ICT, Assessment and the Learning Society[1]

Bertil Roos
Department of Education
Umeå University
S-901 87 Umeå
Sweden

Contact: Bertil.Roos@pedag.umu.se

Abstract

This paper focuses on pedagogy, and the improvement of teaching, learning and assessment. It analyses recent developments in test theory, namely the attention given to assessment as a support for learning by Wiliam, Torrance, Black and Linn. In the process, it draws attention to the difference between formative and summative assessment, high stakes and low stakes testing and divergent and convergent assessment. The paper has three sections: (1) assessment in the recent reform of higher education in Sweden; (2) recent literature on teaching and assessment; and (3) testing, assessment and examinations as social practices to promote rather than control learning. Overall the paper explore possibilities for introducing new ideas about testing into the culture of Swedish higher education.

---

[1] Paper presented to the ICT network at the European Conference on Educational Research (ECER), Lisbon, 11-14 September, 2002)

Assessment and the recent reform of higher education in Sweden.

In the 1930s and 1940s less than one percent of an age cohort studied at universities and institutes of higher education in Sweden, compared with about 40 percent at the end of the millenium (Sörlin & Törnqvist, 2000, p. 34). This difference meant that there were 11 000 students in the beginning of 1940 and over 300 000 by the end of the century (Askling, 1997, p. 36).

In the 1990s universities student-staff ratios increased from 10:1 to 15:1 (Westling, 1999). The number of students during the period increased by 86 percent, whereas the number of staff increased by only 17 percent (Riksdagens revisorer, 2000). Such changes in student-staff ratios have changed the working patterns in universities. Students have problems getting feedback from the teachers and Universities report growing problem of stress and fatigue among staff.

Some authors claim that there is a need for a 'paradigm shift' in assessment, a shift from the current assessment paradigm to the 'problem-solving paradigm', a shift from a testing culture to an assessment culture. Further, this shift is associated with a shift from mental measurement to the assessment of learning (Gipps, 1994, chapter 9, Black & Wiliam, 1998, p. 45).

These paradigm differences can be demonstrated with three distinctions 1) formative- and summative assessment, high- and low stakes testing and 3) divergent and convergent assessment.

Formative versus summative assessment

We can think about the function of assessment partly, at a methodological level, as the 'goals' of the assessment and partly, in a sociological or pedagogical context, as the 'roles' of assessment (Scriven, 1967). The terms formative and summative first appeared in Scriven's article "The Methodology of Evaluation" (1967). Robert Stake has defined the difference in culinary terms: "when the cook tastes the soup, that's formative evaluation; when the guest tastes it, that's summative evaluation" (quoted from Scriven, 1991).

But how does assessment fit the distinction between formative and summative? Black & Wiliam (1998) define formative assessment as follows:

> all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged (p. 2).

If teachers make an effort to develop formative assessment it can take them closer to the students learning and give the students a more active role in their own learning (Black, 2001). Further, such assessment can be seen as integral to learning. Students have to be active in their own assessment and picture their own learning in the light of an understanding of what it means to get better (Black & Wiliam, 1998, p. 22).

Black & Wiliam (p. 14) assert that the core activity of formative assessment lies in the perception, by the learner, of a gap between a desired goal and his or her present state; and in the action taken by the learner to close that gap in order to attain the desired goal.

The practices of formative assessment and feedback overlap. Because of its centrality in formative assessment, it is important to explore and clarify the concept of feedback (Black & Wiliam, 1998, p. 39).

Ramaprasad (1983) defines feedback in terms of two activities: (1) identification of a gap between the actual level and the reference level of a system parameter; and (2) utilization of information which is used to alter the gap in some way (p. 4). For feedback to exist, the information about the gap must be used to alter the gap. If the information is not actually used in altering the gap, then feedback has not taken place (Black & Wiliam, 1998, p. 39).

If the term feedback refers to any information that is provided to the performer, such performance can be evaluated either in its own terms, or by comparing it with a reference standard. Adopting the definition proposed by Sadler (1998), we would argue that the feedback in any assessment serves a formative function only through diagnosis (what do I need to do to get there?). In other words, assessment is formative only when comparison of actual and reference levels yields information which is then used to alter the gap. As Sadler remarks, 'If the information is simply recorded … or is too deeply coded (for example, as a summary grade given by the teacher) to lead to appropriate action, the control loop cannot be closed' (p. 121). The assessment might be formative in purpose but it would not be formative in function. This suggests a basis for distinguishing formative and summative functions of assessment (Black & Wiliam, 1998, p. 45). Summative assessment generates coded information; formative assessment utilises reference- or standard-based information.

Low versus high-stakes assessment

Assessment is one of the most potent forces influencing what teacher should concentrate on in their teaching and what students should concentrate in their learning. Assessment sends a message to students about what is important to learn.

High stakes assessment entails a close association between assessment outcomes and social outcomes. Summative assessment is 'high stakes' if it has important consequences. The stakes are high because the consequences of failure are considerable. Typically, high stakes assessment affects the life chances of students. Moreover, as Amrein and Berliner point out, high stakes assessment is troubled by this close association of assessment outcomes and social outcomes: 'the more important that any quantitative social indicator becomes in social decision-making, the more likely it will be to distort and corrupt the social process it is intended to monitor' (2002, p. 1).

On the other hand low-stakes assessment does not have the same implications. Indeed, its social consequences may be positive, when it is used to guide teaching and learning, rather than to make possibly irreversible decisions about different educational journeys that students can take in the future.

There is a further problem with high stakes assessment. It tends to inflate students' measured achievement – because teachers play the high stakes game and teach to the test. Such improvement, however, does not necessarily imply a corresponding rise in the quality of education, or a better educated student population (Moss, 1992). Through this corruption of teaching and learning, high-stakes assessment stresses basic skills and a narrowing of the curriculum. One result of this negative evaluation of high-stakes assessment is that it has started a wider discussion about assessment. Among other things, it has increased calls for

moving classroom assessment closer to students and their learning (e.g. Linn, 1998, Shepard, 2000).

To enhance the positive impact of assessment and minimize its negative effects, Linn (1998) suggest that it is a necessity that have a variety of ways to assess student. The teachers cannot rely on a single high-stake test when they judge the students. It is important to use multiple indicators when judging the students. The key to long term success is to create a culture that accentuates the intended positive effects and reduces the negative effects of the assessments that are used.

Divergent versus convergent assessment

Convergent and divergent assessment arise 'from teacher's differing views of learning and the relationship of assessment to the process of intervening to support learning' (Torrance & Pryor, 1998, p. 153).

The key issue in convergent assessment is to find out whether the student has a predetermined specific kind of knowledge, understanding or skills. It focuses on students knowledge, understanding and skills in proportion to the curriculum. It uses tick-lists and performance (or performance) statements. It prefers pseudo-open questioning and focuses on contrasting error responses and correct responses. Assessment with these characteristics can be described as behaviouristic, it assesses in a linear way and the assessment is made of the student and executed by the teacher (Torrance & Pryor, 1998).

Divergent assessment, on the other hand, has students' understanding in focus. Divergent assessment has an aim to find out what a student know or can do. The assessment is performed by the teacher and the student together. It is characterize by flexible planning, open forms of recording, emphasis on the learners understanding, open tasks, open questioning and descriptive, qualitative feedback. Divergent assessment strives towards teaching in the zone of proximal development. The theoretical inference from this is that divergent assessment reflects a social constructivist view of education (Torrance & Pryor, 1998).

Testing, assessment and examinations as social practices

Assessment, is far from being merely a technical process. Rather, it is deeply implicated, and may have serious consequences for the lives of those it touches (Johnston et al., 1995, p. 359 in Black & Wiliam, 1998, p. 12).

This point is clearly made by Messick:

> Once it is denied that the intended goals of the proposed test use are the sole basis for judging worth, the value of the testing must depend on the total set of effects it achieves, whether intended or not (Messick, 1989, p. 85).

Linn has made a similar point, arising from his work on high stakes assessment:

> As someone who has spent his entire career doing research, writing, and thinking about educational testing and assessment issues, I would like to conclude by summarising a compelling case showing that the major uses of tests for student and school accountability during the last 50 years have improved education and student learning in dramatic ways. Unfortunately, this is not my conclusion. Instead, I am led to conclude that in most cases the instruments and technology have not been up to the demands that have been placed upon them by high stakes accountability. Assessment systems that are useful monitors lose much of their dependability and credibility for

that purpose when high stakes are attached to them. The unintended negative affects of the high stakes accountability uses often outweigh the intended positive effects. (Linn, R. L. 2000, p. 14).

Assessment processes are, at heart, social processes, taking place in social settings, conducted by, on and for social actors. There are (largely implicit) expectations and agreements negotiated between students and teachers. A feature of such contracts is that they serve to delimit 'legitimate' educational activity by the teacher. For example, in a classroom where the teachers' questioning has always been restricted to 'lower-order' skills, such as the production of correct procedures, students may well see questions about 'understanding' or 'application' as unfair, illegitimate or even meaningless (Schoenfeld, in Black & Wiliam, 1998, p. 47). Thus, all testing has to take account of these social phenomena in the design and administration of its instruments.

ICT and Assessment in the future

The increasing numbers of students in Sweden makes it difficult for teachers to provide individual feedback to the students. Formative online assessment, however be used by students to test out their knowledge and get immediate feedback. Can we use information technology to support teachers and learners with this problems? Teachers can use systems that provide them with tools for analysing and tracking students responses. Teachers can help students with problems. The use of online assessment has the advantage of enabling student responses to be marked and analysed with relative ease and speed. Properly designed online assessment allows students to test their knowledge of a topic and get immediate feedback. Important questions remain, however, about how and whether students organize, structure, and use this information in context to solve more complex problems (Miller, 1999).

It is important that student feedback is of a high enough quality to enhance the learning process. Students need not only feedback on how well they have done but also on what they haven't understood. They also need help to improve their understanding (Ramsden, 1992).

Much attention has been given to ICT as a solution to this problem surrounding teaching and learning in Swedish universities. As late as 1985 computers were used only by a small elite for word processing and simple calculation. Fifteen years later more than 50 percent of the Swedish people have access to the Internet at home or at work. In turn, new forms of on-line assessment have been proposed as a solution to the problem described above. They are being investigated in a EU project coordinated for Umeå university (see www.onlineassessment.nu).

This paper has pointed to problems in Swedish Higher Education. Sweden is committed to the Learning Society and the extension of access to education and knowledge. Historically assessment have been used to separate successful and unsuccessful students. The question is whether assessment can undergo a paradigm shift. Can assessment be used to support learning for all students. International research on assessment, discussed in this paper, has raised the same question. In short, is it possible to replace the assessment or 'audit' society (Power, 1999) with the learning society?

References

Askling, B. (1997). Comments on Traditional and Modern Concepts of Institutional. In *Kvalitet- och förbättringsarbete vid universitet och högskolor*, Högskoleverkets skriftserie 1997:4 S. Stockholm.

Amrein, A. L. & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, *10,* (18). Retrieved August 5, 2002, from http://epaa.asu.edu/epaa/v10n18/

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5* (1), 7-74.

Black, P. (2001). Dreams, strategies and systems: Portraits of assessment past, present and future. *Assessment in Education: Principles, Policy & Practice, 8* (1), 65-85.

Linn, R. L (1998). *Assessments and accountability*. CSE Technical Report 490. CREEST/University of Colorado at Boulder.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher. 29,* 4-16. Retrieved from (http://www.aera.net/pubs/er/pdf/vol29_02/9403AERA004_016a.pdf)

Messick, S. (1989). Validity. In R. LINN (ed) *Educational Measurement* (3rd edition, 13-103) (New York, Macmillan).

Miller, K. (1999). *Which assessment type should be encouraged in professional degree courses - continuous, project-based or final examination-based?* In K. Martin, N. Stanley and N. Davison (Eds), Teaching in the Disciplines/ Learning in Context, 278-281. Proceedings of the 8th Annual Teaching Learning Forum, The University of Western Australia, February 1999. Perth: UWA. Retrieved from http://cleo.murdoch.edu.au/asu/pubs/tlf/tlf99/km/miller.html

Moss, P. A. & Herter, R. J. (1992). *Assessment, Accountability, and Authority in Urban Schools.* Retrieved from http://www.longterm.mslaw.edu/Long%20Term%20View/Vol%201%20No.%204/mossherter.PDF

Power, M. (1999). *The Audit Society: Rituals of verification*. Oxford: Clarendon Press.

Ramsden, P. (1992). *Learning to teach in higher education.* London: Routledge

Riksdagens revisorer. (2000). *Grundutbildningens högskolemässighet* , Rapport 2000/01:1. Stockholm

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

Scriven, M. (1967). The Methodology of Evaluation in *Perspectives of Curriculum Evaluation.* Chicago: Rand McNally & Company.

Scriven, M. (1991). Beyond Formative and Summative Evaluation. In Evaluation and Education: *At Quarter Century*, pp 19-64. The National Society for the Study of Education.

Shepard, L. A. (2000). *The Role of Assessment in a Learning Culture.* Presented at the annual meeting of the American Educational Research Association, New Orleans, April 26, 2000. Retrieved from http://www.aera.net/meeting/am2000/wrap/presadr1.doc

Sörlin, S. & Törnqvist, G. (2000). *Kunskap för välstånd. Universiteten och omvandlingen av Sverige.* SNS Förlag, Stockholm.

Torrance, H. & Pryor, J. (1998). *Investigating Formative Assessment.* Buckingham, Philadelphia: Open University Press,

Westling, H. (Ed.). (1999). *Börjar grundbulten rosta?* En debattskrift om grundutbildningen i högskolan. Stockholm: Rådet för högskoleutbildning.