



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

[Klenowski, Valentina](#) (2009) Raising the stakes : the challenges for teacher assessment. In *AARE International Education Research Conference - 2009*, 29th November - 3rd December 2009, National Convention Centre, Canberra, A.C.T.

This file was downloaded from: <http://eprints.qut.edu.au/43916/>

© Copyright 2009 Valentina Klenowski

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

Raising the Stakes: The Challenges for Teacher Assessment

Val Klenowski

Queensland University of Technology

val.klenowski@qut.edu.au

These are changing times for teachers and their students in Australia with the introduction of a national curriculum and standards driven reform. While countries in Europe such as England, and in Asia such as Singapore, are changing policy to make more use of assessment to support and improve learning it appears that we in Australia are moving towards creating policy that will raise the assessment stakes for the alleged purposes of transparency, accountability and fairness. What can be learnt from countries that have had years of high stakes testing? How can Australia avoid the mistakes of past curriculum and assessment reform efforts? And how can Australian teachers build their capacity to maximise their use of the learning power of assessment? These are the questions that will be addressed in this keynote presentation with reference to innovative research from global networks that have maintained the assessment focus on learning.

Raising the Stakes: The Challenges for Teacher Assessment

Introduction

These are changing times in Australia with the development of a national curriculum, national testing and reporting of school education outcomes. This should be no surprise given the current international context and the heightened accountability demands brought about by the Global Financial Crisis (GFC). As a country that is developing curriculum, assessment and reporting policy at the national level we need to consider the experiences of other countries such as England and Wales where national testing has existed for some time. In particular, we need to study the consequences of national testing, especially the

unintended consequences that have led to the de-skilling of teachers, a focusing on results rather than learning, and teaching to the test. We can learn from these experiences, however, caution is needed in borrowing practices and policies developed in other contexts and for differing purposes.

My argument is that we can avoid the unintended consequences of high stakes testing if we are careful in how we interpret the results and if we avoid the risks of oversimplification by presenting raw data in league tables and the over-interpretation of students' results in terms of innate ability, dispositions and limitations.

Teachers need to be assessment literate and see beyond the raw scores - this includes teachers having an understanding of the related equity issues. They also should continue with, and develop, classroom assessments that directly contribute to learning. This involves teachers developing their theories of learning as the basis for a 'principled' understanding of learning and assessment. Teacher assessment can itself be a source of dependable results through moderation practice.

Governments and policy officers are often seeking quick solutions to complex issues and problems, given the short time frame they have to demonstrate that their particular policies are achieving the intended results. In Australia we need to move beyond "quick fixes", technicist lists of "best practices" and simplistic arguments to support policy change.

Drivers for Educational Change in Australia

Global drivers for curriculum and assessment reform in Australia are apparent from policy makers' responses to international measures of educational attainment such as the results from the Programme for International Student Assessment (PISA), developed by the Organisation for Economic Co-operation and Development (OECD) or the Trends in International Mathematics and Science Study (TIMSS) of the International Association for the Evaluation of

Educational Achievement (IEA). Important questions of whether we are comparing like with like have not always been considered yet governments use the results from international comparisons to justify the introduction of ongoing curriculum change. The rigour of these studies, their methodologies and the way in which the data has been interpreted and used have been debated and critiqued by researchers such as Goldstein & Thomas, (2008); Hopmann, (2008); Wiliam, (2008); Nardi, (2008); Egelund, (2008) and Wagemaker, (2008). In Australia the use of international comparative data, such as TIMSS data, identified significant State and Territory differences. It was therefore predictable that the Labor Government in 2008 would introduce plans for a National Curriculum in Mathematics, Science, History and English in primary and secondary schools by 2011 to be extended to include languages, geography and the arts.

International comparisons have consistently highlighted equity issues for Australia as Indigenous children have scored significantly lower than non-Indigenous children (Klenowski, 2009). Australian schools are not adequately addressing inequalities and when compared with other developed countries, Australia is underperforming. This analysis from the 2003 PISA data suggested that Australia was “over-represented in the lowest categories of maths proficiency and under-represented in the highest” (Australian Council for Educational Research (ACER), 2004, p. xiii). While the achievement of students overall was high there were wide differences between the high and low achieving students.

From the analysis of the data from PISA 2006, that assessed science as the main domain with reading literacy and mathematics as minor domains, Indigenous students continued to be under-represented among the highest scoring students and over-represented among low scoring students. Reports indicate that Australia's lowest-performing students are most likely to come from Indigenous communities, geographically remote areas and poor socioeconomic backgrounds (ibid; Thomson, 2008). In terms of averages, about 40% of

Indigenous students, 23% of students from the lowest category of socioeconomic status, and 27% of students from remote schools, are not meeting a proficiency level in science that the OECD deems necessary for full participation in today's workforce and society. Headlines such as "PISA shows Indigenous students continue to struggle" (ACER, 2007) reflect areas of real inequity in Australia's education system that need to be addressed to ensure access to quality education for all students (Thomson, 2008).

National Levers for Educational Change in Australia

Apart from such global factors as international comparative analyses of achievement data there have also been national drivers for curriculum and assessment reforms in Australia. These developments are derived in part from an earlier investigation of the introduction of an Australian Certificate of Education (ACE) aimed at achieving greater consistency in senior secondary arrangements for curriculum, assessment and certification, more comparable student results across Australia, and more consistent standards of student achievement (Masters, Forster, Matters, & Tognolini, 2006). A further study (Matters & Masters, 2007) investigated what was common content, what was essential curriculum content and whether achievement standards were comparable in the final year of schooling, in English (including Literature), Mathematics, Chemistry, Physics and Australian History.

Significant consistency in what was assessed was identified, however, it was also found that different jurisdictions use different methods of assessment such as external examinations or teacher-devised assessment instruments. This finding raised the important issue of whether achievement standards can be compared across jurisdictions, or whether the existence of different assessment methods hinders comparison. The study recommended that a curriculum 'core' be identified for each nominated senior school subject to specify explicitly what students would be expected to learn no matter where in Australia they live. To achieve a nationally consistent description of how well students are expected to

learn the core in each subject it was recommended that a set of achievement standards be developed.

Other origins for these curriculum and assessment reforms that have been identified include the ministerial agreement on national goals at the Hobart Declaration of 1989, the Adelaide Declaration of 1999 and the National Declaration on Educational Goals for Young Australians of 2008. There is a sense that the nation as a whole can do better than its parts and that the nation's capacity would be greater if all jurisdictions worked together to achieve more efficiency and reduce duplication (McGaw, 2009).

Background

In 1999 the first annual literacy tests (reading and writing) for Year 3 and Year 5 students were conducted and these marked the beginning of benchmark testing in Australia. The nationally agreed literacy and numeracy benchmarks for Years 3, 5 and 7 represent minimum standards of performance. In 2008 the National Assessment Program – Literacy and Numeracy (NAPLAN) was introduced, students in Years 3, 5, 7 and 9 sit the same national tests in reading, writing, spelling, grammar and punctuation and numeracy. In addition, National Assessment Program assessments are also taking place and involve triennial sample assessments in science at Year 6, in civics and citizenship at Years 6 and 10 and in ICT literacy at Years 6 and 10 (Harrington, 2008).

Current Context

The states (6) and territories (2) of Australia developed individual approaches to the use of grades in the implementation of curriculum, assessment and reporting. This occurred because Dr Brendon Nelson, the Federal Liberal Minister of Education at the time 2001-2006, threatened to withhold funds from the states and territories unless they implemented reforms that included benchmark testing for literacy and numeracy and A-E reporting. Given the amount of funding involved the states and territories proceeded to implement these reforms and did not challenge the use of A-E grades. In England the Task Group on Assessment

and Testing (TGAT) that was set up in 1986 to advise on assessment and testing in the National Curriculum recommended progressive levels so that even those students who are making limited progress would still make progress. The grades system was rejected because of the fears related to damage to students. That is a student who was awarded an E grade could continue to receive an E grade which would not offer any real motivation to achieve.

In February 2008 the interim National Curriculum Board was established to set the *core content* and *achievement standards* in Mathematics, Science, History and English from Pre-school to Year 12. By May 2009 the National Curriculum Board had managed the development of four framing papers in the subject areas of English, Mathematics, Science and History. This work was handed over in May 2009 to the new, independent, statutory authority the Australian Curriculum, Assessment and Reporting Authority (ACARA) which now has responsibility for the national curriculum from Kindergarten to Year 12 in the specified learning areas, a national assessment program aligned to the national curriculum that measures students' progress and a national data collection and reporting program. The latter is intended to support analysis, evaluation, research and resource allocation and accountability and reporting on schools and broader national achievement. ACARA is managing the implementation of the national curriculum (to be referred to as the Australian Curriculum), national student assessment and reporting of school education outcomes. There is also an intention to establish a standards-referenced framework to "invigorate a national effort to improve student learning in the selected subjects" (National Curriculum Board, 2008: 3).

The Australian Curriculum is to have a futures orientation and will identify the essential skills, knowledge and capabilities that all young Australians are entitled to learn. A continuum of learning in literacy and numeracy skills will form the foundation for the national curriculum. It will be a web-based document. That is, web technologies will be used to embed links and enable multiple views and

access. The three elements of the national curriculum framework will comprise curriculum content, achievement standards and a reporting framework.

The curriculum content element of the Australian Curriculum will provide teachers with the expectations of what should be taught and what students are expected to learn, that is, knowledge, skills and understanding. Curriculum content will be described for a particular learning area at a particular year level for example, Mathematics, Year 5 (ACARA, 2009).

The achievement standards aim to provide “an expectation of the quality of learning that students should typically demonstrate in relation to the content by a particular point in their schooling (that is, the depth of their understanding, the extent of their knowledge and the sophistication of their skills)” (ACARA, 2009). The aim is to provide achievement standards for each year of schooling across K-10 using a descriptor of the quality of learning that draws together the knowledge, skills and understanding typically expected for that year. The representation of the standards for every year will include a statement of expected learning, a set of generic grade descriptors and a set of work samples that illustrate typical learning (ACARA, 2009).

Course specific standards are to be developed for Years 11-12 with a range of levels of achievement expected of students studying the particular course. The standards aim to assist in reporting to students and parents, to aid consistency of assessment and reporting across Australia and to fulfil the purpose of selection required of assessment for post-school pathways. It is intended that the Year 11-12 standards will be designed to be applicable in jurisdictions with external examinations and with school-based assessment.

Finally the reporting framework aims to provide consistency in nomenclature to describe the quality of achievement associated with each A-E grade for use across K-10. It is intended that the use of the five-point scale will indicate the extent to which a student has met the achievement standard for a particular year of school. To illustrate, students who achieve a grade of C or above will have met

the standard for that year/stage. The grade C would indicate a satisfactory level of achievement while an A grade would indicate an outstanding level of achievement. Conversely a grade of D or E would suggest that follow-up is required and further investigation by teachers, students and parents might be needed (ACARA, 2009).

It is also intended that annotated student work samples will be used to demonstrate the different standards. This collection of work samples will build on the work that is currently established in the Australian states and territories. It is anticipated that this collection will provide a common and national reference point for greater consistency in teacher judgement within and between classrooms, schools, states and territories.

Such changes to curriculum and assessment make considerable demands on teachers who need to be informed, prepared and resourced to implement this level of change. It is most important that teachers are aware of the literacy demands of national curriculum and assessment for the implementation of a national curriculum requires the development of teachers' capacity to use the learning power of assessment to improve the outcomes for all students.

What can be learnt from countries that have had years of high stakes testing?

Patricia Broadfoot in 1996 offered a sociological view of assessment when she stated that one of its functions was "the control of both individual aspirations and systemic functioning" (p.10). Stobart (2008:1) too in his recent book has referred to assessment as "a value-laden social activity" and stated "there is no such thing as 'culture-free' assessment". The title of his book makes his message clear – *Testing Times: The Uses and Abuses of Assessment*.

At this time in Australia it is opportune to consider such messages and perspectives that derive from contexts where issues of power and control in relation to high stakes assessment have been well researched and theorised. Stobart's message (2008:1) that "assessment shapes how we see ourselves" is

particularly important given the ways in which students in Australia, such as some Indigenous students, are currently being labelled as “underperforming”, “continuing to struggle” or “not meeting a proficiency level in science that the OECD deems necessary for full participation in today’s workforce and society.” Such classifications resulting from high stakes testing illustrate how “assessment shapes who and what we are and cannot be treated as a neutral measure of abilities or skills that are independent of society” (Nietzsche as cited by Stobart, 2008:6). Assessment also shapes what we learn and how we learn.

Assessment procedures not only relate to the control of the individual but also extend to include systemic control of the educational system as a whole (Broadfoot, 1996). In Australia this is currently exemplified in the development of a powerful bureaucracy - ACARA. The activities of this authority will regulate the process of education through centralisation and control of the national curriculum and the regulation of the achievement standards. It is helpful to understand the way historical and contemporary developments in assessment policy and practice relate to themes of ‘competence, competition and control’ and how the social, economic and political factors can dictate their priority (ibid). As is apparent from the current Australian context, in times of global economic downturn and competition, accountability looms large.

Accountability

The move to the foregrounding of the accountability purpose of assessment that occurred in England almost twenty years ago, is here today in Australia, with schools and teachers being judged on the published results and placed in league tables. Over the past two years, in particular, accountability testing has assumed a high profile in public education policy as evident from the NAPLAN testing. Doubts exist about what such tests are actually testing and how such tests help to support learning. To illustrate, Willett and Gardiner (2009) drawing on their longitudinal equating study in their critique of the NAPLAN spelling test raise significant questions about the validity and reliability of NAPLAN achievement data of this test. These authors make the crucial point that this type

of test does not assist teachers' practice or student learning rather the information provided to teachers is unhelpful and incorrect and more than likely will have a negative impact by encouraging discredited spelling constructs and teaching methods. They provide examples of how the construction of spelling items is formulaic which they contend derives from the lack of an articulated research-based framework and is also partly due to the desire to keep the item 'pure' by trying to ensure that the items have a single item demand (Willett & Gardiner, 2009: 5).

Table 1: Formula for creation of spelling items (Willett & Gardiner, 2009)

Leave out a letter	craked (cracked), weel (wheel), frends (friends) Used at the syllable junction eg swimming (swimming), disapointed (disappointed)
Add a letter	Used at the syllable junction consumed (consumed), fitness (fitness)
Use a different vowel combination	broun (brown), arownd (around), lowdly (loudly), seet (seat),
Substitute a letter	cumplained (complained), sinse (since)
Reverse a letter sequence	muscel (muscle), marothan (marathon)

The issue here is that constructing items in this formulaic manner contrasts with the authentic student spelling errors. This approach is likely to encourage the teaching of test preparation or 'testwiseness' rather than productive spelling knowledge and skills. These researchers were also able to demonstrate the negative impact of such testing by providing evidence of how the misspelling of the first syllable (com) in complain as cumplain, is not an error that Year 3 students make, yet after exposure to the NAPLAN error, students in their study used this misspelling when attempting to spell the word! This exemplifies the point made that the way students are assessed impacts on the way they learn and what they learn.

In an accountability context, when the stakes are raised, the interpretations of large-scale assessment scores should be treated with caution. The media will too readily provide the public with league tables of raw scores. This reductionist approach provides ambiguous and narrow meaning. “When funding decisions are treated as unambiguous, and when single scores are generalised beyond justification as true characterisations of individuals and systems, the potential for mischief is enormous” (Shavelson, Black, Wiliam, and Coffey, 2004:35). In a context of standards-driven reform and standardisation there is the danger that technical and rationalist approaches will generalise and make superficial the assessment process. There is also the inexorable existence of pressures to pervert practice particularly when results are tied to decisions concerning funding.

In countries, such as England, where for approximately twenty years such testing has existed “... centralisation and tight political control ... left the national ministry and its agencies holding the power. There were age-defined levels of attainment and an inspectorate that played a prominent role in monitoring standards of performance across the whole system. The imposed national curriculum and national testing led to direct and indirect interventions into the way that the curriculum was taught. All in all, there was an increase in the technical elements of teachers’ work and a reduction in the professional.” (Klenowski, 2005:104).

In some states, like Queensland, the state government is keen to raise standards as represented by the results of NAPLAN testing. The Queensland Premier advised schools to sit practice NAPLAN tests in Years 3, 5, 7 and 9 as she was disappointed by the overall results of the 2008 tests which she indicated were designed to assess if students were meeting “national *standards* in numeracy, reading, writing, spelling, punctuation and grammar” (Bligh, 2009, my emphasis). Yet there are no officially endorsed statements about the expected learning of literacy and numeracy as cross-curriculum priorities and no descriptors of standards to inform teachers about the expectations of quality, except for those produced after the testing is complete. Summary statements of skills assessed

to inform parents about their child's report are provided. The standard is referred to as the minimum or benchmark standard representing a level, below which, a student is considered to be at educational risk. Teachers are using practice tests as an opportunity to familiarise students with test conditions and the types of anticipated test questions to measure students' improvement efforts. This is due to the lack of information about expected qualities of performance, or how the testing relates to learning in the curriculum, or to curriculum domain standards more specifically (Klenowski & Wyatt-Smith, 2009).

Headlines such as "Literacy and numeracy problems unchecked for a decade", "Rees schools report card has arrived", "Publish schools results: Rees", "Crunch Time at School: National test results must be publicly available to all states", "Gillard praises state for lead on ratings", " 'Flying squads' to lift state teaching" are now all too familiar. In Queensland the use of "flying squads" of teams of four teachers headed by a school principal evaluating teaching practices and making recommendations to improve them is a clear sign that teachers are being required to account and the assessment stakes have been raised. Such quick fix approaches to improvement are not effective as the overriding goal is now focused on higher grades or results per se rather than the issues related to assessment *and* learning. This \$9 000 000 initiative was proposed subsequent to the Masters' report on the Queensland's education system which was described as lagging behind other states in national exams (Dunleavy, 2009). The Masters' report recommended "that standard science tests be introduced at Years 4, 6, 8 and 10 for school use in identifying students who are not meeting year-level expectations and for monitoring student progress over time."

In England there have been some devastating and lasting effects of the emphasis on testing and examinations. So much so that today there have been some significant changes with a move towards more trust and recognition being given to teachers' judgement. To illustrate the unintended consequences of national testing it is useful to consider international examples where students have experienced years of examinations and tests. In England an expert enquiry

into the Key Stage 2 (students are aged 11) Science tests have recommended that they be abandoned. This decision relates to Professor Philip Adey's research findings that

“The general cognitive foundation of 11 and 12-year-olds has taken a big dip. There has been a continuous decline in the last 30 years and it is carrying on now.” (Griffiths, 2006)

Possible explanations for why this has occurred are numerous but one that is pertinent to the recommendations of the Masters' Report in Queensland is that:

“By stressing the basics - reading and writing - and testing like crazy you reduce the level of cognitive stimulation. Children have the facts but they are not thinking very well. ... And they are not getting hands-on physical experience of the way materials behave.” (ibid)

‘Teaching to the test’ resulted in a reduction in hands on practical tasks which in turn led to students’ conceptual skills actually decreasing over the past ten years. So while the test results improved year on year, the learning and understanding decreased.

In the United States the unintended consequences of testing and the No Child Left Behind Policy has led to a culture of fear of job losses, by teachers and school principals, and a fear of school closures. To avoid such measures teaching to the test has led to test irregularities such as providing answers to exam questions and the reduction in Native language and culture responsive teaching (McCarty, 2009; Patrick, 2008).

How can Australia avoid the mistakes of past curriculum and assessment reform efforts?

Australia can learn from the research conducted in other countries that have implemented national curriculum and assessment systems. The insights concerning intelligent accountability systems and acknowledgement of the

professionalism of teachers are useful. It would appear with the adoption of quick fix approaches as in Queensland and the move towards a transparency agenda limited awareness of the negative impact of such strategies prevails.

In England there is a move to privilege and support teacher judgement in assessment, the reverse could be said to be occurring in Australia. As indicated recently by the Education Minister at the Indigenous Leadership in Education Institute: Stronger Smarter Summit held in Brisbane.

“I have unapologetically pursued a transparency agenda. Much of the debate about transparency has centred on so-called league tables. For me, transparency has always been about information for parents, accountability of governments and schools and being able to target additional resources to schools that need it most. In January next year school performance information for every school in the country will be published on a national website. This is a vitally important and powerful tool for change.” (Gillard, 2009)

She went on to indicate her confidence that some schools with high proportions of Indigenous students will have performed very well “schools that are doing and share this best practice”. She continued that there would also be a number of schools with high proportions of Indigenous students who were not performing well and stated: “It will be completely transparent which schools need to improve their outcomes to deliver for Indigenous students and need help. The transparency agenda comes with significant new resources, with billions of dollars being made available ...” (Gillard, 2009). Such resource allocation and monitoring of achievement will require greater control on the part of the government and accountability on the part of the schools and teachers.

Transparency and Trust

In 2002 Onora O'Neill argued for more intelligent forms of accountability with less focus on ideals of transparency and more on limiting deception. She was unconvinced by heavy-handed accountability that combined managerial targets and requirements for detailed compliance with bureaucratic process. She warned that excessive regulation could undermine professional performance and standards. She suggested that deception could be condoned in the zeal for total transparency.

There is evidence that the power of assessment in high-stakes contexts leads teachers to 'teach to the test' and to prepare students for the test. Indeed, in 2009 in Queensland it was a requirement for teachers to practice the NAPLAN tests in Years 3, 5, 7 and 9. We should heed the lessons learnt elsewhere to ensure that the curriculum is not narrowed or trivialised by such practices and encourage teacher assessment so that it does not focus on the items that are easily tested but rather includes those that will be useful and relevant for students faced with the challenges of globalisation.

The interpretation of test and examination results as unproblematic and predictive implies that:

- assessments measure something innate about the learner that indicates future learning potential
- people's knowledge is available in the same way
- failure or inadequate answers are interpreted as measures of lack of knowledge and or understanding
- the assessment item, and the internal process of individuals responding to it, will be viewed as separate from social and cultural influences
- communication is seen as monologic, meaning is singular
- the receiver is passive and extracts meaning in the words
- assessment constructs are, therefore, stable across children and adults and

- assessment is a process of reading off the knowledge stored 'in the head' (Murphy, 2009).

Such an interpretation of test results has consequences for those who are not succeeding given the current indicators of achievement. This has been exemplified recently in Australia in the categorisation of the lowest-performing students as most likely to come from Indigenous communities, remote regions and low socioeconomic backgrounds. A deficit view of assessment dominates and a more insightful and powerful assessment practice to informing teaching and learning is needed.

A major purpose of assessment has been to establish and raise standards of learning (Stobart, 2008) however in introducing minimum standards that encourage competition it is important to prepare teachers and to inform them of the consequences. For as has been seen the consequences are not just significant for the individual student they are important for the teacher, the school, the education authority and the state. Teachers need to understand the accountability context and appreciate how the practices they engage in are mediated by structures beyond their control such as national or state policy that directs the teacher in terms of what to assess and how that is recorded and reported. It is the wider social and institutional contextual aspects that give meaning and structure to teacher's practice.

Changes because of Unintended Consequences

In October 2008 the government in England abolished national tests for 14 year olds. As expressed by the Chair of the Office of the Qualifications and Examinations Regulator (Ofqual) the government's decision to listen to advice about the impact on schools and the need for change was welcomed. She indicated how new assessments would be subject to validation by Ofqual to ensure that they command confidence, and standards are secure (International Association for Educational Assessment, 2008).

Now in England teacher assessment is the only form of reported judgement at the ages of 5, 7 and 14 and is becoming more significant in the assessment of 7 to 11 year olds. Assessing Pupils' Progress (APP) is a new national approach to assessment that puts the learner at the heart of the assessment process. It is a process of structured periodic assessment for mathematics, reading and writing. APP is linked to national standards, intended to build a well-rounded individual profile of learners' achievements that highlights their strengths and areas for improvement. Over recent years schools in England have invested in 'assessment for learning' practices. APP has been designed to strengthen teachers' assessment capacity by building on their 'assessment for learning' practice in day-to-day teaching and to help them see how assessment is integral to their planning and pedagogy. Teachers in England now have a common language and criteria with which to build clear pupil profiles which are reviewed at intervals that are adequate to allow progress to be recognised and timely to allow the insights gained to be acted upon to make further improvements (Anwyll and Horner, 2009). The aim is to build teachers' assessment capacity that had been stripped away by previous years of political compliance and target setting.

So critique of assessment policy by teachers, academics, unions and parents is not enough. All stakeholders are required to engage in debate and research that informs policy and provides insights into how teachers can use the learning power of assessment in a context that is increasingly calling them to account.

How can we build teachers' capacity to make productive use of the learning power of assessment?

The research tells us that assessment can enhance learning however there is a need for all teachers in Australia to develop their capacity and understanding of assessment practices. For pre-service teachers or practising teachers there are very few assessment courses available that provide opportunities for the practice and development of these skills and understandings to engage in teacher assessment. With the introduction of a national curriculum and achievement

standards, international and local research indicates that teachers need experience and tools to assist them in the productive use of assessment.

In considering assessment in the context of assessment for learning it is important to understand the underpinning theory of learning. It was Denis Lawton (1992) who promoted the view that teachers needed to be more than 'good practitioners'. Theory must underlie their practice and teachers need to theorise about their practice so that theory is grounded and developed from practice. In research conducted with teachers in practices related to assessment a **sociocultural view of learning** has been most helpful in explaining the limitations and necessary supports in a context of achievement standards (Adie, 2008; 2009; Klenowski & Wyatt-Smith, 2008; 2009) and assessment for learning (Willis, 2007; 2008; 2009a).

"Sociocultural theories view learning as central to practice; and all practice is understood to be social. Engaging in the practices of an educational institution or workplace is an example of a social activity" (Murphy, Hall, McCormick and Drury, 2008: 17). From this view learning is a social construction it takes place as people engage in **collective activities** and knowledge is a result of social transaction. Such knowledge increases the learners' ability to participate in social activity to become a particular type of person for example a maths learner or a literate person. Learning is both 'becoming' (increasing in competency) and 'belonging' (transformation of identity) (Murphy, 2009). This way of viewing learning leads to the expectation that **there will be differences** across contexts because learning is understood not just in terms of what individuals can do but what is possible for them to do in certain situations **with certain people** (Murphy, Hall, McCormick and Drury, 2008: 17). Individual development and social/collective development are understood to **be interdependent and complementary** processes. Assessment from this view of learning is therefore seen as a social practice (Stobart, 2008).

Teachers need to develop their **assessment literacy** which is defined here not from a traditional view of "literacy as skills, knowledges and cognitions that reside

within the individual” but more from a view of “**visible social practices** with language, text and discourse” (Gee, 2003). This view aligns with a sociocultural approach to assessment theory and practice. To build teachers’ assessment literacy it is important then for them to understand the **assessment – learning relationship**, practice **teacher assessment** and **social moderation**.

The **assessment - learning relationship is complex**. Assessment is not a separate linear process. The teacher’s view of learning will impact on the approach to assessment (Watkins, 2003; James, nd). For instance, in Queensland the use of NAPLAN practice tests to improve scores on national tests is underpinned by behaviourist assumptions about learning. The message given to teachers is that assessment measures progress through unseen tests with items taken from levels in a skill hierarchy. Those students who underperform require further practice on the incorrect items. This is a position that contrasts with recent developments in the use of assessment to focus on the quality and support of learning rather than the simple measurement of it. Formative assessment aims to promote a discourse about learning in which both learners and their teachers are negotiating future learning. The teacher-student relationship in this context is more supportive of learning as the view of expertise is more horizontal than hierarchical (Murphy, 2009). Assessment for Learning (AFL) a term, often used synonymously, with formative assessment has recently been redefined at an international symposium on AFL. This occurred because the previous definition had been misinterpreted and misunderstood as evident from the way in which it was enacted in policy and practice. There was a lack of adherence to the principles of AFL and the ‘spirit’ of the concept. Too often the practices were superficial without teacher and student active engagement with a focus on learning. Other misunderstandings stemmed from

“... deliberate appropriation, for political ends, of principles that have won significant support from educators. For example, ‘deciding where the learners are in their learning, where they need to go and how best to get there’, has sometimes been (mis)interpreted as an exhortation to teachers

to (summatively) test their students frequently to assess the levels they attain on prescribed national/state scales in order to fix their failings and target the next level. In this scenario, scores, which are intended to be indicators of, or proxies for, learning, become the goals themselves. Real and sustained learning is sacrificed to performance on a test". (James, in press)

The second generation definition that emerged from the Third International Symposium on Assessment for Learning (2009: 2) is as follows:

Assessment for Learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning.

This definition refocuses the practice on the **relationship between the teacher and the student** or between the students themselves in the learning process. In a recent study by Willis (2009b:2) it was found that “the **teacher-student relationship** has a central mediating role as does the complex interrelationship between teacher and student beliefs about learning, the structure of the assessment tasks and the social and cultural interactions and contexts of the classroom, school and policy environment”. A sociocultural perspective that emphasises the participatory, as opposed to an acquisition, view of learning, helped to explain how the students appropriated participatory practices into their traditional cultural narratives of learning, forming ‘entangled’ (Elwood, 2008) learner identities. Entangled in the sense of using the teacher’s discourse related to learning and assessment.

Assessment for learning strategies need to be grounded in an understanding of the importance of the **student-teacher relationship** and for the teacher to have a learning orientation rather than a **telling orientation** to their pedagogy. This requires **a shift in practice to seeing learning as involving participation and negotiation where expertise develops over time through attending to**

practice and engaging with others. From this view of learning meaning cannot be given or handed over - this is a transmission view of teaching and an acquisition view of learning. It is counter to the view that meaning comes from the “dynamic relation of living in the world” (Wenger, 2008 in Murphy and Hall, 2008) so to come to know the world or to learn is a process of **negotiation of meaning.**

Teacher assessment or “the process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students’ learning, based on their professional judgement, and to report at a particular time on their students’ achievements” (Harlen, 2005: 247) is a capacity to be developed in pre-service and in-service education. In Australia teachers do engage in this practice however there is the danger that this important alternative to centrally devised tests could be lost if teachers are continually directed to deliver practice tests and acquiesce. Teacher assessment addresses students’ needs that emerge from a particular context, sociocultural or historical background. **Teacher designed assessment tasks** involve a variety of contexts, a range of modes of assessment, a range of response formats and styles and a range of indicators particularly to address **issues of equity.**

Messages from the Research

Equity or fairness in assessment as defined by (Stobart, 2005: 275) is a “qualitative concern for what is just” and it is more of a **sociocultural issue than a technical one.** Equity involves much more than a consideration of the specific design of tests or tasks. Attention to **whether all students have access to learning, how the curriculum and/or standards are defined and taught and how achievement in the curriculum is interpreted** are equally important considerations.

The differential performance of students from different cultures may not be due to bias in the choice of test content or design alone, but may be attributable to real differences in performance **because of these students’ differing access to**

learning, different social, cultural contexts or real differences in their attainment in the topic under consideration due to their experiences and sociocultural background. The content and mode of the assessment tasks or tests may be outside these students' experiences and may limit their engagement with the tasks as they position them as not knowledgeable or failing in this assessment context.

The intention of **culture-fair assessment** is to design assessments so that no particular culture has an advantage over another. The purpose of culture-fair assessment is to eliminate the privileging of particular groups over others. However, it is difficult to claim that assessments can be completely culturally unbiased. The opportunity to participate in learning (**access issues**) and the opportunity to demonstrate learning (**validity and fairness** in assessment) are deemed fundamental factors in developing **culture-fair assessment**.

The variables identified as possible influences on student performance include:

- the **cultural specificity of how the assessment task is framed**;
- the **cultural specificity of the normative models of child and adolescent development** reflected in the constructs of the assessment or test;
- the **linguistic codes and conventions** of the assessment;
- the **cultural-specificity of content knowledge** (Luke, Woods, Land, Bahr & McFarland, 2004, pp.12-13)

To achieve culture-fair assessment **issues in language, cultural content, developmental sequence, framing, content and interpretation and reporting need to be addressed.** To illustrate, the sampling of the content for assessment needs to offer opportunities for all of the different groups of students who will be taking the test. Assessment interpretations of students' performance need to be contextualized so that what is, or is not, being valued is made explicit as well as the constructs being assessed and the criteria for assessment (Gipps, 1994). **To**

achieve culture-fair assessment the values and perspectives of assessment designers need to be made more public. Further, to understand how culture-fair assessment practice is developed and attained requires a careful study of **how the learning experience is modified by teachers for particular students to achieve engagement, participation and improvement in learning.**

Current research (Klenowski and Gertz, 2009) using a sociocultural perspective to analyse and theorise how culture fair assessment practices support greater equity and success for Indigenous students has helped to explain why it is important for teachers to use culturally-relevant pedagogic practices, that incorporate increased use of oral language and group activity based mathematical problems. Teachers have indicated that the following practices are beneficial to Indigenous students' learning of maths: using maths rotations and maths investigations as regular class work, providing opportunities for peer support, including students explaining to other students how to solve the problems, contextualising the maths problem or maths task so that it is locally and personally meaningful and using examples which relate to Indigenous students' sociocultural backgrounds and their interests such as football. The importance of building and strengthening relationships with Indigenous students so that they feel as though they belong in classes that are predominantly white and western has also been suggested as significant.

When teachers assess students' work at the local level as is happening in these schools it can form part of the state system of assessment of student performance. However **social moderation** processes are considered to be a necessary component of such school-based assessment as teacher assessment is usually deemed to have high validity but questionable reliability.

Social moderation practice at the local level has the potential to fulfil an important role as a process for aiding teachers in ascribing value to student work through the use of standards. It involves groups of teachers meeting to discuss and negotiate assigned gradings of student work with the aim of reaching a

common understanding of the quality of work. The purpose of social moderation is to produce valid and reliable judgements that are consistent with one another and with stated standards of performance. It is a form of quality assurance in terms of achieving comparability of judgements based on evidence of student achievement (Maxwell, 2007).

Moderation is also a social practice that encourages the development of a sense of community of assessors. It is an opportunity for the generation of new knowledge and new ways of knowing. The group's discussion and debate about their interpretation of the quality of the evidence related to the achievement standard awarded is the context for situated interaction. Teachers participate in this context drawing on their individual tacit and explicit knowledge of standards together with the group's explicit and tacit knowledge of standards. For the moderation group, the source of *new* knowledge and knowing lies in the use of knowledge (tacit and explicit) about standards as a tool of knowing. This concept has been usefully called 'the generative dance' (Cook and Brown, 1999).¹

Teachers' engagement in moderation practice and the new knowledge and ways of knowing that are generated include:

- Teachers are able to check that **similar skills** and **levels of skills** are taught and **similar outcomes** are **assessed as equitable** and of a **comparable quality**.
- **Fairness** for all students is extended beyond the classroom or school to between schools spread within the state
- Provides **confidence** for teachers, parents, students, other staff members that common standards are expected and being achieved by a particular year group of students
- Teaching and assessment practices are made transparent.
Teachers' work is made public, open to scrutiny and critique

¹ I am indebted to Patricia Murphy for drawing this concept to my attention.

that helps to address accountability and quality assurance demands. **Gaps or omissions in the teaching programme can be identified**, particularly if the Director of Curriculum or Head of Department participates in the moderation meetings.

- A **sense of community develops** as teachers negotiate their understanding and seek clarification and advice when they are unsure of the standard or the standard of work. There is a shift from individual practice to shared practice and the improvement of shared practice.
- **Engaging in moderation practice focuses teachers' attention on assessment and its place within the teaching and learning programme.** Teachers seem inspired to teach a topic when they realise the results achieved by other teachers using different approaches. In this way teachers learn new ways of teaching a topic, are diversifying their practice to meet the needs of individuals and in so doing are improving practice (Klenowski & Adie, 2009).

In this context standards are intended to be used as the basis for judgements of student achievement; while the results from assessment tasks are meant to both inform the teaching/learning process, and to report and track student progress. In such a system, the role and reliability of teacher judgement takes centre stage.

Internationally moderation has been recognised as an important practice with the Cabinet Secretary for Education and Lifelong Learning in Scotland stating that:

“A national system of quality assurance and moderation of 3-18 will be developed to support teachers in achieving greater consistency and confidence in their professional judgements” (Hyslop, 2009).

In Wales national tests have been abandoned and the value of school-based assessment and teacher moderation practice has been recognised.

Recommendations from the Report on *Future assessment arrangements for Key*

Stages (KS) 2 and 3 included cluster group moderation for transition links with Key Stage 2 and 3 schools. Strengthened assessment at the end of KS 3 takes place by means of external moderation of sample evidence of teachers' understanding and application of the national curriculum level descriptions and verification of school-based systems and recognition of the quality of teacher assessment by awarding schools 'accredited centre' status. Since September 2007, primary school teachers have used school-based moderation, involving suitably robust systems and procedures to ensure that they have appropriate opportunities to discuss their pupils' work and agree a shared understanding of standards.

Conclusion

In Australia the major challenge for the current curriculum-driven reform is for the relationship between the learner, learning and assessment to remain central and for the professionalism of teachers to be sustained through a socio-cultural approach to assessment that views it as educative for teaching and learning. Learning involves negotiation and participation. As a community of researchers and academics I see it as our responsibility to rise to that challenge of informing policy and shifting the agenda along a learning trajectory that is informed by our empirical research and the research of those with years of experience of not just high stakes testing but theorized understandings of how assessment can develop learning for all, and meet accountability demands in an intelligent, fair and just manner.

References

- Adie, L. (2008) *Changing assessment practices: Developing the case for online moderation*, Paper presented at the Australian Association for Research in Education, December 2008, Brisbane.
- Adie, L. (2009) *Consistency, connectedness and creative solutions: An investigation into online moderation*, Paper presented at International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane.

Anwyll, S. & Horner, S. (2009) *Strengthening teacher assessment in schools in England: the development and use of Assessing Pupils' Progress*, Paper presented at International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane.

Assessment Reform Group (ARG) (1999) *Assessment for Learning: Beyond the black box*, Cambridge: University of Cambridge Faculty of Education.

Assessment Reform Group (ARG) (2002) *Assessment for Learning: 10 principles*. Accessed August, 24, 2008, from <http://www.assessment-refom-group.org.uk>

Australian Council for Education Research (ACER) (2004). *Facing the future: A focus on mathematical literacy among Australian 15 year-old students in PISA 2003*. Eds. S. Thomson, J. Cresswell, and L. De Bortoli. Camberwell: ACER.

Australian Curriculum, Assessment and Reporting Authority (2009) *Curriculum Design Paper*, Accessed 16 June, 2009
http://www.acara.edu.au/verve/resources/Curriculum_Design_Paper_.pdf

Australian Government (2008) *Australian Curriculum, Assessment and Reporting Authority Act 2008, No. 136, 2008*, Accessed 12 June, 2009
[http://www.comlaw.gov.au/ComLaw/Legislation/Act1.nsf/0/8F3E822CC1A33B3DCA25751B007E3618/\\$file/1362008.pdf](http://www.comlaw.gov.au/ComLaw/Legislation/Act1.nsf/0/8F3E822CC1A33B3DCA25751B007E3618/$file/1362008.pdf)

Australian Curriculum, Assessment and Reporting Authority (2009) website
<http://www.acara.edu.au/default.asp>

Australian Curriculum, Assessment and Reporting Authority Act (2008), No. 136, 2008, Accessed from <http://www.com.law.gov.au>

Australian Council for Educational Research (2007, December, 4) *PISA shows Indigenous students continue to struggle*, Media Release, Camberwell: ACER.

Bennett, R. (2009) *Formative assessment: Can the claims for Effectiveness be substantiated?* Paper presented at International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane.

Black, P. and Wiliam, D. (1998) *Inside the black box: Raising standards through classroom assessment*, London: King's College.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2002) *Working inside the black box: Assessment for Learning in the classroom*, London: King's College.

Black, P., Harrison, C., Lee, C., Marshall, B. & William, D. (2003) *Assessment for Learning in the classroom: Putting it into practice*, Maidenhead: Open University Press.

Bligh, A. (2008) Letter to Parent, Brisbane: Queensland Government.

Broadfoot, P. (1996) *Education, Assessment and Society*, Buckingham: Open University Press.

Centre for Educational Research and Innovation (CERI) (2005) *Formative assessment; improving learning in secondary classrooms*, Paris: OECD.

Cook, S.D.N. & Brown, J.S. (1999) Bridging epistemologies: the generative dance between organizational knowledge and organizational knowing, *Organization Science*, 10, 4, pp. 381- 400.

Dunleavy, G. (2009) Flying squads to rescue struggling schools, Accessed October, 15, 2009 from <http://www.bribanetimes.com.au/queensland/flying-squads-to-rescue-struggling-schools-20090908-ffco.html>

Earl, L.M. (2003) *Assessment as Learning*, Thousand Oaks: Sage.

Egelund, N. (2008) The value of international comparative studies of achievement – a Danish perspective, *Assessment in Education: Principles, Policy & Practice*, 15, 3, pp. 245-51.

Gardner, J. (ed.) (2006) *Assessment and Learning*, London: Sage.

Gee, J.P. (2003) Opportunity to Learn: a Language-based perspective on assessment, *Assessment in Education: Principles, Policy and Practice*, 10, 1, pp. 27-46.

Gillard, J. (2009) Indigenous Leadership in Education Institute: Stronger Smarter Summit – Brisbane, 28 September, 2009. Accessed October, 12, 2009, from http://www.deewr.gov.au/Ministers/Gillard/Media/Speeches/Pages/Article_090928_130255.aspx

Gipps, C. (1994) *Beyond testing: Towards a theory of educational assessment*, London: Falmer Press.

Glasson, T. (2009) *Improving Student Achievement: A Practical Guide to Assessment for Learning*, Curriculum Corporation: Carlton South.

Goldstein, H. & Thomas, S. (2008) Reflections on the international comparative surveys debate, *Assessment in Education: Principles, Policy and Practice*, 15, 3, pp. 215-222.

Griffiths, S. (2006) Failing to teach them how to handle real life, Times Online, Accessed 13 November, 2009 from <http://www.timesonline.co.uk/tol/news/article721863.ece>

Elwood, J. (2008) 'Gender Issues in Testing and Assessment', in P. Murphy and K. Hall, *Learning and Practice, Agency and Identities*, Sage: London, pp. 87-101.

Harlen, W. (2005) Trusting teachers' judgement: research evidence of the reliability and validity of teachers' assessment used for summative purposes, *Research Papers in Education*, 20, 3, pp. 245-70

Harrington, M. (2008) Australian Curriculum, Assessment and Reporting Authority Bill, Bill's Digest Number 60, 2008-2009, Accessed 12 May, 2009 from <http://www.aph.gov.au/library/Pubs/bd/2008-09/09bd060.htm>

Hopmann, S. T. (2008) *No child, no school, no state left behind: schooling in the age of Accountability*, Essay, Department of Education and Human Development, University of Vienna, Austria.

Hyslop, F. (2009) Assessment for Curriculum for Excellence, Retrieved November 15, 2009 from http://www.ltscotland.org.uk/Images/AssessmentforCfE_tcm4-565505.pdf

International Association for Educational Assessment (2008) *UK: Government abolishes national tests for 14-year-olds*. Retrieved November 21, 2008 <http://www.iaea.info/news.aspx?id=20>

James, M. (nd) *Assessment and Learning*, London: Institute of Education, London, University of London.

James, M. (in press) Assessment for Learning: research and policy in the (dis)United Kingdom, In R. Berry (ed) *Assessment Reform and Educational Change*, Springer.

Klenowski, V. (2005) Honouring teachers' professionalism: reclaiming teaching and claiming assessment, in D. Halpin & P. Walsh (Eds) *Educational commonplaces essays to honour Denis Lawton*, London: Institute of Education, University of London, pp. 103-119.

Klenowski, V. (2009) Australian Indigenous students: addressing equity issues in

assessment, *Teaching Education*, 20, 1, pp. 77-93.

Klenowski, V. and Adie, L. E. (2009) Moderation as judgement practice: Reconciling system level accountability and local level practice. *Curriculum Perspectives*, 29, 1, pp. 10-28.

Klenowski, V. and Wyatt-Smith, C. M. (2008) *Standards Driven Reform Years 1-10: Moderation an Optional Extra?* Paper presented at the Australian Association for Research in Education Conference, Brisbane.

Klenowski, V. and Wyatt-Smith, C. M. (2009) *The Role and Purpose of Standards in the Context of National Curriculum and Assessment Reform for Accountability, Improvement and Equity in Student Learning*, Paper presented at International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane.

Klenowski, V. & Gertz, T. (2009) *Culture-fair assessment: Addressing equity issues in the context of Primary Mathematics Teaching and Learning*, Paper presented at the Australian Council for Educational Research (ACER) Conference Assessment and Student Learning: Collecting, interpreting and using data to inform teaching, Perth, Western Australia, 16-18 August, 2009.

Lawton, D. (1992) *Education and politics in the 1990s: conflict or consensus?* London: Falmer Press

Luke, A., Woods, A., L., R., Bahr, M. and McFarland, M. (2004). *Accountability: Inclusive Assessment, Monitoring and Reporting*. Research Report prepared for the Queensland Indigenous Education Consultative Body. Queensland, Australia.

Masters, G.N., Forster, M., Matters, G.N., & Tognolini, J. (2006). *Australian Certificate of Education: A way forward*. Melbourne: Australian Council for Educational Research.

Matters, G. N., & Masters, G. N. (2007). *Year 12 Curriculum Content and Achievement Standards*. Brisbane: Australian Council for Educational Research.

Maxwell, G. (2007). *Implications for moderation of proposed changes to senior secondary school syllabuses*. Paper commissioned by the Queensland Studies Authority. Brisbane: Queensland Studies Authority.

McCarty, T. (2009) The impact of high-stakes accountability policies on Native American learners: evidence from research, *Teaching Education*, 20, 1, pp.7-29.

McGaw, B. (2009) *Building a national curriculum*, Keynote address at the Queensland Studies Authority senior schooling conference productive Pathways: Successful Futures, 19-20 March, 2009, Accessed April 14, 2009, from <http://www.qsa.qld.edu.au/publications/8206.html>

Murphy, P., Hall, K., McCormick, R. and Drury, R. (2008) *Curriculum, learning and society: investigating practice*, Study Guide, Masters in Education, Maidenhead: Open University.

Murphy, P. & Hall, K. (2008) *Learning and Practice: Agency and Identities*, London: Sage.

Murphy, P. (2009) Applying a sociocultural approach to assessment theory and practice: issues for summative and formative assessment. The Open University, Presentation at Queensland University of Technology, Brisbane, 11, September, 2009.

Nardi, E. (2008) Cultural biases: a non-Anglophone perspective, *Assessment in Education: Principles, Policy & Practice*, 15, 3, pp. 259-66.

National Curriculum Board (2008) *The Shape of the National Curriculum: A Proposal for Discussion*, Retrieved 5 November, 2008
http://www.ncb.org.au/our_work/preparing_for_2009.html

O'Neill, O. (2002) *A Question of Trust: The BBC Reith Lectures*, Cambridge: Cambridge University Press.

Patrick, R. (2008) Perspectives on change: A continued struggle for academic success and cultural relevancy at an American Indian school in the midst of No Child Left Behind. *Journal of American Indian Education*, 47(1), pp. 65-81.

Shavelson, R.J., Black, P.J., William, D., & Coffey, J. (2004). *On linking formative and summative functions in the design of large-scale assessment systems*. Retrieved 3 May, 2007, from www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20Aligning%20Formative%20and%20Summative%20Functions_Submit.doc

Stobart, G. (2005) Fairness in multicultural assessment systems, *Assessment in Education* Vol 12, No 3, pp. 275-287.

Stobart, G. (2008) *Testing times the uses and abuses of assessment*, London Routledge.

Third International Symposium on Assessment for Learning (2009) *Position Paper on Assessment for Learning*, March, Dunedin New Zealand.

Thomson, S (2008) Lessons to learn from high achievers, *The Age*
<http://www.theage.com.au/news/education-news/lessons-to-learn-from-high-achievers/2008/02/01/1201801039731.html> (accessed 30 October, 2008)

Watkins, C. (2003) *Learning a Sense-Maker's Guide*, Association of Teachers and Lecturers: London.

Wagemaker, H. (2008) Choices and trade-offs: reply to McGaw, *Assessment in Education: Principles, Policy & Practice*, 15, 3, pp. 267-78.

Welsh Assembly Government, (2006) Future assessment arrangements for Key Stages 2 and 3. Report on the findings and outcomes of the national consultation, held 31 October 2006 to 12 January 2007. Retrieved 15 November 2009, from
<http://wales.gov.uk/docrepos/40382/4038232/403829/Consultations/870830/future-assesment-ks2ks3-report-e?lang=en>

William, D. (2007) Changing classroom practice, *Educational Leadership*, 65 (4), 36-42.

William, D. (2008) International comparisons and sensitivity to instruction, *Assessment in Education: Principles, Policy & Practice*, 15, 3, pp. 253-57.

Willett, L. & Gardiner, A. (2009) *A critical analysis of the NAPLAN spelling test*, Paper presented at the International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane

Willis, J (2007). Assessment for Learning – why the theory needs the practice. *International journal of Pedagogies and Learning*. 3 (2) pp. 52 – 59

Willis, J (2008). *Assessment for learning: a sociocultural approach*. Paper presented at the Australian Association for Research in Education, December 2008, Brisbane.

Willis, J (2009a). *Building bridges to participation. Sociocultural case studies of AfL*. Paper presented at International Association for Educational Assessment, 35th Annual conference, September, 2009, Brisbane.

Willis, J. (2009b) Case Study One: Collaborative Assessment for Learning, Unpublished Thesis, Queensland University of Technology, Brisbane.