SCHOOL EFFECTIVENESS AND IMPROVEMENT:
AN INTERNATIONAL AND AUSTRALIAN PERSPECTIVE

Peter W. Hill
Centre for Applied Educational Research
Faculty of Education, The University of Melbourne

High performance schooling

In developing economies in the early stages of provision of universal public education, the focus is typically on securing sufficient resources to ensure minimum standards of provision. In Australia, it is widely considered that such standards were reached towards the end of the ‘seventies in the case of primary schooling, following a sustained effort by both Commonwealth and State and Territory governments to accommodate significant growth in student numbers and to overcome serious deficiencies in resourcing levels in Australian schools.

At the post-compulsory level, targets for near-universal completion of Year 12 or its equivalent were set during the middle and late ‘eighties and a rapid increase in retention rates took place during that decade. This numerical expansion took place in a context of dramatic restructuring within the workplace, the collapse of the youth labour market, a prolonged economic recession and a demographic decline in the total school population.

In mature economies, the focus shifts from provision and inputs to ensuring value for money and to justifying increased inputs in terms of the outputs that they generate. In other words, the emphasis is more on effectiveness, productivity, efficiency and quality. The term ‘high performance schooling’ is used here to describe schools that consistently exhibit these characteristics.

Our knowledge base regarding high performance schooling has emerged primarily through research into school effectiveness and improvement. My aim in this address is to review that knowledge base and in particular to supplement findings from international research, which is mostly North America or Europe in origin, with findings from a large-scale Australian study which my colleagues and I have been engaged in over the last three years.

My contention is that the search for the effective school may have been somewhat misguided and that researchers may have been looking in the wrong place for clues to improvement. I say this because a large number of studies are now revealing that it is at the level of classroom teaching that the most powerful knowledge about effectiveness is to be had. On the other hand, I also share the optimism of those such as Sam Stringfield in the USA and David Reynolds in the UK who believe that we are now entering a new phase in which we can seek to reduce that within-school variability and make the notion of high performance schooling a reality rather than an aspiration.
School effectiveness research - a brief history

Inquiry into school effectiveness is concerned with assessing the extent to which schools achieve their goals; and of understanding the characteristics of those schools in which students make greater progress than would be expected from a consideration of their intakes. One might imagine that over the years educators would have developed a detailed understanding of these issues. In fact, research into school effectiveness is of relatively recent origin and the knowledge base is still partial and tentative.

There are two broad categories of questions that have been addressed through research into school effectiveness. The first, prompted particularly by economic considerations and the perceived need to obtain ‘value for money’ from expenditure on education, has been concerned with the issue of “How much?” The second, prompted primarily from sociological considerations and the perceived need for equality of opportunity, has been concerned with the issue of “For whom?”

Historically, the more important of the early studies of school effectiveness -- notably the studies by Coleman and colleagues in 1966 and by Jencks and colleagues in 1972 -- were of the latter kind. They were interested in issues of equity and the influence of the school relative to that of sociologically-determined background characteristics of students. Other studies have tended to reflect economic considerations and pressures for accountability and have thus focussed more on the issue of the magnitude of school effects. Of course, both issues are important. It is important to seek to achieve substantial improvements in outcomes for all students.

Methodologically, measuring school effectiveness is immensely difficult. Schools are large, complex organisations and both logistical and ethical considerations preclude investigation by means of randomised experimentation. Evidence from large-scale quantitative studies has been correlational and has rarely allowed causal conclusions to be made. Studies have tended to be cross-sectional rather than longitudinal, which also makes causal inference problematic. In addition, the hierarchical organisation of education, with students subject to the influences of grouping within classes within schools, has meant that traditional analytic techniques have had to be abandoned in favour of new ‘multilevel’ modelling methods.

Some of the better known insights into effective schools have come not from large-scale quantitative studies, but from case studies of individual schools identified as clearly performing much better than other schools with which they could be legitimately compared, taking into account the background characteristics of students. During the late ‘seventies, studies such as those of Ron Edmonds and Michael Rutter and colleagues were published, which found that a common characteristic of effective schools is that they share an ethos or culture that is oriented towards learning, manifest in terms of high expectations of students, the setting of high standards, an emphasis on the basics, a high level of involvement in decision-making and professionalism amongst teachers, cohesiveness and clear policies on matters such as homework and student behaviour, and so on. They also found that effective schools (at least in the American and British contexts) are characterised by strong educational
leadership, particularly as exercised by the principal and directed at establishing agreed goals, increasing competence and involvement of staff and clarifying roles and expectations.

A limitation of this case study research is that while it provides some rich descriptions of effective schools, it sheds little light on how those schools became effective in the first place. It is now evident that seeking to replicate the characteristics of highly effective schools in less effective schools is an ineffective strategy for bringing about school improvement.

In recent years, researchers have turned their attention to designing so-called ‘state-of-the-art’ studies of school effectiveness. These studies are:

- multi-method in that they make use of both qualitative and quantitative methods;
- multi-level in that they make use of sampling designs and analytic techniques that take into account the organisation of students within classes within schools;
- longitudinal in that they follow students’ progress over several years; and
- multivariate, in that they include measures of a range of student achievements, behaviours and attitudes.

Value-added measures

Central to the notion of school effectiveness is the concept that what is to be measured is the value-added contribution of the school. Gross achievement measures do not provide an accurate indication of this contribution because of disparities in the background characteristics of students entering different schools. This is not to say that gross achievement measures are of little value. Indeed, they can be very useful in terms of giving an indication of absolute standards.

For example, as part of our ongoing research into school and teacher effectiveness, my colleagues Ken Rowe, Philip Holmes-Smith, Jean Russell and I have been collecting information on gross achievement levels for large samples of students based on the Victorian Profiles. Information is now available on achievement levels for all the years of schooling from Preparatory (P) to Year 11 for the various strands of English and mathematics. For example, Figure 1 summarises patterns of achievement for reading. By collecting this kind of information, it is possible to identify key trends, such as:

- the rapid rate of learning over the early years of schooling (Years P-3);
- the increasing spread of performance as students progress through school;
- the negligible increase in performance levels between Years 5 and 8;
- the decline in levels of performance of lower-achieving students as they move from primary to secondary school (Year 6 to Year 7);
- the minimal increase in performance levels of the bottom ten percent of students between Years 4 and 10; and
- the pick-up in progress between Years 10 and 11.
This information can be useful, both at the system level and at the school level, in understanding overall patterns of achievement and in identifying problem areas. For the purposes of assessing school effectiveness, however, it is necessary to obtain measures of achievement that have been adjusted to take into account the intake characteristics of students.

There are at least three kinds of value-added measures that are used in studying school effectiveness\(^7\). The first is what might be called “Unpredicted Achievement”, defined as level of attainment of students adjusted for family background characteristics such as socio-economic status and non-English speaking background, and for student ability, using a general ability measure such as a measure of verbal reasoning. The second is what might be described as “Learning Gain”, defined as the level of attainment of students adjusted for their prior levels of attainment. The third is “Net Progress”, namely level of attainment adjusted for family background, ability and initial level of attainment.

The methodology of obtaining value-added measures involves estimating that part of a student’s score on an outcome measure that cannot be predicted knowing his or her background and/or prior levels of ability and attainment. Much depends upon ensuring that adjustments have been made for all relevant intake characteristics and that each of these have been measured reliably. To the extent that this is not done, under-adjustments are likely to occur that will favour those schools with more advantaged intakes.
Magnitude of school effects

What can the research literature tell us regarding school effects and differences among schools in their effectiveness? The Coleman Report established a kind of benchmark for the magnitude of school effects when it reported that around 9% of the variance in achievement measures was accounted for by differences among schools. This estimate was much smaller than had been expected and led to the view that schools bring little influence to bear on children's achievement that is independent of their background and general social context.

It would seem that this estimate may not have been wide of the mark, however, notwithstanding the analytic techniques available to Coleman and colleagues in the mid-'sixties. In a review of recent studies of school effectiveness, David Reynolds and colleagues conclude that school effects are indeed quite small, with only 8-15% of the variation in student outcomes due to school differences. A systematic meta analysis reported earlier this year by Bosker and Witziers of 75 British and Dutch studies which met strict methodological standards and which were completed after 1986, found that net or value-added school effects account for a mere 7% of the total variance, which is at the lower end of the range given by Reynolds and colleagues.

Evidence recently to hand suggests that within the Victorian context, between-school differences may be somewhat larger than those typically found in the literature and particularly by British and Dutch studies. Over the past three years, my colleagues and I at the Centre for Applied Educational Research have been engaged in carrying out a three-year longitudinal study of school and teacher effectiveness with the financial support of the Directorate of School Education and in cooperation with 90 primary and secondary schools randomly selected from the Government, Catholic and Independent school sectors.

Within the Victorian Quality Schools Project, measures of student progress have been obtained using teacher ratings of student achievement based on the Victorian Profiles for English and mathematics. Adopting a 'learning gain' definition of effectiveness and adjusting achievement levels at the end of 1993 for achievement levels at the end of 1992, school effect sizes of 16-19% have been obtained, with negligible differences between primary and secondary schools but with slightly larger effect sizes for mathematics than for English.

In another ongoing study of the Centre into Year 12 student performance in the 1994 Victorian Certificate of Education (VCE), students' total scores in different studies, were adjusted for their scores on the various subsections of the General Achievement Test (GAT), which is administered half way through Year 12, to provide estimates of instructional effectiveness for the nine VCE studies with the largest enrolments. Some preliminary results are summarised in Table 1, which shows both gross (unadjusted) and net (adjusted) effect sizes. Net school effect sizes range from 12% (English and Biology) to 18% (Accounting).

Clearly, estimates of the magnitude of differences among schools in Victoria are towards the higher end of estimates of effectiveness reported in the international,
literature. The explanation would appear to relate in part to the curriculum-sensitive nature of the outcome measures used in both Victorian studies. Many studies reported in the international literature have employed as outcome measures tests of general academic ability (eg., verbal reasoning ability) which are not closely aligned to what is actually taught in classrooms. In studies which make use of such generalised outcome measures, between-school and between-class differences tend to be underestimated and the effects of home background exaggerated. On the other hand, estimates of effect sizes are higher when measures are used which are proximal and relate directly to school and classroom-based instruction, as do the Victorian Profiles and the various common assessment tasks comprising the VCE studies. This argument could possibly be extended to account for the somewhat higher estimates within the nine VCE studies for Accounting, Chemistry and Mathematics, relative to Human Development, Biology and English.

Table 1. School Effect Sizes in Nine VCE Studies with Large Enrolments

<table>
<thead>
<tr>
<th>Study</th>
<th>Gross</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting</td>
<td>24.1</td>
<td>18.0</td>
</tr>
<tr>
<td>Biology</td>
<td>22.2</td>
<td>11.8</td>
</tr>
<tr>
<td>Chemistry</td>
<td>26.0</td>
<td>16.8</td>
</tr>
<tr>
<td>English</td>
<td>23.0</td>
<td>11.6</td>
</tr>
<tr>
<td>Human Development - Home Economics</td>
<td>18.7</td>
<td>10.3</td>
</tr>
<tr>
<td>Information Processing and Management</td>
<td>22.1</td>
<td>15.1</td>
</tr>
<tr>
<td>Legal Studies</td>
<td>21.7</td>
<td>15.2</td>
</tr>
<tr>
<td>Further Mathematics</td>
<td>26.6</td>
<td>15.1</td>
</tr>
<tr>
<td>Mathematical Methods</td>
<td>21.9</td>
<td>15.3</td>
</tr>
</tbody>
</table>

Magnitude of class effects

To return to the international literature, one of the more powerful conclusions arising from recent research is that much of the variation between schools is in fact due to variation among classes. When the organisation of students in classes is taken into account, the unique variation due to schools over and above that due to differences among classes is very small indeed. This conclusion is reiterated in a recent review of research into education production functions by David Monk, who cites a number of studies in support of the observation that: "One of the recurring and most compelling findings within the corpus of production function research is the demonstration that how much a student learns depends on the identity of the classroom to which that student is assigned." 12

How large are between-class differences? One of the more significant studies to provide evidence regarding the importance of class effects is that of Jaap Scheerens and colleagues. 13 This study presents findings from a secondary analysis of data from the Second International Mathematics Study (SIMS). The findings, as summarised in Table 2, indicate that for the nine countries for which between-class information is available, estimates of within-school class effects in many countries exceed 40%, while school effects are significantly smaller, ranging between 0-15%.
Preliminary results from the Victorian Quality Schools Project confirm this phenomenon. When the data are analysed in such a way as to take into account the organisation of students within classes within schools, estimates of school and class effect sizes as summarised in Table 3 are obtained. Within-school class effects range from 38-45% for English and 53-55% for mathematics, whereas school effects over and above those due to differences among classes shrink to 4-9%. This is not to say that there are not substantial differences among schools in terms of their effectiveness, but rather that these differences are largely accounted for by internal within-school variation among classes and teachers.

Table 2. Class/Teacher- and School-Level Effects in Nine Countries*
(Secondary Mathematics Scores Adjusted for Father's Occupation)

<table>
<thead>
<tr>
<th>Country</th>
<th>Classes (%)</th>
<th>Schools (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Finland</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>France</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>Israel</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>29</td>
<td>15</td>
</tr>
<tr>
<td>New Zealand</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Scotland</td>
<td>31</td>
<td>5</td>
</tr>
<tr>
<td>Sweden</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>USA</td>
<td>45</td>
<td>9</td>
</tr>
</tbody>
</table>

* Source: Scheerens et al. (1989), p.794

Table 3. School and Class Effect Sizes for Victorian Schools:
Achievement Adjusted for Prior Achievement

<table>
<thead>
<tr>
<th>Country</th>
<th>Classes within schools (%)</th>
<th>Among schools (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>45.4</td>
<td>8.6</td>
</tr>
<tr>
<td>Secondary</td>
<td>37.8</td>
<td>7.4</td>
</tr>
<tr>
<td>Mathematics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>54.7</td>
<td>4.1</td>
</tr>
<tr>
<td>Secondary</td>
<td>52.7</td>
<td>8.4</td>
</tr>
</tbody>
</table>

Once again, these estimates are towards the higher end of similar estimates in the international literature. This is likely to be in part due to the curriculum sensitivity of the outcome measures used and in part of the type of adjustment employed. By adopting a 'learning gain' definition of effectiveness and adjusting for achievement obtained one year earlier, it is to be expected that estimates of class effects will be high and school effects reduced. Table 3 nevertheless serves to emphasise that it is at the level of the classroom that learning takes place and that there can be very substantial differences in the progress made by students in different classes within the same school.
**Consistency of effects over outcomes**

Yet another conclusion to be drawn from the international literature is that schools are not usually effective across the board, but tend to have strengths and weaknesses across the range of outcome measures. Considerable evidence for this finding has emerged in UK research into value-added measures of achievement using as outcome measures secondary students' results in GCSE and 'A'-Level public examinations. This conclusion contradicts earlier beliefs that effective schools are consistently effective and points to the fact that within secondary schools there are often significant between-faculty or department differences in effectiveness.

Evidence from the Victorian Quality Schools Project supports this finding. The correlation between value-added measures for 51 primary schools in English and mathematics was 0.64, indicating that primary schools are by no means equally effective across these two core areas of the curriculum. In addition, because in every case the students in these 51 schools were taught by the same teacher for both English and mathematics, it is evident that teachers also are differentially effective in different areas of the curriculum.

Analyses carried out using the Victorian Board of Studies 1994 Year 12 VCE data base reveal the same phenomenon of differential effectiveness across the curriculum. Correlations between school-level value-added measures for those nine subjects with the highest student enrolments range from as low as 0.32 for Accounting and Human Development - Home Economics to 0.71 for Chemistry and Mathematics Methods. For a total of 190 schools, their effectiveness estimates place them in the top decile of schools for one or more of the nine studies, but as Figure 2 indicates, only 22 schools are in the top ten per cent for four or more of these studies.

![Figure 2](attachment://image.png)

**Figure 2. Number of schools or other providers with VCE school effectiveness measures in the top decile in one or more of the nine VCE studies with the highest number of enrolled students**

These findings indicate that the notion of an effective school needs to be treated with caution. At this stage, there is little empirical evidence to suggest that there are schools that are highly effective in all areas of the curriculum. Indeed, such schools
may not exist. Most schools have particular strengths and there is a great deal of internal variation among classes.

**Consistency of effects over cohorts**

Most studies of school effectiveness involve samples of students from one or two year levels. In studies which have obtained data from several cohorts of students it is evident that there is significant variation in the value-added contribution of schools across different cohorts. In other words, there are systematic differences in the progress made by students in different year levels. This is apparent in studies that have analysed data for successive cohorts of final year students. The Victorian Quality Schools Project data indicate that, within primary schools, around 20% of the variance in value-added measures of student achievement in English and mathematics can be accounted for by differences among year cohorts.

**Equal effectiveness for different groups**

In the international literature, there is some evidence to suggest that schools are unequally effective for different groups of students within the same school, such as students of different ethnic groups, ability ranges and socio-economic status. This contradicts earlier beliefs that schools tend to be equally effective or ineffective for all students within the one school.

Evidence from the Victorian Quality Schools Project does not support this finding. In the case of primary school English attainment, girls make greater progress than boys, students from high socio-economic and socio-educational backgrounds make more progress than students from low socio-economic and socio-educational backgrounds, and classes with a high proportion of NESB students make less progress than classes with a low proportion of NESB students, although NESB is not significant at the student level. However, the effect of these factors is the same across all schools and all classes. In the case of primary mathematics attainment, the significant factors are also gender, student socio-economic and socio-educational background and class NESB, but with the difference that girls make rather less progress than boys. At the secondary level the significant factors are student socio-economic and socio-educational background and gender in the case of attainment in English and both student and class socio-economic and socio-educational background in the case of attainment in mathematics.

**Explaining Differences in Student Progress**

Studies consistently indicate that there is enormous variation in the rate of learning of different students. School and teacher effectiveness research attempts to develop models of educational effectiveness which show the influence of those student-, class-, school- and system-level characteristics that best account for these differences. As noted earlier, as a result of the early case studies of Edmonds and others, it was believed that considerable progress had been made in identifying the more salient
characteristics of effective schools. More recent quantitative research has indicated that the notion of a 'blueprint' or 'recipe' for effectiveness is simply not tenable and that researchers still have a long way to go in understanding fully what works and what does not work. There are nonetheless some powerful conclusions to be drawn from the vast body of well-designed studies into aspects of school and teacher effectiveness. This literature has recently been summarised by Jaap Scheerens\(^1\) whose main conclusions are summarised in Table 4.

First, there is multiple empirical research confirmation that structured teaching makes a positive difference. Structured teaching involves clear learning objectives, splitting teaching into manageable units that are then delivered in a planned sequence, the frequent use of prompts and advance organisers, ongoing and systematic monitoring, and feedback and intensive remediation to overcome difficulties. Structured teaching does not imply old-fashioned teaching, since it can occur within a modern, constructivist approach to teaching, but it is of the kind more suited to the early primary years and is indeed what is found at the core of many successful programs familiar to primary school teachers, such as Reading Recovery, the Early Literacy In-Service Course (ELIC), First Steps and others.

### Table 4. Research Evidence Regarding the Efficacy of Key School and Instructional Characteristics

<table>
<thead>
<tr>
<th>Evidence Level</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Considerable Evidence</td>
<td>Structured teaching; Effective learning time</td>
</tr>
<tr>
<td>Reasonable Evidence</td>
<td>Opportunity to learn; Pressure to achieve; High expectations;</td>
</tr>
<tr>
<td></td>
<td>Physical/material school characteristics (marginal effect);</td>
</tr>
<tr>
<td></td>
<td>Parental involvement</td>
</tr>
<tr>
<td>Doubtful Evidence:</td>
<td>Educational leadership; Evaluative potential of the school;</td>
</tr>
<tr>
<td></td>
<td>School climate; Organisational/structural preconditions;</td>
</tr>
<tr>
<td></td>
<td>Descriptive context characteristics (eg. size, location)</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Staff recruitment and staff development; External stimuli</td>
</tr>
</tbody>
</table>

(after Scheerens, 1992, pp.84-95)

Second, there is equally substantial empirical support for the importance of effective learning time, a finding which can be related to the fundamental principle identified by Carroll back in the early 'sixties\(^1\) that learning is in large measure a function of 'time on task'. Aspects of time include the quantum of time devoted to learning as
influenced by the length of the school day and the school year, the setting of homework, increasing the effectiveness of the time available so that as much as possible is on task, and increasing the learning time for a particular subject by reallocating the total time available. Naturally there are limits to the extent to which each of these can be increased or manipulated before such interventions become counterproductive.

There is a reasonable empirical base to the claims that opportunity to learn is important. This is in essence a question of the match between what is taught in the classroom and what is assessed for the purposes of making judgements about student progress. There is also reasonably good evidence regarding the importance of pressure to achieve and of high expectations of students.

In the case of the physical and material character of the school -- characteristics such as class size and per capita spending -- there is reasonable evidence that more resources are associated with improved outcomes, but that the magnitude of the improvement in learning is marginal within the parameters that can be contemplated at a system level. One noteworthy exception, canvassed in a recent UK review of class size by Peter Blachford and Peter Mortimer, may be student-teacher ratios in the very early years of primary schooling, where significantly smaller ratios can lead to significantly improved literacy rates. Unfortunately, as Blachford and Mortimer observe, increases in resources such as smaller classes do not always lead to changes in teaching behaviour such as greater use of one-to-one tutoring. Often in the past, they have been used to make the life of the classroom teacher more comfortable rather than to make teaching more effective. Thus, in the case of resources, the issue is not so much one of the quantum, but how they are used.

On the question of parental involvement, there is reasonable empirical influence that involvement in monitoring the child's progress at school, reading at home, checking up on homework, and so on, have a direct effect on learning.

Moving to those key characteristics for which there is doubtful empirical evidence, educational leadership, particularly as exercised by the school principal, is something for which it is difficult to demonstrate a direct effect on student learning. This is clear from a recent review of relevant studies by Philip Hallinger and Kenneth Leithwood. This should come as no surprise because, as my colleague Brian Caldwell has pointed out, while principals establish the preconditions for effective learning, it is classroom teachers who make it happen. In similar vein, the empirical evidence is doubtful when it comes to the importance of: the evaluative potential of the school, particularly whether it engages in regular assessment of students, appraisal of teaching performance and evaluation of programs; the climate of the school; organisation/structural preconditions, such as the degree of local autonomy and decision-making; and whether the school is large or small, rural or metropolitan. This is not to say that these things are not important but rather that they do not of themselves guarantee effective learning.

Finally, there are a number of characteristics for which there is little or no evidence either for or against their efficacy and therefore their impact remains largely hypothetical at this stage. These include staff recruitment and selection, staff
development programs and external stimuli to make schools effective, including increased accountability requirements, privatisation, output financing and differential pay incentives for effective teaching. These are the focus of educational reforms here in Australia and in systems across the world. While there may be good theoretical grounds for pursuing them, as yet we do not have empirical evidence either way as to their efficacy.

The above is a very brief review of empirical evidence regarding the importance of key school and instructional characteristics on learning. How applicable are these findings to the Australian context? So far within the Victorian Quality Schools Project preliminary findings confirm many of the conclusions summarised by Scheerens. In particular, it is evident that leadership and school climate, while important in terms of teacher morale and how teachers perceive their working lives, does not directly influence student achievement, although where there is a more positive school climate, teachers are more enthusiastic and energised, students in turn are more attentive and attentiveness is a significant predictor of learning. Our findings also indicate that school size and location are not important.

On the other hand, they indicate that student attentiveness and whether or not teachers have participated in structured professional development programs related to the teaching of literacy and numeracy are highly significant predictors of student learning progress. These two characteristics are closely implicated in the issues of 'time on task' and structured teaching and so in one sense this finding comes as no surprise, although it is to be noted that the VQSP is one of the first studies to find a strong relationship between professional development for teachers and effective teaching. The results also provide preliminary evidence suggesting that internal classroom organisation is important, especially for primary students in English, for whom it appears that composite classes may be associated with negative effects on learning. This is something that is being followed up more closely with a new wave of data and will be investigated more fully through detailed case study research.

**Implications for School Improvement**

The key motivation for undertaking research into school effectiveness is, of course, the facilitation of school improvement. In the past, it was often assumed that once a sound knowledge base existed regarding the characteristics of effective schools, all that was required in order to improve effectiveness was to ensure that all schools exhibited the characteristics of effective schools. There is now a more sophisticated understanding of school improvement as a process, and a recognition that successful improvement strategies for failing schools are very different from those that work in moderately effective and effective schools.²¹

In addition, while it is evident that research into school effectiveness cannot be reduced to simple 'blueprints' or 'recipes' for school improvement, there are nevertheless some powerful messages for policy-makers and school administrators seeking dramatic improvements in educational outcomes. The first of these messages relates to the power of information about school effectiveness as a catalyst for improvement and reform. All too frequently schools and systems have lacked credible information.
regarding their performance and effectiveness. This is changing. The trend now is towards the development of indicator systems which enable benchmarking of performance against external standards or reference points. The evidence from systems that have put in place indicator systems and more especially from schools that have begun to collect and use outcomes measures, is that such information is a powerful stimulant to school improvement. As yet, little if any use of value-added measures of effectiveness occurs outside research projects. Indeed, there is great reluctance by some within the profession to countenance any systematic collection of comprehensive and comparable achievement information. This situation is changing rapidly, however, with increasing recognition of the power of information to motivate and shape improvement efforts.

The second message relates to the reason why so many improvement initiatives in education fail to live up to initial expectations. Most reforms in education are directed at the preconditions for learning rather than at influencing behaviours within the classroom. Establishing preconditions is important. For example, it is desirable that schools have flexibility in the ways in which they utilise resources as the school level, including flexibility in the use of staffing resources. Thus present efforts to give schools greater control over their resources are to be commended. But improvements in learning outcomes are not a guaranteed outcome of providing such flexibility. This will only occur if the preconditions, for example staffing flexibility, are then used to effect changes in the way in which students are taught in class. Many reforms stop short of changing what happens beyond the classroom door and thus fail to deliver improved learning outcomes.

The third message to be gained from research into school effectiveness and especially from Australian research, is that schools and especially teachers do make a difference and that it is not so much what students bring with them but what they experience on a day-to-day basis in classrooms that really matters. Without question, consistent high-quality classroom teaching can deliver dramatic improvements in student learning. Quantum improvements in educational effectiveness would result if all classes and teachers were as effective as the most effective classes and teachers. The work of Sam Stringfield into schools as high reliability organisations points to important principles underlying high performance schooling. Programs such as the WA First Steps program and Robert Slavin's Success for All program provide practical illustrations of how internal variability in effectiveness can be reduced and learning outcomes dramatically improved, particularly among those most at risk of experiencing failure.

I recently had the privilege of listening to Dame Professor Marie Clay reflecting upon what she had learnt from many years of seeking to improve early literacy, particularly through her Reading Recovery program. She remarked that dramatic improvements in student learning are not limited by the capacity of students to profit from schooling. Nor are they limited by the capacity of teachers to profit from professional development programs directed at improving the quality and effectiveness of their teaching. Rather, they are dependent primarily upon leaders who have the understandings, commitment and resources to transform ordinary schools into high performance, adaptive learning organisations characterised by purpose, coherence, internal alignment in matters connected with teaching and student learning. This is the challenge that will confront educational leaders in the years ahead. It represents a
welcome change of emphasis away from structural and management reforms and
towards good teaching and instructional leadership as the keys to school improvement.


