PROFESSIONAL VOICE

Winter 2010

The NAPLAN debate

Volume 8
Issue 1

RRP $10

AUSTRALIAN EDUCATION UNION
AEU

EDITOR
Nic Barnard

EDITORIAL COMMITTEE
Ross Dean, John Graham and Nic Barnard

DESIGNER
Peter Lambropoulos

PUBLISHED BY
The Australian Education Union Victorian Branch

PRINTED BY
Total Print

PRINTED ON RECYCLED PAPER
Cover: 310g Re-Art Gloss
Text: 105g Re-Print

ANNUAL SUBSCRIPTIONS
$25 within Australia (free to financial AEU members). Please contact the AEU Publications Unit on **03 9417 2822** for further information. A subscription form is available in the back of this issue or downloadable from **www.aeuvic.asn.au/publications**.

BACK ISSUES
Free to members. $10 to non-members within Australia.

Advertising is accepted according to AEU guidelines. Please contact Lyn Baird for further details. Phone: 03 9418 4879. Email: **lyn.baird@aeuvic.asn.au**.

Editorial correspondence should be directed to Nic Barnard, Professional Voice, Australian Education Union Vic Branch, PO Box 363, Abbotsford VIC 3067. Phone: 03 9418 4841. Email: **nic.barnard@aeuvic.asn.au**.

The AEU Victorian Branch can be contacted at:
PO Box 363, Abbotsford 3067
Tel: 03 9417 2822 **Fax:** 1300 568 078
Web: www.aeuvic.asn.au

ISSN:1445-4165

THE NAPLAN DEBATE

contents

Editorial: The Trouble with My School

JOHN GRAHAM

THE GENESIS OF this edition of *Professional Voice* was a national symposium held in Sydney in July. Its title was "Advice for Ministers and ACARA on NAPLAN, the use of student data, My School and league tables". It was organised by the AEU and the two government school principal associations. The idea of the symposium was to subject claims by politicians about the virtues and benefits of My School to the scrutiny of highly regarded experts in the field. The following six articles encompass the addresses given at the symposium and explore the issues further in a more formal written form.

The introduction of My School was a political high point and an educational low point for the Rudd Labor Government. The site was rushed out at 1am on January 28 to coincide with the beginning of the school year and so gain maximum publicity in the media. The Gillard–Rudd team had decided that the media were on their side on this one and the launch would provide an ideal springboard for an election year.

The half-baked nature of the site was a sure sign that the deadlines for its launch were political in origin. Not only were there myriad errors — missing schools, wrong school locations, ludicrous comparisons of one school with another (large metropolitan private schools with tiny rural primary schools) — but the basis of

school comparisons, the Index of Community Socio-Educational Advantage (ICSEA), was flawed and the promised income levels of each institution did not appear.

The crudity of the website was further evidence of its political purpose. Each school's page was dominated by NAPLAN test numbers. They were presented in a large font with coloured highlighting to focus attention on this section alone. The comparisons between "statistically similar" schools were also based solely on NAPLAN results. The message of the website was that a school's quality can be accurately equated with the performance of its students on literacy and numeracy tests held over three days in May.

A chorus of media approval and a predictable rash of school league tables followed the My School launch. Commercial organisations, including a number of media outlets, used web-scraping tools to extract and restructure the data from the site to praise and shame the "best" and the "worst" schools, all in the interests of "transparency" and the free market. In some ways the most damaging league tables were those compiled by local media comparing all of the schools in their area. The *South Gippsland Times* concluded that a local private school had "blitzed all of its peers" while the "writing was on the wall" for some of the area's public primary schools. The media league tables reduced such niceties as the ICSEA score to a mentioned-in-passing irrelevancy — a sort of politically correct fig leaf. They compared NAPLAN raw score against NAPLAN raw score, sometimes conflating the scores in the different areas of testing into a single average. The unsophisticated consumer of educational statistics (most people?) was presented with an apparently plausible evaluation of the quality of the schools in their area derived from data on an official Federal Government website.

Unlike the media and the Government, teachers and schools overwhelmingly condemned the concept, content and aftermath of My School. Their arguments were educational and professional, rather than political, and were initially shrugged off by Julia Gillard, the then federal Education Minister. The AEU decided that its members would implement a moratorium on any involvement with the NAPLAN testing program — the basis of My School comparisons. After a barrage of rhetoric, the Federal Government moved into damage control and agreed to the setting up of a working party to review the website, including "the use of student performance data and other indicators of school effectiveness". At the same time, the Greens were successful in getting the Senate to establish a parliamentary inquiry into the quality and value of NAPLAN assessment and reporting. The arrival of a federal election suspended, at least temporarily, both of these processes.

The articles in this journal take a few steps back from the flurry of claim and counter-claim which followed the launch of My School, to examine the issues involved and consider the broader implications for educational practice in Australian schools. Alan Reid begins at the beginning, by identifying the three broad purposes of educational practice, which he categorises as "democratic", "individual" and "economic". Under the Howard Government, the individual purpose dominated; under Rudd–Gillard, the focus shifted to a more economic (human capital) purpose. There

is a good fit between both of these purposes and the philosophy and operation of My School. Reid believes that there is not the same compatibility between the democratic public purposes of education and the consumer market model built into the website. The public purposes of education are undermined when schools are treated as consumer items in an education supermarket rather than as community assets linked to the idea of the public good.

The rationale used to introduce My School was its capacity to increase the "transparency" of the country's schooling system. The publication of school results from standardised testing in literacy and numeracy, particularly when statistically similar "like schools" can be compared, would allow consumers to judge the educational quality of each institution. Margaret Wu and Barbara Preston separately argue that the very opposite is the case. Wu examines what is being compared — each school's literacy and numeracy NAPLAN scores — and concludes that the large margins of error in the measurement of any student NAPLAN performance, and the difficulty in separating out student level factors from school level factors in student performance, mean that variations in student performance cannot be reliably attributed to school performance. In these circumstances, the My School comparisons can only be seen as misleading and inaccurate. Given the high stakes involved for the students and teachers of a school judged (via My School) as poorly performing, Margaret Wu's critique alone should lead to the complete restructure or closure of the website.

Barbara Preston examines a second foundation of My School's alleged "transparency": the validity of the instrument (ICSEA) used to determine that one school is statistically similar to another. After My School went online there was an outcry from schools around the country about their ICSEA "like school" comparisons. The index produced a series of groupings which strained all credibility, such as a small public primary school in rural Victoria grouped with Geelong Grammar. ACARA passed these mis-comparisons off as anomalous outriders and contended that ICSEA remained a valid means of fairly comparing school performance. Preston disagrees after finding that the methodology used to construct the ICSEA scores creates a systematic bias against government schools.

According to Preston, the use of census data collected at the district level to establish the index has inflated the ICSEA scores of public schools and deflated those of Catholic and independent schools. The outcome is that public schools are unfairly compared to private schools which enrol higher SES students than their ICSEA scores would indicate.

Julia Gillard was quite open about the unoriginality of the My School concept. She tied its introduction in Australia to the perceived success of a similar system of school comparisons used in New York under the stewardship of Joel Klein. Her enthusiasm was such that she paid for Klein to come to Australia to spruik its success. Since his visit, the wheels have come off the "remarkable outcomes" Ms Gillard ascribed to him. The large gap between the unchanging mediocre performance of New York students on the highly regarded national test (National Assessment of Educational Progress) and the miracle gains made on the state tests led to an independent review of the

state testing program. It concluded that the state's tests had artificially inflated student scores, and new standards were introduced for the 2010 testing program.

As a result, student performance scores plunged. Only 42% passed the English tests in 2010 (compared to 69% in 2009) and the maths score fell from 82% to just 54%. The supposed gains made in reducing the achievement gap between black and white students over the four years of Klein's role as schools chancellor disappeared overnight.

Links between the growing impact of NAPLAN and My School in Australia and the experience of similar approaches to accountability in England and the United States are further explored in Brian Caldwell's article. The dysfunctional effects of high-stakes testing on schools, teaching and education as a whole have resulted in few if any achievement gains for students in these countries. Caldwell quotes the renowned educationist Ken Robinson on the way in which standardised testing is "strangling … what really matters in education".

In our featured interview, Andy Hargreaves points out that England is in fact backing away from standardised testing at the very time that the Australian Government is embracing it. He contrasts Australia's wish to have a "world-class" education system with the adoption of "the simplistic approaches of second-class competitors".

Behind the Federal Government's push for more transparency, high-stakes testing and a "back to basics" national curriculum is the perception that the basics of literacy and numeracy are being neglected in Australian schools. Allan Luke's article reports on several classroom studies which help to dispel this myth. His focus is on the enacted curriculum in the classroom rather than the paper policy set out in official curriculum documents. He has found that, contrary to prevailing political rhetoric, teachers are spending a large and increasing amount of time on "direct instruction in skills likely to be tested and (on) test-taking procedures and formats". This was even more pronounced in low SES schools.

Luke argues that such approaches are indicative of a national "over-correction" to a perceived loss of the basics and will have only negative effects on addressing the equity gap in Australian schooling. Luke's concern is echoed by Hargreaves, who has seen a similar pattern in all countries which have chosen to take this path. The basic testing curriculum simply crowds out innovation, creativity and the "thinking curriculum", particularly in public schools. Hargreaves poses the question: "Are our private schools going to retain these opportunities for more nourishing educational fare while those left in the public system are condemned to a staple diet of bare basics?"

The dilemma now facing teachers and schools in Australia is the same one which has plagued their equivalents in the United States for more than a decade. Steve Seidel (2010) from Harvard University recently spoke about the misalignment between "what counts and what matters" and how for many American teachers what truly matters "doesn't seem to count at all". The higher the stakes, the more teachers feel obliged or are required (despite their professional knowledge and judgement) to concentrate on the narrow tranche of the curriculum linked to standardised testing. Predictably

this approach yields short-term test score gains. This is the measured improvement politicians prize — a concrete achievement which the media can understand at election time — which works to further entrench the high-stakes testing environment. In the 2010 federal election, voters were asked to choose between two new policies: NAPLAN-linked teacher and school performance pay (Labor) and a doubling of NAPLAN testing (Liberal). The overseas evidence that such short-term gains prove largely illusory, that real literacy and numeracy standards do not rise, that equity gaps are not narrowed and that the overall quality of education is diminished — in short, the downside of high-stakes testing — has so far been treated by politicians like an unwelcome guest at a wedding.

The articles in this *Professional Voice* set out a compelling case for a new direction in education policy. One that is based on the best available educational research rather than the opportunism of election cycle politics.

REFERENCE

Seidel S (2010). Teaching Beyond the Tests, *Harvard Gazette,* May 13.

Accountability and the Public Purposes of Education

ALAN REID

IN THIS PAPER, I want to frame My School as an accountability strategy in the context of the purposes of education. I do this because the debate about My School has largely been conducted in the absence of an articulated set of reference points — that is, claims and counter-claims about the importance or the consequences of My School are often made in a vacuum. They are self-evidently good or bad. I will argue that the reference point against which such judgments should be made is the public purposes of education[1].

1. I will indirectly draw from a recently completed Australian Research Council research project report on the public purposes of education which was conducted in partnership with the Australian Government Primary Principals Association (AGPPA) and the Foundation for Young Australians: Reid A, Cranston N, Keating J and Mulford B (2010) Exploring the Public Purposes of Education in Australian Primary Schools (AGPPA). The project involved detailed case studies of a number of primary schools; a nationwide survey of primary principals; and interviews with policy makers.

PART A: AN ARGUMENT FOR THE PUBLIC PURPOSES OF EDUCATION

Educational practice is informed by its purposes; and such purposes are the outcome of political processes, resulting in broad "settlements" which shape educational discourses at particular historical moments. The literature review in our project revealed that there are three broad purposes of education — democratic, individual and economic. These purposes interact and become assertive under different conditions. Given the high status of purpose within education there will always be a dominant purpose.

- The *democratic* purpose is located in a society that expects its schools to prepare all young people to be active and competent participants in democratic life. Since this benefits the society as a whole, it is a public purpose.
- The *individual* purpose aims to advantage the individual in social and economic life. It treats education as a commodity, and supports school choice within an educational consumption approach. It posits education as a private good for private benefits and is therefore a private purpose.
- The *economic* purpose aims to prepare young people as competent economic contributors. Since this combines public economic benefits with private economic benefits, it is a constrained public purpose.

Of course educational purposes are not simply represented in official statements of missions and goals. They are shaped and delivered — both intentionally and unintentionally — through policy and practice in many different ways in schools and education systems which can be grouped into *three modalities of schooling*. These include:

- The *structure of schooling*, such as the ways in which formal schooling is organised and funded, which contain hidden messages about how the society is/should be structured, ordered and maintained
- The *official curriculum*, such as organisation of knowledge, including which knowledge is selected and omitted; assessment and reporting practices; and pedagogy
- The *culture and processes* of education systems and schools, such as social relationships, the nature of decision-making processes, the school ethos and so on — all of which give out messages about what is valued.

A healthy education system is one where there is strong compatibility within and between the modalities of schooling and the stated purposes of education. One way of thinking about the relationship between the modalities and the purposes of schooling is to use the metaphor of the body — where the institution of the school (or the education system itself) is the body; the purposes of schooling are the heart; and the modalities are the arteries. If the arteries get "clogged" as a result of being incompatible with the heart, then the body will suffer. In short, to achieve a healthy education system means ensuring that there is a strong compatibility within and between the modalities of schooling and the stated purposes of education.

Under the Howard Government, the major educational purpose was an individual one, dominated by the ideology of choice and resulting in increased competition

between schools. Education was treated as a commodity. The Rudd/Gillard Government has mainly focused on the economic purposes of education, with the major priority being the preparation of human capital for the labour market. In my view, this dilution of the public purposes of education has had negative impacts on the idea and practise of education as a common good. In a globalised and more complex world, where the nature and role of the nation state are changing, schooling based upon public purpose becomes more important. There is a need to return to a renewed emphasis on democratic *public purposes* for Australian education.

How can the public (democratic) purposes of education be advanced? Since active democracy requires capabilities for its nourishment, the central work of schools in a democratic society is the development of the capacities for social practice. These include capacities for citizenship, work, intercultural understanding, community involvement, communication and so on — capacities that build the common good.

The development of these capacities as a major purpose of schools has implications for the three modalities of schooling referred to earlier. Thus, from the perspective of public (democratic) purposes of education, the structure of schooling would seek to ensure equality of educational opportunity and resources that provide for the needs of all students (it would not tolerate unequal resourcing of schools); the curriculum would ensure that all students are encouraged and enabled to develop the capacities necessary for a democratic society (rather than a narrow and/or stratified curriculum); and the culture and processes of schooling would be based upon and model democratic processes (rather than authoritarian, top-down approaches).

However, it is one thing to argue in the abstract for a renewed focus on the public (democratic) purposes of education — quite another to develop educational policy and practice in ways that are consistent with it. Our research demonstrated that there is a formal commitment in the education community to the public purposes of education — from policy makers, principals and teachers — and that this commitment is represented in the Melbourne Goals of Schooling[2]. However, the research also showed that there is a lack of alignment between the stated goals and intentions of education policy and the strategies that are designed to deliver these, and this makes it difficult for schools to deliver on public purposes. That is, the arteries are clogged. I will use the strategy of My School to ask whether or not it is compatible with the public purposes of education espoused by the Melbourne Goals of Schooling.

PART B: WHAT IS MY SCHOOL AND WHY HAS IT BEEN DEVELOPED?

An education based on public purposes would model democratic practices at the level of systems (policy making) and individual institutions. It would also take seriously the process of review and accountability, seeking to enhance quality on an ongoing basis but ensuring that such processes are not inconsistent with or counterproductive to a public good philosophy. In my view, My School fails this test.

2 The Melbourne Goals of Schooling was the joint declaration made by all of Australia's state, territory and federal education ministers in December 2008.

When then Minister Gillard launched the My School website with great fanfare at the end of January 2010, she argued that it would satisfy the need for "transparent accountability". The dominant information that appears on the current version of My School about each school is its annual NAPLAN (national assessment program — literacy and numeracy) results. Now that My School has been introduced, NAPLAN has moved from being one piece of information which informed schools and education systems about one aspect of the outcomes of schooling, to being a high-stakes test purporting to measure the quality of a whole school and to compare it with other schools. We are told that more information about schools may be added in the future. But since we don't know what that might look like, we can only assess the version of My School that currently exists.

So, why My School? Different reasons have been given for the website — but common to these reasons is the idea of placing information about schools in the public arena. This information is invariably justified in terms of "transparency" and "accountability" in relation to three aspects: information for parental choice; information for quality improvement; and information to inform policy action.

PART C: IS MY SCHOOL CONSISTENT WITH THE PUBLIC PURPOSES OF EDUCATION?

Let's take each of these aspects and assess them against their consistency with the public purposes of education enunciated in the Melbourne Goals of Schooling.

Information and choice: the education market

> *If some (parents) vote with their feet, that's exactly what the system is designed to do.*
>
> — Kevin Rudd, 2009

My School assumes that people should shop around for schools as though they are consumer items like plasma TVs, thus creating the impression that education is a commodity rather than a public good. This diminishes the sense of school community — that feeling of all working together to make this a great school (Our School rather than My School) — because it assumes that if you are unhappy you simply choose somewhere else to send your child.

Further, it promotes the belief that everyone is in a position to choose. In fact, choice is limited to those who can afford the school fees or the relocation or transport costs if the chosen school is a long way away. It ignores the fact that in many states, government schools are zoned. You can't just decide to move to another school if it is not in your zone. It also assumes that the "top" schools will be able to take all comers. They won't be able to, of course, and so people who don't get their choice, or who can't choose, will feel they are being short-changed. In short, choice does not exist for a large percentage of the population.

I believe that the My School approach to accountability actually works against

the public purposes of education, by promoting schools as consumer items that exist to meet an individual's private needs rather than as community assets that build the wellbeing of the community in which they exist and the society as a whole.

Information for quality improvement through competition

> *Of course, we expect this information to provoke frank and robust discussions between parents, their child's teacher and school principals.*
>
> — Julia Gillard, 2010

The idea here is that the publication of school results will bring pressure on schools (principals and teachers) to improve their results and thus the quality of education. This will be achieved by comparing schools and getting the competitive juices going. There has been an attempt to overcome the objection that it is fatuous to compare schools with very different student and resource bases by creating lists of "like schools". It is hoped that there will be a sort of "kick the dog" effect as schools seek to do better than other "like" schools: parents will pressure teachers; teachers will pressure students; principals will pressure teachers; and education departments will pressure principals and so on.

On the surface, this may appear to be about providing the very best schools for all and thus be consistent with the public purposes of education. In fact, the sort of competition that is created must, through its own logic, create winners and losers. There will always be schools at the top of the table and schools at the bottom. If this competition meant that overall there was an increase in "quality", such an approach might even be justified — but unfortunately the evidence suggests that the opposite occurs: that is, it is a form of accountability which has a deleterious effect on the quality of education for a number of reasons.

First, there are major problems with the key technologies that are used to create the My School website: the Index of Community Socio-Educational Advantage (ICSEA) which purports to be a measure that enables meaningful comparisons to be made across schools; and NAPLAN, the testing program which provides almost the sole piece of information about the quality of a school. These concerns are explored in more depth in other papers (see Margaret Wu and Barbara Preston), so I won't deal with the detail here — but suffice to say that if these technologies are flawed, it is hard to say that this approach to accountability is likely to improve quality.

Second, the approach ignores findings from studies conducted in the UK and US, which demonstrate that the use of standardised tests in this way has deleterious effects on teaching, including that it narrows the curriculum and causes teachers to teach for the test; damages the morale and self-image of schools at the foot of the league tables that are inevitably created; and does not over time improve quality. In addition, high-stake approaches like this (league tables, monetary rewards for schools and teachers) cause schools and individual teachers to throw up smokescreens in order to hide rather than reveal issues and problems that confront them in their teaching;

and try to ensure "success" by manipulating the process, such as excluding certain children from tests and focusing on students at the cut-off points, not those at the top or bottom ends.

In short, this approach to improving the quality of education actually has the opposite effect. Far from promoting transparency by encouraging openness, collaboration and rigour across an education system, the high-stakes test approach closes this down, by fostering a competitive jockeying for position between individual schools. Such an approach is the antithesis of the kind of public purposes that are enunciated in the Melbourne Goals of Schooling.

Information to inform policy action

> *So through these new transparency measures we are going to have more information and we will be able to use that to bring these new resources and new reforms so they can make the most difference.*
>
> — Gillard, 2010

The other way in which information on the My School website is claimed to be used is to inform education policy (eg, where to focus resources) at a systems level and inform practice at the level of classrooms and schools. Usually, this aspiration is linked to the concept of equity with the claim that once we have identified "failing" schools or teachers, we will be able to do something about them.

This is a worthy goal. However, apart from the fact that most of this information was already available for these purposes before My School — we know well enough already which groups of students need additional support and resources — the strategy lacks the sophistication that is required to really address the causes of low student achievement.

For a start, the information provided is far too limited a base upon which to make policy or diagnose problems. It doesn't tell us anything about the quality of a school in relation to important aspects of schooling such as social and cultural outcomes — the key public purposes of education. If you think that schools should provide a broad and general education, including the arts, health and physical education, citizenship education and so on, then this narrow focus is a barrier to the enactment of the Melbourne Goals of Schooling. Why is it not just as important to assess and report on the ways in which schools are promoting, say, intercultural understandings or creativity, as it is to report on literacy and numeracy outcomes?

NAPLAN was not designed for the purpose to which it is being put in My School. Understanding NAPLAN results as a representation of the quality of a school is a corruption of its original purpose. Sure, it is possible to argue that parents can gather other data to complete the picture, but this isn't easy. Many people don't have the time or the knowledge to piece together the full range of information about a school, so they will tend to rely on a single source such as My School. After all, it is officially sanctioned information.

More than this, the approach suggests that quality improvement in schools is just a matter of identifying some low NAPLAN scores and throwing money at the problem. Unfortunately it doesn't happen like that. Learning outcomes are influenced by a range of social and cultural as well as educational factors, many of which are deep seated. These have to be identified and worked on over time. There is no quick fix. The fact is that questions about equity and education are incredibly complex. Over the years, many educators have worked long and hard to understand the causes of educational disadvantage. Policies and strategies have been developed and implemented in an effort to make some headway through funding and resourcing, curriculum, pedagogy and assessment, school culture, structures and organisation. Despite failures, new insights have been gained as collectively the profession inches forward in its understandings about the relationship between equity and education, seeking to improve policy and practice. And yet, somehow My School is the magic elixir for what has been evading us for so long!

Sadly, My School might in fact *work against* equity in education. The research in many parts of the world warns about the adverse effects on equity of high-stakes approaches to accountability. For example, Linda McNeil and colleagues summarise their findings in Texas from an extensive study which show that the state's high-stakes accountability system has a direct impact on the severity of the dropout problem, putting:

> *our most vulnerable youth, the poor, the English language learners, and African American and Latino children, at risk of being pushed out of their schools so the school ratings can show "measurable improvement". High-stakes, test-based accountability leads not to equitable educational possibilities for youth, but to avoidable losses of these students from our schools.*
>
> — McNeil et al, 2008
> — see article at **http://epaa.asu.edu/ojs/article/view/28**

Once again, these are not approaches which are consistent with the public purposes of education.

PART D: WHAT WOULD ACCOUNTABILITY LOOK LIKE WHEN BASED ON THE PUBLIC PURPOSES OF EDUCATION?

It is important to carefully analyse the claims that are made for My School as an accountability mechanism in Australian education since, as far as I can tell, no sustained case for the My School website has been made, using research evidence. Rather it is based on a set of apparently common-sense assertions about what information parents want and how to improve quality in education. I have argued that these claims do not withstand serious analysis and that indeed, as it is currently constituted, the My School website may be actually doing damage.

But it is important to go beyond critique and to suggest alternatives. In my view, if

it is genuine about the Melbourne Goals of Schooling, the Federal Government would seek to promote an accountability approach which is consistent with *public purposes* of education. Such an approach would be based on the following sorts of principles. It would:

- *Be developed in consultation with the profession and the wider community*
- *Have a sturdy conceptual basis, grounded in evidence from research about the forms and effects of approaches to accountability in education*
- *Assume that a picture about the quality of a school would encompass a range of data relevant to the capabilities for citizenship demanded by a democratic society*
- *Promote rigorous and open investigation and inquiry into the cause of problems, issues, dilemmas and concerns in schools*
- *Recognise the complexity related to learning outcomes and such factors as socio-economic status and cultural background, and promote approaches which seek to explore causes and trial approaches over time*
- *Recognise that quality learning outcomes are a result of collaborative endeavours within schools and across schools*
- *Promote schools as community centres (public goods) where people have an investment in working together to tackle problems and improve quality.*

A system of accountability based on these principles would certainly support the idea that information about schools should be in the public domain. The question is what sort of information and why. My purpose has been to argue that in its current form, the My School website lacks a sturdy conceptual basis and is at odds with the supposedly flagship document, the Melbourne Goals of Schooling. It is the antithesis of each of the accountability principles I have described. In my view it should be closed down now while further development work is done. I don't think it can be patched up as we go, with the fitful addition of disparate pieces of information as lobby groups argue their case. After all, if a large number of doctors were aware that a drug or some aspect of medical practice might harm patients, would we object to them banning the drug or practice until we are sure there are no toxic effects? If not, why is this different?

The challenge is to shape a commitment to transparency and accountability which truly serves the democratic public purposes of education.

The Inappropriate Use of NAPLAN Data

MARGARET WU

THE GOVERNMENT REPEATEDLY claims that parents can use the My School website to see "how schools are going". At the same time, principals and teachers continue to provide examples of how the website does not reflect school performance. Just who is right? How confidently can we use My School to judge schools? It is important to find answers to these questions, given that the website leads to stress and hardship for many individuals in the teaching profession, as well as its wider impacts on education.

This paper discusses the accuracies of the NAPLAN data which form the core of the My School website, with respect to the kinds of conclusions that can be drawn. It can be easily demonstrated that NAPLAN data do not provide the power to reflect school performance. The red and green bars on the website do not in any way reflect school performance. Consequently, the current My School website should be deemed to provide false information for as long as the public is encouraged to use the information to judge schools.

The Government's key justification for the My School website is that there is a need to be transparent about school performance. The assumption is that student performance, as measured by the NAPLAN tests, can be used to infer school

performance, after student background factors are accounted for. Currently, a socio-economic index (ICSEA) is used to adjust for student background factors. Consequently, when schools with similar ICSEA scores are compared, differences between schools are attributed to school effect. Further, for future NAPLAN results, there is a plan to use student growth measures as a proxy for school effect measures, since growth measures for individual students already control for the variations in student scores, so student background factors will not need to be controlled for.

This paper demonstrates that neither of the two methods (ie, using ICSEA or using growth measures to control for student background factors) provides sufficiently accurate measures of school effect.

There are two main problems with using NAPLAN data to infer school performance. The first is that student results can vary from test to test. This possible variation in test scores leads to a margin of error in the measurement of student abilities. When the test is short, the margin of error is large. That is, in drawing conclusions about student performance, we need to take account of the margin of error, because only a small sample of student performance is captured. The second problem is that, even if student performance is measured very accurately using very long tests, it is not easy to separate out student level factors from school level factors in explaining student performance. Using an index of student socio-economic status, one can explain away some student level factors, but not all. It is extremely difficult to conclude with any degree of confidence that the variation in student results within the so-called "like-schools" groups is directly attributable to school performance.

MARGINS OF ERROR IN MEASURING STUDENT PERFORMANCE

To get a handle on the magnitude of the margin of error in measuring student performance, consider the following example. If David was a Grade 5 student and he obtained 25 out of 40 on the 2008 reading test, we would expect David's test scores to vary by ±5 score points (ie, between 20 and 30, out of a maximum of 40) should similar tests (eg, the 2009 and 2010 tests) be given (Wu, 2010). This variability in a student's test scores when similar tests are given is called Measurement Error. Measurement error does not refer to mistakes made in processing the test results. Rather, measurement error refers to the possible variation of a student's test scores if similar tests are given. The variation in scores is due to the fact that each test only samples a small set of a student's capabilities in a subject domain. There is generally no effective way of reducing measurement error other than to increase the test length.

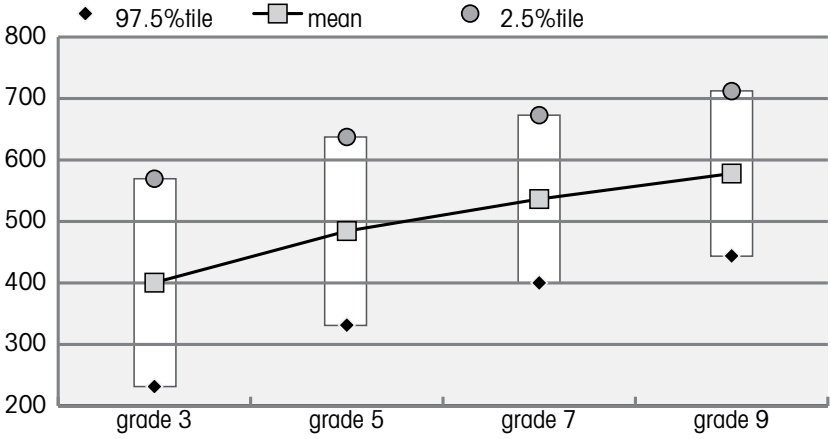
Taking measurement error into account, on the NAPLAN scale, David's reading score is 488 ± 54. That is, should similar reading tests be administered, we would expect David's NAPLAN reading score to vary by ± 54 points (a range of 108 points).[1] This is a very wide range. It reflects the uncertainty with which David's reading ability is measured by the NAPLAN test. To put the magnitude of this range in perspective,

1 The size of the measurement error is taken from the NAPLAN 2008 technical report. The range of ±54 points is the 95% confidence interval for David's "true reading score", where "true reading score" is defined as the test score if an infinitely large number of test questions could be administered.

Figure 1 shows the 2008 average reading scores of Years 3, 5, 7 and 9, and the spread of student scores.

Figure 1 shows student score distributions for Grades 3, 5, 7 and 9 for the NAPLAN 2008 reading assessment. The white rectangular box for each grade shows the range *mean±2×standard deviation* (ie, 95% of the student scores are expected within this range).

FIGURE 1 NAPLAN READING SCORES DISTRIBUTIONS IN 2008



We see that the possible range of David's scores (434 to 542) is so wide that it covers about half of the range of Grade 5 student scores, and that the range (434 to 542) covers Year 4 to the Year 7 average reading scores.

When David takes the Year 7 reading test in 2010, a similarly inaccurate estimate of David's reading score is obtained. A growth measure for David over the two-year period based on the two NAPLAN tests he has taken has a margin of error of ±76 NAPLAN score points. Compare this magnitude of error with the actual expected growth of around 52 NAPLAN score points, and it is clear that a growth measure based on two time points does not provide reliable information on how much a child has really grown. In fact, a small proportion of students will have negative growth measures even when they have made the expected growth in two years.

MARGINS OF ERROR IN MEASURING SCHOOL/TEACHER PERFORMANCE

Research studies on measuring teacher effect have found that a high-performing teacher can raise student standards by one more year of growth as compared to a low-performing teacher (Nye, Konstantopoulos & Hedges, 2004). That is, if you try your best in teaching Grade 3 students, you may raise your students to Grade 5 level in one year, at the most. This is reasonable, as one would not expect that a

teacher can make Grade 3 students grow to Grade 6 and beyond in one year. This means that if teacher effect is measured on the NAPLAN scale, students taught by the highest performing teacher will be around 40 score points[2] higher than if the same students were taught by the lowest performing teacher, after one year of learning. So if individual teacher effects are to be measured, we need precision to within a few NAPLAN score points, given that the range of teacher effects across all teachers is about 40 points. However, research studies have also shown that, based on student growth measures on two testing occasions, the estimates of individual teacher effects have a margin of error of around ±14 points[3] (Leigh, 2010; Leigh, personal communication, 2009). That is, the uncertainty associated with the estimate of a teacher's performance is about as large as the difference between a good and a poor teacher. Consequently, student growth measures obtained from two NAPLAN tests are extremely unreliable measures of individual teacher performance.

It should be noted that the report, *Reporting and Comparing School Performances*, prepared by the Australian Council for Educational Research (Masters et al, 2008) recommended the use of student growth as a better measure of student learning than the use of a socio-economic index for adjusting for student background factors. The argument is that growth measures focus only on the growth and not on the actual levels of achievement; thereby such measures completely control for student background factors that have an impact on student achievement. However, the ACER report failed to point out that many years of growth data are needed to provide the accuracies required to identify individual schools/teachers who are more effective than others. Given that growth measures are considered better measures than using socio-economic index adjustments, and given that growth measures themselves do not provide adequate measures of teacher effectiveness, one can easily conclude that the use of the ICSEA index to construct school performance measures is bound to fail.

THE VALIDITY OF MAKING STATISTICAL INFERENCE

Even if we put aside the issue of margins of error, and we assume that student growth measures are accurately obtained, we still need to make one large leap of faith to make a connection between the amount of growth and quality of teaching. What we have measured is student achievement. But we want to use that measure as a proxy for teacher effectiveness. It is *always* a conjecture to provide substantive interpretations to observed statistical results, for statistics never provide interpretation; statistics only provide numbers. Whenever statistical results are interpreted, statisticians are always cautious in using phrases such as "There is some suggestion that…"; "It is likely…"; or "The results indicate further investigation should be carried out…". Statisticians

2 Growth rates vary by grade level and subject area. Generally, the growth rate reduces as grade level increases. In Figure 1, the growth rate between grades 3 and 5 is around 84 points, so one year's growth is about 40 points. The growth rate is smaller between higher grades. The figures I have quoted are ball-park figures to demonstrate relative magnitudes.

3 Professor Andrew Leigh found that the standard error for individual teacher effectiveness measures is between 0.06 to 0.12 effect size. This translates to about ±10 to ±17 points on the NAPLAN scale for Grade 3 reading.

will never say "Look at this school's test results. It is clear that the school is not performing." This is because statistical results can never be used as proofs. They can only be used as one piece of possible evidence to support a hypothesis.

Sometimes one can make a bold claim based on statistical results, provided the consequences of making a false claim are not serious. A political party may use statistical figures to claim that another political party is at fault. Everyone knows that it's just politics, and we do not overly worry about the validity of such claims. However, when it comes to the My School website, the repeated claim that "the website reveals how schools are going" has serious consequences. Individual teachers and principals are stressed. Parents make judgments of schools and teachers based on misleading information. There are also ramifications for education more broadly. Making unjustified claims can result in law suits for defamation, because individuals are now victims of these false claims. Perhaps many political leaders and others do not realise that they are making false claims. If that is the case, I hope this paper provides the information needed to convince all proponents of the My School website to rethink their support for the website.

SUMMARY AND CONCLUSIONS

With the current NAPLAN design, where there is only one annual test of 40 questions per subject area, student scores contain large margins of error. NAPLAN results do not provide sufficiently accurate information on student performance, student progress or school performance. Further, no amount of additional material (eg, school funding) about schools can enable anyone to make the claim that the My School website reveals school performance. It is educationally unsound to publish the results and to call on parents to judge schools based on these results. We owe the public the real transparency of the NAPLAN results.

Finally, if you think there is conflicting advice from education experts, there is a simple test. Ask any proponent of the My School website to publicly name one underperforming school. You will soon know the degree of confidence the proponents have with regard to NAPLAN results in making inferences about school performance.

REFERENCES

Leigh A (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review,* Vol 29, No 3, June 2010, pp480-488.

Masters G, Rowley G, Ainley J & Khoo S (2008). *Reporting and comparing school performances*. Paper prepared for the MCEETYA Expert Working Group to provide advice on national schools data collection and reporting for school evaluation, accountability and resource allocation. **http://www.curriculum.edu.au/verve/_resources/ACER_Report_on_Reporting_and_comparing_school_performances.pdf**

Nye B, Konstantopoulos S & Hedges L (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis,* Vol 26, No 3 (Autumn, 2004), pp237-257.

Wu ML (2010). Measurement, sampling and equating errors in large-scale assessments. *Educational Measurement: Issues and Practice* (In press: Volume 29 Number 4).

The Systematic Bias of ICSEA

BARBARA PRESTON

THE STAKES ARE HIGH WITH ICSEA

The Index of Community Socio-Educational Advantage (ICSEA) is the index used to determine the socio-educational level of disadvantage (or simply the socio-economic status or SES) of a school so that it can be compared with schools of purportedly similar SES ("statistically similar" schools) on the My School website, according to results from the National Assessment Program — Literacy and Numeracy (NAPLAN) tests.

Whether or not the SES measure is accurate is of vital importance — it is a matter of high stakes for individual schools and school sectors. The former Minister for Education, Julia Gillard, and former Prime Minister, Kevin Rudd, stressed the importance of parents and community members using the information on the My School website to judge the performance of schools, and consequently take action.

For example, in 2008 at the National Press Club, then PM Rudd was asked: "If parents find out that their school is not performing well, won't they do the rational thing: vote with their feet and leave?" Mr Rudd replied:

> *… we would make no apology for that. The whole idea is to make sure that schools are accountable for their performance. And part of accountability*

> *means that the parents and the students know how that school is performing against agreed standards.*

A year later, Ms Gillard spoke to the Brookings Institution in Washington of the proposal that was to become the My School website:

> *A new system of individual school performance reporting will make the results of new mandatory national testing available to policy makers and parents alike — in a way that will emphasise a school's performance relative to like schools.*
>
> *This transparency is crucial. To improve schools that are failing their students, we need information. And we want parents to drive change.*

But what value is "transparency" if the measure to determine like schools is systematically biased? Access to quality data about important matters such as schooling is essential to a vibrant democracy and civil society. But data and other information must be fit for purpose and not misleading.

ICSEA AND ITS APPLICATION

ICSEA has several components, the most important of which is an index based on 14 variables drawn from Australian Bureau of Statistics (ABS) 2006 Census data at the census collection district (CD) level — it is an area-based index. CDs are localities of very roughly 225 dwellings, but can vary greatly in their population and number of school children. The index is applied to CDs, and the value for a school is derived from the weighted average of all the CDs of students' home addresses (for example, if a quarter of the students live in CDs with high scores of 1200 and the other three quarters live in CDs with low scores of 800, the weighted average for the CDs is 900). The final ICSEA value for a school is adjusted for remoteness and the proportion of Indigenous students, and a small proportion of schools have additional adjustments if considered warranted (ACARA, 2010).

A central element of the My School website is the comparison between "statistically similar" schools regarding their NAPLAN results. On each school's web page is a chart setting out the school's average NAPLAN results. Immediately below are the average scores for "statistically similar schools" (SIM) and for all Australian schools (ALL). A bar above each SIM and ALL scores indicates by its colour whether the particular school is "substantially above" (green), "above" (pale green), "close to" (white), "below" (pink), or "substantially below" (red) statistically similar schools and all schools respectively. In addition, each school's main page links to a page that lists up to 60 statistically similar schools and provides similar data and colour coding.

Thus, a scan of the dominant colours can quickly indicate to the viewer whether or not the school with which they are concerned is substantially above or substantially below statistically similar schools, or somewhere between.

THE POSSIBILITY OF THE ECOLOGICAL FALLACY WITH AREA-BASED MEASURES OF SES

Area-based SES measures can be of great value, particularly when the actual nature of the area is important — its physical and social ecology. Area-based measures can also be very administratively convenient. However, they are vulnerable to the ecological fallacy, which involves inappropriately "inferring individual-level relationships from relationships observed at the aggregate level" (Macintyre & Ellaway, 2000, p332).

ABS has strongly cautioned against the inappropriate use of area-based indexes of relative disadvantage. This is in recognition of the heterogeneous nature of areas as small as CDs. While a school's ICSEA score is an aggregate measure, it is derived through a mechanism of individual measures, and thus the ABS cautions are relevant.

The degree to which the ecological fallacy occurs and is a problem depends on the actual empirical situation and the policy purposes (and unintended consequences) of the analysis. An area-based measure such as ICSEA would be reasonable if all schools were broadly comprehensive (thus with no explicit or implicit selectivity or exclusivity on academic or learning ability grounds) and locally universal, and if the purposes of the measure were not high stakes. This is not the case for Australian schooling as a whole, nor for ICSEA's purposes. (However, a measure similar to ICSEA could be reasonable for internal use by school sectors, taking account of a range of other data, for low stakes purposes.)

The alternatives to area-based measures are individual level measures, perhaps based on occupation and education levels of school students' parents. There would still be problems of selectivity and exclusivity by schools according to academic potential or learning and behavioural difficulty, school location and enrolment numbers, and there would be the impact of different resource levels — but the ecological fallacy would not be compounding those problems.

AN INVESTIGATION OF SYSTEMATIC BIAS IN ICSEA

Systematic bias in ICSEA can only be investigated indirectly, and conclusions cannot be applied to specific schools. The Australian Curriculum, Assessment and Reporting Authority (ACARA), which is responsible for ICSEA and My School, has not made the CD-level component of ICSEA available, so a broadly similar, publicly available index, the ABS Index of Education and Occupation (IEO) (ABS, 2008) is used to determine relative SES of CDs.

The dataset of all Australian census districts (approximately 35,000) with their IEO score was matched with datasets of, first, all school students living with their families in each CD, by level of schooling (primary or secondary) and type of school attended (public, Catholic, or independent), by family income (low, medium or high, with approximately one third of all students in each range); and, second, all school students in each CD, by level and type of school, by whether the student had a home internet connection.

The CDs were sorted into nationwide deciles (10 groups with equal numbers of districts in each) of IEO scores from most disadvantaged to most advantaged.

It is very apparent that private schools, both Catholic and independent, tend not to serve low SES communities. Irrespective of family income, more than 80% of primary and secondary students living in districts in the lowest SES decile attend public schools, while in districts in the highest decile, fewer than 60% of primary students and fewer than 35% of secondary students attend public schools.

In even the lowest SES districts there are high-income families, and in even the highest SES districts there are low-income families. In every decile of relative disadvantage, a larger proportion of public school students had low family incomes and a smaller proportion had high family incomes than did private school students.

In the **most disadvantaged** districts, 64% of public primary school students had low family incomes and only 8% had high incomes, while only 50% of the students in Catholic primary schools had low family incomes and 15% had high family incomes (for independent students the figures were 56% low and 13% high).

At secondary level the pattern was even more pronounced — 59% of students at government schools had low family incomes, and 11% had high family incomes, compared to Catholic schools (42% low and 24% high family income) and independent schools (48% low, 20% high).

It is notable that students living in these most disadvantaged districts around Australia and attending Catholic schools tended to have higher incomes than their neighbours attending either public or independent schools.

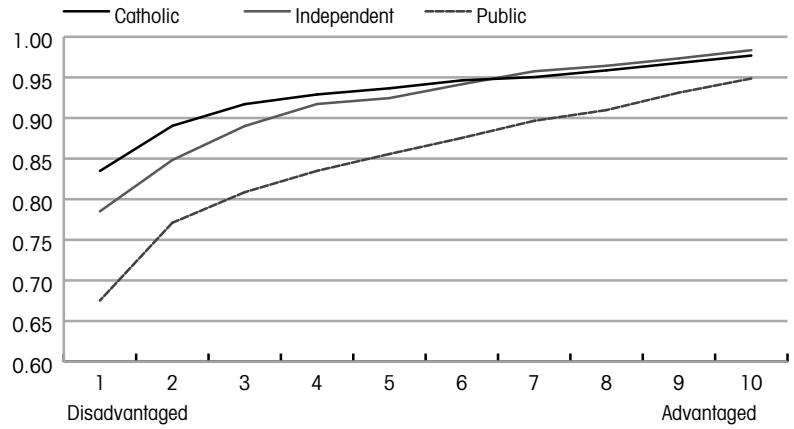
In the **most advantaged** districts, 16% of public primary school students had low family incomes and 65% had high family incomes, against only 11% (low) and 74% (high) in Catholic primary schools and just 8% (low) and 81% (high) in independent schools. At secondary level the pattern was again even more pronounced — 23% low and 54% high in government schools compared to Catholic schools (13% low, 71% high) and independent schools (just 9% low and 80% high family incomes).

In districts at every decile of relative disadvantage, public school students were less likely to have an internet connection at home. In the most disadvantaged decile, only 56% of students attending public primary schools had an internet connection at home, against around 70% of private school students, and only 68% of secondary public school students against 80% of private students. The difference is much less pronounced in the most advantaged deciles because nearly every student had a home internet connection, but still students attending public schools were less likely to be connected (see graph 1).

The national pattern is reflected in Victoria. Statistical local areas (SLAs) can also be ranked according to their IEO score from most disadvantaged to most advantaged. In the most disadvantaged SLAs, such as Hume–Broadmeadows and Melton–Bal, around two-thirds of secondary students attend public schools, while in the most advantaged SLAs, such as Boroondara–Hawthorn and Stonnington–Malvern, fewer than a fifth of secondary students living there with their families (ie, excluding boarders) attend public schools.

Irrespective of the percentage attending public schools, in the SLAs from the most disadvantaged to the most advantaged, a much higher proportion of those attending

GRAPH 1: PERCENTAGE OF PUBLIC, CATHOLIC AND INDEPENDENT SECONDARY SCHOOL STUDENTS LIVING IN EACH DECILE OF DISADVANTAGE WITH HOME INTERNET CONNECTION, AUSTRALIA, 2006



public schools have low family incomes, and a much lower proportion of those attending public schools have high family incomes.

For example, in Hume–Broadmeadows, of secondary students attending public schools, 58% have low family incomes, and only 13% have high family incomes. Of those attending Catholic schools, just 35% are low income and 28% high. (In the most disadvantaged areas, independent secondary students tend to have lower family incomes than those attending Catholic schools.)

For secondary students in the most advantaged area of Booroonda–Hawthorn the difference is even starker: 31% attending public schools have low family incomes, and 34% high compared to Catholic and independent schools, at 9% (low) and 81% (high). (See Table 1 overleaf.)

CONCLUSION

This analysis has been broad and indicative, and no firm conclusions can be drawn about specific schools. However, it is clear that in census districts from the lowest to the highest SES, the type of school attended is not random — students attending public schools tend to have lower family incomes and are less likely to have a home internet connection.

It therefore appears that ICSEA is systematically biased against public schools. That is, it is likely that public schools in general have inflated ICSEA scores, and Catholic and independent schools have deflated ICSEA scores.

Thus, when the My School website compares "statistically similar" schools, public schools are on average being matched with private schools that actually enrol higher SES students than reflected in their ICSEA scores. This would tend to indicate

TABLE 1: SECONDARY STUDENTS LIVING IN SELECTED VICTORIAN STATISTICAL LOCAL AREAS (ORDERED BY PERCENTILE RANK ON THE INDEX OF EDUCATION AND OCCUPATION), TOTAL NUMBER OF SECONDARY STUDENTS, PERCENTAGE OF TOTAL ENROLLED IN EACH SCHOOL SECTOR, AND PERCENTAGE ENROLLED IN EACH SECTOR WITH LOW AND HIGH FAMILY INCOMES

SLA and IEO percentile rank in Victoria		Secondary students who live in each SLA									
		Total number	% of total enrolled in each sector			% enrolled in sector with LOW family income			% enrolled in sector with HIGH family incomes		
			Public	Catholic	Independ	Public	Catholic	Independ	Public	Catholic	Independ
Hume (C) - Broadmeadows	2	3,500	67%	17%	16%	58%	36%	59%	13%	28%	18%
Melton (S) - Bal	6	2,514	68%	16%	16%	45%	29%	21%	23%	37%	52%
Hume (C) - Craigieburn	22	3,844	57%	26%	17%	39%	23%	26%	24%	38%	43%
Hume (C) - Sunbury	46	2,199	64%	24%	13%	29%	18%	11%	34%	49%	68%
Moonee Valley (C) - West	56	1,948	48%	36%	16%	42%	26%	18%	27%	45%	50%
Moonee Valley (C) - Essendon	84	3,145	41%	35%	24%	35%	21%	13%	44%	58%	72%
Stonnington (C) - Prahran	97	880	23%	15%	61%	43%	32%	15%	30%	40%	70%
Stonnington (C) - Malvern	98	2,073	13%	23%	64%	27%	13%	9%	45%	66%	80%
Boroondara (C) - Hawthorn	99	1,165	19%	16%	66%	31%	9%	9%	34%	81%	81%

Note: Data excludes boarders and others who do not live with family

Source: ABS (2008), and ABS 2006 Census TableBuilder

that public schools are "performing" more poorly than in fact they are (whether this is above, at the same level, or below statistically similar schools) and that private schools are performing better than they actually are.

There are, of course, exceptions to this general pattern, though there is no transparent way of determining these exceptions. It may be very clear that a selective school will be shown as performing well on the My School website in large part because of that selective intake. But much less clear is that a school may appear to perform badly because it has lost high-performing students to the selective school. There are many subtle as well as explicit selection and exclusion practices by schools in all sectors, and many other schools affected by these practices — in each case the nature of the student body, which is not reflected in an ICSEA score or an individual-level SES measure, determines NAPLAN results to a greater or lesser degree.

Quality data and information are important, but we should not forget the lessons that led to the development in the 1970s by Donald T Campbell of his "laws":

> *The more any quantitative indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.*
>
> — Campbell 1976, p49

REFERENCES

Australian Bureau of Statistics (2008). *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia — Data only, 2006*. Cat. No. 2033.0.55.001. Canberra: Australian Bureau of Statistics. **http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012006?OpenDocument** (retrieved June 20, 2010).

Australian Curriculum, Assessment and Reporting Authority (2010). *Technical Paper: Index of Community Socio-Educational Advantage*. Sydney: Australian Curriculum, Assessment and Reporting Authority. **http://www.myschool.edu.au/Resources/pdf/My%20School%20ICSEA%20TECHNICAL%20PAPER%20 20091020.pdf** (retrieved June 20, 2010).

Campbell DT (1976). *Assessing the Impact of Planned Social Change*. Hanover, NH, US: Dartmouth College Public Affairs Centre. **http://www.wmich.edu/evalctr/pubs/ops/ops08.pdf**.

Gillard J (2009). A new progressive consensus for Australian schools. Address to the Brookings Institution, Washington DC, June 19. **http://www.brookings.edu/~/media/Files/events/2009/0619_australian_ education/0619_education_gillard.pdf** (retrieved August 9, 2010).

Macintyre S & Ellaway A (2000). Ecological Approaches: Rediscovering the role of the physical and social environment. In Berkman LF & Kawachi I (eds), *Social Epidemiology* (pp332–348). Oxford, UK: Oxford University Press.

Rudd K (2008). Questions and Answers, National Press Club, Canberra, August 27. **http://pmrudd.archive.dpmc.gov.au/node/5620** (retrieved August 9, 2010).

What NAPLAN Doesn't Address
(but could, and should)

JAMES G LADWIG

I WOULD LIKE to push the debates about NAPLAN and My School onto what I think is relatively uncultivated but crucial ground for Australian educators. If we are serious about improving student academic learning outcomes, we must start addressing within-school differentiation in the quality of teaching, curriculum and assessment.

Over the past decade or so, educational policy has started to make use of "multi-level" analyses, so-called because of the nature of the data that comes from studies of students. All students are "nested" — within classrooms, within schools. Whenever we get data of this type it is important to separate our measures into the relevant levels of nesting. If we don't, there is often a good chance that our estimations of each level (and any effect measure of relationships) are inaccurate. In Australia, this form of analysis has already had an impact in policy with the recent strong focus on teachers (within schools).

From these multi-level analyses, both within Australia and internationally, some well-known patterns apply. One is just how much of the differences in achievement between students is related to each level of nesting. This goes some way toward telling us something about just what schools can effect. One problem for current education

policy formulations in Australia is that, in general terms, multi-level analyses have shown us that variance in student outcomes does not match the "school difference" focus of current policy directions.

If we use only three levels in a multi-level analysis, we find (speaking generally) that differences *between* schools account for (roughly) 10–15% of student achievement variance, whereas differences *within* schools account for roughly 40–45% of variances, and differences between individual students account for roughly 40–50%. (Estimates vary from study to study, and statistical model to model.)

The general picture backs up what most teachers will tell you: we know the things schools can most directly leverage are those sources of student achievement that are themselves most closely connected to student learning outcomes: the "enacted curriculum" (learning experiences of kids, in the classrooms, halls, playgrounds and excursions). Even if we assume a fair amount of variance can be attributed to things that schools cannot change (individual student backgrounds prior to entering school, different levels of resources between sectors, etc), there is still a big chunk of the problem that schools can address. (This was the logic and research that pushed the "school effectiveness" agenda along in the late 1980s.)

The policy implications of this are very big indeed. To be blunt, any serious systematic attempt to improve these things has cascading effects well beyond schooling. If we ever get serious about this we will have to dramatically change current and future teaching career structures and working conditions, university budgets, and government priorities relating to social support, job creation, health, science and research funding, to name just a few obvious issues.

In Australia there are several policy conditions of significant consequence. First, since the "devolution" policies of the 1980s and 90s, most (perhaps all) systems of schooling in Australia have essentially purged the main means of monitoring the quality of classroom practice: the inspectorate. Without defending a flawed system, an unpleasant reality has to be acknowledged: no systemic alternative has been put in place. This matters for obvious reasons, but is directly related to any attempt to use measures of student achievement as an indirect proxy measure of teacher performance. The reason many people hope to use NAPLAN for this purpose is a direct response to having few other options.

The second policy condition we need to note is that it is only in the past decade that Australian state and territory systems have developed the database technologies and techniques to match individual students effectively (and find students who move schools in their data).

Thirdly, we have inherited a very limited interpretation of the knowledge about within-school variance. Having realised the relative size of within-school variance, Australia has taken a near singular focus on teachers themselves as the ostensible source. While there is some truth to this, it is only a partial truth. We need to keep in mind that the demographic characteristics and credentials of individual teachers do not determine the quality of what they do. For example, where there may be

some sense in establishing institutes to credential teachers as "competent" or "professional", it would be a mistake to assume that this actually estimates the quality of what teachers do in classrooms all that well. Studies which attempt to measure the effect of teachers on student learning without some means of estimating the quality of what students experience will always be limited in how much they can inform policy.

The fourth policy condition we face is political acceptance of what many educators have long pointed out: that any attempt to deal with socio-economic and racial disparities requires direct policy action. Arguing that universal choice models or market-drive reforms alone will close equity gaps in education is ideological folly. At some point we have to find ways to address these inequities within schools — in and between classrooms — if we are to have any hope of providing equal opportunities to students by the time they reach post-compulsory schooling age.

PUTTING TOGETHER THE PIECES OF THE PUZZLE

To illustrate the general points above, I need to draw on data from the 2006 round of the OECD PISA study, and data from the ARC funded SIPA study into the quality of pedagogy, directed by myself and Jennifer Gore in New South Wales.

First, consider the 2006 PISA study. Taking as an example the results in science, the Australian data shows a "top and tail" distribution, with lots of schools pretty close to each other in the middle, and a small portion on both the top and bottom ends of the distribution. For the majority of schools, the confidence intervals overlap — which indicates that we cannot say (with 95% confidence) that they actually differ from each other. This means that it is very sketchy to make comparisons except at the extremes of the range.

The relatively wide confidence intervals set against the closeness of most schools hints at a lot of "within-school" variance (the two are mathematically related). Using multi-level analyses of two levels of PISA (between schools and between students within each school), the differences between schools accounts for only 19% of the overall variance in outcomes, whereas 81% of the variance is within schools.

This distribution is very common in industrialised countries on just about every standardised test and is the basis for major criticisms of school league tables. There is little doubt that NAPLAN results face the same patterns and problems as PISA.

But the question I would like to raise takes a few steps back: the need to understand the within-school variance better than simply assigning it to individual students and teachers.

Our SIPA study observed and collected data in the classrooms of 322 teachers in 35 NSW schools from 2004–07, using very direct measures of the quality of pedagogy experienced by students. The findings illustrate the same problems of making direct comparisons of mean scores — in this case between teachers.

Here some policy makers might jump to the same temptations that push them to draft school league tables, and suggest this is all about differences between teachers. A two-level analysis of the SIPA data shows that 14% of the variance is between schools and 86% is within schools (between teachers).

But is this within-school variance just "teacher" difference? No doubt some portion of it might be, but not all of it. We have to address one of the main sources of that difference: how schools organise classrooms and students into classes. A majority of teachers in the study were observed in more than one class or subject.

Mapping the data against students' prior achievement (as measured by the NSW basic skills test) we found clear evidence that once students are grouped into classes, there are some pretty big differences based on prior achievement. We also found that the higher the prior achievement, the better the tasks students are given. These two points won't surprise most teachers — but they do raise significant questions about the distribution of quality within schools.

Overseas research has indicated that high-quality, intellectually challenging and interactive pedagogy benefits students irrespective of "ability". For SIPA, one main question was whether or not that was true here. We needed to find classes in which students who traditionally did not do well in school experienced high-quality pedagogical tasks; we found no examples of this.

SOME VERY OPEN QUESTIONS

Clearly there is a large amount of pedagogical variance within schools; however, not all is simply "between teachers". From my own research, I am quite sure that much is between classes, related to how students are grouped. From international research, we also know there is strong social co-linearity between social backgrounds and prior achievement. Unfortunately for Australia there are very few studies of this phenomenon. This begs several questions:

- What are the practices of class grouping based on prior achievement?
- What is the social distribution within "prior achievement" groups?
- Are students "stuck" in groups once placed?
- Over how much time are students within the same groups?
- How much do these groups vary from subject to subject?

Each of these issues carries significant consequences for an assessment of the effect of the "ability" grouping practices of schools — and many of the well documented effects overseas have been negative. We have no answer to perhaps the most important question: What are the effects of "ability" grouping, streaming and tracking in Australia?

One thing is very clear: the practice of grouping students by "ability" is widespread. According to data from the 2006 PISA school questionnaire, 89.5% of Australian schools report ability grouping within or between classes for at least some subject areas, and a further 5.2% report ability grouping for all subjects — either within or between classes, or both. But little more is known about how these practices relate to curriculum, teaching, assessment and ultimately affect student participation and learning in general. We do have ample reason to speculate that these practices account for some portion of the known large within-school achievement variances.

CONCLUSION

In their current regime, NAPLAN and My School will do little to address this situation, but they could. NAPLAN is perhaps one of the most significant data sets on schooling in Australia. However, in its current form, the questions it could potentially help answer cannot be addressed. The difficulty of comparing schools statistically is well recognised by those charged with analysing and reporting NAPLAN results — but these professionals are restricted in what they can (and should) report. Intelligent public debate is hampered by a lack of public reporting of the technical characteristics of NAPLAN and a lack of public education on how to understand such data.

Further, the potential power of the NAPLAN data could be geometrically advanced if it were included in a broader national research agenda, open to a larger body of researchers who know what can be done with it. It would not take much more effort to augment the data collection exercise of NAPLAN and My School and begin analysing the internal curricular and teaching practices of Australian schools in a way that could advance the national goals of improving student learning outcomes far more than any imported policy agenda. The advent of a national curriculum, with its potential risks, makes the need to understand within-school differentiation all the more important.

Much of the current Australian educational debate and policy does not really focus on the main issue if we are to address our main levers for improving student outcomes: within-school curriculum and pedagogy differentiation. While there is a huge amount known internationally, we know very little about how this differentiation plays out in Australia. And that is only the beginning. If international experience holds here, changing this differentiation is a much bigger challenge than measuring students and schools. But we can not even begin the task of developing alternative within-school practices intelligently until we harness our data and research capacity in a more educationally productive manner.

Will the Australian Curriculum Up the Intellectual Ante in Primary Classrooms?

ALLAN LUKE

> *The evidence is now pretty clear. We seem to have figured out how to teach the "basics" to just about everyone. … But we are deeply unsuccessful at our 21st century agenda of moving beyond basic competencies to proficiencies.*
>
> — Lauren B Resnick, Wallace Foundation Distinguished Lecture (*Educational Researcher,* 2010)

WHAT IS SURPRISING in the past two years of debate over an Australian national curriculum has been the failure to engage with research on the enacted curriculum: on what actually goes on in Australian classrooms.

Any official curriculum comes to ground via an *enacted curriculum* of teaching and learning events "lived" by students and teachers. Specific knowledges and skills can only be "named" in official curriculum documents at a level of technical abstraction. They are brought to life in classrooms. The enacted curriculum will be influenced by adjacent policies and practices on assessment, evaluation and credentialing (which define "what will count") — and on school funding, governance and leadership, teacher training and professionalism, and so forth.

In short, there is no direct "hypodermic" effect between the official curriculum and the enacted curriculum.

A CURRICULUM IN SEARCH OF A PROBLEM

The national curriculum, now in its draft stages, remains a solution seeking a robust demonstration of an educational problem. The debate to date has been a forum for contending claims about appropriate ideological and cultural content. These range from divided views on prescribed literary content, the place of new media versus a continued concentration on print, to an ongoing debate over which version of Australian history will be made to count. The social and political process of reaching a settlement can lead to official syllabus statements that are controversial and potentially unstable.

It can also lead to an educationally unwieldy response based on perceived but empirically undemonstrated problems of current teaching and learning. There are two major risks here. First, that the settlement will attempt to incorporate all views on requisite skills and knowledges — leading to bloated or conceptually incoherent syllabi that, simply, cannot be taught. Second, that the settlement might be a response to claims about current classroom practices that are not well substantiated. To date, there has been no comprehensive survey of teachers' uses of official materials.

So perhaps a return to square one is in order: to the core educational problems to be addressed by a national curriculum.

The principled arguments for a national curriculum tabled are: (1) persistent patterns of educational inequality in test score achievement, retention and completion — with students from lower socio-economic backgrounds and Indigenous students' achievement lagging behind that of their middle and upper socio-economic background counterparts; and (2) reportedly inadequate levels of skill, competence and capacity of the "human capital" entering the tertiary training system and workforce.

The patterns of Indigenous achievement in all categories are a national disgrace. And the impacts of socio-economic disadvantage on test score achievement are significant — though the "equity gap" appears to persist in the US, UK and other countries, despite a decade of test-driven policies (Luke & Woods, 2008).

A decade of data from the US *No Child Left Behind* policy indicates that a mandated curriculum emphasis on direct instruction in basic skills can yield improved but limited early performance in standardised tests, as US psychologist Lauren Resnick's comments above suggest.

But the American experience also shows that closure of the equity gap is fraught with more complex problems. Many early gains have ceiling and transient effects; others are not reflected in more robust performance measures such as the National Assessment of Educational Progress, with widespread residualisation of early gains among the most at risk students (the "fourth grade slump"). Documented collateral effects include increased teacher attrition, and decreases in retention and completion, and a diminution of Indigenous curriculum content. "Teaching to the test" is both an

intended and collateral effect, which reportedly has led to a narrowing and diminution of intellectually engaging, challenging and relevant knowledge in many schools.

The empirical literature on what actually occurs in Australian classrooms — specifically, the Queensland School Longitudinal Achievement study (Lingard et al, 2001) and the New South Wales Pedagogies projects (Ladwig, 2005; Ladwig & Gore, 2005) — paints a very different picture to the view propagated in the media and public debates. These studies suggest that the problem is not a lack of basic skills teaching.

Based on systematic observations of large corpi of randomly selected lessons, they corroborate the finding that sustained achievement gains among the most "at risk" students require:

(1) An everyday focus on curriculum content and issues of substantive intellectual demand and depth;
(2) Sustained scaffolded student talk and dialogue around issues of cultural and intellectual substance; and
(3) Visible connections of school knowledge to everyday civic, cultural, political and social life.

Basic skills acquisition is necessary but not sufficient for sustainable engagement and achievement.

In both studies, much of the instruction observed was devoted to basic skills and basic curriculum content. Despite the stereotypes and anecdotes in the press of unprincipled progressivism or politically correct value stances, the studies found high frequencies of lessons where students completed worksheets, copied from the board, answered questions at the end of chapters and engaged in "busy work".

Further, the assessable tasks set by teachers often focused on lower order domains of activity, such as recall, reassembly of existing and provided information, and skill repetition. The overall picture counters the Vygotskian axiom of stretching students beyond what they can readily do.

A VIEW FROM THE FIELD

To fill in this with a picture from the field, I turn to early findings of two ARC-funded projects currently underway.

In a quantitative study of the impacts of social class on Year 1 literacy, Sue Grieshaber, Paul Shield, Shelley MacDonald and I surveyed and interviewed a random stratified sample of 106 Year 1 teachers/classrooms in 44 greater Brisbane metropolitan area schools, with achievement data on 650 of their students.

The three-year study examined social class and home literacy resources of families, teacher self-reported curriculum emphases in early literacy, and results on three outcome measures: standardised reading achievement tests, a "best exemplar" narrative or descriptive writing sample, and teacher-moderated judgments on a language/literacy developmental continuum.

While analysis is still underway, early findings (Luke, Grieshaber, Shield &

MacDonald, 2009) are telling:

- Schools which score low on the Index of Relative Socio-Economic Disadvantage (IRSED) have an overrepresentation of teachers with less than four years' experience; while mid and high IRSED schools have an overrepresentation of teachers with 20 years or more of experience
- Low IRSED parents reported higher expectations that their children would enter Year 1 with some basic reading and writing skills than parents in mid-high and high IRSED groups
- Teachers in low IRSED schools concentrated on direct instruction in alphabetics, including a focus on phonemic awareness, drill in grapheme/phoneme generalisations, alphabetic knowledge, and knowledge of basic print conventions. *Teachers in low IRSED schools reported an average of 550 minutes per week of instructional emphasis on coding*
- Teachers in mid and high IRSED schools reported overall less instructional time on coding and on all aspects of literacy instruction.

These findings raise questions about many of the assumptions driving the current curriculum debate. In Queensland's urban primary schools, the "basics" are a strong curriculum focus.

This is illustrated in the first year of another ARC study, where a QUT and Queensland Teachers Union research team are focusing on the relationships between digital and print literacies across the curriculum in a low socio-economic school with a significant proportion of Aboriginal and Torres Strait Islander students and high overall levels of special needs ascertainment (Luke, Dooley & Woods, in press/2010).

While working in classrooms and undertaking curriculum planning with teachers, we have observed a strong emphasis on the teaching of basic skills in reading and writing. In the past year, with the policy focus on NAPLAN and the My School website, teachers have focused increasingly on: (1) increasing attendance and time-on-task through school-wide behaviour management systems; (2) explicit emphasis on direct instruction in skills likely to be tested and in test-taking procedures and formats.

The current situation, school administrators hope, will enhance student performance on decoding and comprehension sections of NAPLAN tests. But we and the administrators and curriculum leaders have noted little direct engagement with matters of intellectual and cultural substance. Sustained engagement with the sciences, humanities and social sciences is rare; scaffolded dialogue about issues of literary and moral content is infrequent; and community-based cultural knowledge, issues or texts are rarely addressed in the classroom. *The enacted curriculum has a strong focus on basic skills with little sustained or conceptually coherent focus on knowledge*.

We are exploring the specific combination of factors which have led to this situation. For our present purposes, I would offer the following working hypothesis:

- The closure of the "equity gap" cannot be addressed by an emphasis on the teaching and high-stakes testing of basic autonomous skills and behaviours

- Sustained improvement in the performance of students from low socio-economic and Indigenous communities will require an enacted curriculum that features: intellectually challenging, demanding and interesting knowledge; sustained and scaffolded linguistic interaction around and about that knowledge; and demonstrable links between school knowledge and the everyday realities of Australian life, cultures and work.

Calls for a "return to the basics" have been recurrent in Australian education over the past five decades. Yet there is evidence here of a national "over-correction" to a perceived loss of the basics. Commenting on trends in PISA reading achievement, Barry McGaw (2010, p5) recently commented:

> *Australia's rank dropped … primarily because of a decline in performances at the highest level. The reasons for this are not immediately evident from the data but it [is] at least clear that it is due to schools focusing more on basic achievement levels and not so much on the development of sophisticated reading of complex text.*

UPPING THE ANTE

To date, the Australian Curriculum has the hallmarks of the new generic, transnational curriculum settlement that emerged in the late 1990s as a response to new economic and social contexts. This features a focus on basic skills and a reinstatement of a canon in literature, science and history. It attempts to address the emergence of digital cultures and transnational economies through a complex overlay of "new capacities". In this regard, it resembles the curricula of most Canadian provinces, New Zealand, the UK, many US states, Hong Kong and Singapore. All state a commitment to the production of human capital for "new economies", to equitable access, and to quality and competitive "innovation".

International comparative studies such as PISA and TIMMS have spurred convergence towards a *de facto* transnational standardisation (and hence, skill and knowledge description) in early childhood learning, mathematics, sciences, and, indeed, literacy. Whatever might count as distinctively "Australian" (eg, literary content, historical knowledge) would appear within these structural boxes.

This said, the test of the Australian Curriculum is whether it ups the intellectual ante and educational bar. It is worth considering at length Resnick's (2010, p187) retrospective view of the chain of events that have driven US educational policy since the initiation of bipartisan educational reform:

> *Current policy discussions often aim to solve the problem of disappointing levels of learning by investing heavily in theories of performance management. The prescription for better performing schools … is more frequent measures of student performance and greater attention to this "output" data. … This has led to a virtual industry of student measures*

> *that can be administered early and often, in the form of interim, or benchmark, tests. As noted earlier, these tests have come to control the* de facto *curriculum… These items, for reasons of cost, familiarity, and certain psychometric considerations, are mostly simple multiple-choice questions, with little opportunity for the kind of interpretive knowledge work that the Thinking Curriculum calls for. This growing practice… exaggerates the "basics skills" character of the standards movement.*

In Resnick's view, the task at hand is a cognitively demanding, intellectually rich "thinking curriculum", which will require in the US an increased focus on improving teacher knowledge and the enhanced development of pedagogical practices. Her warning is blunt: "Systems that aim to develop extended knowledge and complex forms of argument and reasoning among students will fail if teachers are restricted to scripted lessons that close off discussion" (p196).

A similar policy sequence — already a decade old in the US and UK — is well apace in Australian schools. It is axiomatic that the higher the stakes, the more that teachers, principals and school systems will subordinate or disregard elements of the official curriculum not formally assessed. League tables, standardised testing and media critique of teachers and standards are likely to generate short-term test score gains, especially in the early years. But should the patterns follow the US and UK, longer term results — which happen to extend beyond the immediacy of electoral cycles — are likely to see ceiling and threshold effects, upper primary residualisation of performance among equity groups, ongoing primary/secondary transition issues, and largely unchanged retention and tertiary participation rates.

The test of this national curriculum will be whether it sets the conditions for yet another "back to the basics" movement — with the potential to further narrow, fragment and trivialise the enacted curriculum — or whether it succeeds in focusing systems, bureaucrats, teachers, teacher educators and, ultimately, students' conversations on matters of intellectual demand, cultural meaning and substance, and on an exploration of the complex and critical issues, designs and knowledges for new economies and risky worlds.

Its task will be arduous: for it must supplant a *de facto* enacted curriculum of test preparation and basic skills that many presume is the solution.

NOTE

This is an abridged version of a longer paper that will appear in *Curriculum Perspectives*.

REFERENCES

Ladwig J (2007). Modelling pedagogy in Australian school reform. *Pedagogies* 2, 57–76.

Ladwig JG & Gore JM (2005). Measuring teacher quality and student achievement. *Professional Educator* 4(2), 26–29.

Lingard R, Ladwig J, Mills M, Hayes D, Christie P, Ailwood J, Warry M, Gore J & Luke A (2001). *Queensland School Reform Longitudinal Study: Final Report*. Brisbane: Education Queensland.

Luke A, Dooley K & Woods A (in press/2010). Comprehenion and content: Planning literacy curriculum in low socio-economic and culturally diverse schools. *Australian Educational Researcher*.

Luke A, Grieshaber S, Shield P & MacDonald S (2009). Literacy and social class in Year 1: A descriptive picture. Paper presented at the Futures of Literacy Conference, University of Sydney, 9/09.

Luke A & Woods A (2009). Policy and adolescent literacy. In Christenbury L, Bomer R & Smagorinsky P (eds) *Handbook of Adolescent Literacy Research* (197-219). New York & London: The Guilford Press.

McGaw B (2010). Reducing the impact of social background in education: Fond hope or realistic aim? Murdoch Lecture, Perth, 17/3/10. Retrieved 20/5/10 from: **http://www.education.murdoch.edu.au/pdf_files/McGaw%20-%20Murdoch%20lecture.pdf**

Resnick LB (2010). Nested learning systems for the thinking curriculum. *Educational Researcher* 39, 183-197.

The Impact of High-Stakes Test-Driven Accountability

BRIAN J CALDWELL

I GAVE VOICE to my concerns about high-stakes testing on May 29, 2009 when I delivered the Agitation Hill Lecture in Castlemaine on the topic "Want world-class schools? It's time to agitate" (Caldwell, 2009). I focused in particular on plans for NAPLAN and what we now know as the My School website.

I concluded with the statement: "Unless there is agitation on an epic scale, such as refusal by teachers to administer the tests or by parents to have their children sit for them, it seems that league tables of limited validity that mean little to parents are a *fait accompli*." I added: "Parents and the profession will be standing on the high ground if the language of radical dissent is adopted on this issue."

Everything that has occurred since then leads me to affirm these conclusions today.

THE IMPACT OF HIGH-STAKES TESTING

The use of NAPLAN in the My School website has been based on developments in England and the United States, with Joel Klein, who heads the public school system in New York, called on from time to time to spruik the benefits. It's worthwhile to summarise the evidence of impact in both places.

ENGLAND

The Cambridge Primary Review of policy and practice in England was published in 2009 under the title *Children, Their World, Their Education* (Alexander, 2009). Project director Robin Alexander delivered the Miegunyah Distinguished Lecture at the Melbourne Graduate School of Education on March 10, 2010 on the topic "The Perils of Policy: Success, amnesia and collateral damage in systematic educational reform" (Alexander, 2010). Many of the fears in Australia about the dysfunctional effects of national testing, an excessive focus on and unrealistic expectations for standards, the narrowing of curriculum, and high levels of stress for students and teachers have been borne out in experience in England.

Alexander was careful not to make comparisons or offer recommendations about the implications for Australia, but the message was not lost on his audience. I sensed that many were shell-shocked, especially when he drew comparisons between England and Finland, which has no national tests, has decentralised decision-making, and provides high levels of school and teacher autonomy. Finland has a high-performing school system where students do not start school until they are seven. Finland is in the top rank of nations as far as innovation is concerned. All but about 2% of students attend public (state) schools.

In an eloquent statement at the launch of the Schools First initiative in 2008, Julia Gillard declared: "All children have some gift and even some potential greatness within them. Finding that gift, nurturing it and bringing it to life is the responsibility of every single one of us."

Her words echo those of Sir Ken Robinson, who is a powerful advocate of an intensely personal approach to learning. Writing in *The Element* (Robinson, 2009), he stated:

> *Education doesn't need to be reformed — it needs to be transformed. The key to this transformation is not to standardise education but to personalise it, to build achievement on discovering the individual talents of each child, to put students in an environment where they want to learn and where they can naturally discover their true passions.*

But Robinson warns that the policy framework in England, now being replicated to a large extent in Australia, will impair the nurturing of the "gift" and "potential greatness" of which Gillard spoke:

> *Education is being strangled persistently by the culture of standardised testing. The irony is that these tests are not raising standards except in some very particular areas, and at the expense of most of what really matters in education* (Robinson, 2009).

Anyone in doubt about these effects in England would have been blown away when they accessed *The Times* (UK) website last year, which published the league tables

of primary schools based on national tests (ranked from 1 to more than 13,000). Accessing the website from Australia led to an advertisement for an Australian company offering its coaching services for NAPLAN tests, which in turn led to a link to Beyond Blue, the national organisation that deals with depression and related conditions.

UNITED STATES

One of the most powerful critiques of the impact of high-stakes testing in the United States was offered recently by Diane Ravitch, a distinguished scholar who has supported the approach in the past. She was an adviser in the administration of George W Bush. She has changed her mind.

Referring to the impact of the No Child Left Behind legislation, she described how "many states now claim dramatic improvement in their test scores, but these gains are not reflected on the tests given every other year by the Federal Government":

> *Many schools suspend instruction for months before the state tests, in hopes of boosting scores. Students are drilled on how to answer the precise types of questions that are likely to appear on state tests.*
>
> *As more time is devoted to reading and maths, and as teachers are warned that the scores in these subjects will determine the fate of their school, everything other than reading or maths gets less time* (Ravitch, 2010).

PERSONAL AND POLITICAL PERSPECTIVES

Nothing in the foregoing questions or challenges the importance of literacy and numeracy and of raising levels of achievement in these areas. Nothing in the foregoing precludes a critically important role for state, regional and district levels of government in providing support for schools — but these must be viewed as enablers rather than regulators. All of these things can be accomplished within a national framework that provides the necessary sense of direction.

I am not a "Johnny Come Lately" on these issues. My second book with Jim Spinks, *Leading the Self-Managing School*, was published in 1992. We had the following to say about the impact of high-stakes testing in the light of developments in England where the Conservative Government had introduced national testing in primary and secondary schools and league tables were starting to appear:

> *There are limitations in these approaches to testing, not the least of which is their narrow focus and the resultant distortion which may occur in learning and teaching, especially for testing at the primary level. Highly valued goals may be devalued* (Caldwell & Spinks, 1992, p142).
>
> *There will be great pressure to release the raw data of school-by-school comparisons in a manner that will distort the accountability process. … In*

> *our view the strongest possible stand should be taken against the release of such data when accompanied by claims or implications of relative effectiveness* (ibid, p155).

I became more sympathetic a decade later when improvements in literacy were registered in England as a result of an unrelenting effort by the Blair Government, which expanded the testing and league tables approach of the Conservatives. However, achievement soon plateaued. One of the most powerful advocates, Secretary of State for Education Estelle Morris, resigned when targets for improvement had not been achieved.

The observations I am making and the conclusions I am drawing should not be interpreted in a narrow political frame, especially as this symposium is being held during a federal election campaign. I am alarmed at what the Government is doing in partnership with the states and territories. I am equally if not more alarmed at what has been proposed by the Opposition, which calls for testing at every level from Grade 3 to Grade 10 (as reported by Harrison, 2010). Such a proposal shows no understanding of the nature and purpose of assessment and testing, no understanding of the capacities and potential capacities of the profession and no understanding of what has occurred elsewhere and the evidence of harm.

My conclusion is that despite the popular appeal of the national curriculum, national testing and the My School website, we are unlikely to see more than marginal and short-term improvement in outcomes for all students, or a closing of the gap between high-performing and low-performing students, until such time as we move ahead on a number of fronts. We must open the doors to the creative spirit in our schools that should operate in the future in the broadest of national frameworks. If we can't do this we may make progress in the short term, but other nations are moving faster and further and we'll soon be left behind.

ADVICE TO MINISTERS AND ACARA

The purpose of this symposium is to offer advice to ministers and ACARA. I will adopt a futures perspective in formulating my advice, which is presented as a scenario for 2020.

Writing a scenario is a useful technique for thinking about the future. A scenario is not a prediction. It describes an alternative future, either probable or preferred, with a narrative that credibly explains the pathways from the present to that future. In *Our School Our Future* (Caldwell & Loader, 2010), we provide a template for writing scenarios at the school level.

Here is a narrative for an alternative future to 2020 that describes how we might reverse current and likely trends as far as the impact of high-stakes testing is concerned, while we nurture innovation, creativity and passion.

In 2020 there will be a higher level of transparency and more testing in Australia's schools than in the past. However, approaches associated with NAPLAN and the My School website introduced in 2009 and 2010 have been abandoned because a united profession and the public soon realised that expectations had not been

realised. They inhibited rather than drove the transformation of schools. Australia became isolated in the international community for proceeding with an approach that had been abandoned elsewhere as country after country moved ahead with a more enlightened approach.

Long-overdue reforms in teacher education meant that teachers have become expert in skilful testing, diagnosis of need and immediate support of their students in an unprecedented and comprehensive approach to personalising learning. Every school or community of schools has teachers and other professionals on call who give immediate support to their colleagues to ensure that no student falls behind. A remodelled national agency prepares tests that schools can choose if they wish, but the level of professional skill ensures that most schools design their own. Parents obtain real-time online reports of how their sons and daughters are progressing, and meaningless out-of-date comparisons of schools have been abandoned. Teaching to the test and the narrowing of the curriculum are dysfunctions of the past. The curriculum has been broadened to address the range of knowledge and skills demanded in the 21st century. Schools have far more autonomy than in the past, with many opting for an international rather than national curriculum.

Innovation and creativity flourish and there has been a resurgence in the arts and science. New world-class facilities have been an important factor in attracting able people to the profession. There is a passion that has not been evident for several decades.

REFERENCES

Alexander R (ed) (2009). *Children, Their World, Their Education*. Final report and recommendations of the Cambridge Primary Review. London, UK: Routledge.

Alexander R (2010). The Perils of Policy: Success, amnesia and collateral damage in systematic educational reform. Miegunyah Distinguished Lecture at the Melbourne Graduate School of Education on 10 March. Accessed July 3, 2010 at **www.edfac.unimelb.edu.au/news/lectures/2010series.html**.

Caldwell BJ (2009). Want World-Class Schools? It's Time to Agitate. Invited Agitation Hill Lecture delivered on Agitation Hill, Castlemaine, Victoria, May 29. Available at **www.educationaltransformations.com.au**.

Caldwell BJ & Loader DN (2010). *Our School Our Future*. Canberra: Australian Institute for Teaching and School Leadership (AITSL).

Caldwell BJ & Spinks JM (1992). *Leading the Self-Managing School*. London, UK: Falmer.

Harrison D (2010). Push for more national school tests meets chilly response. *The Age,* 21 July.

Ravitch D (2010) Why I Changed my Mind. *The Nation,* 27 May. Accessed July 19, 2010 at **www.thenation.com/article/why-i-changed-my-mind**.

Robinson K (2009). *The Element*. New York, US: Viking.

Andy Hargreaves and The Fourth Way

INTERVIEW BY JOHN GRAHAM

ANDY HARGREAVES IS ONE OF THE UK'S LEADING EDUCATIONALISTS.

IN THIS CONCLUDING PART OF AN INTERVIEW CONDUCTED BY EMAIL IN APRIL, HE SETS OUT THE DANGERS FOR AUSTRALIA OF HIGH-STAKES TESTING AND TABLES AND THE CHALLENGES TO TEACHERS AND PRINCIPALS IN PURSUING REFORM.

JG: At present the major debate about school education in Australia revolves around the publication for the first time in 2010 of individual school standardised test results on a federal government website ("My School") and the development of school league tables in the media. This development is being heralded by government as a key part of its school (and teacher) improvement strategy. What is your view?

AH: The Australian strategy is an adoption and adaptation of the Third

Way strategies of lower performing Anglo-Saxon countries such as England and parts of the United States. It's an odd choice — a bit like Cricket Australia looking for improvement tips from Bangladesh or the Netherlands! So why would Australia adopt it?

The main answer is that it is a politically plausible and manageable strategy that seems to create improvements in measurable results in literacy and numeracy, at least in the short term, by concentrating almost all effort to a relentless degree on what is tested, using heightened competition, a narrowed and almost obsessive focus, and public transparency as a way to drive up scores.

There are some benefits to this strategy. It does tend to develop a sense of urgency in schools that have cared for poor children but not really pushed them. It can provide data that enable you to find help from higher performing, similarly situated peers. And most Third Way politicians have a sincere wish to boost public confidence in public education so the middle classes will reinvest in public education personally and through their taxes once more.

The drawbacks are dramatic, though. Everywhere, and I mean everywhere, that I have seen these Third Way strategies in operation, the downsides have been the same. Many of the so-called improvements have subsequently turned out to be fake or fabricated — for instance by introducing training for teachers, offering technological assistance for special education students, and "refining" test items only after the baseline tests have been administered, rather than right at the beginning — so that the seeming improvement is actually just a recovery. There has also been a loss of creativity and innovation, of arts and even physical education that are essential for 21st century opportunities — are our private schools going to retain these opportunities for more nourishing and innovative educational fare while those left in the public system are condemned to a staple diet of bare basics?

Teachers are pressured to concentrate disproportionate attention on students just below the borderline of proficiency which gives a political boost to the system's results but provides little benefit for the students themselves and drags attention away from those who really struggle far below the threshold. Except where these strategies have followed on from even more punitive antecedents — as in Ontario or New York City — the increased top-down pressure and prescription have often depressed the status of teaching and made attraction and retention of high quality teachers (and leaders) even more difficult.

Last, while My School can be a really good idea, simplistic measures of poverty or diversity can baffle teachers when they see that their unstable itinerant population is counted as being similar to an equally poor but much more stable

community culture of, say, highly aspirant Bangladeshi families.

In the wake of all these difficulties, England, for example, has substantially backed off its Third Way strategies and abolished all but one of its tests, as well as removing many of its National Curriculum requirements in order to develop more innovation and creativity. Australia seems to be colonising the sinking sands of soulless standardisation for its poorer, publicly educated students that most other nations are leaving behind.

There are ways around these things without going back to the tie-dye Woodstock '70s. For accountability purposes, you can and should test samples, not a census — this prevents people fiddling the results which actually undermines public interest in fair accountability. Scotland, New Zealand and even Papua New Guinea as well as Finland already do this with great success. My School data can and should be confidential to the participating schools (warning though — in other jurisdictions, they started out confidential but this became eroded later). And league tables should be abolished with a public information campaign indicating why. This measure was the very first Act of Northern Ireland's Stormont Parliament.

JG: Australia has one of the highest proportions of students in private schools in the world. What are the implications of this situation for introducing the program of reforms you outline as "the Fourth Way"?

AH: I have never truly understood Australia's fascination with private schools in a society that admirably advocates a "fair go" for everyone.

I am not against private schools. Choice and diversity are important values. The Netherlands has a high degree of private choice but appears able to disconnect this from institutionalisations of advantage and elitism. Ontario has a large Catholic system but offers majority public funding for all Catholic schools (that then have to meet provincial requirements). England has specialist public secondary schools that cater for different interests like sport or visual arts. You can and should have great choice within a public system, especially at the secondary level. You can have a large private system without this being associated with cutthroat competition.

My own objections are when a private system becomes systematically tied up with privilege, unfair opportunity and entitlement by offering tax relief to users of private schools, by failing to invest adequately in the public system so the move to private education is really like an emigration from an impoverished alternative, and by comparing private and public schools on singular measures of performance when they typically serve very different populations.

JG: One of the areas you write about is the

link between community organising and community development and enhanced educational outcomes. Can you explain this connection?

AH: On average, around two thirds of the explanation for student achievement comes from factors outside the school. There are heroic exceptions that fly in the face of this general pattern, but the general pattern still holds. Obviously, therefore, one of the untapped areas for improvement is to influence the things that influence the school — for schools to work more closely together with parents, communities and the wider public outside.

For the individual teacher, this means seeing parents, even maladroit ones, as allies not adversaries. After all, they're the only parents the kids have got. For the school, it's about reaching out to parents and communities in every way possible — making the first move, and the second, and the third, on the community's terms to begin with, at least. In a community as in a classroom, you have to start with where they are and what they've got and move on together — everything from social events, home visits, talking to parents who drive by to pick up their kids, and "snuggle up and read" early years sessions.

And for policy, it's about talking up public schools and teachers, and creating more jobs in schools where people from the community can work alongside teachers for their community. We have seen that this builds trust between community members and teachers, improves student behaviour in schools, and begins to draw people from the community into teaching as a career where they often come back to teach in their community.

JG: Teachers find themselves under pressure on all sides — the blame game from governments, an accountability curriculum, the rise in administrivia, red tape etc. What do you think should be done to enhance the role and improve the quality of the professional life of teachers?

AH: High quality teachers will come and work for an inspiring leader. Good teachers need great leaders who know where they are going and know the people they are leading. They will commit themselves for a greater goal or dream when they are confident that it can, with effort, be achieved.

They need to feel their professional judgment and discretion are respected — that, like doctors, the data matter but so does their intuition; that they are professionals whose judgment is valued, not cubicle employees reduced to standardised and prescribed operations. They need to have pay that is good enough not to distract them from their work or lead them to feel undervalued. They need to feel supported with resources (like the new building program), including human resources.

They need time to meet, talk and plan, to work as a professional community, during the day as well as after it, and in the midst of their

work, not just when it's all over. They need a public and a political class that honours and respects the work they do as those who create the country's future. And they need evidence that their efforts are succeeding around measures that are relevant to what they value, not in relation to alien indicators that are imposed from afar.

JG: *The role of leadership in school improvement is crucial. You describe the dilemma of principals who are caught between being leaders responsive to their own school's needs and "compliant line managers" implementing an endless list of departmental/government initiatives. What can be done to address this dilemma, to make leadership a more attractive option for principals and teachers and to realise its full potential in school improvement processes?*

AH: Leaders in Finland say they are able to achieve a lot of what they do because they are not constantly distracted by having to respond to externally introduced government initiatives. They feel trusted by their teachers on the one hand and their superiors on the other. They regard themselves and are regarded as (barely) first among equals. Their loyalty is to the school and those it serves, not to a higher political authority.

In Finland, England and Alberta, we have seen the immense value of leaders being able to work together. Here, improvement is increasingly driven by resourcing schools so leaders (including teacher leaders) have time and incentive to connect with other schools — helping and being helped by ones that are performing differently. This doesn't mean a return to open-ended networks of earlier times in which enthusiasm was high but results were unclear. It means having networks with robust architectures where there is transparency and collective responsibility for participation and results. All this means a shift. In the Second Way, resources were moved to the schools while control shifted to the top. In the Third Way, control intensifies at the top even more. In the Fourth Way, more control and resources shift from the top to the middle, to support communities among schools where the strong can help those who struggle.

JG: *One of the important ideas you discuss in this book is how to make positive change last and spread. What do you need to do to make school improvement sustainable in a context where there are so many pressures working in the opposite direction?*

AH: First, you may feel it's going in the wrong direction in Australia now, but the public has to know in a vigorous communication campaign that higher performing countries are adopting quite different strategies, and that Australia seems to be cloning the strategies of lower performing, more weakly resourced and more socially divided systems and that this will

put it at great disadvantage in a 21st century economy.

Second, Australians need to know that many countries that have tried this approach — England, Wales and several Canadian provinces — have backed off from the very strategies that Australia is now adopting.

Third, in England, we have studies of highly successful schools and local systems that have many Fourth Way characteristics, but that are operating in Third Way systems. Through effective and assertive leadership (including leadership that will sometimes challenge the system), these schools and local systems prove they can achieve the measured results by getting students and teachers passionately engaged in learning and teaching, by creating lively professional learning communities rather than data-driven drudgery among teachers, and by making massive headway in engaging parents and the community.

I am optimistic, not pessimistic. We can accelerate the pace at which the public can be informed. The strategy has been adopted because it has a track record behind it. The public can also be made more aware more quickly, also because the strategy has a track record.

Not all of it is bad, as I have said. We must acknowledge the urgency, the value of more data, and a sense of national direction that the Third Way brings. And we should never advocate going back to some idyllic past age that older educators might mis-remember. Unions must be bold about the changes they are for, and not only voluble about what they want to defend. It is possible to be a Fourth Way school even in a Third Way system.

But other countries are headed in a different direction. Australia is adopting strategies that don't emulate the best of them. If Australia wants to be world class, why would it copy the simplistic approaches of second-class competitors? That is the question the country now has to answer.

NOTES on contributors

BRIAN CALDWELL IS managing director and principal consultant at Educational Transformations, and Professorial Fellow at the University of Melbourne where he served as Dean of Education from 1998 to 2004. This paper was presented at a national symposium organised by the AEU in association with the Australian Government Primary Principals Association and the Australian Secondary Principals Association on the theme "Advice for Ministers and ACARA on NAPLAN, the use of student data, My School and league tables", University of Technology Sydney (UTS), 23 July 2010.

JOHN GRAHAM IS a research officer at the AEU Victorian branch, with responsibility for researching curriculum and professional developments in education and training. He has written extensively about curriculum change, teachers and teaching as a profession, developments in education at an institutional, state and federal level, and on a range of other matters from funding to organisational review. John has been a teacher in Victorian government secondary schools, a researcher and writer for a national equity program, and a project manager and policy developer for the Department of Education.

ANDY HARGREAVES IS the Thomas More Brennan Chair in Education at the Lynch School of Education at Boston College; the post has a mission to promote social justice and connect theory and practice in education. His teaching and research concentrate on educational change, performing beyond expectations, sustainable leadership and the emotions of teaching. He co-founded the International Center for Educational Change at the Ontario Institute for Studies in Education and is a former Professor of Educational Leadership and Change at the University of Nottingham. He is the author or editor of more than 25 books. *The Fourth Way*, his book with Dennis Shirley, is published by Corwin.

JAMES LADWIG IS Associate Professor in the Education Research Institute at the University of Newcastle. He has designed, advised and evaluated several state and national educational reform initiatives in Australia and overseas. He designed and co-directed the Queensland Longitudinal School Reform Study, the SIPA study in NSW (Systemic Implications of Pedagogy and Achievement in NSW Public Schools) and several other nationally funded research studies relating to schooling and social inequality.

ALLAN LUKE TEACHES literacy education, policy and sociology at Queensland University of Technology. He trained as a primary teacher at Simon Fraser University, Canada, before joining the teacher education staff at James Cook University in 1984. He was deputy director general of education in Queensland (1999–2000), chief advisor to the Minister (2001–02), and established the Centre for Research on Pedagogy and Practice at Singapore's National Institute of Education (2002–05). He is currently researching Aboriginal and Torres Strait Islander school networking and reform, alternative assessment for at-risk adolescents, and transfer of digital to print capacities among working class primary students.

BARBARA PRESTON IS an independent researcher specialising in analysis of the labour markets of school teachers and registered nurses, professional regulation and education, and investigating diverse education policy matters. She worked as a research officer for the Victorian and federal teacher unions for a decade before moving to Canberra in 1991. She is carrying out further research into the systematic bias of ICSEA and similar area-based measures of socio-economic status, and may be contacted at **barbara.preston@netspeed.com.au**.

ALAN REID IS Professor of Education at the University of South Australia. He is involved in a range of national and state professional organisations. His research interests include educational policy, curriculum change, social justice and education, citizenship education and the history and politics of public education. He is also involved in policy development at the state and national levels.

MARGARET WU IS an Associate Professor in Education at the University of Melbourne. She has a background in statistics and educational measurement. Her main interests are in psychometric research. She is a co-author of an item response modelling software program, ConQuest, which is the main analysis tool used for NAPLAN and the OECD project PISA.

BACK ISSUES

The following back issues are still available, free to AEU members from aeunews@aeuvic.asn.au**. Non-members and institutions can order them for $10 each using the form overleaf. Other issues can be found online at** www.aeuvic.asn.au/pv**.**

PV 7.3: New Challenges for Leadership: A look at the shifting ground of leadership, including Andy Hargreaves on "the Fourth Way", John West-Burnham on thinking beyond the school gates and Bill Mulford on tapping our most experienced leaders.

PV 7.2: Beyond Edu-Babble: With education policy paralysed by management speak and business cliches, Guy Claxton, Howard Gardner, Robin Alexander and others attempt to cut through the cant with fresh thinking on the challenges facing education.

PV 7.1: Social Inclusion: How public education from preschool to post-compulsory can tackle social exclusion, including Tony Vinson on funding, Jill Blackmore on inclusive schools, and case studies of programs working with Somali and Indigenous families.

PV 6.2: Early years education: The second of our three issue survey looks at developments including the new early years framework, effective literacy programs and the national reform agenda, plus analyses of early intervention and phonics programs.

PV 6.1: Post-compulsory education: The first of three issues on the phases of education takes on developments in TAFE, higher education, and VET, including Simon Marginson on Howard's funding legacy for higher education; Alan Reid on the need to re-vision post-compulsory education; and Pat Forward on privatisation in TAFE.

PV 4.1: Teacher quality and quality teaching: Includes Andy Hargreaves on teaching in the knowledge society, and the implications for schools of globalisation; Leonie Rowan and Chris Bigum on the challenges of measuring quality in teaching and education; and Lyndsay Connors on the part that class still plays in education.

PV 3.3: Leadership: *PV*'s look at developments in school leadership includes Alan Reid on the need for school leaders to move beyond managing by embracing research and inquiry; Roma Burgess on the challenges facing women in becoming leaders; and collaborations between schools in the UK.

PV 3.2: Pedagogy: This issue analyses the growing focus on pedagogy; with articles by Dahle Suggitt and Peter Cole on new strategies for student improvement; Russell Tytler on how teachers can continually update their practice; the impact of IT on teaching; and how pedagogy can help support students at risk.

ABN: 44 673 398 674 112 Trenerry Crescent, PO Box 363, Abbotsford 3067
Tel: 03 9417 2822 Fax: 1300 568 078 Web: www.aeuvic.asn.au

TAX INVOICE

PV SUBSCRIPTION

Professional Voice is published tri-annually and aims to stimulate debate within the Australian education community. Covering issues of curriculum, pedagogy, leadership, educational practice and related topics, *PV* is positioned to build awareness of innovation, best practice and reform within educational settings from early childhood to TAFE, adult provision and disability centres. *PV* is funded by the AEU Victorian Branch and is available free to financial members by emailing **aeunews@aeuvic.asn.au**. Non-members and institutions can subscribe and order back issues using this form.

Name:.....

Address:.....

.....Postcode:.....

Daytime telephone number:.....

Email:.....

Please start my subscription from the current issue ❒

Please start my subscription from the next issue ❒

Please send me the next issue of *Professional Voice* ❒

Please send me issue(s)_____of *Professional Voice* ❒

Current issue $10 / back issues $10 / one year (3 issues) $25 /
overseas (airmail) +$5 per issue / All prices are GST inclusive

Total amount payable: AUD$_____(includes GST)

Please find enclosed my cheque (in Australian dollars payable to the Australian Education Union Victorian Branch) ❒

Please debit my credit card: Visa ❒ Mastercard ❒ Bankcard ❒

Card Number: . . . . . Expiry: . . . . .

Cardholder's Name: . . . . .

Signature: . . . . . Date: . . . . .

Please use a separate form for each subscription.

ISSN Number:
1445-4165

Featuring essays by

Editorial: The Trouble with My School
John Graham

Accountability and the Public Purposes of Education
Alan Reid

The Inappropriate Use of NAPLAN Data
Margaret Wu

The Systematic Bias of ICSEA
Barbara Preston

What NAPLAN Doesn't Address (But Could, and Should)
James G Ladwig

Will the Australian Curriculum Up the Intellectual Ante in Primary Classrooms?
Allan Luke

The Impact Of High-Stakes Test-Driven Accountability
Brian J. Caldwell

Andy Hargreaves on the Fourth Way
Interview by John Graham