



The Experience of Education: The impacts of high stakes testing on school students and their families

Literature Review

Professor John Polesel, Ms Nicky Dulfer, Dr Malcolm Turnbull

January 2012

AUTHORS

Professor John Polesel

John Polesel is a Professor of Education in the Melbourne Graduate School of Education at the University of Melbourne. He has detailed knowledge of education systems across Australia and has undertaken a range of international and comparative studies. He has played a leading role in winning and conducting over forty major educational research grants and consultancies focussing on the transitions of young people from school to further education, training and the workplace.

He has written or co-written over ninety journal articles, book chapters and commissioned reports, mostly as first author. He has published refereed articles in some of the most prestigious international education journals including: *Oxford Review of Education*, *Comparative Education*; *Journal of Education Policy*; *Australian Journal of Education*; *Journal of Vocational Education & Training*; and *European Journal of Vocational Training*, and chapters in major Springer compilations. He recently co-edited a book in the influential Springer series. He has presented research papers to the JNET Conference at Oxford University in 2007, 2009 and 2011 and to the Faculty of Education at Cambridge University in 2009.

Ms Nicky Dulfer

Nicky Dulfer, a Lecturer within the Education Policy and Leadership Unit (EPL) at the University of Melbourne, has six years experience undertaking a range of research focussed on education policy. Her Masters thesis focussed on the Equity funding model in Victoria, and the ways in which schools use this funding. Since taking on a position at the University she has contributed to a range of projects including a review of the state funding model in Tasmania (2009) and a review of equity practices in Catholic schools in Victoria (2009). In 2008 Nicky was a lead contributor to a series of commissioned literature reviews investigating issues involved with early school leaving dropout, and a commissioned report outlining post-compulsory provision in the Western Metropolitan region of Melbourne.

Dr Malcolm Turnbull

Dr Malcolm Turnbull is a Research Fellow at the Youth Research Centre, Melbourne Graduate School of Education, and a former teacher in State Specialist schools and the ACFE sector. Recent projects include development of a Model of Education for Berry Street (with Dr Helen Stokes) and evaluation of refugee outreach and engagement projects for the Cities of Greater Dandenong and Brimbank. Dr Turnbull also works as a freelance historian with specific interests in the Australian Jewish Community, Classic British Detective Fiction and the Australian Folkmusic Revival.

Editing: The Whitlam Institute

ISBN: 978-1-74108-227-2

Copyright: The Whitlam Institute within the University of Western Sydney, 2012

FOREWORD

There are few things more important to parents than our sons' and daughters' education. Like the pencil markings on a door frame, it's as if each educational milestone becomes a marker of our own child's journey through life: the first flutter of separation at pre-school; teary-eyed independence at the kindy gate on the 'first day'; the claims of friendships over family in the dawning teen years on entering high school; the passage from childhood to the assertive young adult at matriculation.

There is so much wrapped up in schooling and seemingly so much at stake, that schools can become emotional cauldrons and the policies that shape them hotly contested.

It should come as no surprise then that the introduction of a national regime of standardised external testing would become a lightning rod of claim and counter-claim and a battleground for competing educational philosophies. The National Assessment Program – Literacy and Numeracy (NAPLAN) is a substantial educational reform. Its introduction has been a source of debate and argument.

Three years into the implementation of NAPLAN the Whitlam Institute and the Melbourne Graduate School of Education, together with our partner, the Foundation for Young Australians, believed it was timely to take a step back to look at the emergence of high stakes testing in Australia with the benefit of fresh, primary research.

The specific purpose of our project is to examine whether the regime of high stakes testing throughout the school years is in the best interests of the students. Over 1 million students sit NAPLAN tests each year.

We have been struck by the enthusiasm for this explicit focus on the young people themselves from all those consulted on the project and its design. While this might be implied from much of the educational research, apparently it is seldom explicit.

This Literature Review, by John Polesel and his colleagues, is the first project paper. It draws together the relevant, existing research and identifies gaps in the work to date; it does so through this project's particular lens, by drawing out the implications for students.

Though intended primarily to inform the research that is to follow, it is a valuable piece of work that we felt sure would be of interest to others.

We commend it to you.

Eric Sidoti
Director
Whitlam Institute

Jack Keating
Leader of the Education Policy
& Leadership Unit
Melbourne Graduate School of Education

EXECUTIVE SUMMARY

The Experience of Education: The impacts of high stakes testing on school students and their families

This literature review was commissioned by the Whitlam Institute within the University of Western Sydney (UWS) to provide context for the research project *The Experience of Education: The impacts of high stakes testing on school students and their families*. The project is a collaboration between the Whitlam Institute, the University of Melbourne and the Foundation for Young Australians.

The project seeks to examine whether the regime of high stakes testing throughout the school years is in the best interests of the students. The research aims to identify what the impacts of high stakes testing such as NAPLAN are on school students and their families, not only in terms of curriculum and learning but also with regards to students' health and well-being.

As high stakes testing becomes more deeply embedded in the educational landscape it is important that issues such as these be investigated as a basis for better informed policy making.

The literature review explores international and Australian literature on student's experience of standardised testing. It is based on a search and review of the scholarly research published on the implementation and impact of high stakes testing around the world. Much of the research is from the United States and the United Kingdom, as these are two of the nations with the longest histories of standardised high stakes testing and reporting of student achievement. Recent, related studies which have emerged in the Australian context are also included, although these are limited. The literature search focussed on the reported impact of high stakes testing on students (their health, well-being and learning) rather than the technical aspects of the design, implementation and reporting of the testing programs.

- The Australian National Assessment Program – Literacy and Numeracy (NAPLAN) was commenced in 2008 by the Australian Curriculum, Assessment and Reporting Authority (ACARA) in order to assess all students in Years 3, 5, 7 and 9 in Australian schools using national tests in Reading, Writing, Language Conventions (Spelling, Grammar and Punctuation) and Numeracy.
- Although there are several key differences between the Australian NAPLAN/MySchool model and the UK and USA models of student assessment programs, the publication of the results of the NAPLAN program on the MySchool website, with the associated media coverage, means that NAPLAN too may be labelled as a high stakes testing program.
- A range of concerns regarding the impact of high stakes testing is evident in the international literature. These range from the reliability of the tests themselves to their impact on the well-being of children. This impact includes the effect on the nature and quality of the broader learning experiences of children which may result from changes in approaches to learning and teaching, as well as to the structure and nature of the curriculum.
- Concerns in the international literature regarding the reliability of high stakes testing programs are largely centred on their capacity to achieve their own objectives of impartial, reliable and unbiased reporting designed to facilitate student, school and system improvement, without unintended negative consequences for the standing or reputation of particular schools.
- Considerable evidence may be found in the international literature regarding the negative impact of high stakes testing on students' well-being, including its potential to impact on students' self-esteem and lower teachers' expectations of children. There is also evidence of negative effects on service delivery and professional-parent relationships and stress, anxiety, pressure and fear experienced by students.

- Detailed findings such as these are not available in the Australian context, although similar concerns regarding NAPLAN have emerged from various sources, including a recent Australian survey of principals and teachers in independent schools, the recent Senate hearing into NAPLAN testing and reporting and a recent Queensland Studies Authority report, which expressed concern at the capacity of full cohort testing to lower the self-esteem, self-image and long-term confidence of under-performing students, thus widening the gap between them and higher-achieving peers.
- There is considerable evidence in the international literature of the impact that high stakes testing can have on the quality of the learning experience of children. Evidence has emerged that such testing can structure the educational experiences of students in ways that limit the development of the range of skills and literacies needed in the modern world, encouraging low-level thinking and promoting outcome measures rather than the intrinsic processes of learning and acquiring knowledge.
- Research on high stakes testing has also found that these tests may be having a negative impact on teacher pedagogies with a resultant degradation of students' experience of learning. The impact of this may be defined as a shift from a focus on the needs of the child to the needs of the evaluation and reporting process.
- Research also documents the impact of testing on the curriculum, showing that teachers will focus on the areas in which students will be tested, while reducing the proportion of class time devoted to curriculum areas not included in state tests. This influences curricular structures in terms of content, since the content of standardised tests defines what may be regarded as legitimate knowledge, and in the way in which content knowledge is presented in the classroom, with this increasingly aligning to the way it is presented and assessed in the tests, that is, as isolated and largely unconnected facts and pieces of information.
- Some evidence has also emerged in the Australian context of a narrowing of the curriculum as a result of high stakes testing.

What emerges consistently across this range of studies are serious concerns regarding the impact of high stakes testing on student health and well-being, learning, teaching and curriculum. Although much of the literature is focussed on the USA and the UK, the consistency of these findings raises legitimate questions and deep concern regarding the Australian experience.

The introduction of national standardised testing in Australia is a significant educational reform. It is important that such a reform be underpinned by rigorous research to ensure that it advances the interests of students. For this reason, it is important to investigate the extent to which we can extrapolate these findings of the largely negative impact of testing in the international context to the NAPLAN program recently implemented in Australia.

There is a particular need for research that explicitly recognises the best interests of the students as a primary consideration and which collects evidence from a range of stakeholders, including the children themselves.

INTRODUCTION

The Australian National Assessment Program – Literacy and Numeracy (NAPLAN) was commenced in 2008 by the Australian Curriculum, Assessment and Reporting Authority (ACARA), an independent authority “responsible for the development of a national curriculum, a national assessment program and a national data collection and reporting program that supports 21st century learning for all Australian students” (ACARA 2011).

Each year, all students in Years 3, 5, 7 and 9 in Australian schools are assessed using national tests in Reading, Writing, Language Conventions (Spelling, Grammar and Punctuation) and Numeracy. The program supplies individual student level reports designed to enable parents to see their child’s progress over the course of their schooling, and help teachers intervene with informed individual learning opportunities for their students. NAPLAN provides each school in the country with school level aggregated results which help individual schools to identify strengths and weaknesses within their teaching programs. Individual and school level data are also provided to the appropriate school system on the understanding that they can be used to target specific supports and resources to schools that need them most. The NAPLAN results for every school are published on the MySchool website which is accessible to the general public.

ACARA notes that the main purpose of NAPLAN testing is “to identify whether all students have the literacy and numeracy skills and knowledge that provide the critical foundation for other learning and for their productive and rewarding participation in the community” (ACARA 2010). ACARA also suggests that students benefit from these tests since:

NAPLAN tests provide information for parents, teachers and schools on individual student performance. Teachers and schools use this information, in conjunction with other information, to determine how well their students are performing and to identify any areas of need requiring assistance. National testing also enables consistency, comparability and transferability of results across jurisdictions (ACARA 2010).

In the light of these aims, this paper examines the impact of high stakes testing on school students and their families. It interrogates the evidence from the national and international literature on changes in learning, teaching and curriculum practices of schools with respect to the impact on students and evidence on the well-being of young people participating in programs such as NAPLAN.

The concept of testing student achievement is not new, although the current Australian approach may be said to have its origins in current educational policy structures in both the USA and the UK which are focussed around systems of high stakes, standardised testing – that is, systems that use the results of standardised tests to report on the achievement of students and schools in a public manner.

Public reporting and public accountability are central to the issue of what constitutes high stakes testing. Johnson et al. (2008) define high stakes tests in the United States as those which have consequences for student success (e.g. grade promotion or graduation), teacher accountability, the reputation of schools or the funding of schools. Similarly, Marchant (2004), also in the United States, speaks of the consequences of testing for students or their teachers and its impact on the public perception of schools. Ball (2008) notes that the parental perception of schools may be affected by high stakes testing, with a resulting detrimental effect of parental market choice being exerted on low SES schools that work in disadvantaged communities. Howe et al. (2001) also note the detrimental effect of “white flight” from low socio-economic schools in the United States, as a result of competition arising from the publication of high stakes testing results. Howe et al. note the unequal ability of different kinds of parents to “choose”, with middle class parents much better situated to make informed choices. They highlight the parental use of high stakes test results or ‘league tables’ as one of the key pieces of evidence that middle class parents consider when choosing a school for their child. Au (2008, p.501) describes the impact of the results of testing in terms of “sanctions or rewards to students, teachers, administrators, schools, school districts and other official bodies charged with the education of children”, while William (2010) focuses on the idea of accountability as central to the concept of high stakes testing.

In the Australian context, whether NAPLAN constitutes high stakes testing has been subject to some debate. However, Lobascher (2011, pp. 9-10) argues that the publication of the results of the NAPLAN program on the MySchool website, with the associated media coverage, means that NAPLAN too may be labelled as a high stakes testing program. Lingard (2010) supports this view, citing the implementation of the recommendations relating to Queensland’s recent poor NAPLAN performance contained in the Masters (2009) report.

That said, there are several key differences between the Australian NAPLAN/MySchool model and the UK and USA models. Firstly there is currently no suggestion that schools will be found to be ‘failing’ on the basis of their results and face closures. Schools that are found to be underperforming in the Australian context will be offered support and financial assistance under the current Federal Government policy. This is in stark contrast to the USA where the *No Child Left Behind Act* (2001) brought with it the threat of school closures. Secondly NAPLAN test scores are not used to determine grade promotion for any student. In the USA students get ‘held back’ if they have not performed adequately in their high stakes testing. Currently, in Australia, NAPLAN results have no impact on grade promotion. Finally the reporting of NAPLAN results in Australia is currently different to the previous models used in both the UK and the USA. In Australia, it is very difficult (though not impossible) to make a ‘league table’ and a range of contextual data is included on the MySchool website, including socio-economic (SES) data.

While it is hoped that this difference will lead to the avoidance of the 'naming' and 'shaming' regime that has existed in the UK for the past two decades, the absence of official 'league tables' has not prevented media constructing tables of their own nor parents presuming comparability, notwithstanding the questionable validity of the use of the data in this way.

Having made this distinction, there is a commonly cited range of objectives associated with the implementation of high stakes testing programs in education systems. The benefits of comprehensive testing using standardised measures are said to include:

- *Information that can be used for diagnosis (of individual students or teachers, of schools, of school programs)*
- *Efficiencies from alignment, when the tests are matched to curricular standards and teachers teach to those standards*
- *Motivation to study and to attain goals* (Phelps 2006, p.19).

The benefits identified in the research include the maintenance of academic standards and the identification of students not meeting those standards so that remedial programs may be implemented to address individual students whose learning needs are not being met. It is also argued that high stakes standardised testing supplies reliable and unbiased information on student performance outside a student's own school or district (Phelps 2006). Standardised testing is also claimed to be more reliable than teacher grading and testing, which is regarded as susceptible to reliance on non-cognitive influences and outcomes, including student class participation, perceived effort, progress over the period of the course and behaviour (Phelps 2006). Sloane and Kelly (2003) argue that testing may provide students with clearer information regarding their knowledge, skills and potential and may thus motivate them to work harder, although they also acknowledge the potential for testing to frustrate students and discourage them. Armrein and Berliner (2002) also support the arguments that testing may lead to clearer information about what is important in the curriculum and improve student motivation.

There is also evidence that the alignment of the curriculum with high stakes testing may result in greater curriculum consistency within and across schools, ensuring a command of agreed competencies and transferability of experiences across constituencies and across schools (Clarke et al. 2003; Crocker 2004; Jones 2007). These researchers have also argued that testing may result in teachers developing more explicit expectations of what they should be teaching (Clarke et al. 2003; Jones 2007). These studies suggest that the alignment of curriculum with testing requirements may lead to more focussed and consistent curricular approaches.

In Australia, Anderson (2009) points out that the NAPLAN tests can be used to enhance students' thinking skills and confidence as long as teachers avoid the temptation to 'practise' taking tests. Collier (2010) identifies NAPLAN's usefulness as a diagnostic tool for schools while Santiago et al. (2011) acknowledge its credibility in enabling greater consistency, transferability and comparability of results across jurisdictions. They argue, also, that employment of a common scale provides useful information about performance and/or growth in individual student achievement. In a recent submission by the Australian College of Educators (ACE) to the NAPLAN Senate Inquiry, the identification of schools that are 'punching above their weight' was seen as providing a starting point for capacity building or collaboration between schools (ACE 2010).

There is also, however, a range of concerns regarding the impact of testing evident in the international literature. These range from the reliability of the tests themselves (and their usefulness and efficacy in meeting their stated aims) to their impact on the well-being of children. This impact includes the effect on the nature and quality of the broader learning experiences of children which may result from changes in approaches to learning and teaching, as well as to the structure and nature of the curriculum.

CRITIQUES OF TESTING

RELIABILITY

A first question in establishing the impact of high stakes testing on students is to ask whether the tests themselves are reliable, valid and desirable on their own terms as a means of assessment.

To begin, there is a considerable body of well-documented research challenging high stakes testing programs, including both international studies such as the Program for International Student Assessment¹ (PISA) and the Trends in International Mathematics and Science Study² (TIMSS) and national programs of assessment, on the grounds of validity, reliability, usefulness or statistical significance. Such studies include Harlow and Jones (2004), who question whether some of the items in TIMSS can effectively elicit the knowledge held by the student, and Flores and Clark (2003) who argued that the Texas-based Texas Assessment of Academic Skills (TAAS) did not measure what it purported to measure and was used in invalid ways. Meier (2002) outlines the potential for substantial technical unreliability in the collection and analysis of test data, as well as machine and software errors, while Curtis (2007) notes the unreliability and errors in the allocation of grades to children in the Standard Assessment Tests (SATs) in the UK.

In Australia Wu (2010) states that NAPLAN results should not be used to measure school performance because the margins of error for measuring student performance and school performance are too high. Munro (2010, p.3) states that "At best, the test outcomes tell us what students did at a particular time, under particular conditions, on a limited number of tasks. Using them to make cohort, school-level or time comparisons needs to take this into account". Particular concern has been expressed, in this regard, at the NAPLAN's failure to take into account the complexities typically found within Australia's school populations (Australian Education Union 2010).

At a less technical level, Paris (2000) finds that high stakes testing does not provide effective measures of an individual's learning (independent of differences in background and motivation) and that its usefulness as an information source for parents and as a diagnostic tool for teachers are both limited. Similarly, the report of a survey of independent school principals and teachers in Australia found that not more than approximately one half of those surveyed considered NAPLAN test results as providing a useful diagnostic tool (Athanasou 2010).

Caldwell (2010) has also criticised the use of NAPLAN tests as a diagnostic tool, as there is a lag of five months between students being tested and teachers receiving the results. Peter Hill, the CEO of the ACARA, has himself stated that the NAPLAN tests are not able to be used as diagnostic tests due to this time lag (Senate References Committee on Education, Employment & Workplace Relations 2010).

The argument that publishing NAPLAN data will result in public pressure which will drive school improvement has come under considerable criticism with the Australian College of Educators (ACE). ACE notes that schools that are under pressure to perform well in the NAPLAN tests are more likely to encourage certain students not to take the test so that the school results look better than they may really be (ACE 2010). So in effect, publishing data may lead to inaccurate and unreliable reporting of school data rather than providing any motivation or rationale for school improvement. There is also tension surrounding the publication of NAPLAN test results. On the one hand, there is parents' right to information about their child, and on the other a concern that the tests are being used by parents to make comparisons between schools (Senate References Committee on Education, Employment & Workplace Relations 2010). Lingard (2010, p.130) also points out the danger of the "potential for the 'naming' and 'shaming' of poorly performing schools, which most likely will be situated in poor communities and which would fail to recognise the very strong relationship between socio-economic status (SES) and student performance".

These concerns are cited here to provide important context for the examination of the impact of high stakes testing programs and because they will necessarily have an influence on any judgement of the quality of such programs, including their capacity to achieve their own objectives of impartial, reliable and unbiased reporting designed to facilitate student, school and system improvement, without unintended negative consequences for the standing or reputation of particular schools.

The principal concern of this paper, however, is to examine the evidence regarding the impact of testing regimes on learning, teaching and curriculum and how these may impact, directly or indirectly, on students' well-being.

STUDENT HEALTH AND WELL-BEING

Most of the research and literature relating to the impact of high stakes testing on students' health and well-being, like much of the other work in this field, is from the United States and the United Kingdom. The limited research conducted in Australia, which tends not to be directly related to health and well-being, is included at the end of this section. Additional anecdotal evidence is referred to in the conclusion of this literature review, but cannot be cited as hard evidence, given the absence of the controls normally associated with published research (ethics approvals, peer review, and the like).

1 An international study commissioned by the OECD which evaluates education systems by testing the skills and knowledge of 15 year-olds in participating nations.

2 An international study commissioned by the International Association for the Evaluation of Educational Achievement which measures progress in eighth-grade mathematics and science achievement in participating nations.

Certainly, in the international context, concerns regarding the negative impact of high stakes testing on students' well-being date back at least to the early 1990s. Perrone (1991) argued that such testing had the potential to impact on students' self-esteem and lower teachers' expectations of children in a study conducted in the United States. Skrtic (1995), cited in Peters and Oliver (2009, p.273) and also in the USA, describes the operation of the testing regime as a "machine bureaucracy" which operates as a means of regulating and controlling schools, teachers and students, with scant regard to the quality of teaching and learning. Stiggins (1999), describing research conducted in the USA, speaks of the pressure which testing puts on students to do well which, when not accompanied by adequate support, can lead to anxiety and a sense of futility. Similarly, Lewis (2000), in the USA outlines the potential for the anxiety and stress arising from high stakes testing to contribute to dropping out by discouraged students. Schroeder's research (2006), in America aligns with these views, underlining the pressure on teachers and students to perform well, affecting classroom behaviours and social interactions between students and teachers, with a particularly negative impact on children from minorities and children with disabilities. Gregory and Clark (2003), in a study of the impact of testing in Singapore, found a sense of futility among one third of Singaporean children, as well as evidence of excessively long hours of study after school, physical punishment of children by a majority of parents in cases of poor achievement and significant numbers of children requiring psychiatric treatment.

Thomas (2005) describes the "collateral" damage experienced by all the stakeholders in the USA from politicians and administrators to teachers, parents and students, as they attempt to explain instances of failure or poor results. Reporting on the impact of one USA testing regime (the North Carolina ABC Program) as perceived by school counsellors (i.e. the nominated testing co-ordinators within their respective settings), Brown et al. (2004) argue that such positives as increased teacher accountability, greater parental involvement and increased teaching consistency, are more than offset by negative effects on service delivery and professional-parent relationships (at one level) and by the stress, anxiety, pressure and level of fear experienced by students (at another). They cite specific references by their survey respondents to children feeling incompetent or being labelled by their teachers, to an increase in suspensions and problem behaviours, to test avoidance and lowered student self-esteem and confidence. Again in the USA, Paris and McEvoy (2000) describe instances of children "freezing" with fear during tests and experiencing anxiety or physical distress; Flores and Clark (2003) cite instances of emotional, psychological and physical distress, including impact on self-esteem, inability to sleep, confusion, frustration, headaches and throwing up, while Emery and Ohanian (2004) ask why we send our children "to a place that makes them vomit" rather than one where they are nurtured. Madaus et al. (2009) in the USA echo these findings in a study which found evidence of children's exhaustion, frustration and crying emerging as responses to the stress provoked by testing.

Critiques of high stakes testing have tended to rely, as has this paper, on data based on adult perceptions of the emotional and physical effects of testing on students. It must be conceded that much less information has been gathered from the primary stakeholders, i.e. children and young people themselves. Notable exceptions include Wheelock et al. (2002), whose analyses of the drawings of American middle and high school students confirmed a range of emotions associated with testing, encompassing anxiety, boredom, anger, motivation and confidence. Taking the approach a step further, Triplett & Barksdale (2005) have explored the perceptions of 225 American third through sixth graders via a combination of their post-test drawings and writings. On the basis of reiterated artistic and written allusions to nervousness, anxiety, feelings of isolation and alienation and concerns about the consequences of failure, they have highlighted a "prevailing negativity" in the children's responses.

These are stark examples, but they exemplify some of the ways in which children's identifications as learners (Skeggs 1997) are constructed through the assessment process. Cohen (1989) outlines the potential for testing to lead to children labelling themselves as failures at a very early stage of their learning journey. Reay and William (1999) concur in a case study from the UK which describes the way in which one child's reported performance on standardised tests led to her constructing herself as a failure, rather than as a creative and accomplished problem-solver. Reay and Williams (1999) note negative self-perceptions as a result of test results and anxiety about impending tests emerging even in higher-achieving students. Even more disturbingly, children expressed discomfort regarding the impact of test results on their future life prospects, fearing that they might be indicative of failure and hardship (Reay and William 1999). This same study found evidence of aggression and jealousy towards some higher achieving students emerging as a result of the reporting of test outcomes.

Peters and Oliver (2009) in America describe the tendency of testing towards standardisation of educational practices and of students themselves, a tendency leading to a disregard of differences in the needs, talents and achievements of different students – "especially those from minorities and those with disabilities and special education needs" (2009, p.273), a theme echoed by Cunningham and Sanzo (2002), who note the relative lack of home support available to children from low socio-economic status backgrounds. Similarly, Perrone (1991) specifies younger children from minority or socially disadvantaged backgrounds as especially vulnerable to the "deleterious" effects of the testing regimes of the 1980s in the United States.

Detailed findings such as these are not available in the Australian context. However, a recent Australian survey of principals and teachers in independent schools provides some evidence of negative impacts in Australia, citing evidence of pressure on students and teachers as a result of the publication of NAPLAN results and worry about exams causing some primary school children, for example, to lose sleep before the tests (Athanasou 2010).

The recent Senate hearing into NAPLAN testing and reporting has also received a number of submissions from individual parents and schools outlining concerns regarding the labelling of students, demoralisation of staff in schools that appear to be underperforming and the negative impacts of the pressure to perform well on individual students in the high stakes testing regime (Senate References Committee on Education, Employment & Workplace Relations 2010) although these cannot be tested or evaluated in scientific terms. More broadly, the Queensland Studies Authority (2009) has expressed concern at the capacity of full cohort testing to lower the self-esteem, self-image and long-term confidence of underperforming students, thus widening the gap between them and higher-achieving peers.

LEARNING

A recurring theme in the literature is the impact that high stakes testing can have on the quality of the learning experience of children. Au (2008), for example, commenting on the prevalence of high stakes testing in the USA and the UK, has claimed that testing can structure the educational experiences of students in ways that limit the development of the range of skills and literacies needed in the modern world. Anagnostopoulos (2003), in a study based in the United States, observed that the focus on examination requirements in Chicago accountability tests could limit the broader interpretation of a text to a narrow focus on plot and character, excluding important elements of a political and ideological nature in a text such as *To Kill a Mockingbird*. Paris (2000), in a research study also focussed on the United States, argues that testing regimes encourage low-level thinking and promote the valuing of outcome measures rather than the intrinsic processes of learning and acquiring knowledge. Gulek (2003) notes that multiple-choice modes of testing still dominate testing approaches in the United States and argues for considerable care in interpreting such data, as well as the need for a wide variety of testing approaches. Au (2008) ironically notes that, while high stakes testing may benefit middle class children in terms of conferring competitive advantage, it acts against their longer term interests in achieving the flexibility and skills which they need to thrive in a modern knowledge-based economy.

While most of the current evidence in this area comes from the international arena, Lobascher (2011, p.13), in the Australian context, notes concerns expressed by the Queensland Studies Authority (QSA), that testing encourages “methods of teaching that promote shallow and superficial learning rather than deep conceptual understanding and the kinds of complex knowledge and skills needed in modern, information-based societies” (QSA 2009, p.3). Alexander (2010) supports this view, noting that both literacy and numeracy skills are tested within NAPLAN without reference to the subject context in which those skills need to be applied.

These are important considerations, given the broadly accepted contention that education should not be narrowly focussed on limited skills, to the detriment of the child’s broader social and personal development, but rather “shall be directed to the full development of the human personality and the sense of its dignity” (Office of the United Nations High Commissioner for Human Rights 1966), as expressed in Article 13 of the International Covenant on Economic, Social and Cultural Rights, to which Australia is a signatory. These are sentiments expressed in greater detail in the Convention of the Rights of the Child (Office of the United Nations High Commissioner for Human Rights 2001) General Comment Number 1, which includes respect for parents, cultural identity, human rights and freedom, as well as the development of understanding, peace, tolerance and equality as basic aims of the learning process.

TEACHING

Research on high stakes testing has also found that these tests may be having a negative impact on teacher pedagogies with a resultant degradation of students’ experience of learning. Au (2008a) presents the argument that teachers in the US are moving back to more teacher-centred instructional approaches in an effort to transmit the level of content required by the tests, with the result that students are experiencing a more limited range of activities in the classroom and have fewer opportunities to experience excursions and field trips. Cunningham and Sanzo (2002) have also found that high stakes testing impacts negatively on creative and effective teachers, leading to cramming for tests rather than instruction. Reay and William (1999) also note the shift from the supportive and collaborative learning approaches which typify many schools’ current approaches, as exemplified by the emphasis on group work and enquiry-based learning, to increasingly competitive and individualistic attitudes and ways of learning.

In the Australian context, the NAPLAN Senate report notes that a number of submissions to the Inquiry outline concerns regarding schools restricting the amount of enquiry-based learning and an increase in teacher instruction time. Similarly, many submissions were concerned that teachers have increasingly been ‘teaching to the test’ (Senate References Committee on Education, Employment & Workplace Relations 2010). Such findings echo predictions made by Hargreaves (1994), in light of international findings, of teachers increasingly becoming technicians, obliged to deliver a prescribed and narrow product into which they have had little input. Hargreaves has argued that such de-skilling of teachers is part-and-parcel of systematic separation of planning from execution in the workplace, i.e. one of the defining features of post-modernism. More recently, Klenowski (2010, 2011) has speculated on the unintentional impact of NAPLAN in reducing culturally responsive teaching and diminishing trust in teacher professionalism.

Lobascher (2011, pp.15-16), in a summary of the research into the negative effects of high stakes testing on pedagogy, cites a range of studies to support the argument that testing detracts from the creativity of teachers and removes the intrinsic motivation of love of learning in students, replacing this with extrinsic rewards and threats which reduce enjoyment of the teaching and learning experience (Anagnostopoulos 2003; Au 2007; Jones 2007; QSA 2009; Williams 2009). Similarly, Lingard (2009) expresses concern that, notwithstanding claims to the contrary, NAPLAN has indeed become 'high stakes' with predictable negative impacts that include teacher frustration at constraints on their opportunities to practise authentic pedagogies and an undermining of their sense of professional worth. Citing the example of England, Lingard (2009, p.16) decries a "culture of performativity" that affects "the very souls of teachers".

The impact of this culture on the learning experiences of children may be defined as a shift from a focus on the needs of the child (and the associated responsiveness to the child's needs) to a focus on the needs of the system, from "child-focussed" education and the "best interests of the child" (Office of the United Nations High Commissioner for Human Rights 2001) to the needs of the evaluation and reporting process.

CURRICULUM

A related but distinct element of the research relates to the impact of testing on the curriculum itself. The benefits of a broad curriculum that encourages creativity, problem-solving, the development of contemporary skills, and that provides for physical activity and engaged learning, are well established (Ravitch 2010).

A common finding in the literature is that teachers will focus on the areas in which students will be tested, while reducing the proportion of class time devoted to curriculum areas not included in state tests, as Abrams (2004) found in the US context. In the same study, Abrams (2004) noted that a large proportion of teachers were reluctant to use computers to teach writing since the introduction of the state tests, as these tests required students to provide handwritten responses. Also in the USA, David (2011) has argued that the high stakes testing programs spawned by the *No Child Left Behind Act* (2001) have resulted in drastic changes to the proportions of time allocated to different subject areas, with the content of the tests largely determining the basis of the curriculum, especially in those schools identified as being low performing. Jones et al. (2003), cited in Lobascher (2011, p.12), note that the social sciences, science, arts and physical education have all seen reductions in the time allocated to them as a result of the demands of high stakes testing programs, particularly in the United States. Madaus et al. (2009) provide similar evidence of neglect of non-tested subjects, but in addition they find that even recess periods may be abridged to accommodate preparation for the tests. Reay and William (1999) note similar concerns regarding the narrowness of the curriculum among children participating in a study in the UK.

Au (2008a) argues that high stakes testing influences curricular structures in two main areas. The first is content, since the content of standardised tests defines what may be regarded as legitimate knowledge. As Jones et al. (2003) above and others (Renter et al. 2006) have noted, this leads to a reduction of the role played by subjects which do not feature in the tests, such as art, science, and social studies in the United States. The second is the way in which content knowledge is presented in the classroom, with this increasingly aligning to the way it is presented and assessed in the tests, that is, as isolated and largely unconnected facts and pieces of information (Pedulla et al. 2003).

Sabol (2010) found that high stakes testing in the USA had affected students' work in visual arts subjects, with educators claiming that the greater emphasis on literacy and mathematics in all subjects had led to reduced time spent on visual arts curriculum content. This diminishing of the arts is particularly concerning for communities where arts programs have been found to have positive outcome for 'at-risk' students (Ewing 2011). In the USA, Catterall et al. (1999) found that students with high levels of arts education performed better than students with low levels of arts education in all standard test measures, regardless of their socioeconomic status background. They found that "students who were involved in music and drama achieved higher levels of success in mathematics and reading than those who were not" (Catterall et al. 1999).

The sole Australian study with some relevance to this aspect of high stakes testing was a survey carried out on behalf of the Independent Education Union of Australia which found that high proportions of both teachers and principals believed that preparation for NAPLAN affected the time that could be devoted to regular classroom work (Athanasou 2010). However, fears have been expressed by Australian educators that this narrowing of the curriculum as a result of high stakes testing may also be an issue in Australia (Ewing 2011). Furthermore Wyn (2009) has expressed concerns that, in the Australian context, the valuing of narrow assessment and reporting strategies and limiting the subjects offered may contribute to inequitable outcomes for students.

CONCLUSION / SUMMARY

"I conclude that students learn about much more than the three Rs through their experiences with high-stakes testing, and argue that future research should attend to the social dimensions of these experiences" (Booher-Jennings 2008, p.149).

The evidence presented in this paper draws from a range of research studies which have investigated the impact of a range of testing regimes, many of them quite different in their approach. What emerges consistently across this range of studies are serious concerns regarding the impact of high stakes testing on student health and well-being, learning, teaching and curriculum. It is acknowledged that much of this is international research, particularly focussed on the USA and the UK. However, the consistency of these findings raises legitimate questions and deep concern regarding the Australian experience. For this reason, it is important to investigate the extent to which we can extrapolate these findings of the largely negative impact of testing to the NAPLAN program recently implemented in Australia.

As noted in the introduction, there are several key differences between the Australian NAPLAN/MySchool model and the international models. These include differences in the use of sanctions arising from results of testing, such as grade failure, grade promotion or school closure, as well as targetted assistance in the form of support for poorly performing schools. However, the impact on published testing results on parents' selection of schools and the unequal impact of this mechanism on families from different socio-economic status backgrounds, a consistent feature of the international research, cannot be discounted in the Australian context.

Moreover, the international literature and research reflects a much longer experience of high stakes testing (and a weighty body of associated research) which allows a much more considered and thorough analysis of the impact of these programs over a considerable period of time, compared with the relatively recent introduction of these programs in Australia. Most importantly, this substantial body of work reveals consistent and worrying concerns which emerge in almost all of the literature which reports research in this field. A narrowing of curriculum, a restriction in the range of skills and competences learnt by students and a negative impact on the ability of teachers to employ creative and engaging pedagogies are all cited in the extensive body of literature which relates to this field.

With respect to student well-being, the number and range of studies dealing with this issue directly are more limited. The authors of this report were not able to find any scientific evidence relating to the impact of high stakes testing on the well-being of children in the Australian context, although some examples of case studies occur in Athanasou (2010) and there are instances of anecdotal evidence which have been reported in the media. These include claims of stress, bed wetting and anxiety disorders arising from the stress associated with high stakes testing (O'Keeffe, 2011), as well as claims of suicides related to the impact of the Higher School Certificate in NSW (Snow 2006).

The 2010 Senate Inquiry into NAPLAN (Senate References Committee on Education, Employment & Workplace Relations 2010) was provided with a number of submissions from individual teachers, schools and educational organisations, but, again, many of the submissions were based on anecdotal evidence or expressed individual opinion only. At a more general level, the proportion of young people reporting school or study problems as an issue of personal concern has risen to its highest level ever in the annual national survey of young Australians conducted by Mission Australia (Mission Australia 2011).

The introduction of national standardised testing in Australia is a significant educational reform. It will inevitably have a direct bearing on student well-being and a further impact on students' learning and experience of education by virtue of its effects on educational practices. Consequently, it is important that such a reform be underpinned by rigorous research to ensure that in both respects it advances the interests of those students.

The extent of the concerns raised in the international literature and the accompanying lack of evidence relating to the Australian approach contained within the NAPLAN model suggest that research into this issue in Australia is now necessary and overdue. Furthermore, they suggest that the research needs to be rigorous, comprehensive of different levels and sectors of schooling and inclusive of the views of students, parents, teachers, school leaders and policy makers.

In particular, this research needs to focus on a number of key issues. These include:

- The impact of NAPLAN reporting on school enrolments
- The impact of testing on children's health and well-being
- The impact of testing on curriculum
- The impact of testing on teaching approaches
- The impact of testing on children's learning.

There is a particular need for research that explicitly recognises the best interests of the students as a primary consideration.

The range and nature of studies cited in this literature review further suggest that research in this area needs to include both qualitative and quantitative data, including both the views of the stakeholders and quantitative evidence of changes in outcomes for students generally and for different student groups. Given the evidence from the international sphere and given the importance and impact of this major initiative in Australian schooling, it is timely to look carefully at the introduction on a large scale of high stakes testing programs in Australia. In particular, the paucity of evidence based on the perceptions of the most important stakeholders themselves – the children – suggests that a comprehensive study which includes the views of Australian school children is now well overdue.

GLOSSARY

ACARA	Australian Curriculum, Assessment and Reporting Authority
ACE	Australian College of Educators
IEA	International Association for the Evaluation of Educational Achievement
NAPLAN	National Assessment Program – Literacy and Numeracy
PISA	Program for International Student Assessment
QSA	Queensland Studies Authority
SATs	Standard Assessment Tests
SES	Socio-Economic Status
TAAS	Texas Assessment of Academic Skills
TIMSS	Trends in International Mathematics and Science Study

REFERENCES

- Abrams, LM 2004, 'Teachers' Views on High-Stakes Testing: Implications for the Classroom', *Policy Brief: Education Policy Studies Laboratory*, Arizona State University College of Education, Arizona.
- Australian Curriculum, Assessment and Reporting Authority (ACARA) 2011, *Home page*, <http://www.acara.edu.au/default.asp> (accessed 4 November 2011)
- Australian Curriculum, Assessment and Reporting Authority (ACARA) 2010, *NAPLAN Frequently Asked Questions*, [<http://www.naplan.edu.au/faqs/napfaq.html>] (accessed 12 December 2011)
- Australian College of Educators 2010, Submission to the *Inquiry into the Administration and Reporting of NAPLAN Testing*, Senate References Committee on Education, Employment & Workplace Relations, Canberra
- Alexander, RJ (ed.) 2010, *Children, their World, their education: final report of the Cambridge Primary Review*, Routledge, London.
- Anagnostopoulos, D 2003, 'Testing and student engagement with literature in urban classrooms: a multi-layered perspective', *Research in the Teaching of English*, Vol. 38, No. 2, pp. 177–212.
- Anderson, J 2009, 'Using NAPLAN items to develop students' thinking skills and build confidence', *Australian Mathematics Teacher*, Vol. 65, No.4, pp. 17-24.
- Armrein, A & Berliner, D 2002, 'High Stakes testing, Uncertainty and Student Learning', *Education Policy Analysis Archives*, Vol. 10, No. 18, viewed 1 September 2011, <<http://epaa.asu.edu/epaa/v10n18/>>.
- Athanasou, J 2010, 'NAPLAN and MySchool Survey', *Independent Education Union of Australia*, Barton, Australian Capital Territory.
- Au, W 2007, 'High-stakes testing and curricular control: a qualitative metasynthesis', *Educational Researcher*, Vol. 36, No. 5, pp. 258–267.
- Au, W 2008, 'Between education and the economy: high-stakes testing and the contradictory location of the new middle class', *Journal of Education Policy*, Vol. 23, No. 5, pp. 501-513.
- Au, W 2008a, 'Devising inequality: a Bernsteinian analysis of high-stakes testing and social reproduction in education', *British Journal of Sociology of Education*, Vol. 29, No. 6, pp. 639-651.
- Australian Education Union 2010, 'Facts about NAPLAN, MySchool and League Tables', viewed 1 September 2011, <www.sstuwa.org.au/index.php?option=com_content&view=article>.
- Ball, SJ 2008, *The education debate*, Bristol: The Policy Press.
- Booher-Jennings, J 2008, 'Learning to label: socialisation, gender, and the hidden curriculum of high-stakes testing', *British Journal of Sociology of Education*, Vol. 29, Issue 2, pp. 149-160.
- Brown, D, Galassi, J & Alos, P 2004, 'School Counselors' perceptions of the impact of high-stakes testing', *Professional School Counseling*, October, viewed 20 October, 2011 <http://findarticles.com/p/articles/mi_mOKOC/is_1_8/ai_n6335435/?tag=content;col1>
- Caldwell, BJ 2010, 'How can we get off this bus?', paper presented at a symposium on the theme 'NAPLAN and the My School Website: Issues and the Way Forward' at the 2010 International Education Research Conference of the Australian Association of Research in Education (AARE), University of Melbourne, 30 November 2010.
- Catterall, J., Chapleau, R., & Iwanaga, J. 1999 *Involvement in the Arts and Human Development: General Involvement and Intensive Involvement in Music and Theater Arts*, The Imagination Project at UCLA Graduate School of Education and Information Studies, University of California at Los Angeles. <<http://www.aep-arts.org/highlights/coc-release.html>>
- Clarke, M, Shore, A, Rhoades, K, Abrams, L, Miao, J & Li, J 2003, 'Perceived effects of state-mandated testing programs on teaching and learning: findings from interviews with educators in low-, medium-, and high-stakes states', National Board on Educational Testing and Public Policy, Boston College, viewed 24 September 2010, <<http://www.bc.edu/research/nbetpp/statements/nbr1.pdf>>.
- Cohen, D 1989, 'Tests Fail to Pass Their Exams', in G Shrubbs (ed), *Mass Testing Tonic or Toxin?* Counterpoints No. 1, Australian Education Network.
- Collier, J 2010, 'Ranking by NAPLAN results rates a fail', *The Age*, 3 February, viewed 13 September 2011, <<http://www.theage.com.au/opinion/politics/ranking-by-naplan-results-rates-a-fail-20100202-nay4.html>>.
- Crocker, L 2004, 'Teaching For the Test: How and Why Test Preparation is Appropriate', in RP Phelps (ed), *Defending Standardized Testing*, Lawrence Erlbaum Associates, New Jersey, pp. 159–174.
- Cunningham, W & Sanzo, T 2002, 'Is high-stakes testing harming lower socio-economic schools?', *National Association of Secondary School Principals NASSP Bulletin*, June, pp.62-75.
- Curtis, P 2007, 'Test results for third of primary students wrong', *The Guardian*, 2 November, viewed 13 September 2011, <http://www.guardian.co.uk/uk/2007/nov/02/sats.politics>.
- David, J 2011, 'What students need to learn: high-stakes testing narrows the curriculum', *Educational Leadership*, Vol. 68, No. 6, pp. 78-80.
- Elley, W 2004, 'New Zealand literacy standards in a global context: The uses and abuses of international literacy surveys', *English Teaching: Practice & Critique*, Vol. 3, No. 1, pp. 32-45.
- Emery, K & Ohanian, S 2004, *Why is corporate America bashing our public schools?*, Portsmouth, NH.
- Ewing, R 2010a, *Curriculum and assessment: A narrative approach*. Melbourne: Oxford University Press.
- Ewing, R 2011, "The arts and Australian education: realising potential" in *Australian Education Review*, No. 58. <http://research.acer.edu.au/aer/11>
- Flores, B & Clark, E 2003, 'Texas Voices Speak Out About High Stakes Testing: Preservice Teachers, Teachers and Students', *Current Issues in Education* (online), Vol. 6, No. 3, viewed 29th August 2011, <http://cie.asu.edu/volume6/number3/index.html>
- Gregory, K & Clarke, M 2003, 'High Stakes Assessment in England and Singapore', *Theory into Practice*, Vol. 42, No. 1, pp. 66-74.
- Gulek, C 2003, 'Preparing for High-Stakes Testing', *Theory into Practice*, Vol. 42, No. 1, pp.42-50.
- Hargreaves, A 1994, *Changing teachers, changing times: teachers' work and culture in the post-modern age*. Cassell, London.
- Harlow, A, & Jones, A 2004, 'Why Students Answer TIMSS Science Test Items the Way They Do', *Research in Science Education*, Vol. 34, No. 2, pp. 221-238.
- Hilton, M 2006, 'Measuring Standards in Primary English: Issues of Validity and Accountability with Respect to PIRLS and National Curriculum Test Scores', *British Educational Research Journal*, Vol. 32, No. 6, pp. 817-837.
- Howe, K, Eisenhart, M, & Betebenner, D 2001, School Choice Crucible: A Case Study of Boulder Valley. *Phi Delta Kappan*, 83(2): 137-146.
- Johnson, D, Johnson, B, Farenga, S & Ness, D 2008, *Stop High Stakes Testing: an Appeal to America's Conscience*, Lanham, MD.
- Jones, BD 2007, 'The unintended outcomes of high-stakes testing', *Journal of Applied School Psychology*, Vol. 23, No. 2, 65-86.
- Jones, MG, Jones, BD & Hargrove, TY 2003, *The Unintended Consequences of High-Stakes Testing*. Rowman & Littlefield, Maryland.
- Klenowski, V 2010, 'Are Australian assessment reforms fit for the purpose? Lessons from home and abroad', *Queensland Union of Teachers Professional Magazine*, November, pp. 10-15.
- Klenowski, V 2011, 'Assessment for learning in the accountability era: Queensland, Australia', *Studies in Educational Evaluation*, Vol. 37, No. 1, pp. 78-83.

- Lewis, A 2000, 'High Stakes Testing: Trends and Issues', *Mid continent research for education and learning: policy brief*, April 1-8.
- Lingard, B 2009, 'Testing times: the need for new intelligent accountabilities for schooling', *Queensland Teachers Union Professional Magazine*, November, pp. 13-20.
- Lingard, B 2010, Policy borrowing, policy learning: testing times in Australian schooling. *Critical Studies in Education*, Vol. 51, No. 2, pp. 129-147.
- Lobascher, S 2011, What are the Potential Impacts of High-stakes Testing on Literacy Education in Australia? *Australian Journal of Language & Literacy*, Vol. 34, No. 2, pp. 9-19.
- Madaus, G, Russell, M & Higgins, J 2009, *The Paradoxes of High Stakes Testing: how they affect students, their teachers, principals, schools and society*, Charlotte, NC.
- Marchant, GJ 2004, 'What is at Stake with High Stakes Testing? A Discussion of Issues and Research', *Ohio Journal of Science*, Vol. 104, No. 2, pp. 2-7.
- Masters, GN 2009, *A Shared Challenge: Improving Literacy, Numeracy and Science Learning in Queensland Primary Schools*, viewed 27 September 2010, <<http://education.qld.gov.au/mastersreview/pdfs/final-report-masters.pdf>>.
- Meier, D 2002, *In schools we trust: Creating communities of learning in an era of testing and standardization*, Beacon Press, Boston.
- Mission Australia 2011, *National Survey of Young Australians 2011*, The Author, Sydney.
- Munro, J 2010, 'Enhancing students' literacy comprehension using NAPLAN data', *Centre for Strategic Education Seminar Series*, Paper No. 200.
- Office of the United Nations High Commissioner for Human Rights 1966, *International Covenant on Economic, Social and Cultural Rights*. Accessed 24 October 2011 at: <http://www2.ohchr.org/english/law/cescr.htm#art13>
- Office of the United Nations High Commissioner for Human Rights 2001, *Committee on the Rights of the Child – General Comment No. 1, The Aims of Education*. Accessed 24/10/2011 at: <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G01/412/53/PDF/G0141253.pdf?OpenElement>
- O'Keeffe, D 2011, 'NAPLAN nightmares', *Education Review*, 11th August 2011, viewed 14th September 2011, <http://www.educationreview.com.au/pages/section/article.php?s=Features&idArticle=21789>
- Paris, S 2000, 'Trojan Horse in the Schoolyard; the hidden threats in high stakes testing', *Issues in Education*, Vol. 6, No. 1/2, pp. 1-8.
- Paris, S & McEvoy, A 2000, 'Harmful and enduring effects of high stakes testing', *Issues in Education*, Vol. 6, No. 1/2, pp. 145-160.
- Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. 2003 *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
- Perrone, E 1991, 'On Standardised Testing', *Childhood Education*, Vol. 67, pp. 132-142.
- Peters S & Oliver L 2009, 'Achieving Quality and Equity through Inclusive Education in an Era of High-Stakes Testing', *Prospects: Quarterly Review of Comparative Education*, Vol. 39, No. 3, pp. 265-279.
- Phelps, RP 2006, 'Characteristics of an Effective Student Testing System', *Educational Horizons*, Vol. 85, No. 1, pp. 19-29.
- Queensland Studies Authority 2009, *Student assessment regimes: Getting the balance right for Australia (Draft discussion paper)*, viewed 27 September 2010, <http://www.qsa.qld.edu.au/downloads/publications/qa_paper_assess_balance_aust.pdf>.
- Ravitch, D 2010, *The Death and Life of the Great American School System. How Testing and Choice are Undermining Education*, Basic Books, New York.
- Reay, D & Wiliam, D 1999, 'I'll be a nothing': structure, agency and the construction of identity through assessment [1]', *British Educational Research Journal*, Vol. 25, No. 3, pp. 343-354.
- Renter, DS, Scott, C, Kober, N, Chudowsky, N, Chudowsky, V, Joftus, S, & Zabala, D 2006, *From the Capital to the Classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Center on Education Policy.
- Ringarp, J & Rothland, M 2010, 'Is the Grass always Greener? The Effect of the PISA Results on Education Debates in Sweden and Germany', *European Educational Research Journal*, Vol. 9, No. 3, pp. 422-430.
- Sabol, R 2010, *No child left behind: A study of its impact on art education*. Purdue University, Indiana National Art Education Foundation.
- Santiago, P, Donaldson, G, Herman, J and Shrewbridge, C 2011, *OECD Reviews of Evaluation and Assessment in Education: Australia*, OECD.
- Schroeder, L 2006, 'What high-stakes testing means for the emotional wellbeing of students and teachers', unpublished PhD thesis, Graduate University, Claremont.
- Senate References Committee on Education, Employment & Workplace Relations 2010, 'Inquiry into the Administration and reporting of NAPLAN testing', *Commonwealth of Australia*, Canberra, viewed 13 September 2011, http://www.aph.gov.au/senate/committee/eet_ctte/naplan/report/report.pdf
- Skeggs, B 1997, *Formations of Class and Gender: becoming respectable*, Sage, London.
- Skrtec, T 1995, *Disability and democracy: Reconstructing (special) education for postmodernity*, Teachers College Press, New York.
- Sloane, F and Kelly, A 2003, 'Issues in High Stakes Testing', *Theory into Practice*, Vol. 42, No. 1, pp. 12-17.
- Snow, D 2006, 'Teenage Wasteland', *Sydney Morning Herald*, 13 May.
- Stiggins, R 1999, 'Assessment, student confidence and school success', *Phi Delta Kappan*, Vol. 81, No. 3, pp. 191-198.
- Thomas, R 2005, *High Stakes Testing: Coping with Collateral Damage*, Mahwah, N. J.
- Triplett, C. F. and Barksdale, M. A. 2005. Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research* 37(2): 237-260.
- Van Damme, J, Hongqiang, L, Vanhee, L, & Pustjens, H 2010, 'Longitudinal studies at the country level as a new approach to educational effectiveness: explaining change in reading achievement (PIRLS) by change in age, socio-economic status and class size', *Effective Education*, Vol. 2, No. 1, pp. 53-84.
- Wheelock, A, Bebell, D & Haney, W 2002, 'What can drawings tell us about high stakes testing in Massachusetts?', cited in Triplett, CF & Barksdale, MA 2005, <<http://www.westeeep.bc.edu/drawoned/meas/measpaper2.html>>
- William, D 2010, 'Standardized Testing and School Accountability', *Educational Psychologist*, Vol. 30, No. 2, pp. 107-122.
- Williams, P 2009, 'Analytic Report: Factors contributing to and ways of improving Australia's educational performance', Submission to the *Inquiry into the Administration and Reporting of NAPLAN Testing*, Senate References Committee on Education, Employment & Workplace Relations, Canberra
- Wu, M 2010, 'Inadequacies of NAPLAN results for measuring school performance', Submission to the *Inquiry into the Administration and Reporting of NAPLAN Testing*, Senate References Committee on Education, Employment & Workplace Relations, Canberra
- Wyn, J 2009. Touching the future: Building skills for life and work. *Australian Education Review*, No. 55, Melbourne: Australian Council for Educational Research Press.

Whitlam Institute

WITHIN THE UNIVERSITY OF WESTERN SYDNEY

UWS Parramatta Campus, Cnr James Ruse Drive & Victoria Rd Rydalmere NSW Australia 2116
Locked Bag 1797 Penrith NSW Australia 2751
T + 61 2 9685 9187 F + 61 2 9685 9110 E info@whitlam.org W www.whitlam.org

