

Assessing student learning: Why reform is overdue

Geoff N Masters

Australian Council for Educational Research

Advances in our understanding of human learning require new approaches to assessing and monitoring student learning.

Much assessment thinking has changed little over the past fifty years. The field continues to be dominated by twentieth century introductory textbook concepts, including such dichotomies as formative versus summative assessment, criterion-referenced versus norm-referenced testing, quantitative versus qualitative assessment, informal versus formal assessment – distinctions that often hamper rather than promote clear thinking about assessment.

Assessment practice also has changed little over this period. Traditional, high-stakes examinations continue to dominate what is taught and learnt in many of our schools and universities. Greater use is now being made of promising new technologies, including banks of online assessment tasks, computer adaptive tests and technology-based assessments of ‘new’ life skills and attributes. However, while emerging technologies are capable of providing more innovative and informative explorations of student learning, much electronic assessment remains pedestrian and underpinned by traditional assessment thinking.

At the same time, progress in our understanding of learning itself is challenging long-held assumptions and pointing to the need for a paradigm shift in assessment theory and practice.

For example, substantial progress has been made in our understanding of human capacity for learning. It once was believed that individuals differed significantly in their capacity to learn. But research in neuroscience has shown how the plasticity of the brain enables almost all individuals to learn throughout the lifespan. This finding parallels the educational conclusion that, although students are at different points in their learning and are progressing at different rates, almost all students are capable of successful learning if motivated and

if provided with appropriate learning opportunities and support.

Research also is making clear the enormous variability in students’ levels of achievement and progress. Children begin school with very different social, cognitive, psychomotor and language development. Many of these differences do not disappear. In any given year of primary school, differences in reading and mathematics achievement are the equivalent of five or six years of school. And in some areas of learning and development, variability appears to increase across the school grades.

We also know that, in mixed-ability classrooms, students learn best when provided with learning opportunities matched to their varying interests and progress. Learning is maximised when tasks are targeted just beyond individuals’ current levels of attainment – in the region where success is possible, but often only with scaffolding and support.¹

An implication of these observations is that educational assessment is best conceptualised as a process of discovering where learners are in their learning and development. Although it is common to refer to the ‘multiple purposes’ of assessment, assessment has only *one* fundamental purpose: to establish where learners are in their progress at the time of the assessment. This information can then be interpreted and used in a variety of ways. For example, students’ achievements can be interpreted by reference to the performances of other students nationally or internationally, by reference to achievement expectations or standards, or by reference to past performances to study trends or growth over time. The results of assessments can be used to inform starting points for teaching, to evaluate the effectiveness of educational programs and interventions or to award qualifications. For teaching purposes, it sometimes is desirable to obtain more detailed information to diagnose specific student misunderstandings or errors, but once again, the single underlying purpose is to discover where learners are in their learning. Much unnecessary complexity has been introduced into the assessment literature through failure to recognise and begin with this simple truth.

The process of establishing where students are in their learning depends on a thorough understanding of the learning terrain through which they are progressing: typical paths of development; sequences in which understandings normally are established; and side-tracks in the form of common errors, learning difficulties and misunderstandings. Assessment as the discovery of where students are in their learning requires much more than familiarity with the intended curriculum. It depends on expert understanding of how learning occurs in a domain – a reference ‘map’ that is built from research and knowledge about learning itself.ⁱⁱ

Essential to this approach to assessment is an appreciation of learning as ongoing progress. At the heart of all educational effort is the intention of student growth, development or improvement. Rather than being limited to specific courses, semesters or years of school, the progress that students make usually occurs incrementally over extended periods of time. For example, in areas such as reading, mathematics and science, progress typically occurs across the entire period of schooling. The role of assessment should be to establish where students are on these long-term continua of learning and what progress they are making over time.

To establish where students are in their learning, evidence is required, usually in the form of observed performances on classroom activities or assigned assessment tasks. However, individual tasks are rarely, if ever, of intrinsic importance. Students may never again have to read and answer questions about the particular piece of text or solve the particular mathematics problems used in an assessment. Specific tasks are merely convenient but interchangeable vehicles for collecting evidence about what is really of interest – a student’s underlying reading ability, for example, or level of achievement in an area of mathematics. And establishing where students are in their learning always involves an on-balance inference with an accompanying degree of uncertainty.

This conceptualisation of assessment stands in stark contrast to the traditional use of assessment to determine how much of what a teacher has taught each student has successfully learnt. Traditional assessments are made not in relation to an understanding of long-term learning progress, but in relation to a specific corpus of taught content.

The onus is on students to learn this content and the role of assessment – whether during or upon completion of a course – is to judge how well they have done this. Conclusions about ‘how much’ students have learnt commonly are expressed as percentages, which may then be converted to grades to convey the extent of each student’s success (or failure).

Under traditional approaches, it is common to treat ‘curriculum, teaching and assessment’ as separate activities. The role of teachers is to teach the curriculum, the role of students is to learn, and the role of assessment is to judge how much of the taught content students have learnt. By contrast, a view of assessment as professional investigation sees assessment as an integral part of good pedagogy. This view is consistent with the role of assessment in other professional work – for example in medicine and psychology – where the purpose is not so much to judge as to understand for the purpose of making informed decisions.

Research into learning highlights the need for investigative approaches to assessment. Learning is rarely, if ever, a process of passively taking in and storing new information. Even from a very young age, learning is a process of actively trying to make sense of the world. Learners interpret what they see and hear in terms of what they already know. They construct their own mental models and understandings which are sometimes inaccurate or only partially correct. And it is clear that misconceptions, if not identified and addressed, can be significant obstacles to further learning.ⁱⁱⁱ

Research also shows that students sometimes can succeed on traditional forms of assessment while holding fundamental misconceptions. For example, physics students can sometimes recall formulae and substitute numerical values correctly to answer examination questions while holding fundamental misunderstandings about relationships between force and motion.

Studies comparing experts and novices in various fields show that what distinguishes experts from novices is not only extensive knowledge of a field, but also the frames of reference that experts have for organising and making sense of that knowledge. Experts have deep understandings of concepts, principles and big ideas in a field which allow them to see patterns in information

and to transfer their knowledge to new and unseen contexts.

The implications of these research findings are that educational assessments must do more than establish whether students can reproduce what they have been taught, and teachers must be more than deliverers of curriculum content and judges of student success. The investigative process of establishing where students are in their learning must include an exploration of students' understandings of important concepts and principles. An appreciation of learners' own mental models and misunderstandings can provide important starting points for teaching (ie, assessments *for* learning). Assessments of factual and procedural knowledge will continue to be important, but perhaps more important in the future will be the assessment of students' abilities to organise and use this knowledge and to apply their understandings to the solution of complex, real-world problems.

In the past, assessment methods often have been more concerned with judging success and making reliable and fair comparisons of student performances than with investigating and understanding student learning. And the desire for large-scale implementation under standardised conditions, with a quick turnaround of results, often has resulted in assessments requiring only that students reproduce what they have been taught through the provision of 'correct' answers.

Some educators have reacted against assessments of this kind by arguing that 'authentic', *in situ* assessments are always preferable to assessments based on specially-designed assessment tasks, or that 'school-based' assessments made by classroom teachers are always preferable to externally-developed assessments. But these are over-reactions. When the purpose of assessment is to explore and understand where students are in their learning, there must be a willingness to use the methods best able to provide this information, whatever form they take.

Day-to-day observations made by classroom teachers generally provide the richest information for establishing where students are in their learning. Ideally, teachers would have intimate and precise knowledge of each student's progress and learning needs and would use that knowledge to

personalise and focus their teaching efforts, often by grouping students with similar needs.^{iv} As noted already, assessments of this kind depend on expert understandings of the relevant learning domain as well as professional skill in exploring learning progress.

Advances in technology are making it possible to incorporate professional knowledge of this kind into more sophisticated tools for investigating learning. Rather than testing only factual and procedural knowledge, these tools explore student thinking, including by testing hypotheses about misunderstandings and gaps in an individual's learning. Intelligent forms of assessment in the future will be less concerned with judging how much a student has learnt and more concerned with diagnosing and understanding the details of an individual's learning.

Research in neuroscience and cognitive psychology also is revealing the important role of *emotions* in learning.^v People are more likely to learn and to remember if intrinsically motivated and emotionally engaged. In classroom settings, learning is promoted by 'learning cultures' in which all students are expected to learn successfully, are highly engaged and feel safe and supported in their learning. Conversely, negative emotions such as stress and fear of failure have been shown to impede learning and memory. In classroom settings, these emotions can be the result of 'performance cultures' in which learning is extrinsically motivated and students compete with each other for success.^{vi}

Other research has shown the importance of positive attitudes and beliefs about learning. Learners are more likely to learn successfully if they believe that they are capable of learning – in other words, if they have positive views of themselves as learners. They also must believe that effort will result in success. Effective learners are more likely to monitor their own learning, to recognise what they do not know and to be proactive in seeking out what they need to make further progress. Learners are assisted in these processes by relevant and timely feedback that guides action and enables them to see the progress they are making over time.

These research findings relating to emotions, attitudes and beliefs have implications for how

assessments of learning are conducted and how the results of assessments are reported and used.

Some forms of assessment promote ‘performance’ rather than ‘learning’ cultures. For example, one-off, end-of-course examinations usually are designed to judge and compare students on the amount of course content they have learnt – often for the purposes of ranking and selecting students for the next phase of education – rather than to monitor and understand learning progress. In such assessments, learning can be driven more by external pressure for results than by curiosity and intrinsic motivation. And this pressure often distorts teaching and learning by encouraging cramming and creating unacceptable levels of stress for students and their families.

The paradigm shift now required in assessment is from judging how much of a body of taught content students have successfully learnt to establishing where students are in their long-term learning and what progress they are making over time.

For this reason, one-off, high-stakes assessment events probably have a limited future in the assessment of student learning. In some contexts, there will continue to be a need to ensure that minimum performance standards have been met, but such assessments could be undertaken when learners feel ready to be assessed rather than in a single assessment event.

There are significant implications, too, for methods of reporting and monitoring student learning. Traditional reporting methods, such as percentages and grades, are more consistent with ‘performance’ than ‘learning’ cultures. Percentages and grades are used to describe how much of a body of taught content students have learnt. But these reporting methods are incapable of showing learning progress, and indeed usually mask progress. A student who receives a ‘D’ year after year is given no sense of the progress they are actually making. And worse, they are likely to infer from this outmoded method of reporting that there is something stable about their capacity to learn: they are a ‘D’ student.

It sometimes is argued that students and parents ‘understand’ A to E grades; but they do not because course grades usually do not represent consistent, interpretable levels of achievement. Grading is more appropriate for describing the quality of

agricultural produce or the products of industrial manufacturing than for describing learning. The educational challenge is to develop ways of reporting that show where students are in their long-term learning, what progress they are making (ie, assessments *of* learning) and what might be done to support further learning.

Finally, the uses to which assessments are put also can encourage ‘performance’ rather than ‘learning’ cultures. Assessments conducted to understand and promote student learning can be undermined and distorted when the results of those assessments are then used for other, unintended purposes. For example, external attempts to use test results to drive performance inevitably change classroom teachers’ attitudes and behaviours. There is growing evidence that the linking of rewards and sanctions to test results not only fails to produce the desired improvements, but also results in a range of responses that are inconsistent with what we now know about effective teaching and learning.^{vii}

-
- i Vygotsky, L (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
 - ii Pellegrino, JW, Chudowsky, N, and Glaser, R (2001). *Knowing what students know: The science and design of educational assessment*, Washington, DC: National Academy Press.
 - iii Bransford, JD, Brown, AL, & Cocking, RR (2000). *How people learn: Brain, mind, experience and school: Expanded Edition*. Washington, DC: National Research Council.
 - iv Fullan, M, Hill, PW & Crevola, C (2006). *Breakthrough*. Thousand Oaks, CA: Corwin Press.
 - v Organisation for Economic Cooperation and Development (2007). *Understanding the brain: The birth of a learning science*. Paris: OECD.
 - vi Dweck, CS (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
 - vii Hout, M & Elliott, S.W. (Eds.) (2011). *Incentives and test-based accountability in education*. Washington, DC: The National Academies Press.