

Summary

This discussion paper has been developed by the Australian Primary Principals Association (APPA) to draw attention to problems arising from the publication of NAPLAN data on the *My School* website. APPA is concerned that the public availability of data on all schools, combined with the powerful and flexible search tools on *My School*, has shifted NAPLAN into a high stakes environment and had significant unintended negative consequences for schools.

High stakes assessment involves tests that have significant consequences for students, teachers and others in education, and also involves public reporting and public accountability. This paper presents national and international research evidence of the significant negative effects of such testing regimes, including:

- limits on the development of the range of skills and literacies needed in the modern world and the encouragement of low-level thinking;
- negative effects on teacher pedagogy, teacher collaboration and on creative and effective teachers;
- a shift away from supportive and collaborative learning approaches to competitive and individualistic ways of learning;
- stress and tension in students, especially primary students, leading to negative impacts on student well-being, self-esteem and motivation; and
- distortion and narrowing of the curriculum.

The paper then examines the implementation of NAPLAN and the public reporting of NAPLAN results in searchable form on the *My School* website. It concludes that there is clear evidence that NAPLAN has become a high stakes assessment regime, and that the negative effects evident in international research are also evident in Australia, despite the fact that both the high stakes assessment program and research on its effects are at an earlier stage of development than in comparable jurisdictions (e.g. the United Kingdom and United States).

APPA states its support for NAPLAN, recognises NAPLAN's role in school transparency and accountability measures, and values the clear, independent data provided by the assessment regime to schools and school systems. The paper does not see NAPLAN itself as the source of the problems identified in the paper; rather, we argue that the evidence suggests they result in large part from the publication of comprehensive, searchable results on *My School*. This is clear in the publication of league tables of schools in Australia and a public discourse that sees NAPLAN as the key source of information about a school.

APPA proposes that this deteriorating situation can be resolved in the following ways:

- removing school-level NAPLAN results and the related search capacity from *My School*;
- making clear, if it is not already clear in any jurisdiction, that schools should report their NAPLAN results, together with national and their own state or territory results, to students, parents and the community; and
- making clear in public discussion of NAPLAN results both the value and the limitations of the data.

Background

The *My School* website is provided by the Australian Curriculum Assessment and Reporting Authority (ACARA). It provides profiles of about 9500 Australian schools, including information about school context and finances, enrolments, attendance rates and staff numbers, and the performance of the school's students on the National Assessment Program – Literacy and Numeracy (NAPLAN).

The website uses an index of student and school characteristics, known as the Index of Community Socio-Educational Advantage (ICSEA), to identify schools serving students from statistically similar backgrounds. Information on the occupation and education of parents, along with school enrolment data, is used to calculate an ICSEA value for each school. This value represents the average educational advantage level of students at the school. The NAPLAN data from a school can then be compared with results from other Australian schools serving students from what are determined as statistically similar backgrounds according to their ICSEA score.

My School is intended to be used by parents, community members, teachers, researchers, those involved in educational administration and management: anyone who needs access to aggregated information about their own school or about one or more schools. It is intended to assist in providing greater transparency and accountability regarding the performance of schools.

Discussion

The principal issue raised by APPA in this paper concerns the consequences of the public reporting of NAPLAN data for all schools on the *My School* website. It is our contention that this has had serious negative effects on the education of Australian children, especially in primary schools, that these effects will increase, and that they can only be managed by removing NAPLAN data from the website.

The argument is easily misinterpreted. In order to demonstrate the effects we are concerned about, the paper discusses the negative consequences of high stakes testing, of which NAPLAN has become an example. We stress that the argument is not about NAPLAN, which we support, but about the perverse consequences produced by its publication in searchable form on *My School*.

APPA wishes to state its support for NAPLAN, and its recognition of the value of the clear, independent data on school performance provided by the assessment regime. This paper

does not question the value of NAPLAN. We regard NAPLAN as a useful tool for schools and school systems in tracking progress and identifying areas for improvement. We accept that national assessment is one of the tools to provide public information about and strengthen public confidence in schooling and improve accountability and public transparency. We believe it provides one component of the information parents need in making choices about schools for their children.

Despite our support for NAPLAN, we regard the program as pernicious if it is used to produce broad-brush and unwarranted rankings and comparisons of schools. We think it is counter-productive if it drives out and devalues other data and mechanisms for school evaluation and improvement. We think it is damaging if it distorts the curriculum and pushes schools and teachers to spend excessive time in test preparation. We think it can be destructive if it creates anxiety and fear in primary children. These and other difficulties with the present position of NAPLAN in Australian education are functions principally of the publication of NAPLAN data on *My School* and the resulting perception of NAPLAN as a high stakes assessment regime.

Effects of high stakes testing

By 'high stakes testing', we refer to tests that have significant consequences for students, teachers and others involved in education, and involve public reporting and public accountability (AERA, 2000; Polesel et al, 2012; Johnson et al, 2008; Marchant, 2004; William, 2010; Ross, 2013). These are commonly tests (and their results) that:

- are very visible in the community;
- are regarded as key markers of student progress (or lack of it);
- are used to rank, compare and evaluate schools or students;
- are used to establish an apparent hierarchy of school or student performance;
- are or can become a factor in shaping school practice; and
- are or can be used in performance management of school or system personnel.

We argue below that NAPLAN fits this definition.

High stakes testing is argued by many commentators to have significant benefits. It is regarded as improving academic standards, teacher accountability and student motivation, providing more independent measures of performance, providing diagnostic information to improve and target teaching, teachers and schools, being better aligned with curriculum than local assessment and offering reassurance to the community that education is effective (Polesel et al, 2012; Phelps, 2006; Amrein & Berliner, 2002; Nichols, 2007; Ross, 2013).

APPA does not dispute that there are benefits from formal assessment programs. We do dispute that these benefits arise specifically from the high stakes nature of such programs. The principal educational benefits from testing programs (diagnosis, independent information, alignment with curriculum, measurement of progress over time, even student motivation and teacher accountability) can be achieved without the elements that make assessments high stakes. We are concerned that the point at which high quality assessments

tip over into a high stakes environment is the point at which the benefits of formal assessment begin to be undermined and leached away by an accumulation of negative effects.

There is now excellent research¹ internationally and a growing body of work in Australia on the negative impacts of high stakes testing. A major literature review by Polesel et al (2012) on the impact of high stakes testing identified research indicating a wide range of impacts of high stakes testing on student well-being, teaching and curriculum. These include:

- negative impacts on student well-being, including on self-esteem and lower teachers' expectations of students, personal stress and tension in students, undermining of self-esteem and fear of failure (Reddell, 2010; Paris & McEvoy, 2000);
- limiting the development of the range of skills and literacies needed in the modern world and encouragement of low-level thinking (Au, 2008; Paris, 2000);
- negative effects on teacher pedagogy and a negative impact on creative and effective teachers as the focus shifts from students' needs to evaluation and reporting (Cunningham & Sanzo, 2002);
- a shift away from supportive and collaborative learning approaches to competitive and individualistic ways of learning (Reay & William, 1999); and
- narrowing of the curriculum, an increased focus on curriculum areas tested and a reduced proportion of class time devoted to those areas not subject to testing (Abrams, 2004; David, 2011; Madaus et al, 2009).

Ross (2013) examined the international research on the effects of high stakes testing, especially on student wellbeing, in a paper commissioned by APPA. He found evidence of negative impacts including:

- distortion and narrowing of the curriculum, resulting in an expansion of time spent on areas tested (usually mathematics and literacy) and a reduction in areas such as science, social studies, art, music, physical education;
- changes in pedagogy as teachers focus on 'item teaching' based on test items, reduce their use of open-ended approaches and increasingly adopt teacher-centred instruction; and
- reductions in student self-esteem and student motivation.

There is clearly extensive research supporting the view that high stakes testing has substantial negative impacts and that these impacts are likely to lead to a material reduction in the quality of the education offered by teachers and schools, and in the learning of students.

¹ This section draws on the work of Polesel et al in their 2012 literature review of the effects of high stakes testing.

Is NAPLAN a high stakes test?

The view that NAPLAN is a high stakes assessment regime, and therefore likely to lead to the negative outcomes discussed above, is not undisputed. The Chair of the ACARA Board, Professor Barry McGaw, claims 'NAPLAN...tests are not onerous and not high stakes', arguing that the tests occur 'over a few hours spread out over a few days, four times from year 3 to year 9' (McGaw, 2012). He suggests that census assessments of this kind are not new: New South Wales introduced basic skills tests in 1989, followed by other states and territories and eventually national testing.

Professor McGaw says that:

The only high-stakes assessments in Australia now are end-of-year 12 assessments, and the voluntary entry tests for selective schools and scholarship tests for non-government schools (ibid).

Professor McGaw states that the earlier state-based tests were 'well-established and non-controversial'. But he also acknowledges that recent opposition to NAPLAN 'comes because the results are now publicly reported'. And although he does not accept that NAPLAN is a high stakes assessment, he identifies that it is associated with some of the problems linked with high stakes assessment, and suggests that teachers are responsible:

If NAPLAN is being made high-stakes for students, with some reported to be anxious and even ill when the tests approach, this is due to teachers transferring stress to their students (ibid).

Following on from above, parents, believing that NAPLAN results play a significant role in their children's educational future, might also be seen to transfer stress to students. Principals report that, on one hand, there is an increasing number of students being withdrawn from the tests and, on the other, NAPLAN results are increasingly being used in secondary (and selective) school enrolment regimes and in deciding scholarship recipients. The decisions parents are making indicate the high stakes they perceive are involved in NAPLAN testing.

Professor McGaw also notes that the NAPLAN data published on *My School* 'have been irresponsibly used by some parts of the media to publish raw league tables that take no account of a school's context', while claiming that '*My School* does not do that'. He also acknowledges that some teachers and principals 'are now giving students excessive test practice, teaching to the test and narrowing their students' experiences in various ways' (ibid.). APPA takes the view that this is effectively acknowledging the high stakes nature of the program, at least as it is experienced in schools.

Professor McGaw notes that this did not happen when state based precursors were introduced. Why is it happening now? It is APPA's view (supported by Professor McGaw's comments) that the key new variable is publication of results on *My School*, along with easy (and possibly erroneous) school comparisons, the facilitation of the creation of league

tables, and a resulting public discourse that accepts that NAPLAN results are the measure of success or failure of a school (and, by implication, a student).

Professor McGaw has made clear that there is a case that NAPLAN has now become a high-stakes assessment because of public reporting. The issue for APPA is the effect of this change on children, teachers and schools, and the means by which these negative impacts can be ameliorated.

The view that NAPLAN is a high stakes assessment program is widely held. The literature review on the impact of high stakes testing cited above concluded that, while there were differences between NAPLAN and international assessment programs:

the publication of the results of the NAPLAN program on the *My School* website, with the associated media coverage, means that NAPLAN too may be labelled as a high stakes testing program (Polesel et al, 2012: 4).

This conclusion is borne out by the impact of NAPLAN in Australian schools, which matches the kinds of effects noted above for high stakes testing.

Connell points out that although there is no official publication of league tables of schools, the publication of NAPLAN data on *My School* has led to a situation in which:

The statistics are duly processed by the mass media into league tables, ranking schools in order of NAPLAN test performance (Connell, 2013: 282).

This is exemplified in the publication by the *Canberra Times* of tables ranking schools by NAPLAN results for each year level tested (Fowler, 2014) and a similar publication by the *Sydney Morning Herald* (Connell, *ibid.*).

The Whitlam Institute submission to the Senate inquiry into NAPLAN notes that NAPLAN has become the 'default measure for very significant policy purposes': it was the key measure for the National Education and Reform Agreement under the previous Labor Government and a key indicator in the Schooling Resource Standard (Dulfer et al, 2013: 6). Given these policy roles, it is clear that NAPLAN has become a high stakes assessment regime.

Lobascher, arguing that NAPLAN is a high stakes assessment, cites evidence that such testing detracts from the creativity of teachers and removes students' intrinsic motivation, reducing enjoyment of teaching and learning (Lobascher, 2011).

Wu, in a submission to the Senate inquiry into NAPLAN argues that NAPLAN should not be used to measure school performance because the margins of error for measuring student and school performance are too great (Wu, 2010a). She argues that the uncertainty associating with the estimate of a teacher's performance is about as large as the difference between a good and a poor teacher (Wu, 2010b). NAPLAN tests are simply not accurate enough to be used in making judgment about teacher effectiveness:

Since the range of teacher effect is about one year of growth, to separate effective from ineffective teachers our measures need to be accurate to fractions of a month's growth. Achievement tests just don't provide that kind of accuracy (Wu, 2014: 18).

Analysing the use of NAPLAN to make judgments about teacher performance, Wu also argues that:

Attaching high-stake consequences to student performance data can only damage the teaching profession and education more generally (Wu, 2014: 22).

Hardy, discussing the impact of NAPLAN in Queensland, describes 'significant political pressure upon the state government for improved outcomes' because of Queensland's low ranking, and a resulting change to 'testing students' literacy and numeracy capacities more frequently, and spending more time on test-readiness activities'. This and other initiatives have resulted in 'an intensity of focus upon students' results and behaviour in schools in ways which have not previously been the case' (Hardy, 2013: 8, 10). He also notes that some principals were:

placed on contracts, with the potential for bonuses for improved NAPLAN results, but also possible termination of these contracts...should results not improve sufficiently (ibid.: 10).

The Independent Education Union submission to the Senate inquiry on NAPLAN, reporting the results of independent research on attitudes of its members, notes:

The overwhelming conclusion remains that every aspect of teaching and learning has felt the impact of the publication of the NAPLAN data (IEU, 2013).

There is also widespread recognition that participants in NAPLAN, including students, teachers and parents, are experiencing the effects that are associated with high stakes assessments. The absence of detailed and comprehensive research in Australia means that some of this evidence is anecdotal, but there is a growing body of such evidence that primary school children experience anxiety because of NAPLAN (O'Keeffe, 2011; Mahar, 2013; Athanasou, 2010). Other evidence drawn from the Senate hearing on NAPLAN indicates impacts such as:

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 (Polesel et al, 2012: 10).

Similar evidence exists for these and other impacts arising from NAPLAN. The Whitlam Institute submission to the Senate hearing on NAPLAN reports on a large (8300 respondents) online survey that reported:

- narrowing of teaching strategies and of the curriculum
- negative impacts on student health and wellbeing

- negative impacts on staff morale, and
- negative impacts on school reputation and capacity to attract and retain students and staff (Dulfer et al, 2013: 5).

Over half of the respondents reported that NAPLAN impacts the style and content of their teaching, and two-thirds said it has led to a reduction in the time available for other subjects (ibid.: 5-6).

A key survey conducted by Canvass Strategic Opinion Research in 2013 for APPA showed similar findings. The survey of APPA members (primary principals) on attitudes to NAPLAN found that:

- the greatest impact was on student wellbeing: half of all respondents sometimes see students display signs of stress or sickness, express fear of failure or withdraw when faced with the tests. Respondents also reported teacher and parent stress;
- half believe that the burden is greater for Year 3 than for Year 5 students. The greatest negative impact was found in remote and small schools;
- half of the respondents say their schools spend more time on literacy and numeracy in the run-up to NAPLAN, and 60% say they spend less time teaching other subjects, though they regard the impact as relatively slight (Canvass, 2013: 7).

The survey also showed that principals believed parent interest in NAPLAN results was principally in their own child's results, but that about 25% were interested in school results and 16% in teacher performance. Yet school NAPLAN results 'are a drawcard' for parents according to one-third of respondents, and about half of respondents in independent, very large and metropolitan schools say parents cite NAPLAN results as a reason for enrolling their children (ibid.: 8).

We acknowledge that NAPLAN is not accompanied by some of the more dramatic consequences of high stakes regimes in other countries. NAPLAN is not, so far as we know, associated with students being held back or formally used in student promotion decisions; results have not led to school closures or to schools being officially singled out publicly as failures (although we note the increasing use of NAPLAN results for principal performance management). *My School* includes contextual data, and the comparison of like schools is more sophisticated, because of the ICSEA score, than in jurisdictions that support the publication of league tables with no recognition of the different circumstances of schools.

Despite this, many schools and system personnel and students are experiencing the impact of NAPLAN results in just the way that significant impacts are experienced with other high stakes assessments. NAPLAN is now a key element of the policy landscape in Australia. If the definition of a high stakes assessment regime is that it has significant consequences for students, teachers and schools, NAPLAN has clearly crossed this line.

What do we propose?

No assessment is without significance and consequences. Major assessment programs can, however, operate without the full set of negative consequences associated with high stakes assessment. NAPLAN did not always fully meet the criteria for high stakes, and nor did earlier state and territory assessment programs. What has changed is the introduction of *My School* and the publication of NAPLAN data in a form made possible by the website. Our argument is that, to a significant extent, NAPLAN has become a high stakes mechanism because of the publication of results on *My School*.

NAPLAN could meet most of its aims without publication of results on *My School*. Indeed, it is arguable that the level of reliability of NAPLAN results does not justify their use in a highly visible way in ranking and comparing schools or, as Wu suggests (see above), in making judgments about teacher performance. It does, however, provide the kind of information that schools and principals find useful in targeting areas for improvement, as a supplement to the broad range of other data and information generated within a school, including that resulting from school-based student assessment and teacher judgment. APPA is supportive of this approach and of the role of NAPLAN as a component of school reporting to parents and the community.

This can be achieved, though, without *My School* publication. The core purposes of NAPLAN are achieved if schools report NAPLAN results to students, parents and the community; if education systems use NAPLAN as a component in informing school resourcing, helping shape teaching and learning, and in providing feedback to schools and principals; and if parents have access to NAPLAN data from the schools they are considering for their children.

My School publication of results exposes schools to blunt, arguably inaccurate and damaging comparison and ranking. By raising the public profile and stimulating media commentary and league tables based on raw scores, it encourages uninformed and harmful debate.

APPA proposes that this deteriorating situation can be resolved in the following way:

- removing school-level NAPLAN results and the related search capacity from *My School*;
- making clear, if it is not already clear in any jurisdiction, that schools should report their NAPLAN results, together with national and their own state or territory results, to students, parents and the community; and
- making clear in public discussion of NAPLAN results both the value and the limitations of the data.

8 September 2014

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

American Education Research Association (AERA). (2000). *Position Statement on High-Stakes Testing*. Accessed on 17 August, 2014 at:

<http://www.aera.net/AboutAERA/AERARulesPolicies/AERAPolicyStatements/PositionStatementonHighStakesTesting/tabid/11083/Default.aspx>

Amrein, A.L., & Berliner, D.C. (2002). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in states with high school graduation exams*. Accessed on 17 August, 2014 at: <http://epsl.asu.edu/epru/documents/EPSTL-0211-126-EPRU.pdf>

Athanasou, J. (2010). 'NAPLAN and My School Survey', *Independent Education Union of Australia*, Barton, Australian Capital Territory.

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

Canvass Strategic Opinion Research (Canvass). (2013). *Primary Principals: Perspectives on NAPLAN Testing & Assessment*. Australian Primary Principals Association.

Connell, R. (2013). 'Why do market 'reforms' persistently increase inequality?' *Discourse: Studies in the Cultural Politics of Education*, Vol. 34, No. 2 pp. 279-285.

Cunningham, W & Sanzo, T. (2002). 'Is high-stakes testing harming lower socio-economic schools?', *National Association of Secondary School Principals NASSP Bulletin*, June, 2002 pp.62-75

David, J. (2011). 'What students need to learn: high-stakes testing narrows the curriculum', *Educational Leadership*, Vol. 68, No. 6, pp. 78-80

Dulfer, N., Polesel, J. & Rise, S. (2013). *Submission: Senate Inquiry into the Effectiveness of the National Assessment Program—Literacy and Numeracy (NAPLAN); For the Senate Education, Employment and Workplace Relations Committees*.

Fowler, Glenn. (2014). 'League Tables Do Not Help Our Kids', accessed on 17 August, 2014 at: http://www.aeuact.org.au/league_tables_do_not_help_our_kids

Hardy, I. (2013). 'Education as a 'risky business': Theorising student and teacher learning in complex times'. *British Journal of Sociology of Education*, Rapid Online, pp. 1-20

Independent Education Union (IEU). (2013). *A response from the Independent Education Union of Australia to the Senate Standing Committee on Education, Employment and Workplace Relations: Inquiry into the effectiveness of the National Assessment Program—Literacy and Numeracy (NAPLAN)*. Accessed on 16 August 2014 at:
http://www.ieu.asn.au/media/81640/naplan_senate_inquiry_ieu_submission_june_2013.pdf

Johnson, D, Johnson, B, Farenga, S & Ness, D. (2008) *Stop High Stakes Testing: an Appeal to America's Conscience*, Lanham, Rowman & Littlefield

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, 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, North Carolina.

Mahar, Tim. (2013). 'Day in the Life of NAPLAN', accessed on 17 August at:
<http://www.saveourschools.com.au/league-tables/a-day-in-the-life-of-naplan>

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.

McGaw, B. (2012). 'NAPLAN myths: it's not a high-stakes test'. Accessed on 16 August, 2014 at <http://theconversation.com/naplan-myths-its-not-a-high-stakes-test-11057>

Nichols, Sharon L. (2007). *High Stakes Testing: Does It Increase Student Achievement?*, Haworth Press Inc., San Antonio

O'Keeffe, D. (2011). 'NAPLAN nightmares', *Education Review* August 2011

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.

Phelps, Richard. (2006). *Characteristics of an Effective Student Testing System*, Education Horizons.

Polesel, J., Rice, S., & Dulfer, N. (2012). *The Experience of Education: The impacts of high stakes testing on school students and their families*. Whitlam Institute, University of Sydney

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

Reddell, S. (2010). *High stakes testing: Our children at risk*. Accessed on 16 August, 2014 at: <http://files.eric.ed.gov/fulltext/ED513338.pdf>

Ross, Bob. (2013) 'An Examination of High Stakes Testing and its Impact on Primary School Education and Student Wellbeing'. Australian Primary Principals Association

William, D. (2010). 'Standardized Testing and School Accountability', *Educational Psychologist*, Vol. 30, No. 2, pp. 107–122.

Wu, M. (2010a). '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

Wu, M. (2010b). 'The inappropriate use of NAPLAN data'. *Professional Voice*, Vol. 8, Issue 1, 21-25.

Wu, M. (2014). 'Linking student test scores to teacher performance'. *Professional Voice*, Vol. 10, Issue 1, 16-22.