**Translating what? How PISA distorts policy, research and practice in education**

**Professor Alan Reid, University of South Australia.**

***Address to CRED research conference at Magill Campus, University of South Australia, on Thursday November 14, 2013.***

**Introduction: Education policy and the public sphere**

In this paper I want to take the theme of this conference – Research in translation – and explore the way in which an unholy alliance of research and policy is adversely affecting one of the great institutions of Australian society – public education.

Public schools exist in every community in Australia. They are microcosms of their community in that they represent the diversity of the community; they contribute to community building; and they teach young people how to interact across difference. In short they promote the common good.

However, there is a great deal of educational research literature which demonstrates that many contemporary education policies are at odds with the idea of public education and the public good. The dominance of neo-liberal ideology in education policy-making over the past two decades has resulted in a focus on education markets and competition, the construction of parents and students as consumers, the narrowing of the curriculum, and the residualisation of public education - all of which privilege selfish individualism rather than promote community-minded citizens. There is now a move to promote the concept of ‘independent public schools’ (surely an oxymoron) and the new Federal Minister of Education, Christopher Pyne, has boasted that he wants to make public schools more like private schools.

In my view, these policies strike at the heart of the true essence of public education. The international research demonstrates that such agendas frustrate parents for whom choice is a mirage; damage a sense of school community; and, ironically, have not resulted in improved educational outcomes (e.g., Education Policy Response Group, 2012; Ravitch, 2010; Nicols & Berliner 2007; Hursh & Martina, 2005; Amrein & Berliner, 2003). Most importantly they destroy the concept of public education as a public good and for the public good (see Reid 2012 for an elaboration of this argument).

The critique of neo-liberal policy settings has been well canvassed in the research literature, and so I am not going to rehearse it here. Rather I want to ask the question: How has it come to this? Why do we continue to pursue policy directions which erode the essence of public education? And what can we do about it?

Elsewhere (Reid, 2013) I have argued that a major reason for the neo-liberal turn in education policy making has been the unhealthy reliance on data derived from tests scores, particularly international tests scores. I have shown how this data has been picked up by people outside education – politicians, journalists, economists, business people and lawyers, for example – who have begun to make pronouncements about the declining standards of education in Australia, and the strategies that are needed to address the decline. They usually ignore, even dismiss, the expertise of the education profession and peer reviewed educational research. Instead they simplistically read off the League tables of countries that are regularly produced by the OECD, and/or turn to commissioned and unrefereed consultant’s ‘research’ reports as their information base. I used an influential 2012 Grattan Institute Report, *‘Catching Up: Learning from the best school systems in Asia’*, as a case study of this process at work.

My concern is that this focus has produced a narrow and misguided view about the purposes of education, the standards of Australian education, and the policy approaches needed to lift standards. When policy makers, media commentators and some education researchers unproblematically use test scores as the base data for starting-point assumptions about education standards, they legitimate the widely held view that such data is a surrogate measure of education quality. This, in turn, has opened up the space for the neo-liberal policy regime in education, and the consequent erosion of the concept of public education as a common good.

In this paper I will suggest that this policy trend can be resisted if it can be demonstrated that it is based on rotten foundations. This means going right to the heart of the matter – the nature of the data that is being used to sustain public discourse about, and policy making in, education. In particular I will explore the most well-known of the international tests – the Programme for International Student Assessment (PISA) - looking generally at its makeup and nature, impact and influence.

My argument will be that that way in which PISA is being used in this country is having a detrimental effect on education policy and practice; and that unless we face up to this issue it will be difficult, if not impossible, to ensure that public education can continue to serve the public good. The first task is to understand what is happening.

**The PISA Phenomenon**

In the past few years there have been a number of assumptions about education standards in this country which merit, but which are rarely subjected to, critical scrutiny. How many times have you seen a media opinion piece, a government report, a think-tank education report, another politician or a letter to the editor start with the claim that standards in Australian education are declining compared to other developed countries?

Hardly a day passes in which such an assertion is not made in the public arena. It would be interesting to know the genesis of this received truth, and to track the trajectory of the ‘idea’ over time. But the salient point here is that it has become the self-evident starting point for debate, not a matter of opinion requiring evidence.

When evidence is proffered, invariably it is Australia’s performance on PISA tests. So if PISA is at the heart of the matter, it is pretty important to come to grips with the PISA phenomenon.

The Programme for International Student Assessment is an OECD administered test of the performance of students aged 15 years in Maths, Science and Reading. It has been conducted every three years since 2000, with the last results published in 2009. We are currently waiting for the 2012 results which I understand are to be released in December 2013. It purports to test ‘real-life’ skills and knowledge. About 65 countries participate in the tests which are two hours in length and hand written, and involve a stratified sample of students in each country. In 2009 Australia was ranked 9th in reading, 15th in Maths and 10th in Science.

Despite warnings on the PISA site about not using the test scores to make superficial judgments comparing the quality of education in various countries, when PISA results are published, press commentators and politicians in each country blithely ignore these warnings. The ‘winners’ are eulogised while those countries which have slipped a few rungs on the League Table are excoriated

PISA has sparked a burgeoning education industry involving educators and bureaucrats visiting ‘successful’ countries (i.e. those in the top 5 of the League Tables) and identifying the factors that have contributed to the success. These are then written up into education reports, and become the basis for new policy designed to engineer a rise up the league tables. In June 2013, the Gillard Government’s Australian Education Bill enshrined the aspiration for Australia to be in the top five schooling countries by 2025. Given that then PM Gillard made it clear that this aspiration applied to maths, science and reading, obviously PISA is the benchmark for Australian education.

**The PISA effect**

Politicians are currently waiting with bated breath for the release of the 2012. In normal circumstances we would find the Government ready to claim the credit if Australia has advanced a rung or two up the League Table; and the Opposition wanting to take a baseball bat to the government in the event we have slipped down the table. However, given the recent election of the Abbott government, these scenarios may be reversed. Certainly, the results will be used by both political parties to affirm established policy or justify new directions. Either way, what is said will confirm the public perception that PISA is a scientific and objective measure of education quality. This has a number of consequences which I will call the PISA effect.

In recent years, although Australia has stayed in the top ten in Science and Reading, we have slipped a few rungs down the ladder in these areas, and out of the top ten in Maths. This has fuelled the belief that education standards in Australia are declining, and has created a storm of media attacks on educators and policy makers. The response has been a flurry of policy activity designed to address the decline. The situation has also created a sense of educational crisis with schools and teachers bearing the brunt of criticism, with a flow-on negative impact on morale.

But the PISA effect doesn’t stop there. The claims about standards are invariably used by media commentators as a springboard for advocating a favourite policy position. Following the release of the 2009 PISA tests in December 2010 and the usual bemoaning of our perceived decline in education standards, many commentators proposed a range of disparate strategies to improve Australia’s standing in PISA tests – most of them from the neo-liberal policy armoury. They included performance pay for teachers; giving school boards the power to hire and fire teachers; revamping teacher education programs to remove any element of history, philosophy and sociology of education and to ensure a sole focus on the technical elements of teaching; intrusive accountability regimes; narrowing the curriculum; and establishing voucher systems to enable school choice.

It did not seem to trouble the commentators that there was no evidence to support a relationship between the PISA data and the solution they proffered. They simply jumped from apparent ‘problem’ to solution, ignoring important intermediate steps, such as assessing the evidence, clarifying the problem, gathering extra evidence, and making a connection between the solution and the problem. The game of proposing strategies to address the crisis has continued ever since. Watch it at work in the weeks following the release of PISA 2012.

Another version of the PISA effect involves the quest to discover why the top five countries are more successful than Australia. This consists of visiting a country, describing some of its structures, practices and processes, guessing which variables have contributed to its success, aggregating them, and then urging that these practices are adopted in Australia. A classic of this genre is the Grattan Institute’s 2012 report *Catching up: Learning from the best school systems in Asia* (see Reid, 2013).

These various PISA effects influence policy makers and ultimately help to shape education policy. Sometimes this is overt, such as using PISA as the benchmark for Australia’s educational aspirations for the next decade. At other times, its influence is more subtle, such as when PISA outcomes are used in inquiries and reviews to justify particular policy proposals; or when they are used by researchers as a surrogate for educational outcomes.

It is curious that so much store is placed on the results of a test which is conducted every three years, by different companies/groups/agencies in over sixty countries, testing sample groups of students. Clearly a lot of trust is being placed in the test itself. But what would happen if it could be shown that the results of PISA tests, or the ways in which they have been interpreted, need to be taken with a grain of salt, or are faulty? Where would that leave the research based on PISA results or all the strategies designed to address the PISA effect?

**PISA Problems**

There are a number of problems with blindly accepting the PISA outcomes. First, test results at one stage of schooling every three years in only three areas of the curriculum, are too narrow a base upon which to make claims about the quality of Australian education. The fact is that although reading, maths and science are important, they tell us nothing about outcomes in such crucial areas as the arts, history, geography, health and PE, English literature, civics and citizenship, and information and communication technology, to name just a few areas of the formal curriculum. Inevitably this emphasis results in the favoured three curriculum areas becoming the core of the curriculum, attracting most of the allocated curriculum time and resources, with other areas being relegated to the margins. In addition, such a narrow view of education means that we get no sense of how students are faring in such critical domains as problem solving, inquiry, creativity and intercultural understanding. That is, at best the international test results present a very limited picture of student progress.

An example of just how dominant this narrow view of what is valuable in education has become is the title of a recent ACER report: *Measure for Measure: A Review of Outcomes of School Education in Australia (2013).* The report analyses Australia’s performance over the past 20 years in two international tests, PISA and TIMSS, as well as how students have performed in NAPLAN. Apparently the authors assume that the areas of maths, science and reading alone can be used to pass judgment on the outcomes of Australian education. Apparently no other areas of the curriculum count in this determination, let alone such things as the quality of the relationships, the extent of community involvement, or the values practised and espoused by schools!

Second, this narrow picture becomes a real problem when the test outcomes start to become the gold standard for education quality, and when they begin to shape the education aspirations of countries. As countries seek to maintain or improve their PISA league table standings, so the official curricula of these countries start to narrow and converge. Yong Zhao (2012) describes it as a process of ‘global homogenization’ as the less successful countries begin to copy or borrow from the more successful countries. He points out that it is ironic that this homogenisation process is fuelled by global economic competitiveness, and yet actually works against the interests of many countries.

It is worth describing the reasons that Zhao gives for this claim because they add up to a powerful case against the tendency for global curriculum sameness that is produced by the race to the top of the PISA tree. Zhao observes that the traditional notion of an education system designed to develop people with similar skills to meet the demands of the local economy and common citizenry is no longer tenable in a globalising world where there is a need to embrace diversity, creativity and difference. He suggests a number of reasons for rejecting global standardisation, including:

1. At a time when there are incredible differences in labour costs, and jobs can be outsourced to many places in the world, if skill sets of people are the same, then companies will always choose the cheaper option. Surely we should be looking to differentiate our skill sets, rather than to standardise them.
2. Jobs that require routine procedural skills and knowledge will become increasingly automated. In the new economy we need skills which are as much right brain directed (design, art, play etc) as left brain cognitive skills of the sort tested by PISA.
3. The interconnectedness of the world demands such capacities as intercultural understanding etc (not just literacy and numeracy)

These and other reasons point to the need for approaches which focus on creativity, critical thinking, intercultural understanding, problem solving, entrepreneurship and so on.

A third problem with relying so heavily on PISA is that PISA test results are interpreted too simplistically. As England’s Office of Qualifications has pointed out, the differences between countries’ performances are not that large anyway and are usually statistically insignificant; and whether or not a country has moved up or down the league tables is not very meaningful because the absolute difference in scores between countries is not great. What is the real difference between 5th and 10th?

Fourth, even if it was possible to discern fine grained differences between countries on the basis of their position on the league tables, where does that leave us? How can the data be used to identify causes? There are real difficulties associated with ascribing causality to cross-sectional data, and these can only be resolved by exploring the relationship between different variables using longitudinal models with same student cohort.

More than this, there are many obvious problems with the ways in which comparative judgments are being made about the quality of education in various countries. These include the vast differences in contexts between countries, and the fact that it ignores other issues in some countries. In Singapore, for example, there is a concern that although students are successful in tests, their creativity is being stifled by a narrow and strait-jacketed curriculum; and in Shanghai there is considerable concern about the high rate of youth suicide. The point is that although it is useful to share information between countries, uncritically importing policies and practices from other countries is fraught with danger, particularly when the only reference point is an international test.

Fifth, there is now a burgeoning international literature which raises a number of serious questions about the efficacy of the PISA tests and their associated processes. These include, in no particular order, such disparate concerns as:

* the propensity of some systems to prepare students for the tests, including the extensive use of after-hours tutors in some countries. It is interesting to note that Australia seems to be joining this race with the launch this week of a number of PISA resources materials, produced by ACER and designed to introduce students to the kinds of learning skills in maths and science demanded by PISA.
* the difficulties associated with making an international test culturally neutral
* the differences in the ways in which ideas can be translated in various languages
* whether or not some countries endeavour to improve results through selective sampling
* the differences in student populations in various jurisdictions, especially given that some of the successful Asian systems are cities and others are countries
* the unexamined values and beliefs embodied in the test about what constitutes valued knowledge and about curriculum
* the definition of ‘real-life skills’ and the fact that they are only tested with pen and paper
* the narrowness of what the tests measure
* the difficulties associated with interpreting cross sectional data
* the long time period between the taking of the test and the publication of the results, thus removing any diagnostic possibilities from the test results
* the fact that the raw scores do not tell the full educational story. For example, in the 2006 science results, Finland came out on top in cognitive outcomes; and finished nearly bottom in level of student interest in science. One wonders if the results have been gained at the expense of turning students off the study of science, and if so, whether this is something we would want to emulate in Australia?

Now, I am not saying that PISA is a shoddy test – over the past 12 years the OECD has put a lot of resources into making the tests as rigorous as possible. Nor am I saying that such international tests cannot tell us anything. Rather I am making the point that policy, media commentary and research which is premised on PISA test results should at least acknowledge these difficulties and limitations and be much more tentative about using PISA as the sole arbiter of what constitutes quality in education. The use of PISA to assess quality and as the benchmark for our national educational aspirations is fraught with danger.

Instead of accepting PISA outcomes without question, it would surely be more honest to preface a description of PISA outcomes with some qualifications which point out for example that (a) only three areas of the curriculum are being tested, so we can’t generalise from them about overall education quality. Additional data is needed; (b) conducting international tests like this in in different languages is difficult and imprecise work and therefore the results should be treated with caution; (c) the results tables should be read in terms of broad bands of outcomes, not as single rungs on a league table; (d) the results should not be decontextualized from such factors as economic status; ‘race’, and gender; (e) we should be very careful about ascribing causality on the basis of the relationship between test outcomes and a particular policy approach …. and so on.

This might seem obvious, and yet the view of PISA as an objective measure of education quality has become so dominant in the public arena that it is no longer questioned. There are any number of examples I could provide to demonstrate the ubiquity of PISA as THE measure of educational quality in research, the media, public discussion and policy making. Let me provide a case study of how this plays out in practice.

**Case study: PISA in Focus**

For this case study, I will go to the home of PISA – the OECD. Each month the OECD releases a ‘research’ paper titled *‘PISA in Focus’* which purports to answer a pressing question about education policy and teaching and learning. Nearly every month, the new research insights are picked up by the media – especially when it appears that Australia is lagging behind other countries.

Now before we look at an example of such ‘research’ I need to explain what data is used. Each research question draws on the PISA test results in the three areas of Maths, Science and Reading (remember the students are 15 years of age); and the answers to questionnaires by the students who took the tests, and the Principals of their schools.

Many of the research questions refer to equity – the most visible characteristic of which is the socio-economic status of the students who take the tests and which is referred to as the Index of Economic, Social and Cultural Status (ESCS). This is determined from student responses to questions about (a) their parents’ occupations and educational attainment, and (b) their reports on the cultural possessions and educational resources available in the students’ home. One can imagine that there might be some response validity issues with this approach!

It seems that the major methodology of the research in *PISA in Focus* involves the test results being linked to/correlated with some of the student and principal questionnaire answers and cross referenced with the ESCS. The findings are then generalised across all ages, school sectors etc.

Let’s take the latest *PISA in Focus* to see how this process works. No. 32 asks the question: *Do students perform better in schools with orderly classrooms?*

This research draws on data from a question to students in the PISA 2009 process. Students were asked to respond to the overall question: ‘How often do these things happen in your language of instruction lessons?’ by ticking one of the boxes in a scale from (a) never or hardly ever; (b) in some lessons; (c) in most lessons; and (d) in all lessons for each of the following five statements:

* students don’t listen to what the teacher says
* There is noise and disorder
* The teacher has to wait a long time for the students to quieten down
* Students cannot work well
* Students don’t start working for a long time after the lesson begins.

(see Figure 1 below)

**Figure 1**



Let’s assume that all students are motivated to respond to these statements seriously and to tell the truth. Is it not possible that students from across 65 countries might interpret some of these statements differently? For example, that what constitutes ‘noise and disorder’ in one context/culture might differ from another; that for different students, a teacher ‘waiting a long time’ for quiet might vary from 10 seconds to 10 minutes; or that ‘students cannot work well’ might be interpreted by some as ‘I cannot work well’ and by others as ‘they cannot work well’, or that some interpret ‘work well’ to refer to the quality of work rather than the capacity to undertake that work; and so on.

So, right from the start you might expect the researchers to be careful with the answers. Not so. The students’ views on how conducive classrooms are to learning were combined to develop a composite index to measure the disciplinary climate in their schools. These were then converted to percentages and a league table of schools constructed under the heading ‘Order in the classroom’. This table is based on just three of the five questionnaire statements, with those with the ‘best disciplinary climate’ at the top; and the worst at the bottom (see Figure 2 below). Australia came in at 34th out of 65 countries, just above the OECD average.

The final step was to establish a relationship, for each country, between the average disciplinary climate in school and (a) the average socio-economic status of students in school; (b) student performance in reading on the PISA tests; (c) and student performance in reading after accounting for ESCS.

All of this mumbo-jumbo results in the following pearls of wisdom:

* most students enjoy orderly classrooms for their language of instruction lessons;
* Socio-economically disadvantaged students are less likely to enjoy orderly classrooms than advantaged students;
* Orderly classrooms – regardless of the school’s overall socio-economic profile – are related to better performance (PISA 32, 2013)

Now as unexceptional as these finding are, we need to remember that the data relates to a quite specific student cohort who are 15 years old of age, and that there are problems with

* the tests themselves as outlined earlier;
* the way in which the ESCS is constructed;
* Design of the questionnaire.

The fact is that although there may or may not be discipline issues in Australian schools, we can be sure that this research does not establish this either way.

However, such concerns counted for nothing when the research went public in October this year. Commentators immediately jumped on the news. In the Australian, Kevin Donnelly (9/10/13) began his piece with the following:

*Forget arguments about school funding, whether government or non-government schools achieve the strongest results, or what is the best way to reward teachers. The real problem …. is noisy and disruptive classrooms. Australian classrooms were ranked 34th out of 65 countries in a recent OECD survey … It found that Australian classrooms, compared with those in places that achieve the best results in international tests, such as South Korea, Singapore, Japan and Shanghai, are noisier and more disruptive and more time is wasted as teachers try to establish control………. If standards are to improve, especially for disadvantaged students, Australian classrooms must embrace a more disciplined environment where teachers are authority figures who engender respect.*

You can see here how the ‘research’ has been used.

1. The first port of call is the league tables. The numbers can’t be wrong – they are precise and scientific – and Australia is floundering in the middle of the pack.
2. Next is the embellishment of the data. We have noisy and disruptive classrooms in Australia with teachers battling to establish control (there is no mention of the fact that the survey information comes only from the students); there is a total absence of any sense that there might be cultural influences at play; and the research findings based on interviews with 15 year old students have been subtly widened to apply to the entire Australian education system.
3. Finally the solutions are proposed. Donnelly is now able to roll out his educational predilections – teachers must be authority figures controlling more disciplined environments and engendering respect. And this, by the way, will improve educational outcomes for disadvantaged students. How have those in the education community who have spent years grappling with the complex issues of social justice in education missed this silver bullet?

Next day, the Australian editorial (10/10/13) went further. Under the heading ‘*Tone down rowdy classrooms’,* and after bemoaning the fact that Australian schools were placed well behind high performing Asian systems as well as the Britain and the US in the ‘discipline league table’, the editorial claimed that

*A third of Australian students admitted their classes often ignored what teachers were saying and 40% described their classrooms as noisy and disorderly. One in five said conditions were so disruptive they found it difficult to work. Exacerbated by mediocre curriculum content and high truancy rates in some areas, poor outcomes are almost inevitable …..A more detailed reading suggests* ***discipline is more of a problem in state schools***(my emphasis) *than in independent ones, with Catholic schools in between… Only in Finland did high educational performance occur in a noisy setting suggesting cultural and social norms play a part in certain places ….. Australia cannot afford classroom chaos or its gloomy consequences.*

You can see how the research is starting to take on a life of its own and is being embellished with the introduction of new ‘facts’ which are not mentioned in the research itself, such as curriculum content and truancy rates. It is simply another way to berate Australian public schools: they lack discipline, classrooms are in chaos, curriculum is mediocre, and truancy is rife.

This case study is representative of a number of the problems of working with the PISA data. Not only are there difficulties with the ways in which the tests are constructed and interpreted, they also serve to remove complexity from educational issues with their veneer of precision and objectivity. More than this, they dumb down the public discourse about education.

**Conclusion**

In this paper I have raised concerns about the quality of the evidence that is being used to substantiate education policy; and about the quality of public discussion about education. I am arguing that this has made a major contribution to the contemporary policy direction which is counterproductive to education serving the public good. In particular it:

* Provokes a sense of crisis about declining standards in a system which by any measure is one of the best in the world;
* Adversely affects the morale of hard working educators;
* Leads to a misdiagnosis of the causes of educational issues and problems;
* Diverts attention from real issues such as the gap between the advantaged and disadvantaged in Australian education;
* Contributes to the wrong policy directions being set; and thus is counterproductive to enhancing quality;
* Washes out a consideration of complexity in education, by reducing everything to simplistic slogans (eg., the quality of teaching …etc, as though socio-economic status has little bearing)
* Provides justifications for politicians to reduce much needed resources to education.

So what can be done? I will conclude by suggesting some possible policy approaches which move us to a post-Pisa education world

***Post-PISA education policy***

It is important to stress that a critique of PISA is not a defensive educator’s response to adverse data. I believe strongly that we need mechanisms to assess outcomes from our education system, that our schools (and policy makers) must be accountable, and that we should always be striving to improve the quality of Australian education. What I am arguing is that superficial and knee-jerk readings of international test data are more likely to impede than to advance the quality of education in this country. We need a post-Pisa approach to quality improvement.

The starting point for a new approach is to change the existing starting point – narrow and flawed as it is. The best way to do this is to recognise and celebrate all that is so good about Australian education. I spend time in schools and constantly marvel at the creativity, care, breadth, individual attention, warmth, and excitement about learning – to name just a few qualities – that is so evident.

At the same time we need to recognise that there are a number of issues, and these need to be identified, understood and worked on rigorously over the long haul, using evidence, sharing ideas, and tracking progress. The challenge is to do this without adversely affecting all the good work that is being done, without denigrating schools and teachers using flawed evidence, and without driving our schools back to an educational stone age with homogenous policy approaches focusing on very narrow outcomes.

The problematic way in which PISA has been discussed and used in Australia, suggests some alternative strategies for a new agenda. I will outline four of them.

1. *Ensure curriculum breadth*. The Melbourne Goals of Schooling agreement has been agreed to by all states, territories and the Australian Government. It commits Australian education to a broad and comprehensive curriculum and a concern for equity in education. Why should it not also guide the nature of our accountability mechanisms? Why sign up to the Melbourne Goals of Schooling and then privilege just three aspects of the curriculum – reading, maths and science – by confining our educational aspirations to the focus of PISA?
2. *Devise broader and more comprehensive ways to assess outcomes*. Given what is now known about assessment and evaluation in education, surely we can develop more enlightened approaches to assessing education outcomes - both in Australia and internationally - than a two hour, pen and paper test held every three years in just three curriculum areas. New approaches might include some light sampling of a range of subjects and domains across a three to five year period; working with other countries to find ways to assess the development of such important attributes as critical thinking, creativity and intercultural understanding; using a range of mediums for students to demonstrate their learning; and ensuring that methods to assess outcomes reflect agreed goals, and are based on more than just one form of assessment.
3. *Be more rigorous in gathering and analysing data*. Undoubtedly, the information that is garnered from PISA has the potential to contribute to policy making. Unfortunately, as I have argued, commentators and policy makers have misused the data, by removing any of its subtleties and complexities, by jumping from problem to solution, and by making simplistic and superficial claims. If we are to track down some of the issues that the PISA results do highlight, such as the unacceptable differences in educational outcomes between students from affluent backgrounds and those who suffer educational disadvantage such as indigenous students and students from economically disadvantaged backgrounds, then more rigour is needed in our public and professional discussions. This includes asking what the data is telling us, clarifying what it is not saying, and identifying what extra information is required.
4. *Be wary of unproblematically borrowing policy.* I have not argued that we cannot learn from, or compare ourselves with, other countries. Of course it is useful to share and explore ideas and outcomes from other contexts. But we must reject approaches which assume that what has been done in one country can be simply transplanted to another. In making international comparisons we should ensure that we don’t just focus on the good bits in other countries and ignore the negatives; that there is an appreciation of the social and cultural contexts which shape specific approaches; that we try to understand the relationship between variables and not focus on a single strategy as though it exists in isolation; and that we recognise the danger of making judgements about programs in other countries from the perspective of our own values and cultural assumptions.

In a post-PISA education world, education policy would start with, and always use, its agreed goals and purposes as its reference point, including the associated accountability mechanisms. In this policy world an international test like PISA would be seen as just one piece of the evidential jigsaw, not the whole picture. This would help us to identify and understand those aspects of Australian education which need work and improvement, whilst ensuring that we maintain and celebrate the things we are doing well. Most importantly, our policy makers might stop jumping at shadows and instead reacquaint themselves with the powerful notion of education as a common good.