Perils in the meticulous specification of goals and assessment criteria

D. Royce Sadler
Griffith Institute for Higher Education, Griffith University

Abstract The implementation of assessment policies can sometimes achieve almost the reverse of what was originally intended. Studies published in a special issue of *Assessment in Education* provide the basis for this bird’s eye view of particular practices in post-compulsory education and training in the UK. Problems identified include the trend towards extremely fine grained approaches to measuring achievement, and a strong social drive to help learners, some with personal histories of spectacular ‘unsuccess’, obtain a qualification. The assessment practices focus on methods of getting students through – often at the expense of what it really means to learn. The present situation is unlikely to be turned around unless there is clear articulation of appropriate conceptual foundations for both pedagogy and assessment, followed by practices that are consistent with them. Assessment boards could play a critical leadership role in bringing this reform about.

The articles in this Special Issue on Assessment in Post-secondary Education and Training provide a valuable and interesting distillation of the unfolding agenda of how assessment is being conceptualised and implemented in the UK, and of how teachers and students construct and experience assessment. This paper is a personal response, written from the perspective of an academic who has thought and written about assessment for many years. A thoroughgoing comparative analysis of post-compulsory education and training in the UK and Australia would need more experience and appreciation of the history, structure and subtleties of both systems than I have. My work has focused more on schools (and increasingly the higher education sector) than specifically on post-compulsory education and training. It has also been primarily at the conceptual rather than the empirical level. Nevertheless, certain commonalities in assessment themes,
theory and practice run across all educational sectors. I limit most of my observations to these commonalities.

I found the articles in this Issue fascinating to read, and recorded many of my reactions as they occurred. I then organised my thoughts as a bird’s eye view that expresses some things in slightly different language from that used by the various authors, trying not to be opaque or idiosyncratic. A number of core issues characterise the assessment enterprise in all structured learning contexts, and it is important not to divert attention away from them, or to distort what the contributors have said. This paper does not necessarily contain a great deal that is original, except perhaps in its last third. At least one of the authors has at least touched on one or more of the aspects I have singled out for mention.

I observed that many of the terms we use in discourse on assessment and grading are used loosely. By this I mean that we do not always clarify the several meanings a given term may take even in a given context, neither do we necessarily distinguish various terms from one another when they occur in different contexts. For example, the terms ‘criteria’ and ‘standard’ are often used interchangeably. Criteria can mean properties or characteristics. The criterion of readability, for example, is a property that can be used for judging the suitability of a particular piece of text for some purpose. Criteria can also refer to a fixed level above which something qualifies for a particular designation, as when one ‘meets the criteria’ for an award or a grade. In many contexts, the intended meaning is clear; in others, it is not, and needs to be stipulated. Similarly, ‘learning’ can refer to what a person is doing – the purposeful activity. It can also mean a quantum of something that has been learned.

As I read the articles, I found myself constantly identifying, and adjusting to, intended nuances of meaning. Even in such common and apparently transparent phrases as ‘assessment of learning’ and ‘assessment for learning’, the intended distinction is conveyed every bit as much by two different meanings of ‘learning’ – basically, product and process – as by the two prepositions. ‘Assessment as learning’, which I encountered for the first time in Torrance’s (2007) insightful lead article, defines this characterisation of assessment as “the displacement of learning (i.e. understanding) by
procedural compliance: i.e. achievement without understanding” [italics in the original]. It refers to assessment that masquerades as, or substitutes for, learning itself. Torrance’s position is that this approach to operationalising ‘assessment’ has thoroughly permeated the philosophy and practice of teachers, colleges and the broader post-compulsory system, to the point where it has become normalised. Despite the greatest of goodwill towards, and care for, the learner, the implementation of ‘assessment as learning’ manages to cheat not only the learner but others as well. Assessment as learning has a deservedly pejorative ring to it.

The phrase is an apt one, however, given the dominant theme of the papers in this Issue. The backgrounds to many of the situations described provide explanations for sets of practices that have become established in a number of areas in post-compulsory education and training. These practices have been developed incrementally out of what were originally pretty sound ideas. Given their current momentum, these practices will continue essentially unaltered into the future unless a deliberate attempt is made to define different directions, followed by concerted effort to promote and pursue those alternative directions.

In broad terms, the explanations for current implementations involve acknowledgment of the prior experiences of learners, high levels of external societal demand for qualifications and certificates, adoption of criterion-based marking and assessment, and the commitment of teachers to do their very best for students. Each of these, taken at face value, should be a positive contributor to high quality teaching and learning. Having highly motivated teachers, taking students’ previous achievements and attitudes into account, and using criterion-based assessment are usually considered desirable factors. Furthermore, a qualifications system set up to certify achievements provides external goals towards which teachers and learners can strive together. All of these factors could be regarded as providing promising conditions for success.

The substance of the articles reveals that the real state of affairs is, in many supposedly ‘learning’ situations, a long way from this. A problematic mix of mere compliance, and of going ever-so-systematically through all the steps, actually turns out to be instrumental in subverting the goal of assessment.
At the same time, it distorts both the learning itself, and teachers’ and students’ understandings of what learning entails. Many of the authors raised explicit concerns about the discrepancy between promise (as encapsulated in the ideals and design of good assessment processes) and performance (as evidenced in very mixed outcomes, ranging from true learning to dismal minimalism). Together, there appears to be a significant undermining of many of the goals of high quality education and training.

In many of the contexts described, learners entered post-compulsory studies with a history of educational failure. They experienced alienation, disengagement, disillusionment and powerlessness. The indications point to these having been developed during schooling years. The schooling process may have contributed to some extent, but this does not lay blame automatically at the feet of schools and teachers. Multiple factors are always at work, and the situation may well have been worse without the efforts of schools and teachers. But it is important to highlight the plight of many of the learners studied. Irrespective of root cause, their serious psychological handicap was immediately recognised on entry to post-compulsory education. In general, this need was not ignored by teachers, but identified and addressed systematically. Many teachers were inspired to work intensively with and for these learners so they could overcome obstacles and gain a qualification.

External credentials have become critically important in today’s societies, but there is a trap. Extremely strong societal valuing of credentials and awards can mean that the possession of an award, sometimes written into legislation or labour agreements, can come to rule the day. Sitting behind any recognised award should be a substantial body of integrated knowledge and skill into which a person can tap at will in order to engage in a variety of productive work situations. These awards, in the UK as in many other countries, are structured from collections of components. By segmenting the whole into manageable units, progress through the learning process can be facilitated for systems, colleges, teachers and students. But of course, each component can be, and frequently is, broken down into smaller units, and each of these into smaller units again. It is only a small step from this to fully atomised teaching and assessment.
The further this decomposition progresses, the harder it is to make the bits work together as a coherent learning experience that prepares learners to operate in intelligent and flexible ways. The logic of this phenomenon is obvious: if you break something into pieces, whatever originally held it together has to be either supplied or satisfactorily substituted if the sense of the whole is to be restored. The relationships and dependencies that should characterise a good sequence of learning units tend to get lost when course documentation, teaching, and assessment all focus on units, modules, or even specific tasks as self-contained, hermetically sealed elements.

Therein lies the problem referred to by many of the authors. The assessment criteria have been reduced to pea-sized bits to be swallowed one at a time – and for each bit, once only. The practice of coaching over the line has been legitimated by recording each microscopic ‘outcome’ rather than by a focus on the learning that has taken place. This trend has been accentuated, if not driven, by an extreme interpretation of what criterion-based assessment implies.

The matter is compounded by the manner in which many teachers provide the substantial encouragement many learners need. Teachers reward them with marks for effort, or for improvement. Conversely, they may penalise them for not submitting their work on time. These actions may appear not only understandable but also laudable because they shape behaviour or attitude. The appearance is, however, superficial. None of these, and a number of similar components that are often incorporated into marking practices, is a true ‘achievement’ variable. Using marks to represent them as part of a grade or credential muddies what the credential stands for. In the long run, it serves to diminish the very point of having the qualification. The practice further serves to warp the student’s understanding of what it means to learn. Getting through, at all costs, is all that matters. The more teachers do it, the more students expect it. So well accepted has the accumulation of fragments become that the collection provides the functional definition of knowledge, skill or competence. Those terms thereby lose their unique and original significance. To the extent that this occurs, it becomes the new orthodoxy of practice – self-reinforcing, self-legitimating and perceived as quite unproblematic.
Not far below the surface is another positive educational principle that has had its original intention hijacked to another purpose. I refer to the concept of ‘scaffolding’ the learning. Properly understood, it means providing appropriate supports during learning so that learners are better able to bridge the gap between what they bring to the learning task, and where they need to be to achieve a deep level of learning. For many teachers, the scaffolding has become so elaborate, and the level of assistance so comprehensive, that the learner cannot help but ‘succeed’. But scaffolding is supposed to be a temporary arrangement that supports the building process. It loses both its utility and any aesthetic value it may have as soon as the building is finished. After the scaffolding has done its job, it is dismantled. The building then stands on its merits. This brings me to a definition of learning that I am no longer confident is widely shared.

For my money, learners can be said to have learned something when three conditions are satisfied. They must be able to do, on demand, something they could not do before. They have to be able to do it independently of particular others, those others being primarily the teacher and members of a learning group (if any). And they must be able to do it well. Assessment of learning should be directed towards gathering evidence for drawing inferences about capability under these conditions, not the scaffolded conditions.

Coaching a person through doing something once certainly does show that the person is capable of doing it, but only in a very restricted sense. It shows that it was technically possible for the person to do it – given the same conditions under which the demonstration took place, which means complete with supports and coaching. In most regular learning contexts, educators should be only rarely interested in knowing that a specific act has been carried out, just once, by a student. All such a completion can show is that it is literally not impossible for the person to do it.

In theory, we are interested in capability, interpreted as the prospect of successful repeat performances in a context of task variance. Of course, the situation is considerably different for elite athletes, for whom training is often directed towards a singular goal, and ‘achievement’ ultimately is assessed (and remembered) in relation to a superb one-off performance.
that breaks an established record or is better than that of any other athlete in the world at the time. Educators in general, though, should be working with conceptions of both learning and the certification of learning as these are represented by learner outcomes that are reproducible, at a satisfactory standard.

The issue of demonstrating through an (assisted) one-off event needs to be taken even further. The phrase ‘being able to do it, in the general case’ implies a capacity to complete tasks that are variable within reasonable limits, but are still in the same class. A step up again is for the learner to have internalised the basic principles to the point where recognition of class membership is important as well. The person has to understand the essential similarities and differences well enough to know how to tackle, not the same task over and over, but an ‘essentially similar’ one. This recognition process cannot be called recognition if it is evoked little by little by a sympathetic teacher who asks leading questions with minute steps all the way. In some circumstances, that activity may be useful as an initial teaching strategy, but it cannot stop there. When independent recognition occurs, the person is at least at the starting blocks for mixing routinised knowledge, which can be applied with some confidence, with a modicum of tentative or experimental knowledge (with due respect to the attendant risks and cautions), so as to ‘do’ previously unseen tasks.

As soon as the learner is able to codify the recognition and solution processes as propositional knowledge, the stage is set for a different but powerful form of assessment. It is to ask the learner either about how and why they were able to solve non-standard or unseen actual cases, or to ask them to articulate likely-to-be-successful approaches for cases that are still at the abstract or hypothetical level. The gap between the previously-done and the to-be-done can be bridged by drawing from the person’s abstract knowledge base. That is why it makes sense so often to ask students to describe, not necessarily in writing but in understandable words or images, how solution paths could be constructed for both hypothetical and concrete cases, without actually constructing the solutions. This process provides a sound basis for inferences about generalisable knowledge, both ‘knowing that’ and ‘knowing how’. Although teachers have traditionally used these
approaches to assess a learner’s mastery over a domain of knowledge, this
does not amount to a claim that propositional knowledge is, by itself, either
necessary or sufficient. In many contexts, however, it does provide a
satisfactory surrogate for extensive real-life demonstrations of capability.

**Conclusion.** Although I read though the articles in this Special Issue with a
good deal of interest and some puzzlement, my reading was tinged with a
note of despair. In my heart of hearts, I believe that a blinkered
conceptualisation of curriculum, the strong trend towards fine-grained
prescription, atomised assessment, the accumulation of little ‘credits’ like
grains of sand, and intensive coaching towards short-term objectives, are a
long call from the production of truly integrated knowledge and skill. In
Australian post-compulsory education and training, the situation is not too
dissimilar. Too often the focus has been on a multitude of discrete
competencies, rather than on competence. The slight difference in spelling
obscures a world of difference between these two terms in respect of their
implications for practice. Among other things, competence involves being
able to orchestrate a set of competencies.

There is a harsh anomaly in all of this. Many, many teachers are working
sensitively, energetically, caringly, supportingly, and conscientiously for
their students. In the process, they may be inadvertently doing individual
students and the enterprise as a whole a long-term disservice because of all
the fragmentation. The onus is on those at higher levels in the respective
systems to analyse and evaluate the situation, and then take steps to
engage in fundamental reform. In the initial stages, this should be primarily
in conceptualising, clarifying and disseminating the fundamentals. Only later
should it engage in devising extensive and consistent changes to practice. If
assessment authorities and examination boards were to take the lead in
setting the agenda to redefine learning, competence, assessment,
credentials, and awards, an ‘assessment pull’ could help turn the enterprise
around over a period of time, given sufficient resolve and courage, and an
enlightened political environment.

**Reference**
assessment criteria and feedback in post-secondary education and training can come
to dominate learning. *Assessment in Education, XXXXXXXXXX*
Notes on the contributor

D. Royce Sadler is a Professor of Higher Education at Griffith University, Brisbane. His main research interests are in assessment of student learning, achievement standards, grading policies, and how assessment can be used to improve learning. Contact details: Griffith Institute for Higher Education, Mt Gravatt Campus, Griffith University, Brisbane, QLD 4111, Australia. Email: r.sadler@griffith.edu.au