emphasised as a key motivational tool to ensure students engaged. A model developed at the University of Stirling (Career Planning module) and upon which he University of Dundee would be building comprised credit-weighted assessment of various components (job study, CV, application, presentation, interview, action plan).

Jean Gowans and Colin Mason presented a case for an alternative approach in which transferable skills are assessed informally, and do not carry credit. Jean Gowans of the Careers Advisory Service at the University of St Andrews, described an informally assessed, Career Development programme that is attended voluntarily by students, who are awarded a certificate, approved by PriceWaterhouse. Colin Mason presented a further case study involving contract research staff at the University of St Andrews who attend a one-day career development course and are engaged in making both pre and post course self-assessments of a range of their own transferable skills. Statistically significant shifts in scores awarded indicate development in a combination of both their awareness of and actual skills, particularly in articulating their (self-judged, albeit) strengths and weaknesses at interview.

Alternate ways of representing learning, including personal development planning

Rob Ward, Director of the Centre for Recording Achievement (CRA), provided a summary of how seeming incompatibilities of the higher education summative assessment system for representing academic achievements and personal learning can be circumnavigated. He got everyone thinking by presenting a scenario where all records of previous achievements for everyone at the workshop had been lost/destroyed and we were asked to present our own case for future employment. The introduction of progress files, that comprise both a transcript, preferably that conforms to an agreed presentational format, and a way for students to represent their own personal, academic and career development through a personal and professional development planning (PDP) process, has been proposed to offer a representational route that is both more informative and accurate. Information was presented about how the CRA is working with the new Higher Education Academy to promote this, particularly focusing on the potential for recording data using e-portfolios.

Professor David Cole-Hamilton presented the case for the assessment of transferable skills through an integrated approach within the chemistry curriculum at the University of St Andrews. Further, he summarised an ongoing pilot project in the School of Chemistry for implementing both tutor and student mentor-supported personal development planning using both paper-based and electronic resources devised by the Royal Society of Chemistry. He, Fiona Gray and other colleagues are using a tailored system with first year students and third and fourth year student mentors.

Another example of engaging students in identifying transferable skills and the PDP process was provided by the final case study presented by Chris Lusk, Director of Student Support Services at the University of St Andrews. An informal route for acquiring transferable skills is afforded by students who volunteer as part of a network of peer support for their colleagues and engage with personal development planning as a means of recording their experiences. These activities are completely extra-curricular and are both self-assessed and then ultimately tutor assessed informally when provision of a reference is made for students wishing to pursue further related employment. This reference is informed by student-authorised consultation with PDP records.

Breakout discussion groups

The afternoon breakout groups provided an opportunity for a lively and wide ranging discussion and exchange of views and ideas between representatives from a range of institutions and different disciplines. The notes from each breakout discussion group are presented in slightly edited format and an overall summary is presented at the end.

Assessing personal transferable skills

It was thought imperative that personal transferable skills should be made explicit and embedded in the curriculum and be made clear to students. There is a need for unambiguous language. Students need to be able to translate in terms of their self-presentation in context and a frame of reference. It was also felt that this should be a process throughout a degree programme and should not be seen by students as a single moment or event or as 'another' assessment. The group agreed that the development of transferable skills is a process that students must engage with and they must understand the importance of their own skills, why critical reflection is important and the notion of professionalism. There is a need to change the culture typified by the just 'what have I to do to pass' mentality. Inevitably, this turned the debate to the quality of learning and formative feedback. Group members outlined examples of good practice. Many of these were small things that lecturers could build into their everyday teaching for example a short structured reflection exercise at the end of a class. Flexibility must be built in to enable appropriate skills and attributes to be developed. The student is central and must take ownership of the process. This is not something that can be imposed on the student but must be driven by their understanding and hence need for and value of recording or being explicitly aware of their strengths and weaknesses. Students need to know how and what tools and resources they can use to make developmental changes. The content must be relevant to both their stage and the programme of study.

Discipline differences

Employability skills are inherent in some disciplines. One example given by a delegate was in design, where it was thought best to draw attention to these skills but not reasonable to assess them separately. However, it was strongly agreed that all institutions should be allowed to develop appropriate routes or pathways and tools for their particular students.

Assessment - Overall it was agreed this was not essential but that some strategic thinking was necessary for determining what might be the 'carrot' for students to sit down and record their skills and if this should be compulsory, but not graded, to ensure that it is done. Further, student numbers gave rise to issues and problems.

Resources - it was agreed that staff had a role to be the catalyst and facilitator. It was noted that some staff may need further support and development themselves to appreciate the importance of their role in this process. Learning from models in different disciplines particularly in vocational and non-vocational programmes might be very useful and exemplars of good practice should be shared.

Conclusion

It is most important to facilitate individuality within different universities and that a 'one size fits all' approach was not appropriate. However, the exchange of ideas and good practice across institutions as well as interdisciplinary and faculty discussion were valuable. Engaging the student was imperative and that the process of the development of their skills and attributes should be ongoing as an integral part of their degree programme.

Institutional approaches to PDP

- Institutions do not currently have policies for PDP, however there has been recognition that a 'one size fits all' will not be appropriate.
- Many professional/vocational disciplines have been implementing PDP for some time. Institutional approaches need to be fully inclusive and should not generate further work and/or replication.
- Where possible institutions are seeking to enhance existing programme/school level activities rather than developing entirely new approaches to incorporate PDP.
- Portfolio generation (owned by the student), supported by personal tutors and academic programme elements appears to be the overall approach, frequently in the form of an online portfolio supported in an institutional virtual learning environment.
- Student mentoring schemes are perceived to add value to PDP, making good use of an under-utilised resource and building upon the strengths of peer tutoring.

Issues

- The need to provide equal opportunities for engaging with PDP are of concern since many models could be, potentially, discriminatory and/or exclusive. Will all elements of schemes be available to all students? When will this be made available and in what format? Students' lives outside university may restrict their ability to engage with co-curricular activities that frequently contribute to PDP eg volunteering. How will this be catered for?
- PDP will be a voluntary activity hence there may be some issues relating to promoting student engagement. It will not always be appropriate or possible to provide credit for such activity within programmes. How will the status/value of this activity be promoted?
- Students' reflective abilities do not necessarily translate into their practice, which is particularly relevant for some disciplines.
- PDP is frequently perceived to be a further burden for staff.
- The value of PDP may be undermined if it is seen as an 'add-on' rather than an integral part of a student's university experience.
- The issue of whether PDP relates to employability or more broadly to 'graduateness' needs to be clarified.

Approaches

- Making PDP an inherent part of academic programmes (in the same way that transferable skills are now built into programmes).
- Build on existing good practice, making PDP elements of programmes more explicit.
- Provide support for, and opportunities for practice of, reflective capabilities.
- Place emphasis on the developmental element of PDP rather than the personal in order to generate 'buy-in' from academic staff.
- Partnership approaches are required in order to make best use of available expertise within support units and academic departments. Importantly, PDP should not be perceived as the sole domain of careers.
- It will be important to provide a number of alternative approaches to facilitating and supporting PDP, in order to meet the needs of a diverse student body.
- There is a need to link reflection to feedforward, preventing it from becoming a circular activity that looks backwards.
- The value of PDP needs to be made clear to both staff and students - this may become more applicable to staff as they are required to undertake PDP for gaining and retaining good standing with a professional body for learning and teaching (Higher Education Academy).
- There is a need to tap into the co-curriculum (moving away from perceptions that this constitutes extra-curricular activity). Institutions may consider offering a broader range of opportunities. There will be a need to facilitate student engagement with the university community, however, particularly in light of recent research which indicates that widening participation students do not engage with the co-curriculum as extensively as more traditional students.
- PDP needs to be portrayed as the 'glue' which links together experiences gained within various contexts, enabling them to be built upon and translated to new contexts (as required, for example, by employers).
- PDP is a lifelong activity and needs to be linked into the lifelong learning agenda.
- If PDP is to be effective in supporting the development of student employability then universities will need to consider how their values relate to those of employers, in order to ensure that students value elements of their university experience that will benefit them in gaining employment.
- In order to enable students to develop their skills of reflection and to encourage them to consider both successes and failures, they need to 'control' access to their reflective accounts. Students should be able to select which elements of their reflective accounts they wish to make public, particularly for the purposes of assessment.
- PDP needs to be fully inclusive of personal, career and academic development, rather than focusing too prescriptively on employability.

How are skills development and/or training opportunities within the curriculum provided in your institution?

The experience of one member of the group on skills assessment in nursing at The Robert Gordon University was described.

- All skill outcomes are now assessed in the workplace and no longer in simulated conditions in the university; formative assessment of skills is still done in university simulations.
- Getting enough mentors for the students is a problem. Training is given.
- Outcomes to be assessed now include more generic ones with specific content outcomes. There was a challenge here (successfully overcome) of getting all staff to accept that.

General issues

Modularity of programmes

- This has led to problems with developing a cohesive assessment strategy for the development and assessment of transferable skills, particularly for programmes that are multidisciplinary and taught by staff members from several departments/schools or even faculties such as those students undertaking joint degrees.
- Modularity also sometimes results in students being assessed several times (over assessment) in one skill and not in another depending on their choice/pattern of modules.

Staff development issues: How do we engage academic staff?

- It is vital to engage all the academic staff involved in the delivery of PDP and assessing personal transferable skills. Students are often turned off once some staff are seen to show scepticism.
- Preparation for academics is as important as it is for students. Many academics feel uncomfortable with teaching and assessing transferable skills, partly because they have not been trained or assessed in them themselves and indeed are sometimes unaware that they actually have all these skills. Most will never have done PDP themselves, and many are not used to doing or encouraging reflection and self-assessment.
- Support staff (eg library, information technology, careers, educational development) could work with academics first, and then help in the training and assessment of the students alongside academics.
- Institutions need to provide ways of filling the gaps that academics feel unable to cover themselves.
- Line managers and senior university staff need to be involved, both because of the resource issues but also to encourage everyone's commitment.

Use of students' expertise: we need to tap into this

- Students can learn more from student-led than lecturer-led learning, but this is much easier with small numbers of students.
- Involving students in supporting/teaching/assessing other students often has unexpected and highly beneficial outcomes, particularly in the skills area.

- One member reported on their successful student mentoring project in the Accounting and Business Finance Department at the University of Dundee in which year 3 students, in twos or threes, had mentored groups of six year 1 students for the first six weeks of their first year. The year 3 students had got at least as much out of it as the first years, and were motivated to take part by the need for evidence of skills for their CVs. A paper describing this project is available at http://cbs1.gcal.ac.uk/lts/AFox_LStevenson_PeerMent.htm
- Students often need help in interpreting incidents in their workplace experiences, and using them to learn about and develop their skills.

How do your course and institution use undergraduate students' research projects/dissertations to develop generic skills in the curriculum?

Projects usually involve teaching learning and practice in using the literature, abstracting and synthesising data, the preparation of a substantial written report, teamwork, in many cases the learning of new practical skills, giving at least one oral presentation, and in some cases preparing and presenting a poster.

The discussion revolved around the following themes.

- Whether research projects should be scheduled over one semester (short and fat) or over a whole year (long and thin). Preferences for the latter were expressed in terms of the benefits to students' learning. If the former approach was followed, suggestions were made to schedule the literature review into the previous semester, in order for students to most benefit from the task.
- A university's approach to PDP as a crucial contributor. A fully integrated or holistic approach would be of great assistance - at this stage of their studies, students would then be able to include reflection on their learning, and would be motivated to do it (given that PDP would have been embedded in their reflections from first year). The value of encouragement and mentoring from more senior students were seen as beneficial. Utilising the Effective Lifelong Learning Inventory to assist personal tutoring was suggested (for information, Liz Cullen, Education, University of Glasgow).
- The utility of a scaffolded process in earlier years, aiming towards the research project was described. Staff would then understand the line of development and the part played by other activities. Earlier group tasks and mini-projects were seen as contributors.
- In order to assist students reflect, the following questions (from a reflective practitioner model) are helpful: Why did I choose this approach? How did I go about the task? What have I learnt from doing the project? What problems did I encounter, and how did I overcome these? What would I do differently next time? A number of ideas were given for activities to assist student reflection on the task(s) associated with research projects.
    - i Inclusion of a reflective report with the completed project - similar to that done by work placement students, who find the exercise very valuable - but people expect strong resistance from research-embedded colleagues who might fear that such a report would detract from the final product.

ii An oral presentation, a few weeks before the hand-in date (in order to provide time for feedback and incorporation of ideas).

iii Translation of the results of the project into a readable article for the lay person - possibly linked to a student newspaper or web page.

- The marking of projects, and the need for considerations of the processes of feedback to students. The merits and challenges of supervisor as assessor were raised, and consideration of the supervisor assigning some marks to process, oral presentation, following of safe practices needs to be given.

Overall lessons learnt from breakout discussion groups

The development of personal transferable skills including PDP might best be conceived as integrated processes that facilitate subject-based learning throughout the curriculum and degree programmes. Some form of assessment, perhaps predominantly formative, may be necessary to motivate students' engagement, but also permit student ownership. Whether at institutional or subject level, it is clear that a single approach is unlikely to be satisfactory for the whole higher education sector and for either staff or students. Particularly if the approach adopted is integrated, it will be necessary for the outcomes to be made explicit, permitting strategic design of learning opportunities that neither duplicate effort nor leave large gaps where particular skills, abilities and attitudes have few opportunities for development. Particular opportunities are presented by the challenge of research projects and dissertations, and may provide a focus that all universities highlight, the nature of research-informed or research-led teaching.

The whole process will require a range of different levels of support. For students, this may include their peers as co-learners and as mentors; for staff this may mean further support for themselves directly or assistance from careers service staff, educational developers etc to help support design or delivery of appropriate learning opportunities. Finally, the whole process requires support, encouragement and backing from senior staff, providing commitment to this as significant strategic initiative that fosters the development of students as learners, not just as a measure for accommodating an externally driven agenda for implementing PDP or enhancing students' employability.

Quality Assurance Agency for Higher Education

Southgate House
Southgate Street
Gloucester
GL1 1UB

Tel 01452 557000
Fax 01452 557070
Email comms@qaa.ac.uk
www.qaa.ac.uk

QAA 082 05/05