

Challenges in reforming higher education assessment: a perspective from afar

Retos en la reforma de la evaluación en educación superior: una mirada desde la lejanía

Boud, D. 

Deakin University (Melbourne, Australia), University of Technology Sydney (Australia) & Middlesex University (London, United Kingdom)

Abstract

Can we be sure that assessment in higher education meets the need of developing and assuring high quality learning outcomes? Current assessment is typically a collection of conventional practices that have never been seriously questioned. Ten years ago, as part of a national project, representatives from Australian universities came together to identify an agenda for change in assessment. The resulting document—Assessment 2020: Seven propositions for assessment reform in higher education—focused on how assessment needed to change to support long term learning. That is, not how students can pass the next exam, but learning that is useful beyond the point of graduation.

From a learning-centred view this paper examine progress on assessment reform in universities internationally from the perspective of one of the players. It starts by considering Assessment 2020 to see where action is still needed. It reviews some of the major shifts in assessment in higher education and considers their implications. These include the move from comparing students (norm-referencing) to judging outcomes against standards (standards-based); and importantly, the conceptual shift from the single purpose of assessment as certifying students to multiple purposes including aiding learning and building the capacity of students to make their own judgements.

Keywords: Assessment; Assessment reform; Higher Education; Evaluative Judgement

Resumen

¿Podemos estar seguros de que la evaluación en la educación superior satisface la necesidad de desarrollar y garantizar resultados de aprendizaje de alta calidad? La evaluación actual es típicamente una colección de prácticas convencionales que nunca han sido seriamente cuestionadas. Hace diez años, como parte de un proyecto nacional, representantes de universidades australianas se reunieron para identificar una agenda para el cambio en la evaluación. El documento resultante *Evaluación 2020: siete propuestas para la reforma de la evaluación en la educación superior*, se centró en cómo debía cambiar la evaluación para apoyar el aprendizaje a lo largo de la vida. Es decir, no se trata de cómo los estudiantes pueden aprobar el próximo examen, sino de que aprendan lo que será útil más allá del momento de la graduación.

Desde una perspectiva centrada en el aprendizaje, este documento analiza a nivel internacional el progreso que se ha producido en las universidades en torno a la reforma de la evaluación desde la perspectiva de uno de los actores. Se inicia considerando el documento *Evaluación 2020* para ver dónde aún se necesitan acciones. Se revisa algunos de los principales cambios en la evaluación en la educación superior y sus implicaciones. Estos incluyen el tránsito de comparar estudiantes (evaluación referida a normas) a juzgar resultados contra estándares (evaluación basada en estándares); y, lo que es más importante, el cambio conceptual de pasar de un propósito simple de la evaluación, como es el de certificar a los estudiantes, a considerar múltiples propósitos, incluyendo ayudar al aprendizaje y al desarrollo de la capacidad de los estudiantes para emitir sus propios juicios.

Palabras clave: Evaluación; Reforma de la evaluación; Educación superior; Juicio evaluativo

Received/Recibido	2020 May 23	Approved /Aprobado	2020 May 26	Published/Publicado	2020 June 24
--------------------------	-------------	---------------------------	-------------	----------------------------	--------------

We all have an obligation, as part of the role of the university as a public institution, to declare that our graduates have met the academic requirements of their degrees. However, we cannot assume that because it has been done in a particular way in the past, assessment continues to meet the needs of the present. We must assess in ways that do not undermine the very learning we are trying to promote; we must be conscious that anticipation of assessment has a profound effect on what and how students learn.

What specific work does assessment need to do? It is necessary to consider the various important and legitimate roles it has. It is well-established that assessment and assessment related activities have three key purposes (Boud, 2000, Boud & Soler, 2016):

- To *judge and certify* what students have learned. That is to validate that students have met the requirements of the course. This is often termed summative assessment.
- To *help students learn* in order to meet the requirements of the course. This often occurs through the use of various feedback processes and is termed formative assessment
- To *develop students' capacity to judge their own work* beyond the timescale of the course. This involves them progressively developing their own evaluative judgement, and is termed sustainable assessment.

The first of these purposes almost always occurs in some form, but how effective it is in doing so is questionable. For example, what constitutes meeting the requirements of the course? What criteria should be used for making a judgement? These purposes are increasingly viewed in terms of enabling and assuring that the stated learning outcomes of a course have been attained. However, this is commonly done poorly. The way in which assessment is reported does not normally allow a connection between student grades and the declared learning outcomes for a course to be made.

From the point of view of the individual teacher within an academic discipline, changes in assessment often appear to be so slow as to

be imperceptible. However, looking at universities as a whole over several years there have been major changes, and many more can be anticipated. As universities become more influenced by external influence and regulation, the press of globalisation and the demands of students, new trends are apparent worldwide.

Assessment has a legacy of discipline-based practices for which there are no well-articulated rationales, and a set of rule-driven activities that typically are a reaction to previous problems that have been encountered. In response to increased numbers of students, assessment methods have been introduced that, while they supposedly reduce marking, such as multiple-choice tests, they also require considerably greater investment in design and development, that is not taken into account. Excessive amounts of assessment were found in some areas and often arbitrary administrative decisions were made to reduce them with little thinking through of the consequences on student learning. While formative assessment and feedback has recognised in principle following the prompting of Black and Wiliam (1998), most courses were lacking in this regard and the pressures to reduce assessment had the unintended consequence of reducing student practice and feedback opportunities. Vast areas of the curriculum were under-assessed and even larger areas over-assessed. Students could complete degrees through strategic memorisation while retaining fundamental misconceptions of key concepts. These conventional assessment practices had different features in different countries and different disciplines, but none were beyond censure. In countries where there were standard surveys evaluating students' experience, the assessment and feedback scales were typically the least highly-rated of any aspect of their courses.

This paper takes a learning-centred view of the current environment. It addresses some of these changes, and the challenges that have been generated in the process. Its aim is to provide an overview of the changes that have occurred in assessment in higher education and

the various drivers for change. It identifies key areas for further focus on assessment reform and suggests priorities that will benefit all purposes of assessment. The first part of the paper does this through a review of a consensus document created 10 years ago as an agenda for assessment reform in universities that focused on what needed to be done to improve university assessment practices for the longer-term. It offers the author's appraisal of what has been achieved and what still remains to be done.

The second part of the paper focuses on what is needed for the next ten years. It draws on a mix of scholarly work and policy observations. Some of the developments continue the existing agenda, but other issues remain to be explored. It is important to recognise in this discussion that countries, institutions and academic disciplines do not proceed at the same pace over the years. In most countries, it is possible to find in any given institution or even any given faculty, the best and the worst of assessment practice in close proximity. We can only therefore talk in terms of general trends, not specific achievements and, indeed, the timescale of a decade is needed to even recognise the trends. The paper is a personal reflection which does not attempt to review the considerable literature that has sought to make an influence, for example, by major figures such as Royce Sadler and John Hattie. It makes use of studies by the author to illustrate the argument.

Part 1. Review of what has been achieved

A useful starting point for a review of progress on assessment is an initiative from 2010 which led to the document: *Assessment 2020: Seven propositions for assessment reform in higher education*. This was the outcome of a national project involving all Australian universities. The initiative recognised that assessment practice across the system needed to change and that a greater focus was needed both on longer-term learning and on assessment which supported the curriculum. It was based on the idea that: "Universities face substantial change in a

rapidly evolving global context. The challenges of meeting new expectations about academic standards in the next decade and beyond mean that assessment will need to be rethought and renewed" (Boud & Associates, 2010, p. 1).

The project sought to provide a stimulus for those involved in the redevelopment of assessment at all levels. The process it used was to draw on the expertise of a group of highly experienced assessment researchers, academic development practitioners and senior academic managers to identify current best thinking about the ways assessment needs to address immediate and future demands. As a way of gaining commitment from all the stakeholders, the final document was co-authored by representatives from almost all Australian universities and a group of assessment experts. The representatives included many from the most senior levels of those responsible for assessment policy across each institution (Chairs of Academic Board, Pro-Vice-Chancellors, Chairs of University Assessment Committees).

The final product took the form of a consensus document in which those involved in policy development, policy implementation and research on university assessment practice were included. The propositions identified in it were therefore not radical, but represented compromises that reflected both what was accepted as good practice and which could reasonably be adopted across the system. While the document was frequently consulted and was influential in the formulation of assessment policy in some universities, it is difficult to directly link any particular changes to specific propositions.

The principles on which it was based were stated as follows:

1. Assessment ... powerfully frames how students learn and what students achieve. It is one of the most significant influences on students' experience of higher education and all that they gain from it. The reason for ... improving assessment practice is the huge impact it has on the quality of learning.

2. Assessment is the making of judgements about how students' work meets appropriate standards. Teachers, markers and examiners have traditionally been charged with that responsibility. However, students themselves need to develop the capacity to make judgements about both their own work and that of others in order to become effective continuing learners and practitioners.

3. Assessment plays a key role in both fostering learning and the certification of students. However, unless it first satisfies the educational purpose of ensuring students can identify high quality work and can relate this knowledge to their own

work, the likelihood that they will reach high standards themselves is ... reduced. (p.1)

The Propositions

The seven propositions which were constructed aimed to articulate an agenda outlining important directions for assessment policy and practice to take. They were called *Assessment 2020* on the assumption that it would take ten years for significant change to be seen in assessment regimes. The complete set of propositions, excluding the rationale for them, is listed in Table 1.

Table 1. Assessment 2020: Seven propositions for assessment reform in higher education

ASSESSMENT HAS MOST EFFECT WHEN:	
1. ... assessment is used to engage students in learning that is productive.	<ul style="list-style-type: none"> i. ... assessment is designed to focus students on learning ii. ... assessment is recognised as a learning activity that requires engagement on appropriate tasks.
2. ... feedback is used to actively improve student learning.	<ul style="list-style-type: none"> i. ... feedback is informative and supportive and facilitates a positive attitude to future learning. ii. ... students seek and use timely feedback to improve the quality of their learning and work. iii. ... students regularly receive specific information, not just marks and grades, about how to improve the quality of their work.
3. ... students and teachers become responsible partners in learning and assessment.	<ul style="list-style-type: none"> i. ... students progressively take responsibility for assessment and feedback processes. i. ... students develop and demonstrate the ability to judge the quality of their own work and the work of others against agreed standards. ii. ... dialogue and interaction about assessment processes and standards are commonplace between and among staff and students.
4. ... students are inducted into the assessment practices and cultures of higher education.	<ul style="list-style-type: none"> i. ... assessment practices are carefully structured in early stages of courses to ensure students make a successful transition to university study in their chosen field. ii. ... assessment practices respond to the diverse expectations and experiences of entering students.
5. ... assessment for learning is placed at the centre of subject and program design.	<ul style="list-style-type: none"> i. ... assessment design is recognised as an integral part of curriculum planning from the earliest stages of course development. ii. ... assessment is organized holistically across subjects and programs with complementary integrated tasks.
6. ... assessment for learning is a focus for staff and institutional development.	<ul style="list-style-type: none"> i. ... professional and scholarly approaches to assessment by academic staff are developed, deployed, recognised and rewarded by institutions. ii. ... assessment practices and the curriculum should be reviewed in the light of graduate and employer perceptions of the preparedness of graduates. iii. ... assessment of student achievements is judged against consistent national and international standards that are subject to continuing dialogue, review and justification within disciplinary and professional communities.
7. ... assessment provides inclusive and trustworthy representation of student achievement.	<ul style="list-style-type: none"> i. ... interim assessment results used for feedback on learning and progress do not play a significant role in determining students' final grades. ii. ... evidence of overall achievement to determine final grades is based on assessment of integrated learning iii. ... certification accurately and richly portrays graduates' and students' achievements to inform future careers and learning.

The following points should be noted about them. Firstly, they incorporate a focus on all purposes of assessment: summative, formative and sustainable. They suggest that it is not meaningful to focus exclusively on one or other of these purposes of assessment, all must be considered together. A study of university assessment policies published prior to the propositions (Boud, 2007) had identified that overwhelmingly university assessment policies focused on the summative, with just a little on the formative, and hardly at all on the sustainable. This was a concern of the group that formulated the propositions. Secondly, without it stating so explicitly, there was an assumption that assessment now is standards-based and it acts to judge performance against learning outcomes. Norm-referenced assessment is excluded, as it had been discouraged or banned at most Australia universities by then, and this has become established in national quality standards since. Thirdly, the focus throughout is on student learning. Testing which does not have a positive influence on student learning, that is measurement of outcomes without consequential validity, has no place in university assessment. Finally, assessment is not a unilateral practice imposed on students. For assessment to have an influence on learning, students must be actively engaged in assessment processes, not just responding to the non-negotiable prompts of others.

Take up of the Propositions

Over the ten years since they were composed, how much progress has been made in university assessment? The answer to this question varies greatly, and I can only draw on my direct experience with Australian universities and with some elsewhere. A few universities have embraced the propositions substantially and revised their assessment policies in the light of them. For others it is difficult to determine the extent of change as there have also been other drivers of change during this period.

The propositions were more influential at the level of assessment design, both for programs and for specific course units. Changes have occurred with regard to the first three

propositions with the recognition that assessment can positively influence learning, and that students need to be actively engaged with it. Peer assessment activities have proliferated particularly in the formative area. Peer testing (the use of peers to generate formal grades) is increasingly being criticised as not credible, but peer feedback processes are proving valuable and are being widely adopted.

Assessment is seen not as a measurement or judgement that leaves the 'subject' (the student) unaffected, but an act that profoundly influences students' study and their priorities in many ways. Poor choice of assessment activities leads to poor learning and distorts what students end up being able to do. There is still an unfortunate overemphasis on assessing low-level knowledge and an under-emphasis on assessing the application of that knowledge. Tests too often prompt students to memorise rather than understand. There is a much greater awareness that this is a problem now, but excuses are still made that multiple-choice or short answer questions must dominate assessment because of large class sizes. What is forgotten in this process is that while they are easy to mark, it is extremely difficult to design high quality multiple-choice tests. We therefore see a proliferation of badly constructed tests that communicate the message to students that what is important is to learn the low-level knowledge needed to pass them. This fosters poor study habits among students. Illustrations of assessment regimes in courses with large numbers of students (1000+) show that much can be done to improve the quality of learning through assessment design within existing levels of resource: see the cases studies presented at www.feedbackforlearning.org

In relation to the fourth proposition, one of the greatest successes has been in the rise in interest in what has been termed *transition pedagogy* to assist students to be inducted into the culture and practices of higher education learning (Kift et al., 2010; www.transitionpedagogy.com). First year classes are seen not only as introducing students to the disciplines they will study, but the processes of learning and learning-how-to-

learn that they will need throughout their courses. Investment in the quality of first year subjects and first year assessment has not only a long-term benefit for the rest of the course, but enables students to be retained who would otherwise fail.

Assessment design for learning, the topic of the fifth proposition, has been taken more seriously (Bearman et al., 2016; 2017). This has been driven by two changes. Firstly, the focus on standards and learning outcomes and the curriculum mapping that is associated with it has meant that it has become easier to notice that assessments do not address learning outcomes as specified (see www.assuringlearning.com). Institutions have become embarrassed that this transparency has revealed they are not doing what they claim to do in terms of the outcomes achieved by their actual graduates at the point of graduation. Secondly, the desire to have programs accredited by external international agencies (e.g. the AACSB) has focused the attention of business schools in particular on the design of their assessment and the need to align this with explicit learning outcomes. Assessment design was a term used only by education professionals a decade ago, but it is coming into more general use. The topic was given an impetus by the Assessment Design Decisions Project that developed a website with resources to aid academics in the design of assessment: www.assessmentdecisions.org

Assessment as a focus for institutional and staff (faculty) development, the sixth proposition, has become a differentiator of institutions. While there has been a rise of emphasis on development across the board, there has become an increasing gap between those institutions taking educational development seriously and investing in it, centrally and locally, and those who have made token moves. There have been increasing numbers of courses for academics that include a substantial component on assessment, and increasing institutional initiatives to reform courses to use assessment more rationally. In some disciplines, particularly medicine, the notion of programmatic assessment is becoming embedded. This approach takes a program-wide view of assessment,

disaggregates assessment activities from individual course units and assures that all necessary outcomes for the course as a whole have been met (Van der Vleuten et al., 2012).

Least movement has probably occurred in the area of the final proposition. This is not surprising as it focused on the most intractable area for change: summative portrayals of assessment outcomes. Although a few universities in some countries (e.g. the UK) disregard grades awarded in first year for the purposes of determining grade point averages or honours status, there is still an unhealthy and invalid focus on grades. Grades are awarded without assuring learning outcomes, inadequate performance in one area is compensated for by good performance in another, and low grades in early parts of a course still disadvantage students even when they excel in exactly the same area later. Most importantly, the transcript provided for graduates at the end of a program communicates to external parties almost no useful information about what a student knows and can do.

We should not over-estimate the magnitude of the effects in any of the areas of the propositions. There is a considerable way to go, as studies in the European context also exemplify (Panadero et al., 2019; Pereira et al., 2017; Rodríguez-Gómez et al., 2016; Quesada-Serra et al., 2016) particularly in those institutions that have not realised sufficiently that their future is dependent to a significant extent on the quality of their courses not just their research.

Other major changes in assessment

Changes in the areas identified in the Propositions should be seen in the context of major changes taking place over the same time period. Two are worthy of mention. Probably the single greatest change in assessment policies in higher education has been in making them less rule-based and more principle-based. Old policies were often accretions of rules designed to address problems that occurred over the years. They often focused on the conduct of examinations, not assessment more generally. A principles-

based approach means that there is more flexibility in specific assessment procedures—assessment need not be the same across the whole institution or even across a Faculty—but all assessment processes must conform to the principles articulated in the policy. This, for example, may mean that rules requiring certain proportions of marks to be based on examinations are eliminated, and rules about the number of assessment tasks per course unit¹ ditched. Ironically, even where this has occurred, many academics still believe the rules still exist! The move towards a principles-based approach to assessment policy was not driven by the propositions, but by changes in university governance often to protect the institution from unwanted litigation from students.

The second major change in assessment over the past decade has been the rise of authentic assessment. Authentic assessment involves using assessment tasks and processes that represent the kinds of tasks and processes found in the world of practice. Many professional disciplines had already taken this step during the last century, but even there, tasks in foundational or academic subjects as part of a professional degree were often not authentic. This has changed a great deal and there are signs that there has been uptake in authentic assessment in courses that do not directly lead to a profession. While assignments are much more often likely to be authentic, tests and examinations have been more resistant to this trend. Also, the shift to authentic assessment varies substantially across countries. Countries where there has traditionally been great emphasis on tests and examinations show fewer signs of moves to authentic assessment, but there are indications even in countries like Chile that (over)emphasise multiple choice testing, that authentic assessment is being considered (Villarroel et al., 2019). In countries like Australia, the first wave of authentic assessment in all courses has passed and

institutions face the challenge of dealing with academics who resist it.

There have been, of course, changes in the wider field of assessment, but initiatives in assessment from the school sector have exerted very little influence on higher education in most countries. In the Australian context, changes in university assessment were little influenced by these and only in small ways by internationally-known scholars of assessment working within the country. This is common elsewhere.

Part 2. Where to from here?

Despite some important changes, there are considerable matters still to be addressed in assessment in most institutions. Some of these require further investigation before we can provide secure recommendations for change, but many others do not. An outsider would be astounded to discover how much practice which cannot be defended on the basis of any scholarship of assessment still occurs. While some of the changes to be made can be disputed, many other practices have become fixed simply as a matter of habit and inertia. Institutions need a willingness to accept the need to address some of these outstanding matters, and some internal political sensitivity to get them implemented is required (e.g. Deneen & Boud, 2014). Part 2 focuses firstly on unresolved issues that probably do not require further scholarly investigation, and secondly considers those that require further research and development before they can be fully implemented.

Completing the move to a standards-based framework for assessment

Most higher education systems in OECD countries have adopted a standards-based qualifications framework, at least at the highest level. This involves the specification of learning outcomes for all programs, and the use of assessment to assure that these outcomes have been met (Boud, 2014). This framework is often not translated fully into the specifics of

¹ The term course unit or unit is used throughout to refer to the individual subjects or modules of a degree program.

the courses offered. However, in some countries, the standards-based, outcomes-oriented approach has been embedded in legislation used to control universities (e.g. in Australia), in others it has been accepted without any enforcement. It is common, even in countries like Australia where the expectation is a very strong one and institutions accept the shift, that policies still embody assumptions of norm-referencing and grading incompatible with a standards-based approach. Even more common is for faculties and courses to lag well behind institutional policy shifts that have occurred, not primarily because the changes are disputed, but because there is no direct reason to change.

What is required to make assessment work in a standards-based framework? Everything revolves around having meaningful statements of learning outcomes at a suitable level of abstraction. Too many and they tend to become excessively behavioural, too few and they don't encompass the range of outcomes needed. A major trap is to over-emphasise the technical, knowledge features of learning at the expense of the utilisation of knowledge in terms of what students can do. An effective learning outcome should clearly represent what a student can do as a result of a particular course of study. A threshold standard is required to define what all students should be able to demonstrate, and additional levels can indicate achievements beyond this. Students who do not meet the threshold standard should not be allowed pass even if they do well in other parts of the course. Embedded in the learning outcome is of course a considerable body of knowledge, but the detailed knowledge need not be foregrounded, only what it will allow a student to do. If learning outcomes are written well, and often they will require at least three major iterations before they are sufficiently well composed, then assessment activities can be prompted by them.

There are two key levels of learning outcome: for the program as a whole, and for individual units, modules or subjects within it. Traditionally, the focus has been on the individual unit or subject, but increasingly the focus needs to be on the overall program outcomes. Whatever else is assessed, it is these

overall learning outcomes which must be the priority, for it is these that represent the qualification in which a student is enrolled; it is these which must be assured by summative assessment. The dilemma to be faced is whether assessments within specific course units will enable these outcomes to be assured. This can only happen if there is a clear mapping of unit learning outcomes on to course learning outcomes and an agreement that course/program learning outcomes always take priority.

A major difficulty for assessment are the legacy assessment recording systems used by each university, which have not been upgraded from an earlier era. They normally allow for grades (letters or numbers) to be recorded for each student on each task. For each unit a weighted average is determined and averages over units calculated as a grade point average. These various forms of averaging over different learning outcomes are incompatible with a standards-based approach. In such an approach, assessment needs to assure that each learning outcome has been met at the threshold level. Over-achievement on one cannot compensate for under-achievement on another. Averaging across different learning outcomes is meaningless as they represent, or should represent, fundamentally different achievements (Boud, 2017). What we therefore have is a dysfunctional system that undermines the principles on which assessment is now based. Change will only occur when it is possible for assessment record systems to record levels of achievement by learning outcome, without aggregation across different outcomes. This means that if a course unit or subject has three learning outcomes then three grades need to be recorded as it would be inappropriate to average across these.

Even once record keeping systems are fixed, there is still the problem of assuring learning outcomes. Having multiple grades for a single outcome does not address this problem. What needs to be identified is what counts as having attained the outcome? In addition, at what stage is it possible to say that a course (as distinct from a unit) learning outcome has been met? If we are to retain the practice of having assessment located within course units (rather

than at a course or program level), then decisions have to be made about what is the assessment event that assures each one. It is common to aid this process by defining each assessment event for each course unit in terms of whether it represents an introductory level of achievement, an intermediate level or a final assured level. These can be mapped across units to ensure that all outcomes have been addressed and that some are not over-assessed. (See, www.assuringlearning.com). The problem of assurance raises the question of whether introductory or intermediate assessments should ‘count’ at all for graduation purposes. They need to be undertaken as part of the learning process, and they may need to be recorded, but what is the justification for them to appear on a public document? Surely, all that is needed is a validated record of what the student is able to end up doing.

A much greater emphasis on the notion of assessment as assurance will be needed in the next decade of development. There are a number of initiatives that assist in addition to the curriculum mapping referred to in the last paragraph. The first is the disaggregation of teaching and learning from assessment. An initiative at Brunel University in London, for example, allows assessment to only occur in ‘assessment units’ where there is no teaching (<https://bilt.online/wp-content/uploads/2018/11/JPA-Practical-Guide-0918.pdf>). Teaching occurs in ‘teaching units’ in which only formative assessment occurs. ‘Assessment units’ typically cover multiple ‘teaching units’ so that the overall assessment load is reduced, and the focus of assessment is on the overall course learning outcomes. A second initiative mentioned earlier is the notion of programmatic assessment. It takes the idea of separating summative assessment from teaching and judges all assessment in terms of the extent to which it judges student performance against course learning outcomes (Van der Vleuten et al, 2012; www.bradford.ac.uk/pass/).

We should not uncritically accept the focus on standards and criteria and the centrality of learning outcomes. These have been an

important step in assessment reform, but they are not a panacea. While critique is occurring (e.g. Ajjawi et al., 2019), we need to be mindful of not taking a reactionary stance which would take us back to some idealised past which did not exist. The move to standards and learning outcomes was a necessary one in order to deal with the huge expansion of higher education and the accountability of public money. The discipline of focusing on learning outcomes is an important one, but we should be mindful that some of the administrative rigidities that have occurred alongside this move can be counterproductive.

Repositioning students as learners and producers

Probably the major *educational* challenge of higher education is that of shifting students’ identity from that of ‘student’ to that of ‘learner’. To be a student often means doing what is required—attend lectures, take tests, etc.—in other words follow the instructions of others in order to graduate. The passive role is reinforced by the practices of the university. Students are treated as a mass, decisions are made unilaterally about what they should learn, what they should do and how they are assessed. This is a profoundly unhelpful position for students to adopt if they want to become responsible graduates in the world with control over their own direction and their own learning. The ‘student’ identity does not equip them to be an effective practitioner of any kind in today’s society. To be a ‘learner’ is to be proactive, to take the initiative in deciding what to focus on and to judge one’s own performance in the process. For learners, other people can be resources to be used and to be learned from, not directors of their actions.

While there are many things needed in curriculum and pedagogy to address the passive positioning of students, we focus here on assessment. There are three directions to explore and research further and then implement change. The first is with respect to developing students’ evaluative judgement, the second, and related one is to how feedback is conceptualised and conducted. Thirdly, there is that of new forms of portrayal of achievements.

1. Developing students' evaluative judgement

It is not enough that a graduate has certain knowledge and can do certain things. Of course, when these relate to course learning outcomes, they need to be assured. However, more is needed for a graduate to be an effective person in the world. They need to know what they know and what they can do, and know what the limits are to this. All forms of work require practitioners to draw on their knowledge and capabilities. The graduate needs to determine how what they know and can do relates to the problem at hand, and what else they need to know and be able to do in order to address it. When employers criticise universities for not providing useful graduates, they tend not to be concerned about their knowledge and skills, but about their ability to effectively use these in new situations. A fundamental need for higher education (and indeed other sectors as well) is therefore for students to develop the capacity to judge their own learning and that of others, e.g. 'how will I know if this is good enough?', 'am I meeting the standard for this particular task?' This capacity has been referred to as evaluative judgement, which has been defined as, 'the capability to make decisions about the quality of work of self and others' (Tai et al., 2018). It is arguably more important than any particular learning outcomes that a course announces. It is needed both for learning within the course, and for the longer-term.

The development of evaluative judgement is likely to be domain specific and will need to be promoted across the curriculum. It requires considerable practice and involves two key elements: appreciation of what constitutes good work, and the making of judgements of about one's own work and that of others (Boud et al., 2018; Joughin et al., 2019). The former precedes the latter: without knowing what good work looks like, how can it be produced? While the development of evaluative judgement has built on literature from student self and peer assessment, it is not simply a reframing of that: it is a curriculum goal. No matter how poor a student may be about judging their own work or that of others, the aim is for improvement so that students

become less and less dependent on the judgement of teachers as their course progresses.

While we can presently establish the development of evaluative judgement as part of the curriculum, there are many questions to be answered before it can be fully implemented. These include: what are efficient ways of embedding it within courses that enhance learning outcomes? How much practice over how many iterations, over how many course units, over how many years are needed for its development? What is an appropriate balance between teacher judgements, self-judgements and peer-judgements? Does it require quite different implementation strategies for different outcomes and different disciplines? Part of this research has begun in one study in the Spanish context by Ibarra-Sáiz et al. (2020).

2. Reconceptualising and reframing feedback

Feedback is an important feature of learning in general and the development of evaluative judgement in particular. Recent years have witnessed a huge shift in how feedback is conceptualised in higher education, away from the teacher-centred notion of feedback as an input to students (comments on their work) to feedback as a process in which learners necessarily need to be engaged (Boud & Molly, 2013; Carless, 2015). This change has been so great that it has been termed a paradigm shift by Winstone and Carless (2019). The implications of this shift are profound as they influence every part of a course in which feedback is deployed. It also undermines the use of the term feed-forward used in some popular literature as it is superfluous and betrays a misunderstanding of feedback itself.

Feedback is being reframed as a learning-centred idea. Carless and Boud (2018) define it as "as a process through which learners make sense of information from various sources and use it to enhance their work or learning strategies" (p. 1316). The implication of this is that the default for feedback is not the provision of unsolicited information from

teachers to students on their assessed work, as a supplement to grades. Rather, it is a process which can be initiated by students to get information from others, not only from teachers, that they can then process to improve their work. Typical indicators of a learning-centred view are: students propose to others what kinds of comments they want on their work, students are expected to respond to comments with a plan for what they will do, evidence of effective feedback is found in the subsequent work of students. Feedback in this view is a kind of dialogue in which students necessarily play an active role. Examples of this new approach to feedback can be found in the new collection by Henderson, Ajjawi et al. (2019) and suggestions for how student feedback literacy can be built into the curriculum are given in Malecka et al. (2020).

3. *New portrayals of achievements*

The area in which students have been positioned most passively is that of final or summative assessment. Higher education not only has to help students learn more effectively, it also has a responsibility, not only for certifying what students have achieved, but enabling them to portray their achievements well. However, it needs to do so without inhibiting the very processes of independent action that it is seeking to foster. Digital credentials require that records be kept of exactly what a student can do, under what conditions, how that achievement has been judged according to what standards and criteria, and by whom. This credential can be inserted in, say, a LinkedIn profile and officially validated by the university by any person accessing the profile at the click of a button. Such credentials can be extended to not only represent all the learning outcomes met by students, but can be used by students to differentiate themselves from others through displaying their unique achievements (Jorre St Jorre et al., 2019).

Once a digital repository with information about what a student has done has been established, it can be curated by students to enable various forms of portrayal for different purposes: to highlight particular accomplishments to a prospective employer, or

to represent themselves in their professional field. Curation involves combining officially validated achievements with other artefacts including co-curricular accomplishments with a commentary by the graduate that places these in context (Clarke & Boud, 2018). All the building blocks for such assessment portrayals presently exist, but much research and development is needed to produce sufficiently enabling technologies and, importantly to facilitate students to find ways to record and present themselves in the light of developments in social media (Ajjawi et al., 2020).

All three of the issues discussed above have demonstrated considerable promise and may well transform assessment in the future. They each also require considerable further research as there are many features of them that are underdeveloped or unexplored.

Conclusion

While change in assessment seems elusive year to year, a wider perspective shows much is occurring. The challenge is to foster the uptake of the approaches identified above. The main requirement for this to happen is commitment, especially the commitment of institutions whose leadership already is aware that their sustainability involves attracting and meeting the changing needs of students. If this does not occur in a sufficiently rapid way, then some of these institutions themselves will be at risk.

References

- Ajjawi, R., Bearman, M., & Boud, D. (accepted for publication 6 October 2019). Performing standards: A critical perspective on the contemporary use of standards in assessment, *Teaching in Higher Education*. <https://doi.org/10.1080/13562517.2019.1678579>
- Ajjawi, R., Boud, D., & Marshall, D. (2020). Repositioning assessment-as-portrayal: what can we learn from celebrity and persona studies? In Bearman, M., Dawson, P., Ajjawi, R., Tai, J. & Boud, D. (Eds.) *Re-imagining university assessment in a digital world*. Springer.

- Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016) Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education*, 21(5), 545-556. <https://doi.org/10.1080/13562517.2016.1160217>
- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E. Boud, D., & Joughin, G. (2017). How university teachers design assessments: a cross disciplinary study. *Higher Education*, 74 (1), 49-64. <https://doi.org/10.1007/s10734-016-0027-7>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>
- Boud, D. (2007). Reframing assessment as if learning was important. In Boud, D. & Falchikov, N. (Eds.) *Rethinking Assessment for Higher Education: Learning for the Longer Term*. Routledge, 14-25. <https://doi.org/10.4324/9780203964309>
- Boud, D. (2014). Shifting views of assessment: from secret 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* (pp. 13-31). Edinburgh University Press. <https://doi.org/10.3366/edinburgh/9780748694549.003.0002>
- Boud, D. (2017). Standards-based assessment for an era of increasing transparency. In Carless, D., Bridges, S., Chan, C. and Glofcheski, R. (Eds). *Scaling up assessment for learning in higher education* (pp. 19-31). Springer. https://doi.org/10.1007/978-981-10-3045-1_2
- Boud, D., & Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Australian Learning and Teaching Council.
- Boud D., Ajjawi R., Dawson P., & Tai J. (Eds.). (2018). *Developing Evaluative Judgement in Higher Education: Assessment for Knowing and Producing Quality Work*. Routledge. <https://doi.org/10.4324/9781315109251>
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment and Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment and Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>
- Carless, D. (2015). *Excellence in university assessment: Learning from award winning practice*. Routledge. <https://doi.org/10.4324/9781315740621>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Clarke, J. L., & Boud, D. (2018). Refocusing portfolio assessment: curating for feedback and portrayal. *Innovations in Education and Teaching International*, 55(4), 479-486. <https://doi.org/10.1080/14703297.2016.1250664>
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment and Evaluation in Higher Education*, 39(5), 577-591. <https://doi.org/10.1080/02602938.2013.859654>
- Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.) (2019). *The Impact of Feedback in Higher Education: Improving assessment outcomes for learners*. Palgrave Macmillan <https://doi.org/10.1007/978-3-030-25112-3>
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2020). Developing student competence through peer assessment: the role of feedback, self-regulation and evaluative

- judgement. *Higher Education*. <https://doi.org/10.1007/s10734-019-00469-2>
- Jorre St Jorre, T., Boud, D., & Johnson, E. (2019) Assessment for distinctiveness: recognising diversity of accomplishments, *Studies in Higher Education*, <https://doi.org/10.1080/03075079.2019.1689385>
- Joughin, G., Boud, D., & Dawson, P. (2019). Threats to student evaluative judgement and their management, *Higher Education Research and Development*, 38(3), 537-549. <https://doi.org/10.1080/07294360.2018.1544227>
- Kift, S., Nelson, K., & Clarke, J. (2010). Transition pedagogy: A third generation approach to FYE - A case study of policy and practice for the higher education sector. *The International Journal of the First Year in Higher Education*, 1(1), 1-20. <https://doi.org/10.5204/intjfyhe.v1i1.13>
- Malecka, B., Boud, D., & Carless, D. (2020). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*. <https://doi.org/10.1080/13562517.2020.1754784>
- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, 44 (3), 379-397. <https://doi.org/10.1080/02602938.2018.1512553>
- Pereira, D., Niklasson, L., & Flores, M. A. (2017). Students' perceptions of assessment: a comparative analysis between Portugal and Sweden. *Higher Education*, 73(1), 153–173. <https://doi.org/10.1007/s10734-016-0005-0>
- Rodríguez-Gómez, G., Quesada-Serra, V., & Ibarra-Sáiz, M. S. (2016). Learning-oriented e-assessment: the effects of a training and guidance programme on lecturers' perceptions. *Assessment & Evaluation in Higher Education*, 41(1), 35–52. <https://doi.org/10.1080/02602938.2014.979132>
- Quesada-Serra, V., Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in Education and Teaching International*, 53(1), 48-59. <https://doi.org/10.1080/14703297.2014.930353>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, 76, 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- Van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, 34, 205-214. <https://doi.org/10.3109/0142159X.2012.652239>
- Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, 34, 535-557. <https://doi.org/10.1080/14703297.2018.1564882>
- Winstone, N., & Carless, D. (2019). *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. Routledge. <https://doi.org/10.4324/9781351115940>

Author / Autor

Boud, D. (david.boud@deakin.edu.au)  0000-0002-6883-2722

Alfred Deakin Professor and Director of the *Centre for Research in Assessment and Digital Learning* (CRADLE) at Deakin University (Melbourne) and Emeritus Professor at the University of Technology in Sydney. He is also a Professor of Work and Learning at Middlesex University (London). He has written several publications on teaching, learning and assessment in higher and professional education. His current work focuses on the areas of assessment for learning in higher education, academic formation and learning in the workplace. He is one of the most highly cited scholars worldwide in the field of higher education. He has been a pioneer in the development learning-centred approaches to assessment across the disciplines focused on learning, especially in building assessment skills for the long-term learning (*Developing Evaluative Judgement in Higher Education*, Routledge 2018) and designing new approaches to feedback (*Feedback in Higher and Professional Education*, Routledge, 2013). *Re-imagining University Assessment in a Digital World*, by the publisher Springer, will come out during 2020.



Revista ELectrónica de Investigación y EValuación Educativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]



Esta obra tiene [licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).
This work is under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).