

# **Measuring and Rewarding School Improvement**

Geoff N Masters

Australian Council for Educational Research

April 2012

A discussion paper commissioned by the Commonwealth Department of Education,  
Employment and Workplace Relations (DEEWR)



# Contents

<b>Preface</b>	<b>v</b>
<b>Executive Summary</b>	<b>vii</b>
<b>1. Introduction</b>	<b>1</b>
1.1 Promoting school improvement	1
1.2 School contexts, school practices and student outcomes	2
1.3 Rewarding school improvement	6
<b>2. Student Outcome Measures</b>	<b>7</b>
2.1 Test-based measures	7
2.2 Unintended consequences	8
2.3 Multiple outcome measures	9
2.4 Comparing school performances	10
2.5 Measuring improvement over time	13
2.6 Outcome-based incentives	15
<b>3. School Practice Measures</b>	<b>17</b>
3.1 Practice-based measures?	17
3.2 Indicators of quality practice	19
3.3 Constructing measures of school practice	21
<b>4. Australian School Improvement Programs</b>	<b>26</b>
4.1 School improvement frameworks	26
4.2 School review cycles	28
4.3 Conclusions	35
<b>5. Principles for a Reward Scheme</b>	<b>36</b>
5.1 Promoting improvements in practice	36
5.2 Rewarding improved outcomes	37
5.3 Rewarding improved practices	38
5.4 Enhancing existing improvement efforts	38
5.5 Encouraging analysis and self-reflection	39
5.6 Ensuring fair comparisons	40
<b>6. A Way Forward?</b>	<b>41</b>
6.1 School self-monitoring	41
6.2 Student outcome measures	43
6.3 Measures of school practice	45
6.4 Existing school review processes	47
6.5 Fair comparisons	48
<b>Bibliography</b>	<b>49</b>
<b>Attachments</b>	<b>52</b>



## Preface

In August 2010 the Australian Government announced a *Reward for School Improvement (RSI)* initiative. Under this initiative, rewards totalling \$388 million over five years will be provided as ‘a strong incentive for schools to review their own performance against national benchmarks and to put in place appropriate mechanisms to improve outcomes for their students’:

This reform represents the next step in Federal Labor’s plan to drive greater school improvement by rewarding principals and school communities who put in the hard work and generate significantly better results for their students. This reform will help create a school system that is geared towards improvement and results.

(Australian Labor Party, 2010)

According to the Government announcement, the *Reward for School Improvement* initiative will provide reward payments to schools that deliver the greatest improvements across a range of areas, including school attendance; literacy and numeracy performance; Year 12 attainment and results; and post-school destination information, such as the number of students going on to further education, training or work.

The first reward payments are to be made to schools in 2015 based on evidence of improvement in 2014. Individual schools will be eligible for improvement rewards up to \$100 000. Reward payments will be made to 500 schools in 2015, and to 1000 schools in each of the following four years. Principals and school communities will be able to determine the best uses of their reward payments to sustain and drive further improvement.

Schools meriting reward payments are to be identified using ‘a transparent school improvement methodology’ based on: (1) an analysis of national data; and (2) school self-evaluations. Clear metrics are to be used to identify the amount of improvement achieved by each school.

In the context of the *Reward for School Improvement* initiative, Minister Garrett announced in November 2011 that a National School Improvement Framework (NSIF) would be developed during 2012, trialled in 2013 and implemented nationally from the start of 2014. The trialling of the Framework may involve school systems determining how to incorporate the NSIF into their existing school review processes.

The current paper was commissioned by the Commonwealth Department of Education, Employment and Workplace Relations to ‘provide advice on:

- i the best ways to measure school improvement; and
- ii the approaches announced in the *Reward for School Improvement* election commitment, specifically:
  - identification of indicators that are supported by empirical evidence as correlating with the quality of education delivery and school/student outcomes on which schools should self-evaluate performance;
  - measurement of school improvement using information gathered from school self-evaluation reports and an analysis of national data; and
  - measurement of school improvement using an analysis of national data only.’



## Executive Summary

There is now a widely held view that the most effective strategy for improving countries' educational performances is to improve the day-to-day work of schools. This view follows several decades of significant increases in government expenditure on school education in developed countries, often with little or no accompanying evidence of improvements in the quality or equity of educational provision.

A number of countries have introduced incentives – both rewards and sanctions – in an attempt to 'drive' improvements in the work of schools. Many of these incentive schemes have followed the model adopted in business of specifying and measuring desired outcomes, holding employees accountable for delivering those outcomes through a system of rewards and/or sanctions, and leaving it to employees to decide on the best strategies to maximise the desired results.

Two decades of experience in the implementation of incentive schemes of this kind in the United States were reviewed recently by the National Research Council's Committee on Incentives and Test-Based Accountability. That committee concluded that the benefits of incentive programs that hold schools accountable for student outcomes had been 'quite small' and that the research evidence on whether such schemes are capable of producing meaningful increases in student achievement was 'not encouraging'.

Equally concerning was the committee's finding that incentive schemes sometimes distort the work of schools. Distortions occur when an appropriate focusing of effort is replaced by practices designed only to maximise outcome measures. In high-stakes contexts, efforts to 'game' outcome measures can include giving greatly reduced attention to non-assessed areas of the school curriculum; assigning the best teachers to the year levels in which assessments occur; withholding less able students from assessments; providing inappropriate assistance to students; and, in extreme cases, cheating. All these responses on the part of schools produce artificial 'score inflation' – apparent gains in student outcomes that are not reflected in performances on low-stakes assessments of the same curriculum content.

At a fundamental level, the theory of action underpinning incentive schemes of this kind is now being questioned. The general intention of these schemes is to reward schools that 'put in the hard work'. But Richard Elmore argues that people in schools already are working 'pretty reliably at the limit of their existing knowledge and skill' and that giving them information about the consequences of their practice will, in general, not significantly improve that practice. Michael Fullan makes the same point, arguing that only a minority of schools know what changes to make to improve their results. In his view, the 'right drivers' in school improvement efforts operate directly on, and change, work practices and cultures in schools.

At the same time, there is growing questioning of the appropriateness of the business model adopted by outcomes-based incentive schemes. Linking rewards to performance may be an effective motivator for some kinds of work, but professionals usually are not left to find or create strategies to maximise outcomes. Rather, advanced professions typically have standards of practice – evidence-based guidelines that the profession itself expects all practitioners to follow when dealing with particular kinds

of situations and problems. The challenge in professional contexts usually is to ensure that every practitioner is implementing current, evidence-based best practice.

These observations suggest caution in introducing outcomes-based incentive schemes of the kind widely used in the United States over the past two decades. If rewards are to be provided, then they must be tied to matters over which schools have direct control. They must make transparent the relationship between rewards and the work of schools and promote highly effective, evidence-based practices. In particular, the basis of rewards cannot be obscure, ‘black-box’ manipulations of outcome measures to produce rank orders of schools. If schools are left wondering why some schools received rewards and others did not, then the credibility of any reward scheme is likely to be undermined.

This paper argues for the use of ‘practice-based’ measures of school improvement to complement ‘outcomes-based’ measures. Given that the ultimate purpose of rewards for school improvement is to improve practices and processes in schools, it seems logical to use evidence of improved practices and processes in making judgements about school improvement.

The difficulty is that there are currently no agreed practice-based measures of school performance or improvement. Considerable effort over many years has been put into constructing reliable measures of student outcomes, measures of contextual factors such as socioeconomic status, measures of ‘inputs’ including schools’ financial resources, and contextualised value-added measures of school performance. But good measures are not yet available of the core work of schools – for example, the quality of classroom teaching or the quality of school leadership. The development of better practice-based measures is essential if discussions of school improvement are to move beyond debates about test scores and their statistical adjustment.

Nevertheless, strong foundations exist for the development of practice-based measures of school performance and improvement. There is a high level of consensus in the research literature on general characteristics of highly effective schools, classroom teaching and school leadership. Australian education systems identify many of these practices in their current school improvement frameworks. Although these frameworks are of variable quality, there is considerable commonality in the practices they identify. Most systems also have regular school reviews to monitor improvements in practice and regular reports on school progress. However, judgements of the work and progress of schools generally are undertaken only for developmental purposes and do not have strong consequences for schools.

This discussion paper uses judgements made as part of one Australian education system’s school review processes to explore the feasibility of developing ‘practice-based’ measures of school performance and improvement. During 2010, each of the 1257 schools in this system was evaluated on eight different aspects of school practice. For each aspect, practices were judged as Low, Medium, High or Outstanding. Preliminary analyses suggest that useful practice-based measures could be constructed from judgements of this kind.

An opportunity – and a challenge – outlined in this paper is to use the *Reward for School Improvement* initiative to support and promote improvements in school practices more directly than is possible by holding schools accountable for outcomes alone. To do this, evidence concerning the work of schools would have to be judged reliably so that conclusions about schools' performances and improvements were credible, fair and comparable across schools. Here, the processes that Australian education systems already have in place may provide a useful starting point. Judgements of the work of schools are already being made through regular school reviews and the ongoing work of line managers (regional directors; district directors; directors school performance; etc). It may be possible to align national rewards for school improvement with existing school review and improvement processes.

### *Principles and Issues*

On the basis of a review of international research and published experience in the areas of school improvement, school accountability and school incentives, this paper proposes a set of principles to underpin the *RSI* initiative. It proposes that rewards for school improvement have as their underlying purpose ongoing improvements in the day-to-day work of schools and, more specifically, that they:

- encourage schools to evaluate and monitor their own ongoing improvement;
- be based in part on evidence of improved student outcomes;
- be based in part on evidence of improved school practices;
- build on and enhance existing systemic school improvement efforts; and
- be based on fair comparisons of schools.

The paper identifies a number of questions that would have to be answered to develop credible metrics for measuring and comparing school improvements over time.

#### *1. School Self-Monitoring*

A first set of questions relate to the role that schools should play in assembling evidence of improvement and in making their own case for an improvement reward. At one extreme, reward decisions could be based solely on an agreed set of national outcome measures and available external evidence about improvements in a school's practices and processes. Under this scenario, there would be little or no direct involvement of schools and reward decisions would be based on automated analyses of agreed metrics. At the other extreme, schools could play a substantial role in reflecting on and assembling evidence of improvement over time and in making a case for reward funding. Under this scenario, in addition to agreed national metrics, schools might provide locally assembled evidence – for example, results on standardised tests or schools' own survey instruments. And, rather than every school being considered for a reward every year, schools might choose when they wished to be considered for a reward, depending on the strength of the case they believed they were able to make in any given year.

#### *2. Student Outcome Measures*

A second set of questions relate to the metrics to be used to measure improvements in student outcomes. Ideally, there would be a core set of national metrics on which all schools (possibly within defined categories such as primary and secondary schools) were compared. A decision is required about what this core set of metrics should be for each school category. A related question concerns the year levels at which these measures should be made (which and how many?). As noted above, there is also a question about whether schools should be able to

supplement these measures with local evidence of improvement and, if so, what requirements should be imposed on these supplementary measures. And finally, decisions are required about how measures of student outcomes would be combined and used in developing defensible measures of a school's improvement over time. These are non-trivial questions that will require careful exploration.

3. *Measures of School Practice*

A third set of questions relate to the metrics to be used to measure improvements in schools' practices and processes. It has been announced that a National School Improvement Framework (NSIF) will be developed and that assessments of the work of schools will be made against this framework. Decisions will be required about the aspects of school practice to be assessed (the framework 'domains'). Careful consideration also will need to be given to the *reliability* of measures of improving school practice. In a context of significant reward funding, self-assessments made by schools are unlikely to be sufficiently reliable unless accompanied by external validation processes. Finally, consideration will need to be given to the most appropriate way of combining evidence of this kind and of using measures of improved school practices in combination with measures of improved student outcomes.

4. *Existing School Review Processes*

A fourth set of questions relate to the relationship between assessments made for the purposes of the *RSI* initiative and assessments made by education systems as part of their regular school review processes. A possible connection is through the National School Improvement Framework. Ideally, there would be significant overlap between the NSIF and the various school improvement frameworks used across Australia. For example, if the domains of the NSIF were core domains of all systemic frameworks, then national rewards for school improvement would be consistent with the criteria used in systems' own school review processes. A related question is whether rewards for school improvement could be aligned with or integrated into current school review processes. For example, the making or validation of assessments for school reward purposes might be undertaken by line managers or other school reviewers who are trained and accredited in making judgements against the NSIF.

5. *Fair Comparisons*

A fifth set of questions relate to processes for ensuring fair comparisons of schools for the purposes of reward funding. Because different measures of improvement are likely to be appropriate in different kinds of schools (eg, primary and secondary schools), it will probably be appropriate that rewards are given within school categories. Decisions will be required about the categories within which rewards are made, and also about the core student outcome measures to be used within each category.

Prior to implementing the *Reward for School Improvement* initiative, further work will be required to explore options in each of these areas.

# 1 Introduction

The quality of a nation's educational provision depends almost entirely on the capacities and efforts of the staff who work in its educational institutions. Most nations have come to a realisation that increases in expenditure and centrally-driven programs alone do not guarantee improved educational outcomes. In many countries, the doubling or tripling of government expenditure on education in recent decades has produced no measurable improvement in educational performances and in some countries has been accompanied by a decline in student achievement levels (Mourshed *et al.*, 2010). There is also a growing recognition that the reason many large initiatives and programs fail in school education is that they do not impact on, and change, day-to-day practice in schools, which itself is now recognised as remarkably impervious to, and self-protective against, fluctuating external policies and agenda.

For these reasons, almost all countries are engaged in efforts to find ways to improve the quality and effectiveness of the work of school-based staff. System-wide improvements in educational outcomes are now seen to depend on *school improvement*, which in turn is seen to depend on building the capacity and effectiveness of those who work in schools.

## 1.1 Promoting School Improvement

Although school improvement is widely viewed as a key to improving educational outcomes, the strategies that education systems use to promote school improvement, and the theories of action that underpin these strategies, differ significantly from one country to another. At one extreme are outcomes-based 'accountability' systems that provide incentives in the form of rewards and/or sanctions on the basis of test results. At the other extreme are 'development' systems that seek to effect improvement by focusing directly on capacity building and on changing workplace practices.

### *Outcomes-Based Accountability Systems*

Outcomes-based accountability is borrowed directly from the world of business. In business, attempts have been made to drive improvement by clarifying the results being sought, developing unambiguous measures of the extent to which those results are being achieved, and then attaching strong incentives to the achievement of results. The theory of action underlying this approach is that when employees are clear about what is expected of them and are rewarded for the achievement of desired results, improvement will follow. There is little focus on *how* improved results are to be achieved. In fact, a perceived advantage of this approach is that it leaves employees free to develop innovative solutions to problems and to find their own ways of maximising the results being sought.

A small number of countries have adopted outcomes-based approaches to school improvement, including Hungary and the United States. In these countries, each school's 'performance' is defined as, and treated as synonymous with, students' results on standardised tests. In England, standardised test results are used to construct measures of school performance which are then used to hold schools accountable to their local communities and to encourage market competition among schools.

As Richard Elmore has observed, the theory of action underpinning outcomes-based accountability schemes in education is identical to the rationale used in business:

Performance-based accountability systems operate on the theory that measuring performance, when coupled with rewards and sanctions... will cause schools and the individuals who work in them, including students, teachers, and administrators, to work harder and perform at higher levels. (Elmore, 2004, 277)

Section 2 of this report reviews experiences and issues in using measures of student outcomes to evaluate school effectiveness and to monitor and promote school improvement.

### *Practice-Based School Reviews*

Although test results are used by a small number of countries to measure school performance, the majority of countries use direct observations of the work of schools – through inspections and school reviews – as the basis for evaluating school ‘performance’ and improvement over time.

In some countries, such as the Netherlands and England, school inspections are primarily concerned with control and accountability of schools and teachers. In England, school inspections contribute to the more general strategy of encouraging competition among schools (Dedering & Müller, 2010). In New York City, Quality Review teams visit, evaluate and rate schools on how well they are organised to improve student achievement. The resulting Quality Review Scores determine schools’ eligibility for rewards or, in extreme cases, consequences such as leadership changes or closure (Pinkus, 2009).

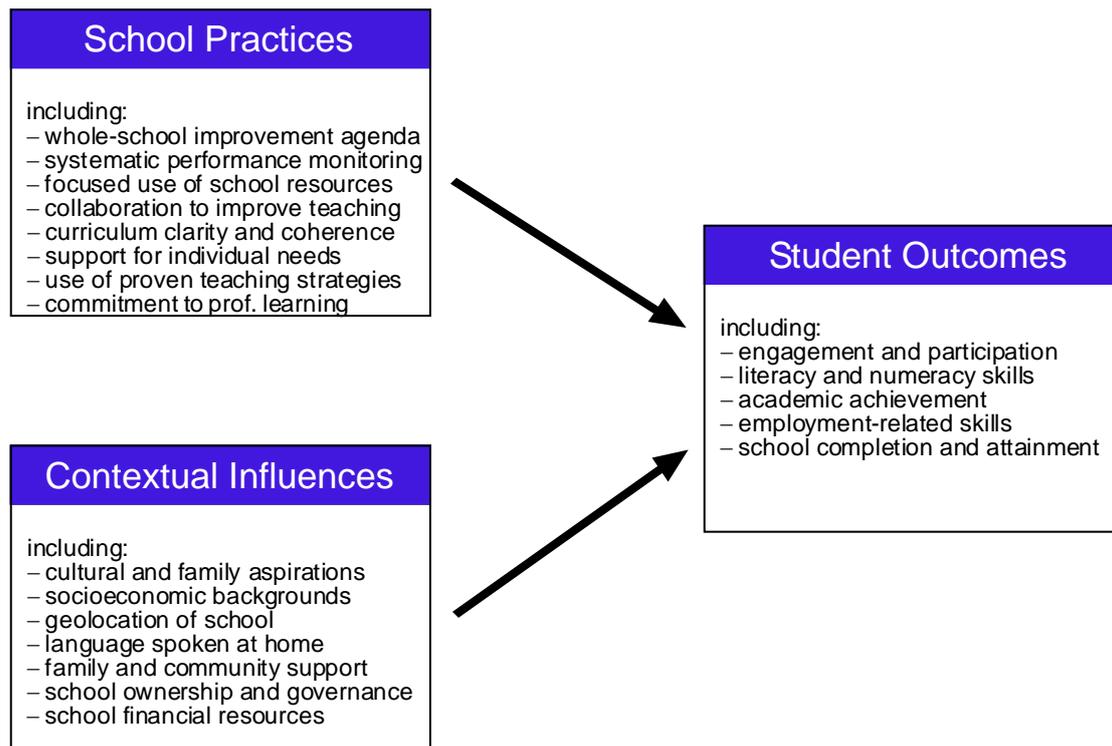
However, in the majority of countries, school reviews are used primarily to provide performance feedback to schools and education authorities for the purposes of ongoing school improvement. This feedback usually is provided directly to schools and may not be made public in its entirety. All of the Australian states have school review processes of some kind in place, as do all federal states in Germany. The theory of action underpinning school review processes is that when areas of school practice in need of further development are internally and externally identified and schools are encouraged and supported to act on this feedback, school improvement will follow.

Section 3 of this paper considers how evidence relating to schools’ practices and processes might be used in measuring school ‘performance’ and monitoring school improvement over time. Section 4 describes current Australian school review processes in some detail.

## **1.2 School Contexts, School Practices and Student Outcomes**

An essential requirement of any attempt to reward schools for improvement is that rewards must be tied to matters over which schools have substantial control. If rewards are to function as incentives for improvement, then there must be a direct and obvious relationship between rewards and the day-to-day work of schools. School staff must believe that rewards fairly reflect their efforts and success. If schools do not perceive a transparent relationship between reward funding and practices and processes under their control, then the credibility and effectiveness of any reward scheme are likely to be seriously undermined.

This is an important observation because schools have limited control over many of the measures that could be proposed to evaluate their effectiveness. For example, the average levels of literacy and numeracy in a school, or the percentages of students achieving national minimum standards in literacy and numeracy, will in part reflect the effectiveness of the school, but also will be influenced by a range of contextual factors including students' home backgrounds and their levels of literacy and numeracy on entry to the school. In general, student outcomes depend both on what a school does (the quality of its practices and processes) and a range of contextual influences, many of which can be thought of as inputs to the school's work (see Figure 1).



**Figure 1. Dependence of student outcomes on school practices and contextual influences**

Because student outcomes are the result of both school practices and contextual influences, it is not possible to infer a school's effectiveness directly from familiar student outcome measures such as mean scores and proficiency levels on standardised tests. Low student outcomes may provide underestimates of school effectiveness; high student outcomes may provide overestimates:

Many schools look much better on performance [outcome] measures than they do upon inspection of their practice. Schools in general, and high-performing schools in particular, produce a large part of their performance with social capital, not with instruction. (Elmore, 2008, 54)

As Raudenbush and Willms (1995) note, in choosing among schools, parents are likely to be interested in the overall quality of education and care that their children will receive. They are likely to look at *contextual* factors such as the quality of a school's facilities and resources and the backgrounds of the children already attending the school; *school practices* such as support for individual learning needs and the quality of the teaching and leadership of the school; and *student outcomes* such as the school's

academic results and the post-school destinations and success of its graduates. Parents are likely to be less interested in separating and understanding what proportion of a school's offerings and effectiveness is due to the efforts of its staff rather than to factors over which the school has little or no control.

In any case, parents may realise that these influences are difficult and sometimes impossible to separate. A school's ability to offer small classes and personalised attention is influenced by its financial resources; quality teaching is easier when parents value education, take an interest in their children's schooling and ensure that they attend school; large schools are able to offer greater subject choice than smaller schools; and so on. It also is not always clear how much some influences on student outcomes – such as school attendance rates, staff qualifications and the students admitted to a school – are under the control of schools.

However, when it comes to *rewarding* schools for their performance or improvement over time, or indeed for holding schools accountable for their performances, an attempt must be made to separate these influences so that schools are rewarded and held accountable for the practices and processes over which they have control. This is a non-trivial problem. Many school systems currently are grappling with this challenge.

This paper considers two approaches to measuring school improvement:

- inferring school improvement from improvements in student outcomes; and
- inferring school improvement from improvements in school practices known to be correlated with student outcomes.

### *Student Outcomes as Evidence of School Improvement*

Because the ultimate purpose of the work of schools is to improve outcomes for students, there is an obvious logic in evaluating schools on the extent to which they succeed in delivering improved student outcomes. In practice, however, the use of student outcomes to evaluate school effectiveness is problematic for two broad reasons.

First, dependable evidence usually is available for only some valued student outcomes. Often the most reliable system-wide evidence takes the form of standardised test scores in a limited range of learning areas (for example, literacy and numeracy). Not only do such measures provide a limited basis for inferring a school's effectiveness, but when incentives are attached, they also have the potential to distort the focus of schools' efforts and, in some cases, encourage inappropriate behaviours to maximise outcome measures without improving student learning. These behaviours include withholding students from testing and providing assistance during tests.

Second, as noted above, student outcome measures inevitably are influenced by factors beyond the control of the school. In statistical terms, contextual influences are 'confounding' factors that make it difficult to infer the quality of school practices from measures of student outcomes. One way of addressing this problem is to use measures that are less influenced by contextual factors in an attempt to measure the value that a school adds over and above its circumstances. For example, the progress that a cohort makes between Year 3 and Year 5 is likely to be less influenced by students'

backgrounds and starting levels of achievement than that cohort's average test score in a given year of school. Another approach to addressing the problem is to attempt to adjust outcome measures to remove the influence of contextual factors. These adjusted measures of student performance, known as 'contextualised value-added' (CVA) measures, are then interpreted as measures of *school* performance.

A difficulty in constructing contextualised value-added measures is in knowing which contextual factors are important and need to be taken into account in adjusting student outcome measures. In some education systems, as many contextual factors as possible are included in the CVA calculation in the belief that this will provide the best estimate of the school's contribution. A second difficulty is that the resulting CVA measures have no substantive meaning. They offer a single measure of each school's 'performance', but there is no transparent way of seeing why one school's CVA score is greater than another's or what a school needs to do to improve its score. A third difficulty is that contextualised value-added measures can mask rather than reveal differences in school performance. For example, if across an education system, the quality of school provision is correlated with socioeconomic background, then standard CVA methods will attribute poorer school practices to the lower socioeconomic backgrounds of students (Raudenbush, 2004).

Section 2 discusses issues and considerations in using student outcome measures to evaluate school performance and improvement over time.

### *School Practices as Evidence of School Improvement*

If schools are to be rewarded for improved performances over time, then there also is logic in basing school rewards on matters over which schools have direct control. These matters include school practices and processes likely to improve student outcomes, including more effective school leadership, higher quality professional development, greater collaboration around improved teaching and learning, more effective teaching strategies, and so on.

A feature of pure outcomes-based accountability systems of the kind implemented in a number of the US states is that they do not collect data on school practices. As noted above, this is consistent with the common business practice of holding employees accountable for results but giving them autonomy to decide how results will be achieved. But the absence of data on school practices means that it is difficult and usually impossible to check the validity of a 'key assumption' of these accountability systems: that school practices are independent of contextual factors, meaning that outcomes-based school 'performance' measures accurately reflect the value that schools add (Raudenbush, 2004).

Section 3 considers the feasibility of developing *practice-based* measures as components of a scheme for rewarding school improvement over time. As Raudenbush and Willms (2004) note, reliable and valid measures of school practice are likely to be much more difficult to construct than measures of student outcomes or measures of socioeconomic status. They also require an explicit theory about the kinds of school practices that are most likely to result in improved student outcomes:

It is far more difficult to adequately measure school practice than to obtain good measures of family background and prior student aptitude or achievement... One

cannot know how effective a school's practice is without a theory of what makes school practice effective. Such a theory and appropriate measures, though difficult to collect, supply a foundation for studying the contributions of school practice. However, even then, the basis for causal inference will be fragile.

(Raudenbush and Willms, 1995, 312-32)

Research into the characteristics and practices of highly effective schools, highly effective school leadership and highly effective classroom teaching provide a starting point in identifying practices that could form the basis of useful measures of school practice. Regular school review processes that operate in many education systems provide further guidance. Some of these external review processes already include external judgements of school practices against pre-specified criteria.

Section 3 discusses issues and considerations in using school practice measures to evaluate school performance and improvement over time.

### 1.3 Rewarding School Improvement

The decision to reward schools for improvement over time raises the obvious question of how school improvement is best measured and recognised. Neither of the forms of evidence considered in this paper – evidence of improved student outcomes and evidence of improved school practices – is ideal as an indicator of school improvement. Available student outcome measures cover a limited part of the work of schools and, when incentives are attached, can be corrupted by behaviours that increase measured results without improving learning and achievement. Reliable and valid measures of school practice have not yet been developed. Such measures would require a theoretical rationale and depend on subjective judgements of schools' practices and processes.

There is also evidence that programs that offer rewards for improved student outcomes have had limited effectiveness in improving those outcomes. A review of two decades of evidence by the US National Research Council's Committee on Incentives and Test-Based Accountability concluded:

The research to date suggests that the benefits of test-based incentive programs over the past two decades have been quite small... The incentive programs that have been tried have involved a number of different incentive designs and substantial numbers of schools, teachers and students... Unfortunately, the guidance offered by this body of evidence is not encouraging about the ability of incentive programs to reliably produce meaningful increases in student achievement.

(Hout & Elliott, 2011, 67)

More generally, research in Psychology shows that incentives can *reduce* intended behaviours if given for behaviours that people would have engaged in anyway. And there is evidence that evaluations of performance are most motivating when they provide specific feedback that points to ways of becoming more effective or competent – without pressure or control (Hout & Elliott, 2011).

Sections 5 and 6 propose an approach to rewarding school improvement using both student outcomes and school practices as evidence of school improvement.

## 2 Student Outcome Measures

### 2.1 Test-Based Measures

Most existing tests were not designed or introduced to provide measures of *school* performance. The purpose of educational tests generally is to provide information about students and their learning. Nevertheless, there has been growing use of tests in recent years to hold schools accountable and to create pressure for improvement through competition, rewards and sanctions.

Some school systems interpret measures of student outcomes as measures of school performance – either directly or after attempting to remove from them the effects of contextual influences and students' prior achievements. In these systems, a school's 'performance' is defined as the performance of its students, typically on standardised tests of English/literacy and mathematics/numeracy.

In the United States, most states and school districts use test scores as measures of school performance. This practice became widespread following the *No Child Left Behind* Act of 2001 which required schools to demonstrate 'adequate yearly progress' in reading and mathematics. More than thirty US states now provide rewards or sanctions to schools based on the test performances of their students. In England, student test scores are interpreted as measures of school performance after being adjusted for a range of contextual influences. Each school's resulting contextualised value-added (CVA) score is assumed to be directly comparable with the CVA score of every other school in the country.

School systems have used a number of different approaches to construct measures of school performance from student test results. In some systems, average test score in a grade is treated as a direct measure of school performance. In other systems, the percentage of students in a grade meeting some pre-specified proficiency standard is interpreted as a measure of a school's performance. Neither of these two uses of test scores takes into account the fact that schools can be differentially effective for students at different levels of achievement, so in a number of US education systems, student results are now reported in four or five proficiency levels (e.g., below basic, basic, proficient, advanced) and schools are compared on the percentage of students in each level (Schwartz, *et al.*, 2011).

Test scores used in these ways generally provide an inadequate basis for measuring and comparing the performances of schools because of the role of contextual influences in determining students' test performances. For this reason, education systems now are making greater use of gain scores (e.g., the average progress that students make in a school between Year 3 and Year 5), growth measures based on more than two time points, or contextualised value-added scores as measures of school performance.

Student test scores also are being used to evaluate schools' contributions to meeting equity targets. This is one of the main objectives of the *No Child Left Behind* Act, which mandates that US states narrow the achievement gap between their highest and lowest performing students (Faubert, 2009).

## 2.2 Unintended Consequences

Changes in the uses to which educational tests are being put – and particularly decisions to attach strong consequences to test results – have influenced the ways in which schools respond to student testing. Educational tests can play a valuable role in clarifying standards that students are expected to achieve, focusing effort on ensuring that all students develop essential skills and knowledge, providing information for educational decision making, and enabling trends to be monitored over time and the effectiveness of initiatives and programs to be evaluated. But there is now significant evidence that the use of tests for high-stakes purposes can distort the work of schools and produce a range of unintended consequences.

Educational tests have the potential to focus teachers' efforts on the outcomes (skills and knowledge) being tested. This can be a positive educational influence – for example, if information is provided about the extent to which students are mastering foundational skills such as reading and numeracy and, as a result, teachers work to ensure that all students achieve adequate levels of these skills. The focusing of teachers' efforts also can be a negative influence – for example, if high-stakes tests result in teachers concentrating on tested subjects or tested content within a subject to the exclusion of other important outcomes that are not tested. There is clear evidence that incentives can encourage surface-level responses on the part of teachers to achieve short-term test gains at the expense of deeper, longer-term improvements in classroom teaching (Hamilton, 2004).

A narrow focus on specific standardised tests can produce improvements in student performances on those tests that are not reflected in other, independent, assessments of the same test content. For example, Koretz and Barron (1998) found that large gains on high-stakes Year 4 reading tests administered by US states between 1992 and 1994 were not reflected in improvements in reading levels on the low-stakes US National Assessment of Educational Progress (NAEP). A number of other studies have found that gains on high-stakes accountability tests do not always generalise to other low-stakes assessments given at approximately the same time in the same subjects (Hout & Elliott, 2011).

Test-based incentives also can encourage other behaviours on the part of schools and teachers. Some schools attempt to maximise test results by assigning better teachers to the year levels in which tests are administered. Teachers in those year levels often report spending more time than teachers in other grades teaching the subjects covered in high-stakes tests. And in systems that reward schools on the basis of the percentage of students meeting specified proficiency levels, there is evidence that teachers target their efforts on students just below the proficiency standard at the expense of students who are well below the standard and those who have already met it. Neal and Schanzenbach (2010) found that the greatest test gains in Chicago schools were made by students in the middle of the distribution near the target proficiency level. The top ten per cent of students showed little or no gain, and the bottom twenty per cent showed no improvement or a decline. Teachers often describe the practices they adopt in response to test-based incentives as inconsistent with their own understandings of good teaching (Pedulla *et al.*, 2003).

Other, still more dubious, school practices can reduce the validity of test scores as measures of school performance. These practices include decisions on the part of schools to discourage less able students from sitting tests. Koretz *et al.* (2001) found

that, in the United States, high-stakes tests occasionally result in various forms of ‘cheating’, including failure to follow test-administration procedures, inappropriate exposure of students to test papers and the alteration of student responses post testing. All of these practices lead to ‘score inflation’ – increases in test scores that do not reflect actual improvements in student learning and achievement. In high-stakes contexts, where rewards and/or sanctions depend on improved student results, and low-performing schools feel pressured to demonstrate improvement, artificial score inflation can provide misleading evidence about actual school improvement:

Teachers and administrators in schools serving poor and minority students may be especially likely to engage in practices designed to raise test scores, including providing extensive test preparation and narrowing the curriculum to focus on tested topics... Without better evidence on the source of score gains among different groups of students, it is impossible to determine whether they represent improved quality of instruction. (Hamilton, 2004, 42)

### 2.3 Multiple Outcome Measures

Tests of basic skills such as literacy and numeracy provide a narrow basis for inferring the effectiveness of a school’s practices or for monitoring improvements in a school’s performance over time. A common criticism of test-based accountability systems is that they fail to take account of other valued areas of the school curriculum. Most tests do not provide evidence about advanced levels of literacy and numeracy, including skills that may be difficult to assess in large-scale tests. They also do not provide evidence of performance in other school subjects such as languages, science, technology and history; performance in non-tested year levels; or the school’s contribution to developing creativity, curiosity, persistence, values, collaboration or socialisation:

Given the broad outcomes that are the goals for education, the necessarily limited coverage of tests, and the ways that indicators constructed from tests focus on particular types of information, it is prudent to consider designing an incentive system that uses multiple performance measures.

(Hout & Elliott, 2011, iii)

There has been growing consensus that the use of a broader range of student outcome measures will improve the quality and fairness of inferences about school practices, provide better guidance to principals and teachers on how to improve schools’ performances, and enhance the acceptability and legitimacy of outcomes-based accountability systems (Chester, 2005; Schwartz et al., 2011).

In practice, some education systems have responded to calls for ‘multiple outcome measures’ not so much by developing measures of a broader range of student outcomes as by developing different measures of the same outcomes or by making repeated measurements of the same outcomes. Other education systems have responded by collecting student and parent perceptions of the quality of education being offered by schools.

Tests of basic skills provide a particularly limited perspective on school performance and improvement in *secondary* schools. Most secondary schools evaluate their own performances not only in terms of literacy and numeracy results, but also in terms of Year 12 results and school completion rates, VET in schools data, student admission

into higher education courses, and successful student transitions into training or employment.

School attendance rates also are important measures for many secondary schools and are predictive of other student outcomes. For example, high school students in Los Angeles who were absent for more than ten days in a school year were only half as likely as other students to complete school (Pinkus, 2009). Some US states have developed composite indicators to establish whether students are on track to complete Year 12. Other outcome measures of interest to secondary schools include work-readiness, the percentage of students completing advanced senior secondary courses, and postsecondary participation and success rates.

Despite efforts to introduce a broader range of outcome measures as part of school accountability and improvement programs, there is currently very little published research or evidence on the quality of these alternative measures or their usefulness as levers for improving school quality or student outcomes (Schwartz *et al.*, 2011).

## 2.4 Comparing School Performances

As noted already, because the ultimate purpose of school education is to improve outcomes for students, there is an obvious logic in attempting to evaluate the performance of a school in terms of the outcomes of its students. In business, a parallel would be to evaluate the performance of a bookstore in terms of the number of books it sells or the profit it makes. A problem with using the most straightforward measures of student outcomes – such as the average level of performance on a test or the percentage of students in a school meeting a particular proficiency level – is that ‘status’ measures of this kind are strongly influenced by contextual factors and students’ prior levels of achievement:

Evidence accumulated over nearly 40 years of educational research indicates that the average level of student outcomes in a given school at a given time is more strongly affected by family background, prior educational experiences out of school, and effects of prior schools than it is affected by the school a student currently attends. (Raudenbush, 2004, 6)

The confounding of status measures by contextual influences makes it difficult to draw inferences about the quality of a school’s practices or to compare one school’s performance with another’s, unless those influences are controlled in some way. One way to do this is to restrict between-school comparisons to schools that operate in similar contexts. Comparisons are more meaningful when restricted to schools that have similar student intakes (eg, socioeconomic backgrounds). Without a restriction of this kind, attempts to infer school performance from status measures inevitably underestimate the performance and contributions of schools serving more disadvantaged students. However, even with these restrictions, it is difficult to be confident that differences in status measures reflect differences in school contributions because many other contextual factors (school finances, rurality, student mobility, etc) influence student outcomes.

The use of status measures to monitor *improvement* over time within the same school is somewhat less problematic because many of the contextual influences are likely to be the same from one student cohort to another. But even here, possible differences in

the background experiences of different cohorts reduce confidence that fluctuations in status measures over time indicate fluctuations in school effectiveness.

### *Gain Scores / Growth Measures*

Another approach to controlling for, or attempting to minimise, the impact of contextual influences is to use measures of student 'gain' (for example, the average improvement in a cohort's reading scores between Year 3 and Year 5) as outcome measures, rather than measures of status at a point in time. By definition, measures of 'gain' are based on two time points; measures of 'growth', on more than two time points. Both kinds of measures require that the same students be assessed on different occasions and that these measures be expressed on the same measurement scale.

Most test-based accountability systems now include measures of gain/growth in place of, or alongside, measures of status. These measures, because they reflect change over time rather than starting levels, are less influenced by student backgrounds and prior levels of attainment and are considered more reliable indicators of the contribution of the school (Schochet & Chiang, 2010). For example, primary schools with different socioeconomic intakes often appear very different on status measures such as mean test scores, but very similar in terms of average rates of progress across the years of school. This may be less true in secondary schools. One US study of mathematics and science test scores found that students in lower socioeconomic areas not only started out behind, but also had lower rates of progress than students in higher socioeconomic areas:

Students entered high-poverty high schools with considerably lower proficiency in math and science than did students entering low-poverty schools. Those differences widened as high-poverty schools displayed lower growth rates than did low-poverty schools. (Raudenbush, 2004, 35)

A further challenge in interpreting gain scores is that they are not always independent of students' starting levels of achievement. As a general rule, children make faster progress in the earlier grades of school than in later grades. This means that, other things being equal, average gains in reading during Year 2 are expected to be greater than average reading gains in Year 5. There is also some evidence that, within the same year of school, students with higher starting levels of achievement tend to make faster progress than students with lower starting levels.

Although gain/growth measures are an improvement on status measures for the purposes of comparing the contributions that schools make to student achievement, they do not necessarily eliminate all confounding by contextual influences. They also do not capture the contributions that schools make to student learning in the years prior to testing (for example, between the commencement of school and Year 3) or post testing (for example, post Year 9), meaning that they may under- or over-estimate a school's actual effectiveness.

Finally, average gains generally are measured less reliably than average status. In other words, there is generally a lower level of confidence about students' true gains. This can be a particular issue in small schools. One way of addressing this problem is to average gain scores across a number of cohorts. However, Schochet and Chiang (2010) found that gains estimated by averaging across three cohorts were still relatively unreliable, with 16 to 21 per cent of schools with truly average gains being identified

as significantly different from average. This misclassification rate only halved when gains were averaged across ten cohorts. In summary, both status measures and gain/growth measures have shortcomings as a basis for measuring school effectiveness and comparing school performances:

Both methods – those based on mean proficiency and those based on [gain scores] – produce estimates with considerable uncertainty and some unknown bias. The logical thing to do in the presence of uncertainty is to seek more information.

(Raudenbush, 2004, 36)

### *Contextualised Value-Added Measures*

Yet another approach to controlling for contextual influences is to attempt to remove these influences statistically from student outcome measures, which are then interpreted as ‘school performance’ measures.

The basic mechanism for creating contextualised value-added measures (CVA) is first to investigate, at the level of an entire education system, the relationships between a range of contextual factors (including students’ prior achievements) and student test scores. These relationships are then used to predict the average test score in each school from the contextual measures for that school. The difference between a school’s actual and predicted average test scores is then treated as a measure of the ‘value’ that the school adds (CVA score).

A perceived advantage of school performance measures constructed in this way is that they allow every school to be compared with every other school on the basis of a single number. However questions remain about whether CVA scores represent an overly simplified approach to measuring school effectiveness and imply unjustified levels of measurement precision.

In England, the contextual measures used to predict a school’s average test results include data on ethnicity, first language, gender, level of special educational need, socioeconomic status (e.g., free school meal), geographical location, date of entry to school, student mobility, absences and exclusions. There has been a tendency to include an ever-greater number of contextual variables:

Month of birth is also taken into account in CVA because by including as many factors as possible outside the school’s control, the residual/difference between the model and the pupil data comes closer theoretically to the school effect.

(Kelly & Downey, 2010, 193)

A difficulty with this approach is that it depends on interpreting residuals (differences between actual and predicted scores in a regression analysis) – a hazardous process at the best of times, but particularly so in high-stakes contexts. A second and related difficulty is that these residuals (CVA scores) have no substantive meaning. As Kelly & Downey (2010) note, in the development of CVA scores in England, ‘slowly the voices in favour of greater complexity began to dominate the stage... Somewhere in the excitement, the importance of access for practitioners was mislaid or underestimated’. In particular, CVA measures are not in a form that can be used for *improvement* purposes. Exactly how a school’s score is derived is not transparent to schools themselves. There is no natural metric that defines how much value-add is appropriate or to be expected and there is no guidance to schools on what they should

do next (Hout & Elliott, 2011). Contextualised value-added measurement seems ‘more aligned to the agenda of public accountability and school performance tables than to the critical process of self-evaluation’ (Kelly & Downey, 2010).

A further difficulty is that, while CVA scores are assumed to represent the ‘value’ that schools add, there is no certainty that this is the case. In practice, CVA scores represent the contribution of the school plus the effects of any other influencing factors that were not included in the prediction process. The CVA process also assumes that the relationships between context variables and test scores are the same for all schools, an assumption that may not be valid and is rarely tested. Finally, as noted earlier, if school practices and contextual influences are positively correlated, the process of constructing CVA scores can mask rather than capture the value that schools add:

If better teachers are able to obtain jobs in schools serving an affluent student population, or if more affluent parents seek the best schools and teachers for their children, demographic and SES variables become proxies for teacher and school quality... If disadvantaged students are systematically assigned to less effective schools and teachers, inclusion of SES as a control can mask genuine differences in school and teacher quality. (Ballou, *et al.*, 2004, 38-9)

It seems clear that attempts to control for the effects of contextual influences on student outcomes – whether by restricting comparisons to schools serving similar student populations; basing decisions on gain scores or growth measures; or by attempting statistically to remove the effects of contextual influences from outcome measures – are only ever likely to be partially successful. Because student outcomes are determined by so many factors, they inevitably provide an imperfect picture of a school’s true effectiveness.

## 2.5 Measuring Improvement over Time

Incentives for school improvement require evidence that school performances have improved over time. If student outcomes are to provide some or all of this evidence, then what is important is that these outcomes have improved. In other words, the focus shifts from point-in-time measures and comparisons of school *performance* (Masters *et al.*, 2008) to measures and comparisons of school *improvement* over time. For example, rather than being interested in the average reading gain that occurs in a school between Year 3 and Year 5 (a performance measure), interest now is in whether the average reading gain between Year 3 and Year 5 increases over time (an improvement measure).

**Table 1. Some Possible Measures of Performance and Improvement**

	<b>Performance</b>	<b>Improvement over Time</b>
<b>Achievement Status</b>	<ul style="list-style-type: none"> <li>• mean score</li> <li>• per cent achieving standard</li> </ul>	<ul style="list-style-type: none"> <li>• change in mean score</li> <li>• change in per cent achieving standard</li> </ul>
<b>Gain/Growth</b>	<ul style="list-style-type: none"> <li>• average gain (e.g., Year 3 to Year 5)</li> </ul>	<ul style="list-style-type: none"> <li>• change in average gain</li> </ul>
<b>Contextualised Value-Add</b>	<ul style="list-style-type: none"> <li>• CVA score</li> </ul>	<ul style="list-style-type: none"> <li>• change in CVA score</li> </ul>

### *Changes in Status Measures over Time*

Many commentators have noted the difficulties in using status measures to compare schools' *performances* unless those measures are controlled in some way for student backgrounds (for example, by restricting comparisons to schools serving similar student populations). However, for the purposes of measuring *improvements* within the same school over time, status measures are somewhat less problematic because many of the contextual factors are the same for different student cohorts. For example, while fluctuations in a school's average Year 5 Numeracy score from one year to the next may simply reflect differences in the backgrounds and prior achievements of the cohorts being compared, longer-term trends in the school's average Year 5 Numeracy score are likely to reflect improved mathematics teaching and learning in the school.

In general, meaningful conclusions about a school's *improvement* depend on evidence from more than two student cohorts. Year-to-year fluctuations, especially in smaller schools, may reflect cohort differences, including factors such as the quality of prior teaching and rates of student attendance and mobility. Related to this is the question of how long it typically takes schools to achieve meaningful improvements. Some improvements in status measures can be achieved relatively quickly through test familiarisation activities and the focusing of teaching effort, but significant, sustainable improvements depend on deeper changes in teaching practice and are likely to occur only over a number of years:

One of the problems in measuring the success of school improvement strategies in individual schools is that improvements in student outcomes will not usually be evident in the first few years. The [Education] Department considers that improving student outcomes could take up to seven years. Given that improvements in student outcomes also need to be sustained, several years of data may be required to know this and to account for student cohort effects.

(Victorian Auditor-General, 2007, 53)

### *Changes in Gain Scores over Time*

An alternative way of evaluating whether a school is 'improving' is to look for improvements in student learning gains. Is average student progress in reading between Year 3 and Year 5 in this school greater now than in the past? There are significant challenges in drawing meaningful conclusions of this kind.

First, simple comparisons of gains can be misleading if they do not also take some account of starting points. For example, student cohort B may make greater average progress between Year 3 and Year 5 than earlier student cohort A because students in cohort B experienced unusually poor teaching in Year 3 and spent the next two years catching up. Under these circumstances, it would be inappropriate to conclude on the basis of the greater gain between Year 3 and Year 5 that the school had 'improved'.

Second, *changes* in gain scores are measured even less reliably than gain scores themselves. Research has consistently found that gain scores are unstable over time, making it difficult, and probably impossible, to draw conclusions about changes in gain scores with any degree of confidence. Only the most extreme changes in gain scores can be identified with any confidence: 'A school in the middle of the distribution can be reliably distinguished only from schools near the extremes' (Raudenbush, 2004).

### *Changes in Contextualised Value-Added Measures over Time*

Changes in contextualised value-added measures are even more difficult to interpret with any confidence because they usually are based on different regression analyses undertaken at different times. In other words, the comparison of CVA scores over time would entail comparing residuals from different regression lines. If CVA scores themselves have no easily understood substantive meaning, then changes in CVA scores as measures of school improvement would be even more difficult to interpret.

In summary, observations that have been made about the difficulties of using status measures to make point-in-time comparisons of school *performance* have led to the development of alternative measures such as gain scores, growth measures and CVA scores. However, these observations about status measures are much less relevant when it comes to making comparisons of school *improvement* over time. This is because cohorts within a school generally will not change significantly from one year to the next. Sustained improvements across multiple cohorts are likely to indicate real improvements in student outcomes. On the other hand, attempts to measure changes in gain/growth scores or CVA scores over time are likely to be unreliable and largely uninterpretable.

## **2.6 Outcome-Based Incentives**

Also of relevance to this discussion of issues in using student outcome measures to reward school improvement is accumulating evidence about the effectiveness of incentive schemes in general. A recent report from the Committee on Incentives and Test-Based Accountability established by the US National Research Council reviewed international research on the impact of test-based incentives in school education. The Committee noted that such schemes have been in place in the United States for two decades, making it possible now to review their effectiveness and to identify any unanticipated consequences of their implementation.

The Committee's review (Hout & Elliott, 2011) included research evidence from economics and psychology on the role and functioning of incentives in human behaviour. A general conclusion from research in economics is that incentives can be effective in focusing effort, but that this focusing also can have a distorting effect:

Organisations often look for performance measures to use in objectively quantifying what each worker is producing. The problem is that these performance measures necessarily focus on the aspects of the job that can be easily quantified and neglect the qualitative aspects of the job that cannot be easily quantified... There are numerous examples of the distortion that results from the use of incentives with performance measures that do not adequately reflect the true value of the work that is being done. (Hout & Elliott, 2011, 7)

Incentives also can encourage behaviours that increase measured performance in a superficial way without changing underlying work practices or providing sustainable long-term improvements in the outcomes being sought. For this reason, most organisations seek and use other evidence about workplace performance, and often base their incentive schemes on subjective rather than objective measures of performance, or some combination of the two.

Research from psychology reveals the complex relationships between incentives and motivation. In particular, it shows that incentives can be de-motivating if they are viewed as part of an attempt at external control or if they are seen as limiting an individual's autonomy. Research in psychology has produced the counterintuitive finding that incentives can reduce the behaviour being rewarded if the reward takes the place of intrinsic motivation and is given for behaviour that a person would have engaged in anyway. Motivation research also shows that evaluations of performance are most effective when they provide feedback on how to improve performance in a non-controlling way (Hout & Elliott, 2011).

In a school context, strong incentives tied to a narrow range of student outcomes (as important as these might be) have the potential to focus effort in ways that distort teaching, including by reducing attention to higher-order outcomes that may be more difficult to measure. As Hout & Elliott (2011) observe, 'it is unreasonable to implement incentives with narrow tests and then criticise teachers for narrowing their instruction to match the tests'. The consequences of such narrowing can be both to overestimate a school's effectiveness and to undervalue areas of the school curriculum that are not tested:

Calculations of the benefits of test-based accountability are likely to be grossly exaggerated if they take test score gains at face value and ignore score inflation and the invisible effects of deemphasising important skills that are not included on the tests. (Hout & Elliott, 2011, 20-21)

There is evidence of score inflation in test-based incentive programs in the United States and limited evidence that these programs have produced significant and sustained improvements in student outcomes. Increases on high-stakes tests with attached incentives have not always been matched by increases on low-stakes tests without incentives.

A common response to the narrowing potential of existing test-based incentive schemes is to propose that incentives be based on a broader range of student outcome measures to better reflect schools' intentions for student learning. Others have proposed the use of student outcome measures in combination with other evidence. For example, Linn (2008) proposes that test results be used as a 'trigger' for the collection of more detailed information about a school's organisational and teaching practices. Evidence of this kind could be used as part of an incentive scheme:

The available evidence does not justify a single-minded focus on test-based incentives as the primary tool of education policy without a complementary focus on other aspects of the system. (Hout & Elliott, 2011, 73)

In summary, because the core work of schools is to improve outcomes for students, evidence of improvements in student outcomes over time will, in general, be a valid indicator that a school is becoming more effective. However, improvements in outcomes may sometimes reflect cohort differences or score inflation resulting from increased test preparation, curriculum narrowing or cheating. When high stakes are attached to test results, school behaviours change. These changes are not always desirable and can result in improved test results without an improvement in underlying student achievements. A solution is to use multiple measures of student outcomes in judging school improvement and to base judgements of improvement on multi-year evidence of sustained trends in outcomes.

### 3 School Practice Measures

A school's 'performance' or 'effectiveness' depends on its practices and processes and, in particular, on practices and processes that result in improved student outcomes. Rewards for school improvement must be tied to matters over which schools have significant control. The purpose of a reward scheme should be to promote and acknowledge school efforts that lead to improved outcomes for students or that have strong potential to produce improved outcomes. This section considers ways of incorporating direct evidence of improvements in *school practices* into rewards for school improvement.

#### 3.1 Practice-Based Measures?

It has to be said at the outset that any attempt to measure school improvement as an improvement in school practices is likely to be difficult. There are no agreed measures of the quality of school practice, the quality of school leadership or the quality of classroom teaching. Much less effort has gone into trying to construct measures of these variables than into measuring student outcomes and socioeconomic status. And part of the reason is that this is more contested territory. There is less agreement on indicators of quality practice than on indicators of quality outcomes. Nevertheless, if the role of practice in school improvement is to be better understood; if differences in school practices are not to be masked by, and attributed to, differences in other variables that are easier to measure; and if discussions of school performance are to be about more than test scores and their statistical manipulation, then some attempt must be made to construct more direct measures of the quality of school practices.

The advent of the school accountability movement is often described as having shifted the focus in education from measures of 'inputs' to measures of 'outcomes'. The inputs referred to usually are contextual influences such as a school's financial resources, the quality of its physical facilities and equipment, the qualifications of its staff, average class sizes, and so on. Most of these 'input' variables are easily quantified and monitored, and most are only weakly related to student outcomes (Cohen *et al.*, 2003).

Consistent with the model commonly adopted by business, this shift to an outcomes focus usually has been accompanied by limited interest in *how* outcomes are to be improved:

The theory of action of [accountability policy] over the past two decades has been based on the idea of holding schools accountable for their outcomes while encouraging local initiative in finding ways to achieve these outcomes. Such a system puts little emphasis on critically examining the quality of organisational and instructional practice. Such a model of accountability relies tremendously on the validity of causal inferences based on the outcome measures.

(Raudenbush, 2004, 37)

Missing from most outcomes-based approaches to measuring school performance and holding schools accountable is an explicit appreciation of the role and importance of *improved professional practice*. Many other professions drive high quality practice throughout the profession by paying close attention to standards of practice. For example, the health professions specify 'standards of care' – guidelines for diagnostic and medical treatment processes that practitioners are expected by the profession to

follow when dealing with particular types of patients, illnesses and clinical situations. These standards of practice are grounded in research and experience and are a significant component of the existing knowledge base of the profession. In developed professions, rather than being left to find or invent their own ways to maximise desired outcomes, practitioners are expected to implement treatments and processes that have been demonstrated by the profession to be effective in practice.

In less developed professions such as teaching and school leadership, standards of practice are as yet largely undeveloped. But there is growing recognition of the desirability of identifying and establishing consensus on practices that characterise and underpin highly effective teaching and school leadership. These 'practices of improvement' (Elmore, 2008) – rather than information on outcomes alone – are seen as key drivers of school improvement:

People in schools are working pretty reliably at the limit of their existing knowledge and skill. Giving them information about the effects of their practice, other things being equal, does not improve their practice. (Elmore, 2008, 41)

High-stakes accountability will only motivate a small percentage of teachers and, even if motivated, only a minority will know what changes to make in instruction to get better results... The right drivers – capacity building, group work, instruction, and systemic solutions – are effective because they work directly on changing the culture of school systems. (Fullan, 2011, 5-9)

It seems logical that a school's performance should be judged, at least in part, on the quality of its practices, and that school improvement should be measured, at least in part, on evidence of improvements in those practices. The questions to which we now turn are: To what extent is there consensus on indicators of quality school practices – and, in particular, practices associated with improved student outcomes? Is it possible to construct meaningful *practice-based* measures of school performance and improvement to sit alongside outcomes-based measures of school performance and improvement? Given the heavy reliance on student outcomes in some school accountability systems, and growing evidence concerning the consequences of this reliance (see Section 2), these are questions that a number of commentators are now beginning to ask:

Most of the policy debate around accountability in recent years has focused on outcomes, but stakes could also be attached to the provision of certain kinds of services or other process-related measures. For example, an indicator system might be designed explicitly to incentivise particular kinds of practices, such as the use of data from formative assessments to alter instruction.

(Schwartz, 2011, 7)

In Michigan, one-third of a school's performance measure is now based on ratings of identified aspects of the school's practices and processes.

Practice-based measures of school performance and improvement, if they can be meaningfully constructed, have the potential to provide schools with feedback to guide further improvement. A criticism of most outcomes-based measures is that they are not actionable and, after being converted into derived measures such as contextualised value-added scores, are uninterpretable and potentially de-professionalising. Practice-based measures also may assist in the interpretation of student outcome measures. For

example, an increase in a school's mathematics scores could be evidence of improved mathematics teaching, but equally could reflect curriculum narrowing and score inflation. 'If this increase were accompanied by evidence of improvements in the quality of teaching as measured by direct observations of instruction, the user of the information might be more confident that an inference of improved instructional quality is warranted' (Schwartz, 2011, 27).

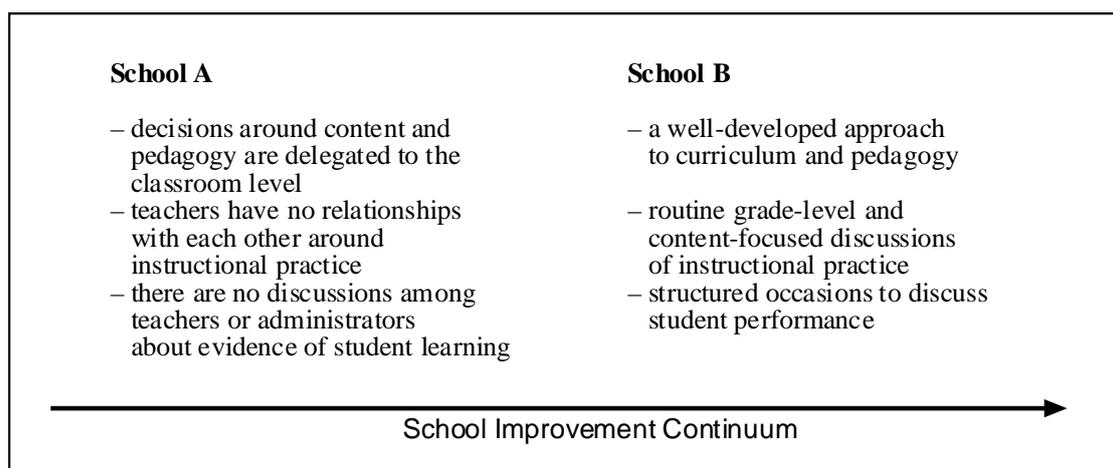
Identifying evidence-based best practices is far more difficult than holding schools accountable for outcomes alone. But this hard work appears essential if schools are to be held accountable in ways that are scientifically defensible, fair, and effective. A mix of evidence based on outcomes and assessments of practice appears essential if accountability is to achieve its potential to improve schools.

(Raudenbush, 2004, 37)

### 3.2 Indicators of Quality Practice

The 'hard work' of building practice-based measures of school performance begins with the identification of indicators of quality practice. There is now a large body of research into effective schooling, quality teaching and effective school leadership. This body of research has identified both general characteristics of effective schools and specific classroom practices that are correlated with improved student learning and achievement (Hattie, 2003). Among the identified characteristics of highly effective schools are a positive learning climate, a strong student achievement orientation, an implemented curriculum of high quality, professional staff collaboration, high quality school leadership, and regular assessments and evaluations of progress (Dedering & Müller, 2010).

Implicit in the notion of school improvement is the belief that schools can continually improve their practices and processes. The assumption is that, no matter where a school is on its improvement journey, it is capable of becoming still more effective. Elmore (2008) illustrates this point by reference to some observable differences between two schools at different stages in their development (Figure 2).



**Figure 2. Two schools at different points on a school improvement continuum (Elmore, 2008)**

We could imagine a number of different points on the continuum, but these two are enough to illustrate the main issue: school improvement is a developmental

process... Like most developmental processes, this one involves more or less predictable stages. Moving a school through these stages requires, first, an understanding that there is a developmental process going on; and second, an understanding of what distinguishes schools at one stage of development from another. (Elmore, 2008, 46)

These understandings – that school improvement is a developmental process, that schools are at different stages on a continuum of development, and that each school's current level of development can be inferred from 'indicators' of progress – are fundamental requirements for the construction of practice-based measures of school performance and improvement.

### *Indicators of Quality Practice in School Reviews*

Many education systems have established school review processes that involve regular visits to schools to examine and judge the quality of their work. These reviews usually include judgements against specific criteria or aspects of practice such as the school curriculum, relationships between staff, and the quality of teaching. For example, following Germany's performance in the Programme for International Student Assessment (PISA) in 2000, all sixteen German states introduced initiatives focused on school improvement, including regular school inspections to evaluate school environments, the quality of teaching and leadership, and student outcomes. The focus of these visits is on supporting ongoing school improvement (Dederig & Müller, 2010). In England, OFSTED provides criteria which are used to evaluate school practices and student outcomes in all school inspections (Faubert, 2009). Quality review teams in New York City visit schools to observe and evaluate classroom teaching, the school's use of data in planning, the quality of school programs and their implementation, and staff collaboration (Pinkus, 2009).

Australian education systems also have introduced regular school improvement and accountability reviews. These processes are described in some detail in Section 4 of this paper. In common with school inspections and reviews in other countries, there is considerable consistency in the kinds of school practices and improvements that are looked for in these external reviews. Not surprisingly, these indicators also tend to be consistent with research findings on effective schools.

School reviews often include school self-assessments against external review criteria. In many countries, schools assemble evidence of student outcomes and school practices and make judgements about their own performances and improvement in preparation for school visits. In some countries, self-assessments are used by external reviewers to make a preliminary appraisal of a school before it is visited (Faubert, 2009). An advantage of school self-assessments is that they encourage schools to engage with the criteria (indicators) and to reflect on their own practices and improvements.

### *An Example: Teaching and Learning School Improvement Framework<sup>1</sup>*

One example of a school review process that includes the examination of documentary evidence and observations made during school visits is the teaching and learning 'audit' conducted by the Queensland Department of Education and Training. The audit process commenced in 2010 when all 1257 government schools were visited by trained 'auditors' and evaluated against the eight domains of the Teaching and Learning School Improvement Framework (TaLSIF).

The TaLSIF Framework (Masters, 2010) draws on research into the practices of highly effective schools and school leaders. The focus of the Framework is not on everything that effective schools and school leaders do, but on those leadership practices that appear to be most directly related to school-wide improvements in teaching and learning. The Framework consists of eight interrelated ‘domains’ representing different aspects of a school’s day-to-day work. These domains are summarised in Figure 3.

On the basis of observations made during school visits, trained auditors make a judgement (or rating) in relation to each of the eight domains. For each domain, a school’s practices are rated as ‘Low’, ‘Medium’, ‘High’ or ‘Outstanding’. The indicators that help define these rating categories are shown in Attachment 1. Before reaching a rating, auditors meet with senior staff of the school to discuss school practices, inspect records and evidence provided by the school and make their own observations through conversations with classroom teachers, students and some parents. Auditors undergo intensive training to maximise the comparability of their ratings across schools. In addition to the eight ratings, each school is given a written report that includes affirmations, commendations and recommendations concerning the school’s practices.

### 3.3 Constructing Measures of School Practice

An interesting question is whether it is possible to construct from judgements of school practices a meaningful ‘measure’ of each school’s performance. As noted above, much less effort has gone into developing practice-based measures than outcomes-based measures of school performance. But if valid and reliable practice-based measures could be constructed, then these measures may provide useful complementary evidence about a school’s performance and improvement over time. In the absence of practice-based measures it is likely that school improvement will continue to be defined and measured solely in terms of more easily measured student outcomes.

To address this question, an attempt has been made to aggregate judgements on the eight domains of the Teaching and Learning School Improvement Framework (TaLSIF) to produce an overall measure of performance for each government school in Queensland in 2010. When the TaLSIF was first developed, there was no intention to aggregate judgements; rather, the focus was on how each school was performing in relation to each domain. However, for present purposes, it is interesting to consider whether useful practice-based measures could be constructed in this way.

As a first step, the four rating categories were assigned scores: Low (1), Medium (2), High (3) and Outstanding (4). These scores were then added across the eight domains so that each school received a total score between  $8 \times 1 = 8$  and  $8 \times 4 = 32$ . The full set of data for the 1257 Queensland schools was then analysed using a measurement model for ordered categories (Masters, 1982) to explore the relationships among the eight domains and to establish the validity of simply adding domain scores in this way. The results of this analysis showed – with one exception – that ratings on the eight domains could be meaningfully combined into a single performance measure. The exception was Domain 6: Systematic Curriculum Delivery. Judgements on this domain were less strongly correlated with judgements on the other domains; more low-performing schools were rated Medium, and more high performing schools were rated Low on this domain than predicted from their ratings on the other seven domains. There may be good substantive reasons for retaining Domain 6 in a measure of school performance, however for present purposes, it has been removed.

**1 An Explicit Improvement Agenda**

The school leadership team has established and is driving a strong improvement agenda for the school, grounded in evidence from research and practice, and couched in terms of improvements in measurable student outcomes, especially in literacy, numeracy and science. Explicit and clear school-wide targets for improvement have been set and communicated, with accompanying timelines.

**2 Analysis and Discussion of Data**

A high priority is given to the school-wide analysis and discussion of systematically collected data on student outcomes, including academic, attendance and behavioural outcomes. Data analyses consider overall school performance as well as the performances of students from identified priority groups; evidence of improvement/regression over time; performances in comparison with similar schools; and, in the case of data from tests such as NAPLAN, measures of growth across the years of school.

**3 A Culture that Promotes Learning**

The school is driven by a deep belief that every student is capable of successful learning. A high priority is given to building and maintaining positive and caring relationships between staff, students and parents. There is a strong collegial culture of mutual trust and support among teachers and school leaders. The school works to maintain a learning environment that is safe, respectful, tolerant, inclusive and that promotes intellectual rigour.

**4 Targeted Use of School Resources**

The school applies its resources (staff time, expertise, funds, facilities, materials) in a targeted manner to meet the learning needs of all students. It has school-wide policies, practices and programs in place to assist in identifying and addressing student needs. Flexible structures and processes enable the school to respond appropriately to the needs of individual learners.

**5 An Expert Teaching Team**

The school has found ways to build a professional team of highly able teachers including teachers who take an active leadership role beyond the classroom. Strong procedures are in place to encourage a school-wide, shared responsibility for student learning and success, and to encourage the development of a culture of continuous professional improvement that includes classroom-based learning, mentoring and coaching arrangements.

**6 Systematic Curriculum Delivery**

The school has a coherent, sequenced plan for curriculum delivery that ensures consistent teaching and learning expectations and a clear reference for monitoring learning across the year levels. The plan, within which evidence-based teaching practices are embedded, and to which assessment and reporting procedures are aligned, has been developed and refined collaboratively to provide a shared vision for curriculum practice. This plan is shared with parents and caregivers.

**7 Differentiated Classroom Learning**

In their day-to-day teaching, classroom teachers place a high priority on identifying and addressing the learning needs of individual students. Teachers closely monitor the progress of individuals, identify learning difficulties and tailor classroom activities to levels of readiness and need.

**8 Effective Teaching Practices**

The school principal and other school leaders recognise that highly effective teaching practices are the key to improving student learning throughout the school. They take a strong leadership role, encouraging the use of research-based teaching practices in all classrooms to ensure that every student is engaged, challenged and learning successfully. All teachers understand and use effective teaching methods – including explicit instruction – to maximise student learning.

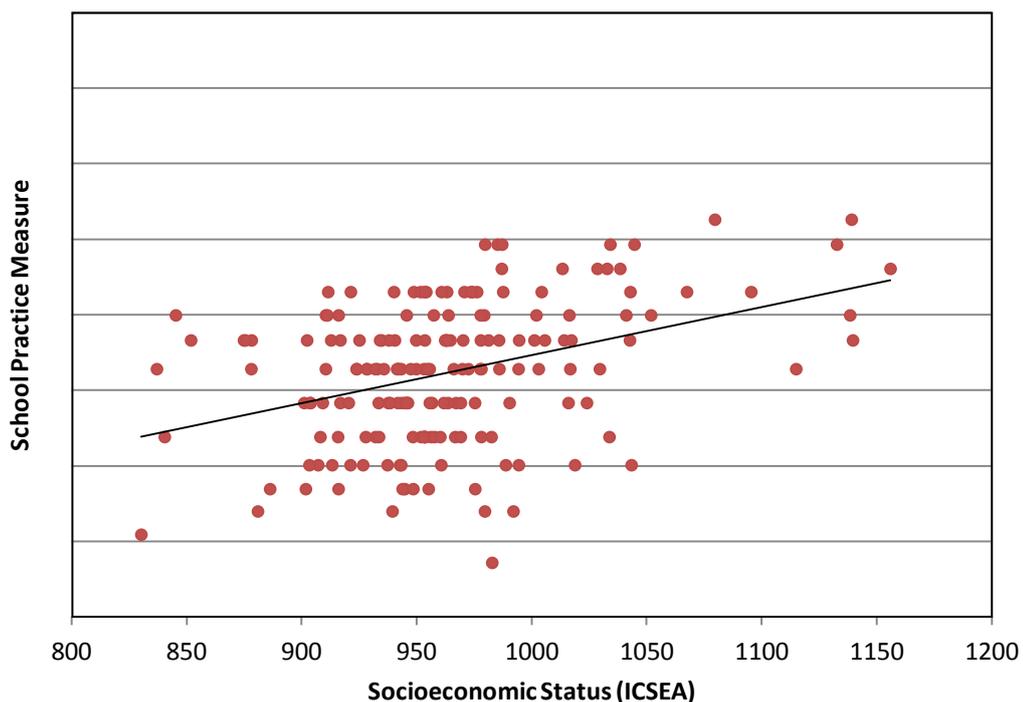
**Figure 3. Domains of the *Teaching and Learning School Improvement Framework***



improvements in one domain sometimes occurring at the expense of another; and – not surprisingly perhaps – least improvement in Domain 8: Effective Teaching Practices.

Judgements made in the same school on different occasions could be used to measure that school's improvement over time (that is, measured progress up the continuum in Figure 4). As multi-year data became available for all schools, improvements over time could be compared across schools and interpreted in the light of statistical confidence intervals. In this way, practice-based measures of school improvement might be used to complement outcomes-based measures.

If meaningful practice-based measures of school performance can be constructed, then it becomes possible to study not only differences in these measures over time and across schools, but also their relationships with other relevant variables. For example, Figure 5 plots the practice-based measures constructed for Queensland secondary schools against a measure of socioeconomic status (ICSEA). The graph shows a weak positive correlation (about .4), meaning that about 15 per cent of the variance in school practice measures in these secondary schools is 'explained' by socioeconomic background. This observation raises the possibility that a percentage of the variance in student outcome measures commonly attributed to students' socioeconomic backgrounds may in reality be due to differences in the quality of school practices.



**Figure 5. School practice measure versus socioeconomic status for all Queensland government secondary schools in 2010**

This section began with the observation that the purpose of a reward scheme for school improvement should be to encourage improvements in matters over which schools have control – namely, their practices and processes. Research into effective schools, school leadership and classroom teaching has identified a number of general characteristics of highly effective practice. These practices increasingly are being

incorporated into professional learning programs for teachers and leaders and are forming the basis of education systems' school reviews and inspections. If evidence relating to a school's practices were collected and evaluated on different occasions, then rewards for school improvement could be based not only on evidence of improved student outcomes, but also on evidence of improvements in the quality of that school's practices.

The analysis of the Queensland audit data illustrates how this might be done. For the purposes of a national rewards scheme, each school could be rated on a small number of domains of practice. Each school's performance in each domain could be rated as Low, Medium, High or Outstanding using nationally agreed definitions (eg, Attachment 1). These ratings could then be aggregated to provide a practice-based measure of a school's performance at a point in time. Subsequent re-assessments – of the kind that occurred in Queensland government schools in 2011 – would provide a basis for studying *improvements* in these practice-based measures over time.

Questions that would have to be addressed in implementing practice-based measures of school improvement are: How feasible is it to reach national agreement on a set of key domains of school practice? And how could reliable, practicable and nationally consistent judgements of schools' practices be made? In seeking answers to these questions, it is instructive to examine how Australian education systems currently monitor school performance and improvement over time.

<sup>1</sup> The permission of the Queensland Department of Education and Training to use data from the 2010 audit of government schools in this section is gratefully acknowledged.

## 4 Australian School Improvement Programs

Most, if not all, Australian education systems have established their own school improvement programs. The purpose of these programs is to describe and promote highly effective school practices and to encourage ongoing school improvement. Common elements of these programs are a system-wide school improvement framework; a school improvement/review cycle (usually 4-year); annual school reports to local communities on progress in implementing improvement strategies; and a formal end-of-cycle review that includes a school visit and associated report.

Australian school improvement programs usually include explicit statements of underpinning principles. These statements make clear that the purpose of school improvement is to improve outcomes for students; that improvement depends on improvement strategies and changed school practices; that whole-school commitment, professional learning and critical reflection on current practice are essential to change; and that improvements in the quality of classroom teaching and school leadership are key drivers of improved student outcomes. The following excerpts from three of Australia's smaller school systems illustrate these beliefs:

School improvement is founded on the belief that active internal debate about educational issues and their impact on student learning by those entrusted with the provision of learning and teaching will enrich the school's understanding of its need for change... The goal of school improvement processes is to enhance student learning by engaging in improvement strategies that draw from the critical reflection of the school community as it responds to the questions: What are we doing well? What do we need to improve? How can we bring about improved learning opportunities for our students?

(Tasmanian Catholic Education Commission, 2011, 4)

Effective school improvement is about change that is driven by commitment to increase the learning outcomes of every student. It requires a supportive environment where relationships and trust are developed. While the critical role of quality teaching in the pursuit of school improvement is irrefutable, the capacity of the school leadership team to build a professional learning environment and lead the change process is equally important.

(ACT Department of Education and Training, 2009, 1)

School improvement requires a professional commitment to problem solving, innovation, critical reflection and continuous professional learning. School improvement is reliant on school leadership that engages people at all levels of the organisation in the learning process by creating a culture of inquiry which develops new capabilities and revolutionises teaching and learning. Strategies must work from the classroom out and have a bottom line of enhancing student progress, achievement and development.

(Northern Territory, 2010, 2)

### 4.1 School Improvement Frameworks

Most Australian education systems have developed school improvement and accountability 'frameworks' to guide school improvement efforts. These frameworks go by various names, but have many features in common. Some of these frameworks are listed in Table 2.

**Table 2. Examples of School Improvement Frameworks**

Australian Capital Territory Department of Education and Training	<i>School Improvement Framework</i>
New South Wales Department of Education and Training	<i>Analytical Framework for Effective Leadership and School Improvement in Literacy and Numeracy</i>
Northern Territory Department of Education and Training	<i>School Accountability and Performance Improvement Framework</i>
Queensland Department of Education and Training	<i>School Planning, Reviewing and Reporting Framework: An improvement and accountability framework for Queensland State Schools</i>
South Australian Department of Education and Children's Services	<i>Improvement and Accountability Framework</i>
Tasmanian Catholic Education Commission	<i>School Improvement for Catholic Schools in Tasmania: School Improvement Framework and Process</i>
Victorian Department of Education and Training	<i>Accountability and Improvement Framework for Victorian Government Schools</i>
Western Australian Department of Education and Training	<i>The School Improvement and Accountability Framework</i>

These frameworks generally identify 'areas' or 'domains' of school practice as priorities for ongoing improvement. Some of these areas/domains are shown in Attachments 2 to 7. Education systems list different numbers of priority areas, but there is a very high level of commonality of content. The following five areas are identified as priorities in all examined frameworks:

#### *Leadership*

This includes establishing a clear vision and direction for the school; promoting a collective sense of responsibility for improving student learning and outcomes; monitoring school performance data; encouraging reflective practices; providing hands-on instructional leadership; and managing and targeting school resources to improve outcomes for students.

#### *Learning*

This includes promoting a belief that all students can learn successfully; enhancing student motivation and engagement; setting explicit and high standards for student learning; exploring and understanding students' starting points and learning needs; and providing flexible programs and pathways to address individual needs.

#### *Teaching*

This includes driving high quality teaching practices; making decisions about what should be taught and ensuring curriculum sequencing and coherence; assessing and monitoring student learning and providing feedback to students and parents; and professional learning focused on continual improvements in teaching effectiveness.

### *Learning Environments*

This includes ensuring safe and supportive learning environments; maintaining positive and respectful relationships; creating classroom cultures in which learning is a priority for all; and celebrating student progress and success.

### *Partnerships*

This includes partnering with parents and families to improve outcomes for students and establishing relationships with local community organisations and businesses in the interests of improved student outcomes.

In the ACT, the intention is that government school leaders and school communities will reflect on the quality of their practices and conduct an annual self-assessment against the framework domains. In NSW, schools undertake a self-assessment against the dimensions and twenty-five statements of the DET analytical framework. In the Northern Territory, schools are expected to monitor and measure school performance against the areas of their framework and to provide an annual performance report to their communities. The frameworks in QLD and SA describe the areas in which schools are expected to work to improve their practices. In VIC, eight aspects of school practice are identified in an 'effective schools model' to assist schools in their evaluation of current school practices and their design of improvement strategies. And in WA, five areas of school practice are identified and schools are encouraged to monitor the effectiveness of their processes and procedures in these areas.

School improvement relies on having sound measuring, monitoring and reporting processes in place for each of the domains (learning and teaching; leading and managing; student environment; community involvement).

(Australian Capital Territory Department of Education and Training, 2009, 3)

## **4.2 School Review Cycles**

Government education systems typically have a four-year school review cycle. Schools are expected to develop a four-year strategic improvement plan that includes goals and associated strategies to improve performance in the areas/domains of practice identified in the system's school improvement and accountability framework.

In the context of their four-year strategic improvement plans, schools usually are expected to develop an annual operating plan that 'describes how the key improvement strategies in the school strategic plan will be put into operation in the following year and how they will be monitored' (Victorian Department of Education and Early Childhood Development, 2011). Schools self-assess and report on progress in an annual report to their local communities. These self-assessments are based on 'evidence related to student achievement and school operations' (Western Australian Department of Education and Training, 2008).

It is common in the fourth year of the cycle for systems to undertake an external review of each school's progress. The report of this review informs the school's planning for the next four-year cycle.

### *Student Outcome Measures*

As part of school improvement processes, schools are expected to set targets for, and to monitor, improvements in student outcomes. Schools are expected to use a range of

student outcome measures to monitor improvement over time; evaluate the effectiveness of their school improvement strategies; and monitor progress towards school, system and national targets for improvement, including equity targets.

Victoria identifies three broad areas in which schools are expected to assess and monitor improvements in student outcomes:

*Student Learning*

Student learning outcomes relate to what students know and can do. All schools develop curriculum and report on achievement in student learning outcomes using a number of frameworks and data sources, including NAPLAN, Victorian Essential Learning Standards, the Victorian Early Years Learning and Development Framework, and senior secondary certificates and programs.

*Student Engagement and Wellbeing*

Student engagement and wellbeing outcomes relate to the extent to which students feel safe, secure and stimulated to learn at school. Measures of student engagement and wellbeing include student attendance, students' ratings of their connectedness to school, motivation to learn or safety (Student Attitudes to School Survey), teacher assessments of student motivation (Staff Opinion Survey) and parents' perceptions of school climate (Parent Opinion Survey).

*Student Pathways and Transitions*

Student pathways and transitions outcomes relate to the quality of student transitions into, through and out of a particular educational context and the extent to which pathways meet the needs of students.

(Victorian Department of Education and Early Childhood Development, 2011, 5)

System documents identify a range of data sources that schools might use to assess and monitor improvements in student outcomes. These include:

- Year 3, 5, 7 and 9 NAPLAN tests
- other commercially available tests
- Year 12 subject results and tertiary entrance ranks
- average student attendance rates
- apparent student retention rates
- surveys of parent, teacher, student perceptions/satisfaction
- post-school destination information
- system-endorsed tools (e.g., First Steps; Literacy and Numeracy Net)

The Western Australian Department observes that measurable improvements in student outcomes often occur only over extended periods of time, making it necessary to persist with improvement strategies and to look for other evidence that progress is occurring:

It is recognised that some strategies may take longer to show improved student achievement than the typical school planning cycle. In such cases, the strategy should continue to be monitored for its effect on student achievement. To do this, schools may need to develop specific milestones that would provide interim evidence that the strategy is on course and should be continued.

(Western Australian Department of Education and Training, 2008, 8)

### *School Self-Assessments*

An essential element of jurisdictions' school improvement programs is schools' evaluations of their own performances and improvement over time:

The professional reflective process of self-evaluation is at the heart of effective school improvement. (Tasmanian Catholic Education Commission, 2011, 10)

School self-evaluation... includes examining teaching and learning strategies, the performance and development culture and other aspects of school operations so they can be strengthened and supported to improve student outcomes. (Victorian Department of Education and Early Childhood Development, 2011, 7)

Self-assessment is fundamental to establishing an effective school improvement cycle... This involves gathering and analysing data and other evidence and making judgements about the standards of student achievement and the effectiveness of school processes and operations. (Western Australian Department of Education and Training, 2008, 6)

Some school systems provide detailed guidance to schools in conducting self-assessments. The Teaching and Learning School Improvement Framework (Attachment 1) is used by many Queensland government schools to evaluate their own current practices. Schools use this framework to make judgements – separate from the external audit process – about whether their performance is Low, Medium, High or Outstanding in each of the eight domains of the framework.

The ACT Department provides a detailed self-assessment matrix based on the domains and elements of its school improvement framework (Attachment 2). The self-assessment matrix provides schools with indicators that describe 'progressive stages in the improvement process' along a 'five-step continuum' (see excerpt in Table 3). The Department considers that these indicators 'also have an aspirational function in that they provide direction to schools by identifying the qualities exhibited by high performing schools. They can assist schools to gauge their stage of development, set goals and track their progress' (ACT Department of Education and Training, 2009).

The Tasmanian Catholic Education Commission's framework has twenty-two 'elements' (see Attachment 5) and provides schools with indicators of outstanding practice in relation to each element. The indicators for the element 'Leading Continuous Change' are shown in Table 4. Schools are encouraged to use these indicators in the evaluation of their current practices and improvement over time, as well as to 'stimulate staff discussion and stretch understanding of what actions might be needed to improve practice':

Evaluations of quality need to be grounded in evidence. So data from observations or canvassing people's opinions or from surveys (or other quantitative strategies) are needed to substantiate the 'quality' judgements being made about the particular practice. (Tasmanian Catholic Education Commission, 2011, 19)

**Table 3. Excerpts from the ACT School Self-Assessment Matrix**

	<b>Expectations about learning are communicated so as to value the quality of work and high standards.</b>	<b>Student learning is personalised to engender motivation, ownership and responsibility.</b>
<b>5</b>	All staff members have high expectations for all learners, with standards discussed, challenged and improved. It is clear to students how these criteria will be used in assessing their work that results in high motivation to achieve.	Students' prior knowledge is always identified and substantially incorporated into the teaching programs in a meaningful way. All students are deeply involved in their learning, taking responsibility for their learning almost all of the time.
<b>4</b>	Most staff members have high expectations for all learners with practices and programs in place to support learners achieve raised standards. Students fully understand what it means to do well.	Nearly all teachers consider students' Background and cultural knowledge when developing their teaching program. Students are broadly engaged in the substance of the learning programs. They take ownership and responsibility for their learning.
<b>3</b>	Staff members have high expectations for learners and put in place practices and programs to support individuals and groups to work towards raised standards. There is little elaboration of what it means to do well.	Teachers often consider the students' background and cultural knowledge when personalising programs. Student engagement is evidenced through a rise in motivation across a range of subjects. Students usually own the results of their learning.
<b>2</b>	Some staff members have high expectations for learners but this is not widely agreed or programmed. Students are unsure of expectations and standards so there is no motivation for raising achievement or performance.	Most students, most of the time, appear indifferent to learning and achieving. Students' background and cultural knowledge is occasionally reflected in teaching programs. Teacher directed learning is the dominant pedagogy.
<b>1</b>	Staff members do not have high expectations for all learners and believe some learners will not achieve. No explicit statements regarding the quality of work or standards are communicated to students. Only technical and procedural criteria are made explicit.	Students' background and cultural knowledge is not considered when programs are planned. Students are generally disengaged and passive participants in the learning process.

**Table 4. Indicators of High Quality for the Element ‘Leading Continuous Change’ (Tasmanian Catholic Education Commission, 2011)**

School leaders:

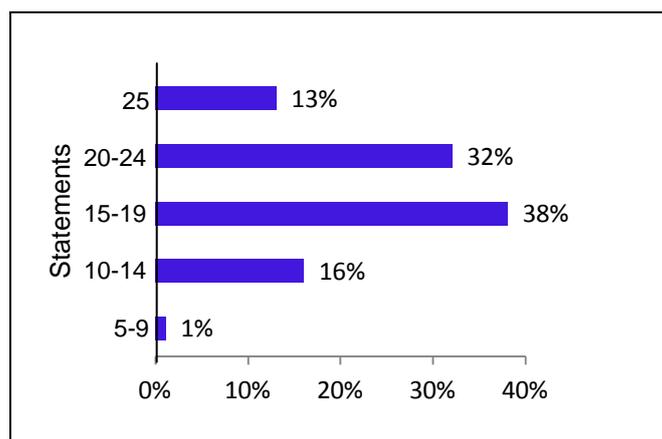
- inspire, and support others in developing a common vision, mission and strategic direction, with long-term goals for how the school can improve over the planning cycle, and a clear focus on improving student learning;
- stimulate and provide opportunities for discussion and debate on broad educational matters;
- encourage and nurture future-focused thinking, creativity and innovation;
- collaboratively develop a cohesive improvement plan that combines mandated government and system priorities with goals derived from the school community’s engagement in ongoing school improvement processes;
- support staff to develop and implement annual action plans, outlining clear evidence-based goals, strategies, indicators of success, key personnel and resources, to ensure the school’s strategic directions and goals are achieved;
- ensure continuity in school improvement by prioritising and fostering a culture of ongoing inquiry into what is working well, what needs to be improved, and what, if improved, would have a strong positive impact on student learning;
- engage in constructive ‘problem talk’ that names, accurately describes and analyses problems as they arise in order to identify possible actions for improvement; and
- engage the school community in a variety of evaluative strategies to ascertain how well the school is achieving its goals.

For each of the twenty-five statements in the NSW Department’s Analytical Framework (Attachment 7), four ‘levels of practice’ have been defined (New South Wales Department of Education and Communities, 2011). Schools using the Analytical Framework self-assess against these four levels. The four levels of practice for Statement 4, ‘There is clarity and consensus about the school’s goals and expectations in literacy and numeracy’, are shown in Table 5.

Of 125 NSW government schools that undertook self-assessments against the Analytical Framework in 2009 and again in 2010, 16 schools (13%) reported that they had improved by a level on each of the 25 statements in the framework over that twelve month period. A further 40 schools (32%) reported that they had improved by a level on between 20 and 24 of the 25 statements (see Figure 6).

**Table 5. Four ‘Levels of Practice’ in the NSW Analytical Framework (shown for Statement 4)**

Level	Statement 4: ‘There is clarity and consensus about the school’s goals and expectations in literacy and numeracy.’
4	All staff members collaborate to ensure that the school’s goals and expectations for literacy and numeracy are embedded in school routines and classroom routines and procedures. The articulation of the school’s literacy and numeracy goals is central to promoting and sustaining a strong learning culture.
3	There is staff clarity and consensus about the school’s goals and expectations in literacy and numeracy. All staff members can articulate the school’s literacy and numeracy goals.
2	All staff members are aware of the school’s goals and expectations in literacy and numeracy.
1	A few staff members are aware of the school’s goals and expectations in literacy and numeracy.



**Figure 6. Percentage of schools reporting improvement by one level on statements on the NSW Analytical Framework between 2009 and 2010 (n=125)**

### *Annual Reports*

School self-assessments usually are incorporated into schools’ annual reports to their school communities, the system and the general public. Typically, systems require these reports to appear on school websites. Some departments provide templates for schools outlining what the annual report should contain. In Victoria, the annual report provides an opportunity for the school council, leadership team, staff and students to reflect on the success of their improvement strategies and the allocation of resources, and informs planning for improvement in the following year.

The School Report provides the community with information about school performance. It gives parents and other members of the community a clear sense of how students in the school are progressing and what is being done to maximise student achievement.

(Western Australian Department of Education and Training, 2008, 10)

### *External Reviews*

School review cycles generally include an external review of schools' progress in improving student outcomes and school practices. These reviews normally occur as part of the four-year cycle, however some systems, including Victoria and Western Australia, have different categories of reviews and may initiate reviews at any time if there is a perceived need.

External reviews involve an external reviewer (or review panel) evaluating evidence about a school's performance and improvement over the review cycle. External reviews commonly are preceded by schools' self-assessments, and the results of self-assessments may form a key part of the evidence considered by reviewers.

The external validation process reviews the judgements and decisions made by the school in relation to its progress over the school improvement cycle ... The external validation panel may request supplementary evidence across the four domains of school improvement.

(Australian Capital Territory Department of Education and Training, 2009, 14)

The external school review is conducted by an independent external reviewer with expertise in school improvement. The external school reviewer receives the school self-reflection report at least two weeks before the panel meeting and will organise a two-day visit to the school to conduct the external review.

(Catholic Education Office Melbourne, 2009, 3)

The school review follows directly from the school self-evaluation... During the review, the reviewer analyses performance data, the school self-evaluation and additional information to formulate recommendations for the school via a review report.

(Victorian Department of Education and Early Childhood Development, 2011, 7)

External school reviewers receive training in the review process. In some cases, including the Queensland audit process, this includes training to ensure that judgments are made consistently across schools.

The review process includes the preparation of a report. In the ACT, this report contains 'a brief analysis of each domain for school improvement, and commendations and recommendations for future school planning purposes'. The Queensland audit process also results in a report containing 'affirmations, commendations and recommendations' in relation to each school's practices. In Western Australia, if there are concerns about a school's performance, the review team provides a report of the team's findings and a set of prescribed improvement strategies.

### 4.3 Conclusions

Australian school systems already place a high priority on school improvement. Existing school improvement frameworks vary in quality and in the extent to which they are focused on evidence-based practices for improving student outcomes. Nevertheless, there is very significant overlap in the teaching and leadership practices identified in these frameworks, possibly reflecting a degree of cross-referencing. Given the extent of overlap, it would appear to be relatively straightforward to develop a set of nationally-agreed core school improvement domains.

Some school systems already have taken the step of developing schemes for rating school practices in relation to the domains of their frameworks. The ACT rating scheme (Table 3) is used primarily for school self-assessments. The Queensland rating scheme in Attachment 1 is used both for school self-assessments and in the external audit of school practices. The Northern Territory is in the process of introducing ratings almost identical to those in Queensland. A national rewards program would require a common rating scheme and nationally consistent judgements of schools' practices.

Current school review processes suggest ways in which school practices might be judged. The line managers of school principals (regional and district directors; directors school performance; etc) already play a central role in judging school performances and improvements and probably should be integral to a national rewards process. Evidence from current system processes – including schools' four-year improvement plans, annual school self-assessments, progress reports, and reports of formal school review visits – could inform national rewards. The *Reward for School Improvement* initiative probably should build on and enhance the processes already in place around the nation and be seen as an opportunity to strengthen capacity in schools and systems and to promulgate and encourage the national take-up of existing best practices in school improvement.

## 5 Principles for a Reward Scheme

This section uses the foregoing reviews of research, published experience and current Australian school improvement efforts to develop a set of principles that could underpin a system of rewards for school improvement.

### 5.1 Promoting Improvements in Practice

*Principle 1. Rewards should be designed to promote improvements in practice.*

Many reward and award schemes are designed to recognise past achievements; for example, a lifetime contribution to a profession or an outstanding performance under difficult circumstances. Such schemes usually are not intended to encourage or to guide future performance – at least not of reward/award recipients themselves.

A reward scheme for school improvement should be designed not only to recognise retrospectively the improvements that some schools have made, but also to encourage further improvement, both among recipient and non-recipient schools. In other words, rewards should be more than a ‘prize’ for a job well done; they should be based on an explicit recognition that all schools are on an improvement journey and be provided as incentives and support for further improvement.

Related to this point, a reward scheme should unashamedly be designed to drive improvements in school practices and processes. The purpose of a reward scheme should be to improve outcomes for students by improving the quality of classroom teaching and school leadership. To this end, a reward scheme should provide a direct and transparent connection between rewards and practice. The basis of rewards cannot be a ‘black box’ that leaves schools wondering why some schools received a reward for improvement and others did not. Schools must believe that rewards fairly recognise effort and improvement and are based on matters over which they have direct control. In the absence of such transparency, rewards are unlikely to drive improvement or to be seen by schools as credible or fair.

A reward scheme will not lead to overall improvement if it distorts the work of schools. It is clear from recent international experience that, if incentive schemes are not well designed and implemented, they have the capacity to distort rather than to improve practices in schools. Distortion results when strong incentives – rewards or sanctions – are based on a narrow range of evidence; for example, when large financial rewards are tied to results on a small number of test results. Under these conditions, a desirable level of attention and focusing of effort in schools can be replaced by an undesirable lack of attention to anything not tested, including other school subjects and higher-order skills in the areas tested. In some schools, distortion may include allocating large amounts of time to training students in test-taking strategies, assigning the best teachers to the tested year levels, or withholding less able students from testing. In extreme cases, distortion can take the form of dishonest practices such as exposing students to test papers prior to testing, providing access to answers (e.g., on classroom walls) or changing students’ test responses after testing. All of these practices have been observed in high-stakes testing situations.

Distortions of practice inevitably lead to score inflation: increases in test results that are not reflected in increases in students’ underlying skills and knowledge. In

education systems in which strong incentives have been tied to student results, there is evidence that score increases on high-stakes tests have not been matched by increases on other, low-stakes tests constructed to assess the same curriculum content. This finding is consistent with observations about incentive schemes in general. If rewards or sanctions are based only on easily measured outcomes such as a company's annual profit, then superficial – and occasionally dubious – practices are undertaken to maximise these outcomes, often at the expense of unmeasured activities that would have been more important to the long-term profitability of the company.

The way to promote genuine improvement and to minimise distortions of practice is to ensure that reward schemes are based not only on easily-measured outcomes, but also on rich evidence that encourages real and sustainable improvements in the work of schools.

## 5.2 Rewarding Improved Outcomes

*Principle 2. Rewards should be given in part for evidence of improved outcomes.*

Evidence of improved student outcomes must form part of the evidence for judging a school's improvement over time. The core purpose of schools and their work is to improve outcomes for students, so evidence of improved outcomes over time usually will be strong evidence of improving school effectiveness.

Student outcomes are useful indicators only if they are not artificially inflated. In practice, this is less likely to occur if multiple outcome measures are used. Most Australian schools already monitor a range of outcomes, including results on national literacy and numeracy tests at Years 3, 5, 7 and 9; results on other commercially available tests; Year 12 subject results and tertiary entrance ranks; average student attendance rates; apparent student retention rates; surveys of parent, teacher, student perceptions/satisfaction; post-school destination information; and performances on system-provided assessment materials. Rewards for school improvement should be open to a wide range of evidence, although the ability to interpret and judge evidence for the purposes of school rewards will depend on how comparable it is across schools.

Another shortcoming of outcome measures is that they usually are confounded by contextual influences such as students' socioeconomic backgrounds and their levels of achievement upon entering a school. Much has been written about the difficulties of untangling a school's effectiveness from the influence of student backgrounds to infer the 'value' that a school adds over and above its circumstances. This is a particular concern when simple measures of 'status' – such as average test scores and the percentage of students meeting a specified proficiency level – are used to measure a school's performance. Much of the literature on this topic has been concerned with the comparison of schools in terms of their 'value-add'. A number of strategies have been proposed for making more meaningful comparisons of schools' *performances*, including restricting comparisons to schools with similar student intakes and using alternative student outcome measures such as gain scores, growth measures or contextualised value-added measures (see Masters *et al.*, 2008).

However, for the purposes of measuring school *improvement* over time, simple status measures are less problematic because many of the contextual influences are the same or very similar from one student cohort to the next. Because year-to-year fluctuations in status measures may reflect minor differences in student backgrounds and/or prior

achievements, evidence of sustained improvement over a number of years generally will be required to infer improving school effectiveness. Although improvements in gain scores or contextualised value-added measures also could be considered as measures of school improvement, these measures would be less reliable and less interpretable than improvements in simple status measures (see Section 2.5).

### 5.3 Rewarding Improved Practices

*Principle 3. Rewards should be given in part for evidence of improved practices.*

Because the purpose of rewards for school improvement is to improve the work and effectiveness of schools, it makes sense to use direct evidence of improving school practices and processes in judging school improvement. The inclusion of such evidence has the advantage of being consistent with existing Australian school improvement frameworks and school review processes, and of creating more transparent links between rewards and effort than usually is possible through student outcome measures alone – particularly when those measures have been statistically adjusted in some way.

Another reason for using evidence of improved school practices to complement outcomes-based evidence is the research finding that incentive schemes based on student outcomes alone have had limited impact on practice (Hout & Elliott, 2011). Effective drivers of improvement in schools work by building staff capacity and changing the nature of day-to-day work (Fullan, 2011). Most outcomes-based schemes provide very little feedback and guidance to schools on the kinds of changes required to deliver better student outcomes.

It was seen in Sections 3 and 4 of this paper that there is a reasonably high level of agreement among those working in the field of school improvement on the kinds of school and leadership practices required to improve student outcomes. These practices include establishing an improvement agenda for the school; promoting a collective sense of responsibility for improving student learning and outcomes; monitoring school performance data; encouraging the analysis of, and reflection on, professional practice; creating safe and supportive learning environments; and implementing effective teaching strategies throughout the school. Practices of this kind are identified in systems' school improvement frameworks which sometimes include schemes for rating the quality of these practices (see rubrics developed by ACT, NSW and QLD in Tables 3 and 5 and Attachment 1).

This paper also explored the feasibility of aggregating judgements of a school's practices to derive a practice-based 'measure' of school performance (and, potentially, practice-based measures of school improvement over time). This brief exploration suggested that judgements could usefully be aggregated to provide practice-based measures of school performance and improvement to sit alongside outcomes-based measures.

### 5.4 Enhancing Existing Improvement Efforts

*Principle 4. Rewards should build on and enhance existing school improvement efforts.*

A national scheme for rewarding school improvement should not operate in isolation from education systems' current school improvement efforts. Many schools already have in place school improvement plans as part of three-year or four-year review

cycles. These plans include strategies for improvement in the areas/domains of practice identified in school improvement frameworks, annual progress reports to local school communities, and regular formal school visits and reviews. The *Reward for School Improvement* initiative provides an opportunity to support, build on and enhance existing school improvement efforts.

If a National School Improvement Framework were introduced alongside current State and Territory school improvement frameworks, and if schools had strong reasons to address priorities in both national and system frameworks separately, then confusion and duplication of effort would likely ensue. Instead, it would be preferable for school systems to agree on a single National School Improvement Framework in place of existing system frameworks, or at least to adopt a common set of ‘national’ domains as the core elements of all system frameworks and the basis for national rewards.

It also may be desirable that school system personnel already involved in monitoring the performances and improvements of schools are closely involved in school evaluations for the purposes of national rewards. In most cases, these will be the line managers of school principals. As a general principle, wherever possible, duplication of evidence collection should be minimised. Observations and judgements made for system purposes ideally also should provide evidence for national rewards. This would mean that system personnel would have to make judgements in relation to the same domains of school practice, using the same rating categories (e.g., Low, Medium, High, Outstanding), and be trained to make nationally consistent judgements.

An ideal scenario may be one in which national rewards for school improvement were based on nationally agreed processes for reviewing and evaluating the work of schools. Given the high level of commonality of systems’ current school improvement frameworks and school review processes, this may be a realistic aspiration. It would have the advantage of providing a seamlessness between national rewards and State/Territory school improvement efforts.

## 5.5 Encouraging Analysis and Self-Reflection

*Principle 5. Rewards should encourage self-analysis and the critique of improvement.*

A national rewards scheme will be most effective in improving school practices and processes if it engages schools directly in analysing, reflecting on, and presenting evidence of how they are making improvements over time.

Psychological research into the role of incentives in human motivation highlights the importance of providing individuals with a sense of autonomy and control. Incentives can be de-motivating if they are seen as a form of external control and if they are offered in place of intrinsic motivation for behaviours that a person would have engaged in anyway. There is also evidence that external evaluations of performance are most effective when they provide feedback on how to improve future performance in a non-controlling way.

The implications of these observations are that a national reward scheme should not be something that is done *to* schools. For example, if rewards were the result of taking data already available for each school (literacy and numeracy test scores; attendance data; Year 12 results; etc) and statistically manipulating these data to produce a rank order of schools, then the entire process would be taken out of the hands of schools.

There would be no opportunity or encouragement for schools' own analysis of, and reflection on, the improvements they had made. There would be a risk of the entire process becoming an external and obscure 'black box', and there almost certainly would be no feedback to schools on what they needed to do to improve in the future.

An alternative would be a rewards scheme in which schools were deeply involved in analysing and presenting evidence of their improvements. Because schools work in different circumstances and face different challenges, they often have different improvement priorities. For some schools, improving student attendance is a high priority. For others, increasing parental involvement and perceptions is a priority. For still other schools, an important improvement measure is the number of students making successful transitions into vocational training and work. There would be value in giving schools opportunities to identify their improvement priorities and to assemble and reflect on evidence of improvements in their practices (for example, improvements in the strategies they use to collect, analyse and interpret data) as well as evidence of improvements in student outcomes.

Many schools already are familiar with the processes of assembling, reflecting on, and presenting evidence of their work both as part of systems' school review processes and as part of national awards schemes. For example, the Australian Government National Awards for Quality Schooling (2004-2008) invited award nominations from individuals and from schools, including in the categories 'Excellence in School Improvement' and 'Excellence in Family-School Partnerships'. The NAB Schools First awards are open to all schools in Australia and provide significant reward money to schools (\$100,000 to State/Territory winners; \$500,000 to the national winner) for evidence of successful school-community partnerships. These processes sometimes are based only on documentary evidence assembled by schools, but often include both documentary evidence and school visits.

## 5.6 Ensuring Fair Comparisons

*Principle 6. Rewards should be based on fair comparisons of schools.*

Because Australian schools operate in widely different circumstances, improvements in outcomes and practices can be quite different in different schools. Progress in some schools is measured in terms of the number of students attending school and the proportion of those students reaching national minimum standards in basic skills such as literacy and numeracy. In other schools, progress is measured in terms of Year 12 academic results and the proportion of students admitted to prestigious university courses. If schools are to compete for national reward funding, then an attempt should be made to ensure fair comparisons. This probably means not attempting to compare improvements in small, very remote primary schools with improvements in large, urban senior secondary colleges. Any competition for reward funding should be based on comparisons of schools in broadly similar circumstances facing broadly similar challenges.

## 6 A Way Forward?

This final section considers some of the questions and issues that will have to be addressed to develop an effective national system of rewards for school improvement.

The principles outlined in Section 5 of this paper identify a number of desirable features of a national *Reward for School Improvement* scheme. Such a scheme would have as its fundamental purpose the improvement of student outcomes through improved school practices and processes. Ideally, financial rewards provided to schools would:

- encourage schools to evaluate and monitor their own ongoing improvement;
- be based in part on evidence of improved student outcomes;
- be based in part on evidence of improved school practices;
- build on and enhance existing systemic school improvement efforts; and
- be based on fair comparisons of schools.

Decisions will be required in relation to each of these five features of a reward scheme.

### 6.1 School Self-Monitoring

The objectives of the *Reward for School Improvement* initiative are more likely to be achieved if the initiative results in schools developing greater clarity about their own improvement agenda. Ideally, through this initiative schools will develop a clearer understanding of where they currently stand (both in terms of student outcomes and current school practices and processes) and will put in place strategies for further improvement. In other words, all schools will be engaged in ongoing processes of self-evaluation and self-monitoring. These processes will include identifying areas for further development, setting targets and timelines for improvement, identifying and implementing improvement strategies, collecting and analysing evidence of progress, and evaluating the effectiveness of improvement strategies and initiatives. Many schools are already deeply engaged in these processes. An aim of the *Reward for School Improvement* initiative should be to promote these processes in all schools.

A characteristic of improving schools is that they typically have an explicit agenda for improvement. These schools can describe the improvements they wish to see – particularly in student outcomes such as literacy and numeracy levels, Year 12 results, school attendance rates and post-school destinations. This improvement agenda is understood by the whole school community and is regularly reinforced in staff meetings, school newsletters and meetings with parents. The school closely analyses and monitors its performance over time and usually sets targets for improvement on clearly specified timelines.

In such schools, a culture of high expectations has been established. There is a strong underlying belief that further improvement is possible. This belief is promoted by the school leadership team and is shared by the staff of the school. High expectations are set for student behaviour, attitudes and academic achievement and there is low tolerance for attitudes and behaviours that are not consistent with these high expectations.

Importantly, there is also a strong collegial culture. Ongoing improvement is seen as the business of everybody in the school. Interactions between school leaders, staff and

parents are focused on continually improving the school's ability to meet the needs of all students. The school is characterised by high levels of trust and a sense of belonging and pride in the school.

These observations about the importance of high expectations and self-monitoring raise a question about the role school communities should play in assembling their own evidence of improvement and in making their own case for reward funding.

A reward scheme could operate simply by using pre-specified school performance data and deciding on the basis of these data which schools are eligible for rewards. These data may already be available nationally (eg, from national assessment programs) or be collected from schools in nationally standardised ways (eg, parent/community perceptions data). Once the required data are available for all schools, reward decisions could be largely automated.

An advantage of such an approach is that rewards would be based on evidence of improvement on reliable, nationally agreed and nationally comparable measures. The use of a common set of measures (at least within categories of schools) would enable direct comparisons of schools. A further advantage is that reward decisions could be made relatively quickly and inexpensively on the basis of statistical analyses of available data.

The disadvantages of such an approach are that school improvement would be defined and measured only in terms of available, nationally comparable measures. In the first instance, these are likely to be measures of only a limited range of student outcomes. There would be no opportunity for schools to identify and provide evidence in relation to specific, local improvement priorities. There also is likely to be a very limited role for schools in assembling and reflecting on their own data and evidence of improvement. And, depending on how the data are then statistically manipulated, such a scheme may lack transparency, making it unclear why some schools received rewards and others did not.

An alternative approach would be to have schools assemble a case for reward funding. For example, schools may be required to submit evidence in a number of pre-defined categories, demonstrating improvements in school practices and student outcomes over time as well as an explicit ongoing school improvement plan. In this way, schools would be more directly involved in the collection and analysis of evidence of improvement. Table 6 illustrates what a pre-specified set of categories of evidence might look like. The nine categories of evidence listed here are for illustrative purposes only; the actual categories would need to be determined through a national consultative process.

For some categories of evidence, schools may be required to include specified data from national assessment programs or other standardised instruments. For other categories, submitted evidence might take the form of judgements of school practices made as part of external school audit/review processes. But schools also would be given the opportunity to select and include locally-collected evidence of improvement (eg, school measures of non-academic outcomes) as part of their case for a reward. All evidence could be submitted online for evaluation. And not all schools may choose to submit evidence in a particular year, especially if they felt that there was not yet sufficient evidence to support their case for an improvement reward.

**Table 6. Examples of Possible Categories of Evidence**

No.	Category of Evidence
<i>Student Outcomes</i>	
1	Academic outcomes
2	Non-academic outcomes
<i>School Practices</i>	
3	Improvement agenda
4	Analysis and use of data
5	Culture that promotes learning
6	Targeted use of resources
7	Expert teaching team
8	Effective teaching practices
<i>Action Plans</i>	
9	School improvement plan

An advantage of this alternative approach is that school communities would be encouraged to take ownership of their school improvement agenda and to assemble and reflect on evidence of improving outcomes and practices. There is also the possibility of greater flexibility with schools also providing evidence relevant to their own improvement priorities (eg, increased parental engagement; reduced suspensions; improved punctuality).

The disadvantages of this alternative approach are that it would require judgements about the quality of submitted evidence (perhaps similar to judgements made as part of other school awards programs). These judgements would be more resource intensive than automated decisions based on pre-specified metrics. And although all schools would be judged on evidence relating to the same set of categories, it may be more difficult to compare reliably some of the evidence submitted within these common categories.

#### **DECISION POINT**

Should *Rewards for School Improvement* be based only on statistical analyses of pre-specified, nationally agreed performance measures or should schools be able to choose and assemble evidence of improvement and make their own case for reward funding?

## **6.2 Student Outcome Measures**

However the *Reward for School Improvement* program operates, rewards to schools must be based in part on reliable evidence of improving student outcomes. Because the focus of the rewards initiative is on improvement, reward funding should go to schools showing the greatest improvements in student outcomes. This is an important point – reward funding should not go to the best performing schools in an absolute sense (eg,

schools with the highest NAPLAN test results) or to schools that have outstanding performances given the backgrounds of their student intakes (ie, schools with the greatest contextualised ‘value-add’), but to those schools in which the greatest *improvements* in student outcomes have been made over time.

Measures of improvement require data collection on more than one occasion. They also require measurements that can be compared on different occasions, either because the same instrument has been used on these different occasions or because instruments have been ‘equated’ so that their results are expressed on the same measurement scale. NAPLAN tests are equated in this way, enabling results in each year to be reported on the same measurement scale and changes in test results to be compared from one year to the next. If instruments are not statistically equated, then there is no way of knowing whether differences from one occasion to another reflect changes in performance levels or differences due to the instruments themselves. These requirements for the reliable measurement of improvement mean that there are likely to be only a small number of nationally agreed metrics on which improvements in student outcomes can be compared across schools.

Sections 1.3, 2.2 and 2.6 of this paper identified some of the risks associated with using only a narrow range of student outcome measures as a basis for school rewards. When high stakes are attached to a narrow range of measures, the focus of schools and teachers can be narrowed to the outcomes measured. Teaching practices can be distorted, and in extreme cases corrupted, by the prospect of financial rewards. The outcome is often ‘score inflation’ – improvements in test scores that are not matched by real changes in student achievement levels.

A common response to this observation is to attempt to broaden the range of measures on which rewards are based. In the current context, this could be done either by developing a broader range of nationally agreed student outcome metrics and/or by allowing schools to broaden the range of measures by including other evidence of improvement (eg, results on commercial tests or local survey instruments that schools administer each year).

A further challenge in measuring improvement in a school from one year to the next is that student cohorts can differ in their backgrounds and starting points. For example, improvements in Year 7 numeracy results in a school may reflect a more able Year 7 cohort rather than an improvement in the quality of teaching. It is sometimes proposed that a better way of measuring the impact of the school would be to consider the gains that students make (eg, between Year 3 and Year 5) or to use contextualised value-added (cva) measures that adjust for students’ differing backgrounds and starting points. These options were discussed in Sections 2.4 and 2.5. Although these approaches may be useful for making point-in-time comparisons of the performances of different schools, they are less useful for measuring improvement over time. Gain scores are already relatively unreliable; *changes* in gain scores and cva measures over time would provide even less reliable measures of school improvement.

For the purposes of the *Reward for School Improvement* initiative, an ideal solution to measuring improved student outcomes may be to:

- base measures of improvement either on the same instrument administered on different occasions or on ‘equated’ instruments (ie, results expressed on the same measurement scale);
- measure improvement on a number of nationally agreed student outcome metrics;

- expand the range of available measures of this kind over time;
- allow schools to supplement these measures with other evidence of improvement (eg, based on commercial tests and/or local survey instruments);
- use changes in performance measures (rather than changes in gain scores or cva's) as the basis for measuring improvement;
- base conclusions on evidence from multiple year levels; and
- base conclusions on sustained trends in student performance over a number of years.

#### **DECISION POINTS**

Should schools (possibly within categories) be compared on a common set of nationally agreed student outcome measures? If so, what should these be?

Should schools be able to supplement these measures with evidence of improvement on other measures (eg, standardised tests; local school surveys)?

How should this evidence be used to decide on the schools making the greatest improvements in student outcomes?

### **6.3 Measures of School Practice**

An essential feature of rewards for school improvement is that they should be linked to matters over which schools have significant direct control. Because many factors outside the control of schools can influence student outcomes, rewards for school improvement should be based in part on direct evidence of improvements in what schools are doing (that is, improvements in school practices and processes).

Currently there are not well developed measures of schools' practices. However, an inspection of systemic school review processes (Sections 3 and 4) reveals a high level of commonality in the kinds of practices and processes that systems are expecting of schools. These expectations are generally consistent with research evidence on effective schooling, suggesting that it should not be difficult to reach national agreement on the broad parameters of a National School Improvement Framework. Such a framework could then be used as a basis for evaluating school practices and monitoring improvements in those practices over time.

Section 3.3 provided an example of a School Improvement Framework and considered how such a framework might be used to develop measures of school practice. The starting point is the identification of a set of *domains* of school practice and a rubric against which individual schools can be evaluated (see Attachment 1). Clearly, consistency would be required in the interpretation and use of such a framework. Some education systems currently provide training in the consistent application of frameworks of this kind and implement moderation processes to ensure an adequate level of consistency of interpretation and use across schools.

Following the development of a National School Improvement Framework, schools could be asked to provide self-evaluations against the domains of this framework. There would be considerable value in doing this from the point of view of encouraging self-reflection and self-monitoring. However, in a high-stakes context in which significant reward funding rests on judgements of schools' practices and improvements

over time, school self-evaluations alone are unlikely to provide the required level of reliability and would require an accompanying validation process of some kind.

There are a number of options for collecting and evaluating evidence of schools' practices against a National School Improvement Framework:

1. *ask schools to provide self-evaluations against the framework*  
As noted, while this is one possibility, in a high-stakes context with significant attached reward funding, self-assessments which are not externally validated are unlikely to meet the required level of reliability.
2. *ask schools to provide self-evaluations with supporting documentation*  
Under this option, schools would provide evidence that could be inspected as part of an effort to externally validate schools' self-evaluations. This is preferable to no external validation, but relies on schools' descriptions of their own practices and processes.
3. *require self-evaluations to be confirmed through a school visit*  
Under this option, schools' self-evaluations would be validated by having a trained reviewer visit the school. The purpose would be to inspect and provide independent judgements of the school's practices and processes. The process would be expected to be predominantly confirmatory, with the possibility of non-confirmation and alternative judgements being provided.
4. *base judgements only on evidence from an external review*  
Under this option, evidence relating to school practices would be based only on the judgements of external school reviewers. An example would be the evidence currently collected through external school audits/reviews in some education systems.

Whether based on school self-evaluations or external audits/reviews, for the purposes of the *Reward for School Improvement* initiative, evidence relating to the framework domains will need to be brought together and combined in some way (see example in Section 3.3). A decision also will be required about how this evidence is then used in combination with student outcome data. Further work is required on these questions.

And for the purposes of measuring *improvement* over time, judgements of schools' practices at different times will be required. For example, 25 per cent of Queensland government schools were evaluated against the Teaching and Learning School Improvement Framework (Attachment 1) in 2010 and again in 2011, making it possible to identify changes in some schools' practices over this twelve month period.

#### **DECISION POINTS**

Which aspects of school practice should be identified in a National School Improvement Framework? (These would become the framework 'domains'.)

For the purposes of the *Reward for School Improvement* initiative, what processes should be used to ensure reliable evidence of improving school practices?

How should evidence relating to school practices be combined and used in combination with student outcome data for the purposes of comparing schools and rewarding improvement?

## 6.4 Existing School Review Processes

As noted in Section 5.4, the *Reward for School Improvement* initiative provides an opportunity to support, build on and enhance existing school improvement efforts. Most schools have regular school review cycles and, within education systems, these cycles often include formal school visits and evaluations.

The proposed National School Improvement Framework (NSIF) almost certainly will incorporate areas of school practice that are already the focus of school improvement efforts across the country. Consideration will need to be given to the relationship between the proposed NSIF and existing school improvement frameworks (see Section 4.1). Ideally, the NSIF might identify a core set of domains that eventually would become common to all system frameworks. In this way, consistency would be established between rewards for school improvement and systemic school improvement efforts.

A related question is whether greater alignment might be possible between existing school review processes and evidence collection for the purposes of the *Reward for School Improvement* initiative. Ideally, system staff responsible for reviewing schools and ongoing school improvement also might be involved in the external review/validation processes for school rewards.

One possible strategy would be to identify accredited school evaluators for the *Reward for School Improvement* program. These evaluators could be responsible for either auditing/reviewing schools or for validating schools' self-assessments against the NSIF. In education systems, most line managers might be accredited for this purpose. There might also be other accredited evaluators (eg, experienced school principals; retired principals) who could visit schools to undertake reviews or to validate schools' self-assessments. Some level of training and accreditation in making judgements of schools' practices against the NSIF would be required. This strategy would provide a level of national consistency and comparability, but also could provide valuable professional development experiences for those involved.

### DECISION POINTS

Should the collection of evidence for the *Reward for School Improvement* initiative, including judgements of schools' practices and processes, be integrated with education systems' school review processes?

For the purposes of making reliable comparisons of schools' practices and processes, should school reviews/validations be undertaken by trained and accredited *RSI* evaluators?

## 6.5 Fair Comparisons

School improvement will be differently defined in different types of schools and different school contexts. The student outcome measures relevant in primary schools will generally be different from outcome measures in senior secondary colleges. Improvements in school attendance, community perceptions and parental engagement will be significant objectives for some schools, but not for others. The nature of improvement in a small, very-remote school may be quite different from improvement in a large, urban school. For these reasons, in providing rewards for school improvement, it will be important to recognise the different circumstances and challenges faced by schools of different sizes and in different parts of the country, and the different definitions of ‘improvement’ that may apply in these schools.

One obvious approach to maximising the fairness of comparisons is to make comparisons and to provide reward funding within categories of schools. Separate reward categories for primary and secondary schools are likely to be appropriate given the different outcome measures relevant to these phases of schooling. In primary schools, data on literacy and numeracy levels will be especially important. In secondary schools, outcome measures might include Year 12 results, VET in schools data and post-school destination data.

Consideration also might be given to providing rewards in other categories of schools. For example, significant improvements in student achievement levels may be harder to achieve in already high-performing schools than in very low-performing schools. Fair comparisons of schools may require comparisons only of schools in similar circumstances. This is an issue that requires further investigation.

### **DECISION POINTS**

Given that different student outcomes are likely to be relevant in different kinds of schools (eg, primary and secondary), should rewards for school improvement be given within categories of schools?

If so, what should these school categories be, and what student outcome measures should contribute to reward decisions in each category?

## Bibliography

- Australian Capital Territory Department of Education and Training. (2009). *School Improvement Framework*. Canberra: Department of Education and Training.
- Australian Labor Party (2010). *Rewarding Improvement so Every School Becomes a Great School*. Reward for School Improvement Fact Sheet. Barton, ACT.
- Catholic Education Office Melbourne (2009). *School Improvement Framework*. Melbourne: Catholic Education Office, Archdiocese of Melbourne.
- Chester, D. (2005). Making valid and consistent inferences about school effectiveness from multiple measures. *Educational Measurement: Issues and Practice*, 24(4), 40-52.
- Cohen, D.K., Raudenbush, S.W. & Loewenberg Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119-42.
- Dederling, Kathrin, and Sabine Müller, School Improvement through Inspections? First Empirical Insights from Germany. *Journal of Educational Change*, December 3, 2010.
- Elmore, R. F. (2004). The problem of stakes in performance-based accountability systems. In S.H. Fuhrman & R.F. Elmore (Eds.), *Redesigning Accountability Systems for Education*. New York: Teachers College Press.
- Elmore, R.F. (2008). Leadership as the practice of improvement. In Pont, B., Nusche, D. & Hopkins, D. (Eds). *Improving School Leadership*, Vol 2: Case Studies on System Leadership. Paris: OECD.
- Faubert, V. (2009), School Evaluation: Current Practices in OECD Countries and a Literature Review. *OECD Education Working Papers*, No. 42, OECD Publishing.
- Fullan, M. (2011). Choosing the wrong drivers for whole school reform. Seminar Series 204. Melbourne: Centre for Strategic Education.
- Hamilton, L. (2004). Assessment as a policy tool. *Review of Research in Education*, 27, 25-68.
- Hattie, J. (2003). Teachers make a difference, what is the research evidence? *Building Teacher Quality: What Does the Research Tell Us?* Proceedings of Research Conference 2003. Melbourne: Australian Council for Educational Research.
- Hill, R. K. & DePascale, C. A. (2003). Reliability of No Child Left Behind accountability designs. *Educational Measurement: Issues and Practice*, 22(3), 12-20.
- Hout, M & Elliott, S.W. (Eds.) (2011). *Incentives and Test-Based Accountability in Education*. Report of the Committee on Incentives and Test-Based Accountability in Public Education. Washington: The National Academies Press.
- Kelly, A. & Downey, C. (2010). Value-added measures for schools in England: looking inside the 'black box' of complex metrics. *Educational Assessment, Evaluation and Accountability*, 22, 181-198.

- Koretz, D., & Barron, S.I. (1998). *The Validity of Gains on the Kentucky Instructional Results Information System (KIRIS)*. MR-279-PCT/FF. Santa Monica, CA:RAND.
- Koretz, D.M., McCaffey, D.F., & Hamilton, L.S. (2001). *Toward a framework for validating gains under high-stakes conditions* (CSE Tech. Rep. 551). Los Angeles: Centre for Research on Evaluation, Standards, and Student Testing.
- Linn, R.L. (2008). Educational accountability systems. In K.E. Ryan and L.A. Shepard, (Eds.), *The Future of Test-Based Accountability*. New York: Routledge.
- Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- Masters, G.N. (2010). *Teaching and Learning School Improvement Framework*. Melbourne: Australian Council for Educational Research.
- Masters, G.N., Rowley, G., Ainley, J. & Khoo, S.T. (2008). *Reporting and Comparing School Performances*. Canberra: Dept of Education, Employment and Workplace Relations.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Meyer, R.H. & Dokumaci, E. (2011). Value-added models and the next generation of assessments. Paper presented at the annual conference of the Association for Education Finance and Policy, Seattle Washington, 24-26 March.
- Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the World's Most Improved School Systems Keep Getting Better*. New York: McKinsey & Company.
- Neal, D. & Schanzenbach, D.W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*, 92(2), 263-283.
- New South Wales Department of Education and Communities. (2011). *Developing School Improvement Capacity using the NSW DET Analytical Framework for Effective Leadership and School Improvement in Literacy and Numeracy*. Sydney: Department of Education and Communities.
- New South Wales Department of Education and Training. (2011). *Framework for School Development and Accountability*. Sydney: Department of Education and Training.
- Northern Territory Department of Education and Training. (2010). *School Accountability and Performance Improvement Framework*. Darwin: Department of Education and Training.
- Pedulla, J.J., Abrams, L.M., Madaus, G.F., Russell, M.K., Ramos, M.A., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. National Board on Educational Testing and Public Policy.
- Pinkus, L.M., (2009). Moving beyond AYP: high school performance indicators. *Policy Brief*. Washington, DC: Alliance for Excellent Education.
- Queensland Department of Education and Training (2011). *School Planning, Reviewing and Reporting Framework: An improvement and accountability framework for Queensland State Schools*. Brisbane: Department of Education and Training.

- Raudenbush, S.W. (2004). *Schooling, Statistics and Poverty: Can we Measure School Improvement?* The ninth annual William H. Angoff memorial lecture. Educational Testing Service, Princeton, New Jersey.
- Raudenbush, S., & Willms, D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307-335.
- Robinson, V. (2007). The impact of leadership on student outcomes – making sense of the evidence. *The Leadership Challenge: Improving Learning in Schools*. Proceedings of Research Conference 2007. Melbourne: Australian Council for Educational Research.
- Sammons, P., Hillman, L. and Mortimore, P. (1995) *Key Characteristics of Effective Schools: A Review of School Effectiveness Research*. London: Institute of Education.
- Schochet, P.Z. & Chiang, H.S. (2010). *Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains*. (NCEE 2010-4004). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Schwartz, H.L., Hamilton, L.S., Stecher, B.M. and Steele, J.L. (2011). *Expanded Measures of School Performance*. Report prepared for the Sandler Foundation. Rand Education: Santa Monica, California.
- South Australian Department of Education and Children's Services (2007). *Improvement and Accountability Framework*. Adelaide: Department of Education and Children's Services.
- Tasmanian Catholic Education Commission (2011). *School Improvement for Catholic Schools in Tasmania: School Improvement Framework and Process*. Hobart: Catholic Education Commission.
- Victorian Auditor-General (2007). *Improving Our Schools: Monitoring and Support*. Melbourne: Victorian Government Printer.
- Victorian Department of Education and Early Childhood Development (2011). *Accountability and Improvement Framework for Victorian Government Schools*. Melbourne: Department of Education and Training.
- Western Australian Department of Education and Training (2008). *The School Improvement and Accountability Framework*. Perth: Department of Education and Training.

## Attachment 1

### **Teaching and Learning School Improvement Framework Queensland Department of Education and Training**

#### **Domains**

- 1 An Explicit Improvement Agenda
- 2 Analysis and Discussion of Data
- 3 A Culture that Promotes Learning
- 4 Targeted Use of School Resources
- 5 An Expert Teaching Team
- 6 Systematic Curriculum Delivery
- 7 Differentiated Classroom Learning
- 8 Effective Teaching Practices

**Domain 1: An Explicit Improvement Agenda****OUTSTANDING**

The principal and other school leaders have developed and are driving an explicit and detailed local school improvement agenda. This agenda is couched in terms of specific improvements sought in student performances, is aligned with national or system-wide improvement priorities and includes clear targets with accompanying timelines which are rigorously actioned. The school improvement agenda has been effective in focusing, and to some extent narrowing and sharpening, the whole school's attention on core learning priorities. There is a strong and optimistic commitment by all staff to the school improvement strategy and a clear belief that further improvement is possible. Teachers take responsibility for the changes in their practice required to achieve school targets and are using data on a regular basis to monitor the effectiveness of their own efforts to meet those targets.

**HIGH**

The school has developed an agenda for improvement and school leaders can describe the improvements they wish to see in student behaviours and outcomes. This agenda is communicated in staff meetings, school newsletters, parent-teacher meetings and on the school website using a variety of formats to suit local needs. The principal and other school leaders have analysed school performance data over a number of years and are aware of trends in student achievement levels. Targets for improvement are clear and accompanied by timelines. The school leadership team is clearly committed to finding ways to improve on current student outcomes. This is reflected in an eagerness to learn from research evidence, international experience and from other schools that have achieved significant improvements. There is evidence of a school-wide commitment to every student's success and staff of the school tell stories of significant student improvement.

**MEDIUM**

The principal and other school leaders articulate a shared commitment to improvement, but limited attention has been given to specifying detail or to developing a school-wide approach (eg, plans for improvement may lack coherence, be short term or without a whole-school focus). Plans for improvement do not appear to have been clearly communicated, widely implemented or to have impacted significantly on teachers' day-to-day work. Targets for improvement are not specific (eg, not accompanied by timelines). The school's focus on data is driven more by external requirements (eg, NAPLAN) than by an internal desire for good information to guide school decision making and to monitor progress. Although there is an expressed commitment to improvement, this is not reflected in a high level of enthusiasm for personal change on the part of staff. The communication of performance data to the school community tends to be sporadic and/or is limited only to information that the school is required to report.

**LOW**

There is no obvious plan for improving on current achievement levels. The principal appears to be more focused on day-to-day operational matters than on analysing and understanding school data, setting targets for whole-school improvement or communicating an improvement agenda to the school community. Minimal attention is paid to data (eg, NAPLAN results) and there is very limited communication of school results or of intentions for improvement to the wider school community. Expectations for significant school improvement are low and staff tend to 'explain' current achievement levels in terms of students' socioeconomic backgrounds and/or geographical location. There is little evidence that the staff of the school have a shared commitment to improving outcomes for every student, and this appears to be contributing to a culture of underperformance. There is little evidence that the school is looking to external sources to identify evidence-based strategies for improvement.

**Domain 2: Analysis and Discussion of Data****OUTSTANDING**

The principal and other school leaders clearly articulate their belief that reliable data on student outcomes are crucial to the school's improvement agenda. The school has established and is implementing a systematic plan for the collection, analysis and use of student achievement data. Test data in literacy, numeracy and science are key elements of this plan. Data are used throughout the school to identify gaps in student learning, to monitor improvement over time and to monitor growth across the years of school. A high priority has been given to professional development aimed at building teachers' and leaders' data literacy skills. Staff conversations and language reflect a sophisticated understanding of data concepts (eg, value-added; growth; improvement; statistical significance). Teachers are given test data (including NAPLAN) for their classes electronically and are provided with, and use, software (eg, *Excel*) to analyse, display and communicate data on individual and class performances and progress, including comparisons of pre- and post-test results. Teachers routinely use objective data on student achievement as evidence of successful teaching.

**HIGH**

The school has developed an agenda for improvement and school leaders can describe the improvements they have made. There is evidence that the principal and other school leaders view reliable, timely student data as essential to their effective leadership of the school. There is a documented school plan and timetable for the annual collection of student outcome data. One or more members of staff have been assigned responsibility for implementing the annual plan, analysing the full range of school data, and summarising, displaying and communicating student outcome data for the school. The school has ensured that appropriate software is available and that the assigned staff have been trained to undertake data analyses. Time is set aside (eg, on pupil free days and in staff meetings) for the discussion of data and the implications of data for school policies and classroom practices. These discussions occur at whole-school and team levels. The school can illustrate through case studies, meeting minutes and project plans how data have been used to identify priorities, take action and monitor progress.

**MEDIUM**

School leaders pay close attention to data provided to them about the performance of the school (eg, NAPLAN results; Year 12 results) and identify areas in which the school is performing relatively poorly or well. Tests (eg, commercially available reading tests) may be used by some teachers, but generally are not used as part of a whole-school assessment strategy. An ad hoc approach exists to building staff skills in the analysis, interpretation and use of classroom data. Software may be used for the analysis of school results, including the performances of priority groups, but analyses generally do not extend to studies of improvement or growth. School data are presented to staff in meetings, but presentations tend to be 'for information' rather than a trigger for in-depth discussions of teaching practices and school processes. Information about the school's performance is communicated to the school community, but may lack explanation or analysis.

**LOW**

There is very little evidence of school leaders' practical use of school-wide student outcome data. There is either no annual data collection plan for the school or the plan is being implemented in a minimalist fashion. The school makes little or no use of tests beyond those that the school is required to use. Teachers do not systematically analyse test and other data for their classes and teachers make little use of data to reflect on their teaching. The school is unable to demonstrate how data have been used in meetings to analyse and discuss current achievement levels and strategies for improvement.

**Domain 3: A Culture that Promotes Learning****OUTSTANDING**

The school ethos is built around high expectations and a commitment to academic excellence. There is an expectation that every student will learn and achieve positive outcomes. Classrooms are calm but busy and interruptions to teaching time are kept to a minimum. There are no obvious behavioural problems and staff morale is high. There is a happy, optimistic feel to the school. High levels of trust are apparent across the school community. Interactions are focused on the learning and wellbeing of students and on continually improving the school's ability to meet the needs of all students. Parents, school leaders and teachers work together in a mutually supportive way and there is a strong sense of belonging and pride in the school. A strong collegial culture has been established. Teachers have an overt and shared commitment to the improvement of teaching and an openness to critique by colleagues. This is reflected in the fact that teachers regularly invite leaders and colleagues to visit their classrooms to observe their teaching.

**HIGH**

The 'tone' of the school reflects a school-wide commitment to purposeful, successful learning. There are very few obvious behavioural problems and behaviour management takes up very little, if any, time of school leaders and classroom teachers. There is a strong focus on quality learning and on the creation of a culture in which all students are expected to learn successfully, in their own ways and at their own pace. Individual talents are valued. Class 'busy work' is kept to a minimum, and an attempt is made to ensure that all students are engaged in challenging, meaningful learning. Respectful and caring relationships are reflected in the ways in which staff, students and parents interact and in the language they use in both formal and informal settings. Parents are encouraged to take a genuine and close interest in the work of the school and are welcomed as partners in their children's learning. There are agreed guidelines on such matters as greeting visitors, taking messages, and responding to queries promptly and respectfully. Staff morale is generally high.

**MEDIUM**

Classrooms are generally orderly, although some are more so than others. Non-attendance is an issue for a small minority of students. However, many other students appear to be minimally engaged in productive learning activities. The school effectively implements its policies, for example, by ensuring that disruptive behaviour, bullying and harassment are dealt with promptly. The school has clear expectations for how students should behave and interact with one another, and in the main, relationships are caring and respectful. Some staff time is taken up dealing with behaviour problems. Most parents take an obvious interest in their children's learning. Engagement is primarily through regularly scheduled parent-teacher interviews. Staff morale is satisfactory.

**LOW**

Behavioural problems, disengagement and non-attendance are issues for a significant proportion of students. In a number of classrooms students are clearly not engaged in productive learning activities. The school may have policies and agreed procedures relating to student behaviour but these appear to have had little impact in practice. Much of the time of school leaders and teachers is taken up dealing with inappropriate behaviour. Interactions between parents, staff and students are not always productive and respectful. Some teachers appear to work in isolation from colleagues. Staff morale is low and staff turnover is high.

**Domain 4: Targeted Use of School Resources****OUTSTANDING**

The principal and other school leaders have given a very high priority to understanding and addressing the learning needs of all students in the school. This is reflected in the implementation of systematic strategies for identifying student needs and the development of creative school-wide solutions for addressing those needs. A school-wide process has been established for identifying specific student learning needs. This process includes systematic testing to establish learning gaps and special needs. Records of individual student needs, achievements and progress are maintained centrally and shared across year levels. A range of initiatives (eg, across-class and across-grade groupings for literacy and numeracy; the sharing of specialist teachers with neighbouring schools) is being implemented to make more effective use of available resources to enhance teaching and learning. The school deploys staff in ways that make best use of their expertise (eg, specialist reading/science teachers).

**HIGH**

The principal and other school leaders have introduced programs and strategies to identify and address the needs of students in the school and are sourcing and applying available resources to meet those needs. The school has developed processes (eg, systematic testing) for identifying student learning needs, although there may not always be good central records of student achievement and progress. Programs to meet individual learning needs (eg, programs for gifted students, students with learning difficulties, students for whom English is a second language, Indigenous students, refugees) are prioritised, where possible, in the school budget. Physical spaces and technology are used effectively to maximise student learning. Learning spaces are organised for whole group work, small group work and individual work.

**MEDIUM**

The school uses its human and physical resources to address the needs of students, although this may not be preceded by a systematic analysis of those needs. Specialist diagnostic testing is organised for a small number of students with special needs, but teachers do not routinely administer tests to better understand specific learning difficulties (eg, problems in learning to read) or individual learning needs (eg, LBOTE, gifted). There are very few school-wide programs or policies designed to address the learning needs of particular student groups (eg, gifted students, students with disabilities, students for whom English is a second language). School leaders encourage teachers to address individual learning needs in classrooms, but there are very few agreed school-wide strategies for doing this. Physical learning spaces are used creatively and technology is accessible to the majority of staff and students.

**LOW**

The improvement of student outcomes does not appear to be the driving consideration in the allocation of school resources (eg, the use of discretionary school funds). There is very little, if any, systematic testing of students to identify individual learning needs. The school does not always make best use of available staff expertise. School leaders have developed very few, if any, school-wide policies or programs to address individual needs, which are left to classroom teachers. School learning spaces tend to be used traditionally, with limited flexibility to support different kinds of learners and learning.

**Domain 5: An Expert Teaching Team****OUTSTANDING**

The teaching staff of the school are experts in the subjects they teach and have very high levels of pedagogical knowledge and skill, including expert knowledge of evidence-based teaching strategies. Teachers and school leaders take personal and collective responsibility for improving student learning, working as a team and learning from each other's practices. In team meetings there is an emphasis on the joint analysis of student work and on teaching strategies for improving student learning. Teachers collaboratively plan, deliver and review the effectiveness of lessons. School leaders place a very high priority on the ongoing professional learning of all staff and on the development of a school-wide, self-reflective culture focused on improving classroom teaching. School leaders participate in professional learning activities, learning alongside teachers, and the school supports teachers to continue formal study and celebrates professional success.

**HIGH**

There is evidence that the principal and other school leaders see the development of staff into an expert and coherent school-wide teaching team as central to improving outcomes for all students. There is a documented professional learning plan and the school has arrangements in place for mentoring and coaching. Teachers visit each other's classrooms and welcome opportunities to have principals and other school leaders observe and discuss their work with them. Attention is paid to strengths and weaknesses in the school-wide team, with strategies in place to recruit staff with particular expertise, to train staff to address particular needs, and to support staff who find it difficult to commit to the school's improvement agenda. The school provides opportunities for teachers to take on leadership roles outside the classroom.

**MEDIUM**

The school undertakes professional learning activities, although these may not always focus on the development of knowledge and skills required to improve student learning and there may not be a coherent, documented learning plan. The principal and leadership team are seen as supportive of, but not generally involved in, the day-to-day practice and learning of teachers. Teachers are open to constructive feedback and provide feedback to colleagues, although there may not be formal mentoring or coaching arrangements in place. The school is implementing a formal process for conducting professional discussions with staff. The school's professional learning agenda is made explicit to staff at induction, and in staff handbooks. Where it is necessary to manage unsatisfactory staff performance, this is done professionally and effectively, and in accordance with agreed guidelines.

**LOW**

The development of a professional school-wide team does not appear to be a driving consideration of the principal or other school leaders (eg, no reference is made to professional standards for teachers, there are no mentoring arrangements in place, teachers work largely in isolation from one another 'behind closed doors'). There is little evidence that school leaders are proactive in the recruitment and retention of staff. There is little sense of a whole-school coordinated approach to professional learning and a low priority is given to enhancing staff performance.

**Domain 6: Systematic Curriculum Delivery****OUTSTANDING**

The principal and other school leaders have developed and are driving an explicit and detailed local school plan for curriculum delivery. The school has a clearly documented whole school plan for curriculum delivery. This plan is aligned with national and/or system curriculum frameworks and makes explicit what (and when) teachers should teach and students should learn. The curriculum delivery plan is being implemented throughout the school. A strong alignment has been achieved between the overall curriculum delivery plan, term and unit plans, classroom teaching and the regular assessment of student progress in relation to curriculum expectations. Considerable attention has been given to ensuring 'vertical' alignment of the curriculum so that there is continuity and progression of learning across the years of school, with teaching in each year building on to and extending learning in previous years.

**HIGH**

The school's curriculum delivery plan identifies curriculum, teaching and learning priorities and requirements. The curriculum delivery plan reflects a shared vision (principal, school leadership team, and teachers) for the school, and provides a context for delivering the curriculum as detailed in national and/or system frameworks. The school curriculum plan and curriculum delivery (including the time allocated to particular learning) balances requirements to address all learning areas, to give priority to English, mathematics and science, and to embed the fundamental skills of literacy, numeracy and higher order thinking in all school subjects. The school leadership team ensures that the enacted curriculum remains a focus for discussion among, and collaboration between, teachers and that the curriculum plan is the reference against which flexible delivery (to meet the needs of the range of students within each year level) is designed, assessment tasks are developed and student learning is reported.

**MEDIUM**

The school has a documented plan for curriculum delivery that includes year level and term plans, but the progression of learning from year to year is not always obvious and the relationship between the pieces of the plan (the year, term and unit plans) would benefit from further clarification. School leaders talk about embedding the fundamental skills of literacy, numeracy and higher order thinking within all subjects, but there is little evidence that school-wide strategies are in place to drive a consistent approach. Literacy tends to be seen as the responsibility of English teachers and numeracy the responsibility of mathematics teachers. Discussions about curriculum delivery tend to be sporadic and reactive with a year level focus rather than being driven by a leadership team with a whole school approach.

**LOW**

School leaders and teachers do not appear to be familiar with national or system-wide curriculum documents. The school may have a documented plan for curriculum delivery but there is little evidence that the whole school plan drives the lesson plans of individual teachers. The enacted school curriculum is not seen as a central concern of all teachers (eg, it is not a regular topic of conversation, a focus for assessment design or a framework against which student learning is reported).

**Domain 7: Differentiated Classroom Learning****OUTSTANDING**

The principal and other school leaders actively promote the use of differentiated teaching as a strategy for ensuring that every student is engaged and learning successfully. It is recognised throughout the school that some students require significant adjustments to their learning programs (eg, accelerated programs, special support) if they are to be optimally engaged and challenged, and individual learning plans have been developed for a least some students. Differentiation is a priority of the school and a feature of every teacher's practice. Regular data on the achievements, progress, strengths and weaknesses of individual students are used in all classrooms to make judgements about individual needs, to identify appropriate starting points for teaching and to personalise teaching and learning activities.

**HIGH**

School leaders explicitly encourage teachers to tailor their teaching to student needs and readiness. This includes the systematic use of assessment instruments (standardised assessment tasks and teacher developed assessment tools) to establish where individuals are up to in their learning and to identify skill gaps and misunderstandings. Teachers also are encouraged to respond to differences in cultural knowledge and experiences and to cater for individual differences by offering multiple means of representation, engagement and expression. Planning shows how the different needs of students are addressed, and how multiple opportunities to learn are provided, including multiple pathways for transition to external studies (eg, apprenticeships) for students in Years 10-12. Students' workbooks also illustrate differentiated tasks and feedback. Reports to parents show progress over time and include suggestions for ways in which parents can support their children's learning.

**MEDIUM**

School leaders are committed to success for all, but do not drive a strong classroom agenda to assess and identify individual learning needs or to differentiate teaching according to students' needs. Some use is made of assessment instruments to identify individual strengths and weaknesses and starting points for teaching, but this appears to be at the initiative of individual teachers rather than a school-wide expectation. Some use is made of differentiated teaching (eg, differentiated reading groups in the early primary years), but in most classes teachers teach the same curriculum to all students with similar levels of individual support. Regular assessments of student learning are undertaken, but these often are summative and disconnected (eg, relating to different topics) rather than exploring progress in skill development and understanding over time. Reports to parents generally do not show progress over time or provide guidance to parents on actions they might take.

**LOW**

School leaders do not place a high priority on teachers identifying and addressing individual learning needs, but are more focused on ensuring that all teachers are teaching the core year level curriculum. Little or no classroom use is made of assessment instruments to establish starting points for teaching. Assessments tend to be used only to establish summatively how much of the taught content students have learnt. Teachers tend to teach to the middle of the class, with the expectation that some students will not master the content, and finding ways to occupy more able students who finish work early. Reports to parents tend to be summative reports of how students have performed, with little guidance on what parents might do to assist in their children's learning.

**Domain 8: Effective Teaching Practices****OUTSTANDING**

The principal and other school leaders have accepted personal responsibility for driving improvements in teaching throughout the school. They demonstrate a strong conviction that improved teaching is the key to improved student learning and have clear and well-known positions on the kinds of teaching that they wish to see occurring. All teachers and leaders are committed to identifying and implementing better teaching methods, and a high priority is given to evidence-based teaching strategies (strategies that have been demonstrated through research and practice to be highly effective). School leaders spend time working with teachers to improve their teaching practices, including modelling, evaluating and providing feedback on classroom teaching.

**HIGH**

School leaders are committed to continuous improvement in teaching practices throughout the school and expect team leaders and teachers to identify ways of doing this, although the principal and other senior leaders may not themselves have clear positions on the kinds of teaching they wish to see occurring across the school or be 'hands on' in driving improved teaching practices. There is a particular focus on improved teaching methods in reading, writing, mathematics and science, and professional learning activities are focused on building teachers' understandings of highly effective teaching strategies in these areas. Clarity about what students are expected to learn and be able to do, high expectations of every student's learning, explicit teaching of skills and content, individualised attention as required, and timely feedback to guide student action are key elements of the school's push for improved teaching and learning.

**MEDIUM**

School leaders are explicit about their desire to see effective teaching occurring throughout the school, but are less clear about what this might look like. They do not appear to be driving a strong agenda to improve and/or enhance teaching practices across the school, except perhaps indirectly through a focus on school results and targets. School leaders take a close interest in the school's literacy and numeracy results, but generally do not engage in discussions with staff about effective literacy and numeracy teaching strategies. Discussions of specific teaching practices are rare and generally occur only in the context of concerns about a teacher's performance. There is some clarity about what students are expected to learn, but a lack of balance in teaching methods (eg, over-reliance on whole-group teaching or very little explicit teaching).

**LOW**

School leaders do not appear to have strong views on the characteristics of highly effective teaching. There is little evidence that they are driving an agenda to change or enhance teaching practices across the school. The principal and other school leaders spend very little time on issues related to teaching. Questions of pedagogy appear to be viewed solely as the responsibility of teachers. There is little obvious school-wide discussion or consideration of highly effective teaching methods. There may be a lack of explicitness about what students are expected to know and be able to do as a result of classroom activities, very little explicit teaching, limited attention to individual learning needs, or low expectations on the part of teachers that all students in the room will master the content of lessons. Classroom activities frequently have the appearance of 'busy work'.

## Attachment 2

### School Improvement Framework Australian Capital Territory (2009)

The domains and elements of school improvement

Domain	Description
<b>Learning and Teaching</b>	<p>The learning and teaching domain describes the context in which the curriculum is delivered. High quality learning occurs when teachers make appropriate decisions about what is taught, how to engage students in meaningful experiences and how progress will be assessed to inform future actions.</p> <ul style="list-style-type: none"> <li>• teachers apply their contemporary and professional knowledge to establish highly effective learning environments</li> <li>• teachers set expectations, plan for success and assess learning outcomes</li> <li>• school curriculum design and delivery establishes explicit and high standards for learning.</li> </ul>
<b>Leading and Managing</b>	<p>The leading and managing domain is concerned with communicating a clear vision for a school and establishing effective management structures. Leaders set directions and guide the school community in alignment of its purpose and practice. Effective leadership within the school is collegial, student centred and teacher focused, promoting a collective responsibility for improvement.</p> <ul style="list-style-type: none"> <li>• school vision is collaboratively developed to be realistic, challenging and futures oriented</li> <li>• leaders use reflective practices to appropriately manage people to achieve improvements to teaching and learning</li> <li>• the school's leadership team demonstrates effective resource management to achieve results.</li> </ul>
<b>Student Environment</b>	<p>The student environment domain describes the promotion of positive and respectful relationships which are stable, welcoming and inclusive. In safe and productive learning environments students willingly engage and participate in the broad range of learning opportunities. They contribute to decisions about their learning and their contributions are valued.</p> <ul style="list-style-type: none"> <li>• quality learning environments are created to focus on student needs and foster potential skills and interests</li> <li>• schools create opportunities for students to develop into self-regulating learners within and beyond the classroom</li> <li>• schools value participation, and encourage student expression of new knowledge and understanding.</li> </ul>
<b>Community Involvement</b>	<p>The community involvement domain describes the development of quality ongoing community partnerships and networks. Schools are responsive to community expectations, value diversity and encourage contribution. Positive futures and cultures of success are promoted as educational outcomes.</p> <ul style="list-style-type: none"> <li>• schools develop effective relationships with parents/carers to support student engagement with learning</li> <li>• the school enriches the curriculum through partnerships and activities involving the local community and resources</li> <li>• the school celebrates successful learning outcomes and promotes its achievements across the wider community.</li> </ul>

## Attachment 3

**School Accountability and Performance Improvement Framework  
Northern Territory (2010)**

<b>Key Result Areas</b>	<b>Outputs and Milestones</b>	<b>Outcomes and targets</b>
<b>Teaching and Learning</b>	<ul style="list-style-type: none"> <li>- Curriculum planning including: strategies for Literacy and Numeracy and ICT for Learning</li> <li>- Aboriginal and Torres Strait Islander Education Individual Learning Strategies</li> <li>- Educational Adjustment Plans</li> <li>- NTCET Assessment Plans</li> </ul>	<ul style="list-style-type: none"> <li>- NAPLAN for Years 3, 5, 7 and 9</li> <li>- NT Curriculum Framework achievement</li> <li>- T-9 net</li> <li>- VET in schools qualifications by level</li> <li>- NT Certificate of Education and Training achievement</li> <li>- Diagnostic/school-based assessment</li> </ul>
<b>Wellbeing</b>	<ul style="list-style-type: none"> <li>- Priorities identified in student and staff perception surveys</li> <li>- Relevant well-being policies and/or programs</li> <li>- Celebrations of success and progress</li> </ul>	<ul style="list-style-type: none"> <li>- Student perception data</li> <li>- Staff perception data</li> <li>- Student suspension data</li> <li>- Student mobility data</li> <li>- Staff Absences data</li> <li>- Staff turnover and retention data</li> </ul>
<b>Transitions and Pathways</b>	<ul style="list-style-type: none"> <li>- Range of options and flexible learning programs for individuals and groups of students including students with special needs.</li> <li>- Student destination survey</li> <li>- Attendance strategy</li> </ul>	<ul style="list-style-type: none"> <li>- Australian Early Development Index data (AEDI)</li> <li>- Assessment of Student Competencies (ASC) data</li> <li>- Student attendance data</li> <li>- Real and apparent retention data</li> <li>- VETiS and NTCET completions</li> <li>- Australian Tertiary Assessment Rank</li> </ul>
<b>Partnerships</b>	<ul style="list-style-type: none"> <li>- Aboriginal and Torres Strait Islander School Community Partnership Agreement</li> <li>- Memorandums of Understanding</li> <li>- Service Level Agreements</li> <li>- Community and/or industry engagement events</li> <li>- Effective partnerships</li> <li>- Priorities identified in parent and community perception surveys</li> </ul>	<ul style="list-style-type: none"> <li>- Parent perception data</li> <li>- Community perception data</li> </ul>
<b>Leadership</b>	<ul style="list-style-type: none"> <li>- School mission/vision and values statements</li> <li>- 360<sup>o</sup> feedback</li> <li>- Interpersonal relationships</li> <li>- School review</li> <li>- Performance Management processes</li> <li>- Professional Development data</li> <li>- Leadership and PD components of School Literacy and Numeracy and ICT self-assessments</li> <li>- Alignment of expenditure with strategic directions</li> <li>- Expenditure of supplementary funding in period of agreement</li> </ul>	<ul style="list-style-type: none"> <li>- Occupational health and safety data</li> <li>- Principal and leadership turnover (system level monitoring)</li> </ul>

## Attachment 4

### Improvement and Accountability Framework South Australia (2007)

## Underpinning principles

<b>Principle</b>	<b>Description</b>
<b>Focus on Learning</b>	A focus on learning requires us to create a learner centred environment with high expectations for all learners. Each learner, with the right support and direction, can achieve success in terms of their learning and wellbeing. Quality teaching lies at the heart of learner achievement. As a system we work to continuously improve the services that promote and support teaching and learning and build effective learning communities.
<b>Think Systemically</b>	Thinking systemically means understanding the complex relationships between all parts of the systems within which we work. It requires us to work together to develop effective processes and strategies, involving the broader community, to support the needs and aspirations of our children and young people.
<b>Share Leadership</b>	Sharing leadership requires us to build leadership capacity and expertise throughout the system. The most effective sites and systems are those where leadership is purposefully developed and distributed within and across the local community.
<b>Attend to Culture</b>	Attending to culture requires us to build a positive learning culture focused on continuous improvement and growth for all. By clarifying our expectations of behaviours and practices we create an environment where stakeholders are involved in transforming the capacity of their system to improve outcomes.
<b>Listen and Respond</b>	Listening and responding requires us to be customer focused by purposefully developing processes to understand the current and future needs of learners and stakeholders. In connecting to and understanding these needs and aspirations we are able to take action, direct our improvement efforts and target strategies to support improvement.
<b>Make Data Count</b>	We make data count when we collect and interrogate the key data required to strategically evaluate and improve outcomes. Inquiring into multiple measures of data enables us to consider aspects of our context, practices and performance to direct future actions.
<b>Set Direction</b>	By setting directions based on agreed values, vision and purpose, developed with stakeholders, we identify priorities and actions to continuously improve outcomes. Effective planning processes support the achievement of DECS directions and local priorities.
<b>Target Resources</b>	Targeting resources requires us to align resources effectively and innovatively to reflect our priorities and provide the best opportunity to achieve desired outcomes.
<b>Continuously Improve</b>	We continuously improve when we seek to improve processes and programs across the systems in which we work. Through developing successful and known processes we are better able to ensure continuous improvement for individuals, the system and a sustainable future.

## Attachment 5

**School Improvement Framework  
Tasmanian Catholic Education Commission (2011)**

Five domains and twenty-two elements of the framework

Domain	Description
<b>Catholic Culture and Ethos</b>	<p>Centred on Christ, Catholic culture and ethos is the starting point for all growth and development in a Catholic school. The school draws life from its Vision and Mission, the Church and the school's own tradition. It focuses on building a positive school climate and engagement with the community, and is enlivened by the spiritual formation of all who form part of the school.</p> <p>elements: <i>vision and mission; tradition; school climate; community; spiritual formation</i></p>
<b>Leadership</b>	<p>Leadership in a Catholic school encompasses positional or designated leaders, staff, students and members of the wider school community working together, sharing responsibility for improving the school so that all, and most particularly students, flourish and learn. A Catholic school requires strong Religious and Spiritual Leadership, characterised by self awareness and interpersonal and managerial capabilities that harness the dreams, drive and energy of the school.</p> <p>elements: <i>leading the catholic school community; leading self; leading others; leading continuous change; leading the school organisation</i></p>
<b>Learning Environment</b>	<p>The learning environment of the Catholic school focuses on the wellbeing of all, the engagement of students in their learning, support for development of positive behaviours and the recognition of parents/ caregivers as the first educators of their children. Positive relationships and stimulating learning opportunities are essential for a quality learning environment.</p> <p>elements: <i>student, staff and community wellbeing; parent/school partnership; student engagement in learning; positive behaviours</i></p>
<b>Curriculum and Teaching</b>	<p>Curriculum and teaching provide the stimulus, support and structure that students need to make optimum learning progress. Quality teaching, rich, appropriate learning programs, a curriculum that has relevance, breadth and depth in all learning areas, including Religious Education, combine to provide students with multiple, high quality learning experiences.</p> <p>elements: <i>quality teaching; learning and teaching programs; curriculum; religious education; staff professional learning</i></p>
<b>Student Learning and Achievement</b>	<p>Student learning and achievement is at the heart of all schools – it is why they exist. The belief and expectation that all students can learn is fundamental to the school providing inclusive, effective learning opportunities for all. Students achieve when they are active participants and see the relevance of their learning, when they are provided with ongoing feedback on their learning progress and when supported by parents informed about their child's strengths and areas of need.</p> <p>elements: <i>student achievement; assessment, reporting &amp; feedback; equity</i></p>

## Attachment 6

**The School Improvement and Accountability Framework  
Western Australia (2008)**

Areas of focus for effective school operations

Area of Focus	Description
<b>Teaching</b>	High quality teaching is a pre-requisite for improving student achievement. It is therefore the responsibility of schools, with system support, to develop teachers with a purposeful pedagogy that challenges but supports students and connects to their experience, stage of development and background.
<b>Learning Environment</b>	A learning environment that is safe, caring and inclusive is also a pre-requisite for improving the standards of student achievement. It is the responsibility of schools, with system support, to establish a learning environment where students feel safe, where student behaviour is well managed, where pastoral care programs promote student wellbeing and where attendance, retention and engagement are priorities. Attendance and behaviour data (including suspension and exclusion data) will be important indicators of the health of the learning environment.
<b>Leadership</b>	School leadership is an important enabler of improved student achievement. Leaders who focus on teacher pedagogy and practice are likely to have the greatest impact on student achievement. By encouraging expert teachers to operate collaboratively and share their skills and understandings, good school leaders nurture, develop and expand quality teaching in the school.
<b>Resources</b>	The strategic deployment of school resources is another important enabler of improved student achievement. This will be measured by the extent to which resources are targeted, through school improvement planning, to maximise student achievement. It is expected that school resources are managed in accordance with legislative and policy imperatives. This is subject to audit processes, and reports identifying deficiencies will be addressed through performance management.
<b>Relationships</b>	Building strong internal and external relationships is another enabler of improved student achievement. Effective internal relationships between staff members and between staff and students are important in supporting improved student achievement because they facilitate the alignment of school goals and the capacity to enact whole-school approaches and consistent practice. Public schools are part of the community in which they are located and because they provide a crucial service to that community it is both a requirement and advantageous for schools to build a strong relationship with their community capable of responding to local needs and shaping a sense of pride in the school. Community relationships are not limited to parents. Local government, business, other educational institutions, and government and non-government agencies can all be important partners with schools.

## Attachment 7

### The NSW DET Analytical Framework for Effective Leadership and School Improvement in Literacy and Numeracy (2011)

Five leadership dimensions (based on Robinson, 2007) and twenty-five statements

Dimension	Statement
<b>Strategic Resourcing</b>	<ul style="list-style-type: none"> <li>• Human resources are allocated and managed across the whole school so they support the continuous improvement of student achievements in literacy and numeracy.</li> <li>• Decisions about school/class timetables and calendars are made to maximise instructional time for literacy and numeracy and to facilitate staff collaboration.</li> <li>• Teaching resources including ICT materials are allocated and managed across the whole school so they support the continuous improvement of student achievements in literacy and numeracy.</li> </ul>
<b>Establishing Goals and Expectations</b>	<ul style="list-style-type: none"> <li>• There is clarity and consensus about the school's goals and expectations in literacy and numeracy.</li> <li>• A school plan is developed collaboratively with staff and community members with an emphasis on literacy and numeracy.</li> <li>• Whole school goals, targets and priorities in literacy and numeracy are translated into explicit policies and classroom practices.</li> <li>• Literacy and numeracy targets in school plans are effective and SMART —Specific, Measurable, Achievable, Relevant and Time framed.</li> <li>• The school plan sets improvement priorities in literacy based on a thorough evaluation of student performance data.</li> <li>• The school plan sets improvement priorities in numeracy based on a thorough evaluation of student performance data.</li> </ul>
<b>Ensuring an Orderly and Supportive Environment</b>	<ul style="list-style-type: none"> <li>• The school addresses the literacy and numeracy needs of students in an orderly and supportive environment.</li> <li>• The school recognises, values and supports the literacy and numeracy needs of a diversity of student learners including Aboriginal and Torres Strait Islander students, students with special needs and non-English speaking students.</li> </ul>
<b>Planning, Coordinating and Evaluating Teaching and the Curriculum</b>	<ul style="list-style-type: none"> <li>• Quality literacy practices are identified and implemented across the school.</li> <li>• Quality numeracy practices are identified and implemented across the school.</li> <li>• Literacy skills are taught across all key learning areas.</li> <li>• Numeracy skills are taught across all key learning areas.</li> <li>• School leaders monitor the collection and analysis of students' literacy and numeracy performance data.</li> <li>• Staff members collect and analyse students' literacy and numeracy performance data.</li> </ul>
<b>Promoting and Participating in Teacher Learning and Development</b>	<ul style="list-style-type: none"> <li>• School leaders actively participate in learning about literacy and numeracy.</li> <li>• School leaders promote staff discussion and dialogue about literacy and numeracy to have an impact on programs and practices.</li> <li>• School leaders are seen by staff as a source of advice and are knowledgeable about literacy and numeracy.</li> <li>• Professional learning in literacy and numeracy is strategically planned.</li> <li>• The school actively seeks and promotes opportunities to develop leadership capacity in literacy and numeracy.</li> </ul>
	<ul style="list-style-type: none"> <li>• The school develops effective partnerships with parents and carers to develop and promote literacy and numeracy.</li> <li>• Student achievements in literacy and numeracy are reported to parents and carers.</li> <li>• School planning and whole school improvements in literacy and numeracy are shared with the school community.</li> </ul>