

Re-assessing assessment

Centre for Education Statistics and Evaluation



Publication and contact details

Re-assessing Assessment

Centre for Education Statistics and Evaluation
Sydney, NSW
September 2015

For more information about this report, please contact:

Centre for Education Statistics and Evaluation
Department of Education
GPO Box 33
SYDNEY NSW 2001

Email: cese@det.nsw.edu.au

Telephone: +61 2 9561 1211

Web: www.cese.nsw.gov.au

This report can be accessed online at:

www.cese.nsw.gov.au

Contents

Chapter 1:	Introduction	3
Chapter 2:	What is effective assessment?	4
	Formative versus summative assessment	
	Integrating assessment with teaching and learning	
	Quality assessment systems	
	Equity and assessment	
	Effective assessment tools	
	Elements of effective assessment	
Chapter 3:	Assessment tools and systems	8
	Portfolio assessment	
	Criterion-referenced assessment	
	Computer-administered testing	
	Use of technology to engage and assess	
	Diagnostic feedback	
	Student learning accounts	
Chapter 4:	Examples of local and international practice	11
	Europe	
	Asia	
	United States	
	New Zealand	
	Australia	
	NSW Higher School Certificate	
Chapter 5:	Conclusion	15
References		16

1. Introduction

What gives assessment a bad name? What is effective assessment? And what innovative tools are making assessment more effective? This paper examines developments in assessment around the world, and highlights cases of innovation and best practice.

We measure what we value. In education, assessment signals to learners the knowledge that is valued. Educational assessment has undergone much change in recent years. This is both as a result of conceptual shifts in thinking about the nature of assessment and its integration with teaching and learning, but also the increasing role that digital technologies play in education, and the opportunities technology affords for innovation. Assessment is a powerful strategic tool for educators, but it can be done badly. Research shows both positive and negative effects of assessment in improving student learning, regardless of whether assessment is occurring at a systemic level at a school level, or within classrooms. The OECD notes that ‘getting assessment right will be one of the most important priorities for education systems in OECD countries over the near term’ (Looney 2009, p. 7.)

This paper considers:

1. what is effective assessment
2. tools and systems that can be used in educational assessment
3. examples of assessment practice both locally and internationally.

It draws on research evidence about the nature and importance of effective assessment. The paper ends with a series of questions that attempt to explore implications of an altered approach to assessment in the Australian context.

2. What is effective assessment?

Assessment is a broad term that can describe formal, high stakes examinations (like the HSC); standardised diagnostic tools (NAPLAN); class-based quizzes and tests; or the informal questions, teacher judgements, and observations that occur in routine classroom interactions.

Assessment in one form or another, has always been a feature of education. Nonetheless, despite the long history of educational assessment, it can be difficult to define exactly what 'effective' assessment is. This is because there are many different views on the nature and purposes of assessment.

The OECD (2015) states that assessment can have a multitude of purposes such as:

- contributing to improving student outcomes and the quality and equity of the education system
- providing stakeholders at both the system and school level (student, parents, teachers and policy makers) with information on what students know and should know and steps for further learning and improvement
- enabling comparability across the education system and across schools.

On the other hand, Masters (2013) asserts that all assessments have the same fundamental purpose - 'to establish where learners are in their learning at the time of assessment' - and that the difference is the varying use to which assessment information can be put.

Generally speaking, assessment in a contemporary education setting can be understood as:

...the broad name for the collection and evaluation of evidence of a student's learning. It is integral to teaching and learning and has multiple purposes. Assessment can enhance student engagement and motivation, particularly when it incorporates interaction with teachers, other students and a range of resources. (NSW BOSTES n.d.)

Formative versus summative assessment

Assessment is often distinguished as being either 'formative' or 'summative' (Scriven 1967); terminology that delineates goals and purposes of assessment.

Formative assessment is a term for the variety of methods teachers use to monitor student learning and to identify concepts that students are struggling to understand or skills they are having difficulty acquiring. In other words, it is a method of assessment for learning. Formative assessment is often characterised as being low stakes and may include methods such as essays, projects, lab work or journal entries.

Black and Wiliam (1998) analysed over 250 studies and found that formative assessment leads to significant learning gains. The formative assessment experiments they studied produced a typical effect size of between 0.4 and 0.7 which is a larger effect size than most education interventions. They also found that improved formative assessment helps 'low attainers' more than other students and so reduces inequities in attainment whilst also raising attainment overall.

Summative assessment, on the other hand, is typically used to evaluate student learning, skill acquisition, and academic achievement at the conclusion of a defined instructional period. It is an assessment of learning. Summative assessment is often considered high stakes. Summative assessment methods may include exams and quizzes.

Summative assessment may be delivered using standardised tests. Standardised tests are used to assess a student's understanding of skills and knowledge in comparison to other students. Standardised tests are usually constructed against clearly defined instructional or curriculum standards. They are objective and able to provide accurate comparisons between students and against standards.

Integrating assessment with teaching and learning

One of the greatest shifts to have occurred in educational assessment theory and practice in recent years is recognition that assessment is most effective when it is integrated with teaching and learning.

A major cause of resistance to assessment has been the use of standardised tests for high stakes accountability (usually summative) purposes. In some parts of the world – notably the US, but not only there – high-stakes standardised tests have been seen to reduce assessment to a series of multiple choice questions which may not allow for any in-depth exploration of what a student knows or can do and which are blamed for perceived poor practices such as ‘teaching to the test’ and ‘narrowing the curriculum’.

Irrespective of the standardised and credentialing testing regimes in some places, assessment is becoming understood as an essential tool for teachers who need better information about their students’ learning. Only with effective assessment can teachers know if learning is taking place. Evidence shows that good feedback to students yields greater learning gains (Hattie and Timperley 2007; Hattie 2009). Teachers need to assess to capture information that can be shared with the student as feedback, and can inform the teaching that ensues (formative assessment).

Like a doctor trying to identify what treatment patients need to improve their health, teachers need to identify what teaching their students need now to improve their learning. (Goss et al 2015)

Evidence shows that the most effective forms of feedback are feedback about a student’s process or effort, and feedback that encourages students’ self-regulation (CESE 2014). Data from PISA 2009 showed that Australian students whose teachers discussed students’ work after they had finished assignments in most or all classes, performed better than students whose teacher did so hardly ever or only in some cases (CESE 2014).

Hattie (2009) in his synthesis of meta-analyses relating to educational achievement found that the average effect size of feedback was 0.73. This is a higher effect size than other influences on achievement such as the quality of teaching, school effects or classroom cohesion. Hattie maintains that all assessment should be primarily concerned with providing teachers and/or students with feedback, and thus should address the following three questions:

- what are the learning intentions?
- how am I going?
- where to next?

Quality assessment systems

Education systems place different emphases on different types of assessment, usually aiming for a balance between formative and summative approaches. Clarke (2012) states that the most important features are ensuring that: the enabling context is supportive of the assessment (i.e. assessment activities are in line with broader legislative and policy frameworks); there is system alignment (i.e. there is a connection between assessment, the broader curriculum and pre- and in-service teacher training); and there is assessment quality (both in terms of reliability and validity).

Masters (2013) stresses the importance of having an overall assessment system. He says that rather than beginning with the belief that one form of assessment is intrinsically superior to another, it should instead be recognised that the most appropriate assessment method(s) are those that provide the most practicable, valid and reliable information about where learners are in their learning. He suggests a learning assessment system that encompasses the following five principles:

- Assessments should be guided by an empirically based understanding of the relevant learning domain
- Assessment methods should be selected for their ability to provide useful information about where students are in their learning within the domain
- Responses to, or performance on, assessment tasks should be recorded using one or more 'task' rubrics
- Available assessment evidence should be used to draw a conclusion about where learners are in their progress within the learning domain
- Feedback and reports of assessments should show where learners are in their learning at the time of assessment and, ideally, what progress they have made over time.

In their report for the Grattan Institute, Goss et al (2015 p. 23) advocate a balanced system of assessment which should: assess both knowledge and understanding; assess what we want students to learn; and test and track progress across the curriculum (not just in literacy and numeracy).

Equity and assessment

'Fairness' in assessment is often understood as validity and reliability of measures used. In standardised measures, such as NAPLAN, items are trialled and analysed for their 'DIF' (Differential Item Function) to ensure questions do not advantage or disadvantage certain population groups (for example, girls and boys). But 'fairness' can also be a way of considering equity issues in assessment. Equitable access to assessment for students with disabilities has been much addressed in research, but less attention has been paid to more holistic equity concerns associated with students' sociocultural capital.¹

Some research into the issue of equity in assessment has suggested that NAPLAN is inequitable because students from low SES backgrounds perform comparatively poorly on the tests. The performance of the students does not mean the test is unfair *per se* – indeed it may be argued that NAPLAN is successfully exposing inequitable opportunities to learn, for students from low socioeconomic backgrounds. That is, low SES students do poorly on some assessments because they do not have the expected skills and knowledge, not because the assessment is poor or unfair.

From an equity perspective then, effective assessment is assessment which:

- is accessible by all students
- provides a fair measure of learning
- does not tacitly or explicitly privilege students from high socio-economic backgrounds.

¹ In Bourdieu's analysis (1986), students from working class backgrounds are denied success because they lack cultural capital and social connections to gain successful employment outcomes. Bourdieu maintained that schooling reproduces social inequality by rewarding the cultural capital of the elite and devaluing that of the lower classes.

Effective assessment tools

As long as individual assessment tools have some degree of validity (i.e. the assessment tool actually measures what it sets out to measure) and reliability (i.e. the assessment tool is consistent in measuring what it sets out to measure), then they can be considered useful within the constructs of an effective overall assessment system.

There are some tools, however, that can be seen to be particularly useful when it comes to embedding assessment within a teaching and learning framework, and ensuring that feedback is an integral part of this framework. Examples of such tools are:

- graphic organisers – or tools that visually represent thoughts, ideas, knowledge and concepts
- review and reflection tools which enable students to review and reflect on their knowledge, progress, and what they have learnt and achieved during a unit, topic or project and
- feedback tools which enable students to provide feedback on their work and performance (Victoria Department of Education and Training, 2013).

Digital technology has enabled the development of computer-based assessment tools which are further discussed later in this report.

Elements of effective assessments

The following table presents a summary of the elements that are most strongly supported by available research as being important features of effective assessment.

Elements of effective assessments		Summative purpose	Formative purpose
Equity	The assessments are fair, equitable and accessible to all	✓	✓
Validity and Reliability	The assessments are psychometrically robust	✓	✓
Consistency	Scoring of the assessments is consistent	✓	✓
Transparency	The purpose, meaning and requirements of the assessments are well understood by all students	✓	✓
Engagement	When appropriate, tasks can reliably assess more complex skills (for example, analysis and synthesis, not just recall)	✓	✓
Timeliness	The assessments are administered at a time that is best placed to inform teaching plans and integrate with teaching and learning		✓
Feedback	Assessments lead to timely, personalised constructive and comprehensive feedback and/or reports that assist learning		✓
Alignment	When appropriate, tasks can reliably assess more complex skills (for example, such as analysis and synthesis, not just recall)		✓

3. Assessment tools and systems

Two assessment tools that have particular currency and broad application are portfolio-based assessment and criterion-based assessment. In addition to these, there is a plethora of new and innovative assessment tools available as a result of advances in digital technology. Technology-based systems can provide rich sources of information about individual student learning across wide segments of the curriculum over extended periods of time (Pellegrino and Quellmalz 2010).

Digital assessment tools lend themselves to a whole-of-system approach that allows assessment to be woven seamlessly into the learning environment so that it is almost invisible and supports real-time, just-in-time instruction (Shute 2015).

Portfolio assessment

In portfolio-based assessment, students are assessed on a portfolio of work. Portfolios are a systematic collection of student work and related material that depict a student's activities, accomplishments, and achievements in one or more school subjects. A portfolio can be both a product and a process-oriented tool that enables learners to monitor and reflect on their learning performance and receive feedback (UNSW 2014). Portfolios may take various forms, including:

- showcase portfolios which highlight the best products over a period of time
- process portfolios which concentrate on the 'journey' of learning
- evaluation portfolios which exhibit a series of evaluations over a course and the learning or accomplishments of the student in regard to previously determined criteria or goals (Fernsten 2009).

The advantages of using portfolios as an assessment tool are that they allow for presentation of a range of work by individual students; they provide material for dialogue and feedback between teachers and students; and they allow students to evaluate their own work through guided reflective practice (Fernsten 2009). Disadvantages of portfolio-based assessment include: it is difficult for teachers to make consistent assessments of material that differs from one portfolio to the next; portfolios can contain a lot of material that is 'busy work' and does not help an assessment of learning; and they do not allow very nimble responses to learning needs that may be identified, i.e. they have limited use as formative assessment.

Criterion-referenced assessment

Criterion-referenced assessment allows students to be uniformly assessed against a set of common criteria. Students' work is evaluated against criteria which describe the desirable qualities or expected attributes of a student's performance. This contrasts to norm-referenced assessment which compares students' performance with that of their peers (e.g. ranking students on a 'bell curve'). Criterion-referenced assessment may be used in high-stakes or low-stakes assessment.

In the best uses of criterion-referenced assessment, the criteria are well known to the students so that their work is performed in full knowledge as to how it will be judged; the criteria are sufficiently well described that two teachers using the same criteria would score the same piece of work identically; and the criteria provide the basis for detailed feedback to students.

Criterion-referenced assessments can be seen to be 'fairer' to students than norm-referenced tests because they evaluate achievement against a common and consistently applied set of criteria; hold all students to the same high expectations; and suit measurement of learning progress over time, giving teachers information they can use to improve teaching and school performance (Great School Partnership 2014b). On the other hand, some expertise is required to design good tasks and criteria, and teachers new to the approach usually need some training in how to apply the criteria.

Computer-administered testing

One of the most obvious applications of digital technology is the replacement of paper-based tests, with computer-administered assessments. In Australia, for example, NAPLAN moves online from 2017;² and the ESSA science test in NSW has been delivered online since 2008.

Research has shown that the mode of delivery does not affect student test scores greatly (Wang et al 2007; Wang et al 2008). A recent study into comparability across tablets and computers for high school students also found no statistically significant differences in student performance across device type for any content area or item type (Davis et al 2015).

Computer adaptive technology (CAT) takes the idea of computer-administered assessments significantly further. CAT assessments are designed to adjust their level of difficulty – based on the responses provided – to match the knowledge and ability of the test taker. The advantages of this type of assessment are that they:

- offer a shorter testing session with a smaller number of questions needed to hone in on student achievement levels
- tailor each question to the knowledge and ability of the test taker, thereby removing the need for students to deal with questions that are too hard or too easy
- provide more precise and quickly available information on student learning needs
- test security is improved as not all test takers see the same items.

Concerns about this form of assessment include:

- limited ability for computer adaptive technology to work for open-ended and essay questions
- students with lower technological literacy and/or less access to digital technology could be disadvantaged
- schools without adequate technological resources and/or support may also be disadvantaged (Great School Partnership 2013b).

Both computer-administered testing and adaptive testing allow for the development of large-scale student assessment which can be rapidly and reliably scored. This improves the speed and frequency of feedback given to both students and teachers (Csapo et al 2012). Online assessment also implies better management and use of data (OECD 2013) as educators have at their fingertips a variety of digital data that can be used for analysis and fed into teaching and learning systems. Furthermore, computer-based assessment allows for precision in detecting the actual values of the observed variables (Csapo et al 2012). In other words, a standard system is used to mark all assessments against agreed variables, meaning there is no variability in outcomes.

An example of a computer-administered testing system is e-asTTle, a New Zealand online assessment tool developed to assess student's achievement and progress in reading, mathematics, writing and Maori languages for students in Years 1-10. Teachers can choose the length of the test, whether to use it in adaptive mode or not, whether multi-choice questions are automatically marked and whether they would like to mark open-ended questions online with access to correct examples. Both teachers and students can access a wide range of reports and analyses once assessment tasks have been completed. Not only does it provide school leaders with information that can be used to inform learning programs and apply teaching practice that maximises individual student learning, it is also a tool that is being used by teachers for planning, for helping students to understand their progress and for involving parents in discussions about how well their children are doing (New Zealand Ministry of Education n.d. [a]).

2 Note that the delivery of NAPLAN online is proposed as 'semi-adaptive' or 'tailored testing design'; items will be delivered in 'testlets' of items at comparable levels of difficulty, rather than every item being presented on the basis of a student's response to the item preceding.

Use of technology to engage and assess

Technology has the potential to engage students more than traditional assessment tools may. For example, digital technology can allow assessment in problem-solving, conceptual understanding and critical thinking. In Minnesota in the US, students participate in an online science test with tasks engaging students in simulated laboratory experiments or the investigation of phenomena such as weather. Similarly, since 2014, the Technology and Engineering Literacy Assessment of the US National Assessment of Educational Progress (NAEP) has been entirely computer-administered and includes interactive, simulation-based tasks involving problem solving, communication and collaboration (Pellegrino and Quellmalz 2010). In NSW, a partnership between the Department of Education, ACARA and others has been investigating computer-based collaborative assessments⁴ in which students use technology to work collaboratively in real time to undertake assessment tasks. To date, the application of digital technology to assessment tools in terms of problem solving and engagement appears largely to have been seen in the STEM subjects (Pellegrino and Quellmalz 2010), but they could potentially be extended more broadly to other subjects.

Diagnostic feedback

Digital technologies support effective assessment through enhanced and faster provision of feedback on performance. The importance of explicit, timely feedback in teaching and learning is well established (as described earlier in this paper), but it is not always well implemented. Digital technologies such as the DIAGNOSER project provide a way to facilitate quality feedback. DIAGNOSER is an internet-based tool for classroom instruction of physics. It delivers continuous formative assessment and feedback to high school physics students and their teachers, and diagnoses misconceptions that underlie wrong answers, such as a confusion of velocity with acceleration (Thissen-Roe et al 2004).

ePortfolios are another form of digital tool that can promote and capture feedback. A digitised form of student portfolios (discussed earlier in this paper), ePortfolios may consist of text-based, graphic or multimedia components which are stored on a website or a portable electronic device such as a USB stick. They allow for enhanced feedback by enabling students, teachers, peers and parents to connect and collaboratively change and adapt ideas and feedback to a learner's developing needs and goals. They also provide the opportunity to capture sound, image and motion which may be beneficial to students who are not as comfortable with text (Rate 2008). New Zealand has promoted the benefits of ePortfolios extensively at the school level, citing the use of an ePortfolio as a support structure for teaching and learning. Examples of how schools in New Zealand have used ePortfolios in teaching, learning and assessment can be found in a document produced by the NZ Ministry of Education called *Digital Portfolios: Guidelines for Beginners* (2011).

Student learning accounts

Student learning accounts are another digital tool for supporting assessment in schools and integrating assessment into teaching and learning. These are currently used in Queensland for senior secondary students, and by the online Khan Academy.³ A student learning account is an online collation of evidence about learning achievements. For example, in Queensland, a student's learning account could include a Senior Statement, Queensland Certificate of Education, Tertiary Entrance Statement and/or Queensland Certificate of Individual Achievement (Queensland Curriculum and Assessment Authority 2015). A Khan Academy account includes a record of a student's own learning achievements (a 'personalised learning dashboard'), as well as a 'coach dashboard' where parents or teachers can see a summary of class performance as well as detailed student profiles (Khan Academy 2015). The idea of these accounts is that all student achievements are stored digitally in one place. Not only can a student easily track their record, their learning achievements can also be shared with others such as teachers and parents. This way teachers or parents can provide personalised feedback, see gaps in learning achievement, and set and plan assessment tasks that fit in with an individual student's learning needs and requirements.

3 The Khan Academy is a free, not-for-profit online learning forum where students can study a variety of subjects including maths, science, economics and finance, arts and humanities, and computing. The site offers practice exercise, instructional videos and a personalised learning dashboard.

4 This work has been informed by Professor Patrick Griffin and others involved in the Assessment and Teaching of 21st Century Skills project.

4. Examples of local and international practice

There is growing perception in many countries around the world of the need for more sophisticated approaches to assessment (Darling-Hammond 2012). High performing countries tend to embed assessment within a holistic system that sees assessment, teaching and learning feeding into and from one another.

Europe

In Europe, while there has been growing attention paid to digital assessment, its implementation across countries has been uneven (Scheuermann and Bjornsson 2009). For example in Finland, nearly one-third of Finnish pupils never or hardly ever use a computer at school. This is beginning to be seen as an impediment to educational attainment in some circles. In relation to the nexus between assessment and learning, the Finnish innovation organisation Sitra, has warned that people need to be able to gather data related to their own learning in one place (Berner et al 2015).

Finland is often used in education as an example of what a 'good' education system should look like, due to its outstanding results on PISA assessments in maths, science and reading. Educational leaders in Finland attribute their results to their intensive investments in teacher education, and major overhaul of the curriculum and assessment system. Finland has moved from a highly centralised education system, to one where local schools have more autonomy and prospective teachers are competitively selected from the pool of college graduates. All prospective teachers must attain a Masters degree before they can practice teaching (Darling-Hammond 2012).

The only standardised testing in Finland occurs in the final year of high school (Sahlberg 2014). All other assessment is school-based, student-centred and focused on open-ended tasks which are embedded in the curriculum. Assessment is used to cultivate students' active learning skills by posing complex problems and helping students to address these problems. Teachers have extensive decision-making authority in the areas of both curriculum and assessment (Darling-Hammond 2012).

Norway, on the other hand, is one of the most advanced users of e-assessment at a national level in Europe. Norway began experimenting with computer-based testing in 2002 and, following a subsequent trial in 2009 where Norwegian students took their exams on laptops, Norway was the first country in Europe to implement e-assessment on a national level. Tests for reading, English and maths are now taken digitally (Smith 2014). Similarly in Denmark, all national tests for primary and junior secondary students are online, using adaptive technology. The final exams for high school in Denmark are also taken online using a laptop (Victorian Department of Education and Early Childhood Development 2011). The emphasis of digital assessment in Denmark is on a student's ability to search for, sift through, and analyse information (Bargh 2011).

Asia

In some Asian countries such as Singapore, Japan and Hong Kong, there has been a strong push towards preparing students for the knowledge economy. However, Csapo (2012) observes that technology-based means of assessment delivery have not been a focus of development in any of the Asian countries at the system level and that technology-based assessment innovation is rare in these countries to date. Certainly there may be examples of individual digital assessment tools in place, but there has been little push at a government or policy level to introduce digital assessment systems that integrate teaching, learning and assessment.

Singapore's students also score exceptionally well on international tests such as TIMSS and PISA. Intensive investment and reform over thirty years has transformed the education system, broadening access and increasing equality. Like Finland, the alignment of curriculum, assessment and instruction in Singapore is exceptionally strong (Hogan 2014). Unlike Finland, traditionally the Singaporean assessment system has been based on one-off, timed, high-stakes summative assessments, although in recent years, there has been increasing emphasis placed on school-based assessment⁵ as well as the existing large-scale testing systems.

Teachers are encouraged to engage in continual assessment in the classroom, using a variety of modes such as classroom observation, oral communication, written assignments and practical and investigative tasks. Emphasis is placed on assessing problem solving and developing 'independent' learning (Darling-Hammond 2012). It is important to note that this shift in focus for Singapore does not affect a fundamental systemic commitment to ensuring that foundation skills are in place for every student, as a prerequisite for other learning.

United States

One nation that has not performed as well on international measures of student achievement, is the United States. Darling-Hammond and Adamson (2010) note that while student scores in the US have been rising on state tests used for accountability purposes, scores have been declining on tests that gauge students' ability to apply knowledge to novel problems (e.g. PISA). Some observers believe that high stakes, standardised testing has led to narrow and shallow classroom instruction (Goss et al 2015). Neither PISA nor the National Assessment of Educational Progress have shown any significant improvement in student learning in the US since the introduction of high-stakes testing. Some educators and education researchers in the US have proposed increased use of 'performance assessment', where students must construct an answer, produce a product, or perform an activity; and where assessment is embedded with teaching practices rather than only undertaken as a means to fulfil accountability or performance requirements (see, for example, Conley 2014; Darling-Hammond and Adamson 2010; Pellegrino 1999).

In the US there has been considerable interest in, and uptake of, digital assessment technologies both at local and system-wide levels. This has been driven by the perceived need to improve the learning outcomes of US students and the belief that technology can achieve this aim by integrating assessment more closely with teaching and learning (Bargh 2011). In 2010, the US Department of Education, issued a call for proposals under the 'Race to the Top' assessment program. The winning consortia were the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (Smarter Balanced). Both consortia proposed one system for summative assessment and another to serve individual-learning purposes (Bennett 2015). The Common Core Standards (CCS) which were introduced in the US in 2010 and adopted by most states, are facilitating the use of computer adaptive technology through the state-led consortia of PARCC Smarter Balanced. These standardised new assessment systems, will have embedded within them accessibility features including text-to-speech, and sign language (Bennett 2015). These consortia are working to develop digital assessments across K-12 that aim to provide meaningful feedback to students (Common Core State Standards Initiative 2015).

⁵ See, for example, Singapore's *Teach Less, Learn More* policy introduced in 2004 which heralded a move towards the 'quality' of learning and away from the 'quantity' of learning and exam preparation.

New Zealand

New Zealand appears to have a comprehensive plan in place, not only to embed assessment within teaching and learning frameworks, but also to make use of the newest digital technologies to do this. In New Zealand, the digital assessment system is focused on the student and assessment, not the technology - the technology is simply the enabler. It allows teaching and learning to be reflected in assessment; and enables student needs to drive the program, not assessment. Technology is the opportunity for assessment to be integrated into the learning process with a focus on problem solving and creation of new knowledge applied to real world situations.

New Zealand has in place an online computer-assisted assessment system - e-asTTle, as described earlier in this paper. In addition to this, there are numerous resources provided by the New Zealand Ministry of Education on effective assessment, embedding assessment within teaching and learning, and digital tools available. See, for example, the Ministry's 'Assessment Tools and Resources' site (n.d. [b]). These resources include an online assessment tool selector which is a resource for teachers and schools to help them select the most appropriate assessment tool to suit their particular purpose. The selector provides information about assessment tools most frequently used in New Zealand schools, in every area of the curriculum up to and including Year 10. Tools can also be compared to see which is the more suitable (New Zealand Ministry of Education 2015b). By 2018, the New Zealand Qualifications Authority aims to have online examinations in place across the board, as well as external moderation of internal assessments online. By 2020, they envisage offering a wide range of digital assessments, with the ultimate aim of assessment being online, anywhere, anytime (New Zealand Qualifications Authority 2015).

Australia

Australia tends to sit in the middle when comparisons are made between it and other countries, both in terms of the performance on international assessments; and in terms of an assessment system which is balanced between classroom level assessment and standardised assessment, with a reasonably strong focus on classroom assessment. For example, student survey data suggests that NSW teachers are more inclined to use formative assessment (including feedback to students) than the average of OECD teachers (CESE 2015). In most states and territories there has also been a long history of whole cohort testing in literacy and numeracy (since 1988 in NSW). And in some states and territories there are exit credential tests, such as the Higher School Certificate (HSC) in NSW (see below). Since 2008 Australia has had a national assessment – NAPLAN – which is a set of standardised tests of aspects of literacy and numeracy. NAPLAN is a rare example of a standardised test that is useful at the system level, as well as at school and student levels: it provides data for system monitoring, and can also point to areas of strength and weakness for students in a classroom.

All NSW schools are required to have policies and procedures in place to assess student progress and report to parents. Teachers are encouraged to use a variety of assessment techniques that are valid, reliable and appropriate to the age and stage of learning. The development of these teaching skills has been a major focus of teacher professional learning opportunities in NSW schools over the past decade. Teacher survey data from OECD's Teaching and Learning International Survey (TALIS) indicates that 80 per cent of NSW teachers who participated in professional development relating to student evaluation and assessment practices reported that it had a moderate to large impact on their teaching (CESE 2015).

Nonetheless, there is evidence that Australia has some way to go to ensure that teachers understand how to interpret and understand assessment data and effectively embed assessment within a framework of teaching and learning. For example, an OECD review of Australian assessment practices undertaken in 2011 found that when teachers graded against national A-E standards, the consistency of their judgements within a school was weak (Goss et al 2015). A 2013 Staff in Australia's Schools (SiAS) survey reported that 25.7 per cent of primary teachers identified the need for more professional learning in 'making effective use of student assessment information'. The findings were similar for secondary teachers (CESE 2014).

The NSW Higher School Certificate

The Higher School Certificate (HSC) is a locally, nationally and internationally recognised qualification for students who successfully complete secondary education in NSW. The HSC aims to deliver clear and comprehensive reports about what students know and can do, using a comparative backdrop of statewide standards.

Traditionally, credentialing and awarding bodies such as the Board of Studies, Teaching and Education Standards (BOSTES) in NSW have relied on summative assessments: typically unseen, timed examinations, externally marked. This is because the awards made as a result of the assessments are high stakes and the community expects the assessments made to be rigorous, fair and consistent across subjects and years.

Assessment for the HSC in NSW, for example, is largely shaped by its underpinning measurement framework. The need to translate student learning into scores on a scale has dictated the nature of the assessment. The HSC exams are a time-efficient way to assess all Year 12 students in all subjects in a way that is perceived as fair – if demanding – and psychometrically valid. The HSC also delivers a credential that is understood and valued by the community.⁶

The Higher School Certificate (HSC) is the culmination of your school career and the highest educational award you can achieve at secondary school in New South Wales...You will sit written exams at the same time as everyone else in the state studying the same course...Once the exams are finished, the BOSTES employs several thousand experienced teachers to mark all the exam papers...(NSW students online, BOSTES 2015)

Twenty-five years ago the HSC examination system moved from being purely exam-based to including moderated school-based assessments. These were also summative in intent, but were designed to mitigate some of the perceived negative effects of the exam-based system: the stress of exams for adolescent candidates; the challenge of a three hour exam capturing the breadth or depth of a student's learning; the consequences of a bad performance on the day.

However, there are critics of the school-based assessment approach:

...there is [now] not enough emphasis on tests, and too many marks are taken from take-home assessment tasks, which students can cheat on. These assessment tasks which are taken home can have input from a variety of people, such as tutors and parents, and do not often reflect the student's true ability... Tests ensure that the student gets the marks that they independently deserve...⁷

There is a risk that students with better resources and greater access to tutors or parents who can help can drive success in school-based assessment. This creates fairness and equity concerns, which must be addressed when the assessment process is as high stakes as the HSC. (See Equity and assessment, p. 5)

⁶ Other examples of assessments designed for delivering credentials and awards exist nationally and internationally, including the International Baccalaureate which is also available in some Australian schools. The IB is largely externally assessed. Though the assessment is not moderated in any way the assessments are varied in type and include criterion-referenced assessments (see page 7).

⁷ This comment comes from the student-operated and maintained Boredofstudies website, which was established 'to provide current year 12 and year 11 students with all the resources and support they need to succeed during the 2 years that form the most important part of their education and subsequently, their futures'.

5. Conclusion

Assessment is an area worthy of attention. Effective assessment is a critical tool for teachers and for systems to understand student learning, so as to better target policy at the system level, and practice at the classroom level. Attitudes and practices around assessment have undergone much change recently and technology has brought us the possibility of still greater change.

This paper has highlighted the inherent links between assessment and broader teaching and learning practices, given examples of current practices around the world, and highlighted some of the tools and ways that digital technology are changing assessment.

However, there are still many unanswered questions, which could form the basis of any further exploration of this issue at the state and national levels:

- What are the assessment needs of different years of students? What tools could be particularly beneficial at different year levels?
- What types of assessment most suit the purpose of credentialing or exit testing senior secondary students?
- How can we ensure that assessment tools used by teachers are effective?
- How do we ensure assessments are equitable?
- What needs to happen in initial teacher education to ensure new teachers graduate with good skills designing, implementing and using data from effective assessment tools? And how do we make sure that practising teachers are confident users of effective assessment tools?
- What is the relationship between assessment and reporting? How should reporting change as assessment practices change?
- Should NSW aim to develop an assessment framework and digital system that meets formative and summative, high stakes and low stakes assessment needs?

References

- Australian Council for Educational Research (2015a) *Computer adaptive testing at ACER*. Accessed 29 June 2015: <http://www.acer.edu.au/cari/computer-adaptive-testing-at-ac>
- Australian Council for Educational Research (2015b) *Centre for Assessment Reform and Innovation*. Accessed 30 June 2015: <http://www.acer.edu.au/cari/new-technology>
- Bargh, S. (2011) The future of secondary school examinations: The use of technology in the New Zealand context. In Hodis, M. and Kaiser, S. (eds) *Proceedings of the Symposium on Assessment and Learner Outcomes*, 1-3 September, Victoria University, Wellington, New Zealand
- Bennett, R. (2015) The changing nature of educational assessment. In *Review of Research in Education*. 39
- Berner, A-S., Laaksolahti, H. Kopola, R. (ed.) (2015) *A land of people who love to learn*. Sitra: Helsinki
- Black, P. & Wiliam, D. (1998) *Inside the Black Box: Raising standards through classroom assessment*, King's College, London
- BOSTES (n.d.) *Principles of effective assessment*. Accessed 18 June 2015: <http://syllabus.bos.nsw.edu.au/support-materials/principles-of-effective-assessment/>
- Bourdieu, P. (1986). The forms of cultural capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). New York: Greenwood Press.
- Csapo, B., Ainley, J., Bennett, R., Latour, T. and Law, N. (2012) Technological issues for computer-based assessment. In, Griffin, P., McGaw, B. and Care, E. (eds) *Assessment and teaching of 21st century skills*. Springer: Melbourne
- Centre for Education Statistics and Evaluation (2014) *What works best: Evidence-based practices to help improve NSW student performance*. NSW Department of Education and Communities: Sydney
- Centre for Education Statistics and Evaluation (2015) *State of the NSW Teaching Profession: 2014*. NSW Department of Education and Communities: Sydney
- Clarke, M. (2012) *What matters most for student assessment systems: A framework paper*. The World Bank: Washington D.C.
- Common Core State Standards Initiative (2015) *Frequently Asked Questions*. Accessed 30 June 2015: <http://www.corestandards.org/about-the-standards/frequently-asked-questions/>
- Conley, D. (2014) A new era for educational assessment. *Deeper Learning Research Series*. Jobs for the Future: Boston. Washington D.C.
- Darling-Hammond, L. and Adamson, F. (2010) *Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning*. Stanford CA: Stanford University, Stanford Center for Opportunity Policy in Education
- Darling-Hammond, L. (2012) Policy frameworks for new assessments. In, Griffin, P., McGaw, B. and Care, E. (eds) *Assessment and teaching of 21st century skills*. Springer: Melbourne
- David, L., Kong, X. and McBride, Y. (2015) *Device comparability of tablets and computers for assessment purposes*. National Council on Measurement in Education: Chicago
- Fernsten, L. (2009) Portfolio assessment. Education.com. Accessed 6 August 2015: <http://www.education.com/reference/article/portfolio-assessment/>
- Goss, P., Hunter, J., Romanes, D., Parsonage, H. (2015) *Targeted teaching: how better use of data can improve student learning*. Grattan Institute

- Great School Partnership (2013a) Summative assessment. *The Glossary of Education Reform*. Accessed 22 June 2015: <http://edglossary.org/summative-assessment/>
- Great School Partnership (2013b) Computer-adaptive test. *The Glossary of Education Reform*. Accessed 29 June 2015: <http://edglossary.org/summative-assessment/>
- Great School Partnership (2014a) Formative assessment. *The Glossary of Education Reform*. Accessed 22 June 2015: <http://edglossary.org/formative-assessment/>
- Great School Partnership (2014b) Criterion-referenced test. *The Glossary of Education Reform*. Accessed 6 August 2015: <http://edglossary.org/criterion-referenced-test/>
- Griffin, P., Care, E. and McGaw, B. (2012) The changing role of education and schools. In, Griffin, P., McGaw, B. and Care, E. (eds) *Assessment and teaching of 21st century skills*. Springer: Melbourne
- Hattie, J. (2003) Who says formative assessment matters? *Formative and summative interpretations of assessment information*. University of Auckland. Accessed 22 June 2015: [https://cdn.auckland.ac.nz/assets/education/hattie/docs/formative-and-summative-assessment-\(2003\).pdf](https://cdn.auckland.ac.nz/assets/education/hattie/docs/formative-and-summative-assessment-(2003).pdf)
- Hattie, J. and Timperley, H. (2007) The power of feedback. *Review of Educational Research*. 77(1)
- Hattie, J. (2009) *Visible learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge: Oxon
- Hogan, D. (2014) Why is Singapore's school system so successful, and is it a model for the West? In *The Conversation*. February 12, 2014
- Khan Academy (2015) About. Accessed 7 August 2015: <https://www.khanacademy.org/about>
- Kingston, N. and Nash, B. (2011) Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*. 30(4)
- Looney, J. (2009) Assessment and innovation in education. *OECD Education Working Paper No. 24*. OECD: Paris
- Masters, G. (2013) Reforming Education Assessment: Imperatives, principles and challenges. *Australian Education Review*. Australian Council for Educational Research: Camberwell
- Moss, Pamela et.al. (eds, 2008), *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press: Cambridge; New York
- New Zealand Ministry of Education (n.d._a) *Welcome to e-asTTle*. Accessed 6 August 2015: <http://e-asTTle.tki.org.nz/>
- New Zealand Ministry of Education (n.d._b) *Assessment tools and resources*. Accessed 7 August 2015: <http://assessment.tki.org.nz/Assessment-tools-resources>
- New Zealand Ministry of Education (2011) *Digital portfolios: Guidelines for beginners*. New Zealand Ministry of Education: Wellington
- New Zealand Ministry of Education (2015) *Assessment tool selector*. Accessed 7 August 2015: <http://assessment.tki.org.nz/Assessment-tools-resources/Assessment-tool-selector>
- New Zealand Qualifications Authority (2015) *Digital assessment*. Accessed 30 June 2015: <http://www.nzqa.govt.nz/assets/About-us/Our-role/innovation/DA-Factsheet-May15.pdf>
- NSW Department of Education and Communities (2015) *Government schools of NSW from 1848*. Accessed 18 June 2015: <http://www.governmentsschools.det.nsw.edu.au/examinations.shtm>
- OECD (2013) *Synergies for better learning: An international perspective on evaluation and assessment*. OECD Reviews of Evaluation and Assessment in Education. OECD: Paris
- OECD (2015) *Education Policy Outlook 2015: Making Reforms Happen*, OECD Publishing. <http://dx.doi.org/10.1787/9789264225442-en>
- Pellegrino, J. (1999) *The evolution of educational assessment: Considering the past and imagining the future*. Paper presented at Educational Testing Service on November 17, 1999. Policy, Evaluation and Research Center: Princeton
- Pellegrino, J. and Quellmalz, E. (2010) Perspectives on the integration of technology and assessment. In *Journal of Research on Technology in Education*. 43(2)

- Queensland Curriculum and Assessment Authority (2015) *My learning account*. Accessed 7 August 2015: <https://studentconnect.qcaa.qld.edu.au/learning.html>
- Rate, N. (2008) *Assessment for learning and ePortfolios: What are the formative benefits of ePortfolios*. New Zealand Ministry of Education.
- Sahlberg, P. (2014) Why Finland's schools are top-notch? *Pasi Sahlberg Blog*. Accessed 28 July 2015: <http://pasisahlberg.com/why-finlands-schools-are-top-notch/>
- Scheuermmann and Bjornsson (2009) *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing*. European Commission Joint Research Centre Institute for the Protection and Security of the Citizen. Ispra, Italy
- Shute, V. (2015) Quoted in: Assessment by stealth in a digital learning environment. *Research Developments*, ACER. Accessed 30 June 2015: <http://rd.acer.edu.au/article/assessment-by-stealth-in-a-digital-learning-environment>
- Smith, K. (2014) Digital assessment: What's happening around the world? *AQA policy blog*. Accessed 30 June 2015: <http://policyblog.aqa.org.uk/2014/11/17/digital-assessment-whats-happening-around-the-world/>
- Thissen-Roe, A., Hunt, E. and Minstrell, J. (2004) The DIAGNOSER project: Combining assessment and learning. In *Behaviour Research Methods, Instruments and Computers*. 36(2)
- University of New South Wales (2014) Assessing by portfolio. Accessed 6 August 2015: <https://teaching.unsw.edu.au/assessing-portfolio>
- Victoria Department of Education and Early Childhood Development (2011) Assessment in online learning environments. *Digital Learning Platforms Research Series* Paper No. 3. October 2011
- Victoria Department of Education and Training (2013a) *Assessment advice*. Accessed 22 June 2015: <http://www.education.vic.gov.au/school/teachers/support/pages/advice.aspx>
- Victoria Department of Education and Training (2013b) *Assessment tools*. Accessed 6 August 2015: <http://www.education.vic.gov.au/school/teachers/support/pages/tools.aspx>
- Wang, S., Jiao, H., Young, M., Brooks, T. and Olson, J. (2007) A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, 67(2)
- Wang, S., Jiao, H., Young, M., Brooks, T. and Olson, J. (2008) Comparability of computer-based and paper-and-pencil testing in K-12 reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, 68(1)
- William, D. (2011) *Embedded formative assessment: practical strategies and tools for K-12 teachers*. Solution Tree Press: Bloomington

Centre for Education Statistics and Evaluation
GPO Box 33
Sydney NSW 2001
Australia

 02 9561 1211

 cese@det.nsw.edu.au

 www.cese.nsw.gov.au

NSW Department of Education
September 2015

