DID GONSKI GET IT RIGHT? School Funding and Performance

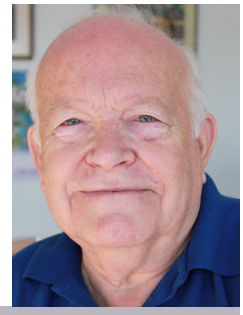
The idea that more funding will lead to better academic results is not supported by evidence, reports **Ken Gannicott**

It has become almost an article of faith that Australian education is under-funded. Much of that belief has been driven by the 2011 Gonski Report on school funding,[1] with the 'I give a Gonski' slogan remarkably effective in raising public awareness of the topic. Despite political debate about the cost of the Report's recommendations in an era of public budget restraint, Gonski's argument for across-the-board increases in funding has not been seriously challenged.

Gonski provided much valuable analysis of the existing method of funding schools. The Report's review of disadvantage is judicious. More problematic is that the Report evolved from a focus on mechanisms for allocating funds to an argument for increased funding across the board. The case for a system-wide increase seemed to be buttressed by the Report's finding that as a proportion of GDP Australia's spending on pre-tertiary education in 2008 was below the OECD average. Those often useful international comparisons can sometimes be a false friend. Latest data show that in fact Australia's expenditure in relation to GDP was slightly above the OECD average in 2008, and by 2012 Australia was spending substantially more than that average.[2]

Gonski was also alarmed by the fact that Australian students are falling behind those in Shanghai, Singapore, Hong Kong and Canada. Achievement scores on the OECD's Program for International Student Assessment (PISA) show that the proportion of Australian students in the highest achievement band has fallen, and our lowest performing students are not meeting minimum standards.[3] Gonski linked alleged underfunding and poor performance outcomes with the claim that 'a significant increase in funding is required across all schooling sectors'[4] and stated that 'the proposed funding arrangements outlined in the report are required to drive improved outcomes for all Australian students'.[5]

Targeted expenditure on specific interventions is one thing: even the most successful school system may have pockets of disadvantage that require money for extra classrooms, teacher bonuses for remote locations, specialist programs for migrants, and the like.[6] Claims that Australia needs to spend more at an overall system level are quite another. Robust international evidence has consistently found that increased overall school spending in advanced countries does not necessarily lead to improved performance.[7] While it is possible that Australia is an exception to the international evidence, nowhere in the Gonski Report is there any



Ken Gannicott was previously Professor of Education and Head of the Graduate School of Education at Wollongong University. He now works as a consultant.

specific test of the relationship between spending and performance in Australian schools.

This article examines the Australian evidence. The testing relies on statistical analysis and is inevitably technical, but the conclusions and policy implications are easily stated: whatever else explains Australia's declining educational performance, there is no evidence from either primary or secondary schools that across-the-board increased funding is associated with improved academic outcomes.

The approach

My approach was straightforward, relating a school's performance as measured by its average NAPLAN score, to its recurrent income per student. NAPLAN (the National Assessment Program—Literacy and Numeracy) tests every student in years 3, 5 7, and 9 in reading, writing, spelling, grammar and punctuation, and numeracy (the five 'domains' in NAPLAN jargon). Measuring primary and secondary schools separately, each school's NAPLAN average is calculated from its aggregate score in years 3 and 5 (primary schools) and 7 and 9 (secondary). Measuring results in all domains across the two years of testing is likely to give a reliable indicator of student academic performance. It hardly needs to be emphasised that there is much more to schooling than the three Rs, but literacy and numeracy are the foundation of further learning. Many countries around the world now administer similar tests, as well as taking part in international assessments such as PISA.

I used the latest NAPLAN data for 2015 to construct a sample directly from the My School website of one in five primary schools and one in two secondary schools.[8] Combined and special schools were not included (see Box 1). After removing schools with missing data and those with very small enrolments, the final tally was 1,208 primary and 712 secondary schools.[9] These are very large sample sizes. The Australian education system is diverse: a large sample generates a strong probability that my sample is representative of that diversity. In contrast to most work in this area, which for data reasons is often limited to a small sample of top-performing Year 12 students or schools, these large samples were drawn from across the spectrum of NAPLAN scores and funding differences; they covered every state; they included government and non-government schools; and they permit correction for the socio-economic background of students in each school. [10,11]

Box 1: Some issues in measuring funding

Assessments such as NAPLAN or PISA have long been contentious, but the expenditure side of performance measurement is not always straightforward either.

It cannot be assumed that all recurrent income is used for explicitly instructional purposes, so funding comparisons between schools should be interpreted carefully. Some studies make a bottom-up estimation of specifically instructional costs, but this has its own disadvantage of rarely going beyond partial measures such as staff salaries.

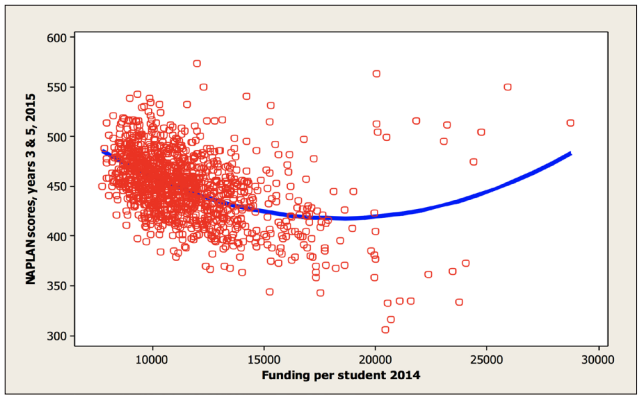
Schools in Australia are required to report school-wide income, and the My School website does not show differences in expenditure between the primary and secondary streams of combined schools. Some combined schools may operate a seamlessly integrated system in which differences between primary and secondary costs are minor. Because of more specialised subject teaching in the upper years of secondary, most schools would face higher costs for teaching the secondary stream. On average, OECD countries spend 15% more per secondary student (OECD, *Education at a Glance*, 2015, p.207). Given that this study requires accurate attribution of funding, combined schools have been excluded.

By definition, recurrent income is intended to be spent in the year in which it is received, but that accounting convention leaves unresolved the educational question of how long it takes expenditure to have a measurable impact on performance. Recurrent funding in Year 1 may be used to hire an extra teacher, but improvements in student performance may not become apparent until Years 2 or 3. My NAPLAN scores refer to 2015; the funding data refer to 2014. While this is the pragmatic result of data availability, these dates do at least pay slightly more than lip service to the notion of lags between expenditure and educational outcomes.

An overview of funding and performance

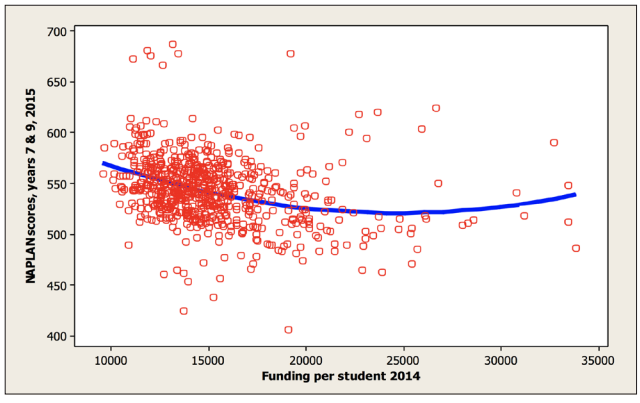
I start by showing the overall relationship between funding per student and NAPLAN scores. Charts 1a and 1b show that overall picture separately for primary and secondary schools.

Chart 1a: Primary funding and NAPLAN scores



Equation for the fitted regression line:
NAPLAN primary = 621.29 - 0.0222 Funding + 0.000001 Funding^2
Adjusted R^2 = 21.6%

Chart 1b: Secondary funding and NAPLAN scores



Equation for the fitted regression line:
NAPLAN secondary = 653.3 - 0.0107 Funding + 0.00000022 Funding^2
Adjusted R^2 = 10.8%

The results are both striking and quickly summarised:

- For both primary and secondary schools increased funding per student is negatively associated with NAPLAN performance for the majority of schools shown in that dense cluster of observations in each Chart. The regression lines turn positive at very high levels of expenditure, as the mathematical functions plot a path between a handful of expensive and high-scoring schools and a small number of equally expensive but low-performing schools.[12]

- Looking right across the spectrum of funding, a simple comparison between the beginning and end points of the regression lines indicates that on average NAPLAN outcomes are no better at very high levels of funding than at lower levels.

- By taking a vertical 'slice' through each scatter diagram it becomes clear that there is a wide range of performance outcomes at every funding level.

Refining the tests

While the Charts convey the overall relationship, they also make clear the limitations of the fitted regression lines. Despite individually significant coefficients, there is much left unexplained. The very wide scatter of observations in both Charts indicates that funding alone explains only a small percentage of variation in NAPLAN performance. Other explanatory factors are involved. For example, there is a modest inverse relationship between a school's funding and its socio-educational index (known as ICSEA). The correlation between funding and ICSEA is -0.33 and -0.38 for primary and secondary schools respectively. These mid-range correlations are high enough to confirm the need to check whether the negative role of funding continues to be statistically significant once we take additional variables into account.[13]

In a series of extended regressions, I added variables for each school's socio-educational background (ICSEA) and type of school (Government or Independent).[14] The results are shown in Box 2.

These extended tests are significant in more than the formal statistical sense. The regressions now measure the association between funding and NAPLAN after taking student and school characteristics into account, and the previous negative results for funding are sustained. Funding continues to show an overall negative (and

Box 2: Extended regressions for schools

	Regression Coefficient	
Variable	**Primary**	**Secondary**
Intercept	135.8	247.2
Funding	-0.0047***	-0.0039***
Funding^2	0.00000014***	0.00000005^{NS}
Government or Independent	-0.047^{NS}	8.76***
Socio-Economic Index ICSEA	0.345***	0.337***
	Adjusted R^2 = 79.5%	Adjusted R^2 = 61.8%
Statistical significance	*** = coefficient highly significant at better than 1% level. ** = coefficient is significant at better than 5% level. * = coefficient has low significance at 10% level. NS = not significant	

significant) association with NAPLAN for primary schools. At secondary level, the funding variables also show the same mainly negative pattern as in the simpler scatter Chart 1b. Socio-economic standing has a strong, positive and significant association with NAPLAN. Government schools show a significantly positive relationship with NAPLAN for secondary schools.[15]

The conclusion from both the simple scatter charts and the extended regressions is that there is no evidence that system-wide extra funding is necessary for good NAPLAN performance. Across the range of two very large samples, even when adjusting for important variables such as social background, I cannot conclude that additional funding per student is associated with better performance. Nor is there a close association between NAPLAN and any given level of funding: any given level of funding is associated with a wide range of NAPLAN outcomes.

Measuring under-performance

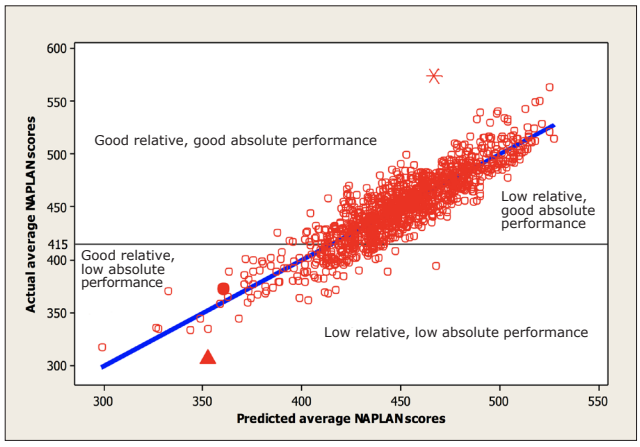
Concern about student performance in Australia is not misplaced. While all countries vary in their social and cultural structure—and this structure has an important influence on educational outcomes—results from PISA are taken seriously because comparative PISA scores offer a standard benchmark of attainment, whatever the role of socio-cultural differences. Results from PISA 2012 indicate that by international standards the proportion of lowest-performing Australian students has reached 20% in mathematics. On average, Australian students aged 15 have fallen 18 months behind those in Shanghai in reading, and seven months behind those in British Columbia. At the other end of the performance scale, the proportion of Australian students aged 15 in the top international maths categories fell from 20% in 2003 to 15% in 2012.[16]

It is not possible to analyse these issues with my NAPLAN data, but it is possible to shed some light on the lowest-performing schools. One of the powerful features of regression analysis is that the results can be used not just to measure association between variables. Regression equations can also be used for predictive purposes. Once the statistical determinants of NAPLAN performance have been calculated (as shown in Box 2) those coefficients can be combined with each school's endowment of funding, socio-educational index, and so on. It then needs no more than high school arithmetic to solve the equation for each school and estimate the performance that 'should' result from any school in the sample, given its endowments.

It needs to be emphasised that 'predicted performance' is a statistical construct subject to the usual caveats about confidence intervals and margins of error, but in light of current public debate it is useful to add an additional approach. Chart 2 (overleaf) shows the results for primary schools.

Looking first at the diagonal regression line, this shows average performance for the primary sample as a whole. Schools located along or near that line are those whose actual performance is just what would be expected, given their funding, social background and sector. Observations above the diagonal line are the over-achievers, those whose actual performance exceeds what would be predicted. In today's jargon, they punch above their weight. Some schools, such as that identified with an asterisk, are doing spectacularly well: that school's NAPLAN score far exceeds what would be predicted from its funding and student background characteristics (predicted NAPLAN is 467).

Chart 2: Predicted versus actual NAPLAN scores, primary schools



Schools below the line are not achieving their potential. They are under-performing against what would be predicted. On the basis of its background characteristics, the school identified with the triangle is predicted to achieve a NAPLAN score of 353. Its actual score in 2015 was barely 300.

It is important to emphasise that the calculations underpinning Chart 2 take into account the socio-educational background of the students in each school, as measured by ICSEA and other variables. They are what My School would call 'fair' comparisons, with NAPLAN scores in effect corrected so that like is compared with like. But there is a problem. Such estimates measure under-performance in a purely relative sense, comparing performance by schools in the light of each school's characteristics. Relative performance can be thought of as a type of 'personal best'. This can be useful, but it is not the same as measuring performance against an absolute or external benchmark of achievement.

Looking at the school identified by the solid circle in Chart 2, it is in relative terms a satisfactory school: it is located 'above the line', doing somewhat better in NAPLAN than would be expected from its characteristics. But in absolute terms its actual score is well below the primary sample average of 450. Statistically speaking, that school has done well given its funding and social characteristics, but suppose that low score means that its graduates are leaving school with a level of literacy or numeracy too low for effective functioning in contemporary Australia?

Establishing a benchmark for that effective functioning cannot be done with any precision with my data, but by using assumptions that researchers like to call 'heroic', I can derive some indicative evidence. This is where the horizontal reference line in Chart 2 comes into the story.

When marking NAPLAN assessments, raw scores are converted to a scale consisting of ten bands. These bands are intended to represent the increasing complexity of knowledge and skills assessed by NAPLAN from Years 3 to 9. Students with results in the lowest band have not achieved the national minimum standard of knowledge and skills. Without that standard a student will have difficulty making sufficient progress. The second lowest band on the achievement scale represents those who are only just meeting the national minimum standard.

My sample consists of schools, not students, and all schools have a mix of students scoring in the different band levels. For this exercise the school is the appropriate unit. From the point of view of national policy, it is schools, not students, which will be accountable for non-performance, and schools will be the instrument of improvement. The proportion of students achieving each band at the national level is known from NAPLAN reports[17] and I have extended that proportion to apply to the distribution of schools. Nationally, across the five testing domains in Years 3 and 5, 14.2% of students scored in the lowest two bands in 2015. While it is very much an approximation, it is plausible to suppose that 14.2% of schools can also be characterised as meeting only the lowest two bands of achievement. Given my sample size, that benchmark means that the lowest-scoring 172 primary schools have an overall performance level which puts them in the lowest two bands of being below or only just reaching the national minimum standard. This is the equivalent in our primary sample of an average NAPLAN of 415.

The horizontal reference line in Chart 2 shows this cut-off. All schools located below this line are not meeting, or are only just meeting, national minimum standards. The reference line also works in conjunction with the diagonal line to define four areas of relative and absolute performance. Going back to the school identified with the solid

circle, Chart 2 now confirms that it is performing adequately in relative terms, but not well enough to meet minimum standards of academic achievement. More importantly, it is apparent from the low relative, low absolute performance segment of Chart 2 that those schools with NAPLAN scores so low that they are not meeting, or barely meeting, national minimum standards, are overwhelmingly the under-achievers. They not only fail to meet absolute standards of performance, but also do not make the most of their characteristics, including their existing level of funding.

This finding suggests that improved performance is largely a matter of raising standards of the lowest-performing schools and students so that they reach their potential, an inference which is consistent with international evidence. In its review of PISA trends, OECD noted that 'in nearly all the countries that showed improved performance during the period, the percentage of low performers dropped, meaning that the number of students who scored below the PISA baseline reading proficiency Level 2 was significantly smaller in 2009 than in 2000'.[18]

To come full circle with the analysis of funding, a check of the sub-sample of 172 primary schools whose NAPLAN scores put them at or below the national minimum standard reveals that the average funding per student is $14,226. For the 1,036 primary schools which perform above that minimum standard the average funding in 2014 was $11,447. The reasons for low performance in the 172 schools are no doubt complex, but funding is clearly not the major explanation.[19] There is no reason to suppose that additional funding will in itself raise their levels of achievement to acceptable levels.

Explaining the findings

The analysis in this article produces a result that runs utterly counter to intuition. But it is consistent with a large body of international evidence. As recently as 2007 it was possible to note, with the frustration clearly apparent, that 'education reform is top of the agenda in almost every country of the world. Yet despite massive increases in spending . . . and ambitious attempts at reform, the performance of many school systems has barely improved in decades'.[20]

There seemed to be something missing in the policy approach to school performance. The explanation was neatly expressed by Eric Hanushek in his pioneering work: 'if we think of schools as maximising student achievement, the . . . evidence indicates that schools are economically inefficient, because they pay for attributes that are not systematically related to achievement.'[21] Reduced teacher-pupil ratios, smaller classes, longer pre-service training, graduate education for teachers, training teachers in inappropriate instructional methods or ineffective techniques of reading, and funding more teachers even if this increased supply drives down their quality have—until recently—formed the reality of how most education money was spent in many OECD countries, including Australia. Yet there is little evidence from the international literature that expenditure on any of those items leads to improvements in student performance.

This finding suggests that improved performance is largely a matter of raising standards of the lowest-performing schools and students.

In recent years there has been a perceptible evidence-based shift in international research. There is now a large and growing literature, including contributions from Australia.[22] With (i) reduced quantitative pressures because most children are now enrolled in school in OECD countries; (ii) the enhanced interest in performance as a result of often dismal TIMSS (Trends in International Mathematics and Science Study) or PISA results; and (iii) the realisation that increased public expenditure does not in itself produce more effective schools, education policy has taken a new direction, focusing upon specific practices in the school and classroom.

This article is not the place for a review of that literature, but a single example will illustrate the change. In past years, one of the main tasks for education policy was to ensure that all teachers received pre-service training and were certificated

as teachers. We now understand that this is a necessary but not sufficient condition. Just like any other occupational group, there is a distribution of effectiveness among qualified teachers. Evidence from rigorous investigation finds consistent and significant results for more effective teachers. According to UK evidence, one year with a very effective teacher adds 25-45% of an average school year to a pupil's maths score. The effects of high-quality teaching are especially significant for pupils from disadvantaged backgrounds. Over a school year, these pupils gain 1.5 years' worth of learning with very effective teachers, compared with only 0.5 years with poorly performing teachers. In other words, for pupils from a low socio-economic background the difference between a good teacher and a bad teacher is a whole year's learning.[23]

It is not a question of how much money, it is a question of how it is spent.

The policy implications of such evidence for how education funding is spent are dramatic. One recent plea for a change of approach noted: 'paper qualifications and personal characteristics tell us very little. Gender, race, teaching experience, undergraduate university attended, advanced degrees, teacher certification and tenure explain less than 8% of teacher quality. The underlying facts . . . that the difference between good and bad teachers is very large . . . should change the way we think about selection into teaching, the nature of teacher training, the professional development of teachers, and the management of under-performing teachers.'[24] Policies consistent with these recent research findings will cost money, but it will be targeted, performance-relevant expenditure that differs from older and out-of-date notions of how educational funding should be spent. It seems highly likely that the four areas of relative/absolute performance defined in Chart 2 will each require a different set of tailored policies.

Conclusions

There is no uniquely correct way to test the role of funding in educational performance. Even major international exercises such as PISA are driven by the pragmatics of data availability as well as specification of the relationship between variables. This limitation is true also for this article. The estimates for absolute levels of under-performance are more an illustration of a method than decisive evidence. Nor does the analysis measure the impact of funding increases over time. It looks only at the relationship between funding and performance at a point in time. With several years of NAPLAN scores now available, future analysis could usefully examine 'gain scores' or changes in performance over time.

These limitations do not, however, detract from the broad lessons to be learned. The first and dominant finding is that there is no evidence from either primary or secondary schools that higher funding is associated with improved academic outcomes. No matter whether we use a simple test of the relationship between school funding and NAPLAN scores, or whether we include additional explanatory variables such as the socio-educational background of the school and its students, the result is the same: school performance is negatively related to funding.

This leads directly to the second major inference to be drawn from this article. There has to be a change of focus. A striking feature of the popular debate about Gonski has been that increased funding for schools is taken to be self-evidently beneficial. It is rare to find any discussion of how the money is to be spent, but that should be the starting point of funding decisions. It is not a question of how much money, it is a question of how it is spent. Future funding decisions should not be decided separately from the evidence supporting the newer approach to education policy. If we are serious about raising academic performance in Australian schools, it will require evidence-based policy review about the number of admissions and methods of selection into teacher training faculties, a fresh approach to training teachers in effective methods of instruction, and incentives and rewards for more effective teachers.[25] More of the same is not going to work.

Finally, if we are serious about low performance, it is time to review the Australian Curriculum, Assessment and Reporting Authority (ACARA)'s approach to information about that performance. There is much of value to parents in the My School

website administered by ACARA. Development of the index of socio-educational background (ICSEA) is a major contribution. Despite the limitations of a single figure to represent funding in both primary and secondary streams of combined schools, the development of a standardised approach to compiling data on school income is genuinely useful.

The fact remains, however, that ACARA publishes only limited information about school performance using its own definition of statistically comparable or 'fair' comparisons. In other words, it relies on relative notions of performance rather than providing an overall view of whether schools are meeting the absolute criterion of national minimum standards. This is a serious omission in today's climate of concern about performance and minimum standards. It overlooks the fact that it can also be regarded as 'fair' that every child should reach minimum standards of literacy and numeracy, regardless of economic and social background. As noted earlier, improved performance may be largely a matter of raising the standards of the lowest performers. There is a legitimate public interest in detecting schools which are not meeting minimum standards.

This argument is closely analogous to the publication of PISA scores. Those scores are often unwelcome, even embarrassing. It is now well understood that PISA scores are partly an outcome of a country's social and cultural circumstances. But PISA publishes its results without any adjustment for a country's socio-cultural background, implicitly and correctly acknowledging that a benchmark of absolute performance is the critical foundation for understanding whether student achievement reaches appropriate standards, not just a 'personal best'. Much school information depends upon cooperation from the schools, and no-one suggests that there is much to gain from a tabloid orgy of 'name and shame'. There is a balance to be found between protecting the goodwill of schools and protecting the public interest. Given the findings in this article, it can be argued that ACARA is implementing governmental decisions which have tipped that balance too far away from supporting the public interest in improved performance by our schools.

Endnotes

1 David Gonski (Chair), *Review of Funding for Schooling—Final Report* (Australian Government, December 2011).

2 The latest and revised data show that in 2008 Australia's pre-tertiary spending, public and private, amounted to 3.7% of GDP, slightly above the OECD average of 3.6%. By 2012, the Australian percentage had risen to 4.0%, well above the OECD figure of 3.7% (OECD, *Education at a Glance 2015: OECD Indicators,* Paris: OECD Publishing, Table B2.2). Taking a longer term perspective, Buckingham has pointed out that expenditure on schools has risen faster than student numbers over several decades, with the result that per student expenditure has also grown substantially, even after adjusting for inflation (Jennifer Buckingham, *School Funding on a Budget*, Target 30 series, T30.09, Sydney: The Centre for Independent Studies, 2015).

3 Ben Jensen, Julie Sonnemann, Katie Roberts-Hull and Amélie Hunter, *Beyond PD: Teacher Professional Learning in High-Performing Systems, Australian Edition* (Washington, DC: National Center on Education and the Economy, 2016). See also OECD, *PISA 2012 Results: What Students Know and Can Do: Student Performance in Mathematics, Reading and Science Volume 1*, (OECD, 2013), http://www.oecd.org/pisa/keyfindings/pisa-2012- results-volume-I.pdf

4 Gonski Report, p. xv.

5 As above, p. xi.

6 Natasha Bita, 'Gonski Fertilises Green Shoots of Success in Disadvantaged Suburbs', *The Weekend Australian* (12-13 March 2016). Bita describes improvements in literacy and numeracy from increased funding targeted on the basis of need since 2014.

7 The literature on this topic is extensive. The issue was kicked off by Eric Hanushek ('The Economics of Schooling: Production and Efficiency in Public Schools', *Journal of Economic Literature* (September 1986) and this is still a good starting point. A more recent perspective can be found in Eric Hanushek and Ludger Woessmann, 'Education and Economic Growth', in Penelope Peterson, Eva Baker and Barry McGaw (eds), *International Encyclopedia of Education*, vol. 2 (Oxford: Elsevier, 2010). Jennifer Buckingham, in *School Funding on a Budget*, Target 30 series, T30.09 (Sydney: The Centre for Independent Studies, 2015), figure 6, contributes supporting evidence from Australia using PISA scores. By contrast, ABC reporter Alison Branley, in a recent post on the ABC's online discussion forum ('Explained: How Funding Affects Educational Results', *The Drum*, 7 April 2016, http://www.abc.net.au/news/2016-04-07/branley-how-funding-affects-educational-results/7307692?WT.mc_id=newsmail) argues that there is a positive correlation in Australia between funding and improvement, but the methodology is flawed and the results do not support her argument. The claimed relationship between state school funding per student in 2012-13 and 2014 NAPLAN mean scale scores for reading, year 5, in fact produces a *negative* correlation of -0.36. The attempt to compare student gain in reading between years 3 and 5 founders on the fact that it is not meaningful to

compare only one component of NAPLAN from two different years with a single year estimate of funding for the whole school. It is no surprise that although the correlation is modestly positive at 0.36, its statistical probability is 0.39. This so far exceeds the conventional probability level of 0.05 that we cannot conclude that the correlation has any statistical significance. In plain language, it is a result that could have occurred by chance.

8 The My School website provides information on school performance to help parents, teachers and policymakers. See https://www.myschool.edu.au. The website is managed by the Australian Curriculum, Assessment and Reporting Authority (ACARA). The author would like to acknowledge ACARA's My School website as the source of the NAPLAN and funding data used in this article. Use of the data does not imply any endorsement or agreement by ACARA with the analysis or conclusions of the article.

9 In recent years *The Australian* has constructed its own database of NAPLAN results, publishing lists of the 100 top primary schools, the top 50 in reading and writing, and so on. Because of the volatility of any given year's test results when student numbers are small, that database excludes primary schools with fewer than 100 enrolments and secondary schools with fewer than 200 students. This seems sensible, and I have followed the same procedure. I made only limited correction for statistical outliers. Outliers are extreme or very atypical data observations. They are always problematic. Outliers can exert a disproportionate leverage on statistical results, resulting in a misleading interpretation of the main body of data, but care has to be taken not to manipulate the evidence by excluding inconvenient observations. It makes little sense to compile a large sample in the interests of capturing the diversity of Australian schools, only to delete large numbers of schools which fall outside the main pattern. The usual benchmark for outliers is a standardised residual of 2 or 3. I adopted an ultra-cautious decision rule, deleting what by any standard were extreme observations (two observations from the primary sample, with residuals exceeding 6 and 9, and four from the secondary sample, with two residuals of 4.8 and two of 5).

10 Socio-economic background was measured by the Index of Community Socio-Educational Advantage (ICSEA) for each school. The index was developed by ACARA as a measure of educational (dis)advantage, and includes—for each school—information on parents' occupation, parents' education, geographic location of the school, and the proportion of indigenous students. The index is defined to have an average of 1,000.

11 Despite the large samples and the data I was able to include, it is worth emphasis that this analysis does not pretend to measure all the many variables which may contribute to improved educational performance. Statistical textbooks point out that omitting important variables can produce misleading results. The size of the bias is related to the importance of the omitted variable and its correlation with the included variables. On the other hand, building a 'kitchen sink' model and relying on the computer to select significant variables may not help our understanding very much. For example, the days have long gone when regression studies of funding and performance would routinely include variables for class size or teacher ratios. We now know from more recent work that what matters for performance are genuinely effective teachers, a culture of high expectations, teacher collaboration, regular assessment feedback to students, and the method of teaching. None of these things can be measured with the data we have available. It should also be noted that some plausibly important variables were deliberately excluded. The My School website indicates the location of each school, such as metropolitan, provincial, or remote. It is redundant to include such variables when the variable for socio-educational background (ICSEA) already contains a component for geographic location. Similarly, the importance of measuring the role of indigenous students cannot be denied, but no separate allowance for this is necessary when the proportion of indigenous students is already included in the calculation of ICSEA for each school.

12 With today's computing power it is a simple matter to test a wide variety of mathematical forms for the best fit of variables in a scatter diagram. These 'fishing expeditions' can be useful if we do not have prior information or expectation about the nature of the relationship between X and Y. That procedure was followed in this case. It is clear from Chart 1a and 1b that the funding/performance relationship is approximately linear across the funding experience of the vast majority of schools (the dense cluster of data points in the graphs). For those schools a declining linear function ($Y = a - bX$) would provide a simple but adequate fit. However, tests established that a quadratic function provided a better statistical fit when taking into account the full range of schools in my sample. I also tested whether funding and NAPLAN are related through a percentage rather than dollar value relationship. I estimated this by transforming the relevant data to logarithms. Neither log-linear nor double log formulations approached statistical significance.

13 It might be argued that the observed negative relationship between funding and performance is a consequence of lower performing schools receiving extra funding because they have more students from a lower socio-economic background. Feedback loops such as this are very common in social or economic activities, making it difficult to identify the direction of explanatory relationships. Mechanisms such as the Average Government School Recurrent Costs (AGSRC) contained a component of needs-based funding, but the major thrust of the Gonski Report was to highlight the inadequacies, limited impact and inter-state inconsistencies of the socio-economic components in overall funding. Feedback effects are therefore likely to be small. A greater emphasis on needs-based funding was introduced under the Australian Education Act (2013) and

full implementation of the needs-based funding proposed by Gonski could make corrective procedures such as two-stage least squares more important in future analysis.

14 I also tested for interactions between ICSEA and type of school and between funding and type of school. I could identify no such interactions.

15 This was an unexpected result, and may be a consequence of the omission of combined schools. A strong majority of combined schools are Independent. Their exclusion meant that, particularly at secondary level, the sample is biased towards government schools. As explained in Box 1, on a fine balance of argument I judged that the problems of attributing costs between primary and secondary streams in the combined schools justified their exclusion from the sample. In line with standard statistical practice, I measured type of school as a dummy or indicator variable. With government schools coded 1 and Independent schools coded 0, the dummy variable for sector is interpreted as the coefficient for government schools relative to independent schools.

16 Jensen et al, *Beyond PD* (see note 3). Peter Goss, Julie Sonnemann, Jordana Hunter, Cameron Chisholm and Lucy Nelson, *Widening Gaps: What NAPLAN Tells Us About Student Progress* (Grattan Institute, March 2016), add convincing evidence about low performance, commenting that 'NAPLAN national minimum standards (NMS) are set *very* low. A Year 9 student can meet NMS even if they are performing below the typical Year 5 student. They can be a stunning four years behind their peers' (p. 2).

17 Australian Curriculum, Assessment and Reporting Authority (ACARA), *NAPLAN Achievement in Reading, Persuasive Writing, Language Conventions and Numeracy: National Report for 2015* (Sydney: ACARA, 2015), pp.1-110.

18 OECD, *Improving Performance: Leading from the Bottom*, PISA in Focus 2 2011/2 March (OECD, 2011).

19 It might be argued that this funding differential is misleading, because many of the 172 low-performing schools are small, in remote areas, often with high proportions of indigenous students, and with families drawn from lower-income social groups. On this line of argument, such schools are more expensive to operate and dollar for dollar may not have the same performance payoff as other schools. This broad line of argument may be correct, but the whole point of using ICSEA to correct for socio-educational standing of each school is to standardise across the sample for these differences in background, and very small schools were in any case removed from my sample.

20 Michael Barber and Mona Mourshed, *How the World's Best-Performing School Systems Come Out On Top* (London: McKinsey & Company, 2007) was among the first to look in depth at how the consistently good performers in international tests actually achieved their success. This was followed in 2010 by Michael Barber, *How the World's Most Improved School Systems Keep Getting Better* (London: McKinsey & Company, 2010).

21 Eric Hanushek, 'The Economics of Schooling', (see note 7).

22 The McKinsey & Company approach (see note 20) of focusing on analysis of what lies behind the PISA numbers and 'what works' in producing high student performance has been followed by the OECD in its diagnostic follow-up to the PISA tests. Of the many publications, OECD, *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education* (Paris: OECD Publishing, 2011), is among the best, and is not exclusively concerned with the United States. More detailed classroom-based evidence can be found in the following sources. Sutton Trust, *Improving the Impact of Teachers on Pupil Achievement in the UK—Interim Findings* (UK: The Sutton Trust, September 2011); John Hattie, *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement* (London and New York: Routledge, 2009); Jodene Dunleavy et al, *The Relationship Between Student Engagement and Academic Outcomes*, *What Did You Do in School Today?* Research Series Report Number One (Toronto: Canadian Education Association, 2012); Paul A. Kirschner, John Sweller and Richard E. Clark, 'Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-based, Experiential, and Inquiry-based Teaching', *Educational Psychologist* 41:2 (2006); NSW Centre for Statistics and Evaluation, *What Works Best: Evidence-based Practices to Help Improve NSW Student Performance* (NSW Government, Education and Communities, 2015)*;* Ben Jensen, *Turning Around Schools: It Can be Done* (Grattan Institute, 2014); Ben Jensen, *Catching Up: Learning from the Best School Systems in Asia* (Grattan Institute, 2012); Robert Coe, Cesare Aloisi, Steve Higgins and Lee Elliot Major, *What Makes Great Teaching?: Review of the Underpinning Research* (UK: Sutton Trust, October 2014) reviews over 200 pieces of research to identify the elements of teachers with the strongest evidence of improving attainment, dispelling many myths along the way.

23 See the Sutton Trust, *Improving the Impact of Teachers on Pupil Achievement in the UK—Interim Findings* (UK: The Sutton Trust, September 2011).

24 As above.

25 Some of these recommendations for a new approach are under active consideration as part of proposals in the Federal Budget 2016. As part of a $1.2 billion increase in funding over the period 2018-2020, teacher pay will be linked to performance rather than seniority, states will be required to meet basic standards in literacy and numeracy to qualify for additional funding, teacher trainees will be taught literacy and numeracy teaching as a core skill, and age-appropriate testing will help identify students who start school and need extra help with early-grade tasks such as writing their name or simple counting.