On May 7th 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 America, 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 summarize the essential themes and recommendations which emerged over the course of the day through references to the papers presented by all involved (available elsewhere on the QAA web-site), 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 HEI student associations conducted by the QAA during February/March 2004.

The first part of the conversation: the speakers’ perspective

“Assessment is a big jig-saw 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 feed forward, 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 his/her 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 break-out 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 the QAA. 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 Break-out groups – Session 1

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

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, whilst 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 as 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 they 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 e.g. placements etc.)
• Enabling students to view their marks individually e.g. through an online portal (rather than posting 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).

“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 – e.g. 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 un-biased results, would be extremely difficult to achieve.

“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 (approximately 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 s/he 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 s/he couldn’t, then the tutor should consult with the student as to what s/he would see as an alternative, and endeavour to provide it

• Alternative methods of presentation should be explored such as video, PowerPoint etc.

The Break-out Groups - Session 2

“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 amongst staff and students.

____________________
“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 an 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 vet school was cited - this particular department have 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 amongst 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 e.g. 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.

"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 adhoc/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 & 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 the QAA 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.”

(Dr. Simon van Heyningen [Chair, QET 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


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