
The Problem of Simplification: Think-tanks, Recipes, Equity and ‘Turning around low-performing schools’

Tony Loughland Sydney University & Greg Thompson Murdoch University

Abstract

Non-government actors such as think-tanks are playing an important role in Australian policy work. As governments increasingly outsource policy work previously done by education departments and academics to these new policy actors, more think-tanks have emerged that represent a wide range of political views and ideological positions. This paper looks at the emergence of the Grattan Institute as one significant player in Australian education policy with a particular emphasis on Grattan’s report ‘Turning around low-performing schools’. Grattan exemplifies many of the facets of Barber’s ‘deliverology’, as they produce reports designed to be easily digested, simply actioned and provide reassurance that there is an answer, often through focusing on ‘what works’ recipes. ‘Turning around low-performing schools’ is a perfect example of this deliverology. However, a close analysis of the Report suggests that it contains four major problems which seriously impact its usefulness for schools and policymakers: it ignores data that may be more important in explaining the turn-around of schools, the Report is overly reliant on NAPLAN data, there are reasons to be suspicious about the evidence assembled, and finally the Report falls into a classic trap of logic – the post hoc fallacy.

Think-tanks as policy actors

Australian education policy, as in many places in the world, has shifted from defining education through social democratic principles to seeing it as a driver of economic productivity. Furthermore, Australian education policy is more influenced by global trends than perhaps any other time in its history. This global space is typified by ‘policy borrowing’ (Lingard, 2010) as nations, states and jurisdictions engage in a commensurate, and commensurable, policy space of measurement and evaluation (Rizvi & Lingard, 2010). One common feature of this commensurable policy space is certain ‘articles of faith’, such as the firm conviction that accountability testing improves education outcomes, competition between individuals, schools, and systems improves efficiency and that educational equity is essentially a problem of teaching quality. In the Australian context, this concern regarding the ‘quality’ of schooling has led to a focus on international rankings on standardised tests such as the OECD’s Program for International Student Assessment (PISA). In 2013, for example, this ‘PISA envy’ led the Australian Parliament to legislate as a national education goal to be ranked top 5 by 2025 on the PISA tests (Gillard, 2012).

However, while many criticise national governments for simplistic policy responses to complex educational realities, what is often ignored is that policy work, in particular creating, researching and agenda-setting, are not being conducted in the same ways as in the past. A feature of the global policy apparatus is the emergence and/or increased importance of non-traditional policy actors such as think-tanks in the public policy cycle. Think-tanks, as Ball (2012, p. 9) argues, exemplify new policy networks that “bring into play in the policy process new sources of authorities and indeed a new ‘market of authorities’”. The rise of think-tanks in Australia bears striking similarities to what has previously occurred in England and the US. This rise is symptomatic of shifts in government and governance “towards informal policy networks and the concomitant marginalisation of traditional ‘partners’ – local authorities, teaching unions and the civil service, and academia” (Ball & Exley,
This paper contributes to concerns about the role of think-tanks in Australian education policymaking by focusing on one ‘superstar’, the Grattan Institute, with a particular emphasis on their report *Turning around schools: it can be done* (TAS). While there has been a number of studies considering how education policy is borrowed within the Australian context (Lingard & Sellar, 2013; Mockler, 2014; Thompson & Harbaugh, 2013), not as much attention has been paid to the ways that policymaking, analysis, and agenda setting have also been reconfigured.

**The emergence of think-tanks**

There appears to be a common perception that think-tanks in Australia are associated with conservative politics. However, this is not the case. There are many think-tanks, often recently created, that represent a variety of conservative, centrist and progressive political traditions. Whatever the ideological slant of their work, the proliferation, visibility and impact of think-tanks is constituting a new ‘market of authorities’ for policy work beyond that of the traditional academic and government players. Think-tanks profit from the privatisation and outsourcing of policy work previously done by government departments and commissioned academics. The importance of think-tanks in policy work must be understood alongside the rise of the so-called evidence-based policy movement (Wiseman, 2010) and the increasing privatisation of education policy making (Mahony, Hextall & Menter, 2004). Think-tank research is usually characterised by a preference for more immediate, ‘commonsense’, ‘what works’, and politically popular interventions and contributions to policy work. If education policy is an attempt to simplify the complexity of education’s ‘wicked problems’ (Dale, 2007), then think-tank reports invariably attempt to arrive at ‘commonsense’ (and cost-effective) responses to those problems. These solutions are designed to be easily digested, simply actioned and provide reassurance that there is an answer, if only individuals, institutions and systems would follow a simple, prescribed recipe. The danger is that the strong logic of ‘what works’ recipes “privileges that which is simple and easy to measure over the more complex and untidy dimensions of this very human enterprise” of schooling (Mockler, 2011, p. 518).

One effect of this privatisation of policy work is that it circumvents the difficult relationship between academic knowledge systems (such as peer review of research findings) and education bureaucracies. The increased authority of think-tanks in policy work, legitimated in various ways by governments, mainstream media and political bureaucrats represents a distrust that many elected and appointed officials hold of public service research, development departments and university-based academic research. This relationship between new policy elites and evidenced-based policy, where evidence is defined in the narrowest of terms, has had a significant impact on the types of debates we have about education. For example, Sellar and Lingard (2014) argue than an effect of this new policy game is the absent presence of SES inequalities in policy, such that subsequent discussion and agenda setting doesn’t make mention of, nor proffer solutions to, these structural inequalities.

Think-tanks, along with various NGOs, IGOs and international corporations blur “the boundaries between state, economic and civil society” evidencing “one type of the new ‘social’, involving particular kinds of social relationships, flows and movements” (Ball, 2012, pp. 5,9). This new social necessitates a critical examination of the role of third party actors like think-tanks, or what Gunter et al. (2015, p.519) refer to as the consultocracy, “whereby non-elected consultants are replacing
political debate conducted by publicly accountable politicians” within networked organisations interwoven in the new social, economic and political articulations, authorities and relationships. One tool of the consultocracy is ‘deliverology’. Deliverology was coined by Michael Barber, one of New Labour’s policy advisors who is now Chief Education Advisor for Pearson. This awkward neologism refers to the consultocracy’s role in producing policy knowledge in a timely manner for the government that allows it to achieve ‘small wins’ within the electoral cycle (Barber, Moffit, & Kihn, 2010). The narrow electoral cycle militates against the type of careful and thorough policy making that is required to tackle structural inequalities and issues within the time frame. This has specific implications in Australia where the Federal electoral cycle is only three years which places extreme pressure on first term governments (such as the Rudd/Gillard Labor government of 2007-2010 and the Abbott Coalition government of 2013-2015) to achieve small wins in policy areas like education. Governments turn to new policy actors like think-tanks, and their philosophy of deliverology, to provide policy solutions that are timely, achievable and politically expedient rather than necessarily educationally desirable. As a sidenote, the politics of deliverology explains to some extent the challenges facing the Gonski review of school funding in 2010 and its derailment because the educational and political benefits of structural change to Australian school funding would not be realised until many years later.

Deliverology is particularly evident in the reconstitution of matters of equity and education, as education equity is “melted into the domain of economic governance” where “the same market technologies that have been widely critiqued in Australian and international research for exacerbating inequalities” are expected to provide the solutions (Savage, 2013, p.187). In practical terms, this is evident in the focus on the ‘heroic’ role of the quality teacher and the effective school in addressing educational disadvantage, to the exclusion of policy that addresses structural issues such as poverty and school funding on educational disadvantage. Of course, this is not to claim that schools and individuals should not aspire to provide rich teaching and learning experiences, rather it is to stress that learning experiences are always impacted by the complex social realities of advantage and disadvantage. This “treatment of ‘equity’ and ‘excellence’ underscore[s] a particular understanding of equity – not as a collective, systemic obligation, but as an individualised and localised competitive pursuit” (Loughland & Srirakash, 2014, p. 11).

A critical view of the process is that the consultocracy does the knowledge work that the government demands. Of course this need not be an overt request. Given the transparent ideological positions of think-tanks that they use to market themselves as different, the choice by governments and private organisations as to which think-tanks to select for specific research tasks is made with some advance sense as to what their findings will be. Medvetz (2008) argues this imperative to deliver ideologically consistent, politically expedient and popular solutions is because think-tanks are essentially “constitutively hybrid organisations” operating across multiple fields of influence. Think-tanks work hard to present themselves as independent of, but proximally located within, academic fields, political fields, media fields and economic fields. They must also work to accrue “the very forms of authority that derive from an association with these institutions” (Medvetz, 2008, p. 6). The problem with these negotiations is that think-tanks inevitably find themselves divided and conflicted by competing academic, political, economic, and media imperatives so that
...the space of think-tanks is a world divided against itself... the considerations of academic rigor and exactitude will run up against a set of powerful limitations, the main ones being the need to maximize political access and immediate policy relevance, the need to sustain the budget, and the need to garner continuous publicity. Each of these demands exacts a particular cost on the think-tank’s intellectual production, curtailing its ability to do long-term research, to gather background knowledge, to incubate ideas – in a word, to think (Medvetz, 2008, p. 10).

One result of these competing demands is that careful, precise and methodologically sound research work is not always possible and, indeed, may not be desirable. This is because, first, each field of influence competes within each report. For example, careful academic research takes time and this is always under pressure by the need for the think-tank to contribute to issues of the political and media moment. Second, each individual report functions as a link in the chain, the more think-tanks generate media and political traction for their research, the more opportunity in the future for further research, political and media opportunities to further contribute to policy debates. This explains the tendency for think-tank research to ignore context, complexity and nuance in favour of recipes and uplifting vignettes because these are media friendly, politically expedient and conceal ideological assumptions within commonsense narratives.

The Rise of the Grattan Institute

Within this milieu that sees the rise of new policy actors and networks, and intense competition between these new actors for market share, there are some that have been highly successful. One of these is the Grattan Institute, whose uses its website to describes itself as an “independent think-tank dedicated to developing high quality public policy for Australia’s future”. It was formed in 2008 “in response to a widespread view in government and business that Australia needed a non-partisan think-tank providing independent, rigorous and practical solutions to some of the country’s most pressing problems”. Grattan applies its “independent, rigorous and practical” insight to seven specific policy programs: Australian Perspectives, Cities, Energy, Health, Higher Education, Productivity Growth and School Education. Grattan was the recipient at its inception of a $15 million investment by the Rudd Federal Labor Government, matched by further $15 million invested by the Victorian State Government. The University of Melbourne supplies in-kind support each year valued at $350,000 in Grattan’s financial reports. Since 2008 there has been a yearly donation of $200,000 by the National Australia Bank. Grattan is governed by a Board of Directors who are “leaders in business, academia, government and the law”.

From 2010-2014, the Schools Education Program led by the ex-OECD economist Dr Ben Jensen\(^1\), produced a series of research reports, media opinion pieces and participated in radio, television and newspaper interviews. While it is not possible in this article to analyse each contribution in detail, as Ashenden (2013) points out, Grattan’s policy work is informed by the school effectiveness tradition,

\(^1\) Since leaving the Grattan Institute, Dr Jensen has created another consultant organisation Learning First, whose website states that it: “produces informed, extensively researched reports on key education policy issues that provide effective solutions to improve student learning. Many of these reports provide a toolkit for schools and policy makers to help guide reforms. The reports gather the lessons from the world’s best-performing education systems and the research on effective practices to build an evidence base governments, policy makers, school leaders and teachers can use to implement changes to improve student learning”. 
typified by a belief in (quantifiable) evidence of student achievement, and ‘what works’ solutions. In particular, Grattan’s policy has emphasised the importance of education for economic productivity, the cost to the economy of declining student achievement on international tests, how teacher quality and effectiveness impacts student achievement and the need for urgent reform to improve the quality of applicants to the teaching profession; improve the quality of initial education and training; develop teachers’ skills once they enter the profession and are working in our schools; and promote, recognise and retain effective teachers and move on ineffective teachers who have been unable to increase their effectiveness through development programs (Jensen, 2010, p.10).

Grattan has been a highly proficient media presence commentating on education policy debates, often producing reports that compare Australia’s education service and delivery to ‘high-performing’ Asian neighbours (see for example Jensen, Hunter, Sonnemann, & Cooper, 2014). One key example was the media and political traction Grattan garnered for their argument against decreasing class sizes because “we should not follow low-performing systems around the world that have tried to improve schools by decreasing class sizes”; instead, Australian policy should focus on making “time for programs that develop teacher skills and deliver great teaching” such as those found in Asian countries (Jensen, Hunter, Sonnemann, & Cooper, 2014, p.3). Grattan is also critical of the idea that increased funding is the answer to improved student outcomes, arguing that while “funding matters, particularly for some schools... it will never be the policy instrument that significantly lifts our school systems” (Jensen, 2013, n.p.). The ‘deliverology’ of Grattan proposes ostensibly appealing solutions to powerful interests such as shifting “the debate from school funding to a focus on school leadership and quality teaching” (Jensen, 2013, n.p.).

The Grattan reports contain powerful discourses about Australian schools and schooling. This first of these is that there are a number of crises confronting Australian education requiring reform. The evidence for this crisis is often international test results, particularly PISA, which shows that Australian students are not performing as well as they did in 2000, and are falling behind high achieving school systems in places like East Asia. Secondly, the biggest cause of this decline, based on comparative claims on test data, is a crisis of teacher quality. While there are many great teachers,

...research consistently shows that quality teachers are the most significant influence on student performance. With an excellent teacher, a student can achieve in half a year what would take a full year with a less effective teacher. And the impact is cumulative: students with effective teachers for several years in a row out-perform students with poor teachers by as much as 50 percentile points over three years (Jensen, 2010, p. 4).

There are a number of clarifying clauses missed that might make this statement somewhat more valid. The claim should include the caveat on ‘on testing data’ – the absence of which suggests a belief that tests, and the results generated, are accurate, objective, comparable measures of the quality and purposes of education that each student receives. Also, while teachers have a significant impact on student achievement, social variables beyond the control of the school have a far greater impact (Hattie, 2008, pp. x-xi; Snook, et al., 2009, p.98). For example a recent report from the American Statistical Association suggests that teachers only account for between 1-14% of the
variance on student test scores and that “the majority of opportunities for quality improvement are found in the system-level conditions” (American Statistical Association, 2014).

However, the elision of these structural problems for educational equity has led Grattan to propose a number of solutions to the problem of teacher quality, such as improving teacher evaluation methods, tying teacher pay promotion to evaluation, creating more time for teachers to engage with, and undertake their own, research through increasing class sizes, using value added measures of student progress (VAM) as a way to make schools accountable for their student achievement and improving the use of testing data. While these solutions may be politically appealing it is doubtful that, without addressing these structural issues through policy interventions, the desired improvement in student achievement will be either achieved or sustained.

A further key assumption that flows through Grattan’s work is that schooling largely serves an economic, rather than a democratic or social justice, purpose. This leads to some curious statements, such as the claim: “Improving teacher effectiveness also has substantial economic benefits for individuals. Young people who stay in school and invest in further education can expect to earn an additional 8-10% per year for each additional year of education they undertake” (Jensen, 2010, p. 4). The claim that “accurate measures of school performance are vital to improvement” also assumes that the performances being measured by tests like PISA are capable of objective categorisations that bear comparison both between and within countries. However, the reduction of learning to performance on standardised tests is conceptually problematic for a variety of reasons. Firstly, it ignores the breadth of learning research that suggests learning is more diverse than standardised tests can cover, secondly it assumes that the tests are valid and not liable to distortion by a variety of social pressures (such as the accountability uses that Grattan advocate for) and thirdly it reduces the complexity of the aims and purposes, and therefore what counts as quality, to a series of tests that entail approximately 40 questions in a variety of domains. These assumptions are evident in Grattan’s 2014 report that suggested simple strategies that low performing schools could implement to turn around low student achievement.

Turning Around Schools (TAS)

The TAS Report, published by Grattan in February 2014, builds an argument that “low-performing schools” can be turned around by policy changes that follow five steps “to change the behaviour and practices of leaders, teachers and students” (Jensen & Sonnemann, 2014, p. 1). From the outset, we must admit to being somewhat troubled by the use of the language of the ‘turn-around’, because of Lipman’s (2011) analysis of turn-around schools in the US as a policy ‘trojan horse’ that has seen for-profit companies take-over aspects of public education in Chicago and New Orleans. Grattan’s Report argues that “Defining the size of the problem in Australia depends on how you define a ‘low-performing’ school... These questions are complex and this report does not seek to develop strict measurement criteria for school turnaround policies” (Jensen & Sonnemann, 2014, p. 1). Instead, Grattan uses a simple reading of Year 9 NAPLAN results – a low-performing school is a school where NAPLAN results are two years below the national average, and this two-year gap has persisted for the past five years. No mention is made of what defines a low-performing primary school in their methodology (Jensen & Sonnemann, 2014, p.3). Using the above criteria, the Report identified 156 Australian secondary schools (about 6%) that satisfied the criteria. The Report outlines a rationale for identifying these schools, the 156 high schools identified is “a small enough number for policy to
intensively and effectively target a narrow group of schools for turnaround” (Jensen & Sonnemann, 2014, p. 3). The argument is that targeting these ‘low-performing’ schools in order to boost their performance will improve Australia’s declining literacy and numeracy rankings in comparison to other jurisdictions on PISA. The assumed goal, that we should want to compete on PISA, that we are in an education race and that these rankings should be how we understand the purposes of schooling, is troubling.

The TAS Report explicitly frames the equity issue in Australian schools as occurring at the school level. This framing attributes inequality to “behaviour problems, low expectations and poor staff morale” (Jensen & Sonnemann, 2014, p.1). The turn-around solution is also framed at the school level and is unashamedly behaviourist, for example, “A focus on continually changing behaviours in schools and classroom is a marked change from many Australian education policies” (Jensen & Sonnemann, 2014, p.1). On the strength of this five point plan, the authors of the TAS Report are prepared to argue that “it will make a huge dent in inequality and enrich the lives of the students who need it most” (Jensen & Sonnemann, 2014, p.1). That is a big claim to make, given the scale of the problem, but it is more instructive for what it reveals about how the Grattan Institute frames the equity issue in Australian education as a school malaise that is best treated at the school level by quality, and possibly heroic, individuals.

The rationale for the TAS Report resides in the claim that Australia has a universal equity problem in education that is amenable to the simple policy prescription to improve quality. This problem is presented in unfavourable comparisons to the East Asian PISA elite (Jensen & Sonnemann, 2014, p.19). However, the basis for this claim is simplistic, when the PISA test data itself are examined in detail, rather than through the simple representation of national, aggregated averages, the picture that emerges is far more complex:

Comparing Australian jurisdiction results with those of the East Asian PISA elite, ACT is ranked fifth internationally by mean score. So a part of Australia is already in the ‘PISA top five’! Indeed, ACT’s performance is not statistically significantly different from second-ranking Korea. WA is ranked eighth internationally by mean score, and it is also not significantly different to Korea’s performance. On the other hand, TAS and NT both have mean scores significantly below the OECD mean, with rankings close to those of Greece and Spain. With such a diverse range of mean scores between Australian jurisdictions, a focus on Australia’s international ranking is not a very useful way to assess Australian education systems (Gorur & Wu, 2014, p. 8).

The data also reveals that the differences within jurisdictions are more significant. For example, In New South Wales and the Northern Territory “the variation between schools is larger than on average across the OECD, while the amount of variation within schools is still higher than the OECD average” (Thomson et al., 2013, p.314). NSW also has the largest gap between advantaged and disadvantaged schools: “111 score points, or more than 3 years of schooling” (Thomson et al., 2013, p.214). These data suggest that educational policy in Australia should pay attention to what is happening in each of these jurisdictions and this might then explain this disparity in educational outcomes. Of course, this is not a politically expedient solution, as it does not fit with the “use of moral panics by Australian politicians in the shaping of public discourse” central to a deliverology aimed at reframing of problems of equity with problems of quality (Mockler, 2014, p. 3).
Notwithstanding these problems, after identifying the 156 high schools based on patterns of low NAPLAN results, the Report identifies 4 case study schools who have managed to ‘turn it around’. The Report suggests that there is a causal relationship between 5 factors and the improved results at these schools. These are “strong leadership that raises expectations; effective teaching with teachers learning from each other; development and measurement of student learning; development of a positive school culture; and engagement of parents and the community” (Jensen & Sonnemann, 2014, p. 1). In the Report, these 5 factors are supported by a variety of reports, policy briefs and syntheses published by groups such as the US Department of Education, the OECD, Social Ventures Australia and consultancy companies and think-tanks such as McKinsey, Mass Insight, and the Institute of Education Sciences. As a sidenote, we find it troubling that so much emphasis is still attached to the McKinsey Report (2007), particularly given the excoriating analysis provided by Coffield (2012).

The five TAS ‘turn-around’ factors are exemplified by four case studies. How these schools were selected as case sites is not made clear in the Report. The Report builds a case that there are 156 low performing high schools in Australia that “are two years (or more) behind the national average on reading and numeracy in the Year 9 National Assessment Program Literacy and Numeracy (NAPLAN) tests. Further, the two-year gap has persisted for over five years” (Jensen & Sonnemann, 2014, p.3). However, two of the case sites are primary schools, as they would not sit Year 9 NAPLAN tests it is not clear on how the criteria for selection works for primary schools. Methodologically, there is no evidence as to how the data collected from each case site was collected, how it was analysed and what methodological design was constructed to improve the trustworthiness of the qualitative research. Readers remain unclear as to whether the empirical warrant of the Report is constituted by more than a one-off phone conversation with 4 school principals as these appear to be the only sources that would constitute fieldwork quoted in the Report.

The extent to which any case can, and should be, generalised is always dependent upon a number of factors, which include the strength of the methodology used, the understanding of the wider literature and what we may call a healthy scepticism, a qualitative version of the statisticians null hypothesis if you will. While there may be theoretical flaws that we can debate, a report of ‘what works’ stands and falls on the data that it assembles and uses. In this paper we aim to challenge the Grattan Report primarily for its use of data, that is, to engage with it on its own terms as providing evidence that asserts a causal relationship between these five schools as exemplars of a research recipe that other schools should follow. We will do this based on 4 problems: firstly that the Report ignores data that may be more important in explaining the turn-around of those schools, secondly the Report is overly reliant on test data, thirdly we have reason to be suspicious about the empirical warrant of the evidence assembled, and fourthly that the Report falls into a classic trap of logic – the post hoc fallacy.

**Problem 1: School funding**

Throughout its body of work, a continued theme that comes out of the Grattan Institute is ambivalence towards the claim that one of the problems confronting low performing schools is that of an inequitable funding mechanism. The Grattan line has been that while

...many have jumped on the latest PISA results as proof that a particular school sector is superior to others or that more money should go to a specific sector. There is no evidence to
support these claims. It is the same tired old game. Education Minister Christopher Pyne is right to try to shift the debate from school funding to a focus on school leadership and quality teaching (Jensen, 2013, n.p.).

Ignoring the fact that the Gonski review, to which Dr Jensen alludes, proposed a sector-blind needs-based funding model, as Table 1 shows, the four schools highlighted as ‘turning it around’ all received extra funding as part of the previous Labor government’s National Partnership2 scheme. The scheme, subsequently abolished by the LNP government, represented a significant redistribution of funding, in the case of the Low Socio-economic Status School Communities funding, each school received up to $800 per year per student.

<table>
<thead>
<tr>
<th>School</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellenbrook Primary School</td>
<td>Received funding from the Australian Government and the WA Government under the Literacy and Numeracy National Partnership4.</td>
</tr>
<tr>
<td>Ravenswood Heights</td>
<td>Received funding from the Australian Government and the TAS Government under the Literacy and Numeracy National Partnership and the Low Socio-economic Status School Communities National Partnership5.</td>
</tr>
<tr>
<td>Sunshine College</td>
<td>Received funding from the Australian Government and the VIC Government under the Low Socio-economic Status School Communities National Partnership.</td>
</tr>
<tr>
<td>Holroyd High School</td>
<td>Received funding from the Australian Government and the NSW Government under the Low Socio-economic Status School Communities National Partnership.</td>
</tr>
</tbody>
</table>

Table 1: Case schools and the National Partnership funding received

We note the Australian National Audit Office Report from 2012 which found that overall there were problems with the National Partnerships program and that they had not made a “statistically significant improvement, in any state, on the average NAPLAN results of schools that received LNNP funding, when compared to schools that did not receive funding” (Australian National Audit Office,

2 National Partnerships were one of the first intergovernmental agreements between Federal and State governments to improve literacy and numeracy. From 2008–09 to 2011–12 $540 million of Australian Government funding was devoted to supporting programs to improve literacy and numeracy. For more details visit [http://smarterschools.gov.au/discover-participating-schools](http://smarterschools.gov.au/discover-participating-schools)

3 In the Report, Grattan uses the actual names of schools and school personnel. Obviously the individuals, schools and relevant authorities must have given permission for identification, but given it is a general ethical guideline that research participants remain anonymous, we think this unfairly sets up participants for criticism. However, given that the identification is already in the public record, we thought it made no sense to use pseudonyms. We want to stress that our critique is of the Report, not the schools and individuals named.

4 Literacy and Numeracy Partnerships provided funding (which came from both the State and Federal Governments) for schools to select from a range of strategies designed to achieve improved student outcomes in literacy and numeracy. These included funding for specialist literacy and numeracy specialist teachers to support classroom teachers, access to phonics-based programs such as MultiLit and the requirement that all participating schools started on-entry diagnostic testing in pre-primary. For example, the WA guidelines can be found here [http://det.wa.edu.au/partnershipschools/detcms/navigation/literacy-and-numeracy/projects/](http://det.wa.edu.au/partnershipschools/detcms/navigation/literacy-and-numeracy/projects/)

5 The Low Socio-economic Status School Communities National Partnership funding was “designed to transform the way that schooling takes place in participating schools to address the complex and interconnected challenges facing students in disadvantaged communities, and improve student engagement, educational attainment and well-being in participating schools and make inroads to entrenched disadvantage” [http://det.wa.edu.au/partnershipschools/detcms/navigation/low-ses-school-communities/overview/](http://det.wa.edu.au/partnershipschools/detcms/navigation/low-ses-school-communities/overview/)
2012). However, it is likely that this is based on a spread of impacts, positive impact at some sites, minimal impact at others, and potentially a negative impact at others. The importance of case study work is to provide deep analysis of how things work in specific sites. There seems to be a *prima facie* case that this funding would have some impact at each of these sites, and could potentially be the most important intervention and indeed, impact on the other 5 factors. Not examining this possibility limits the explanatory power of the other 5 factors that Grattan argues led to these schools turning around.

**Problem 2: Using NAPLAN as a comparative measure**

There is widespread criticism of using NAPLAN data as a comparative measure of quality between schools. Partly, this criticism can be attributed to widespread concern over the statistical reliability of the tests at the school level as a comparison tool, because “the over-emphasis on student level and school level NAPLAN results should be discouraged due to the large margins of error associated with these results” (Wu, 2010, p. 41). Furthermore, Wu argues that the statistical reliability is often ignored, and as a result people “are likely to misinterpret the results and make inferences that are not justified” (2010, p.41), for example, improvements or decline might simply be within the margin of error. Other concerns address the consequences of using the data as a tool for comparison, that can lead to ‘gaming’ the data (Thompson & Cook, 2014), excessive test preparation and problems of validity of the assessment (Thompson, 2013). Despite these reservations, the Report has a significant reliance on NAPLAN data as its comparative measure of quality; firstly in identifying low performing schools, and secondly in highlighting those schools that have turned it around. Table 2 outlines how NAPLAN data is used as a comparative tool for each school in the Report:

<table>
<thead>
<tr>
<th>School</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellenbrook Primary School</td>
<td>“In 2008, student results were below, often substantially, the NAPLAN national average in all areas of literacy and numeracy. By 2012, the situation had changed dramatically” (p.8). “In numeracy, students are progressing at a much faster rate at Ellenbrook than other schools across the country” (p.8).</td>
</tr>
<tr>
<td>Ravenswood Heights</td>
<td>“The school’s reform efforts over the past four years are now starting to be seen in student results. In 2013, Year 3 NAPLAN numeracy scores were - for the first time - above the Tasmanian average and equal to the Australian average score” (p.11). “The school is adding real value to student learning. In 2012, student results rose much faster between Year 3 and Year 5 compared to other students with the same initial low scores in Year 3” (p.11).</td>
</tr>
<tr>
<td>Sunshine College</td>
<td>“Improved student performance has been sustained over time. For the last two years students progressed at a faster rate in literacy and numeracy compared to other schools across the country” (p.13).</td>
</tr>
</tbody>
</table>
| Holroyd High School      | “The high number of recent immigrant students at Holroyd High School means it is difficult to compare growth in NAPLAN results over time. Many students (44 per cent in 2012) have been in the country for less than three years, so their progress at the school
cannot easily be tracked. However, for those whose results are comparable over time, the school is adding significant value. These students improve at a faster rate than the average state student between Year 7 to Year 9 in both literacy and numeracy” (p.16).

Table 2: A summary of how NAPLAN data is used to characterise each school

Problem 3: Cherry-picking evidence

One of the problems for the Report concerns how it typifies each school as turning around, and the inconsistent application of data. Consider Holroyd High School vs Ellenbrook Primary School. For Ellenbrook, the evidence for the fact that this was a school that used “strong leadership that raises expectations; effective teaching with teachers learning from each other; development and measurement of student learning; development of a positive school culture; and engagement of parents and the community” to drive improvement rested solely on the NAPLAN results. The My School website shows Holroyd’s NAPLAN results have fluctuated widely, generally showing improvement in 2011, a significant decline in 2012, and an improvement in 2013. Using Grattan’s logic, this would seem to support the argument that there is significant doubt as to the reliability of the indicator, assuming that the principal’s strong leadership, given that she has worked at the school since 1995, was also being displayed in 2012, it seems doubtful that the trend that emerges is anything more than a cohort effect. However, the Report glosses over what is effectively a problem of method to emphasise the post-school destinations of its students. In other words, is NAPLAN a measure of quality only when it suits the Grattan?

For Ellenbrook, the picture is even more confusing. There is no other evidence offered that Ellenbrook is a turn-around school except for NAPLAN results. Unlike Holroyd attendance statistics are not offered although these are readily available. Instead, Ellenbrook is praised for improved NAPLAN results, due to a principal who “has high expectations of his students and the teaching that enables them to perform at their best” (Jensen & Sonnemann, 2014, p.9), the introduction of “explicit instruction methods (a highly structured teaching approach) alongside a greater focus on literacy and numeracy” (Jensen & Sonnemann, 2014, p.9) and a variety of strategies that aim to improve school-community relations such as each morning the school administration “stand at the school gate to greet students and their parents, and to talk about their children’s progress” (p.9). The Report makes a causal link between improved student outcomes and these interventions.

The Report published the following graph based on Year 3 reading scores at Ellenbrook.
Table 3: Grattan’s analysis of Ellenbrook’s Year 3 Reading results on NAPLAN 2008-2012

The results do show remarkable improvement but we stress that this improvement coincides with the National Partnership funding, curiously missing from the turning it around recipe. More worryingly is the Grattan Report neglects to give all of the evidence available as to the reliability of the NAPLAN data as a tool for comparison. Put simply, as the average Year 3 Reading score at Ellenbrook has increased, participation in the test has declined. The My School website, which provided the data for the Report, also published the following participation data on the same page:

Table 4: Year 3 Reading Participation Rates on NAPLAN 2008-2013
(NOTE: The scale has been adjusted to correspond with Table 3 above)
In other words, there is a relationship between decreasing participation on NAPLAN and improving results. While it may be easy to jump to the conclusion that the school is somewhat complicit in the manipulation of the data, we have gathered no data to support that conclusion. We simply note that there is a correlation between the increased average student achievement from 2008-2012 and decreased participation in the 2012 tests. This pattern continued into 2013. Thus, the 21% non-participation rate in Year 3 Reading in 2012, made up by 20% withdrawn students and 1% absent students, is significantly higher than the national average of 5% in Year 3 Reading in 2012, made up of 3% absent and 2% withdrawn students. It could be that the 21% non-participation in 2012 represent a cross-section of student achievement, perhaps influenced by national movements such as the ‘opt out’ movement. It could also mean that the 21% represent a disproportionate number of high achievers, in which case the Ellenbrook achievement may be even more remarkable. However, international accounts suggest that it is most likely that the majority of students withdrawn from tests tend to be those whose results are below average (Berliner, 2011; Nichols & Berliner, 2007).

Furthermore, as the My School website data shows, the increase in the student NAPLAN results from 2008-2012 expressed in the Grattan Report correlates with changing demographic and enrolment patterns at the school, from an ICSEA of 966 in 2008 (below the national average of 1000) to 1005 in 2013 (slightly above the national average). This is highlighted by a shift in distribution of the Year 3 student cohort; for example, in 2010 41% of students were in the lowest quartile (bottom 25%) based on SES, while by 2012 this had decreased to 16%. This is a significant omission, particularly when peer reviewed, quantitative analysis of NAPLAN data has found that the “ICSEA variable was the main explanatory variable across all school sectors and years of study” in explaining differences in student achievement (Miller & Voon, 2012, p. 163). Our point, and we must stress this, is that as researchers Grattan should have identified that there were problems with the data. Our analysis is in comparing Ellenbrook’s NAPLAN Year 3 Reading results with other schools we are more likely to be comparing participation rates and demographic patterns, rather than literacy achievement, particularly given the national average participation rates during the period have remained between 97% to 95%. We must be suspicious about drawing inferences on school improvement from these data.

**Problem 4: Methodological issues of comparison and the ‘smoothing’ of difference**

A continued challenge for comparative studies in education is that of context. To an extent, the challenge is always to identify the differences that make the difference, with the continued caveat that the smaller the focus of the inquiry, the more likely we are to be subject to the problems of generalisability and the impact of context. As Gorur and Wu (2014, p.4) have argued in response to Grattan’s 2012 Report *Catching up: learning from the best school systems in East Asia*: “Rushing off to observe the practices of high-performing systems and then concluding that these practices are the reason for their success can lead to erroneous conclusions; the same practices could well be prevalent in low-performing systems as well”. A similar criticism can be applied to the TAS logic, while the five ‘turn-around’ factors identified appear highly desirable, no attempt is made to identify schools whose data patterns have not turned around, but who also show “strong leadership that raises expectations; effective teaching with teachers learning from each other; development and measurement of student learning; development of a positive school culture; and engagement of parents and the community”.
This may be attributed, in some part, to the tendency for economists when looking at education to see quantitative performance data generated as objective measures of educational value. Biesta (2007) argues the focus on outputs and test data as measures of quality promotes a “cookbook” approach to educational interventions that fails to recognise the complexity of educational systems, endeavours and practices. This is a form of epistemological reductionism, that cannot account for the multiple purposes of education, the possibility of intelligent problem-solving by professionals in specific contexts, and the limitations that deciding what works best imposes on identifying problems and their solutions. Cookbooks or recipe approaches to research “can tell us what worked but cannot tell us what works” (Biesta, 2007, p.16).

One way of thinking of the problem of this style of comparison is the post hoc ergo propter hoc (Latin for ‘after this therefore because of this’) fallacy. Basically, the post hoc fallacy occurs when an assumption is made based on the sequence of events – so the assumption is made that because one thing occurs after another, it must be caused by it. While it may appear desirable, and probably politically expedient, to draw up a list of ‘commonsense’ characteristics that emphasise individual and collective responsibility for turning around low performing schools, this overly simplifies the complexity of forces at play within schools, and ignores other possibly more significant factors that impact on these vignettes. Furthermore, the over-generalisation of these specific case sites, a problem identified in much of the introductory case study research methodology, elides the significance of context. Even if the (disputed) evidence built a stronger case for this recipe the Report, because it uses data that elides context, specificity and histories (both individual and institutional), may actually result in schools in different contexts wasting their attempts to reconfigure and reconceptualise what they do. This is because the differences that make the difference may be a) context dependent or b) not recognisable due to the data collected. If you really want to know how schools operate, listening to teachers and students will present richer data than a series of standardised tests administered infrequently because “NAPLAN and PISA... obscure complex social, cultural, economic and historical factors, which differentially position young people in relation to education and achievement in the first place” (Lingard, Sellar & Savage, 2014, p.16).

Discussion/Conclusion

How societies educate their least advantaged is an important aspect of the democratic purposes of education. Equally, this extends to important skills such as literacy and numeracy in post-Fordist economies. We see this as the motivating idea behind Grattan’s TAS Report, and indeed much of their policy work around schools. In these debates, Grattan has made contributes to debates about the aims and purposes of education in Australia. That said, because it is so important to consider how policy work frames and responds to problems of inequity, we feel that critique of Grattan’s deliverology is essential if we are ever to move beyond simplistic debates and simplified solutions. Our analysis is that there are problems with the simple policy prescriptions offered by Grattan’s TAS Report.

Furthermore, while we agree that “strong leadership that raises expectations; effective teaching with teachers learning from each other; development and measurement of student learning; development of a positive school culture; and engagement of parents and the community” are things that all schools should be doing, in our experience schools already devote a significant part of their planning, development and day-to-day practice to developing each of these areas. However,
we doubt that these are enough to combat systemic, structural problems that impact student outcomes in Australia. The TAS Report makes exaggerated claims on the efficacy of their five turn-around factors on educational disadvantage, given their meagre, and as we have shown, less than rigorous, basis in their data.

Partly, the problems with the simple solutions espoused represent problems associated with a ‘what works’ approach to education and reform. This analysis has identified problems around the causal claims of the Report. The first of these is the logical fallacy known as the post hoc. Even if we accept that the 5 factors were the reasons that these schools had ‘turned it around’, it does not follow that these 5 factors would produce the same effect in different school contexts. Perhaps more importantly, there is an implied causality between what schools do and the measures that are used to portray those schools. Biesta (2007) sums up our critique of such approaches by stressing the noncausal/complex nature of education as a symbolic or symbolically mediated interaction, which a simple focus on test results as ‘the story’, can never adequately represent.

The simple—and by now we should actually say: simplistic—idea is to assume that interventions are causes and results effects and that, under optimal conditions, the causes will necessarily generate the effects. This is a kind of ‘magic bullet’ notion of causality which, if possible at all in the social domain, actually only exists under very special conditions (Biesta, 2010, p. 496).

So the question of what works, must always be tempered by the concerns of how it works, in which contexts, with some tentative sense of why it may work. TAS does little to address these questions, and as a result, has ended up giving a superficial treatment of the complexity of schools and learning that will do little to assist policymakers in formulating responses that make a positive difference to schools working in our least advantaged communities.

Medvetz’ (2008) conclusion that policy think-tanks, because of the pressing need for deliverology and the multiple fields that they seek to mediate, remain divided against themselves and this is most obvious in the lack of academic rigour of the work produced. Our concern remains that, while the recipe of turning around low-performing schools may appeal in its simplicity, it promotes a view that those schools who may not manifest better NAPLAN data, either as aggregated or value-added scores, are not already demonstrating “strong leadership that raises expectations; effective teaching with teachers learning from each other; development and measurement of student learning; development of a positive school culture; and engagement of parents and the community” (Jensen & Sonnemann, 2014, p.7). Furthermore, this narrative of the heroic individual whose commitment, perseverance and implementation of evidence-based practice, elides the role of the state, governments and appropriate policy levers in combatting those entrenched, structural factors are more significant in explaining student achievement than teacher quality approaches seek to rectify.

The rise of think-tanks like Grattan represent a broader reconfiguration of the relationships between states, citizens and expertise as we have moved beyond social welfare agendas, and are finding the market fundamentalism that replaced it is no Shangri-La. Habermas’ (1976) crisis of legitimation seems equally to apply to this new expertise. How do we understand the role of think-tanks, what are the opportunities and limitations? How might we understand the role of academic research and its contribution to policy debates in a 24 hour media cycle where sound bites seem to have greater
impact than careful analysis? Perhaps most importantly, given the significance of think-tanks in policymaking and the expectation that they are here to stay, how do we hold them to account?

References


