

Accepted version of Boud, D. (2014). Shifting views of assessment: from teacher's business to sustaining learning. In Kreber, C., Anderson, C., Entwistle, N. and McArthur, J. (Eds) *Advances and Innovations in University Assessment and Feedback*. Edinburgh: Edinburgh University Press, 13-31.

Shifting views of assessment: from secret teachers' business to sustaining learning

David Boud, University of Technology, Sydney

Despite common assumptions that assessment practices in higher education are enduring, the past forty years has seen remarkable changes. A key change has been from the dominance of unseen end-of-year examinations, through the move to 'continuous assessment' and on to a range of diverse assessment approaches. Another notable change has been from assessment weightings being regarded as confidential, to the transparency of assessment standards and criteria of today. Assessment has thus shifted in some positive directions. Unfortunately, during the same period the emphasis on what we now call formative assessment has lessened: from a general acceptance that all courses involved the production of considerable written work that was used purely to aid learning, we now have regimes based on the assumption that all assessment must 'count' towards final marks or a grade point average.

My aim in this chapter is to briefly sketch these developments with the intention of projecting forward to explore emergent assessment practices. These are practices that move beyond current innovations in areas such as authentic assessment, self and peer assessment and improved feedback to students. They represent new views of assessment based upon developing students' capacity for judgement and involve practices which emphasise an active role for students beyond the production of work. The chapter will explore this emerging agenda and consider what change might be possible given the continuing dominance of accountability mechanisms that have had the effect of constraining the development of assessment for learning.

Assessment as taken-for-granted

One of the problems of discussing assessment is that we all have an early conception of what it is and a reaction to it laid down in our earlier experiences. Sometimes assessment has touched us deeply, sometimes it has left bruises (Falchikov and Boud 2007). While changes may occur around us, our point of reference is often what assessment was like when we were most influenced by it. This conception can easily get locked in to provide a personal yardstick against which we judge assessment now. It is important to surface these events as they influence what we regard as legitimate.

In some cases we see these as the gold standard of assessment, in others we resolve never to subject our students to the practices that impacted badly on us.

Many of the changes in assessment that have occurred over the past half century are reflected in my own biography. First, as an undergraduate student in England and then later as an academic, mainly in Australia. There have been minor differences of emphasis between the two countries from time to time, but the main trajectory of assessment is similar. When I entered university there were two main activities that we now label as assessment. First, were set tasks which were completed mostly out of class time, but occasionally within it. These were commonly handed in and 'marked'. A record may have been kept of these marks, but as students we were not very conscious of them beyond the point of the return of our work. Marking usually involved assigning numbers or grades with some brief comment. Work of this kind was commonplace and we completed it because it was the normal expectation of what students did at that time. Secondly, there were examinations. These occurred at the end of the year and sometimes at the end of each term. These were unseen tests undertaken under examination conditions. No notes were allowed and all information needed had to be recalled from memory (Black 1968). No details of examination performance other than the final mark was made available to students. Degree performance was based predominantly on final year examination results. The course I took in physics was part of a wave of innovations in the 1960s in that it also included a final year project which was also marked and contributed in part to the final degree classification.

While different disciplines and different universities used variants on this approach, these variations were minor and the mix of regular marked work with modest comments that did not count towards final grades, and grading based on examinations was commonplace. In our language now, there was a clear separation between formative and summative assessment.

For me, the most influential assessment events were not exams, but the ones that didn't feel like assessment at all at the time. The first, was a project conducted over many weeks in pairs in a laboratory during the final semester of first year in which we had to design and conduct experimental work on a problem the solution to which was unknown, or at least not easily located in texts by undergraduates. It was far from the stereotypical recipe-like lab problem that was common at the time. The second assessment event was a substantial final year project in theoretical physics which involved me in exploring a new way of looking at statistical mechanics. What these activities did for me was to give me a taste of the world of research, not just learning more subject matter. It showed me that it was possible to contribute to the building of

knowledge in a small way and not just to absorb it. While it led to a resolve not to do physics research it was influential in my becoming a researcher.

During my undergraduate years there was a substantial degree of secrecy about assessment processes. We weren't told what the criteria for marking would be and how different subjects would be weighted into final results was confidential. The Head of the Department in which I studied (Lewis Elton) first took the then daring step of formally disclosing the weightings of the elements that would comprise our degree classification in the year of my graduation. Assessment was secret teachers' business: it was not the position of students to understand the basis on which they would be judged.

Over the late 1960s and the early 1970s, a campaign for what was known as 'continuous assessment' was mounted by student organisations (Rowntree 1977). Their argument was that, (a) it was unfair to base degree performance on a limited number of examinations taken at the end as examination anxiety among some students led them not to perform well, (b) assessment should be spread throughout a course to even out workload and anxiety, (c) multiple points of judgement should be used and final grades should be a weighted accumulation of assessments taken across the curriculum, which should be disclosed to students. Assessment for certification moved from one or two points late in a program to a continuous sampling over the years. The battle for 'continuous assessment' was comprehensively won by students and this practice is now so universal that the term for it is fading from our common language. Later, the massive expansion of higher education that occurred without a commensurate increase in unit resources meant that the amount of regular course work that could be marked was severely reduced. Continuous assessment commonly transformed into two or three events within each subject per semester. In the Western world, it appears that Oxford and Cambridge universities now provide some of the rare exceptions that keep to the old ways.

Every change in assessment has unintended consequences and the move to continuous assessment has had quite profound ones. Firstly, students have come to expect that all tasks they complete will contribute in some way towards their final grades. The production of work for purposes of learning alone, with no extrinsic end, has been inhibited. "Will it count?" is a phrase commonly heard when asking students to complete a task. This shift also indicates a change in the relationship and contract between teachers and students. Trust that work suggested by teachers will necessarily be worthwhile has disappeared in an economy of grades. Secondly, having separate events for formative assessment and summative assessment has become unsustainable. When all work is summative, space for the formative is diminished. Poor work and mistakes from which students could have learned in the past and then surpassed are

now inscribed on the record and are weighted in a grade point average. Space for learning is eroded when all work is *de facto* final work. The dominance of the summative is well illustrated by the curious phenomenon pervasive in the US literature of referring to everything other than tests and examinations as ‘alternative assessment’ or ‘classroom assessment’, as if tests and examinations are the gold standard that defines the concept of assessment. Anything else is not quite the real thing, they are alternatives or confined to what a teacher might do in the classroom.

The educational measurement revolution

Alongside the primarily social change to continuous assessment, other forces outside the immediate community of higher education were influencing assessment and seeking to position it quite differently. In the 1960s and 70s the impact of the educational measurement revolution (eg. Ebel 1972) began to influence higher education assessment. The proposition articulated by educational testing specialists from a psychometric background was a simple one. In summary, they regarded student assessment as a folk practice ripe for a scientific approach. If only assessment could be treated as a form of psychological measurement, then a vast array of systematic techniques and strategies could be applied to it. Measurement assumptions were brought to bear on it. The prime one was the assumption of a normal distribution of performance within any given group. Whatever qualities were being measured, the results must follow the pattern of a bell curve. If this assumption could be made then all the power of parametric statistics could be applied to assessment.

The educational measurement revolution was taken up with more enthusiasm in some quarters rather than others. The impact on psychology departments was high and later medical education was strongly influenced by this tradition, but many disciplines were not touched at all. I recall in the late 1970s joining the University of New South Wales and discovering that grades in each subject in the School of Psychology not only were required to fit a normal distribution but that this included an expectation that a requisite number of students needed to fail each subject in conformity to the normal distribution. (It took many years to acknowledge that the selection process into higher education and into later years of the programme, severely skewed the distribution and made assumptions invalid). While not all disciplines shared the enthusiasm of the psychologists, norm-referenced assessment became firmly established. Students were judged against the performance of other students in a cohort, not against a fixed standard.

The impact of educational measurement lingers to today. Notions of reliability and validity are commonly used in assessment discussions and the multiple-choice test—a key technique from this period—has become ubiquitous. The metaphor of

measurement became entrenched for a considerable time and is only recently being displaced. For example, it is interesting to note that between the first (2000) and second editions (2006) of the *UK Quality Assurance Agency Code of Practice on Assessment of Students*, the use of measurement in the definition of assessment had been removed.

A less obvious influence of this period is that once student assessment became the subject of scrutiny beyond the immediate context of its use, it would forever be the object of critical gaze. Assessment was no longer a taken-for-granted adjunct to teaching, it deserved consideration of its own as a separate feature. Assessment became to be discussed independently of disciplinary content or the teaching that preceded it. Assessment *methods* became the focus of attention, as if they were free-standing and the most important element of the process.

Widening the agenda

Following the period of influence of educational measurement, there have been a number of other shifts of emphasis of a greater or lesser effect. The first major ones were the incremental moves from norm-referenced testing to a criterion-referenced and standards-based approach. It is impossible to date this precisely, or even cite key events that mark the transition, but there has been a comprehensive shift, at least at the level of university policy, from judging students against each other to judging them against a standard using explicit criteria. Desired learning outcomes which have to be disclosed to students are widespread as a required feature of course and programme documentation, and increasingly assessment tasks are expected to be accompanied by explicit criteria used to judge whether standards have been met. Prompted by the OECD agenda to create minimum standards for equivalent courses across countries, there have been national (www.olt.gov.au/system/files/resources/Resources_to_assist_discipline_communities_define_TLOs.pdf) and international initiatives to document threshold program standards (<http://www.unideusto.org/tuning/>).

Linked to this, the second shift of emphasis has been the influence of an outcomes-oriented approach and a focus on what students can do as a result of their higher education. Until the 1990s, assessment had focused strongly on what students knew. Students were judged primarily on their understanding of the knowledge domain of the subjects they were studying. There was some emphasis on practical activities in professional courses and project work in later years, but the main emphasis was on what knowledge and academic skills students could demonstrate through assessment tasks. In some quarters, this has been represented as bringing the debate about

competencies and capabilities into higher education. Vocational education and training systems have brought a very strong view about organising learning around explicit operational competencies, but higher education has taken a weaker view which takes a more holistic approach to competencies. The focus has been on outcomes, but not on reducing these to a behavioural level of detail.

Various initiatives have led to an emphasis on transferable skills or generic attributes or core competencies (eg. Hughes and Barrie 2010), that is, skills that all graduates should develop irrespective of discipline of enrolment. There has also been an increased emphasis on making assessment tasks more authentic (Wiggins 1989). That is, creating tasks that have more of the look and feel of the world of practice, than activities that would only be found within an educational institution. For example, replacing the essay with a range of different writing tasks that involve academic writing adapted for particular contexts. Students learn not to perfect the standard academic essay, but to write in different genres for different audiences. This emphasis on authenticity has also permeated beyond vocational courses to ones which may be regarded as more conventionally academic.

These changes to widen the notion of assessment have positioned it as an indicator not of what students know, but of what they can do. And, not only what they can do, but what they can do in a variety of different contexts. What is important here are not the various facets of learning, but how they can be put together into meaningful and complex tasks, the kind of tasks after graduation that professional practitioners encounter.

Dilemmas and contradictions in assessment practice

So, today we have a single term—assessment—that is normally used without qualification to refer to ideas with quite different purposes. It means the grading and certification of students to provide a public record of achievement in summary form—summative assessment. It also means the engagement of students in activities from which they will derive information that will guide their subsequent learning—formative assessment. However, the tasks associated with each have collapsed together. All set tasks seem now to have a dual purpose.

Severe problems are created by this arrangement as it serves neither end very well. Let us take two examples. Firstly, for purposes of certification it may be satisfactory for grades to be used to summarise quite complex judgements about what has been achieved. These can be recorded on a transcript and provide a simple overview of patterns of accomplishment. However, this does not work for formative purposes. A grade or mark has little informational content. A ‘C’ on one assignment tells the

student nothing in itself about what might be needed for a 'B' to be gained in the next. Even when detailed grade descriptors are added, they only reveal what has been done, not what is needed to be done. For purposes of aiding learning, rich and detailed information is needed about what criteria have and have not been met, what is required for better subsequent performance and what steps a student might need to get there. For certification, summary grades are normally sufficient; for learning, much more detail is needed. Indeed, there is a suggestion in the research literature (eg. Black and Wiliam 1998) that the provision of a grade may distract students from engaging with more detailed information about their work.

There is a second tension between the two purposes of assessment. It involves the timing of assessment. For purposes of certification, in decision-making for graduation, employment and scholarships, assessment needs to represent what a student can do on the completion of their studies. Difficulties a student may have had in earlier stages of their course, and which have been fully overcome, should not effect the representation of what a student ends up being able to do. The implication of this thinking is that assessment for certification should occur late in the process of study. Returning to assessment for learning, does late assessment help? The answer is clearly, no. Information for improvement is needed during the process of study, not when it is finished. Indeed, early information is most needed to ensure misconceptions are not entrenched and academic skills can be developed effectively. For certification purposes, assessment needs to be loaded later in courses, for learning, earlier.

While logic might demand that these two purposes be separated so that both can be done well without the compromises that are required when each is associated with the other, it is now unrealistic to imagine that we can reverse the clock and return to a simpler time when different activities were used for formative and summative purposes. The demands that summative assessment cover a much wider range of outcomes than in the past, along with reductions in resources per student, mean that there may be little scope for additional formative activities.

This pessimistic view of the overwhelming dominance of assessment for certification purposes needs to be balanced by the rediscovery and re-emergence of discussion on the formative. The review paper by Black and Wiliam (1998) on formative assessment was one of the very few from the wider realm of educational research that has had an impact on higher education. Many authors took the momentum of this initiative to seek to reinstate the importance of formative assessment (eg. Yorke 2003). It is difficult to know the extent to which the considerable discussions of formative assessment in the higher education literature have embedded themselves into courses though. Like many innovations that have been well canvassed with positive outcomes

(eg. self and peer assessment) there are many reports of practice but, unlike the initiatives mentioned above, little sense that uptake of the idea has been extensive.

In summary, it is apparent that the present state of assessment in practice is often a messy compromise between incompatible ends. Understanding assessment now involves appreciating the tensions and dilemmas between the demands of contradictory purposes.

The emerging agendas of assessment

Feedback

Notwithstanding the dilemmas and contradictions of two different purposes of assessment operating together, where is the assessment agenda moving and why might it be moving in that direction? If we look at what students are saying, we could conclude that the greatest issue for them is feedback; or rather, their perceptions of its inadequacy. In student surveys across universities in both Australia and the UK the top concern is assessment and feedback (Krause et al 2009, HEFC 2011). This is commonly taken to mean that students are dissatisfied by the extent, nature and timing of the comments made on their work. We should be wary though of coming too readily to an interpretation of what they mean. As an illustration, surprisingly students at the University of Oxford also complained of lack of feedback, even though they were getting prompt, detailed and useful comments on the work (a characteristic of the Oxford tutorial system). However, they were concerned that the formative information received, while helping them improve their work, did not enable them to judge how well they were tracking for the entirely separate formal examinations conducted at the end of their second and third years (Oxford Learning Institute 2013).

This concern with feedback has led to a range of responses. At the crudest level, there are stories in circulation of Pro-Vice-Chancellors urging teaching staff to ensure that they use the word feedback at every opportunity when commenting on anything that might be used by students to help them in assessed tasks, so that they remember this when filling in evaluation surveys! More importantly, the concern has prompted researchers to explore differences in interpretation between staff and students in what they mean by feedback (Adcroft 2011). More substantially again, in some cases the concern has led universities to appoint senior personnel to drive improvement and mount systematic and research-based interventions designed to improve feedback in many forms (Hounsell 2007). The initiatives and substantial website developed by Dai Hounsell and his colleagues is particularly notable in this regard (<http://www.enhancingfeedback.ed.ac.uk/>). Most important for the present discussion, it has prompted scholars to revisit the origins of feedback and to question what

feedback is and how it might be conducted effectively (Nicol and Macfarlane-Dick 2006, Hattie and Timperly 2007, Boud and Molloy 2013, Merry et al 2013).

The use of language in assessment lets us down again. As discussed earlier, assessment is used in everyday language to mean quite different things, but for feedback the problem is more severe. We use the term feedback in the world of teaching and learning to refer to the information provided to students, mainly by teachers, about their work. This use of the word appears ignorant of the defining characteristic of feedback when used in disciplines such as engineering or biology. Feedback is not just an input into a system, such as teacher comments on student work. A signal can only be termed feedback if it influences the system and this influence can be detected. Thus, a heating system switches on when the temperature falls below a given level and switches off when a higher temperature is reached. The signal from the thermometer to the heater can only be called part of a feedback system if it detectably influences the output. If we apply this example to teaching and learning, we can only call the 'hopefully useful information' transmitted from teacher to student, feedback, when it results in some change in student behaviour which is then manifest in what students subsequently do. The present emphasis on what the teacher writes and when they give it to the student, needs to be replaced with a view of feedback that considers what students do with this information and how this changes their future work (Hounsell et al 2008). Feedback does not exist if students do not use information provided to them (Boud and Molloy 2012).

We might speculate on how it is that academics who appreciate what feedback is in their own disciplinary area manage to so thoroughly change their understanding of it in teaching, however, it is more fruitful to focus on what are the implications of a clearer conception of feedback on higher education practice. The first implication is that we must focus attention on the student as an entity to be influenced, rather than solely on the teacher seeking to influence. The second is that we must focus not just on the single act of information provision at a point in time (important as that might still be), but on what occurs subsequently: what effects are produced, when might they expect to be seen? Thirdly, we must be conscious that students are not passive systems, but conscious, thinking agents of their own destiny. How can student agency influence the processes of feedback?

My view is that we are at a point of fruitful innovation that will lead from a reappraisal of what we mean by feedback to changes in practice. The starting point should focus on course design. Changing what teachers do when confronted with student work in itself presents limited options. Looking at where tasks occur, what their nature is, how they relate to learning outcomes and what follows in future tasks during the semester may start to make a difference. Most importantly, change will

only occur if there are active steps to monitor what changes take place in students' work following input from others. Information on the performance of later tasks gives invaluable information to teachers about the effectiveness of their earlier inputs. The design of courses might start by consideration of how feedback loops can be adequately incorporated throughout the semester so that, for example, students can learn, demonstrate their learning, receive comments about it, act on these comments and produce subsequent work on multiple occasions. Claims by students that feedback is inadequate can only be effectively countered by showing that it does make a difference.

Developing judgement

While feedback might be prominent publicly, a more fundamental issue concerning assessment is emerging. As it becomes increasingly apparent that assessment is probably the most powerful shaper of what students do and how they see themselves, the question arises: does assessment have the effects on students that are desired for a higher education? As there are two purposes, it is likely to have two kinds of effect. Firstly, for certificatory purposes does it adequately and fairly portray students' learning outcomes from a course? Secondly, does it lead students to focus their attention and their efforts on what they will most need on graduation? While the generic attributes agenda is focusing attention on features of a graduate needed in all programs and how these might be developed, there is an underpinning issue that affects all other outcomes. It is, how does assessment influence the judgements students make about their own work? Clearly graduates need to know things and do things and be able to work with others to get things done. But, they also need to be able to know the limits of their knowledge and capability and be able to tell whether they can or cannot address the tasks with which they are faced. In other words, they need to be able to make judgements about their own work (Joughin 2009). Without this capacity to make effective judgements, they cannot plan and monitor their own learning, or work with others to do the work at hand. How well do their courses equip them to do this? In particular, what contribution does assessment make?

A precursor to this focus on student judgement is found in the literature on student self-assessment and self-directed learning (Knowles 1973). Particularly since John Heron's seminal article in 1981 (Heron 1988), there has been a flourishing of studies about student self-assessment, often in conjunction with peers. Unfortunately, much of this literature has been preoccupied with seeking to demonstrate that students can make similar judgements of their grades as their teachers (Boud 1995). While many students can do this reasonably well, research has identified that students in introductory classes and those whose performance is weaker than average tend to overrate themselves and that students in advanced classes and those above average

tend to underrate themselves. Regrettably, this research focus is an outcome of thinking about certification, not about learning. The implicit, and sometimes explicit, aim of this research appears to be trying to judge if student marks could be substituted for teachers marks. This is in the end a fruitless endeavour for at least two reasons. Firstly, there are likely to be different marks generated by students depending on how great the consequences will be for them. Secondly, the generation of marks does not address exactly what students are and are not able to judge in their own work.

Viewed from the perspective of assessment for learning, the problem changes dramatically. Of course, students on first encounter with new material will not be good judges of their own performance. As at the start they will not sufficiently appreciate the criteria they need to apply to their work, it is understandable that they may err on the side of generosity to themselves—they just don't know that their work is not good enough. As they understand better what they are studying, they will increasingly appreciate criteria for good work and be able to apply them to their own work. Once they are sufficiently aware of the complexity of what they are doing and are conscious of their own limitations prior to having reached a level of mastery, they will be sparing in their own judgements and tend to underrate themselves. Self-assessment then should be seen as a marker of how well students are tracking in developing the capacity to judge their own work. We should not be dismayed that they are over or underrating as that is just an indicator of progress on a more important process: that of calibrating their own judgement (Boud, Lawson and Thompson 2013).

Assessment as equipping students for future challenges

This leads us to a view about where assessment for learning is heading. Unlike those who seek to set standards and articulate competencies to be achieved, my starting point, like that of Barnett (1997), is that the future is unknown and necessarily unknowable to us (see also Kreber in this volume). Acceptance of this creates constraints and possibilities for what we do in higher education. Of course, new knowledge, skills and dispositions will be required by our students in the future that cannot possibly be acquired now. So, whatever else, we must prepare students to cope with the unknown and build their capacity to learn when the props of a course—curriculum, assignments, teachers, academic resources—are withdrawn. What then does that imply for what and how we assess?

Returning to our original distinction between the purposes of assessment as certifying achievement (summative assessment) and aiding learning (formative assessment), it is possible to see a third purpose, that of fostering lifelong learning. It could be reasonably argued that this latter purpose is merely a subset of formative assessment;

however, there are merits in separating it out. Formative assessment, both in the literature, and in practice, is predominantly concerned with assisting students to cope with the immediate learning demands required to enable students to address what they will encounter in summative assessment. Acknowledgement may be given to its longer-term importance, but strategies to support that are not so common.

To provide a focus for this idea, I coined the term ‘sustainable assessment’. Following the form of a well-known definition of sustainable development, sustainable assessment was described as “Assessment that meets the needs of the present without compromising the ability of students to meet their own future learning needs.” (Boud 2000). This does not refer to assessment being sustainable for the staff who have to carry the marking load, though that is desirable also. It clearly positions the sustainability in terms of student learning. It focuses on assessment tasks that not only fulfil their present requirement—to generate grades or to direct immediate learning—but also contribute to the building of a student’s capacity to learn and assess themselves, beyond the current task. Thus, for example, a sustainable assessment task might involve students in identifying for themselves criteria appropriate for judging the task in hand and planning what further activities are required in the light of their current performance against these criteria. It does not imply that they receive no assistance in these processes, but it does mean that they are not specified in advance by a teacher. Sustainable assessment may not involve wholesale changes in assessment tasks, but it does require changes in the pedagogic practices that accompany them, especially with regard to feedback. Hounsell, Carless and others have taken the notion of sustainable assessment and further applied it to the practices of feedback (Hounsell 2007, Carless, Salter, Yam and Lang 2011).

An agenda for assessment change

If we were to generate a set of ideas for assessment that would respond to the demands of coping with uncertainty and prepare students for a future in which learning was a central feature of their lives, what might they include?

Feature 1. Sustainable

As discussed above, acts of assessment will need to look beyond the immediate content and context to what is required beyond the end of the course. A view is needed that is not a simple projection of present content and practices, but what is required for students to face new challenges. A key element of this must be a strong focus to avoid creating dependency on current staff or courses. While assessment to please the teacher has supposedly ended in the past, any residues of this must be addressed. The more insidious challenge is to ensure that assessment doesn’t involve

always looking to teachers for judgement. Multiple sources to aid judgement must become a normal part of assessment regimes.

Feature 2. Develops informed judgement

As argued before, students must develop the capacity to make judgments about their own learning otherwise they cannot be effective learners now or in the future. This means that assessment should focus on the process of informing students' own judgements, as well as on others making judgements on their work for summative purposes (Boud and Falchikov 2007). The development of informed judgement thus becomes the *sine qua non* of assessment. Whatever else it might do, this is needed to ensure graduates cope well with the future. We should be aware that summative assessment alone is too risky and does not equip students for new challenges. Of its nature, it tends to be backward looking and it is not a strong predictor of future accomplishment. Assessment then is more important than grading and must be evaluated on the basis of what kinds of students it produces. Of course, opportunities for developing informed judgement need to be staged across a program as isolated opportunities to practice this will not plausibly lead to its development. Therefore thinking about assessment across course modules becomes not just desirable, but essential.

Feature 3. Constructs reflexive learners

We have come a long way from when assessment was a secretive business to which students were blind. Transparency is now a key feature. However, there is a difference between openness and involvement. If students are to develop their own judgements and if assessment events are the focus of such judgements, students need to be active players in it. They need to understand what constitutes good work, be able to demonstrate this and judge if what they have produced meets appropriate standards. Students must necessarily be involved in assessment because they need to know how to do it for themselves. Assessment then needs to position students to see themselves as learners who are pro-active and generative and drive their own learning. An example of this is in the use of rubrics. The provision of a rubric for a task that specifies learning outcomes and criteria associated with each to be marked by tutors provides limited scope for reflexivity: students follow the path of others. However in contrast to this, if the task prompts students to construct and use a rubric, they become actively involved in making decisions about what constitutes suitable criteria. This can require them to demonstrate that a learning outcome has been addressed, identify signs in completed work that indicate what has and has not been achieved and the gather evidence from other parties through seeking and utilizing feedback. Fostering reflexivity and self-regulation is not something to be relegated to a limited number of tasks, but should be manifest through every aspect of a course. The programme and all its components needs to construct the reflexive learner.

Feature 4. Forms the becoming practitioner

Finally, assessment needs to shape the becoming practitioner. All students become practitioners of one kind or another. It is only in particular professional or vocational courses that it is known what kind of practitioner they are likely to become. There are common characteristics of all those who practice in society: they address issues, they formulate them in terms of addressable problems and they make judgements.

Assessment to this end needs to help students calibrate their own judgements (Boud, Lawson and Thompson 2013). Learners act on their belief in their own judgements; if these are flawed it is more serious than them having particular knowledge gaps.

Assessment then needs to contribute to students developing confidence and skills to manage their own learning and assessment. Understanding is not sufficient; showing that they can perform certain tasks is not enough. Capable beginning practitioners need to be able to become increasingly sophisticated in judging their work. In particular, they need to be able to do so when working effectively with others to assist each others' learning and mutually develop informed judgement.

This view provides a substantive agenda for further changes in assessment. Any one of the particular elements mentioned above can be seen in the literature, but they are rarely seen in concert, and even less so across a curriculum.

Conclusion

In conclusion, what would assessment that helped meet future challenges look like? It would start by focusing on the impact of assessment on learning as an essential assessment characteristic. It would position students as active learners, seeking understanding of standards and feedback. It would develop their capacity to make judgments about learning, including that of others. It would involve treating students more as partners and less as subjects in assessment discussions. And, it would contribute to building learning and assessment skills beyond the course.

Of course, the first question to be asked is: would such assessment be more demanding for teachers? It would require us to think much more clearly about what changes in students we expect courses to influence. It also would require an initial investment in redesign and a redistribution of where we focus our effort. But following this it could lead to a more satisfying use of time with less time spent on repetitious tasks and more on what really makes a difference.

References

- Adcroft, A. 2011. The mythology of feedback, *Higher Education Research and Development*, 30, 4: 405-419.
- Barnett, R. 1997. *Higher education: A critical business*. Buckingham: The Society for Research into Higher Education and Open University Press.
- Black, P. J. 1968. University examinations, *Physics Education*, 3: 93-99.
- Black, P., and Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Education*, 5, 1: 7-74.
- Boud, D. 1995. *Enhancing learning through self assessment*, London: Kogan Page.
- Boud, D. 2000. Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22, 2: 151-167.
- Boud, D. and Falchikov, N. 2007. Developing assessment for informing judgement. In *Rethinking assessment for higher education: Learning for the longer term*, ed. D. Boud, and N. Falchikov. 181-197. London: Routledge.
- Boud, D. and Molloy, E. eds. 2013. *Feedback in higher and professional education*. London: Routledge.
- Boud, D., Lawson, R. and Thompson, D. 2013. Does student engagement in self-assessment calibrate their judgement over time? *Assessment and Evaluation in Higher Education*, DOI:10.1080/02602938.2013.769198
- Carless, D., Salter, D., Yang, M. and Lam, J. 2011. Developing sustainable feedback practices, *Studies in Higher Education*, 36, 5: 395-407.
- Ebel, R. L. 1972. *Essentials of educational measurement*. Englewood Cliffs, New Jersey: Prentice-Hall. Originally published 1965.
- Falchikov, N. and Boud, D. 2007. Assessment and emotion: the impact of being assessed. In *Rethinking assessment for higher education: Learning for the longer term*, ed. D. Boud, and N. Falchikov. 114-156. London: Routledge.
- Hattie, J. and Timperley, H. 2007. The power of feedback, *Review of Educational Research*, 77: 81-112.
- Heron, J. (1988). Assessment revisited. In *Developing student autonomy in learning*. ed. D. Boud. London: Kogan Page.
- Higher Education Funding Council for England. 2011. *The national student survey: Findings and trends 2006–2010*. Bristol: Higher Education Funding Council for England.
- Hounsell, D. 2007. Towards more sustainable feedback to students. In *Rethinking assessment for higher education: Learning for the longer term*, ed. D. Boud, and N. Falchikov. 101-133. London: Routledge.
- Hounsell, D., McCune, V., Hounsell, J. and Litjens, J. 2008. The quality of guidance and feedback to students, *Higher Education Research and Development*, 27, 1: 55-67.
- Hughes, C. and Barrie, S. 2010. Influences on the assessment of graduate attributes in higher education, *Assessment and Evaluation in Higher Education*, 35, 3: 325-334.
- Joughin, G. ed. 2009. *Assessment, learning and judgement in higher education*. Dordrecht: Springer.

- Knowles, M. S. 1973. *The Adult Learner: A neglected species*. Houston: [Gulf Publishing Company](#)
- Krause, K., R. Hartley, R. James, and C. McInnis. 2009. The first year experience in Australian universities: Findings from a decade of national studies. http://www.cshe.unimelb.edu.au/research/experience/docs/FYE_Report_1994_to_2009.pdf
- Merry, S., Price, M., Carless, D. and Taras, M., eds. 2013. *Reconceptualising feedback in higher education*. London: Routledge.
- Nicol, D. and Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice, *Studies in Higher Education*, 31, 2: 199–218.
- Oxford Learning Institute. 2013. Assessment and feedback, <http://www.learning.ox.ac.uk/support/teaching/resources/assess/> [viewed 16 May 2013].
- Rowntree, D. 1977. *Assessing students: How should we know them?* London: Harper and Row.
- Wiggins, G. 1989. A true test: toward more authentic and equitable assessment, *Phi Delta Kappan*, 70: 9 (May).
- Yorke, M. 2003. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice, *Higher Education*, 45, 4: 477-501.