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)