Grattan Submission

Productivity Commission
Education and Training Workforce Study

19 August 2011

Table of Contents

1. Introduction

In developed nations like Australia, most wealth is held in the form of human capital.[1] Schools, clearly, are an essential ingredient in generating this wealth. In Australia, our school system is fortunate to benefit from the efforts of many talented and dedicated people. But despite these efforts, the system is producing some worrying results. In tests that measure meaningful capabilities, Australian students are falling behind our regional peers.[2] There is a growing disconnect between stagnating performance on the one hand, and real increases in per student expenditure on the other.[3] And inequality of outcomes continues to be a system-wide challenge.[4]

The single biggest determinant of how well schools function is the effectiveness of learning and teaching.[5] Differences in teacher effectiveness account for a large proportion of differences in student outcomes – far larger than differences between schools.[6] Conservative estimates suggest that a student with an excellent teacher (at the 90th percentile) would achieve in half a year what a student with a less effective teacher (at the 10th percentile) will learn in a full year.[7]

The goal of improving teaching is both laudable and urgent. It's also very broad. Behind this goal sit complex, interrelated decisions and the substantial challenge of implementation. There are at least five mechanisms to lift teacher effectiveness and improve learning and teaching in schools:

1. Improving the quality of applicants to the teaching profession;
2. Improving the quality of teachers' initial education and training;
3. Appraising and providing feedback to improve teachers once they enter the profession and are working in our schools;
4. Recognising and rewarding effective teachers; and
5. Addressing under-performance.

These mechanisms, in turn, depend on a series of related initiatives, including: collecting and making the most of more accurate data; providing meaningful teacher appraisal linked to teacher development; and reforming how high-quality teaching is recognised.

[1] Barlevy and Neal (2011)
[2] Thomson *et al.* (2010)
[3] Jensen *et al.* (2011)
[4] Jensen (2010b) p.6
[5] Jensen (2010c) p.8
[6] Jensen and Reichl (2011) and references on p.6
[7] Leigh (2010); Hanushek (1992)

This submission canvasses some of these issues as they relate to the questions outlined in the Commission's *Schools Workforce Issues Paper*. The submission is based largely on five pieces of research undertaken by the Grattan Institute's School Education Research Program in the last 18 months.

Grattan reports about the school workforce

- ***Measuring What Matters: Student Progress***
  The report details the benefits (and costs) of collecting and analysing value-added scores, and was published in Jan 2010

- ***What Teachers Want: Better Teacher Management***
  The report analyses the views of teachers collected in the OECD's TALIS survey, and was published in April 2010

- ***Investing in Our Teachers; Investing in Our Economy***
  The report looks at the economic implications of improving teacher effectiveness, and was published in Nov 2011

- ***Better Teacher Appraisal and Feedback: Improving Performance***
  The report proposes a new system of teacher appraisal and development, and was published in April 2011

- ***The Real Issue in School Funding: An Analysis of Increasing Government School Expenditure and Declining Performance.***
  The report provides a high-level analysis of cost drivers in the school system, and will be published in the AER in Sep 2011

2

These reports tackle some of the biggest issues that affect how students learn. Thoughtful, major changes to areas like teacher appraisal offer significant benefits. Incrementalism, although politically more palatable, is not what is required if we're serious about making Australia's school system the best in the world.

Although many voices contribute to the debate on education reform, **a clear agenda to take our students to the best in the world is yet to emerge**. School autonomy, for example, can offer a range of benefits – and as this submission suggests, can compliment important reforms. It can, for example, help to make the most of improved data on student and teacher performance by empowering principles and teachers to *use* the information in decision making. However, very little research suggests that this prominent part of the reform agenda is a central driver – either directly or indirectly – of

how students progress.[8] In general then, a significant role the Commission can play would be to strip away ideas on the periphery of student performance, and highlight those reforms which are most important to increasing student learning.

Similarly the Commission can play a role in reintroducing necessary complexity to valuable reforms. 'Performance-based pay', for example, is often taken to mean a simple system of bonuses based solely on test scores. Such a system is relatively easy to characterise as arbitrary and counterproductive. But this simple characterisation misses the variety of ways in which teacher effectiveness can and should be appraised, of which test-scores are just one element. More importantly, it misses the bigger picture: giving teachers clearer and better-structured career progressions linked to significant increases in pay. Promoting a more sophisticated understanding of how reforms such as this might operate will lower some of the barriers to change.

The Commission will no doubt be interested in ways to increase flexibility in the teacher labour market. There are a number of issues here that should be considered:

- Centralised wages agreements: clearly these impede on flexibility and the opportunity to differentiate pay based on performance. Centralised agreements also fail to recognise that there are numerous labour markets for school teachers, with differences stemming from subject and year level taught. Treating these labour markets as homogenous creates both surpluses and shortages in particular areas.

- Student-teacher ratios and class size: these are central requirements in many systems and have a large budgetary impact. This impact restricts flexibility in teacher salaries given the budget commitment to what are often reduced class sizes.

- Poor information: this is an issue that is unfortunately often overlooked. There is a considerable lack of information on teachers' skills (including the effectiveness of the education and training they have received) and their effectiveness within schools and classrooms. This restricts labour market flexibility by increasing transaction costs and the difficulties with recognising performance and addressing under-performance. The discussion of teacher appraisal and feedback in this paper highlights this issue.

[8] See, for example, Hattie (2009)

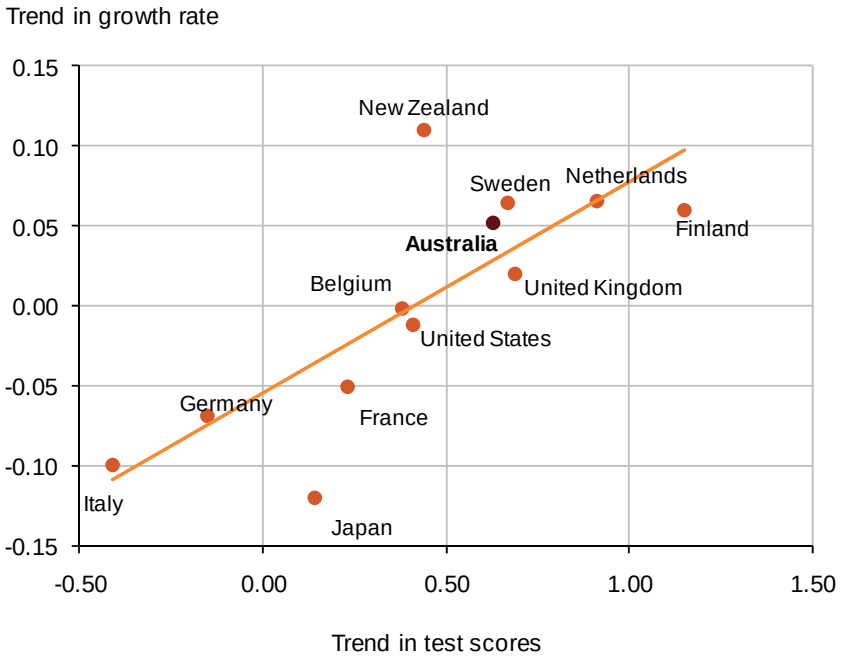
2. Data and benchmarks

2.1 Using international data

The commission will be well aware of Australia's performance in international student assessments such as PISA, TIMMS and PIRLS. But it is worth emphasising the recent research showing the strong relationship between performance and economic growth.

Assessments such as PISA are not narrow, or curriculum based. They measure something meaningful and useful. A strong indication that tests such as PISA provide an important measure of human capital is the degree to which student performance correlates with economic growth (illustrated in Figure 1).[9] There was a much stronger correlation than previously thought with older human-capital models. The issue of the relevance of student performance on tests – and the relationship between these measures and real per capita GDP growth – is discussed in *Investing in our Teachers, Investing in Our Economy.*[10]

Figure 1 - Trends in educational performance and economic growth rates



Note: Only 12 countries have participated in international tests over a sufficiently long period to look at trends over a 30 year period. In the chart, the 'trend in growth rate' is simply a bivariate regression of test scores on time. Trends in test scores are similarly derived. The plot provides the pattern of slopes from the test regressions.

Source: OECD (2010), originally presented in Hanushek & Wössmann (2007)

[9] This issue is also covered in detail in OECD (2010b).

[10] Jensen (2010a). See p.18 for a summary of the research linking human capital (as measured by student performance on tests such as PISA) with growth in real GDP per capita.

There are a number of important implications for using international assessments. First, data from international tests can be used to identify and understand policy-relevant issues. Using meaningful, credible benchmarks helps shine a light on important issues that may not be emphasised by inward-looking analysis. An example with significant policy relevance is inequality. By providing an international perspective, the PISA results emphasise that there is a relatively large gap between high-performing and low-performing Australian students. In addition, the data strongly suggest that there is not a trade-off between high performance and promoting equality. The top performing school systems (Shanghai, Korea and Finland) have very low levels of inequality. In short, evidence suggests that inequality need not stand in the way of performance improvements. Data from tests such as PISA can also be used to start to unpick the nature of unequal outcomes. In this case the vast majority (81%) of the variance in performance amongst Australian students was shown to be *within schools*, and not between schools or states.[11]

Second, comparisons across countries and across time can identify high-performing systems from whom we can learn. A prime example here is the rapid increase in performance of East Asian school systems such as Hong Kong, Korea, Shanghai, and Singapore. The recent upward trajectory of this cluster of education systems – even compared to historical high achievers such as Finland – has been too stark to ignore. Many fields of endeavour benefit from clusters of innovation, where each fresh achievement offers goals and lessons for others. The successes of East Asia suggest that the region may be a source of innovation and best practice for the rest of the world. Clearly, policies have to be appropriate for the Australian context, but to disregard this success on the basis that benchmarks lack cultural relevance would be to miss an opportunity to learn from systems that are striding forward.

Third, in an increasingly knowledge-based and integrated global economy **international benchmarks also matter because they measure the standards that other countries set**. Once again, the prime example for Australia is the success of economies like Hong Kong, Korea, Singapore and Shanghai. These places are our economic and geographical peers, and success in Australian education must in part be defined relative to the achievements of these peers. We can ill-afford to ignore the standards being set in our region. The shift in the world's economic centre of gravity to Asia only heightens the need to improve performance.

[11] Thomson and De Bortoli (2008)

2.2 Using local data

Collecting and analysing accurate and granular data on domestic outcomes is critical to teacher quality. Data can help identify areas where improvement is needed both across the system and within schools. It can also be used for individualised instruction to improve outcomes for particular students.[12] To perform these roles, data needs to be accurate, unbiased, and timely.

Recent developments in data collection and analysis have been positive, but the area remains a relatively untapped resource. Of particular benefit would be the measurement and use of **value-added scores.** The benefits (and costs) of using value-added scores are discused in detail in *Measuring What Matters: Student Progress.* In short, measuring value-added scores is the best method available to isolate the contribution of schools from other factors that affect student performance. Value-added is fairer than other methods – which can often have significant bias, particularly against schools serving more disadvantaged communities.[13]

Effective implementation of value-added scores also requires **more school- and student-level data**. In addition to the high-quality NAPLAN assessments, reliable student-level data is an important input to value-added measures of school performance.[14] As discussed in *Measuring What Matters,* at present there are some problems with the quality and coverage of this data.

Making the most out of good data

Measuring value-added scores is of limited use unless these data are used as a basis for action. School principals and teachers should be empowered to *use* value-added measures to improve instruction and school programs. Resources should be provided for teachers and school principals to better analyse value-added scores, develop programs to address problems, and disseminate best practice. This would involve:

- A user-friendly information technology system should be developed that allows school principals and teachers to better analyse and then act upon their own performance data (not just value-added data).

[12] Griffin (1990)
[13] See Jensen (2010b) and references therein on p.17
[14] Goldstein *et al.* (2008)

- Education and training for school principals and teachers to incorporate value-added measures and other assessment information into instruction and school programs should be provided.

2.3 Program evaluation

One of the great benefits of data is that it can be used to inform decisions with evidence. It is especially useful for spending and investment decisions where the options are many and the stakes are high. Unfortunately, very little rigorous evaluation of programs and cost-effectiveness analysis is done within the Australian school system. Decisions are made without the best possible understanding of what works in different contexts, or – critically – which programs achieve results at the lowest cost.

The need to improve the information available to decision makers is underscored by trends in performance and spending. For decades, real expenditure per student has increased with little or no improvement in student performance.[15] As the paper *The Real Issue in School Funding* illustrates, this long-term trend has continued over the past 10 years. From 2000 to 2009, real expenditure per student in government schools increased 17%.[16] During this period there was, on average, a 13 PISA-point decline in performance.[17] Australia was one of only four OECD countries to experience a statistically significant decline over the period.[18]

Given expenditure increases and the lack of impact on performance, the first step in developing and understanding of cost-effectiveness is simply to analyse which costs have grown. *The Real Issue in School Funding* presents a preliminary high-level analysis of the main cost drivers. Unsurprisingly, changes in teacher expenditure made up the largest percentage of the total cost increases from 2000 to 2009. *The Real Issue in School Funding* identifies three factors behind this rise in teacher costs:

1. Decreases in student-teacher ratios – which account for roughly half the increases in total teacher costs.
2. Real increases in teacher wages – which had negligable impact on total teacher cost.
3. Other factors (largely the natural ageing of the teacher cohort) – which account for roughly half the increases in teacher costs.

[15] Leigh and Ryan (2011)
[16] These data are not highlighted to suggest that government schools in particular are inefficient – but because at the time of writing Jensen *et al.* (2011) they were the most pubicly available cost figures.
[17] Jensen *et al.* (2011)
[18] Ibid.

The example of class-size reduction

The significance of student-teacher ratios suggests that cost-effectiveness work should begin with a closer analysis of class-size reductions.[19] As discused in *Investing in our Teachers; Investing in our Economy* many studies find that the effects of modest class-size reductions are either small, or non-existent.[20] Small classes can be valuable in specific classes for specific students, but across-the-board class size reductions have not proven to increase performance.

On the other side of the ledger is cost. While the precise costs of class size reduction varies depending on the system and the method by which class sizes are reduced, reduction of even just few students per class can be very costly. More teachers are often required to teach the greater number of smaller classes and teacher salaries comprise around two-thirds of all education expenditure in Australia OECD (2010a). For many individual schools, the proportion of their budget dedicated to teacher salaries is higher, often closer to 90%.[21]

Therefore, even if there were positive outcomes – as may well be the case in classes of younger and disadvantaged students[22] – the costs involved raise the urgent question: does shrinking class size represent value for money? The continued drive to spend money on these programs in the context of rising costs and slipping performance highlights the need for more widespread use of cost-effectiveness analysis.

[19] Although there is a strong link between the student-teacher ratios and class size, it is important to note the differences between the two. These differences stem from variability of instruction time, teachers' overall working time, the number of classes for which a teacher is responsible, and the degree of team teaching. See OECD (2010a).
[20] See Jensen (2010a) and references therein on p.8
[21] Jensen (2010a) p.9
[22] As was the finding of the much debated Project STAR in Tennessee. See http://www.heros-inc.org/star99.pdf

3. Appraisal and teacher development

Teacher appraisal and feedback matters. Systems of appraisal and feedback that are directly aimed at improving learning can increase teacher effectiveness by as much as 20 to 30%.[23]

Done well, appraisal and feedback makes significant contributions to:

- improving teachers once they enter the profession;
- recognising and rewarding effective teachers; and
- addressing ineffective teachers who have not benefited from improvement programs.

Considerable resources are already devoted to school evaluations, teacher appraisal, and teacher development. Despite these efforts, however, the current system of appraisal and feedback still has some serious shortcomings and has little impact upon teachers' careers or, more importantly, learning and teaching in classrooms. Previous analysis of teacher evaluation in Australia shows that virtually all teachers receive satisfactory ratings and progress along their career structure so that teacher salaries essentially depend on their tenure.[24]

Teachers themselves have identified their desire for better appraisal and development. Central evidence here comes from the TALIS survey which is analysed and discussed in Grattan's report *What Teachers Want: Better Teacher Management*.

Teachers say that:

- Evaluation is not a determinant of progression:
  - *83% of teachers report that the evaluation of their work has no impact on the likelihood of their career advancement.*[25]
- Evaluation often has little impact on how teachers teach and is not linked to development:

[23] Jensen and Reichl (2011)
[24] BCG (2003); Ingvarson (2007)
[25] Jensen (2010c) p.4

- *63% of teachers report that the evaluation of their work is largely done simply to fulfil administrative requirements. 61% of teachers report that the evaluation of teachers' work has little impact on the way they teach in the classroom.*[26]

- Evaluation and development is not addressing ineffective teaching:
  - *92% of teachers work in schools where the school principal never reduces the annual pay increases of an under-performing teacher. 71% of teachers report that teachers with sustained poor performance will not be dismissed in their school.*[27]
- High performing teachers are not recognised. Neither is innovation, nor improvement:
  - *91% of Australian teachers report that the most effective teachers do not receive the greatest recognition.*
  - *92% of Australian teachers report that if they improved the quality of their teaching they would not receive any recognition in their school.*
  - *91% of Australian teachers report that if they are more innovative in their teaching they would not receive any recognition in their school.*[28]

There are ways to improve the appraisal processes for teachers

The Grattan report *Better Teacher Appraisal and Feedback* outlines a new approach to evaluation and development. The system outlined in the report suggests that **teachers should be assessed through a range of methods**. It requires schools to use a range of assessments that draw a direct line to effective teaching and learning.

Schools would choose four of eight methods of appraisal (listed below). Each school would be required to include student test scores and assessments among the four.[29]

1. Student test-scores and assessments;
2. Peer observation and collaboration;
3. Direct observation of classroom teaching and learning;
4. Student surveys and feedback;

[26] Jensen (2010c) p.4 and p.18
[27] Jensen (2010c) p.17
[28] Jensen (2010c) p.15
[29] The extent to which these requirements can be effectively put in place may vary between school sectors. Schools may choose to place less emphasis on self-assessment and parent surveys, given that they often provide little feedback that draws a direct line to improved student performance.

5. 360-degree assessment and feedback;
6. Self-assessment;
7. Parent surveys and feedback; and
8. External observation.

Explanations of how these eight methods work in practice, along with case study examples, are presented in *Better Teacher Appraisal and Feedback.* These eight methods are all focused on improving learning in schools. Mistakes are made in teacher policy when the emphasis is on teachers rather than learning.

Organisation of teacher appraisal and feedback.

There is considerable scope to improve current teacher appraisal and development. Combining multiple assessment methods, and using accurate, unbiased data, has the potential to provide more meaningful and constructive measures of teacher quality. But no system is perfect and no 'one size fits all'.

As discussed in *Better Teacher Appraisal and Feedback*, school principals and teachers should be able to fashion a system of teacher appraisal and feedback to suit the context and direction of their school. Part of the system's success will depend on whether principals and teachers have ownership and responsibility for their own evaluation and development.[30] A decentralised approach is also compelling as principals and teachers have the best information about their school, students and teachers.

Successful implementation of an improved system of appraisal would require a range of steps involving teachers, principals, and governments. The roles of different actors are discussed in detail in *Better Teacher Appraisal and Feedback.*

There have been recent developments related to improving teacher appraisal. The National Professional Standards for Teachers, for example, creates a common language and understanding, which is potentially valuable. However, the standards are not a tool for appraising teachers. While it is expected that the Standards will be linked to performance management processes in schools, it is unclear how they will be used for appraisal purposes. There is a danger that if the Standards are adopted directly as a framework for teacher appraisal, the process will become unwieldy and time consuming. Adopting the Standards as an appraisal framework would require teachers to address all 37 descriptors of their career stage. This information is useful but if used incorrectly can

[30] Caldwell and Spinks (1998)

harm the teaching profession and school effectiveness. Every school is different. It is therefore important that teachers and principals discuss what the national Standards mean for teaching at their school. This will promote conversations about effective teaching and provide teachers with a greater sense of ownership over effective teaching in their school. These conversations should shape teacher appraisal in schools, informing both what is appraised and how it should be appraised.

Some would advocate that the National Standards be linked to centralised accreditation programs and courses. These are not the answer, with recent research showing minimal if no impact on teaching and learning of certification programs such as the National Board of Professional teaching Standards (NBPTS) in the USA.[31] Pushing teacher appraisal and feedback away from the classroom not only ignores but can hinder the learning gains that come from meaningful systems of appraisal and feedback *within* schools.

[31] See For example Goldhaber and Anthony (2007).

4. References

Grattan Reports

Jensen, B. (2010a). Investing in Our Teachers, Investing in Our Economy, Grattan Institute.

Jensen, B. (2010b). Measuring What Matters: Student Progress, Grattan Institute.

Jensen, B. (2010c). What Teachers Want: Better Teacher Management, Grattan Institute.

Jensen, B. and J. Reichl (2011). Better teacher appraisal and feedback: Improving performance. Melbourne, Grattan Institute.

Jensen, B., J. Reichl, et al. (2011). The Real Issue in School Funding: An Analysis of Increasing Government Expenditure and Declining Performance. Article to be published in the Australian Economic Review.

Other References

Barlevy, G. and D. Neal (2011). Pay for Percentile. Working Paper 17194, National Bureau of Economic Research.

BCG (2003). Schools Strategy Workforce Development.

Caldwell, B. J. and J. M. Spinks (1998). Beyond the self-managing school. London, Falmer Press.

DEECD (2009). Measuring Performance Fairly: Developing Intake Adjusted Performance Measures in Victorian Government Schools, Published by the Data, Outcomes and Evaluation Division Office for Children and Portfolio Coordination, Department of Education and Early Childhood Development, Victorian Government, Melbourne.

Ferrão, M. (2007). "Sensitivity of VAM Specifications: Measuring Socio-Economic Status: A Background Report for the OECD Project on the Development of Value-added Models in Education Systems."

Griffin, P. 1990, 'Profiling literacy development: Monitoring the accumulation of reading skills.' Australian Journal of Education, 34 (3): 290-311.

Goldhaber, D. & Anthony, E. (2007) "Can Teacher Quality be Effectively Assessed? National Board Certification as a Signal of Effective Teaching." Review of Economics and Statistics, 89 (1): 134-150.

Goldstein, H., D. Kounali, et al. (2008). "Modelling Measurement Errors and Category Miss-classifications in Multilevel Models." Statistical Modelling **8**(3): 243-261.

Hægeland, T. and L. Kirkebøen (2008). "School Performance and Value-Added Indicators – What is the Importance of Controlling for Socioeconomic Background? A Background Report for the OECD Project on the Development of Value-added Models in Education Systems."

Hanushek, E. (1992). "The trade-off between child quantity and quality." Journal of Political Economy **100**: 84-117.

Hanushek, E. and L. Wößmann (2007). Education Quality and Economic Growth. Washington D.C., World Bank.

Hattie, J. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. Milton Park, UK, Routledge.

Ingvarson, L. (2007). "Research on Performance Pay for Teachers." Educational Research.

Leigh, A. (2010). "Estimating teacher effectiveness from two-year changes in students' test scores." Economics of Education Review **29**: 480-488.

Leigh, A. and C. Ryan (2011). "Long-run trends in school productivity: Evidence from Australia." Education Finance and Policy **6**(1): 105-135.

Mishel and Rothstein (2002). "The Class Size Debate." Economic Policy Institute.

OECD (2010a). Education at a Glance 2010.

OECD (2010b). The High Cost of Low Educational Performance.

Thomson, S. and L. De Bortoli (2008). Exploring Scientific Literacy: How Australia Measures Up, ACER.

Thomson, S., L. De Bortoli, et al. (2010). Challenges for Australian Education: Results from PISA 2009. Melbourne, Australian Council for Education Research.