

## The quality of assessment

**Link to publication record in USC Research Bank:**

<http://research.usc.edu.au/vital/access/manager/Repository/usc:11265>

**Document Version:**

Author accepted manuscript (postprint)

**Citation for published version:**

Christie, Michael, Nordlund, Lars (2006) The quality of assessment. in Christie, Michael (ed), *Shifting perspectives in engineering education*, Chalmers University of Technology, Sweden, pp.272-277.

**Copyright Statement:**

Copyright © 2006 Michael Christie and Lars Nordlund. Reproduced with permission.

**General Rights:**

Copyright for the publications made accessible via the USC Research Bank is retained by the author(s) and / or the copyright owners and it is a condition of accessing these publications that users recognize and abide by the legal requirements associated with these rights.

**Take down policy**

The University of the Sunshine Coast has made every reasonable effort to ensure that USC Research Bank content complies with copyright legislation. If you believe that the public display of this file breaches copyright please contact [research-repository@usc.edu.au](mailto:research-repository@usc.edu.au) providing details, and we will remove the work immediately and investigate your claim.

## The quality of assessment: who decides?

Michael Christie and Lars Nordlund

### Introduction

In this chapter we want to argue that if the purpose of a university is about initiating, fostering and promoting learning then a focus on the quality of a university's assessment system will help us to better understand both the notion of 'quality' and how we can assure it. Our position is based on the fact that assessment drives much of what happens in a university. It involves the same stakeholders mentioned in another chapter in this book, namely the government and the taxpayers they represent, industry, university administrators, teachers and learners, although this may not be obvious at first glance. A focus on assessment also allows us to discuss some of the definitions of quality that we have mentioned in the chapter on quality assurance. These definitions included the notions of excellence, the best of its kind, a desirable standard, fitness for purpose and value for money. We also introduce a few more, namely continuous and/or transformative improvement.

The definition 'fitness for purpose' forces us to ask 'What is the purpose of assessment in higher education? The simple answer should be to discover if students undertaking a course of study really understand and can apply what they have learnt. This is the pedagogical purpose. But of course it can have a political or a social or an economic purpose. Exactly what those purposes are depends on whom you ask. The government, the taxpayer, industry, the university, the dean of a faculty, the teacher or the student.

### The purpose of assessment

Their answers will be reminiscent of those given in the quality assurance chapter where the question at issue was the purpose of the university itself. In regard to assessment of learning government and industry would like lecturers to assess students so that their marks reveal whether or not they can be employed to do a job as engineers, teachers, social workers, medical staff and so on. If graduates are graded this is even better since it can be used in the job selection process. Assessment procedures that produce statistics which can be compared across institutions are particularly appreciated. Summative, norm referenced examinations in a secure, controlled environment are ideal for this.

The more pragmatic teachers mentioned above are often in agreement with government and industry. So are those who lack job security or are keen not to rock the boat. Researchers who have some teaching duties also tend to 'go with the flow', especially when the flow is along a deep channel dug out by generations of closed-book, end-of-course exams. Assessment for many of these people is about grading students and complying with institutional and government requirements. The system works, why change it.

Teachers who are fascinated with their subject and are interested in promoting learning for learning's sake tend to favour formative, criteria referenced assessment. The more varied and more continuous it is then the better it fits their purpose. Of course this takes more time, breaks with tradition and is not always popular with other teachers who may have an unwritten agreement with their students that traditional exams are the simplest way of grading or gaining accreditation. The latter would argue that closed-book, end-of-course exams provide an even playing field in what is basically a competitive system. They often feel that essays, projects and take-home exams open the door to cheating. Such fears are fed by the fact that there are already internet sites that are willing (for a price) to help students do just that.

'Fitness for purpose' forces us to define the group we are talking about. It also forces us to see that quality assessment for one group may be poor quality for another. But definitions like 'goodness, excellence or the best of its kind' take us a step further. What is the best reason for assessing learning? Which sort of assessment is the best of its kind? If the answer is assessment that tests for a deep understanding of a subject, that encourages meaningful, transferable learning, that diagnoses gaps or mistakes in a student's knowledge, that builds a learner's confidence, that discovers innovative or creative answers then we can begin to list some of the defining qualities of good assessment.

### **Formative assessment**

A defining quality of excellent assessment, given the above, is that it should be formative. In other words students should get immediate feedback and be given the chance of learning from their mistakes. Other qualities are that it should be continuous or at least spread evenly over the whole course. It should be valid. In other words it should serve the purpose we intend, whether that is formative or summative. In the case of an achievement test it should provide a fair sample of what has been learnt. It should also be reliable. A similar assessment taken at different times ought to yield similar results.

To demand formative assessment in its true sense flies in the face of the competitive nature of higher education, at least as it is currently constituted. The present system is designed to sort the sheep from the goats. With formative assessment learners compete against themselves rather than others. One could argue that industry, government and universities should set their own summative exams – this already happens with university entrance exams. Let the employer assess capability for the job or the government set an annual exam to decide on the funding of student places. Teachers could then get on with providing 'quality' assessment, in the sense that it fits the excellent purpose of enhancing learning.

It is interesting to note that Higher Education quality assurance agencies in Sweden and the United Kingdom of Great Britain would probably agree with our definition of quality assessment. They even ask for evidence of it when they carry out their audits. Ironically the audit itself smacks of summative rather than a formative evaluation of quality. The intention is to give feedback to the particular department and university under review. But this purpose is muddled by the need to compare universities. Because audits have to be published a bad report will damage a university's reputation whether or not it rapidly repairs the defects identified in the report. Audits, like end-of-course exams, are highly motivating. But they also encourage a type of compliance that looks more like rote learning than deep understanding. Like their own students, university staff prepare for their particular exam by looking at previous audits, by asking other departments what they did to pass and by anxiously gathering as much evidence as possible to ensure the highest grade from the quality assessors. Like their students they want

to give the auditors what they want. And, like other summative examinations, the audit's main function becomes one of inspection and grading.

The danger of this is that teachers within the university could concede the responsibility for defining and assuring the quality of examination and other assessment practices to others. Had universities taken more care with quality assurance issues earlier they might have avoided the pressures of external reviews. However external reviews are here to stay and in many ways boost the quality of teaching, learning and assessment. Nevertheless, researchers, teachers and students - those who have most at stake, given the espoused purpose of the university – often feel disempowered by external quality assurance. It is a feeling not dissimilar from that which students feel when they study hard for a term, sit an exam and receive a number grade as their only feedback. The purposes that the different groups mentioned above have for assessing students are numerous and varied. In this chapter we make a plea for giving the experts in this field (teachers and students) a chance to agree on the main purpose of assessment. A chance to identify what makes particular assessment practices excellent or the best of their kind.

### **Chalmers case studies**

In an effort to better understand what teachers and students think about assessment we undertook a number of case studies and surveys at Chalmers Technical University in Göteborg. This work was part of the Chalmers strategic effort to improve learning and teaching at Chalmers, mentioned in the introduction to this book. This large scale pedagogical project was begun in 2000 and scheduled to run over five years. Initially 17 action-learning projects were funded. Our project was entitled 'Appropriate ways of assessing learning' and aimed, among other things to:

- Analyse and describe the specific and generic capabilities that students at Chalmers should develop as a result of their undergraduate degree in engineering
- Identify, analyse, promote and, where possible, implement appropriate ways of developing and assessing such capabilities both within subject specific areas and across the curriculum.

Our focus was very much on the ways that assessment can facilitate these aims and we began by mapping the sorts of assessment that were currently in use at Chalmers. The project team was composed of teachers from each faculty at the university and they used emails, interviews, and net searches to determine the types of assessment each faculty used. This was a pilot survey only. Nevertheless we gained an overview of what was happening in regards to assessment. Those faculties that were closer to the humanities end of the academic spectrum (architecture, industrial management, environmental studies) were already using assessment practices that gave students a greater say in demonstrating that they had understood and could apply the knowledge and skills they had learnt.

The more scientifically oriented faculties, which were in the majority, made more use of traditional assessment. By that we mean that the most common way of assessing students was by means of an end-of-course, closed-book exam. Any change and innovation that was occurring was usually in the latter years of these courses and especially in the international masters courses. There was some irony in this given that the Swedish system pays universities according to the previous year's pass rate. In other words the first year courses bring in more money but often have a poorer quality assessment system.

In a follow up survey we asked 50 teachers and 50 doctoral students to write down an example of 'quality assessment'. Our aim was to analyse their examples to see if there was any agreement about the defining qualities of good assessment. The doctoral students had some

teaching duties, carried out research but also studied some subjects by course work. This small piece of action research had interesting results. It demonstrated clearly that conceptions of quality varied, often in line with one's own particular aims and interests. A majority of the examples were of good, traditional practice. This was understandable given what we noted above about assessment patterns at Chalmers. But there were also examples of more innovative practice. One reason for this is that the course work subjects that the doctoral students took or the teachers ran involved smaller numbers and so it was easier to introduce more appropriate forms of assessment.

### **Examples of assessment**

An example of quality assessment in one such course consisted of the students reading the latest journal articles in the field, discussing them and writing critiques of a couple of articles that most interested them. Their critiques were then presenting verbally in a 'conference format'. The mark they received for their efforts was based on a combination of self, peer and teacher assessment. The examples of good, traditional assessment, focused on the extent to which real understanding of the subject matter was tested. The rationale behind the praise for this process was that it countered the guessing game. Students who simply swatted and checked past exam papers for clues as to the exam's content gained some advantage but only if they really understood the concepts and could apply them. In these exams the working out of the problem as well as the final solution were marked.

The number of examples that involved the sort of formative assessment we mention above were in a clear minority. This can partly be explained by the fact that at Chalmers such practice is still apparently rare. We say apparently because during our discussion of the incidents it was clear that other factors influenced the types of assessment used in Swedish universities. As part of its commitment to social justice and individual rights university students have the right to re-sit exams a number of times. This means that there is a type of long-term formative assessment system in place. The effect this has on assessment practice is that teachers are loathe to give too many 'real exams' during a course because any student who fails them has the right to a new exam. The work involved in writing these second-chance exams would be considerable and re-sitting them disruptive for ongoing teaching and learning in the course.

Teachers who want to continuously assess do so by optional exercises and mini exams that give bonus points. These points can lift a student's performance in the final, 'real exam' from a pass to a credit or from a credit to a distinction. The second author of this chapter introduce continuous assessment, mainly in the form of regular take-home exams backed up by shorter oral exams. The student's response to this change of assessment practice was extremely positive. A common comment was that they had to work harder but they also discussed the subject more with their fellow students and felt they had a better understanding to the subject matter.

### **Conclusion**

We are aware that assessment does not occur in a vacuum. As we have demonstrated above different groups have different perspectives on the purpose of assessment. Some of them may be at cross purposes. We have argued throughout this chapter that the first people to be consulted in any discussion on appropriate ways of assessing learning should be the teachers and learners involved in a particular learning situation. They are the experts. Consensus among this group is essential. To ensure a healthy resistance to the sort of collusion that may occur in

the face of outside pressure the best sort of consensus is one that involves the majority of teachers and learners in that subject, across a representative number of universities. This is where the Swedish Higher Education Agency can play a positive role but its consultative rather than its punitive powers must be highlighted in assisting such a process. Our hope is that when such consensus is sought pedagogical ideals such as the meaningfulness and transferability of learning will outweigh political, economic or social pragmatism.

## References

- Ball, C (1985) What the hell is quality? In D Urwin (ed), *Fitness for Purpose: Essays in Higher Education*, SRJE and NFER-Nelson, Guildford.
- Boyle, P and Bowden, J A (1997) Educational quality assurance in universities: an enhanced model. *Assessment and Evaluation in Higher Education*, 22 (2), 111-21
- Harvey, L and Green, D (1993) Assessing quality in higher education: a transbinary research project. *Assessment and Evaluation in Higher Education*, 18 (2), pp 143-148
- Högskoleverkets rapportserie (2001:9) *From quality audit to quality assessment: the new evaluation approach for Swedish higher education*. Stockholm: Högskoleverket
- Newton, J (2000) Feeding the beast or improving quality? *Quality in Higher Education*, 6, 153-163.