

People within the educational community—policymakers, schools, administrators, teachers and students—use assessments for different purposes. This issue of *CDTL Brief* presents some discussions on the issues surrounding **Assessment**.

Assessing Education Quality: Measures and Processes*

Professor William F. Massy

Professor Emeritus of Education and Business Administration, Stanford University
President, The Jackson Hole Higher Education Group, Inc

The means of assessing the quality of education offered by an institution, faculty, or department has recently become a hot issue in many countries. The three approaches to quality assessment are accreditation, evaluation and academic audit.

Accreditation

Accreditation asks whether academic credits can be trusted and, in the United States, whether an institution's students should be eligible for federal financial aid. America's accreditation system originally addressed the threshold standards needed to assess credit transfer and financial aid eligibility. Does an institution have the resources needed to provide threshold quality levels (e.g. inputs like the quantity and quality of faculty, libraries, physical plant)? Does it comply with accepted criteria for diversity, academic freedom and the like?

U.S. accreditors are redesigning their processes to evaluate and improve quality at all levels as well as to certify threshold standards. For example, the Western Association of Schools and Colleges (WASC) now defines itself as a 'capacity building organization with regulatory functions' where 'capacity' includes the ability to improve and assure education quality. In addition to WASC's use of audit principles, the North Central Association's new system uses criteria from the Malcolm Baldrige National Quality Award and other regional and disciplinary accreditors are experimenting along similar lines.

Evaluation

Evaluation of delivered quality at the subject level by teams of outside experts has been the mainstay in the UK and a number of other European countries. European oversight agencies tend toward direct evaluations of quality by outside experts. Britain calls its evaluations

'subject-level assessments' to emphasise that the focus of attention is the department rather than the institution. The programme began with the polytechnic universities in the late 1980s and was extended to all of the country's 150 or so universities a few years later. (The UK's new programme uses evaluation only on an exception basis, if an institution fails an audit). The Netherlands introduced evaluation at about the same time, Denmark started in 1992, and a number of other countries have implemented or are considering similar programmes.

Education quality evaluation requires that a team of experts in the discipline visit the department and gauge the quality of education actually delivered to students. Self-studies help focus the visit, but team members also review syllabi and examination papers, interview faculty and students, and visit classrooms. The team forms its own conclusions rather than relying mainly on the self-evaluation.

Evaluation provides incentives for improvement, but this is offset by its intrusiveness. One cannot evaluate delivered quality without 'getting inside' the department. Once inside, the outside experts may second-guess professors on what are essentially matters of judgment. This problem is particularly severe when the experts have less research prestige than local faculty. Research prowess doesn't necessarily confer good pedagogical judgment, but experience in the UK, particularly, demonstrates how the lack of a strong research record can undermine an evaluator's credibility.

* Source of material: Massy, William, F. (2003). *Honoring the Trust: Quality and Cost Containment in Higher Education*. Bolton, MA: Anker Publishing Company, Inc. © 2002 by Massy, William F. Permission for use by the National University of Singapore is granted.

Opinions differ on whether the advantages of external quality evaluation outweigh its disadvantages. On the positive side, evaluation provides credible information to further the sponsoring entity's accountability, public reassurance, and market information goals. On the negative side, its intrusiveness undermines the improvement agenda. Direct evaluation of individual subjects also is very expensive, and many commentators question whether site visit teams can produce meaningful evaluations in the time available to them.

Academic Audit

Academic audit, originally developed in the UK, is a central feature of quality oversight in Hong Kong, Sweden, New Zealand, Australia, and (again, since mid-2002) the UK. In the United States, audit is being tested by WASC (the accreditor for west-coast universities) and the University of Missouri.

Academic audit assesses education quality processes rather than education quality itself. It asks whether a unit is performing the activities necessary to produce, assure, and regularly improve quality. Therefore, audit is less intrusive and less expensive than quality evaluation.

By not disempowering professors, audit spurs improvement as well as generates information. It elicits structured conversations about education quality processes: first within the unit and then with the site visit team. Conversations are important because that is how actions get launched in academe. Structure is important because it focuses attention on the following five process domains that are important for quality.

a) *Determination of Desired Learning Outcomes*

What should a student who successfully completes the course or program know and be able to do? How will the course or program build on the student's prior knowledge and capability? How will it contribute to the student's future employment opportunities, capacity to make social contributions, and quality of life?

b) *Design of Curricula*

What will be taught, in what order, and from what perspective? How will the teaching contribute to desired learning outcomes? What course materials will be used? How will these materials relate to other parts of the student's programme?

c) *Design of Teaching and Learning Processes*

How will teaching and learning be organised? What methods will be used for first exposure to material, for answering questions and providing interpretation, for stimulating involvement, and for providing feedback on student work? What roles and responsibilities will the faculty need to assume? What

other resources will be required and how will they be used?

d) *Student Learning Assessment*

What measures and indicators will be used to assess student learning? Will they compare performance at the beginning and end of the term, or simply look at the end result? How will the long-term outcomes of the educational experience be determined? Will baseline and trend information be available? Who will be responsible for assessment? How will the assessment results be used?

e) *Implementation Quality Assurance*

How will faculty assure themselves and others that content is delivered as intended, teaching and learning processes are being implemented consistently, and assessments are performed as planned and their results used effectively?

Deans, provosts, and external oversight bodies can conduct academic audits for units within their jurisdiction. The process is flexible and inexpensive, and it can spur improvement as well as produce oversight. Professors can participate as auditors outside their disciplines, which helps spread exemplary practices. Those who do so are pleased with the process—a feeling that is shared by professors in the units being audited. ■

Prof Massy is a specialist in the economic and leadership issues confronting today's colleges and universities. In the 1970s and 1980s, he held senior administrative positions at Stanford University, and in the 1990s, he was a Senior Scholar at Stanford's National Center for Postsecondary Improvement. His research focuses on academic quality assurance and improvement, performance indicators, faculty roles and responsibilities, cost analysis, processes for resource allocation, and mathematical modelling of higher education institutions—including a full-scale computer simulation of university behaviour recently released under the title, "Virtual U". His latest book, *Honoring the Trust: Quality and Cost Containment in Higher Education* was published in February 2003 (Anker Publishing Company, Inc).

Addressing Students' Fears about Examinations

Associate Professor Alice Christudason, Department of Real Estate/Associate Director, CDTL

Assistant Professor Anne Magdaline Netto, Department of Building

Background

It was a bright Tuesday morning and we were conducting a scheduled 2-hour Law tutorial. We were already past the middle of the first semester and in about three weeks' time, the students would be facing their semestral examinations.

As the tutorial progressed, it became increasingly clear to us that the students were paying scant attention to the topic before them. From the presentations the sub-groups were making, the students seemed to be confusing fundamental principles and the issues before them. We found this puzzling as at previous tutorials, they had been enthusiastic learners, had done their 'homework', were motivated to contribute and actively offer their views. In addition, the topic under discussion was one of the most fascinating branches of the law.

So perhaps there was something more to (or behind) the sense of gloom that seemed to pervade the class?

We popped a question to the students: "What's wrong?" Silence.

Eventually, the group representative piped up: "We are afraid of the exams." "Yes, that's right", said another. And another, and another and soon about 25 of them were in agreement as the buzz of consensus gained momentum.

"Oh," we said, "so is *that* the problem?" We left the class that morning feeling rather unsettled and unfulfilled. Since 'fear' denoted a negative emotion, as their teachers we realised that we certainly needed to find out more about it; and if possible, we felt we had a duty to address (if not eliminate) their fears. Hopefully, they could then face the examinations with confidence and this would in turn enhance their learning experience and increase their interest in the subject.

This is what led to our preliminary survey of the issue.

The Fear Factor

Fear has been said to arise due to the 'anticipation of pain'; it has been described as the "*hypothetical state of the brain or neuro-endocrine system arising under certain conditions and eventuating in certain forms of behaviour*" (Gray, 1987; our emphasis in bold). Adapting the foregoing definition to our students' behavioural manifestations, we set out to identify the

'conditions' under which fear arose and 'the forms of behaviour' it gave rise to.

- *Arising under certain conditions*

The conditions under which their fears arose were 'examination' conditions. Performing under examination conditions obviously requires an assimilation of various skills. Typically, students are required to read, understand, analyse questions, apply their knowledge to issues/problems raised, and then present an organised answer to those questions (Messick, 1999). That in itself may not have been a problem, except that all that had to be done *within a limited time!*

- *Forms of behaviour*

It is well documented that fear leads one to freeze, fight or flight (Gray, 1987). In the case of our students, we certainly had evidence of 'freeze' and 'flight'; however, we were curious about how they sought to 'fight' their fear. The answer to this may be gleaned from their responses considered below.

The Questionnaire

We devised a very simple questionnaire (to start off with) and sent it out to about 200 students in October 2002. Responses were received via email. The following questions were posed:

1. Do you fear examinations?
2. What are the reasons for your answer in (1)?
3. Do you have the same fear about assignments?
4. How can your fears be allayed?
5. What are the effects of fear of examinations?

Students' Responses

While it is not possible here to go into specifics, inferences will be drawn and comments made about their responses. As expected, most of the students answered 'Yes' to Question 1 and 'No' to Question 3. However, we will consider some of the responses to Questions 2 and 4 here.

- *Question 2: What are the reasons for your answer in (1)?*
 - Fear of failure
 - Language problem

- Fear of ‘black-out’
- Unable to understand/answer questions
- Unable to obtain good grades
- *Question 4: How can your fears be allayed?*

We found that the responses could be divided into two main categories: (a) where the respondents saw that the *teacher* had a more significant role to play; and (b) where the *student* should assume greater responsibility. Also as expected, (a) had a bigger majority. To elaborate, respondents felt that teachers could assist to allay their fears by:

- Focusing more on examination-type questions
- Explaining and revising important concepts
- Narrowing down the syllabus
- Providing clearer outlines
- Having more Question & Answer sessions

It is noted (with disappointment but not surprise) that from the students’ viewpoint, it appears that getting past the examination is their primary concern. As for (2) above, students also recognised their own role in addressing their own fear of examinations (Clarkson, 1994).

Some Offbeat Responses

There was a small category of students whose response to (1) was ‘No’. These were some of the reasons they provided for *NOT* having a fear of examinations:

- They had faced worse trials before (e.g. a major illness).
- They were more afraid of the results (!)
- Examinations were not the be-all and the end-all of university education.

Surprise, Surprise

We were pleasantly surprised with their responses to Question 5. We suspect that the students themselves were pleasantly surprised. A large majority of students pointed out that they recognised that there were *positive* effects to their fear of examinations. To summarise, they felt that it ‘compelled’ them to prepare and gain mastery over the subject, manage their time better and be conscientious and focused. This is consistent with established theories on so-called “aggressive test takers” (Sherman, 1982).

Other Issues

We realise, of course, that there are a host of other variables and dimensions that need to be taken into account in order to gain a better and deeper understanding of an issue as complex as this. Some of these include matters such as how students’ prior knowledge, perceptions, experiences and intelligences, the role which external stimuli (other than their teacher) affect their inclination and performance in a subject. In other words, there is the particular chicken and egg problem posed to psychologists—which came first, the stimuli or the response? (Gray, 1987) Despite that, through this preliminary survey, we obtained a better insight into the minds of students and gained a better understanding of how we could help them in this joyous yet arduous journey of learning, discovery and “getting an education”.

References

- Gray, Jeffrey Alan. (1987). *The Psychology of Fear and Stress* (2nd ed.). Cambridge: Cambridge University Press.
- Sherman, T.M. & Wildman, T.M. (1982). *Proven Strategies for Successful Test Taking*. Columbus: Charles E. Merrill Publishing.
- Messick, S.J. (Ed.). (1999). *Assessment in Higher Education, Issues of Access, Quality, Student Development, and Public Policy*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Clarkson, P. (1999). *The Achilles Syndrome Overcoming the Secret Fear of Failure*. Shaftesbury, Dorset: Element. ■

Student Assessment in Problem-based Learning: A Challenge Beyond Reliability and Validity

Associate Professor Tan Chay Hoon, Department of Pharmacology & Medical Education Unit

Professor Matthew C.E. Gwee, Department of Pharmacology & Medical Education Unit/
Associate Director, CDTL

“Elaborating an assessment plan that respects PBL principles, is reliable and valid, and has no negative steering effect remains a challenging task” (Nendaz & Tekian, 1999).

Introduction

Problem-based learning (PBL) is an innovative pathway to learning in medical education. The PBL curriculum

provides students with a more holistic approach to medical education that goes beyond just content knowledge acquisition. Value-added outcomes in PBL include the acquisition of educational process/life skills such as:

- Critical thinking;
- Problem-solving and clinical reasoning skills;
- Self-directed learning;
- Interpersonal/social skills; and
- Group/team-work skills.

Student assessment in PBL should therefore ensure “a match between assessment procedures used and the curricular tenets of PBL” (Nendaz & Tekian, 1999). An overview of some of the more useful current strategies and test instruments used will be discussed.

Multiple-choice Questions (MCQs): Need for Testing Knowledge Acquisition

“I do not believe that there is any real evidence to support the claim for problem-solving skills independent of knowledge. [The] evidence from the last two decades of research overwhelmingly argues against this premise” (Norman, 1997).

One of the major aims of PBL is to improve the problem-solving skills of students that can be applied to a wide variety of clinical situations. As problem solving is dependent on sound knowledge, it is therefore imperative that assessment strategies must also test for knowledge. From a psychometric perspective then, how best can we assess knowledge gained from a PBL environment? Norman (1997) clearly states that he is “unequivocally on the side of multiple-choice questions”, on the presumption that the design construction of the MCQ test instrument includes “a rich clinical stem, [that]...involve higher order skills, and hence... more discriminating”. MCQs offer the advantage of high consistency and reliability as it allows for sampling of broad content areas, as well as high validity if appropriately constructed.

The Progress Test (PT)

“[Progress test]...reflects the end objectives of the curriculum and samples knowledge across all disciplines and content areas in medicine” (Nendaz & Tekian, 1999).

The primary objective of PT is to overcome the potential negative steering effect often associated with summative examinations. The PT consists of items drawn from all areas of medicine and is administered several times in a year.

There is convincing evidence that the PT does not have a negative steering effect on student learning “either at

the level of individual learning approaches, learning style (memorisation vs. concept learning) or tutorial function. [PT]...is a reliable measure of student knowledge acquisition, with test re-test reliability...of the order of 0.6–0.7. [Besides, PT]...also has predictive validity, demonstrated by a correlation of about 0.6...” (Blake, *et al.*, 1996; Norman, 1997). PT also has high validity with high correlation ($r = 0.93$) in respect of testing clinical reasoning skills (Boshuizen, *et al.*; 1997). Thus, the PT provides a valuable test instrument to monitor the learning progress of students and can be used for both formative and summative assessments.

Process-oriented Test Strategies (POTS)

The acquisition of process skills by students is also an important educational objective in PBL. Several strategies have been used to assess one or more process skills. Thus, process-oriented test strategies are an essential component in the overall PBL assessment as they can “provide a positive steering effect on learning and useful education”. (Nendaz & Tekian, 1999). However, on psychometric grounds, POTS are generally considered to be less rigorous than the more outcome-oriented tests and not recommended for use in isolation for summative decision-making.

• Tutor, Peer and Self-assessment

“Tutor, peer and self-assessment develop the ability to give and receive feedback and to appraise one’s own needs, which are required in the daily activities of the physician. They also allow detection of potential interpersonal problems that would have remained unnoticed otherwise” (Nendaz & Tekian, 1999).

Tutor, peer and self-assessments are aimed primarily at assessing students’ ability to give and receive appropriate feedback that can be reflected through the following qualities:

- Self-awareness and internal motivation;
- Professional attitudes and behaviours (e.g. mutual trust and respect for and responsibility to each other);
- Critical thinking and self-directed learning skills; and
- Interpersonal, communication and team skills.

The Tutotest, developed by the University of Sheerbrooke to assess the skills and attitudes of medical students working in tutorials in a problem-based curriculum, was found to have high reliability (Cronbach’s coefficient $\alpha = 0.98$); a correlation coefficient of 0.64 was obtained with tutor global assessments and 0.39 with students’ written examinations (Hebert & Bravo, 1996).

- *Four Step Assessment Test (4 SAT)*

The 4 SAT, recently implemented by the University of Queensland, consists of:

- Solving a case scenario individually and in writing through identifying key features, generating hypotheses, explaining symptoms, defining hypotheses from requested additional data, and formulating learning issues;
- Repeating the above processes at the group level with presentation of new information (with observers assessing the tutorial process);
- Undertaking a period of self-directed learning; and
- Taking a written exam testing content knowledge and based on the "top 10" learning issues identified by all groups.

The 4 SAT is aimed at assessing individual knowledge, clinical reasoning and group process skills. Inter-rater agreement was found to be greater than 80% with good correlation ($r = 0.49$) between 4 SAT scores and those from other objective test instruments (Zimitat & Alexander, 1998).

- *Triple-Jump Exercise (TJE)*

The TJE is organised as a structured three-part oral assessment that reflects the learning process in PBL, but under more controlled and standardised conditions. The TJE can be used for formative and summative assessments, but its reliability and validity are generally considered to be low. Inter-rater correlations can be low and errors of measurement resulting from the use of a single case pose serious concerns to the use of the TJE.

Formative Assessments

The use of formative assessments to provide regular, informative and detailed feedback to students on their progress and performance at various stages during a given course, is an essential component of the PBL educational strategy. Such assessments will enable

students, whenever necessary, to undertake effective and timely remedial action that is either self-initiated or upon the advice of the tutor.

Conclusion

"Despite the large range of assessment methodologies used in PBL settings, no single choice emerges, and the triangulation of diverse instruments is required to obtain a fair judgment about students" (Nendaz & Tekian, 1999).

The intended educational outcomes of PBL go beyond just the acquisition of content knowledge. The acquisition of educational process skills, that contribute to the development of clinical competence and desired professional attitudes and behaviours in medical practice, is also an important educational objective of PBL. Student assessment in PBL, from a psychometric standpoint, will continue to pose a challenge to medical educators in the selection of test instruments that can ensure high consistency, reliability and validity to meet the educational demands and curricular tenets of the overall PBL curriculum.

References

- Boshuizen, H.P.; van der Vleuten C.P.; Schmidt H.G. & Machiels-Bongaerts, M. (1997). 'Measuring and Clinical Reasoning Skills in a Problem-Based Curriculum'. *Medical Education*, Vol. 31, pp. 115-121.
- Hebert, R. & Bravo, G. (1996). 'Development and Validation of an Evaluation Instrument for Medical Students in Tutorials'. *Academic Medicine*, Vol. 71, pp. 488-494.
- Nendaz, M.R. & Tekian, A. (1999). 'Assessment in Problem-Based Learning Medical Schools: A Literature Review'. *Teaching and Learning in Medicine*, Vol. 11, pp. 232-243.
- Blake, J. M.; Norman, G.R.; Keane, G.R.; Mueller, C.B.; Cunningham, J.P.W. & Didyk, N. (1996). 'Introducing Progress Testing in McMaster University's Problem Based Curriculum: Psychometric Properties and Effect on Learning'. *Academic Medicine*, Vol. 71, pp. 1002-1007.
- Norman, G.R. (1997). 'Assessment in Problem-Based Learning'. In *The Challenge of Problem-Based Learning*, Boud, D. & Feletti, G. I. (Eds.). London: Kogan Page Ltd. pp. 263-268.
- Zimitat, C. & Alexander, H. (1998). 'A New Assessment Instrument for Large Classes in PBL Curricula'. Abstract from 8th *Ottawa International Conference on Medical Education*, Philadelphia, PA: National Board of Medical Examiners. ■

Implementing Effective Peer Assessment

Lim Yuen Lie, Lisa-Angelique
Research Assistant, CDTL

Peer assessment refers to students' critical evaluations of peers' performance, whether for writing, oral or visual presentations. Peers may be evaluated in terms of their contribution to the group, their product, or both. When effectively implemented, peer assessment promotes critical thinking and learner autonomy (Race, 2001; Zariski, 1996), both desired characteristics of life-long learners which Singapore aims to cultivate in its

graduates (Lim & Chan, 2000; Poh, 1999). Please refer to Race (2001) and Dochy, *et al.* (1999) for more thorough reviews on the topic.

This article highlights possible difficulties in implementing peer assessment and suggests practical solutions to facilitate effective peer assessment.

Setting up Peer Assessment

To promote effective peer assessment, several issues must be addressed:

1. Validity & Reliability

Some problems of validity and reliability include:

- a) Peer over-marking, where peers tend to give higher marks than would tutors (Falchikov, 2002; MacKenzie, 2000; Roach, 1999);
- b) Too wide a range of marks such that tutors have to moderate the marks for the whole class (Bostock, 2000); and
- c) Too narrow a range of marks, making it difficult to differentiate between good, average and weak performers (Cheng & Warren, 1999; MacKenzie, 2000; Zariski, 1996).

The problems listed above are attributable to students' inexperience and lack of confidence in marking (Hanrahan & Isaacs, 2001). Practising and exposing students to the peer marking procedure regularly can improve its validity and reliability. Time-conscious educators could model the marking process first, highlighting their rationale along the way, so that students can understand the thought processes behind the marking and apply them to their own peer-evaluations.

2. Development & Use of Criteria

Related to the issues of validity and reliability, unsuitable or misused criteria can also invalidate assessment. These problems more likely occur when criteria are simply given to students beforehand; if students themselves do not consider what is important in grading a piece of work, they tend to use the criteria less thoughtfully (Bostock, 2000). Even mere discussion of the criteria does not seem as effective as getting students to develop them independently (Zariski, 1996).

There are two suggested solutions:

- a) Get students to first work out suitable criteria for assessment, and train them to apply the benchmarks through analysing and discussing their own answers to sample questions (Stanton, 1999), or the work of their predecessors (Smith, Cooper, & Lancaster, 2002). Students are compelled to exercise their critical thinking skills by generating criteria.
- b) Clear and objective criteria will enable students to mark their peers' work more confidently (Purchase, 2000). Race (2001:Appendix 1) suggests that students reduce obscurity of assessment criteria by employing simple language, and formulating them in terms of checklist questions. For instance, students could re-phrase

the criterion, 'Provides clarity in illustrations', as 'How clearly were illustrations given?', to increase precision in peer-judgments.

Making students develop their own criteria helps them analyse and think critically as they methodically assess and evaluate a piece of work to determine good and bad qualities. Criteria setting is an essential stage in the peer assessment process if the benefits of critical thinking are to be maximised.

3. Level of Formality

Level of formality refers to the accountability of the assessment done by peers: to what extent and how peer-given marks would be included in students' final course results.

a) *Formative, not summative*

Formative assessment focused on suggesting improvements is preferred to summative assessment that is done solely for computing final marks. This aspect also increases student acceptance of this assessment process, as the direct benefits are concrete (Falchikov, 2002). Not only would students have a chance to improve on their work before submission for a final grade, tutors also benefit from receiving work that has been fine-tuned, removing some of the tedium in marking.

b) *Negotiate, not negate*

Tutors should not override peer marks but strive to use peer marking as an opportunity for negotiating differing peer opinions. There would probably be concerns over the implications of 'significant outliers' in cases where a peer-given mark is arbitrated because it differed significantly from the average of the marks given by other groups (Purchase, 2000), or from the tutor-given mark (Lim & Chan, 1999). However, it is interesting that marks given by peers are mostly consistent with the tutor's marks. Lim and Chan (1999) attribute this to the fact that criteria were discussed and made explicit to students, which again highlights the importance of students' active engagement in generating marking criteria and understanding how to apply them in critical evaluations. Should significant outliers appear, students should discuss with the tutor and the class why they choose to allocate a certain mark. Such discussions would allow tutors to give useful feedback on students' thought processes.

The tutor's involvement as arbitrator instils in students a sense of accountability in peer assessment, with the assurance that the tutor will be there to ensure fairness of given marks. More importantly, by having to account for the marks given, students are again

engaged in critical thinking and learning to verbalise their arguments.

c) *Contribution, not content*

Since fellow group members are probably in the best position to judge individual contributions, peers could assess each other's contributions to the group, leaving the assessment of product content to the tutor. The individual then gets an overall mark based on some weighted combination of both marks (Crockett & Peter 2002; Crowe & Pemberton, 2000). Weighting procedures can also counter different peer-marker scenarios, for example, those who are overgenerous, and those who 'conspire' to penalise a member (Li, 2001).

To conclude this section, given that the issue of marks is a sensitive one, it is best for tutors to still be involved in mark determination (Sher 2001), while making students accountable for it. More importantly, both tutors and students should strive to keep their focus on the process of critical evaluation and not on the outcome of the marks.

4. Student Attitudes Towards Peer Assessment

Peer assessment can be met with negative initial responses from students, such as scepticism, lack of confidence, and fears of being discriminated against by peers (Sher, 2001). Students are also known to argue that peer assessment "is too demanding" (Lapham & Webster, 1999), or that assessment is "the tutor's job" (Crockett & Peters, 2002; Crowe & Pemberton, 2000). Such responses might manifest in attitudes of hostility or even refusal to participate in the process (Bostock, 2000; Zariski, 1996)!

Although these are initial reactions which students do eventually get over, it helps to inform students early of the rationale and benefits of peer assessment (Crowe & Pemberton, 2000; Hanrahan & Isaacs, 2001) and its formative aspects (Sher, 2001). Most students appreciate feedback so that they can improve on their work before it is actually graded. Additionally, it is important to make the procedure clear to students (Smith, Cooper & Lancaster, 2002). Communication with students also includes discussing assessment criteria and allowing students to discuss and negotiate peer-given marks. Giving students feedback on their feedback (Hanrahan & Isaacs, 2001) helps keep them on track and increases their confidence in the process. Ultimately, students must become actively engaged in the thinking processes to reap the benefits in assessing each other, so that they will be empowered to carry it out more seriously.

Final Recommendations

Peer assessment is an important tool to develop critical thinking and autonomous learning—skills that are valued

in today's society. A possible concern for NUS educators in implementing peer assessment is discovering the most time-efficient way of carrying out this procedure without compromising on its benefits. One suggestion is to use students' presentations of course topics as a form of peer assessment, integrating assessment with course coverage and getting students actively involved in thinking and learning (Lim & Chan, 1999). Another alternative is to take the assessment part of the procedure out of class-time. Students could perform the actual assessment online, and still be made accountable for it (Bostock, 2000).

Conclusion

Implemented effectively, peer assessment fosters critical thinking. Having raised some possible problems and solutions in this article, it is hoped that tutors and students can focus more on the processes of criteria setting, marking and negotiating stages in the procedure, and work towards effective peer assessment.

References

- Bostock, S. (2000). *Computer-Assisted Assessments—Experiments in Three Courses*. From Learning Technology website, Keele University. http://www.keele.ac.uk/depts/cs/Stephen_Bostock/docs/caa-ktm.htm. (Last Accessed: 23 December 2002).
- Cheng, W. & Warren, M. (1999). 'Peer and Teacher Assessment of the Oral and Written Tasks of a Group Project'. *Assessment & Evaluation in Higher Education*, Vol. 24, No. 3, pp. 301–314.
- Crockett, G. & Peter, V. (2002). 'Peer Assessment and Team Work as a Professional Skill in a Second Year Economics Unit'. In *Focusing on the Student*. Proceedings of the 11th Annual Teaching Learning Forum, 5–6 February 2002. Perth: Edith Cowan University. <http://cea.curtin.edu.au/tlf/tlf2002/crockett.html>. (Last Accessed: 21 December 2002).
- Crowe, C. & Pemberton, A. (2000). 'But That's Your Job!: Peer Assessment in Collaborative Learning Projects'. *Proceedings of the 3rd Effective Teaching and Learning at University Conference, 9–10 November 2000*. Brisbane: University of Queensland. http://www.tedi.uq.edu.au/conferences/teach_conference00/papers/crowe-pemberton-1.html. (Last Accessed: 21 December 2002).
- Dochy, F.; Segers, M. & Sluijsmans, D. (1999). 'The Use of Self-, Peer and Co-assessment in Higher Education: a Review'. *Studies in Higher Education*, Vol. 24, No. 3, pp. 331–350.
- Falchikov, N. (2002). 'Unpacking' Peer Assessment'. In P. Schwartz & G. Webb (Eds.), *Assessment: Case Studies, Experience & Practice from Higher Education*. London: Kogan Page.
- Gregory, A. & Yeomans, L. (2002). *Peer Assessment and Enhancing Students' Learning*. Leeds Metropolitan University. <http://www.business.ltsn.ac.uk/publications/Best%20Stories.htm>. (Last Accessed: 23 December 2002).
- Hanrahan, S.J. & Isaacs, G. (2001). 'Assessing Self- and Peer-assessment: the Students' Views'. *Higher Education and Development*, Vol. 20, No. 1, pp. 53–70.
- Jordan, S. (1999). 'Self-Assessment & Peer-Assessment'. In Brown, S. & Glasner, A. (Eds.), *Assessment Matters in Higher Education*. Buckingham [England]; Philadelphia, PA: Society for Research into Higher Education & Open University Press. pp. 172–182.
- Lapham, A. & Webster, R. (1999). 'Peer Assessment of Undergraduate Seminar Presentations: Motivations, Reflection and Future Directions'. In *Assessment Matters in Higher Education*. Buckingham [England]; Philadelphia, PA: Society for Research into Higher Education & Open University Press. pp. 183–190.

- Li, L.K.Y. (2001). 'Some Refinements on Peer Assessments of Group Projects'. *Assessment & Evaluation in Higher Education*. Vol. 26, No. 1, pp. 5–18.
- Lim, L. & Chan, S. (2000). 'Practical Ways to Develop Lifelong Learning Skills in SP Students'. *Journal of Teaching Practice, Singapore Polytechnic*. http://www.vc.sp.edu.sg/journals/journals_intro.htm#Practical%20Ways. (Last Accessed: 21 December 2002).
- MacKenzie, L. (2000). 'Occupation Therapy Students as Assessors in Viva Examinations'. *Assessment & Evaluation in Higher Education*, Vol. 25, No. 2, pp. 135–147.
- Poh, S.H. (1999). 'Assessment Issues in Singapore'. *Educational Measurement: Issues and Practice*, Vol. 18, No. 3, pp. 31–32.
- Purchase, H.C. (2000). 'Learning About Interface Design Through Peer Assessment'. *Assessment & Evaluation in Higher Education*, Vol. 25, No. 4, pp. 341–352.
- Race, P. (2001). *A Briefing on Self, Peer and Group Assessment*. Assessment Series No. 9. The Generic Centre, Learning and Teaching Support Network. <http://www.ltsn.ac.uk/genericcentre/index.asp?docid=17219>. (Last Accessed: 21 December 2002).
- Roach, P. (1999). 'Using Peer Assessment and Self-Assessment for the First Time'. In *Assessment Matters in Higher Education*. Buckingham [England]; Philadelphia, PA: Society for Research into Higher Education & Open University Press. pp. 191–201.
- Smith, H.; Cooper, A. & Lancaster, L. (2002). 'Improving the Quality of Undergraduate Peer Assessment: A Case for Student and Staff Development'. *Innovations in Education and Teaching International*, Vol. 39, No. 1, pp. 71–81.
- Sher, W. (2001). *Peer Assessment in the Design & Construction of a Reinforced Concrete Lintel*. From Assessment Case Studies in the Centre for Education in the Built Environment website. Loughborough University. <http://cebe.cf.ac.uk/learning/assess/case-studies/index.html>. (Last Accessed: 21 December 2002).
- Stanton, K. (1999). 'Involving Students in Setting Criteria for Assessing Written Work'. In Hinett, K. & Thomas, J. (Eds.). *Staff Guide to Self and Peer Assessment*. Oxford: Oxford Brookes University. pp. 51–53.
- Zariski, A. (1996). 'Student Peer Assessment in Tertiary Education: Promise, Perils & Practice'. In Abbott, J. & Willcoxson, L. (Eds.), *Teaching and Learning Within and Across Disciplines*, pp. 189–200. *The Proceedings of the 5th Annual Teaching Learning Forum*. Murdoch University. <http://cea.curtin.edu.au/tlf/tlf1996/contents.html>. (Last Accessed: 21 December 2002). ■

Self- and Peer-assessments— Vehicles to Improve Learning

Dr Laina Ho

Centre for English Language Communication

Are feedback exercises becoming the bane of students' lives? It may seem so considering the multiple assessments required of NUS students and the response of one of my students who remarked that it was 'feedback fatigue' and not 'exam fatigue' that riled her. Likewise the apathy of my foreign graduate students towards peer-assessment; they would give inaccurate feedback by ticking the same column for all aspects of an oral presentation skills performance. Thus the peer-assessments that I had conducted in the past semesters had failed miserably. Consequently, I was confronted with the question of how to motivate my students to be more discriminating and enhance their learning at the same time.

Using peer- as well as self-assessments to contribute towards a student's final grade is one of the ways to encourage students to exercise critical thinking and take responsibility for their learning. From the voluminous research done on peer- and self-assessments, the general consensus is that the two assessments are measures of improving students' grades through students' motivation to learn, particularly in tertiary institutes (Boud, 1981; Falchikov, 1986; Falchikov, 1988). Self-assessment is defined as where the learner

them to consider the characteristics of say, a good essay or practical report, or performance skills in a practical exercise" (Boud, 1995:12). Thus self-assessment could enhance learning if learners take responsibility for their grades. It encourages critical thinking and critical assessment that is more objective as opposed to subjective assessment made by a single assessor.

Bearing in mind the advantages of peer- and self-assessments (Boud, 1981; Falchikov, 1986; Falchikov, 1988), I introduced the two assessments together with tutor assessment for a 100% continual assessment course in oral presentation skills that is compulsory for foreign doctorate graduates to read as part of their English language and communication skills requirements. Implemented on a total of 30 students in semester two (AY 2001/02) and semester one (AY 2002/03), the final grade¹ was obtained from average of the three scores from the three assessments. Students were trained to assess themselves using the same criteria in both peer- and tutor assessments through class discussions, lectures, previewing professional public speakers, viewing and learning from professional teaching videotapes, critiquing and assessing the performance of previous students' oral presentation skills. Each student had to do a five-minute

1. Confidentiality applies to the university's grading system but not to the final grade I use.

oral presentation that was videotaped for the individual and the group to practise self- and peer-assessments and to calculate their final scores that included the tutor's ratings. They then individually did a 20-minute videotaped oral presentation where all the three assessments were made.

Observations from the sample group of 30 students in the two semesters showed that self- and peer-assessments contributed significantly towards students' motivation to improve. There were minimal deviations in the marks given by the tutor, the peers and self. This concurred with studies of Filene (1969) and Falchikov (1986) that the older the students the more accurate the peer marks with the tutors, thus addressing the possible problem of inaccuracies in a single teacher's rating. Students came to recognise that the three kinds of assessments were more accurate (reliable) as opposed to evaluations done by a single assessor, which could be subjective. The observation verifies that of Magin & Helmsore (2001) who showed that in oral presentation skills, where there is a single teacher's rating, peer-assessment should be included to act as a benchmark in order to obtain a more accurate overall rating. Instead of being nonchalant, students learned to be more constructive in their peer feedback because they felt that they and their peers needed encouragement to improve. In short, they learned to take responsibility for their learning. In particular, foreign students from a different education system experienced a sense of fulfilment when they attempted to achieve their target marks. In addition, they felt honoured to be given some responsibility for the grading system. In a questionnaire on this exercise, 75% of the foreign graduates agreed that these assessments were fair and accurate. Most importantly, this exercise had shown that self-assessment was a good way of reinforcing assessment procedures and patterns, by "providing an opportunity to renegotiate, in a controlled way, certain aspects of the marking process" (Taras, 2001).

To ensure that all three assessments were accurate and fair, it was necessary to devote time and energy to train the students to do a proper assessment. This may discourage tutors who are constrained by insufficient time for their modules and a large class. However, the

time-consuming and labour-intensive training sessions can be modified and various methods can be adopted to prevent feedback fatigue and subjective assessment. For example, I taught and trained another group of foreign graduates to assess themselves and their peers by checking on specific criteria items in the evaluation forms, without being videotaped for their presentations. The result showed accuracy of peer and tutor marks, i.e. minimal deviation of peers' marks from the tutors'. It also showed that the students were equally motivated to learn and improve on their oral presentation skills.

Self- and peer-assessments of students' course work (particularly those that involve performance skills in laboratories, clinics, field work, work attachments and any practical projects) are more likely to achieve reliability of marks. Studies have also shown that tertiary students have benefited from self- and peer-assessments because they can foster critical self-assessment—a skill highly valued by both the academic as well as the corporate world. However, changing the mindsets of tutors to make students responsible for their final marks could be more difficult than the preparations required on the part of the tutor.

References

- Boud, D. (Ed.). (1981). *Developing Student Autonomy in Learning*. London: Kogan Page.
- Boud, D. (1995). *Enhancing Learning Through Self-Assessment*. London: Kogan Page.
- Falchikov, N. (1986). 'Product Comparisons and Process Benefits of Collaborative Peer Group and Self-assessment'. *Assessment & Evaluation in Higher Education*. Vol. 11, No. 2, pp. 146–166.
- Falchikov, N. (1988). 'Self and Peer Assessment of a Group Project Designed to Promote the Skills of Capability'. *Programmed Learning & Educational Technology*. Vol. 26, No. 4, pp. 327–339.
- Filene, P. G. (1969). 'Self-grading: An Experiment in Learning'. *Journal of Higher Education*. Vol. 40, No. 6, pp. 451–458.
- Magin, D. & Helmsore, P. (2001). 'Peer and Teacher Assessments of Oral Presentation Skills: How Reliable are They?'. *Studies in Higher Education*. Vol. 26, No. 3, pp. 287–298.
- Taras, M. (2001). 'The Use of Tutor Feedback and Student Self-assessment in Summative Assessment Tasks: Towards Transparency for Students and for Tutors'. *Assessment & Evaluation in Higher Education*. Vol. 26, No. 6, pp. 605–614. ■

What is Quality in Assessment Practice?

Mr Dennis Sale,

Section Head, Educational & Staff Development Department,
Singapore Polytechnic

Introduction & Context

One of the most significant developments in education over recent years has been an increasing focus on the importance of assessment and the need to ensure quality

in assessment practice. A number of interrelated factors have contributed to this development.

Firstly, greater pressure on educational institutions to be accountable for their products has made assessment high

profile. We need to be able to justify public expenditure in terms of value for money outputs. The quality of teaching and the cost-effective use of resources are rightly important issues in this context. However, it is assessment that largely defines the value of accredited educational programmes. If assessment practice lacks quality, what value can be placed on the qualifications accredited?

Secondly, major curriculum developments in competency-based programmes and flexible learning initiatives, have shifted the curriculum emphasis from issues of delivery to that of valid, reliable and cost-effective assessment. It is increasingly recognised that people learn from a variety of contexts and media in their own ways and at their own pace. What is important is being able to ensure that we can measure learning accurately, irrespective of how that learning has been derived.

Thirdly, as institutions compete for student numbers or 'types of student', we need to show clearly that student needs and interests will be professionally catered for. Assessment, in particular, must be seen to be well constituted and fair. In assessing learners, we are making claims to know them in important ways. If our judgements are not accurate, what levels of disservice are we doing to these learners as well as to the community we serve?

Finally, in this introduction, it is important to emphasise the key role that assessment plays in the learning process (Boud, 1988; Ramsden, 1992). People will typically learn what they are assessed on and learn in ways to meet those assessment demands. As Ramsden (1992) once wrote:

From our student's point of view, assessment always defines the actual curriculum...Assessment sends messages about the standard and amount of work required, and what aspects of the syllabus are most important (pp. 187–8).

Given this relationship, quality in assessment is inevitably concerned with 'quality learning outcomes'—however defined—because these determine the teaching and learning processes. For example, if our assessment activities lack interest for learners and encourage rote learning, this is likely to lead to 'surface approaches' to learning as documented by Marton, *et al.* (1984). In these situations learners will do what is necessary for the purposes of assessment, but are unlikely to derive both a real understanding of the subject or a genuine interest in it. Similarly, teachers are likely to teach to such requirements and largely transmit knowledge through didactic means using rote-learning activities for reinforcement.

As a result, this has led to greater advocacy for more authentic forms of assessment (Tombari & Borich, 1999), based on premises of more meaningful learning and improvements in the conditions in which learning occurs

(e.g. learner motivation, integration with instruction, ongoing formative assessment). Hence, when we talk about quality in assessment practice, there are equally valid concerns about the nature and usefulness of what is being learned as well as the constitution and conduct of the assessment systems that direct assessment practices at the institutional level.

In this short paper, it is not possible to explore all quality issues in detail. Consequently, I will only outline an organising framework from which standards can be derived for promoting, evaluating and improving the quality of assessment practices.

A Framework for Developing Quality in Assessment

The assessment process typically encompasses three interrelated stages:

- Producing and reviewing a scheme of assessment
- Judging and making decisions based on assessment evidence
- Providing feedback on assessment decisions

In developing quality in assessment practices, it is necessary to identify and implement thoughtfully the activities and processes involved in each of these stages. The following criteria, based on established knowledge of assessment practice and my own experience in developing assessment standards in the UK, are offered:

1. Producing and reviewing a scheme of assessment

- The scheme specifies the assessment methods to be used, their purpose, the marks to be allocated, and the timing of assessments.
- The selected assessment methods are valid for assessing the knowledge, skills and attitudinal components specified, and at the appropriate levels.
- The assessment methods provide fair and reliable assessment opportunities.
- The assessment methods are well constructed, meeting established design criteria for each item type.
- The assessment methods employed are sufficiently varied to encourage learner motivation and enable them to display their learning in different ways.
- The assessment methods are planned to make effective use of time and resources in producing sufficiency of evidence.
- The key aspects of the assessment scheme are explained to learners.

- Opportunities are provided for learners to seek clarification on assessment requirements.
- Ways to ensure the authenticity of assessment evidence are identified.
- The scheme is reviewed at agreed times and updated as necessary.

2. Judging and making decisions based on assessment evidence

- Learners are provided with clear access to assessment.
- The assessment evidence is judged accurately against the agreed assessment criteria.
- Only the criteria specified for the assessment are used to judge assessment evidence.
- The assessment decisions are based on all relevant assessment evidence available.
- Inconsistencies in assessment evidence are clarified and resolved.
- The requirements to ensure authenticity are maintained.

3. Providing feedback on assessment decisions

- The assessment decisions are promptly communicated to learners.
- Feedback to learners is clear, constructive and seeks to promote future learning.
- Formative assessment is planned and ongoing.
- Learners are encouraged to seek clarification and advice.
- The assessment decisions are appropriately recorded to meet verification requirements.
- Records are legible, accurate, stored securely and promptly passed to the next stage of the recording/certification process.

Summary

In this paper, I have argued the necessity for quality in assessment, both for the credibility of the curriculum we offer and for the quality of student learning in those curricula. Hopefully, the framework provided offers a valid and practical base from which to produce standards for quality in assessment and to explore some of the wider concerns relating to the quality of learning outcomes.

References

- Boud, D. (Ed.). (1988). *Developing Student Autonomy in Learning*. London: Kogan Page.
- Marton, F. (1984). *The Experience of Learning*. Edinburgh: Scottish Academic Press.
- Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.
- Tombari, M. & Borich, G. (1999). *Authentic Assessment in the Classroom*. New Jersey: Prentice-Hall. ■

Look out for our next **CDTL Brief**
on the theme of 'Cultivating Leaders'
coming out in April 2003



The Centre for Development of Teaching and Learning (CDTL) engages in a wide range of activities to promote good teaching and learning at the National University of Singapore, including professional development, teaching and learning support, research on educational issues, and instructional design and development.

contributors

William F. Massy
Alice Christudason
Anne Magdaline Netto
Tan Chay Hoon
Matthew C.E. Gwee
Lim Yuen Lie, Lisa-Angelique
Laina Ho
Dennis Sale

advisor

Daphne Pan

editors

Verena Tay
Teo Siok Tuan

layout

Ma Lin Lin

© 2003 CDTL Brief

is published by the

Centre for Development of
Teaching and Learning.

Reproduction in whole or in part
of any material in this publication
without the written permission
of CDTL is expressly prohibited.

The views expressed or implied
in *CDTL Brief* do not necessarily
reflect the views of CDTL.

An online version is available
at our website.

Comments, suggestions and
contributions should be addressed to:

The Editor, *CDTL Brief*
Centre for Development of
Teaching and Learning
Central Library Annexe, Level 6
National University of Singapore
10 Kent Ridge Crescent
Singapore 119260

Tel: (65) 6874-3052

Fax: (65) 6777-0342

Email: cdtpost@nus.edu.sg

<http://www.cdtl.nus.edu.sg>

Printed in Singapore by First Printers Pte Ltd.