Rationality, Reporting and Indicators: Improving School and Systemic Effectiveness Through Better Information Management

Tim Wyatt
BA, Dip Ed, LittB (NE) MA, MEd (Syd)

A portfolio submitted in fulfilment of the requirements for the degree of Doctor of Education,

University of Western Sydney, 1997.
PLEAS NOTE

The greatest amount of care has been taken while scanning this thesis,

and the best possible result has been obtained.
Contents

Acknowledgments .............................................................................................................. 4

Overview ......................................................................................................................... 5

Trends in the Education Indicators Literature ............................................................... 24

The Role of Indicators in School-Level Evaluation and Development ..................... 65

Some Observations on School-Based Management Programs in the United States ...... 92

An Evaluation of the Effects of School Based Management on Science and Math Education and on Students .......................................................... 105

Evaluating Outcomes for Customers in School Education ......................................... 134

Informing School Improvement Efforts: The Legacy of Quality Assurance Reviews ............................................................................................................................. 150

What Students Expect of Their Schools: Findings from Quality Assurance School Reviews ................................................................................................................. 174

The Achievements of Schools as Seen by Parents: Findings From Quality Assurance School Reviews ............................................................................................................................. 183

What Parents Want from Further Participation and Communication: Findings from Quality Assurance School Reviews ............................................................................................................................. 192

The Senior Curriculum and Subject Choice: Findings from Quality Assurance School Reviews ............................................................................................................................. 201

School effectiveness research: dead end, damp squib or smouldering fuse? ............ 211

Understanding Value Added .......................................................................................... 249

Recent Developments in Annual School Performance Reporting in NSW ............... 266
Acknowledgments

The papers presented in this portfolio are the product of work which has taken place over the better part of the past decade. Over that time my thinking has been influenced by many people, and it would be impossible to list them all. Many of the papers have benefited from suggestions, proof-reading and critiques by colleagues and anonymous referees. Despite their help, the remaining errors are mine alone.

I would particularly like to acknowledge the support and contributions of some key sources of influence. Alan Ruby and I collaborated on the development of models for using performance indicators in the mid to late 1980s. Without Alan’s leadership of the Reporting on Educational Progress project, none of the ensuing work is likely to have eventuated. While not involved in any of the papers here, colleagues Brian Davies and Sarah Shrub helped turn the ideas generated in this project into practice. Dr Jeanne Griffith, Dr Gary Phillips and Dr Harry Hatry in the United States all provided me with opportunities to further develop these ideas in the international context, and involved me in projects which allowed me to grow professionally, and provided a source for some of the papers included in the portfolio. Throughout all of this period, John Watt has been a sounding board for ideas and a source of sound advice and encouragement.

The majority of the papers included here were written during the time that I was employed in the Quality Assurance Directorate in the NSW Department of School Education. I am particularly grateful for the professional freedom and encouragement provide by Dr Neville Hightett, whose leadership allowed me (and others in the directorate) to attempt things which had not been possible in this school system, including the developmental work which underpins the current systemic reporting and outcomes analysis models. The contribution of Dr Peter Cuttance in establishing the intellectual framework for the NSW Quality Assurance program is also acknowledged.

I have benefited greatly from being part of the team working on the development of the school accountability and improvement program, which in many ways is the practical culmination of all this work. I wish to acknowledge the contribution of Patrick Dias, Dr Max Smith, Eric Jamieson, Dave Wasson and Bill Clark to the development of this model.

Finally, I would like to thank Professor Jim Walker and Professor Neil Baumgart, who as co-supervisors, have guided me in the completion of the portfolio. Jim, in particular, has provided feedback on early drafts of several of the papers, and constantly reminded me of the larger philosophical and conceptual frameworks within which the ideas contained in the papers operate.
Overview

Introduction

Throughout my career, firstly as a classroom teacher, then as a researcher and evaluator, and later as a senior executive, I have been troubled by the apparently widespread practice of arbitrary and capricious decision making throughout education systems, whether at the classroom, school, system or political level. As an illustration, as would-be classroom teachers, many of us believed that appointments to public schools were determined by the dartboard method – some clerk in a backroom threw a dart at a map of the state, and where it landed was where you went. Any such frivolous method, if it actually occurred, would obviously prove as unsatisfactory for the aspiring teacher as for the schools and students they are allocated to. Some 15 years later, and with the benefit of insider knowledge, I am not convinced that the New South Wales Department of School Education’s rigid staffing rules, so tortuously negotiated with the NSW Teachers’ Federation, have been any more successful than the apocryphal dartboard method, if the aim has been to deploy the state’s resources to achieve high level and equitable outcomes for all students.

This anecdote serves to introduce one of the major themes underpinning the work presented in this portfolio – the belief that for quantum improvement in the outcomes achieved by students to occur, decision making at all levels must be informed by better collection, analysis and application of information. This, of course, is not a new idea. The notion that actions should have a rational basis has a very long history, with origins traceable to Platonic thought (Fasano, 1994). The intention here is to present a collection of works which demonstrate the application and testing of this notion, drawing on a number of research traditions and settings to do so.
This same theme can be discerned in concepts as diverse as the reflective practitioner (Schon, 1983); learning organisations, (Senge, 1992); action research (Henry & Kemmis, 1985); school choice (Chubb & Moe, 1990), and Total Quality Management (see for example, Demming, 1986). It is also present in the myriad works concerning social indicators, performance indicators, corporate and strategic planning, and school choice and accountability.

Ruby (1994) notes that in the hearts of policy makers, statisticians and social scientists, there is a belief that information changes things for the better. The improvement comes because information enhances the capacity to control, monitor and evaluate, and, as a consequence, produce better outcomes. This belief has shaped the management practices of most public and private enterprises throughout the 20th century. Levin et al. (1990) argue that the historically dominant influence on educational management this century has been “scientific management”. Its first principle, as enunciated by F.W. Taylor in 1911, was that decision making should be scientific rather than intuitive. It should be based on observation, analysis and inference and guided by the goals of efficiency and rationality. While vastly different in style and substance of implementation, this principle is echoed in the total quality movement (Demming, 1986), the social indicators movement of the 1950s and 1960s, and the performance indicators movement in the 1980s and 1990s. As scientific management has given way to an “applied science” model in the 1980s (Levin et al., 1990), the emphasis has shifted to the use of information to identify and define problems and then select a course of action. This model uses scientific information, as well as common sense and practical knowledge. It links information and the quality of education and “extends the demand for the development of more extensive, and more imaginative, education information systems than planners have used in the past” (p.78).

The papers in this portfolio are part of my efforts to “make things better” in education, and are the products of over a decade’s thinking and research. While most of the papers are written from within the context of the state school system in NSW, their
messages apply equally well to school systems elsewhere. The papers reflect an eclectic mix of research perspectives and methodological paradigms.

Schools and school systems are awash with data. Ashenden (1987) notes that for more than a century, education departments and ministries have amassed volumes of statistics on pupil enrolment numbers, attendance and so on. Likewise, in the United States, Kaagan and Smith (1985) observe that each year, the U.S. Department of Education produces a three-foot thick stack of reports that contain thousands of statistics, but that useful education indicators were missing. In their words, the situation was “a shambles”.¹ In NSW schools, teachers have been prodigious recorders of all kinds of test results, observational checklist, and anecdotal records about student’s behaviour and performance. Yet most of this data does not pass beyond the classroom, is rarely shared with other teachers in the school, and almost never beyond the school (including transfer of records from primary to secondary schools).²

A consistent finding in the literature on organisational change (eg. Kanter, 1985), is that improvement is dependent on a shared understanding of goals, desired outcomes and progress in achieving outcomes. Joyce et al. (1993) for example, identify the following as characteristics of effective large scale school improvement initiatives:

- All have focused on specific student learning goals. None have had only general goals of the “to make exam results go up” variety;
- All have employed strategies tailored to their goals and backed by rationales grounded in theory or research or a combination of these;
- All have measured learning outcomes on a formative and summative basis, collecting information about student gains on a regular basis and not leaving evaluation to a yearly examination; and
- All have employed substantial amounts of staff development in recognition that the initiative involved teacher and student learning.

¹ Since then a considerable amount of resources and effort have been devoted to turning this situation around. The US National Center for Education Statistics has re-oriented the focus of its work to concentrate on the production of indicator information.

² For further discussion of assessment practices in NSW schools see Chapter 7 in this portfolio.
Information systems which ultimately inform operational change or affect behaviour in an organisation are essential if an improvement culture is to take hold. A more advanced variant on this theme envisages schools developing a quality systems approach, using both qualitative and quantitative data systems, rather than the current ad hoc arrangements. However, as Murgatroyd and Morgan (1993) suggest, many schools know little about their operations in measurable terms. The data systems they do have are poor at informing quality improvement, and many exist only to meet the demands of record keeping.

Hopkins and Leask (1989) see the enhancement of the quality of education as dependent on measuring two things – student outcomes and the internal conditions of schools, such as the capacity for change. However, as the recent OECD report “Schools and Quality” makes clear, student learning outcomes have remained remarkably impervious to the efforts of educational authorities to change. Three decades of production function research (eg Hanushek, 1986) demonstrates that the relationship between education inputs, such as class sizes, and educational outcomes is extremely weak. The two variables that appear to make the most difference are the quality of the interaction between a particular student and a particular teacher (Rowe et al., 1995) and the existence of preconditions which establish an improvement culture within the school (Stoll and Fink, 1993).

This analysis suggests that student learning outcomes may be improved if schools specify goals and objectives, make the necessary resources available and establish accountability mechanisms for monitoring and measuring performance. In other words, schools should know exactly what they are trying to achieve, they know what success looks like, they map their progress, measure outcomes against objectives, and use this information to improve. Data systems become information systems when they are subjected to analysis and synthesis and are used to focus on developmental purposes. The analysis and synthesis generates critical knowledge, that is, knowledge which is explanatory and interpretive (see Sirotnik, 1987). The desire to assist schools
(and the school system as a whole) to achieve a position where they are able to (firstly) assemble their own data systems, and secondly to conduct their own analysis of this data and report this to the school community has been the common theme underpinning all of my professional work.

The first paper included here is a synthesis of the literature on education indicators\(^3\). It provides a short history of the performance indicator movement, summarises the definitions, potential uses and limitations of indicators as tools for improved accountability and school development, and concludes with a brief description of some of the attempts to put indicators into action around the world. This paper is a distillation and extension of the knowledge gained through the *Reporting on Educational Progress* project, which was jointly funded by the (then) Australian Conference of Directors–General of Education and the Commonwealth Department of Employment, Education and Training. (This project produced a series of publications around the theme of performance indicators in education, which are not included here. See Ruby *et al.*, 1989 for further details of this project).

This paper provides a context for the works that follow, and illustrates and encapsulates the essence of the application of the rationalist paradigm which underpins the indicators and reporting approach.

The second paper included in this set was initially delivered to a conference on school-level evaluation in Lillehammer, Norway, and subsequently published as conference proceedings by Oppland College.\(^4\) In contrast to the earlier paper, which was written from the perspective of school systems at the macro-level, this paper attempts

---


to link indicators with evaluation and planning at the micro-level. It argues that well-designed indicators have a legitimate place as tools for school self-evaluation, as part of annual planning cycles, and for the continuous monitoring of progress towards school objectives and for reporting of school achievements. A number of projects conducted in NSW which have attempted to implement these principles are then described. The important contribution of this paper is its articulation of the belief that the process of developing indicators at school level is just as important, or more so, than the indicators themselves. As the examples cited illustrate, the requirement for schools to develop indicators served in many cases as the first real opportunity for school staff, as a body, to discuss priorities and articulate objectives, and to develop a shared understanding of the school’s purpose and intentions. The paper argues that this power to stimulate educational debate is weakened when schools must simply comply with requests for data to satisfy the requirements of a centrally determined list of indicators.

The power of this process has been diluted over time, as the inclusion of performance indicators (or at least a column headed “performance indicators”) has become standard practice in the strategic plans of NSW public schools. Evidence from a statewide evaluation of school based planning, and from a synthesis of the relevant findings of over 1,500 school reviews, suggests that there has been no real attempt to use indicators as part of a wider system for monitoring school performance or for driving improvement. In producing these plans, there has been little discussion of the meaning of the indicators or evaluation of their adequacy. Despite the hopes, and fears, that the development of indicators will lead to a greater valuing of what is measured (David, 1988), little evidence can be found that this is what has transpired in NSW.

The third paper included here is used to illustrate the weakness of reform initiatives that do not have explicit mechanisms for defining the objectives of the reform and for monitoring achievement of progress towards these reforms. The paper draws on some of the findings of an evaluation of school based management (SBM) in several school districts across the United States, of which the author was a member of the
evaluation team. Proponents of reform of school organisation and management argue that change is needed to raise the level of student outcomes (see for example, Picot, 1988; Scott, 1989). Yet, in the school districts included in this study, the linkage between improved outcomes and administrative change was far from clear: indeed the majority of schools had not communicated any expectations for change to the school community, and had not developed any means by which differential outcomes could be tracked, monitored or reported. This is not to say that there were no improvements to either the efficiency or quality of the schooling process as a result of the introduction of SBM (indeed, in what were previously tightly controlled and rule-bound systems there were many), but the overall impression gained was that a great deal of time and energy was wasted because of the lack of a specific focus.

The papers presented thus far have argued the case for greater use of one particular means of information transmission – performance indicators which are largely quantitative, and most often displayed in graphical format. Later papers in this portfolio return to this theme. The fourth paper presented here describes an alternative approach, based on qualitative data analysis and reporting, to show how this has been used in attempts to influence policy making. The paper discusses early attempts to synthesise the findings and recommendations of the first year of Quality Assurance School Reviews, and describes the development of computer software to assist with the large volumes of data generated through the QA school reviews. Since the QA reports are public documents, and matters of historical record, collectively they provide a rich database for the analysis of the strengths and areas for development in one school system at a particular point in time. This paper uses an analysis of one important

---

5 Published as Chapter 14 in Hatry, H.P., Morley, E., Ashford, B. and Wyatt, T. (1993). Implementing School Based Management. Washington DC: Urban Institute Press. In this presentation, an unpublished paper is used to provide a context for the specific findings about the outcomes of SBM.

element of education practice – assessment – to illustrate the kinds of conclusions that can be drawn.

Influencing policy making involves more than simply putting data in front of decision makers and hoping that they will draw the “right” conclusions, which then dictates the “correct” course of action. The indicators movement has attempted to ensure that consideration of data will be at least part of the decision-making agenda. Recent efforts have focused attention on technical concerns about the way the data are formatted and presented to increase the likelihood that these data will be incorporated into the decision maker’s understanding of the issues at hand. Yet even the most ardent of indicator enthusiasts would concede that policy making is a largely irrational process. Policy makers base their decisions on their own prejudices, hunches, opinions and guesses (or phog as Cavenagh (1995) refers to it). This phog may be partly informed by formal sources of knowledge, but is more influenced by the policy-maker’s own observations (however limited), by second or third-hand rumours, innuendo and comments expressed by others, and by the dictates of political ideology or organisational philosophy (see Smithson, 1989 for a more sophisticated discussion of the epistemology of ignorance).

Several authors have discussed the problem of influencing change in school practices when teachers’ tacit knowledge, (gained through experience) conflicts with formal knowledge, gained through research and theory. The continued popularity of grade repetition (Shepard & Smith, 1989) and tracking (or streaming as it is sometimes called) (Oakes, 1992), despite the overwhelming research evidence of the harmful effects of these practices, are often cited as significant examples of this phenomenon. Policy makers are no different from teachers in this regard: when confronted with evidence which conflicts with their own world view their reactions will be quite predictable. These reactions will include denial, calls for rejection of the evidence, attacks on the credibility of the messenger, the methodology of the data gathering, and validity of the findings. Any serious attempt to influence policy making must therefore pay attention
to the mechanisms that create and sustain the world views of the decision makers. The implications of this for proponents of information-driven reform are first, that the "knowledge" that will ultimately reside in the hearts and minds of policy makers must also be made available to the sources of influence on these policy makers, which will include a wide constituency of citizens, media operators and so on; second, that multiple forms of information transmission will be required to reach each of these audiences; and third, the information must be presented in a form relevant to how policy makers see their problems.

One of the features that distinguished the NSW Quality Assurance (QA) program from other school inspection and review schemes elsewhere was the attempt made to bring information from the classroom into the public domain. The school review reports themselves served this purpose at the micro-level, but more powerful information was generated from the systemic meta-evaluation of these findings. As described in Manefield and Wyatt (1995) and Wyatt (1994), the results of these meta-analyses took several forms, including significant input into the DSEs annual planning process. The impact of this first hand information about the status of schooling – available to decision makers in NSW for the first time – is reflected in the choice of the resulting strategic priorities (see DSE Agenda 94; 95). The primary purpose of these meta-analyses was to provide accountability to the community (see DSE Quality Assurance Review Report, 1993). Part of the QA philosophy was also to ensure that the information generated from the meta-analyses was made available to teachers and parents. In terms of the preceding discussion concerning influences on policy, by ensuring that all stakeholders in the education process are aware of the issues confronting the school system, a milieu in which consistent messages would reach policy-makers would be reinforced. If this was not an explicit goal, it may well have been a useful by-product of a series of short papers written in 1995. Three of these were published⁷, and are included here together with one unpublished paper. With the

⁷ DSE (1995). Quality Assurance 20/20 Reports (No.1) What students expect of their schools; (No.2) The
demise of the QA program in 1996, the series came to a premature end, despite the two original papers being well received.  

Ideas from several movements impacting on the operation of both private and public organisations in the early 1990s have slowly begun to impact on education (although perhaps with less enthusiasm than elsewhere). Among the more influential of these ideas (as far as the NSW government was concerned) was the increased focus on customer service and on continuous improvement, both as raison d’être and as a fundamental operating principle for the public sector (see, for example, Premier’s Memorandum 95-45, 1995). As obvious as these notions may appear, for bureaucracies which for more than a century had seen their mission to be administration of pre-existing rules and regulations, the change was radical. Some agencies (particularly those which already had a commercial focus) embraced the change willingly. Other agencies, including schools, have been slower to respond.

The next paper included here draws attention to the need for education accountability schemes to acknowledge the customer service dimension\(^9\). Some aspects of customer service are technically simple to measure\(^10\), but the paper argues that in education, the identification of “customers” can be problematic. A change in cultural orientation will be needed before substantial improvement can occur in many schools (see Hargreaves, 1995).

\(^{\text{achievements of schools as seen by parents (No.3) What parents want from further participation and communication; and (No.4) The senior curriculum and subject choice (unpublished).}}\)

\(^{8}\) The published 20/20 Reports are now beginning to be cited in various places, such as the DSE Literacy Strategy.


This cultural change will not occur until the teaching profession (at least that in NSW) is able to confront squarely the issue of the effectiveness of teaching and school organisation. Despite several decades of international research on these issues (see for example, Witrock, (1986) on teacher effectiveness, and Sammons, et al., (1994) for reviews of the school effectiveness research), until relatively recently, NSW schools have not had access to the kinds of data which would allow them to make judgements about their effectiveness.

The next paper reviews the school effectiveness literature with a view to identifying a research agenda for NSW. As others have noted, a feature of this literature is the relatively small number of empirical works compared to critical and review articles. Of these empirical studies, many of the seminal works (eg Rutter et al., 1979) were published before the widespread availability of appropriate statistical techniques for estimating the unique contributions of schools and classrooms to student learning outcomes (Goldstein, 1987) Failure to take account of the multilevel structure of schooling has long been recognised as a source of considerable error in estimating school effects (Burstein, 1980), suggesting the need for caution in drawing specific conclusions from the earlier works. While the general consensus in the more recent research is that schools do make a difference (Scheerens, 1992), the lack of consistency in findings of the size and stability of effects suggests that findings may be very much context dependent.

Arising from the definition of school “effectiveness” as “schools in which pupils make greater progress than would be expected from consideration of its intake”, the primary research tool for exploring this issue has been the so-called “value added” measures. While the rationale for the value-added approach is discussed in this paper, a gap in the literature is the lack of a simple explanation of how value added measures are derived, why they are a measure of school effectiveness, and how they can be used to

---

inform school improvement. Most treatments on this topic assume a degree of technical knowledge not often found in most schools. The following paper is an attempt to fill this gap with a non-technical description of this topic.\textsuperscript{12}

As the above papers make clear, the data sets potentially available in NSW, compiled from the various external assessments conducted at years 3, 5, 10 and 12, provide a unique opportunity to contribute to scholarly understanding of the unresolved issues in school effectiveness research (Reynolds & Packer, 1992). The datasets also provide a unique opportunity to provide a sound underpinning for systemic accountability and improvement efforts. The model of accountability and improvement that will be implemented in NSW is still under development and final details are still the subject of negotiation, but the general principles are articulated in the final paper and in DSE documents (NSW DSE, 1996). What distinguishes this model from most others in place around the world, is the reporting of value-added measures of school performance at aggregate and course level, and the integration of school self-evaluation, annual reporting and external review. While drawing on the school effectiveness research to drive aspects of the scheme, it has a strong bias towards school improvement (of the kind described by Hopkins, Ainscow & West, 1994), rather than the application of five-point programs popular among early adherents of the effective schools movement (Edmonds, 1979). At the time of writing, the reports are not expected to present the outcomes data in the form of graphs, as described in the paper.

In many respects, the theoretical orientation of the accountability and improvement model described in this paper marks a return to that inherent in the indicators model described in paper 2. This latest model is advanced only by the details of its application – the basic rationalist paradigm remains intact. What then, are the prospects for success of any of these programs and models described in these collected

\textsuperscript{12} Understanding Value Added. To be published by the DSE as a Corporate Performance Directorate Occasional Paper.
prospects for success of any of these programs and models described in these collected papers. As recounted in paper 5, many have already fallen by the wayside, despite their apparent short term successes. They have failed to take hold in any systematic way (in NSW at least), or to bring about any lasting improvements in student learning outcomes. However, as numerous papers presented at the 1997 American Education Research Association attest, the ideas refuse to die.

The chances of success of this new model are enhanced by the coalescence of a number of forces. First, and most powerful, is the ever-growing demand for public accountability. While this has been a common refrain for more than a decade, in NSW the recent demonstrations of significant gaps in the accountability processes within education by the Woods Royal Commission, unremitting media pressure for publication of examination results, greater competition between the public and private school sectors encouraged by changes to the Commonwealth government’s funding formula and new schools policy, and revised state government procedures for examination of budget bids have all contributed to an environment in which the public demonstration of outcomes and value for money are necessary. Second, advances in technology, and the development of computer software that makes multi-level analysis feasible (Woodhouse, 1996), and in the NSW context, the investment of resources in developing longitudinal sets of student outcomes data means that it is possible to provide the indicator data on a large scale in a regular and timely manner. Third, the growing willingness of parents and the press to litigate means that policy makers and senior bureaucrats are being increasingly pressured to use data to justify their decisions. Fourth, there is a discernible shift in the previously defensive culture of decision makers in education (and throughout the public sector in NSW) that kept bad news secret, rather than openly acknowledged and responded to. Finally, there a recognition that improvements in student learning cannot be mandated. Programs that empower teachers and principals, and in which schools are supported as well as pressured, are far more likely to deliver change. By giving schools the responsibility for analysing and using information (within a systematic framework), the probability that the process will be
 accorded a seriousness of purpose within schools that is absent from schemes that demand only compliance with the publication of certain data is enhanced. It also ensures that the focus is on issues relevant to each school’s context rather than the sometimes abstract nature of statewide priorities.

The kind of integrated, holistic approach to school accountability and reporting envisaged here is a far cry from the league table approach adopted in the United Kingdom. When the (then) British Government published rankings of school’s raw examination results in 1992, the league tables were widely criticised within the education profession as being crude, unfair, and misleading. Much of the criticism of league tables is unique to the British context, and much of this criticism is ill-informed, unsubstantiated and as ideologically driven as the publication of league tables themselves. Nevertheless, league tables can be validly criticised on several grounds. The scheme outlined in the last paper hopefully avoids these pitfalls.

The lack of fairness in league tables comes about not from the data contained in the tables, but from the implicit school–with–school comparisons the tables invite, and in the conclusions which might be drawn from the tables. Assembling any data into a ranking assumes that the entities are alike in a fundamental way. A ranking of school examination performance contains the implicit assumption that we expect all schools to perform in the same way. In a differentiated school system, for example where there are specialist schools, this is clearly not the case. Schools are also not all alike in the amount of relative effort required to produce the same exam results. League tables of raw results do not provide an indication of the amount of improvement made by students, nor do they take account of schools’ different starting positions.

Because the league tables do not take account of the value added by schools, incorrect conclusions may be drawn. According to Barber (cited in Johnson, 1997)) the U.K. league tables imply that schools in poorer parts of the country which are performing relatively well, are failing. At the same time, they make schools in wealthier areas seem much better than they actually are. However, Goldstein and Thomas, (1995)
are equally critical of value-added league tables, because of (1) the inability to separate
the majority of schools – on any numerical scale – with any real accuracy; (2) the
retrospective nature of any school measures means the data is always out of date; (3)
the existence of differential effectiveness – some schools will enhance the performance
of certain kinds of students, say girls, and not other, and typically departments within a
school will differ; and (4) the variability of results over time. For these reasons, a single
year’s value added data should be treated as cautiously as a single year’s raw
examination results.

Many authors have recognised the danger of trying to reduce a school’s
performance to a single number (see for example Shavelson, 1987). Several indicators are
necessary. All education systems and all schools have multiple goals, which are not well
represented by simplistic rankings of exam results. This premise underlies the
construction of all major indicator systems throughout the world, illustrated by the
OECD’s indicator publication, Education at a Glance, and the U.S. Department of
Education’s Condition of Education.

While the British Secretary of State for Education, Gillian Shephard, argues that
publication of performance tables have been one of the most significant and profound
changes introduced to the education system over the past 10 years and have played a
major part in driving up standards, the consensus of opinion amongst school people is
not so positive (Johnston, 1997). Some teachers fear that the explicit focus on
performance will force schools to concentrate on high achievers while neglecting the less
able. If the motivational benefits of league tables are unclear, there is no evidence that
this is so either. More fundamentally, school performance rankings are useful at best at
reporting a school’s status (and only in a limited way), they can tell us nothing about
how the school got to be that way, and nothing about what needs to be done to improve
or how to improve.

The potential for performance indicators to deflect attention from what is really
important is also well known. This is summed up in the conclusion reached by an
editorial in The Times Education Supplement: “The danger of published [league tables] is that they focus too much effort on narrow targets and not enough on broader objectives. The answer to that is to make the indicators more comprehensive and more useful. The challenge to the public service, then, is to keep the indicators in perspective, to treat them as tools, not punishments and to help improve them next time round”.

There is thus a certain inevitability of the continued development and implementation of schemes of the kind outlined in these papers. Whether they are successful or not will depend on a range of factors, but the experience from other sectors suggests that educators too, will have to become more skilled and more rigorous in the systematic use of data in decision making. If the lessons offered by the programs described in the following papers is heeded, there is cause for some optimism.

There are others, though, who would argue that the seeds of self-destruction are sown in the managerialist paradigm inherent in the conceptualisation of any data-driven improvement initiative. While the approaches described in the various papers are some distance removed from Taylor’s principles of “scientific management, they do draw on the similar (perhaps naive) belief that, when confronted by data, school people will want to respond to it, to make things better. There are similar embedded assumptions that there will be common recognition of the legitimacy of the data in question, and shared understandings about the meaning and interpretation of data by all stakeholders. Experience suggests that, in NSW schools at least, this is far from evident (as demonstrated by the series of articles in the NSW Teacher’s Federation journal, *Education* during 1997).

I have argued earlier that when systems and schools are encouraged to use information as a means to improvement, they are being asked to operate in a fundamentally different way, in which the organisational culture changes from one concerned with administration of policies and curriculum to the management of people and resources. This is a change which many teachers and school communities find deeply problematic. The notion of “teacher as manager” or “principal as manager” is not one which sits very comfortably with the world view developed by
many of those currently teaching in NSW public schools, who learnt their trade at a
time when “teacher as professional” was the dominant model. Teachers are
notoriously resistant to change, as evidenced by the reluctance to abandon practices
shown by research to be harmful rather than helpful to children (such as grade
repetition and streaming). The psychological literature on cognitive dissonance
provides ample warning, at the individual level, of why information-driven
approaches may not be successful. There is a very common tendency for people to
simply reject, or ignore information which does not conform to their world view.

There is an acknowledged weakness to the collected papers, in that many of
them arose in the context of the systemic improvement efforts within a state education
bureaucracy, and as such consciously or unconsciously reflect the values and beliefs
inherent in such an organisation, although there is a more critical edge to those papers
not published by the employer, and the positions advocated in the various papers are
not endorsed by the NSW Department of Education. It is possible to detect, for
example, the inference that the school system acts rationally and the periphery (ie
schools) less so. Given the complex interplay between political control, ideology, and
bureaucracy which impacts on decision making about almost every educational issue,
such an assumption would appear unwarranted. The evidence concerning the ability of
school systems to plan and manage effectively using sophisticated information
systems is also mixed.

Why this is so is not well addressed in the current literature, although it is
possible to find interesting parallels in the literature on the utilisation of social
science research and evaluation (see for example, Weiss, 1974: Patton, 1978). As
Ernest House has argued “... producing data is one thing, getting it used is quite
another.” Many researchers expect that their findings should cause a 180 degree
reversal in the way a program is headed. This rarely happens, because decision-
making just doesn’t work that way. Utilisation of research findings is not something
which happens at one distinct moment in time. Rather, utilisation is a diffuse and
gradual process of reducing decision maker uncertainty within an existing social
context. It provides a background of empirical generalisations and ideas that gradually
creep into policy deliberations.
The managerialist approach to school improvement and accountability has often been characterised as soul-less and mechanistic by its detractors, in comparison to the supposedly warmer and more humanistic alternative, where the accumulated wisdom and connoisseurship of the professional teacher is paramount. The operation of information-driven school management is sometimes described as being akin to flying by auto-pilot, where certain inputs (such as change in wind direction) dictate automatic responses. Such a comparison is as unrealistic, and reflects the same inherent problem contained in the research utilisation phenomenon.

In practice, the experience is somewhat different. In using “value-added” outcome indicators, for example, the data has been used by school improvement officers in NSW simply as a means for initiating discussion and identifying trends and patterns which require further exploration and explanation, rather than to label, praise or criticise school performance.

This “softer” approach to indicator usage, while not retreating at all from the fundamental requirement to identify intended outcomes and devise appropriate means for measuring their achievement, acknowledges the flawed nature of most measurements in education, the limited ability of examinations and tests to represent all that is important in schooling outcomes, and the questionable validity and reliability of some sources of educational outcomes data. In this sense, being rational also involves being reasonable.

The works included in this portfolio document an evolving process, and one which is far from concluded. There is much further work that could be done. An analysis of the macro and micro-politics of school improvement processes, for example, would be a useful contribution to the literature. For the present purposes, there is a practical need to limit the scope and number of the works presented. The works included all address, in some way, the common theme of improving school and systemic effectiveness through better information management. Hopefully, they have made some small contribution to better understanding the phenomenon, and will in turn positively impact on the performance of schools and school systems.
Bibliography


Chapter 1

Trends in the Education Indicators Literature

Trends in the Education Indicators Literature

Introduction

This paper examines some of the trends in the now voluminous literature on education indicators. It is neither an exhaustive review nor a chronology of events in the literature but an attempt to chart some of the major issues that have arisen during the course of the debate. The paper focuses primarily on the literature concerning school-level education. While some of the concepts to be discussed apply across the educational spectrum, issues relating specifically to the higher education sector, such as institutional management and research and development output measurement are generally excluded. The literature concerning higher education indicators is itself large, but has been well covered in syntheses such as Frackman, (1987) and in the supplement and addendum to the Frascati Manual (OECD, 1989).

With the number of publications increasing almost daily, it would be almost impossible to develop a comprehensive catalogue of all current indicator publications, especially those reporting the results of particular indicator schemes, which may not be distributed beyond the system from which they originate. A second problem is limiting the scope of what to include and what to exclude from discussion; the concept of indicators intersects with so many other themes; evaluation, testing and measurement, accountability, school improvement and reform, data analysis and information systems being only a few of these.
However, two relatively complete bibliographies are available which give
coverage to the main conceptual papers, in the English language at least. The first of
these, *Educational Quality Indicators* is an annotated bibliography of 230 citations
published by the Alberta (Canada) Ministry of Education. Now in its second edition,
this volume is organised thematically.

The second source is a bibliography prepared as part of the *Reporting on
Educational Progress* project of the Australian Conference of Directors-General of
Education, last published in 1989. An updated version of the bibliography is appended
to this paper.

This paper is organised as follows; first a brief review of the literature is given to
provide some background to the major forces which have shaped the current situation
and to outline common themes which have emerged. Each of the themes is then examined
in turn, drawing on major review articles to identify trends and developments in theory
and practice. The paper concludes with a brief consideration of the prospects for future
developments.

**Antecedents of education indicators**

The need for an efficient means to summarise the performance of education
systems is not new. Attempts to evaluate the effectiveness of schools extend back
almost to the beginnings of mass education systems. In many cases, these early
evaluations arose in the form of examination of “standards” obtained by pupils and the
associated “payment by results” schemes decried by Matthew Arnold in *The Twice
Revisited Code* as early as 1862. Throughout the late nineteenth and early twentieth
century Minister’s Reports to Parliament, Inspectors’ Reports and reports to district,
county and state boards of education were important vehicles for making public
information about the condition of schooling in many countries.
Madaus, (1981) has observed that whenever there are perceptions of falling levels of achievement, the traditional response has been a call for greater accountability and the imposition of higher standards. Shavelson et al. (1987) note that the concerns raised and solutions proposed during the Common School Movement in the 1880s, and the rationale for the establishment of the U.S. National Center for Education Statistics are similar to those expressed today.

Present ideas about performance indicators and education indicators can be seen as an extension of the research on social indicators that emerged as an active area among social scientists in the mid-1960s. This work has its own origins in attempts to measure social change by William O’Brien and colleagues at the University of Chicago in the 1930s and 1940s. The movement began with high hopes and ended with a retreat from ambition. Detailed accounts of the development of the social indicators movement and its predecessors have been presented by Land (1975), de Neufville (1975), Carley (1981) and Broad (1983) to which the interested reader may refer for a more complete description. Wyatt (1988) and Rockwell (1989) have reviewed the literature on social indicators and identify a number of lessons that can be drawn from the history of that movement.

Interest in social indicators was partly a reaction to the successful use of economic indicators in the early 1960s to guide government policy (de Neufville, 1975). The inability of economic indicators to provide evaluations of wider social welfare considerations, such as qualitative aspects of life, equity and the side effects of economic prosperity, generated a demand for a systematic framework of social accounting to guide government policy.

Further interest in social indicators in the USA came in 1962 from the American Academy of Arts and Sciences’ examination of the second order effects of the space exploration program. The research team turned its attention to the general issue of monitoring socio-economic change, the result being the influential book Social Indicators (Bauer, 1966). This book discussed the development of social indicators, their
relationship to social goals and policy-making and the need for systematic social accounts and improved statistical information.

Other significant papers followed, including a series of studies concerned with the conceptual and methodological problems of monitoring, including *Indicators of Social Change: Concepts and Measurements* (Sheldon and Moore, 1968) and *The Human Meaning of Social Change* (Campbell and Converse, 1972). These two papers proposed different approaches to the subject: the first with objective, socio-structural indicators, the second with subjective indicators of attitudes, expectations, aspirations and values.

Other major influences were *Towards a Social Report* by the U.S. Department of Health, Education and Welfare (1969) which set out the parameters and requirements for the development of a comprehensive social report and the U.K. Central Statistical Office bulletin *Social Trends* (1970 and 1971). In Australia, comprehensive work was undertaken by the Australian Bureau of Statistics, which led in 1976 to the publication of a bulletin titled *Social Indicators*. Japan, Sweden and the Federal Republic of Germany also produced high quality social reports.

The interest of national government agencies was complemented by the work of international organisations. The OECD and UNESCO both fostered the development of social indicators. Of particular significance was the OECD study which dealt with indicators for measuring the impact of education on society (OECD, 1973: 6). UNESCO, on the other hand, tended to view indicators as essential components of evaluation exercises.

Social indicators were applied to many areas in the early 1970s, including education, health, recreation, quality of life and employment. By the late 1970s enthusiasm for social indicators had waned. Their early promise had for the most part not been realised and policy-makers had made very little systematic use of indicators.
Whilst great potential for education indicators was foreseen by the early writers, Jaeger later observed:

Unfortunately, the literature does not abound with examples of the successful application of social or educational indicators (1978:301).

By the late 1970s enthusiasm for social indicators had declined since their promise had, for the most part, not been realised, especially in relation to satisfying the data needs of policy-makers and planners. Some bureaucrats and administrators dismissed social indicators as useless or even misleading (Carley, 1981:20). Many academics turned from a wide concern with social information for government decision-making to a narrower perspective involving complex statistical manipulations. For many, social indicators were an idea whose time had come and gone.

Carley (1981) puts forward three reasons to explain what went wrong. Firstly, expectations of social indicators were naive in relation to the time needed for their development and in relation to the state of development of social theorising. Secondly, anxiety to supply social indicator data to assist in pressing policy decisions led in many cases to the provision of ad hoc information which was often poor in quality, or even wrongly articulated, such that bureaucratic confidence in this kind of information was undermined. Thirdly, there have been insufficient attempts to relate indicators explicitly to policy objectives.

Current thinking about education indicators has also been influenced by the success of economic indicators such as the Gross National Product, the Consumer Price Index, the exchange rate, the inflation rate, and the unemployment rate and so on. These indicators now have a well established and indispensable role in public and private decision making. There are many similarities in the historical development of economic and educational indicators. Both, for example, grew out of necessity or perceived crisis. Measurement of employment, for example became crucial during the great depression of the 1930s, while the need for better measures of student outcomes was sparked by the
fear that falling standards were harming America’s international economic competitiveness.

Originally, the primary purpose of economic indicators was to track the magnitude and direction of changes in specific types of economic activities at the national level. However, their usefulness has expanded beyond this narrow purpose to include predictive and planning capabilities, analytical capabilities and evaluative capabilities in measuring progress towards economic targets and objectives. The expanded capabilities of the indicators was both a consequence and a response to a growing dependence on the measures of national economic performance (Horn and Winter, 1989).

With these antecedents providing a conceptual framework for indicators, a number of separate phenomena occurred which lead to the current regenesis of interest in indicators. The revival of interest in indicators coincides with the scrutiny of U.S. education that begun with the National Commission on Excellence in Education report A Nation at Risk (1983). One of the outcomes of the review was the realisation of how little was known about schools and schooling in that country. Solid information to substantiate propositions about declining levels of achievement were not – and are still not – collected regularly. Further interest was generated in the following year after the publication of the Wall Chart by the then Education Secretary Bell. The purpose of the Wall Chart was to provide comparative data that indicated key features – inputs, processes and outcomes – of each of the 50 states in the U.S.. The generally negative reaction to the Wall Chart in the scientific community sparked a number of influential studies, notably those by the RAND Corporation and the Center for Policy Research in Education at Rutgers University.

The paper by Jeannie Oakes – Education Indicators: A Guide for Policy Makers – was particularly influential. Since the definitions and criteria proposed for indicators in this paper rapidly became widely accepted, much of the effort dissipated by the social indicators movement in its search for purpose was able to be avoided, and more fruitful
directions able to be pursued in the research which followed. So comprehensive was the coverage of the issues in Oakes paper, that in one sense, all of the papers written since then have merely expanded on or implemented the principles expounded (see for example, the collection of articles on various indicator issues edited by Susan Furhman in Phi Delta Kappan, 1988).

Further legitimacy to the concept of indicators as vehicles for accountability and public reporting was added when the National Centre for Education Statistics decided to publish its annual Report *Conditions of Education* in the form of indicators. This has been followed by a further publication in 1988 and 1989 of a volume on *Youth Indicators*.

Also in part as a response to the wall chart, the Council of Chief State School Officers reversed their stand of 20 years or more to create the State Education Assessment Center to produce new, accurate and valid comparative information on the state education systems. The improvement of the National Assessment of Educational Progress has been an important part of that exercise (see Alexander and James, 1987). The Center has, in addition, been at the forefront of research issues related to indicator development and use, including most recently, how to present data that appropriately accounts for different socio-demographic contexts. Meanwhile, many states had begun to develop their own indicator schemes (see for example, Honig, 1985; Fetler, 1986).

At the international level, publication of the results of the IEA international Science study aroused considerable interest, rekindling an examination of many of the issues involved in making cross-national comparisons. The OECD conferences in Washington and Poitiers, which led to the current international indicators project were influential in sustaining and institutionalising this interest. Although the conferences did not lead to agreement on specific indicators or even indicator areas as the organisers had hoped, it did direct attention to important issues emerging at that time, and provoked more in-depth investigation as to how a national picture of the status of education could be constructed. The conferences also highlighted the differences in the extent to which
national systems already collected, processed and used data in the management of their education systems. The OECD indicators project itself has provided a stimulus for conceptual development of indicator methodology and thought (see for example the special issue of the *International Journal of Education Research* edited by Bottani and Delfau, 1990).

In the USA the most tangible and powerful symbol of recent political interest in education was the presidential Summit held late in 1989 at Charlottesville, where the President and all the State Governors met together for only the third time in American constitutional history, and to discuss, for the first time, education (Griffith, 1990). Following on from the Summit was the creation of the Indicators Panel, charged with developing a means of reporting indicators to Congress on an annual basis. The work of the panel has generated a large number of papers on many indicator themes, although these are not yet widely available.

**Trends in the literature**

During the 1980s public education systems in many countries have been faced with contracting budgets, coupled with calls for greater accountability for the use of these scarce resources and the results of their actions. What was needed was some means that could simply, accurately and in a timely manner discharge these demands for accountability. Indicators (by a variety of names) have been proposed as one solution. Ruby *et al.* (1989) in Australia and Cuttance (1989) in Scotland provide comprehensive reviews of the education indicator debate.13

Much of the indicators literature addresses this theme of accountability, why we believe that indicators are suitable for the task, what is meant by indicators, and how they might work.

---

13 See also Nuttall (1989), and MacCollum and Turnbull (1989) for reasonably comprehensive summaries of this literature.
The 1980s have also been a decade of educational and bureaucratic reform in many countries. Commonly, these reforms have lead to decentralisation and devolution of responsibility to the school level. Indicators have been proposed as a means of monitoring these reforms. A second major theme to emerge in the indicators literature thus concerns their use at the school level. This literature emanates most frequently from the United Kingdom, Australia, New Zealand, The Netherlands and France. Within this literature a number of sub-strands can be identified. The first is essentially a reflection of the concerns of central governments to know what is happening in their decentralised system and calls for schools to demonstrate accountability on a number of, usually, centrally determined dimensions. The debate revolves around what the criteria should be, how they should be measured, and how they should be reported.

The second sub-strand concerns how schools might evaluate themselves. The effective schools literature provides one focus for this use of indicators. Another focus emphasises the use of locally determined indicators in the school management process.

A third strand concerns the use of indicators to monitor specific policy objectives in schools. Monitoring of academic achievement has been a strong focus in the United States. Another focus has been on curriculum reform particularly in mathematics and science courses, teacher supply and demand, participation and drop out rates. Indicators which monitor progress towards equity goals has also been an important focus.

Throughout the literature there are also a number of themes which intersect with those described above. These themes include the problems and limitations with indicators and how indicators might be improved through better methodology, better data collection and better analysis and presentation of the information.

---

14 See for example, the CIPPHA report, commentary on the Education Reform Bill in England, and Tomorrows Schools in New Zealand.

15 See Cuttance (1989); Nuttall (1989) and so on.
Each of these themes is examined below, drawing on the major works addressing the issues. The themes were selected as those most commonly recurring throughout the literature. The works drawn from are those which are most readily available and most often cited by others, but are by no means. Further sources are referred to in footnotes or are listed in the bibliography.

The rationale for using indicators

The question of why indicators have achieved such prominence in a very short space of time has not often been explicitly addressed in the literature. The traditional explanation – that it is a combination of demands for information for various purposes by stakeholders in the education process, combined with the development of the technological capability to do so, put forward by Stern (1986) has seldom been analysed and is taken as an article of faith by many who follow.

However, many note that in the U.S. at least, there was increasing frustration with the quality of national data on education. They also note that calls for better data meant not only more data, but data that answered other questions – not how big is the education system but how well is it doing? Kaagan and Coley (1989) note that Governors, legislators and state education agency leaders are no longer operating on the edges of school business, but instead are wrestling with matters of school effectiveness, teacher policy, curriculum emphases and student performance. This higher level of involvement itself has created demands for more, and more specific, information.

Yet the concept of indicators has also attained popularity in countries such as Sweden and the Netherlands, which have had comprehensive education statistical collections for some time. Part of the reason for this interest, identified in a deeper analysis of why so many have been willing to invest in indicator development (Ruby,

---

1989) derives from the technical and conceptual criteria required of indicators as now defined. Not only should indicators provide information which is policy relevant and problem orientated, but should also be able to be understood by broad audiences (with the implication that creative and visual methods of presentation are very important in the reporting of the information) and be able to be delivered in a manner timely enough to influence decision-making. Ruby believes that there seems to be, in the decision-making process, and in the minds of people trying to make policy decisions or resource allocations, a need for a reference point. Decision-makers are looking for ways to help them sift and organise massive amounts of data and information. In these circumstances, the decision makers are often overloaded with data and are looking for reference points which will allow them to close off some sources of influence, or to appeal to some seemingly objective measure to make or defend a judgement.

Four other trends in the kinds of information needs of policy makers can be found with remarkable similarity across OECD countries: information to define the skills of the workforce, in answer to questions about international economic competitiveness; to track and monitor reforms; to support ideological commitments; and to effect qualitative changes in teaching and learning.

Informed decisions require information which is timely, accurate and accessible. As summary devices, performance indicators may have the ability to meet the diffuse needs of governments, policy makers educators and the general public.

Whether they prove to do so in the longer term remains to be seen. Few attempts to evaluate the effectiveness of indicator schemes have yet to reach the literature, except in the form of reports of changes to methodology (for example Fetler describes the evolution of the California accountability program) and the difficulty of moving from theory into practice (Wyatt, 1990). Although many point to the potential problems with indicators (discussed in more detail later in this paper) few challenge the conceptual need for them. Eide (1988) goes closest in reminding us that most decisions are not made on the basis of statistical information. Qualitative information has always played a
major role in decision-making and will continue to do so. Whether this makes indicators invalid for the task can only be evaluated following more comprehensive trials.

What are education indicators?

Defining terms is the traditional starting point in many academic papers. Definitions of what an indicator is vary widely. So do the names by which they are known – performance indicators, education indicators, education performance indicators, quality indicators, workload indicators, management indicators, indicators of success, management indicators – being only a few of the commonly used terms (Ashenden, 1987a; Wholey, 1983). There appears to be little consistency in how they are applied, and few reports provide a rationale for their choice of term. Exceptions are Ruby and Wyatt (1988), who prefer the term education indicator to emphasise that the concern is not only on outcomes but for the whole of the education process. Cuttance (1989), on the other hand argues that since education is multi-level, with input at one level being output from another, performance may therefore be viewed as a process, then the term performance indicator is appropriate.

Definitions are important, as they determine what can be measured and why, but it is often hard to determine whether the differences in definition used by various authors are due to real differences in conceptualisation, or merely to differences in emphasis or terminology. Jaeger (1978), in a review of a dozen definitions found the situation to be anything but clear and consistent. There was much that was contradictory and little that was concise or illuminating. Much the same is still true today, but perhaps for different reasons.

This lack of consensus comes in part from the influence of the social indicator movement, where there was also little agreement as to definition, scope and purpose. This is illustrated by the variance between major approaches, such as those of the OECD and United Nations, and in the syntheses of Land (1977) and Jaeger (1978) both
of whom found only three common elements; that indicators should be quantitative, should measure social conditions, and be time-related or in time-series.

Jaeger 1978) recommended that:

... all variables that (1) represent the aggregate status or change in status of any group of persons, objects, institutions or elements under study, and that (2) are essential to a report of status or change of status of the entities under study or to an understanding of the condition of the entities under study, should be termed indicators. I would not require that reports of status or change in status be in quantitative form, for narrative is often a better aid to comprehension and understanding of phenomena than is a numeric report.

Jaeger’s recommendation that the definition of an indicator be left open and be determined on pragmatic rather than strict definitional terms appears to have gained some following, although most now refer to indicators either explicitly or implicitly as statistics.

Stern and Hall (1988) for example, regard indicators as “... usually derived statistics which may be either single statistics, such as average teacher salary and average student achievement scores on standardised tests, or composite statistics formed by combining two or more related variables.” More complex measures can be created in this way, in the form of ratios or indices. Johnson (1983) goes to some length to distinguish between indicators and variables, which to him are only the building blocks from which true (that is, composite ) indicators are built.

Current reviews (for example, MacCollum and Turnbull, 1990) tend to present definitions based on amalgamations of previous work. Some however refer only to indicators as “pieces of information” (Ashenden, 1987a).

An emerging view, based on the new ideas of what the purpose of an indicator should be regards indicators as statistics that report the “health” of the education system – what students know and can do, the quality of the system’s operations and whether these conditions are improving or declining over time (Smith, 1984). A distinction is drawn between the more common education measures of size or level (eg.
numbers of students, schools or teachers) which, while they may be interesting facts about education, do not directly address the questions of “health”.

More explicitly, others propose that indicators are to aid in policy analysis. This is achieved by developing indicators which are capable of being contrasted with a standard criterion level; contrasted with themselves over time; compared across systems; and contrasted with other indicators in a cost-benefit model (Kaagan and Smith, 1985). Whether this final addition, the need for a reference point is part of the statistical calculation formula for the indicator or merely refers to how the information is presented is not always clear.

The most frequently cited definition in recent times derives from the composite definition of Oakes (1986), who argues that indicators must provide at least one of the following types of information:

- information that describes the education system's performance in achieving desired educational conditions and outcomes. The indicator is thus linked to the goals of the system, and provides a benchmark for measuring progress.
- information about features known through research to be linked with desired outcomes. Such indicators have predictive value because when they change, other changes can be expected to follow.
- information that describes central features of the system (eg. inputs) in order to understand how the system works.
- information that is problem-oriented.
- information that is policy relevant. Indicators should describe educational conditions of particular concern to policy-makers and be amenable to change by policy decisions.

In addition to these substantive criteria, Oakes believes that indicators should have the following technical characteristics:

- Indicators should measure ubiquitous features of schooling, that can be found in some form throughout the system, so that information can be compared across diverse settings.
- Indicators should measure enduring features of the system so that trends over time can be analysed.
Indicators should be readily understood by broad audiences.

Indicators should be feasible in terms of time, cost and expertise required for collection.

Indicators should be generally accepted as valid and reliable statistics.

From these different perspectives Ruby et al. (1989) draw out some of the common attributes of education indicators and identify some of the problems in the common conception of indicators. In seeking to monitor many aspects of human well-being and to assess the impacts of major developments in society, indicators are concerned with measuring end results or ultimate outputs. Several concepts flow from this output orientation.

One is the normative or goal-based character of indicators. Using indicators to measure progress towards a goal often creates controversy, as there is often disagreement about what the goal is and to what extent performance is good or bad, or is an improvement or degradation.

The relevance of an indicator is a combination of its validity and its accuracy. Since educational output tends to be a vague and subjective concept, there will always be some uncertainty about the validity of any indicator that focuses on such a concept. Equally, any education indicator system that omits the qualitative variables will be a partial system that may distort policy-making.

Indicators also need to be representative. An indicator is not simply a measure of itself but acts as a summary variable for some broader concept. Obviously, indicators do not tell everything about education systems. Instead, they provide an “at a glance” profile of current conditions. However, there has always been a major conceptual problem with indicators in that at a broad level they can mask widely divergent situations within subgroups. Indicators should therefore be capable of disaggregation to allow specific examination of sub-populations.
There is little, if any disagreement with this position in the current literature. As a consequence, much of the ultimately unproductive effort spent in search of the ultimate definition of an indicator during the 1970s has been channelled into action. As Nuttall (1989) concludes, the experience of earlier efforts to develop social indicators demonstrated that technical quality is not sufficient to guarantee their use and continuation. Indicator systems must also produce information useful to the policy community if they are to survive as publicly supported endeavours. There thus seems little prospect of any radical change in definition in the near future, nor does it seem that one is necessary.

**Indicator systems and models**

Almost as much space in the literature as that devoted to definitions is given over to explanations of the need for indicator systems, and the closely linked topic of conceptual models of indicators.

With a great deal of consensus being reached about the definition of education indicators as “statistics which reveal something about its health or performance, which describes its core features and that are useful for making policy decisions” (e.g. Smith, 1988) there has been almost universal rejection of the once popular attempts to ascertain for education one general or comprehensive index equivalent to the GNP – or what Fanchette (1974) describes as some sort of Gross Happiness Product\(^\text{17}\). While D’Agostino (1974) sees such a comprehensive indicator as providing more stability than single values and a means of reducing the complexity of the data, most now concur with Odden (1990) that a single indicator or even a large number of indicators by themselves cannot begin to describe the complexities of the schooling process or an educational system.

\(^{17}\) See Gooler (1975); D’Agostino(1974); Ferriss (1969); Drenowski (1970); and Adelman and Morriss (1972).
Nuttall, (1989) reflects the current consensus that:

... given the complexity and the diversity of education systems it is obvious that an individual indicator conveys limited information. It is necessary to build a system of indicators, that is, a coherent set of indicators that together provide a valid representation of the condition of the education system, not just a collection of statistics. Ideally, an indicator system will provide information about how individual components work together to produce an overall effect. In other words, the policy and interpretive value to be gained from a system of indicators is greater than the sum of its parts”.

There appears to be little disagreement with these objectives. The arguments come about what the elements of the system should be. Shavelson (1989) writes that “... past experience with social indicators has demonstrated the need for indicator systems to be firmly grounded in a working theory or model of how the social system being measured actually operates”. The model may be very simple and intuitive or it may be complex but it must represent the phenomena of interest, identify its most important components and the relationships among them. Only with such a model can we have a context for interpreting indicators and for exploring the trends reported by the indicator systems

Carlisle (1970), Oakes (1986), Darling–Hammond (1987), Anderson (1986) and others emphasise the need to construct a model of how the education system works, in order to specify which indicators should be selected. More than any other factor, the model chosen as the basis for indicators will influence what information an indicator system will provide.

Scheerens (1989) identifies three trends in the development of these approaches to conceptualising education indicator systems. These trends mirror to a large extent changes in the emphasis in projected uses for indicators. The first trend in the development of education indicators is the transition from descriptive statistics, largely input and resource measures to those attempting to measure performance or outcomes, or more generally, a shift towards figures with evaluative importance. This trend reflects
the calls for policy makers to know what was happening in their systems and for schools and systems to become more accountable.

The second trend can be characterised by a movement towards more comprehensive indicator systems, with the addition of output outcome measure and context measures, and a growing interest in manipulative–input factors or process characteristics (see Stern, 1986; Teuber, 1987). This interest follows from what Kaagan and Coley (1989) observe to be a higher level of intervention in the top management of education by policy makers in recent years. They want to know not only what is happening, but why, and what they can do about it. Process indicators generally refer to characteristics of educational systems that can be manipulated. Adding process measures therefore enhance the policy relevance of indicator systems.

When we think of process indicators as referring to the procedures or techniques that determine the transition of inputs into outputs, interest in the process indicators leads to an interest in what happens in schools. The third trend in conceptualising indicators is a concern to measure data at more than one aggregation level (Scheerens, 1989). What emerges is that the context – input – process – output model is still the most useful analytic scheme to systematise thinking about indicator systems.

Nuttall (1989) points out that in the development of a system of international education indicators, the choice of a suitable model is particularly difficult since the educational and political values of nations vary, as do their institutions and procedures. Van Herpen (1989) provides a more extended review of issues related to conceptual models of education indicators. Selden (in this volume) also takes up these issues in some depth. The question of which model is most appropriate in which circumstance, and the level of complexity required are likely to continue to attract attention. As the drive to produce indicators itself leads to better and more comprehensive data, part of the problem of defining chains of causality and relationships between the elements that make up the education process may be better understood.
Potential uses for indicators

The question of how indicators could potentially be used also receives substantial coverage in the literature. Conceptualisations of how indicators may be used also follows closely from how they are defined. Shelden (1975), De Neufville (1975) and Macrae (1985) saw the overarching purposes of indicators as characterising the nature of a system through its components, their relationships and their changes over time. While these purposes still hold true today, these early writers saw indicators primarily as providing an information base, with social reporting being the key element in their use. Simply providing information, albeit with an evaluative slant is still important, but with growing acceptance in the mid 1980s of definitions emphasising the role of indicators in gauging the health of the system came a greater emphasis on accountability, both in terms of giving an account of what was being achieved, but also in holding people responsible for performance.

Most of the literature from the U.S. at this time was primarily concerned with measuring student performance at the macro-level, that is at system, state or national level. Kaagan and Smith (1985) are typical in proposing that indicators may help educational agencies to further their reform efforts by:

- monitoring changes in key variables such as the quality of teaching and student performance which would identify impending problems;
- assessing the impact of educational reform efforts;
- encourage better performance by comparisons with other nations and states; and
- focusing attention on areas or institutions which require improvement.

More recent publications (eg Cuttance (1989), Odden (1990) and Ruby et al. (1990) restate the expanded list of uses for indicators given by Oakes (1986), as follows:

- report the status of schooling
- monitor changes over time;
- explain the causes of various conditions and changes
- predict likely changes in the future
profile the strengths and weaknesses of the system;
inform policy makers of the most effective ways to improve the system
inform decision making and management
define educational objectives.

Cuttance (1989) discusses these possible uses at some length. Whilst all of these uses appear to be theoretically possible, Oakes (1986) considers that while some may be achievable, others may ultimately be unrealistic. This theme is also taken up by others, and has helped to focus attention of the practical application of indicators. Shavelson et al., (1987) for example considers that indicators cannot be used to set goals and priorities (although they can inform these objectives, they are only one factor in the decision making process); to evaluate programs (they do not provide the level of rigour or detail necessary) or develop a balance sheet, since social indicators lack the common referent available to economic indicators. What they can do, he believes is to describe and state problems more clearly, signal new problems more quickly, and obtain clues about promising new endeavours.

Two other shifts in focus from the earlier literature can be detected here. Not only is there the suggestion that indicators can do more than simply providing information in a form which can allow comparison (for readers to make up their own minds about what they mean and what they should do about it – a theme particularly strong in the U.K literature) but can and should play a more active part in the process. This belief manifests in three main forms. The first, which is an extension of the earlier accountability orientation, is the overt linking of performance on certain indicators to rewards and sanctions (see for example Honig 1985; Pipho, 1988; and Kaagan and Coley, 1989)

The second, which might be characterised as an orientation towards improvement, is the more subtle, (and some would say more sinister) use for focusing
attention on what is important and worth doing\textsuperscript{18}. Indicator systems in education both influence and are themselves influenced by the social system in which they reside.

The uses outlined above have most often been envisaged most often for system-level use or in what Oakes describes as “top-down” development. In the past two years there has been growing body of literature which addresses the "bottom-up" use of indicators at the school level, with a focus on school improvement through self-evaluation, including better planning and financial management. Anderson, (1987) argues that if we are really interested in understanding relations in education, then there are both statistical and practical reasons why the locus of attention should be as close as possible to where the service is delivered, that is the school or even classroom levels.

Scheerens (1989) discusses how school process indicators might be derived from the findings of the effective schools literature (and how these might then be of value in contributing to national level indicators). Another approach which shows some promise sees indicators forming a checklist against which schools can evaluate themselves (see for example Hargreaves, 1988; Hopkins and Leask, 1989; and Wakefield, 1988), or by outside agencies such as an inspectorate (Cuttance, 1989). A further refinement places greater emphasis on the process of indicator development as much as the indicators themselves, which forces schools and their communities to clarify their goals and reflect on their outcomes (see Marshall, 1988; Ashenden, 1987b; and Wyatt and Ruby, 1989). Caldwell and Spinks (1988) also make extensive use of indicators in the program budgeting cycle, which requires attention to be paid to how inputs contribute to processes and outputs, as a means of creating self managing schools.

One final use for indicators receiving growing popularity in the commercial world, but which has yet to receive much attention in the educational literature concerns

\textsuperscript{18} See Porter (1988); David (1988) for discussion of political influences on indicator selection, David (1988) for discussion of possible threats to local control over the improvement agenda, and the critics of minimum competency testing, eg Wise (1979); Stedman (1985); Madaus (1981); Haney and Madaus (1978), Gipps (1988); Goldstein and Cuttance(1988); Gray, Jesson and Jones (1986) and so on.
their application in staff appraisal schemes. That which does address this issue places it in the context of revivals of “payment by results” or more usually these days “merit pay” in the U.S (see Murmane and Pauly, 1986), although McCormack and Stevens (1990) describes a pilot project to develop indicators for principals of Catholic schools.

In summary, education indicators may be used in four main ways: to meet accountability requirements; to defend or legitimate what has been done; to help in decision making and as a means of focusing effort, normally by setting targets for achievement. The particular emphasis given to these uses is still emerging and varies from country to country, and over time. It seems unlikely that other than refinements to these uses for indicators identified thus far will emerge in the near future. What appears more likely is that as these ideas are more widely implemented, the literature will address more fully and more analytically the causes of success or failure of various schemes, leading to a clearer understanding of what is both heuristically and conceptually possible. Looking to the future, it appears as if meeting accountability requirements may continue to be the dominant theme, and therefore this may need to be balanced by attention to indicators for developmental and planning purposes.

**Indicators in action**

Publications describing the application of the various uses for indicators noted above fall into three main categories: those which are regular publications reporting indicator data, those describing work in progress, and those analysing these efforts. These publications account for a large proportion of the most recent citations.

A number of authors analyse the style, content and outcomes of these reports. Wyatt (1988) provides information about some of the indicator systems in use in a number of countries. For each indicator system reviewed, a description of the major publication reporting the indicator is given (where appropriate) together with a specimen of their presentation. Comments regarding the particular strengths and
weaknesses of the publication are also given. One trend discernible in the reports reviewed is a growing sophistication in presentation. Early efforts tended to consist of pages of tables, often without commentary or with the narrative not contiguous with the tables. In later efforts (eg those from Quebec, PACE and Youth Indicators from NCES) each indicator is typically presented as a double page, with one page given over to graphical representations, with the facing page giving an interpretation of the data, describing the data source and discussing any data issues. These changes reflect not only advances in technological capabilities, but also changes in the conceptualisation of what indicators are and what they are meant to achieve.


Griffiths (1990) describes the changes made to the flagship publications of the U.S. National Center for Education Statistics: The Condition of Education; Youth Indicators and the Wall Chart. Wohlstetter and Baruch analyse in some depth the earlier editions of the Conditions of Education and point out many of its shortcomings, some of which have been addressed in later editions. Many have commented on the substance of the Wall Chart, particularly in relation to the reporting of SAT scores (eg Powell and Steelman, 1984).

California has been a leader in the United States in education indicator initiatives. Odden, (1990) discusses three of the California indicator reports: Conditions of California Education, published by Policy Analysis for California Education (PACE); the Quality Indicators Reports of the state Education Department, and the School Accountability Report Cards required to be produced by Californian schools. While policy imperatives have been the major factor behind many of these indicator projects,
over the years California has created a data infra-structure that allows it to produce indicator systems. It has a comprehensive school–by–school data system that can provide detailed information on students, teachers, school context and curriculum, instruction and student performance. Without these kinds of data, a comprehensive indicator system cannot be produced.

Inman, Pribesh and Salganik (1990)\(^9\) provide the most recent and comprehensive review of state education indicator systems. Their review recognises that the field is a rapidly expanding one, and even in the short time since publication there may have been additional developments. They develop for each state a profile which identifies the legislative base or reform initiative responsible for the indicator scheme, the major publication reporting the indicator, the indicators covered, and any development likely to occur in the future. The following example illustrates the stage of development at present. Each of the Illinios school districts must submit a school report card assessing the performance of its schools and students. The report card is an index of school performance measured against statewide and local standards and provides information to make prior year comparisons and set future year targets. The report card includes the following indicators:

- district and statewide student performance;
- percent of students placed in top and bottom quartiles of nationally normed achievement tests;
- composite and subtest means for college bound students;
- student attendance rates;
- count of chronic truants;
- percent of students not promoted to next grade;
- graduation rates;
- student mobility;
- class size;

---

\(^9\) See also Pipho (1988) for a discussion of several initiatives in various individual state systems in the US.
percent of enrolments in high school maths, science, English and social science;

amount of time devoted each day to maths, science, English and social science;

pupil-teacher ratios;

operating expenditure per pupil;

per capita tuition charge;

district expenditure;

average administrators salary;

average teacher salary.

From their study of 50 states, Kaagan and Coley (1989) conclude that there is premature pressure to use indicator results to hold local school systems accountable, there is insufficient investment to ensure high quality for all measures that become part of the state indicator system; and there is a reluctance to make sense of the indicators, that is, to postulate relationships among inputs and outcomes for the purpose of recommending policy action.

These are problematic issues, because there is a real need to be clear about the purpose of an education indicator system and not to use it for inappropriate purposes (Odden, 1990). Both McDonnell (1989) and Kaagan and Coley argue that using indicators for accountability is inappropriate at present, because few systems provide enough information to do so and because there is still too much to learn about the linkages among the indicators in order to make accurate conclusions about cause and effect for use in accountability contexts. Indicator systems also provide information about the system at levels too far removed from the classroom to be appropriate for improving local problems. These analysts conclude that indicator systems are thus best used to provide broad overviews of the status of the education system, a function which is of greatest interest to policy makers.

Selden, (1990) analyse these developments in state indicator systems. Almost all systems reviewed now go beyond mere reporting of student achievement testing
programs. They reflect in varying degrees professional developments that have occurred recently, but they also reflect the individual histories and circumstances of the states, particularly the balance of state and local control of policy-making. Selden notes that the state education indicator systems take many forms and serve many purposes. Some seem intended to use public interest in quality to press schools and school districts to improve, through the publicity surrounding test score results and other data. Some are designed to reward good schools and districts and to identify low performing ones for potential action. Other systems are intended to provide state-level measures of quality over time, sometimes as evidence of effectiveness of state-level reform, or are intended to provide diagnostic insights into the workings of educational programs.

Selden goes on to identify four trends in the development of these systems. Firstly, he believes that the trend towards centralisation or state initiation of the systems appears to be virtually complete. Even those states reporting no indicator programs do report comparative achievement data on local schools and districts, or are moving towards them. Second, in fewer than ten states are these reports not part of an integrated indicators system. Third, the state indicator systems do not reflect convergence on common approaches or models. There is a great deal of divergence among the states in the number of indicators reported, the kinds of indicators and the way in which they are organised. Finally, the earlier trend towards more extensive indicator systems with more direct policy consequences appears to be levelling off. The most recent projects do not reflect a continuation of the direction towards firm consequences based on local performance monitored through indicator systems.

In contrast to these top-down accountability systems predominant in the United States has been the "bottom-up" developments in Australia, New Zealand, the United Kingdom and other places. The two national conferences on Indicators in Education have provided a forum for discussion of some of these. Wyatt, (1989)20 provides an

---

20 See also Marshall, 1988; Ashenden, 1988; and Brown, 1988 for discussion of state level initiatives of various kinds.
overview of these developments within government school systems. Looking across the different systems of education there is clear evidence that while the form of accountability is very much influenced by the local environment, the Australian and New Zealand schemes are all attempting to link accountability with improvement in the quality of schooling. All are attempting to provide information that will assist decision makers at the school level improve on current performance.

Irving (1990), Osborne (1990) and Hocking and Langford (1990) all provide useful information on the practical arrangements and the issues and factors that have shaped them in three different educational systems. Concerns for equity have also been a major focus. Shrubb (1990) describes a project to develop indicators for gender equity. Millan (1990) outlines a proposal to monitor the outcomes of special purpose programs for disadvantaged schools and for aboriginal students. Keefe (1990) provides a provocative analysis of ideological resistance to evaluation of outcomes in aboriginal education, illustrated with examples from a case study in higher education. He argues that this evidence demonstrates an urgent need for the use of strategies such as institutional profiles and indicators to concentrate efforts on the achievement of equity goals. Stevens (1990) gives details of the approach to developing indicators that combines school-level and system level evaluation of a program to increase participation in schooling.  

All this work is important. It shows that indicators and related techniques can go past the point of theory to application and that this is possible in many fields. In practical terms there is much to be learned from these accounts. They offer insights into how to introduce and institutionalise indicators and other evaluation techniques most effectively and suggest some pitfalls that can be avoided. In addition, they provide a basis for assessing the efficacy of wider or more general applications.

---

21 See also Lawton et al (1989) for a review of the potential role for indicators to monitor student retention and transition in Ontario High Schools.
Equally important is the contribution these experiences can make to the theoretical work on indicators. The papers point to the need for more attention to defining what we mean by indicators. This is also evident in the work from the United States. Some do not offer explicit definitions and use the term to mean pieces of data. The strengths and weaknesses of the various definitions can also be inferred from these accounts of work in progress. At a technical level it is also apparent that practitioners, and others, are moving away from definitions when constructing individual indicators. This gap between definition and measures confuses the emerging indicators methodology with education statistics, management by objectives and general evaluation strategies. It could also reduce the benefits that flow from investment in indicators if the products of the work do not serve the policy and monitoring ends which led to the initial investment.

In summary, these are only a few of the many applications of indicators underway at present. The selection here serves to illustrate the potential uses proposed in the theoretical literature. It appears likely that this area will continue to expand in the near future. In particular, reports of school level improvement programs, which take some time before results become apparent, are likely soon to appear.

**Developing better indicators**

Almost all of the writers who have contributed to the conceptual development of indicators have mentioned the potential problems and limitations they present. These experiences are not confined only to education, but are also found in other areas including the electricity industry (Curran, 1988), the airline industry (David, 1988), other nationalised industries (Woodward, 1986) and other public service areas (Bourne, 1984). The point of appreciating these difficulties is not to yield to them or to argue for the wholesale rejection of indicators, but to see how problems can be overcome or at least diminished in effect.

These problems can be summarised as follows:
indicators provide limited information;
problems with simple models
problems with the collection and analysis of indicator data;
indicators affect subsequent performance;
indicators create political pressures (Ruby and Wyatt, 1988).

To these, Oakes (1986) adds the following issues: (1) the level of information collected; (2) the challenges of making comparisons; (3) the costs and benefits of extensive indicator systems; (4) the political pressures that the existence of an indicator system will bring; and (5) the vital question of who makes the design decisions for any indicator systems.

Various commentators examine these issues in greater depth. Porter (1988), David (1988) and Garbutcheon-Singh (1989) discuss some of the political implications of indicator systems, in terms of possible effects on local control over teaching, standardisation of the curriculum and discouragement of innovation. They point to the risk that staff time and money will be spent on collecting data which will then not be used, and that indicators then become ends in themselves, driving instructional practices in unintended directions, whether or not they are used by policy-makers.

Consideration of conceptual problems receives less coverage, except as has already been outlined in earlier sections (see Eide, 1987) and Shavelson, (1989), although information useful for both operationalising indicator development and shaping conceptual frameworks has been generated from comparison with earlier work on social indicators (Rockwell, 1989) and economic indicators (Murnane, 1987; Horn and Winter, 1989).

---

22 See Wood and Power (1984) for discussion of the rejection of National Testing in Australia, which was partly due to political forces and partly due to the fact that teachers do not use this kind of information in their practices, and discussion of the experiences of the APU in England and the NAEP in the USA.
Many of the same problems facing attempts to improve education indicators needed also to be confronted in attempting to improve economic indicators. Understanding the nature of these common problems may help us appreciate the progress that has been made in improving education indicators.

Murnane (1987) concludes that in both cases, it is extremely difficult to develop operational measures of performance that are closely related to the underlying outcomes of concern. In both cases, there are trade-offs to be made in determining appropriate levels of disaggregation to use in reporting performance measures. In both cases it is difficult to determine the subgroup breakdowns most likely to be helpful in explaining performance trends. The comparison between economics and education illustrates that many of the problems in developing social indicators are generic, and are not unique to education. Recognition of this general principle may encourage analysts working to improve education indicators to learn from the experience of indicator development in other sectors.

Murnane and Pauly (1988) draw three lessons from this experience. Firstly, it is important to develop multiple measures. No single set of test scores provides a reliable basis for assessing the performance of the education system, just as no single measure such as the employment rate alone provides a reliable measure of the economy. Secondly, indicators make us aware of new and puzzling questions. They do not provide all the answers nor should we expect them to.

Thirdly it is important that we become knowledgeable users of indicators. In the economic sphere, a rich set of indicators has been available for over 40 years. Users of these indicators have become quite sophisticated in piecing together a complete picture of economic performance by examining trends in more than one indicator. The development of a comparable set of education indicators is still evolving in many important ways. As a result, most users are still in the early stages of learning how to use them, nor do we yet know how to interpret many of the patterns appearing in the current data.
Horn and Winter (1989) also draw a number of lessons from their analysis of the evolution of economic indicators. They point out that the most successful economic indicators in use today have a well defined purpose in the policy formulation process and are linked to specific policies, programs or legislation. When payment of millions of dollars (eg for welfare programs) is tied to a particular indicator (such as the CPI) there is substantial pressure to ensure its validity. Modern economic indicators have developed interactively over a substantial time period. It is unlikely that a new indicator could come into practice in less than a decade.

Successful indicators are supported by sound methodological underpinnings. Economic indicators have gained the support of the research community in part because they have come to be based on sound methodological procedures and a sophisticated survey techniques. The validity and reliability of the economic data are obtained from many different sources. As a result, the system is not explicitly subject to the influence of individual agencies or data reporting entities. Reliability is also a function of the fact that the individual data reporting agents have no reason to misinterpret information.

Substantial financial resources have been and continue to be invested in developing economic indicators. Start up costs for a new indicator are substantial. Consequently, it may require considerable political support to garner the resources necessary for the development and maintenance of a new indicator.

The experiences of the social indicator movement suggest a number of lessons for those involved in developing performance indicators for education. Firstly, definitions of what indicators are and how they are composed must be simple, clear, and consistent. If measures are to be aggregated, then the definitions of terms governing those measures must also be consistently applied.

Secondly, as Carley suggests, the temptation towards an ever-increasing aggregation that results in a single index of educational output should be resisted. It is important to develop multiple indicators. No single set of test scores provides a reliable
basis for assessing the performance of the education system, just as the employment rate by itself does not provide a reliable measure of the performance of the economy.

Avoiding overstatement of both short-term and long-term benefits is a third point for consideration. Indicators can have an important part to play in both the planning process and accountability schemes, but they are by no means the sole answer to an organisation’s problems. Indicators at best can only point to the existence of problems or successes, but of themselves cannot suggest solutions for problems. The technical limitations to the construction and interpretation of indicators also needs to be recognised.

Equal attention needs to be given to work on interpretation as well as construction of indicators. As Murnane and Pauly (1988) conclude, it is important that we become knowledgeable users of indicators.

Finally, it is most important that education indicators maintain policy relevance. If indicators do not influence policy-making it will be difficult to justify the resources expended on them and the costs in lost opportunities they will exact. Indicators are more likely to be successful if they are designed to reflect specific conditions and concerns which have been the focus of particular policy initiatives. Caution needs to be exercised here, of course. Indicators which are specific to current concerns may be of little use over time.

The literature on technical problems can be grouped into two main sub-categories; those addressing the issue of how to make fair comparisons, and those concerned with the adequacy of available measures. The literature on each of these areas is quite substantial, but while all have some relevance to the question of how to develop better indicators and indicator systems, not all directly address the concept of indicators themselves.

methods of making judgements about the relative effectiveness of schools. A trend observable in these papers is a move away from earlier "standards" models to methods which attempt to measure the "value added" by schooling. Discussion focuses around how changes over time can be measured, whether and how differences in student performance should be adjusted to factor out influences not due to the effects of schooling, and how to deal with the multi-level nature of the education process.

The sophisticated statistical techniques typically involved in these methods lose much of their power if the basic data they are applied to is less than adequate. A long-standing criticism of many of the early (and existing) national and state accountability schemes mentioned earlier has been that indicators of student outcomes have been limited in the areas they cover and the skills they measure, and that are lacking in diagnostic value (see Alexander and James, 1987). We have fairly good paper and pencil tests of the most commonly taught basic knowledge and skills, but we lack adequate measures of children’s ability to think critically, to apply knowledge and to solve problems (Oakes, 1986). These shortcomings have, in part, led to attempts to make greater use of teacher-assessed tasks in the National Assessment scheme in the U.K. (see Black, 1988).

It is not only student assessment that lacks adequate measures. Oakes notes that while most agree that the quality of the teaching force should be a key indicator, teacher quality can at present only be approached through indirect measures such as qualifications and years of experience, factors which correlate only weakly with student outcomes. It remains to be seen whether this reliance on "proxy" measures is an obstacle able to be overcome or an intrinsic weakness in the indicator model.

A further issue of concern to those attempting to develop state, national and internationally comparable indicators is the lack of consistency in the definitions, and thus common reference points of even the most basic concepts, such as "what is a school" and "who is a drop-out". This has been a recurrent theme in each of the OECD sponsored conferences on indicators, (Suter and Sherman, 1989; and Ruby, 1990) and
one which shows little sign of resolution without a great deal of compromise (Burstein et al., 1989).

The analyses of the weaknesses of indicator systems seldom directly challenge their conceptual validity. Rather, they tend to serve as a warning of what can realistically be expected from them, how they can be methodologically improved, and by implication, how they should be presented. None of these problems are insurmountable, and reports of successful solutions may soon be forthcoming. Yet even these are not likely to lead to any ultimate answer, as the issue most frequently canvassed are in a sense only extrapolations from ongoing debates in more traditional areas of inquiry. This debate has proved to be fertile in the past, and is likely to continue.

A deeper problem, not often addressed, concerns the proper role of indicators in the organisation of the delivery of schooling implicit in the conceptualisation of indicator use presented in this paper. The managerial model dominant today is a relative newcomer to public administration, and is yet to prove more successful than the previous bureaucratic and professional models. Fasano (1994) explores this issue further.

**Conclusion**

This paper has attempted to trace the major trends emerging in the literature on performance indicators in education. It has shown that the concept is both an old and a new one. In its most recent phase it has proved to be both dynamic and productive. Yet we should legitimately ask what, if anything, is new, what has changed since the concept re-emerged with much fanfare. Even though the literature is prolific, it is not surprising that certain key figures and developments have played such an important part in furthering our understanding of what indicators are all about, and thus recur frequently both in the literature and in this review.
Further work on definitions and the continued analysis of attempts to apply indicators will clearly be important areas of investigation over the next twelve months. In addition to looking at implementation issues an analysis of work in progress should address the pressing question of whether and how the outcomes of the emerging work at the school level can be aggregated and selectively assembled into an indicator system capable of representing system-level interests. This type of analysis would underpin the larger scale developmental work on models of education required for a comprehensive indicators system.

Related to this work is the need for a critical examination of the models of accountability implicit in definitions of indicators and in the shape and structure of indicator systems. Work of this kind is needed to clarify the assumptions and define more rigorously the relationships between accountability, quality and indicators.

Greater attention to the data collection and selection issues that need to be resolved before indicator systems can be fully implemented is also necessary. Work on such issues as standardisation and compatibility of elements within education data systems need to be undertaken. There are also major data management questions that need to be resolved.

Combined, these areas present substantial challenges to those working on education indicators and accountability. Without this research and analysis the considerable benefits promised by indicators will not be realised. In the mean time, indicators seem set to continue to be a hot topic in the education literature.
Bibliography


Chapter 2

The Role of Indicators in School-Level Evaluation and Development

The Role of Indicators in School-Level Evaluation and Development

This paper presents a brief overview of the role of performance indicators in school-level evaluation. It describes several approaches to developing school-level indicators, including one based on school effectiveness criteria and a more generic strategic planning model. The paper argues that indicators can be useful tools if they are construed as part of a more comprehensive system of education management. Further, they are likely to be particularly useful in a devolved system in which school principals have increased management responsibilities and accountabilities.

Examples from two different approaches are illustrated: the first, a “top down” approach in which schools are mandated to publish certain types of indicators; the second, “bottom up” approach in which schools determine appropriate indicators according to their own goals and objectives. The paper concludes with a brief analysis of the lessons learned from these experiences and some suggestions as to how systems can help create an environment in which indicators can achieve their promise.

Introduction

The concept of performance indicators has been a hot topic in the education arena in the last decade. In that time, both the volume of literature surrounding them and examples of their implementation have grown enormously (see Wyatt, 1991 for a comprehensive review). But while much has been written about indicators at the international level (see for example Bottani, 1990; Ruby, 1991; and OECD, 1991) and at the national or system level (see for example, Griffith, 1990; Cuttance, 1990; Alkin, 1989), relatively few sources address the issue of how indicators can be used at the school level in any practical detail.
In many OECD countries, increasing demands for reform in education and enhanced accountability have dominated the domestic political agendas for several years. These demands for greater accountability for the use of public funds, debate about national standards of student achievement and structural reform have in turn all generated demands for more and improved information. Teachers, schools and Ministries of Education and other educational organisations have long been fond of collecting vast amounts of data on pupil enrolments, number of teachers, and educational budgets and so on. However, much of this existing data was about inputs, reflecting the priorities of earlier times, when improved outcomes seemed possible by making substantial investments in resources. More recently, there has been a shift in demand towards information about outcomes. This shift from quantity to quality carries with it a trend for more summary information, focussed on the key elements of decisions, and requires a change in the way in which that information is presented. As decision-making processes in education become more diffuse, as more and more responsibility is decentralised to schools, the need for information to simplify decisions and to justify decisions also increases.

There has also been greater emphasis on information that can be used at the site level to shape, make and monitor qualitative improvements. This increased emphasis is associated with arguments about teacher professionalism, school improvement programs, the governance of schools and the need to make qualitative improvements by redeploying resources at the school level.

This paper is structured around the theme of this conference, (Theory and Practice in School Based Evaluation), and so will present a brief overview of some of the theory and some examples of practice in the use of indicators as evaluative tools at the school level. In doing so it will draw on experience with indicators in Australia and the United States.
Theoretical underpinnings of indicators at the school level

The current thinking about indicators in education owes much to the earlier social indicators movement of the 1960s and 1970s (see Rockwell, 1989). It also has some similarities to the much earlier ideas, such as Frederick Taylor's theory of scientific management and business models of management by objectives, in that they are all rooted in the basic belief that better decisions (about educational issues) will be enhanced when those decisions are based on sound information or hard evidence. Improvement is expected because information enhances the capacity to control, monitor and evaluate, and as a consequence, make better decisions and produce better outcomes. How this evidence is derived and used for indicators is somewhat different from the earlier models.

Real interest in applying this "rationalist" approach to education reform in the modern era dates from the 1984 publication of the Wall Chart by the U.S. Department of Education Secretary, Terrell Bell. For the first time, comparative data on key features of educational inputs, processes and outcomes for each of the 50 states in the U.S. was easily accessible. The reaction to the Wall Chart sparked a number of studies, notably those by the RAND Corporation and the Center for Policy Research in Education at Rutgers University.

Particularly influential was the paper by Jeannie Oakes (1986) Education Indicators: a guide for policy-makers. The uses, definitions, models, standards and criteria for indicators proposed in this paper have become widely accepted and underpin much of the conceptual work conducted since then. These definitions are shown in Exhibit 1.

While it may seem somewhat pedantic and trivial to dwell on definitions, our experiences have shown that we need to continually attend to what we mean by indicators. Looking across a number of reports of indicator projects, it is clear that many do not offer explicit definitions and use the term indicator to simply refer to almost any
piece of data. At a technical level, it is apparent that practitioners move away from definitions when constructing individual indicators. This gap between definition and measures confuses the emerging indicators methodology with education statistics and general evaluation strategies. It can also reduce the benefits that flow from investment in indicators if the products of the work do not serve the policy and monitoring ends that led to the initial investment (Wyatt and Ruby, 1990). In our work, we have stressed that policy usefulness and links with key elements in education as the distinguishing features of indicators.

### Exhibit 1

**WHAT IS AN INDICATOR?**

- Information that describes the education system’s performance in achieving educational conditions and outcomes; the indicator is thus linked to the goals of the system, and provides a benchmark for measuring progress;
- Information about features known through research to be linked with desired outcomes; such indicators have predictive value because when they change, other changes can be expected to follow;
- Information that describes central features of the system, in order to understand how the system works;
- Information that is problem-oriented;
- Information that is policy relevant: indicators should describe educational conditions of particular concern to policy makers and be amenable to change by policy decisions.
- Indicators should measure enduring features of the system that can be compared across time;
- Indicators should be readily understood by broad audiences;
- Indicators should be feasible in terms of time, cost and expertise required for collection;
- Indicators should be generally accepted as valid and reliable statistics.

Another important theoretical consideration, and one which occupies almost as much space in the literature as that devoted to definitions, concerns the need for conceptual models of indicators (eg. Van Herpen, 1989). The current consensus is that because individual indicators convey only limited information, it is necessary to put together a *system* of indicators, that is, a coherent set of indicators that together provide a valid representation of the condition of the organisation, and not just a collection of statistics. Ideally, an indicator system will provide information about how individual
components work together to produce an overall effect. There seems to be little disagreement on this point. The arguments concern what the elements of the system should be, and what organisational framework is most appropriate. Murnane (1987) for example, believes that education indicators should focus exclusively on school outputs. Oakes (1989) disagrees, arguing that information on school context should also be provided. Resolution of this argument will depend largely on what kinds of information are found to be most useful.

The dominant paradigm in management theory at present holds that decisions are best made by those closest to where those decisions are to be implemented. Anderson (1986), for example, argues that if we are really interested in understanding relations in education, then there are both statistical and practical reasons why the locus of attention should be as close as possible to where the service is delivered, that is, at the school or classroom level.

The dilemma we face is how to balance the information needs of individual schools with those of the larger systems to which most public schools belong. One way of meeting increasing demands for accountability may be to set up better centralised data collection systems. In such a scheme, schools would need to send in agreed statistical data to a centralised unit responsible for some form of public reporting. If such an approach was adopted, there would be some benefits to the system. The existence of a centralised data base could have a useful role to play in increasing public confidence in standards. It may, for example, provide some evidence for use by system administrators to counter anecdotally-based criticisms of falling standards. It may be a useful adjunct to Governments in the objective assessment of some of the system’s strengths and weaknesses. However, such a database would have little appeal to school staff, would put substantial extra strain on their time and would be unlikely, per se, to improve educational standards. Experience in other school improvement efforts suggest that the most powerful educational effects are wrought by information that is localised, relevant,
and owned by the school itself and which causes change in teacher behaviour (Brown, 1988).

A more powerful way of meeting these competing demands is to regard use of performance indicators as an integral part of a broader, improvement oriented management strategy. Good management in education involves components such as establishing management responsibility, determining clear goals and standards, systematically using information on the achievement of goals and standards, and planning more effective resource management. Judicious use of performance indicators is a vital component of such a management strategy. This approach does not preclude the existence of a centralised database, but it takes a broader and more sophisticated view of systems and how they change.

The evolving model of a professional school (Eraut, 1991) is characterised by a series of cyclical development and evaluation processes, in which needs assessment (itself an evaluative process) gives direction to development which is later modified by formative evaluation. At the school level, there are the processes of formulating and implementing a school development plan, updating it, monitoring all areas of school activity, and conducting occasional, searching reviews. At the teacher level there are both individual and collegial aspects of staff development, teacher appraisal and the setting of personal targets. At the pupil level, the most flexible modes of teaching and learning are used, involving agreement on objectives, selection of learning activities, providing of feedback and recording achievement.

In keeping with these proposals, our model for using indicators for school improvement, (described in Ruby and Wyatt, 1988; Ruby et al., 1989) therefore emphasises better use of information as the key to better management. Use of performance indicators can be part of the response schools make to school development, providing a means by which the school can demonstrate to itself, and to the system, its achievement of goals and standards. Having management responsibility, the school
would nevertheless, be free to decide how it might go about achieving these goals in a way that suits its own circumstances.

While indicators can be used in many ways, we see them as being most valuable when they are part of a planning cycle, where their function is to monitor, report and communicate the progress of the plan. This cycle provides an overarching approach to planning which serves many purposes. It has a generic nature and wide applicability, and can incorporate other activities, such as program budgeting and collaborative needs assessment which may already exist within the school.

The cycle has five phases – with no clear beginning or end – but with specific activities and products at each stage. It derives from the decision-making model developed in the early 1970s (Stufflebeam, 1972), known by the acronym CIPP (Context, Input, Process and Product). To these four elements there is the added element of reporting.

This approach using indicators for monitoring has, we believe, several advantages over traditional evaluation methods. First, while evaluations can be formative or summative, in practice the majority tend to be of the later kind, thus having little impact on program implementation. Second, formal program evaluations tend to be complex in methodology and lengthy in time to prepare and these in turn are lengthy to conduct and prepare. Often, action proceeds in spite of, or before the evaluation report has been finalised. Third because they contain such a wealth of information, evaluation reports also tend to be of greater interest to technically informed audiences. The principal virtues that indicators have to offer are (1) because they are established a priori, they deliver information in a timely fashion, (2) their simplicity (and emphasis on graphic presentation) allows them to communicate to broad audiences, and (3) this simplicity also provides focus to the key areas requiring decisions.
At the same time, it is also important to keep in mind the limitations of indicators. They cannot tell all there is to know about a program, they do not provide explanations for observed phenomena, and what they show requires interpretation. They do not, therefore, replace other forms of evaluation, but supplement them. They function best by providing “early warning signals” of areas where other, more searching questions should be asked.

The ability of this process to result in any continued improvement is based on one key assumption: that school communities, teachers and decision-makers are responsive to information, and that they do have the capacity and/or authority to respond. Until recently, few of the management practices which have evolved in education over the last 100 years reflect this orientation. The process sketched above appears at first glance to be a simple one, but it also raises important questions about the nature of accountability and wider policy questions, such as who controls information about the functioning of schools and where responsibility for poor performance should be located (Ruby, 1991) and the professionalism of teachers and school management (Eraut, 1991).

There are several other approaches to the self-directed, goal oriented model for deriving indicators outlined above. Porter (1991) argues strongly that since schools do
not directly produce student learning, it is important to know about the nature of educational opportunity as a direct policy output of schools. A further motivation for indicators of school processes is to provide explanatory information when student output goals are not reached. School process indicators may point to possible causes and thus to possible solutions for inadequacies in school outputs. The primary domains of school processes that Porter sees as worthy of developing indicators are organisational characteristics and instructional characteristics, with the later being divided into curriculum quality and teaching quality.

Braithwaite and Low (1988), and Scheerens (1989) discuss how school process indicators might be derived from the findings of the effective schools literature, as a means of identifying promising process variables. He looked at recent studies that involved characteristics measured not only at the school level but also aspects of instruction at the classroom level. Some potential indicators that emerged are shown in Exhibit 2.
Exhibit 2

INDICATORS BASED ON EFFECTIVE SCHOOLS MODEL

Educational leadership

- the amount of time head teachers spend on educational matters, as opposed to administrative and other tasks;
- whether head teachers do or do not discuss test results on pupils' progress with teachers;
- the amount of instructional issues on the agenda of staff meetings;

Achievement oriented policy

- the amount of overt statements in official school documents that express an achievement oriented emphasis in school policy;

Orderly and safe climate

- statistics on absenteeism, lesson drop out, suspensions and expulsions, and delinquency as instances of the degree of order in the schools;
- ratings of school discipline by teachers, parents and students;

Clear objectives

- whether or not explicit school curricula, stating educational objectives and levels of achievement are available;

High expectations

- estimates by teachers and/or head teacher of the percentage of students that will complete their secondary schooling;
- student estimates of their further educational career;

Monitoring/evaluation of pupils' progress

- the frequency and use of curriculum specific tests at each grade level;
- the frequency and use of standardised achievement tests;
- whether or not the school uses a (computerised) system to monitor pupils' progress
It is not hard to imagine how it would be possible to collect information in these areas. While some would involve intensive data collection and what Scheerens terms "high inference" measures, some could be collected using relatively uncomplicated scales or questionnaires. Other characteristics could be assessed by unobtrusive observations by researchers, inspectors or other observers, for example, the presence or absence of graffiti in and around the school and the noise level in school corridors (if these were factors thought to be related to school effectiveness). At the classroom level, detailed observation and reporting by teachers and pupils may be required. Some creativity is necessary to find simple, low inference proxy measures for those effective school characteristics that are not amenable to direct measurement (Davies, 1990).

Another approach that shows promise sees indicators forming a checklist against which schools can evaluate themselves (see for example Hopkins and Leask 1989; and Wakefield, 1988), or be evaluated by an outside agency such as an inspectorate (Hargreaves, 1988; Cuttance, 1989). This approach is most common in the United Kingdom. Not all of these indicators result in easily quantified and unambiguous statistics, such as percent of students absent. Some require more subjective judgements, set against some sort of rating scale which gives at least some benchmark against which change can be assessed. An example of the latter might be "the extent to which recent examples of pupils' work are displayed throughout the school". Here, to attempt to quantify this objective by counting the number of displays or calculating the percent of wall space devoted to pupils' work would be a spurious attempt at precision. It is far more preferable to construct a scale (of whatever length) anchored against criteria that make sense for each circumstance.

**Indicators in practice**

There are now many reports which document examples of indicators in use at the school level (summarised in Wyatt, 1988 and 1990b). These examples tend to take one of two forms: those in which a set of indicators is mandated by the school system,
primarily for monitoring and accountability, and those in which the indicators are defined and developed by the schools themselves. The former tend to be the most common, and while perhaps of less interest for this conference, are a legitimate and important use of indicators that deserves some discussion.

California has been a leader in the United States in education indicator initiatives. Odden, (1990) discusses three of the California indicator reports: Conditions of California Education, published by Policy Analysis for California Education (PACE); the Quality Indicator Reports of the California State Education Department, and the School Accountability Report Cards required of Californian schools. While policy imperatives have been the major factor behind many of these indicator projects, over the years California has created a data infrastructure that allows it to produce such indicators. This database has a comprehensive school–by–school capability that can provide detailed information on students, teachers, school context and curriculum, instruction and student performance.

Inman, Pribesh and Salganik (1990) provide a comprehensive review of state education indicator systems. Their review recognises that the field is constantly and rapidly changing. They develop for each state a profile which identifies the legislative base or reform initiative responsible for the indicator scheme, the major publication reporting the indicator, the indicator covered, and any developments likely to occur in the near future. The following example from the state of Illinios illustrates the kind of indicators required. Each of the Illinios school districts must submit a school report card assessing the performance of its schools and students. The report card is an index of school performance measured against state and local standards, and provides information allowing comparisons with prior years and for setting future targets. The report card includes the following indicators:

Selden, (1990) analyses these developments in state education systems. Almost all systems reviewed go beyond mere reporting of student achievement testing programs. They reflect in varying degrees professional developments that have occurred
recently, but they also reflect the individual histories and circumstances of the states, particularly the balance of state and local control of policy making. Selden notes that the state education indicator systems take many forms and serve many purposes. Some seem intended to public interest in quality to press schools and school districts to improve, through the publicity surrounding test score results and other data. Some are designed to reward “good” schools and districts and to identify low-performing ones for potential action. Other systems are intended to provide state-level measures of quality over time, sometimes as evidence of effectiveness of state-level reform, or are intended to provide diagnostic insights into the workings of educational programs.

Exhibit 3

ELEMENTS IN THE ILLINIOS SCHOOL DISTRICT REPORT CARD

- district and statewide student performance
- percent of students placed in top and bottom quartiles of nationally normed achievement tests
- composite and sub-test means for college bound students
- student attendance rates
- count of chronic truants
- percent of pupils not promoted to next grade
- graduation rates
- student mobility
- class size
- percent of enrolments in high school maths, science, English and social science
- amount of time devoted each day to maths, science, English and social science
- pupil-teacher ratios
- operating expenditure per pupil.
- per capita tuition charge
- district expenditure
- administrator’s average salary
- teacher’s average salary
Selden goes on to identify four trends in the development of these indicator systems. Firstly he believes that the trend towards centralisation or state initiation of the systems seems to be complete. Even those states reporting no indicator programs do report comparative achievement data on local schools and districts, or are moving towards them. Second, these reports are part of an integrated indicators system in all but ten states. Third, the indicator systems show little convergence in the number of indicators reported, the kinds of indicators and the way in which they are organised. Finally, the trend towards more extensive indicator systems with more direct policy consequences appears to be levelling off.

Kaagan and Coley (1989) also studied the 50 state indicator systems, and concluded that there is premature pressure to use indicator results to hold local schools systems accountable; that there is insufficient investment to ensure high quality measures that become part of the state indicator system; and there is a reluctance to make sense of the indicators. These are problematic issues, because there is a real need to be clear about the purpose of an indicator system and not use it for inappropriate purposes (Odden, 1991). Both McDonnell (1989) and Kaagan and Coley argue that using indicators for accountability is inappropriate at present, because few systems provide enough information to do so and because there is still too much to learn about cause and effect for use in accountability contexts. These analysts conclude that indicator systems are thus best used to provide broad overviews of the status of the education system, a function which is of greatest interest to policy makers. Whether these views overstate the case for caution, given later evidence, is debatable.

In contrast to these top–down accountability systems, is the bottom–up developments in Australia, New Zealand, the United Kingdom and other places. While the form of accountability is very much influenced by the local environment, the Australian and New Zealand schemes all attempt to link accountability with improvement in the quality of schooling. All are attempting to provide information that will assist decision–makers at the school level to improve on current performance. Some
of these involved pilot or demonstration projects, and some were, or have become, fully fledged, mainstream and system wide initiatives\textsuperscript{23}. These efforts also varied widely in their focus and methods, from those which address system level needs to those affecting only a single school. Two examples of different projects are summarised in the exhibits which follow. Other examples can be found in papers by Shrub, Cuttance, Millan, Keefe, Dean, Rose \textit{et al.}, Hocking and Langford and others presented at the \textit{National Conference on Indicators for Quality, Accountability and Better Practice} (Wyatt and Ruby, 1990; Wyatt, 1990b).

\textbf{Lessons learned from this experience}

While these programs have not been formally evaluated, the experiences illustrate the kinds of difficulties in moving from theory into practice that are likely to be found by others attempting to use indicators at the school level.

This work shows that indicators can be applied in many fields. In practical terms there is much to be learned from these accounts. They offer insights into how to introduce and institutionalise indicators and other evaluation techniques most effectively and suggest some pitfalls that can be avoided. In addition, they provide a basis for assessing the efficacy of wider or more general applications.

The first and most powerful lesson that we have learned, is that when you begin to encourage schools to use indicators, you are asking them to operate in a fundamentally different way, one in which the organisational culture changes from one concerned with the administration of policies and curriculum to the management of people and resources. In many schools, this was a radically new way of thinking and acting. Done properly, these initiatives were no small undertaking, but which hopefully had lasting and positive effects.

\textsuperscript{23} Indeed, one of the criteria for success of these projects has been whether they do enter into mainstream practice.
In essence, what schools were being asked to do in each of these projects was to state clearly what they were trying to achieve (i.e. to identify goals in fairly concrete terms), to identify the means by which they would know whether they are achieving these goals, gather evidence about the extent of their achievements, then use this information to inform and improve practices. This process is, of course, nothing new. The only new thing is the use of indicators as the means for focusing, organising, and presenting the information.

It came as somewhat as a surprise that after more than a decade of exposure to these similar evaluation and curriculum planning models, that a major barrier to be overcome was that many schools and programs did not have clearly stated purposes, and even more infrequently had they identified goals or objectives for achievement. All schools are, of course, officially required to have documented policy statements setting out their goals, but more often than not these are couched in vague and fuzzy terms, such as “to contribute to the all round development of each child’s potential”. Let me take some more specific examples. A goal common to many schools is “to raise students’ self-esteem”. This is a worthy goal, one no one would argue with, but when questioned, teachers were hard pressed to state how they would know when this goal was being reached. Rarely had they thought about, much less written down what they mean by self esteem or what behaviours they would expect from students with high or low self-esteem. The same sort of problem was also evident in many other areas of educational endeavour.

A further problem was that even when schools had written goals and objectives, they were seldom linked to any specific strategies for achieving them. Even more rare in Australian schools at that time, was any attempt to link strategies or goals with costs.\(^4\)

\(^4\) In recent times this has become more common place as systems have introduced forms of School-Based Decision Making or Local Management of Schools.
This is not entirely the fault of teachers – they had rarely had the opportunity to control what they do.

Also, while our experience has been that almost all teachers and school administrators acknowledge that evaluation is important, they had seldom considered evaluation methods, or targets or criteria for success at the time of setting their goals. Too often, evaluations tended to be anecdotal descriptions, left until the end of the program, and seen as an extra, additional imposition rather than an essential and integral part of the process.

The second important lesson to emerge from this is that the process doesn’t just happen. It requires a mechanism and a framework within which to operate. The most successful mechanism, but by no means the only one, was when indicators were embedded in a formal planning process leading to the production of a School Improvement Plan (by whatever name), through a collaborative process involving a school-based committee. This planning process also serves to reinforce that indicators are a means, rather than an end in themselves.

Third, an unintended benefit we found was that the process of trying to establish indicators provides a focus for thinking about the mission of the organisation and is an opportunity to bring staff together to reflect on issues affecting the school as a whole. Secondary teachers, in particular, are accustomed to thinking in terms of subject departments rather than from the larger perspective. In trying to document the outcomes of these discussions a critical issue often raised by teachers was how detailed do the plans need to be. They asked “aren’t these just behavioural objectives again?” This is a legitimate question. Indicators have been criticised because they provide a simplistic view of a complex reality. To adequately describe the complex, diffuse and long term objectives of education, a large number of indicators are required. At the same time, each individual indicator, taken out of context, can appear mechanistic and trivial. Another lament is that if all of the programs operating in the school are to have indicators
developed for them, you could have hundreds of indicators, and therefore little time for anything else besides collecting, analysing and reporting indicators.

Obviously, common sense must prevail to provide the necessary balance between parsimony and completeness. Our answer to these criticisms has been to stress that indicator selection is an evolutionary process. The original list of indicators will need to be amended as some are found to be useful and others are not. As priorities change, so will the indicators used. The system must above all be flexible. We have also strongly supported the notion of sets of indicators to counter arguments about their reductionist potential, while at the same time keeping in view the fundamental difference between indicators and other classes of statistics.

The reluctance to evaluate programs and practices rigorously is the result of many factors, including the prevalence in the 1970s of \textit{laissez-faire} and child-centred theories of education, in which assessment and testing were reviled, and a lack of time in which teachers could properly plan, implement and reflect on what was required. The pace of change in Australian schools has been very rapid for many years now, and teachers have little enthusiasm for more change without a clear perception that it will be beneficial to themselves and the children they teach.

Unfortunately, there will always be pressure for innovations to be “up and running” and to produce results quickly, but this particular initiative requires considerable time for staff to come to terms with indicator concepts. Our approach has been to hasten slowly, in the belief that unless the important political issues are addressed first, we will not be able to build the commitment needed to sustain long-term development.

The indicator schemes that have shown most promise have accordingly been those which have concentrated a great deal of effort on staff development and have been able to provide continuing support. While initial interest has often been high, because there is a long lead time between planning and results, this interest often flags along the
way. However, despite reservations about the extra work involved, the response of most involved, once begun, has been supportive and enthusiastic.

Creating an environment in which indicators can work

Statistics, data or information, on their own, no matter how presented, are unlikely to lead to significant improvements in schooling or its outcomes for students. Unless an environment is created that allows teachers to use this information as reflective practitioners, any change is likely to be minimal. Teachers are cynical that indicators may be just the latest fad or somebody’s bright idea. The lead in creating such an environment must come from the system.

For many school people, the kind of strategic planning implied in this process is a new way of operating and requires the exercise of new skills. It is unrealistic to expect that teachers and principals can automatically step into these roles. They need to be trained in the use of data, in the interpretation of data, and in the communication of data, and so on. Schools also need help in establishing effective, valid, inexpensive and unobtrusive measurement systems and data collection and processing systems. Most are only in the early stages of using technology, but there are examples of good practice that can be drawn on. For example, some schools have computer terminals on each teacher’s desk, and the appropriate software that can issue an updated report on the progress and achievement of each student at any time. For this to work also means a change in the work habits of teachers – data entry needs to become a regular and routine part of their day. Teachers also need technical assistance in gaining access to and selecting appropriate assessment instruments, particular for measuring non-cognitive outcomes, such as school climate, self esteem and so on.

But more than this is the need for teachers to believe that there is need for change, they need to believe that using indicators and information systems and planning can help, that the extra work required, especially in getting started will in the longer term be worthwhile in making their job easier and more productive.
Systems can help by fostering a vision of teachers as part of a team – working for the good of the school and for students as a whole, and not just conveyors of subject matter in isolation from others. It is important to realise that not everyone will be comfortable with this vision, and that school authorities need to be able to deal with this appropriately. Making the classroom walls transparent can be threatening to those who have been accustomed to pursuing their roles unchallenged. This raises even further questions about the role of indicators in creating expectations for conformity and divergence, and setting agendas of what is important.

As I have suggested earlier, the process of developing indicators can be useful in itself, by forcing teachers to ask questions about their practices, such as: Why are we doing this? Is this the best way? How else might it be done? The biggest enemy in this process is lack of time (and thus money) to allow it to happen. In our pilot projects we have been fortunate that we have been able to provide grants to schools that enables them to release teachers for a time from their regular duties. But in the longer term, some means must be found which creates the flexibility in teacher’s schedules that will allow them to exercise these important functions.

Finally, systems have an important role to play in the wider community in promoting understanding about indicators. Hargreaves (1988) discusses the reasonable fears that parents will focus on the wrong things, and the importance of encouraging them to be educated consumers of indicator information.

Conclusion

Whether school administrations are able to come to grips with this means of synthesising sophisticated information into the decision making process is an open question. The use of goal based measures, the strong ties between key policy variables and the emphasis on contextual and outcome variables all contribute to the relevance of these strategies. Just how much they will contribute to decision-making will not depend solely on the utility or validity of the information. Even when the information is timely,
accessible and pertinent it will still only be one element in the process of policy formulation. Information alone does not change opinions and attitudes grounded in ideology and experience.

While it is important to establish technical validity and reliability of indicators, we have devoted a great deal of our energy into developing a process of using indicators. This process is not entirely unfamiliar to teachers (even if not well applied), and this focus encourages a sense of purpose and action rather than just another new theory.

Using indicators at the school level is still relatively new, and it is too soon to make judgements about how successful they will be. What is certain, is that the concept continues to have great popularity with policy-makers. We have already seen a large number of innovative applications of indicators, and they appear poised to make a useful contribution to the betterment of schooling.
Exhibit 4

THE NSW SCHOOL DEVELOPMENT AND EVALUATION MODEL

The cornerstone of this model is the school level development and evaluation segment. It is based on the belief that an educational organisation which is motivated and given the resources to develop and evaluate itself, with ongoing advice and regular, cyclic external evaluation should be a healthy one, with everybody, including teachers, parents and the community working together with the explicit aim of improving students' learning.

During 1989, nine schools of varying sizes and levels volunteered to participate in a trial of the model. Each school was visited and the model explained in detail to all staff, using a set of guidelines which covered all aspects of the model and the proposed trial. Each of the participating schools was required to agree to:

- develop a Mission/Statement of Purpose/Philosophy for the school;
- develop/revise School Goals/Objectives in line with the stated mission;
- develop a School Development and Evaluation Plan (SDEP) for the next 3-5 years, covering all aspects of the schools enterprise;
- complete a school level evaluation in accordance with the plan (ie. implement the plan);
- progressively develop a range of Education Performance Indicators, linked to School Mission, Goals, Objectives, Policies, Programs, Practices and Procedures;
- be willing to be involved in a trial External School Evaluation, in which a team of school administrators, teachers and community members, including some from within the school and some from other schools would evaluate the progress towards achieving the plan.

The role of indicators in the plan

Education performance indicators are for the individual school to develop and utilise as a tool in the evaluation of both processes and outcomes, taking cognisance of contexts and inputs, over time. Correctly used, these school level indicators will provide a vital link between development and evaluation, giving a more rigorous approach to the latter. They are only one aspect of the model, one evaluative tool and a focus rather than the focus. Schools are encouraged to attach achievable but at the same time challenging targets to key policy and program objectives. The application of the indicators will form a focus for discussion during the monitoring and external evaluation process.

(Source: Osborne, 1990).
EDUCATION INDICATORS AND THE STAYING ON PROGRAM

Staying On is essentially both an equity and a school effectiveness program. Its impetus was widespread concern that too many students were not completing a full secondary education, were not taking up further studies, leaving themselves with severely limited employment prospects. The program began in 1988 and eventually included 50 schools across the state. Schools were selected for inclusion in the program according to a number of criteria; low retention rates in Years 9-12; low SES; large populations of Aboriginal or non-English speaking students; geographic isolation, and some administrative considerations. Program elements include staff supplements, a second Deputy Principal responsible for enhancing school effectiveness; appointment of Community Liaison Officers; some cash grants, consultancy support and professional development funding.

The move to incorporate indicators into the evaluation of Staying On focused on the program level, but the nature of the program dictated that if this approach was to be used it also had to be valuable to schools and help inform the school improvement process. Accordingly, a plan was sketched to introduce indicators to staying on schools on a pilot basis and to develop them for the program wide context. The program level goal was to develop a framework of indicators to help organise and analyse evaluation information about the program, and inform its ongoing planning and operations. At the school level the use of indicators was thought to have the benefits of discovering of trends, patterns and relationships which can help the program by pointing the way to better curriculum delivery structures and hopefully better pedagogy.

To explore how and what indicators might apply to staying on, we started with a two dimensional matrix, with the Staying On objectives (participation, curriculum, school ethos, teacher effectiveness, course and career advise, parent participation/community involvement) on the vertical axis and the four indicator categories (context, input, output, processes) on the horizontal axis. We then went through a protracted process of suggesting, refining and rejecting or accepting information which could help evaluate the program. The resulting list contained some items not adopted because they were uninterpretable. This meant that there were some objectives where we could not assign an indicator, for example teacher effects. The following indicator areas were selected for the pilot. Together, they form a comprehensive coverage of the program objectives.

- Percent Disadvantaged students
- Total Fiscal Effort
- student attitudes
- subject choice
- post-school destinations
- retention in grades 9–12.

(Source: Stevens, 1990).
Bibliography


Chapter 3

Some Observations on School-Based Management Programs in the United States

Unpublished paper
Some Observations on School-Based Management Programs in the United States

Introduction

It is some years now since the report Schools Renewal (the Scott Report) heralded widespread changes to the education system in NSW. One of the cornerstones of the Schools Renewal strategy has been the devolution and decentralization of responsibilities away from the central Head Office to the regional and school level, based largely on the belief that bringing control closer to the site of service delivery will lead to more appropriate and more effective decision-making and thus better outcomes for students.

This trend towards decentralization is a world-wide phenomenon, and counterparts to the NSW experience can be found in countries as diverse as Canada, New Zealand, France, Norway, the United Kingdom and the United States. Within these countries, the formal mechanisms for implementing devolution have taken many directions, and have many names and acronyms: School-Based Management, Site-Based Management, School-Centered Decision-Making, Local Management of Schools and Shared Decision Making being among the most widely used. Some commentators use these terms synonymously, others claim there are real differences in the philosophy and thus the practical implications of these different approaches. The same could be said for the concepts of devolution and decentralization.

The literature on the theory underlying SBM has grown extensively in recent
years, and will not be revisited here. The standard reference in Australia, Caldwell and Spinks’ *The Self-Managing School*, gives good coverage of the issues, and more importantly was seminal in influencing the direction of decentralization plans in most state systems. Other extensive reviews of the underlying principles of SBM can be found in Clune and White (1988), David (1989), Guthrie (1986), for example. While most commentators agree on the basic principles of how SBM should work, the implementation of these principles has varied enormously in practice.

This paper describes some findings of an investigation into the processes and outcomes of such programs in the United States. While recognizing the potential differences in the meanings of these various terminologies, this paper uses the generic term School-Based Management (SBM) for the sake of brevity.

The present investigation took place in two stages. First, school districts which had implemented some form of SBM were identified from a review of the literature and leads provided by researchers in this field. In each of these districts, a senior administrator (usually the superintendent or person responsible for overseeing SBM) was interviewed, and information collected about the extent of SBM, how it was organized, how it originated and what they felt had been achieved. Additional documentation describing the SBM program in the district were also gathered during this interview. In the second stage, this list was narrowed down to a final selection of 19 school districts representative of schooling conditions across the nation. In each of these districts one high school (typically grades 9-12) and one middle school (grades 6-8) were nominated for intensive study. The criteria for selection of these schools were that: (1) they had participated in SBM programs for more than 2 years; (2) were known to have attempted to use their delegated powers to make changes within the school; (3) were willing to participate in the study. In each school visited, the principal and other school leaders, department heads, and a sample of mathematics and science teachers were interviewed separately or in small groups. Relevant school documents were also collected and analysed. At district offices, staff responsible for supervising mathematics
and science, directors of personnel, professional development and budgeting/finance were also interviewed. Information was also gathered from focus group interviews conducted with mathematics and science teachers from other schools in the district.

This research design does not allow assertion of causality of any effects found to any particular source. Indeed given that many other, often overlapping reforms including reporting and accountability schemes, “effective schools” programs, the introduction of “teaming” in middle schools and national curriculum frameworks were also being implemented at the same time, it would be extremely difficult if not impractical to evaluate the separate impact of such complex initiatives. The focus in this study was on gathering information about the process of SBM implementation and in particular, how it had impacted on mathematics and science teaching. However, across the diverse sites, some common experiences stand out which have implications for the improvement of practice elsewhere.

Before these findings are discussed, a note on context is necessary. The concept of “local control”, as practiced in the United States, is often not well understood. Unlike Australian education systems, where legal authority for schooling is vested in state governments, in the United States such powers are the province of school districts, governed by elected or appointed School Boards. Districts are configured in many and varied ways, sometimes overlapping with city, county or municipal boundaries, sometimes are consolidations of several local government areas, but sometimes consist of schools in a single suburb. The school districts range in size from those with only a single school to those with over a million students and up to 300 schools. Each school district is responsible for its own curriculum, standards, policies, budget and is the employing authority for teachers and administrative staff. What exists in many areas are

25 The study reported here was sponsored by the National Science Foundation (NSF), and was principally concerned with the effects of SBM on science and mathematics education. Full details of the study and its findings are reported in Hatry, et al. (1993). The views expressed in this paper are those of the author and not those of the Urban Institute of the NSF.
highly politicized and micro-managed "local tyrannies". The extent to which where the school district, through the School Board, prescribes not only policy but also what is taught, when it is taught and how it is taught been seen in NSW schools for many decades. Thus any scheme which turns control over these things to the school level, even in very limited ways, can be a radical departure from previous practice.

Looking across the school districts studied, the most evident feature is the diversity of practices which are labelled as SBM, and that almost nowhere is devolution as complete as the theoretical models suggest it should. Central offices still generally placed considerable restrictions on how budgets can be spent, set pupil/teacher ratios (and thus determined hiring and firing policies), controlled utilities and maintenance accounts, and by mandating assessment objectives and textbooks selection, exerted strong *de facto* control over curriculum and instruction. In all districts, the amount of discretionary money available to schools was limited, and since many districts were experiencing fiscal crisis, the amount of discretionary money available was sometimes less than before SBM. Since teacher salaries are always the largest component of school budgets, without substantial outside supplementation, there will seldom be sufficient funds to mount major or radically new programs unless teacher salaries are redistributed. Few schools had the ability to trade-off teacher positions for other goods or services, and of those that could do so, even fewer had exercised this option.

Districts differed in their interpretation of what decentralization meant: in some it was simply the delegation of more power and responsibility directly to the principal, in others a site-based management council or team became the ultimate decision-making body. In others the council was merely advisory to the principal. While the SBM model in every district called for the establishment of some form of site council, how these councils were constituted, how they made their decisions, and what kinds of issues they considered also varied greatly. Some functioned as an executive group elected from the faculty, some were composed only of department heads, while in others the whole school staff was involved in making every decision. Which of these modes of operation
was chosen was typically left to each school to decide, as were the by-laws or procedures under which they operated. The prevailing philosophy was that SBM was meaningless unless schools could make these fundamental decisions for themselves, yet in some instances, discussion about these operational on establishment details had not been finalised after a year of meetings for over a year.

Despite this great variation, no one type of organizational structure appeared to be more or less successful than any other. What did emerge was that at every level of the school system and in almost every district there was considerable dissatisfaction that the lines of authority were clear. Principals and central office supervisors often expressed the view that they were no longer sure who was responsible for what. Some teachers were frustrated that they had not gained control over school decision-making as they believed was the intent of SBM. Part of the explanation for this was that in many districts the philosophies, rules and procedures had not been fully worked out before implementation had begun, and part is also due to the lack of training (and lack of depth of training) of all school staff in the initial stages. Both of these themes were raised repeatedly in widely different sites.

Across all districts, there was little evidence that becoming self-managing had led to any sustained improvement in student outcomes (in terms of traditional measures). That no across-the-board improvements for students are evident was not surprising, as few schools had used any additional flexibility that SBM offered to make any significant changes to their curriculum or instructional practices, with a few notable exceptions. This is not to say that nothing of value had been achieved or that SBM itself is fundamentally flawed in conception, but that an infinitely more subtle interplay of forces are at work where success is achieved than the simple decentralization model would lead us to believe.

The vignette drawn from the experiences of one school shown in Exhibit 1 demonstrates the kinds of successes that are possible. This brief overview does not do
justice to what is a very rich set of findings, but the experiences help to underscore some important lessons:

1. Process issues should be decided early and should be decided early and should be understood by all, such as how many members the site council has, how they are elected and who they represent, how often they meet and for how long, and whether they make decisions by consensus or by majority vote. These things have practical and symbolic importance, but the means should not out shadow the end which they are meant to achieve. The level at which certain decisions are to be made has also to be established early are important. It may well be appropriate to leave a lot of the decision-making to the department level rather than at the whole school level, but who does what has to be clearly established and agreed to. The issues in setting up and running SBM are very different for high schools than primary schools, and even different for very large schools than small schools, for example, in how the opinions of all teachers are canvassed and communicated in an efficient and timely manner.

2. Good initial training (perhaps by non-educators, such as business managers) outside of the school setting for administrators and council members was highly regarded and seems to be effective in “breaking the ice” and getting the team started quickly. Principals and teachers, in particular, have seldom had the extensive training or experience in financial management that SBM demands. Although costly, outside training has the benefits that outsiders often carry more credibility, being out of the school setting means that distractions are minimal, and offers opportunity for increased communication across levels that might not have happened before. This may be particularly important in schools where things are not going well and there is not a lot of trust or confidence in the leadership.

3. Ongoing training and deliberate efforts to renew commitment to SBM is also needed so that it is seen to be part of everybody’s responsibilities. Successful programs sometimes come at considerable personal cost to individuals, and there is a limit to how long they can keep this up.
4. Time for meetings and other SBM activities has to be found within the school day, or council members must be otherwise compensated. SBM has to become part of the regular operation of the school and not seen as just another “extra”.

5. Attention has to be given by the school as to how their newly-gained SBM powers can be directed to making specific or coherent attempts to improve student outcomes. While the decision-making process itself is important, more important is the need for participants to really understand and internalize that the ultimate focus of SBM must be on improving student outcomes and not simply the retribution of the locus and means of decision-making. While SBM might empower some (but at the same time can disenfranchise others), on its own it does little to improve the educational experiences of students.

6. There needs to be some formal mechanism for linking ideas to goals to budgets and to implementation. These links can be made explicit through a School Improvement Plan, which must be seen as a working document, part of an ongoing process, not an end in itself and not just more paperwork. Similarly the plan must be monitored and evaluated constantly in such a way that it can provide advice as to inform revisions to the plan. This also helps to make SBM more forward looking and not simply responsive to ad hoc requests, and not the province of a few individuals.

7. There needs to be clear understanding of what decision-making power and authority all the players have to prevent confusion and to save time. It makes no sense for the school council to make every decision - some of these are clearly decisions that should be made by the administration, or by the departments. There needs to be some ordering of priorities for what gets discussed by the council. Clear division of responsibility helps to streamline the process – let SBM worry about the big things and the long term.

8. SBM has to be seen to be sufficiently successful in the short term if it is to sustain long term interest. SBM will appear to be most successful when it undertakes projects with clear, if limited goals. At the same time, it needs to avoid “projectitis”, the tendency
towards undertaking more and more programs with little or no coherence and which address problems in a piecemeal and sporadic fashion rather than looking at root causes.

9. Effective programs don’t have to be high priced – but they do need to have the budget flexibility (and possibly seed money) to get started. Relatively minor changes can have important consequences, even in “difficult” schools, and even if these changes are at first more symbolic than substantive. Improvement in particular subject areas for both regular students and those at risk can be brought about indirectly and accumulate over time, for example, by bringing about improvements in discipline and school environment leading to increased motivation and better attitudes.

10. There should be no expectation that will instantly solve all of the problems of education or school management. Plain good management is also needed to keep everything going – like keeping communication open, establishing rules and policies and enforcing them fairly and rigorously, keeping student achievement and welfare as the highest priority, and treating teachers as professional adults are all needed to make a school work effectively. SBM is only a tool that can help these things happen. With weak and uncommitted leadership and faculty, SBM will not make any difference at all.

11. Personalities matter a great deal. While SBM usually means principals giving up some control, this doesn’t necessarily mean a diminution in their power. Principals can still be dynamic and flamboyant but they also need to be able to lead rather than dictate, to manage rather than carry everything themselves, to delegate and to work with and for people. Total commitment and support from the administration of the school is essential, but at the same time it is not impossible for a strong and charismatic principal to empower others and seek input from others.

Conclusion

Our findings from this investigation correspond closely with those of others who have examined the literature on the outcomes of school-based management in practice. In
an extensive review and synthesis of the literature, Malen, Ogawa and Kranz (1990) also found that, taken as a whole, there was little evidence that school-based management alters influence relationships, renews school organizations, or develops the qualities of academically effective schools. Indeed, as Purkey (1990) concludes, school-based management:

"... is not a panacea, and certainly not a magic wand. By itself, it can not improve on math and geography tests, keep at-risk students from leaving school if they believe that success in school is irrelevant or unobtainable; suddenly convince college seniors dreaming of expense account lunches, BMWs and high status jobs to forsake all that to enter classrooms as teachers in central city schools... School-based management will not necessarily empower teachers, boost morale or make them engage in pedagogically progressive instructional practices. Nor is there much reason to think, even theoretically, that school-based management will turn ineffective schools into effective schools, bring eternal peace and harmony to union and district relations, or result in outpourings of public gratitude and demonstrations of appreciation for a job well done by the public schools." (p. 373).

School-based management is simply a tool that faculties, but does not guarantee, a change in the distribution of power and authority over matters that can affect the lives and work of teachers and students. But like all tools, unless the end which it is meant to accomplish is kept clearly in sight, and it is used appropriately towards that end, school-based management will not achieve its intended purpose.
Bibliography


Exhibit 1.

A Case Study in the Introduction of School-Based Management

Palm HS is located in a largely suburban area of southern Florida. The school has 2400 students in grades 9-12, of whom 19% are black, 14% Hispanic, and 60% white; There are 165 teachers on the staff. The student population is also diverse in other ways: some are very affluent and some very poor, some students achieve highly on standardized tests and college entrance exams, and yet there are also many who cannot read. (The name of the school has been changed, but all other details are correct).

When the district implemented SBM in 1987–88, interested schools were invited to develop a proposal setting out their plans, which were supposed to span over a three year period. For the plan to be accepted, two thirds of the faculty had to agree to it. Developing the proposal for Palm took 5–6 months. Teachers were initially cynical that this would be “just another program”, and a great deal of preparatory work by the principal was necessary to convince them that change could be beneficial. Meetings were held at the Sub and Faculty level, led by the principal and central office people. Existing parent organizations were also invited into the process, as was the Student Governance Council.

Getting organized and setting rules and policies took up most of the first year. Once the proposal had been accepted, a large number of the faculty wanted to be part of the decision-making process. Not wanting to discourage anyone, a very large council (32 members in all) was established, known as the SBM cadre The work of the council was supposed to be to set policy, allocate the budget and allocate resources, review school procedures. The council was supposed to give everybody a say in decision-making for the school. They soon recognized that this was not a very efficient means of operating, so a steering group was elected who would be responsible for research, communication and dissemination and monitoring the plan. A number of task-forces were set up, whose function was to develop recommendations which were then voted on by full council.

An important role for the SBM Council was to develop the school improvement plan. A self-assessment of each department was the starting point for their plan of action, considering everything from the physical plant to the curriculum, leading to recommendations which were prioritized by the steering committee and translated into the budget. These self assessments were presented to the entire staff at a dinner meeting, which for many was the first opportunity to meet others from other subject areas, and for my first insight into the operation of other disciplines. With the introduction of SBM, relationships changed completely in the school. In the spirit of increasing democracy, the school decided that department heads would be elected from the faculty rather than appointed by the principal. Communications needed to be increased sideways. Department Heads and cadre members had to talk to
each other, rather than directly with the principal. The result was a flattening of the hierarchical structure. In the process of presenting the results of the needs assessment, the department heads had to present the results to the faculty. This was the first time that many of them had ever listened to each others problems. Most teachers, for example, had no idea of the amount or cost of the consumables that the science department required. A much greater respect and appreciation for each others abilities developed by bringing parents into the cadre and task-force meetings. SBM turned out to be less of a problem than expected. Some mistakes were made. Some people who were good politicians were found not to be necessarily good leaders. New teachers had to be made aware that “this is the way we do things here”; they had to understand that SBM required them to make a commitment, they couldn’t hide away in their classrooms.

One of the outcomes of the needs assessment was the recognition that achievement levels and completion rates for certain sub-groups of students were particularly low. Teachers proposed as one solution a school-within-a-school for 11th-graders identified as at risk of not graduating. A team of teachers were assigned to teach these students in the core subjects and allowed a great deal of flexibility, for example, in scheduling lessons, not constrained by the regular timetable. Another change was an attempt to make the curriculum for the at-risk students more “hands-on” and focussed on problem solving in real-life settings rather than the tradition approach of drill and practice in basic skills. Students were required to develop their own programs of study, by selecting and researching a local issue and propose a solution and presenting the project to the entire staff in a formal presentation. This program was based on the belief that while these students may sometimes be lacking in academic skills are not unintelligent, but do not respond in the traditional classroom environment for one reason or another. To ensure that the program could deliver the individual attention needed, the school hired some part-time teachers using money reallocated from other areas, and used community volunteers to help with the placement of students while doing their research.

The purpose of these programs was to address the school’s attendance and drop-out problems. The regular school setting was not working for these students, some of whom would be 20 years old before they would finish in the regular program. The outcomes of the program was to significantly increase attendance and to bring students back to school who had dropped out. Achievement scores for these students increased, as did those of students in the mainstream, because less classroom time had to be spent on discipline. Students and teachers responded positively to the more safe and orderly climate that followed. The program has continued to evolve and expand (now up to 80 students are involved), and has progressed beyond this to an alternative program where students can work part-time at school, part-time at a job, and complete their academic requirements at night school.
Chapter 4

An Evaluation of the Effects of School Based Management on Science and Math Education and on Students

An Evaluation of the Effects of School Based Management on Science and Math Education and on Students

Introduction

As noted in the previous chapter, this study was primarily concerned with investigating the extent to which school based management efforts had impacted on science and math education, and to a lesser extent how this in turn may have impacted on outcomes achieved by students. "Math and science education" is a loose term, encompassing complex interactions between school organizational structures, state, district and local school policies, the content of instruction, teaching styles and interpersonal dynamics. To fully capture the dynamics of these interactions with SBM in all its forms was beyond the scope of this study. Our focus instead, was limited to just two areas – what is taught (the curriculum) and how it is taught (instructional practices). Although curriculum and instruction are closely related conceptually and in practice, they are separated here for ease of reporting. We asked respondents only about changes to mathematics and science education. There may or may not have been changes in other subject areas resulting from school-based decisions in particular schools.

i. Changes to mathematics and science curriculum

SBM appears to have had little impact on mathematics and science curriculum as yet: 14 of the 19 schools studied reported that their curriculum was still largely determined by district or state requirements (in the form of guidelines, frameworks, objectives or standardized tests). However, within these constraints, most schools had
some ability to decide on course offerings, course content and the sequencing of material within courses, either on their own authority or after approval from the district. The following comments are typical:

The science curriculum is determined by the Curriculum Action Plans (CAP) and Criterion Reference Tests originating from the central office, but we have always had some options over the curriculum. The county gives the general topics and objectives to be covered (in the CAP) but how and when they are implemented is up to the teachers. To make best use of limited resources, different teachers at the same grade level do not teach the same topics at the same time. For example, not every class in a grade can study microbes at the same time, because there are not enough microscopes to go around everyone.

District-wide committees will continue to develop core learning objectives/content, district assessment instruments and review texts/learning materials to identify one or more best suited for teaching core objectives as recommended texts/materials. Schools have the option of developing materials themselves or making their own study of available materials. Schools remain responsible for teaching core objectives and for preparing students for the district-wide assessment. Core objectives represent 70% or less of curriculum, which provides schools or teachers leeway for in-depth study of content or inclusion of related content. Superintendent says school can study, at school level, the most effective way of teaching the core.

Many schools had the ability to make these kinds of decisions prior to SBM, as can schools without SBM, although who makes the decisions and how they are made differed in SBM schools. Previously, the initiative to introduce a new course, for example, would have been made by the principal or other administrators, but now recommendations or decision to do so are usually made by the departments or a school-wide curriculum committee.

Only two schools in our sample, both from New Mexico had really made radical departures from past practice in terms of curriculum, and in both cases this resulted from being part of a wider Relearning/Essential Schools initiative (although the decision to be part of the project was a school-based decision). Other than these two exceptions, none of those interviewed said that they are now able to make curriculum decisions
which would have been impossible for them to have made before, but the process has been made easier, and a climate conducive to change and experimentation has been created.

However, teachers in several schools noted that while there had not yet been any substantial changes in curriculum, they had become more aware of the need to change, less accepting of the status quo, and had begun discussing possibilities for doing things differently.

While curriculum is set by the county, teachers said that they have always had freedom in how to teach. Math teachers are bound by the class sequencing – Algebra 1, Geometry, Algebra 2 – but some would like to be able to alter the sequences or to develop math courses that are not handed down and geared more to their students, e.g. a nine week elective course on the techniques of graphing, even though this is integrated into other courses. Math teachers said that if nothing else, being able to make these decisions makes them feel better about themselves and more professional. Science teachers said they had discussed different ways in which they could do things, such as having smaller lab classes but larger discussion classes, and more health-oriented or more use-oriented courses for students at the lower level. Before we would not have even done this because there was the feeling they really couldn’t achieve anything.

We would like to be able to rearrange the school day, adjust course outlines, have more interaction among the disciplines, for example longer science lab periods and advanced math and computer labs. We would like to teach less, but teach better rather than giving superficial coverage.

SBM, and more particularly the philosophy of shared decision-making, has also brought important changes to the way that curriculum is determined at the district level. This is especially important in smaller districts, where there may only be one or a few high schools. (It may be misleading to say that the curriculum is determined by the district in these circumstances, unless the administration imports the curriculum from elsewhere without consultation).

There are opportunities for teachers to be involved in the curriculum as part of the development committees, and for recent curriculum revisions (such as the new reading curriculum) considerable efforts were taken to distribute draft versions to all teachers in the
district for comment before final adoption. Involvement of particular teachers in these committees is in part a school decision, and is connected with career ladder status.

Teachers in the Poway school system are involved in making curriculum decisions through their participation in a district-wide curriculum committee, which acts as an advisory body to the central staff in suggesting and effecting curriculum change. Through this committee teachers have been able to affect changes in courses offered, course content, and textbooks selected. For example, the committee was able to change the content of an Algebra/Geometry course that would be offered at both the middle and high school level, with the intent of bridging between the two levels.

About half of the schools surveyed said that they had added to or deleted from the courses they were able to offer. Three middle schools introduced algebra or pre-algebra courses (or made them available to the majority of students rather than select individuals. Two schools re-introduced lower level courses. One middle school lowered the county requirement for all students to take pre-algebra for students in a drop-out prevention program, instead allowing them to take general math. Some high schools had obtained waivers from the district to alter graduation requirements:

The school requires all students to take 4 credits/years of math to graduate. The district and the state only require 3 credits. Math is compulsory for all students in the school, including seniors. As a result, the school has a very large number of math classes. The key thing is that teachers had input into this decision – teachers make decisions about what courses were offered, not administrators.

General math is not offered at the school any more, because we thought that the standard was too low. General math was felt to be about a 4th or 5th grade standard. If students haven’t learned it by high school they probably never will. Instead we offer a more practical, applied course.

At the insistence of the math teachers the number of advanced level math courses has been increased – two full sections of calculus are now offered as opposed to one. The Basic Math 2 and Math 3 courses have been changed to blend mathematical functions incorporating graphing, courses are more technology focused.
Schools which developed or introduced new courses most often did so to draw on the particular talents of the teachers in the school at that particular time.

West HS added an aviation course proposed by a member of the science faculty (a former pilot), which opened up an area of interest for students, some of whom would never have been exposed to the level of learning and mathematical and scientific concepts involved (like weather patterns) otherwise. Extra credibility and concrete linkage to the real world was provided by requiring students to take as their final exam for the course one set by the FAA.

The school also established an astronomy course, using a newly appointed teacher who had skills in Astronomy. Likewise, a science teacher at the school was developing a wildlife biology course in conjunction with the Bureau of Land Management, Forestry Department and Utah State University. As a science teacher at the schoolsaid: “Teachers have great flexibility to introduce topics and courses so long as they fit the core objectives.”

Another West HS example: the faculty decided that the technical and vocational education curriculum was not meeting the needs of the students so they developed their own, and established an “Academy of Technology”. Students are required to have good grades in the basic subjects to be able to enrol. These courses have concurrent credit with the High School and Community College and count towards graduation in either institution. Teachers from the school were also accepted as adjunct staff at the college. These activities were initiated within the department but went through the School Improvement Council.

The Capital HS science department introduced a new science course entitled Science for Today’s World – a 9th grade science course added as a pre-requisite for taking biology, chemistry and physics.

Other schools introduced new courses based on a their perceptions of the needs of their particular students or weaknesses in the existing curriculum. In a few schools, this called for significant rearrangement of the school timetable to accommodate these curriculum changes. (The subject of arrangements needed to allow teaching teams to operate is a separate topic, discussed in the instructional practices section). The process used to develop introduce new courses at Sandia HS is typical:
Curriculum changes are normally discussed first in department meetings, then presented to the curriculum committee by the department head (in this school the curriculum committee consists of all department heads) and are acted upon by the committee or the management council. Depending on the issue, some suggestions for curricular change may be overseen by central level personnel—particularly if waivers or policy changes are required. Prior to SBM implementation, any school-based curriculum decisions were made by the principal and assistant principals.

At Fred Lyn MS, after introducing SBM, faculty began discussing what they needed to do to make students better learners. They established some new keyboarding and computer operations courses. Teachers were paid from school funds to work over the summer to develop detailed lesson plans for these courses.

Highlands MS has developed some new science courses. An Environmental science was piloted in the 6th grade in 1992 but after review was to be offered in the 7th and 8th grades in future. An advanced life science course was developed for 8th grade (a one semester elective). The process of developing these courses illustrates what the principal saw as the main role of SBM—involving everybody in the school in discussion and thinking about what students needed in science and how this related to other subjects. The other major change is that they have been able to determine which courses they will offer. They have “untracked”, that is, they now have only one math course at 6th grade and two courses (basic and advanced) in 8th grade rather than several math courses at each level.

At Rancho Bernardo HS, science teachers reported being able to change the course content of a new physical science class offered to all freshmen across the district. All teachers had the opportunity to be involved in this activity. Schools had offered different physical science courses that varied in content. Teachers from all three schools in the district were involved in standardizing the content to adopt a thematic approach to teaching the class and in finding a textbook to use.

Two schools also mentioned other problems concerning SBM and curriculum. Teachers in one school said that while they have input into the overall program, they no longer have time to meet and discuss particular subject issues. "Morning time is taken up by committee meetings of one kind or other, and the afternoon is taken up by the activity program. SBM has siphoned too much time away from curriculum issues into whole school issues. It has been negative in this respect". In another school, teachers were frustrated that their attempts to make curriculum changes had been blocked by the
district. An example was the math department’s attempt to introduce a probability/statistics course at the pre-algebra level. A proposal to introduce the course was developed twice, and submitted to the superintendent’s office but was rejected each time.

ii. Textbook Choice

Given the importance of textbooks in determining what students are taught, the ability of teachers to select the textbook they use is another important element in determining curriculum. For various reasons, control over textbook selection has become a key issue for many school districts, and few schools in our survey had gained much additional flexibility in textbook selection as a result of SBM. In the majority of districts, schools must choose from among a limited number of texts that have been adopted by the district, usually on the advice of a committee comprised of administrators and teachers. In some cases districts adopt from a similarly developed state sanctioned list. Each subject is usually reviewed every 3-5 years.

Similar to their involvement in district-wide curriculum development committees, participation in district-wide textbook adoption committees is an important vehicle for teacher input in smaller districts. For example, at one of three high schools in Adams 12 District, several math teachers wanted to change the textbook they used to teach calculus. The math teachers presented a proposal to the school’s curriculum committee, the issue was then presented to the School Board, and the suggested textbook added to the district approved list and ultimately adopted for use at the school.

The majority of schools in our sample had not changed texts or the way texts are selected as a result of SBM. At the school level, the decision to select a particular text was most often made by consensus within departments with the concurrence of the principal, although sometimes principals retained veto power (but seldom exercised it). SBM’s impact on text selection, has been to formalize the previously widely practiced
informal delegation to teachers of substantial input into text selection from those on the State’s and/or District’s lists.

Textbook selection often hinged as heavily on budgetary considerations as on instructional issues. There were several instances, however, where extra money was allocated to allow purchase of class sets of books that would not have been possible without the added budgetary flexibility allowed by SBM. Some schools, however, were severely hampered in their ability to impact on the textbook selection process because of restrictive categorical funding arrangements.

Central office administrators raised several concerns about the implications of unrestricted textbook selection by schools. The first is an equity issue, particularly in districts with high student mobility. Supervisors were concerned that students moving from school to school even within the same district could have vastly different experiences. They also worried about the extent to which this lack of standardization would impact on district-wide examinations and state minimum competency tests. The second area of concern was the impact that text selection had on district curriculum reform efforts. Some districts were trying to revise curriculum, moving away from a content or text-based approach to a performance or outcomes based approach. According to supervisors, allowing schools to choose outside the county list could be a problem in implementing NCTM standards, if schools attempted to save money or redirect it elsewhere and keep old textbooks. Other problems mentioned were: evaluation/testing comparisons across schools; preparation of students for subsequent courses; especially in math; moving from school to school could be more difficult for students if material covered by each school was different; and additional costs would be incurred in having multiple texts, which could be wasted if the only teacher using the text went elsewhere. (Inventories need to be higher so that each text has spares. Most books cost $30 or $40 each).
iii. Changes in Instructional Practices

Instructional practices, loosely defined, refers to all the activities that teachers/schools do to implement the curriculum. Many national and district reforms have been enacted that directly address the ways in which teachers attempt to engage students in learning. This is also a time of rapid technological change, with many more options for using technology now available to teachers.

These reforms, and indeed any innovative teaching practice, can and do occur in schools quite independently of SBM. The purpose of our investigation was not simply to catalogue innovative practices, but to determine the extent to which SBM had contributed to their implementation. We found that in almost every school, teachers said that they already had considerable flexibility over their instructional/teaching methods. (This is not a finding unique to this study).

SBM appears to be useful in creating an environment permissive of change, such as by increasing teachers’ feelings of professionalism and collegiality, and influencing the pace of change by allowing schools to more appropriately target resources. In reading the findings presented below a note of caution is necessary: while we found (isolated) examples of how SBM had impacted on instruction in more than half of the schools in our sample, the overall impact has been small. A major exception occurs in team-teaching used primarily by middle schools, which in effect is an important form of school decentralization, whether or not it has been encouraged by a formal SBM program. Lacking an appropriate set of non-SBM comparison schools, we are unable to say with certainty whether these increase in instructional practices were significantly different or encouraged by SBM.

The most widespread change in instructional practice encouraged by SBM was the introduction of a broader range of teaching styles, including the use of cooperative learning, small group discussions and tutorials, an increased focus on problem solving and research courses, greater use of manipulatives in mathematics classes, and more
hands-on lessons and supposed real world applications generally. These are all in line with the changes suggested in the NCTM standards.

West HS, for example, used a discovery learning model to teach 8th grade science. Students were given the State’s objectives at the beginning of the course, and were asked how they could reach them, what projects and research they would need do, how they would report on this and how they will be graded. Several schools had also begun using different assessment techniques, such as buddy exams and portfolio assessments to compliment these new teaching styles. SBM’s contribution to this was through allowing teachers access to professional development (e.g. one school bought each math teacher a copy of the NCTM standards), in some cases by allowing resources to be allocated (for purchase of manipulatives, for example), but perhaps most importantly in changing attitudes within schools and the district. We were told in some schools, for example, that teachers were more likely to use equipment that they had ordered for themselves than when it was supplied to them. In one school we observed that the district had supplied each math class with “thousands of dollars” worth of manipulatives that teachers did not use, because they complained, they had not been trained in how to use them. Two schools in New Mexico had used SBM procedures to opt into the Relearning model, which reinforced the use of all the above teaching practices.

iv. Scheduling flexibility

Some schools used the SBM process to make significant changes to the school time table to allow the above changes to take place. Five high schools had adopted some form of block scheduling, allowing double period lab classes to be held, trading off extra time for one subject with shorter time in another. This is particularly important in science classes where the teacher has to set up for the students “Students don’t get to learn nearly as much as when they do it for themselves”.
Four high schools adopted team teaching arrangements similar to the approach common in middle schools. Typically, teaming in high schools operated in only one or two grades, or in school–within–school programs. Teams usually consisted of the teachers of the “core” subjects (English, math, science and social studies). Each team was assigned about 100 to 120 students. Students usually took each subject each day (plus electives). Each teacher had one individual planning period and a common planning period with the other teachers in the team. Teachers said that they primarily used these joint planning periods to address students with special needs, and attitude and behavior problems. They can also use the planning period to design collaborative, integrative subject matter. The teams are also free to alter the time within their “block” to lengthen or shorten a particular lesson by mutual consent.

v. Integrating material

More than half of the schools observed had attempted to integrate material covered in two or more courses. For example, in Northwestern HS, English classes asked students to write a biography of a mathematician: and a social studies teacher incorporated map reading, teaching coordinates and some graph reading at the same time these topics were being covered in math class. At Roosevelt MS, the math teachers gave their math vocabulary list to the language arts teacher so that the vocabulary was also taught in language arts classes that week. Two West High School science teachers taught a Biochemistry unit by combining Chemistry and Biology Classes. This freed up one teacher for additional preparation time and allowed swapping of ideas. Also, the teachers felt that students were better prepared in Chemistry because teachers sat down and identified mathematics concepts in Chemistry so that Math teachers could also cover them in math classes to reinforce learning.

Thematic units, where each class worked on a particular aspect of a larger project are also popular. The description of the process from Horizon HS below better illustrates how SBM institutionalized such efforts.
Several teachers at Horizon HS, after being introduced to interdisciplinary teaching methods conducted research on it, developed a structure for it in their particular school, and discussed it with the principal. The issue was then presented to the curriculum committee and adopted for use in the school. This teaching method encompassed both block/alternative scheduling and subject integration and placed heavy emphasis on project-oriented instruction and cooperative learning. The curriculum is based on thematic units (for example, Patterns, Change, Diversity, and Interdependence). Within these units, topics were studied across the four disciplines (English, Math, Science, Social Studies) in the core program.

c. **Expanded/faster introduction of new technology**

The role of SBM was most obvious in increasing the level and usage of technology in mathematics and science teaching, because increased flexibility in the use of school dollars is one of the clearest effects of SBM.

For example, while many of the schools had already used graphing calculators, computers, laser disks, video microscopes, OHP calculators to one extent or another, about half the schools responded that SBM budgeting procedures had allowed them to either purchase more equipment (such as whole-class sets of calculators) or more extensive and appropriate software. Most often, these equipment purchases originated as requests from within a department that were approved by the site council (such as the construction of the computer lab at Morningside MS, the purchase of an Algebra Item Bank by the math department of American HS).

At Highlands MS, the greatest impact on instructional methods appeared to come through the technology program. This is funded separately from the SBM program, and was a district-wide activity. However, how each school implemented the program and spent its money were purely local decisions. In some classrooms, teachers changed their role from lecturer/presenter to facilitator/coordinator of student learning. For example, students worked at computers and then came back to the group and shared what they had learned. This is an attempt to make students become more involved in their own learning. According to the principal, this is a good example of how the local decision-making process can work when you have the money to support it. "Technology is supposed to thread throughout the curriculum, it is not just bringing in hardware and
hoping that it works. To do this well, teachers also need inservice in how to make it
(technology) more useful in the classroom.”

In the schools we visited, we were told of many other examples of how SBM had given teachers, either individually or as departments or interdisciplinary teams, the freedom to try new things or provided the money to implement them. These included:

- funding field trips, excursions and in-school performances;
- establishing cooperative programs between high schools and their feeder schools, with universities, vocational and community;
- establishing business/industry links or greater use of volunteers or consultants; and
- instituting systematic mechanisms for recognizing and rewarding students for academic performance and behavior.

While these practices were not radically innovative, they were helpful and inexpensive, and being instigated by the schools themselves as responses to their own recognized needs. All were credited with helping to create positive environments and raising staff morale.

d. Some overall observations on SBM effects on instructional practices

SBM has allowed and seemed to encourage all the examples of changes in instructional practice noted above. All appeared to be beneficial in and of themselves. While many of the innovations above occurred in isolation, the most promising examples appeared to occur as part of broader, whole school development, as the following example illustrates.

In the past, the curriculum in all schools in the district was the same, the texts were the same, and students could have expected to transfer between schools without great disruption. Under SBM, there appears to be growing diversity between schools, and greater use of supplementary and alternative texts. This is likely to increase in future.

The district has a district-wide instructional process model that hadn’t changed with the introduction of SBM. However, the principal wanted the teachers to know more about differences in learning styles in their students, and had bought some National Association
of Secondary School Principals (NASSP) materials about this with school managed funds. This has implications for how students are grouped and so on, which will impact on instructional practices in the future.

The school had also used school funds to obtain professional development and inservice opportunities for teachers to make further changes in instruction. The science department head had attended a summer school about co-operative learning and integration of science/math/technology education. The department head shared this with other teachers in the school. For example, lasers are becoming more important in Technology Education, and the science department have made an effort to ensure that physics teaching about lasers is complementary.

**Recommendations**

- School districts should pilot new techniques that involve substantial cost and other risks in one school, evaluate the pilot, and quickly disseminate the findings to other schools. When ideas originate within one of the school, central office staff should encourage systematic evaluations of the innovations undertaken, so that other schools can learn from the experience, even if the trial is not successful. Evaluations should identify strengths and weaknesses, and the dollar cost and staff time required to learn how to use the new instructional practice.

- Principals and site councils should explore ways in which SBM could help to encourage individual teachers to learn about, and try out new instructional practices. For example, they could allocate funds and make time available to expose teachers to new approaches through inservice training, or encourage external suppliers to provide accessible briefings to faculty. These efforts should be coordinated with central science and math efforts to seek out new science and math instructional approaches (both procedural and equipment-based approaches), provide for evaluations of them, and bring those approaches with potential to the attention of school faculty, perhaps by helping arrange vendor visits to schools.

- Science and math faculty should be encouraged to consider how the purchase of new equipment could impact on instructional practices and the need for training when using the added flexibility over the use of school funds.

- After a new instructional technique has been tested and adopted by one or more faculty, encourage those faculty to inform (tactfully) other faculty of the procedure's strengths and weaknesses, and provide learning opportunities to these other faculty members. In general, other faculty should not be pressured to conform to the new practices but should be given ample opportunities to learn about and experiment with them.
What have been the outcomes of SBM for students?

One of the main arguments for introducing SBM has been the belief that the perceived poor performance of students (particularly in math and science) is in large part caused by the inability of teachers and schools to solve their own problems because they are hamstrung by rigid central bureaucracy. Freeing schools from central control is supposed to unleash the creative potential of teachers to address the particular needs of their students in ways that far-off administrators cannot. It would seem reasonable that in self-managed schools, that there would in fact be significantly improved outcomes for students.

As noted in the introduction, this study focussed more on the processes of SBM and its effects on mathematics and science education rather than outcomes in particular, and the study design made it impossible to attribute causality to any finding. Even if comparable pre- and post-intervention measures had been available, the uncontrolled influence of other contemporaneous and sometimes inter-related reform efforts, such as increased state and district accountability requirements, the growth of the middle school philosophy, the effective schools movement and the widespread influence of the development of national standards for mathematics would make interpretation difficult. Also, changes in student performance often lag considerably behind reform efforts, and insufficient time may have elapsed for these effects to begin to become evident. The introduction of SBM has seldom been a single event, and typically occurs as a series of sometimes partial and fragmented delegations of authority. Attribution of changes in student performance to the generic concept of “SBM” may therefore require caution.

Nevertheless, we did ask teachers, department chairs, principals and central office personnel their perceptions about what effects SBM had on students in the following areas:

- mathematics and science grades;
• mathematics and science standardized achievement tests;
• interest and motivation in science;
• interest and participation in math and science extra curricular activities; and
• interest in pursuing math and science careers.

Surprisingly, the most common reactions of respondents at all levels was that they did not know whether SBM had impacted student learning or not. This issue was clearly not a high priority for the majority of those interviewed. There seemed to be no differences between districts that had been engaged in SBM for relatively short or longer periods of time as to their perceptions or efforts to measure student outcomes.

This lack of awareness was most prominent at the central level. While we infrequently interviewed those at the very top level of management (Superintendents) or those explicitly involved in evaluation and testing, we did regularly interview those responsible for mathematics and science education. We were frequently told that monitoring standardized test results of schools was a regular part of their role, so data should have been available to them.

A few of the central office personnel voiced our own concerns about attributing causality, given the other reforms also operating in their districts. However, one administrator responded

"Why would you expect it [SBM] to have any effect. That's not what SBM is about."

Another believed that "it will take 13 years (presumably the time for a child to pass completely through the system) for SBM to have any significant effect." Still others expressed the reasonable view that the types of learning encouraged by SBM are not necessarily the skills measured by standardized tests, and that other types of outcome measures (not yet taken) would be needed to draw appropriate conclusions. But the majority of those interviewed at the central level either expressed no view, or thought that perhaps student's test scores or grades had improved (or expected
that they should have), but had no evidence to confirm it. This lack of knowledge was not dependent on any verifiable improvement in the schools in the district.

This lack of ability and willingness of mid-level district staff to speak about SBM impacts suggests:

1. In implementing SBM, districts have not given high enough priority to projecting a vision of what SBM is ultimately supposed to achieve for students. If such a view has been articulated by the superintendent and board, it has not been adequately communicated to all within the district administration and certainly not internalized.

2. In implementing SBM, districts have been overly concerned with processes rather than outcomes (this is not surprising, because this is what districts can impact on). This concern for processes is also evident in the roles that district personnel have played in monitoring the introduction of SBM. Typically, they have been most concerned with school's adherence to union contracts, state and district board policies and financial regulations, and whether there had been problems in implementation such as disputes over decision-making, rather than whether any of the newly-given powers had been used to achieve anything substantive. Few districts seem to have published evaluations of the outcomes of implementing SBM. If districts have conducted quantitative or qualitative outcome analyses of SBM, these have not been adequately communicated to all central administrators.

3. Central administrators often expressed the view that they needed more training in SBM, and many spoke of their uncertainty about their role vis a vis SBM. Many were uncertain of their future and whether they would lose their positions if SBM was taken to its extremes (although in few districts had their been any central office positions eliminated that could be attributed to SBM). Issues concerning the relationship of central office personnel and schools are discussed elsewhere in this report and will not be dwelt on here. However, the relationship between the role of each central office position and self-management of schools appears to have received
little attention in all but a few districts, where a philosophy of shared decision making had begun to permeate the whole district prior to SBM.

This also suggests there is a problem inherent in the way in which districts have implemented SBM on a pilot or optional basis. While the need for caution is obvious, the partial, slow doling out of additional powers reinforces feelings of distrust, and allows too many people to believe that it really doesn’t concern them – or foster any commitment to district wide change. This is especially true for large districts where there are more places to hide. In one county several people interviewed appeared to know nothing about SBM.

At the school level, the picture is somewhat brighter, about half of the schools in our sample able to point to some positive outcomes for students that resulted from some school-based decision (a broader definition than the full application of school-based decision-making procedures). Respondents in a further six schools related anecdotal evidence that there might have been improvements, but were not able to provide any convincing documentation of these. However, in seven of 19 schools, the majority of respondents did not believe that there had been any impacts on students or did not have an opinion. At least there had been no negative impacts for students reported in any of the schools visited, but clearly “SBM” as implemented in these districts is no silver bullet or panacea for across the board improvements.

Not surprisingly, within each school there were differences in perception of positive outcomes (perhaps reflecting access to information), with principals more often asserting positive changes than department heads or teachers.

While it would be expected that it takes some time for the effects of SBM to become apparent, no general trends were observed in how effective schools with different lengths of experience had been. Some of the most successful schools had only been self-managing for a little over a year, while some that had experienced some form of SBM for over four years still showed no evidence of change. Nor was there any
consistent connection between the extent of decentralization of budgetary, staffing or curriculum powers and significant gains in achievement (note that in many schools the extent of instructional freedom has not changed as teachers already had significant discretion over how they taught). Some schools which had been given relatively little added responsibility achieved dramatic results, others had made effective use of their extensive autonomy, while yet others which had significant delegation of responsibility and had significantly changed the delivery of instruction had achieved little gain.

Specific findings

A detailed analysis linking initiatives and outcomes is beyond the scope of this report, but some specific examples are worthy of mention.

Changes in student grades and standardized test results

We asked respondents whether there had been any effects on student grades, including Advanced Placement scores or on standardized test results that they would attribute to SBM. Standardized test scores include SAT and ACT test results, as well as district-wide or school administered normative or criterion referenced tests. While many schools were required to have yearly tests for some grades in mathematics, only a few districts had regular science testing programs.

None of the faculty interviewed indicated that grades had improved since the introduction of SBM. Some teachers felt that there had been an improvement in student interest and motivation which will ultimately affect grades. The department head anticipates that student grades in a new, teacher-developed course – Science in Today’s World – will be higher and that students will have a better comprehension of scientific principles. (This course is more ‘hands-on’ and less theoretical than other courses and is pitched to the middle-lower end of the ability range). No change in standardized test results was noted.

No improvement in grades was evident to any of those interviewed. Most responded that they either did not know if there had been any changes in test scores or that there was no discernable change. However, one science teacher did say that SAT (unspecified) scores had risen by 10 points over the last few years, but was extremely cautious about linking
this to any one factor. Neither math teacher reported any discernable changes in SAT scores.

Marginal improvements in both math and science were reported, but in both cases the respondents (ST, MDH) thought that this was probably more directly related to increased teacher enthusiasm and innovation (which may be a by-product of SBM). The biology teacher noticed a slight increase in SAT scores. This was not confirmed by others interviewed, but there was a slight increase in ACT scores reported.

No evidence of improvement in grades was provided, and respondents differed in their perceptions of whether there had been any change. The principal felt that “scores go up a little each year” – but didn’t know why. Other teachers felt that teaming and cooperative learning had benefitted some students but not others. The SDH felt that restructuring focus and the movement towards hands-on collaborative learning, less lecturing from the book should have a big impact, but hasn’t been around long enough to assess the impact. No change in standardized test scores noted.

In one school there did appear to be positive changes, but the effects on students in general have been indirect, and mostly result from improvements in student attendance, attitude and discipline (discussed below). “Nothing that the senate has dealt with has had a different impact on math than other subjects.” However, the school has developed some specific programs targeted at particular groups of students which were thought to be successful. Summer schools, funded and organized through the SBM process are used for advancement as well as make-up, which was credited with allowing greater enrolment in AP courses, and also increases in AP scores. Also, a school within a school was created, which allowed a team of teachers to focus intensively on a small group of students. Graduation rates and thus grades had been increased for these at-risk students. On the negative side to this, teachers said that while attendance might be higher, failure rates may also be higher because of the deficits of past absences. The other trade-off is that discipline is harder because they are now retaining troublesome students who used to be absent.

Teachers believed that as a result of their special programs, there had been an improvement in achievement and test results for students in the PALS program and regular students, because the improvement in classroom discipline allowed teachers to concentrate on teaching and freed students from distraction – although math is least improved.

Neither the principal or teachers could generalize about whether there had been any increase in grades because of SBM, and no study had yet been done to determine impacts.
This school has always had high levels of achievement, but is changing demographically. However, some teachers did point to specific areas of improvement that they were familiar with that resulted from school initiated programs. Having the additional equipment (graphing calculators, computers and overhead projecting calculators for math) had raised the AP scores, but most of this is was attributed to the quality of the teachers. Having the tutorials (for Spanish speaking students and SAT preparation classes was expected to help these students, but they haven't seen any impact yet.

The principal believed that the collection of changes made in the school, such as the tighter supervision of teachers (many of these seemed to be the actions of the principal rather than the site council) had indirectly allowed improvements in student performance in math and science. However, the school has always had a reputation for high achievement, which some teachers attributed more to the level of input they give in class rather than anything to do with SBM. Others felt that student performance has probable increased because of the programs they have been able to introduce through SBM money. They set high standards for the students which they seem to be meeting. They have evidence that in the last few years that more students are competing homework and fewer are getting incomplete grades, however, the school's failure free policy complicates this analysis (the lowest grade a student can receive is a C-). Otherwise they get an Incomplete or No Grade (usually for non-attendance), or P (working to capacity but never going to pass). Students with an Incomplete have to go to an 8th period (after school) and summer school for remediation. The state allowed the school to use a trimester system on a trial basis, and test scores improved 13 points (this was three years ago). The principal believes the students are better off because they have more electives, and teachers are made to be more efficient because they have to cover the material more quickly.

According to the principal, there have been great gains in kids at the upper level for those who have taken the opportunities. They have been very successful in state competitions, also students taking AP calculus exams and so on. The school seems to be above the state and county average. Principal attributes this 75% to the change in relationship through the reassignment of duties and the tutorial program and the change in philosophy that follows from the other changes.

*Student interest, motivation, selection of electives and participation in extra curricular activities.*

In addition to perceptions about changes in academic performance attributed to SBM, we also asked about changes in attitudes, behavior, interest and motivation to study mathematics and science, and whether there had been any increase in enrollment in
elective classes or participation in math or science related extra curricula activities, such as math clubs, science fairs, summer schools and so on.

We also asked about students choices of mathematics or science careers. None of the schools were able to comment on this, however, one high school was in the process of establishing a tracking system to follow former students through higher education and employment, as part of an effort to monitor whether their courses were appropriate or relevant to their students.

In contrast to academic improvements, the majority of schools visited commented on significant improvements in students interest and motivation in math and science that could be attributed to specific programs established as a result of school centered decision making. However, we must again express caution about interpreting middle school findings, as many of the positive changes are attributed to teaming as much as anything.

Sandia has sponsored a Science Fair for the last few years, in which student participation is reported to be excellent. Sandia students have also continued to do well in school-sponsored and regional and national science events. This continued interest in participation in these events is probably related to collaboration with Sandia Laboratories (Sandia Lab scientists teach an advanced level Physics course, Sandia labs supplying the high school with supplies and materials, and teachers volunteering extra time to help students with their projects, and active grant writing efforts by the school to get extra money for supplies and equipment. This is at least indirectly linked to SBM which has resulted in teachers having a greater sense of empowerment and responsibility for student outcomes.

Both math and science teachers reported noticeable changes in student interest and participation in extra curricular activities – participation in the Nature Guide Program and Woodlands project for which 50 students had signed up was thought to be exceptional.

Average attendance has increased from 85% to 90%. This has shown up primarily in grades 9 and 10 (where team teaching has been introduced, but is beginning to show up in later grades (presumably as students who have been team taught progress into those grades). The science teachers report that the number of students in AP science courses has
increased from 22 to 66 over the last two years, which they attribute to greater encouragement from the team approach.

The principal noted that more students are participating in programs sponsored by the university – and they are spending an increased amount of their free time working in the computer lab. He also noted that the recently sponsored math Fair was well supported and exceptionally well attended by students (over 300). More students are being retained in the college prep program.

Principal says enrollment in higher level courses (pre-algebra and algebra) and enrollment in general math has decreased (in the past two years only about 40 students each in grades 7 and 8 takes general math – in previous years it has been more like 100. Principal relates this to the SIP focus on opportunity to learn. Teachers have been “deprogrammed” – and now believe 7th graders are ready to take pre-algebra. Their school cut the red tape – the county guideline says that a student can take pre-algebra if their teacher recommends it, they get a B or better in 6th grade, and score at a specified level on the CRT (county test). The school guideline is that if students meet one of these criteria rather than all three they do pre-algebra. I doubt whether this is an SBM effect – it is not an example of students independently signing up for advanced electives on the basis of better experiences in basic courses). Note suspensions decreased by 56% over the last year, but teachers complain about the disruptive effects of students they can’t get rid of, which makes the school look good in the district’s eyes, but has a negative impact on teachers ability to teach their classes.

As a result of school-initiated programs, including tutoring and teaming, students know that teachers care about them as a person, and they have seen increased participation in all school activities, science fairs, essay writing, math competitions. They claim to have improvements in student’s self esteem as well as academic areas.

The major outcome so far has been increased attendance in general, and in particular for the school within a school, which also has a higher graduation rate. The MDH believes that SBM affects teachers more than students. There have also been increased enrollments in AP courses and in calculus and trigonometry classes – without the summer school classes they would not have been able to build the students up to this level, Math club membership has increased (80 participants). Club sponsors are paid, which was not possible before SBM.

The science department head thought that there had been a positive impact on student’s interest and enjoyment of science because it is now more hands on, not just lecturing. “Students are more involved and using more equipment. They are learning more because
they are doing more.” This has not translated into improved test scores for all students yet, but she feels that it will in future because students now have an opportunity to do things they couldn’t before.

Science teachers interviewed thought that working in teams will impact on performance, but they have no data to back this up (there is no standardized test of science across the district). Class sizes are large, and there are certain lab activities that you cannot do. The quality and participation in the science fair has not really improved — participation is compulsory for advanced students, extra credit for regular students. The Faculty Council has tracked Reading levels as part of the school plan.

The major impact has been on increasing attendance. Discipline has improved, kids know that something will be done, improving behavior outside the classroom has led to better classroom behavior. This more than anything has improved morale.

One of the outcomes of the needs assessment done when developing the SIP at the outset of SBM was that a school—within—a-school was developed for 11th graders who were at risk of not graduating. A team of teachers were assigned to teach a 4-hour block of English, math, science and social studies, which allowed greater flexibility on the part of the teachers. They found that they had a difficult time with the math part of the scheme, because students ranged in ability from Algebra II to basic levels. They ended up transferring this part to regular classes with a support structure. Teachers felt that the students would have been better off in heterogeneous groups with positive role models.

While the program was felt to be successful, it was changed to become an elective course, taught by counsellors teaching life skills rather than academic subjects. The program has continued to evolve and expand, and now has progressed beyond this to an alternative program where students can work part-time at school, part-time at a job, and complete their academic requirements at night school.

Another change was an attempt to make the curriculum for the at-risk students more “hands-on” and focussed on problem solving in real-life settings. Students were required to present an issue to the faculty, and propose a solution (but not necessarily solve it). The school hired some part-time teachers and community volunteers to help with the placement of students while doing their research and so on. The purpose of all these programs were to address their attendance and drop-out problem. The regular school setting was not working for these students, some of whom would be 20 before they would finish in the regular program.

The outcomes of the program was to increase attendance and to bring kids back to school who had dropped out (started with 50 kids, now 75-80). In the first group of students, 15
of 17 graduated. They found that if students last a year in the program, they tend to stay, otherwise they lose them very fast. Much of the success of the program comes from setting higher expectations for students.

Effects on minority students and girls

We also asked the respondents whether there had been any special effects of SBM on minority groups or girls. We found that in the vast majority of schools there were no different outcomes other than as a result of programs specifically targeted for special groups (in a few schools, the enrolment of Hispanic and black students was greater than 85%, so general effects apply across all groups). In a few schools, the number of girls enrolling in math and science elective courses was said to be increasing (although only marginally at best, and we have no hard evidence), and there is no obvious and direct link that this is a direct result of SBM. No differences or effects were noted in 14 schools. Some of these schools do have high minority enrollments, so while there may have been positive outcome for minority students these are not different from other students.

The primary concern of the district at the moment is equity, given the rising number of minority students entering the system. The district has created profiles for each school, looking at the extent to which they provide equity and access. Every school is sent a copy of their profile, and a district profile is aggregated from the results from each school. This is part of the district’s efforts at reducing disproportionate achievement within and between schools. One elementary school used its equity money (see above) to hire mentors – community members who can contribute to the school. Another school has looked at race and attendance issues. The district is making a very concerted effort to evaluate their performance in this area and they are trying to do something about it. School profiling is the primary means by which schools are being made to confront student performance and (through the self studies) to develop improvement plans to address their particular needs.

The school has concentrated on racial and diversity issues as one of its goals for the past year. The school has the highest minority enrollment in the Bellevue district and have made some programmatic changes. For example, they now mainstream all of the special ed. students into the math program. A grading policy has been adopted in which students
either pass or fail, which (according to the Principal) forces teachers to look more closely at students, especially those on the borderline.

PALS has helped all students, but a large proportion of the students in this program are minority. They are responding to positive role models provided by (minority) teachers. Girls in the PALS program have shown improved attitudes because they have been challenged more by requiring them to do pre-algebra.

The school plan addresses specific needs of disadvantaged students, e.g. they plan twice weekly after-school lessons and tutorials and provide transportation for these students; they hold extra SAT preparation classes for disadvantaged students. Staff members are paid using SBM money to hold these courses – these things were requested by students, parents and faculty. Through ESL and tutorials, students find it more comfortable in asking questions and enrolling in higher courses. There was an increase in SAT scores for students in the tutorial program. Teaming allows so much more to be done, to work much faster than before, to group heterogeneously, allow students to work together, and increase self-esteem, kids know each other better. A lot of girls take upper math, but they cannot attribute anything to SBM. A plan for increasing performance of female students in AP physics has been included in the budget.

No direct effects on outcomes for students seem likely just because a school makes its own decisions or as a general result from improvements in teaching methods in the short term. If schools want to address equity issues, then they are most likely to as a result when programs are planned, monitored and resources applied specifically for the purpose. Changes don’t have to be revolutionary or costly to be worthwhile. When all else is basically sound, changes in attitude and motivation can be just as effective for increasing achievement as other direct interventions in math and science curriculum.

Recommendations

District level

- greater attention needs to be given at all levels to articulating the ultimate impacts of SBM on students – that is what it is supposed to be for. That SBM is a tool rather than an end in itself needs to be reinforced.
- district evaluation and monitoring plans should have a greater focus on outcomes (within the limits of reasonableness), and should be
communicated to all. Specific examples of successful practices should be circulated.

**School level**

- Each school should develop a vision of what the school wants to accomplish and should communicate this to the whole school community. The vision should address the objectives of the decentralisation approach and its ultimate purpose of improving student interest and achievement.

- Schools should get the decision-making structure sorted out quickly, and make sure everyone is focussed on the why they are doing things rather than how.

- All planning and budgeting activities should be coordinated (as do the committees working on them) and must be explicitly focused on how they will impact on students. Many schools have school goals, mission statements, School Improvement Plans, SBM councils, advisory councils, budget committees, but often these were isolated from each other. In schools which had achieved significant outcomes, the relationship between these things was explicit, documented and communicated to faculty.

- Evaluation of initiatives should be built into the process and should be an explicit part of the site council’s role.
Chapter 5

Evaluating Outcomes for Customers in School Education

Evaluating Outcomes for Customers in School Education

The importance of evaluation of the success of customer service efforts has received increasing recognition over the past few years. Customer service measurement efforts are widespread in the business community (Jones and Sasser, 1995) and in many public sector agencies. Yet in school education, these activities have proven problematical when it comes to the core business of schools – serving the needs of students and their parents. This paper briefly examines the approaches to evaluation taken within the NSW Department of School Education in terms of their ability to deliver the kinds of information necessary to lead to substantial improvements in the outcomes for clients. It argues that the heart of the problem lies as much with the inability of school systems to clearly articulate who their customers really are (leading to a school culture dominated by provider values) as it is with the evaluation methodologies used.

Introduction

The NSW Government has, since 1995, pursued an objective of creating a public sector which is world class in its delivery of value for money and quality services to its customers and the community. The initiative builds on previous efforts to develop a greater customer focus in public sector agencies and deliver improved services for customers through the Quality Customer Service, Guarantee of Service and Frontline Complaints Handling projects. While agencies will be expected to continue to achieve measurable improvements in outcomes for customers, they will also be required to demonstrate that they are putting in place strategies aimed at measurably improving the quality of their performance across all dimensions of their operations.
The notion that public sector managers should identify their clients and evaluate the effectiveness of the delivery of services to them has been widely embraced by the Australian public sector in recent years, arguably to a greater extent than most other countries. But despite the efforts of central agencies\textsuperscript{26}, a burgeoning academic literature and a thriving professional body\textsuperscript{27}, there is much less agreement as to how such evaluations should be conducted, for what purposes, and how they should be reported. Nor is there common agreement that the various performance evaluation regimes instituted by governments and central agencies have actually led to the improvement of the delivery of services to public sector clients. Guthrie (1994) notes:

Where trust was once placed in the service provided by public service professionals, faith is now placed in the evaluation of those services. However, in some cases, the promises and claims made for the "new" management techniques have continually not been realised in practice.

This paper briefly reviews the effectiveness of the various approaches to evaluation of service provision in the NSW Department of School Education in recent years, particularly as it applies to schooling outcomes, and discusses some areas for further improvement.

The discrepancy between rhetoric and reality of performance evaluation is clearly evident in the recent history of evaluation practices in NSW school education sector. The debate about the most appropriate means for evaluating educational performance touches on issues including the conceptualisation of professionalism, the ownership of knowledge, the legitimate role of government, the community and parental involvement in educational decision-making, and issues of power and control in these relationships. The arguments are seldom clearly articulated, but find expression in the

\textsuperscript{26} See for example, Department of Finance (1989) \textit{Program Evaluation: A Guide for Managers}. Department of Finance, Canberra and the Program Evaluation Bulletin series of the NSW Public Service Board.

\textsuperscript{27} The Australasian Evaluation Society has almost 1,000 members.
decisions of the NSW Teachers’ Federation to oppose the Australian Studies in School Performance studies in the 1970s (Bourke, 1981), their opposition to the introduction of Basic Skills Testing in the late 1980s and the banning of Quality Assurance school reviews (Cavanagh, 1995).

Many of the leading evaluation theorists, such as Tyler (1950), Scriven (1972), Stake (1980), House (1980), Guba (1978), Eisner (1991) and others have written primarily in the educational context, and some claim that the development of evaluation mechanisms in education systems have been well ahead of those in other government enterprises and business. Paradoxically, while knowledge about evaluation may be high in the education profession as a whole, the practice of evaluation is identified as a major weakness in school planning and management in several reports (see for example Eltis, 1995; QA Review Report, 1994).

A difficulty faced by managers of the public school system is identifying who the “clients” of education are, and the nature of the services offered to them. There has been a general reluctance over the years for the education system to articulate a clear answer to the simple question of “who are we here for”. The reality is that public school systems must serve multiple clients simultaneously, as reflected in the broad ranging mission contained in the Department’s Strategic Plan:

“Our mission is to educate the public school students of NSW for the benefit of each individual, the community and the nation.”

All levels of the system share some clients in common, but the primary client may vary from level to level. At the broadest (systemic) level, since investment in education is universally considered a “public good”, in the sense that it is expected to contribute to national development and prosperity as well as materially benefitting the individual recipients (Hanushek, 1986), the primary client for school systems (and thus the public servants accountable for the performance of the system) is the wider
community at large, represented by the elected government of the day and Minister. At the same time, the individual schools, teachers and other staff employed by public school systems are also clients of the system in that they are recipients of centrally determined and provided policies, programs, resources and corporate services. Within the bureaucracy itself, various units, branches and directorates may have a range of internal as well as external clients. Most education bureaucrats of today would also acknowledge that parents and students are also their clients, even if they do not often deal with them face-to-face.

While the clients of the service delivery level of education, (ie schools and classrooms) may seem obvious, teachers also express concerns about determining priorities in deciding whose needs they are to satisfy. Is their first responsibility to the students, they ask, or is it the students’ parents for whom they act in loco parentis? Teachers and school communities are also keenly aware that employers and other community members – the ultimate consumers of their “product” (that is, an educated student) – have their own sets of needs and demands that are made vociferously known, especially when the “product” is perceived to be of a lower standard than expected or desired.

Overlaid on this need to be all things to all people has been a failure to define exactly what services are implied by the term “to educate”, especially in terms of the outcomes that might be expected. While school systems have long been able to describe in detail their inputs into the education process (for example in terms of pupil/teacher ratios, per pupil expenditure, number of library books per head), they have not been able, or willing, to describe (much less guarantee) the outcomes expected given these inputs. The legislation and syllabus documents concerning public schooling for most of

---

28 In this sense, school education is no different from other public sector agencies. This notion of responsibility to the citizenry is perhaps more clearly demonstrated in the public sector arrangements in New Zealand, where the Minister of the Crown, representing the people, contracts with individual departments for specified outputs.
this century have been couched in terms of what will be given, (ie, inputs) rather than what will be achieved (ie, outcomes). This orientation has contributed to the development of deeply entrenched attitudes and traditions which have seen schools become provider cultures, whose services are provided on a "take it or leave it" basis. Students are compelled to accept the standard curriculum on offer, sit common examinations, attain common credentials, dress similarly, and behave similarly. The clear message is "conform to our expectations, or go elsewhere". The end result of this tradition is that some teachers think that they know the needs of their clients better than their clients do themselves. It has also contributed to a culture, particularly prevalent in senior secondary years, in which many teachers see their role to be the arbiters of "standards", and in which student outcomes result from the student's native ability and effort, rather than as a result of their teaching. Still others believe that student outcomes are nothing more than the product of their social circumstances (Cavanagh, 1995), again reflecting a view that implicitly accepts that teaching has no impact on learning.

These attitudes and traditions have significant implications for understandings the need for evaluation and the forms such evaluations might take, and it is in this context that the effectiveness of evaluative approaches can be considered.

A further problem in measuring educational performance is defining what "performance in education" means. Even if the sole purpose of an education system were to teach children a finite set of facts and skills, there are problems with attempting to isolate the contribution of the education system by testing children when they come to the end of their schooling. Scores that children obtain do not necessarily reflect the contribution of the education system alone. What students know and can do reflects many other influences, including the home, the church, peer groups, the media and so on. More importantly, children do not begin school on equal terms, and do not enjoy equal assistance during their school years. If no other data are used apart from test scores, there is no way of knowing how much their successes and failures reflect the input of the education system and how much they reflect other factors.
Approaches to evaluation

While there has been a long history of evaluation and review activities conducted by the NSW Department of Education (the Knibbs–Turner Report of 1904 a well known early example), the particular strategies and structures established systemically follow closely the history of the development of evaluation methodology as a whole (see Owen, 1993 for an overview). In the 1960s and early 1970s, for example, most studies followed the classical experimental design paradigm and were conducted from a research perspective. Program evaluation as a separate discipline emerged in NSW in the late 1970s and in the early 1980s. As in the United States, the major impetus for the further development of evaluation came from the need to formally demonstrate the worth of Federally funded programs. A specialist evaluation unit was established within the DSE in 1985, with a staff of up to a dozen officers conducting and advising on a range of evaluation projects (such as the Participation and Equity Program). This unit was well placed to respond to the NSW Premiers Department initiative to institutionalise the requirement for government departments to formally evaluate their entire operation over a five year period and produce an annual five year rolling plan for program evaluations (See Allan, 1987). Evaluation teams, usually led by Inspectors of Schools and containing officers with expertise in evaluation methods were established to conduct reviews of specific programs. A Program Evaluation Coordinating Committee was established, chaired by the Deputy Director–General, to ensure that all current evaluative and review activities were well planned and met management requirements. Given the breadth of the Department’s activities, a variety of structures and processes to assess progress within programs was necessary, including reviews of curriculum, management reviews, internal audits, evaluations of school effectiveness, assessment of teachers, and monitoring of schools' management and organisational procedures by the Inspectorate.

The Department of Education published its first policy on program evaluation and formalised its expectations for schools to evaluate their programs in 1987. In the
five year period 1985–1990, nearly 100 programs had been subjected to intensive evaluation. Despite this impressive record of activity, it is difficult to find evidence that these evaluations contributed anything that was able to substantially improve the quality of outcomes achieved by students. This finding casts no aspersions against those involved in the evaluation process or the quality of the work done, but reflects wider problems with the models of evaluation prevalent at the time. Few of these evaluation reports were ever made public, decisions to abolish programs were made before many evaluations were completed, even fewer were able to connect programmatic causes with quantifiable effects, or were able to comment on the cost effectiveness of programs. In the aftermath of the Management Review of the Education portfolio (Scott, 1989) the Program Evaluation Unit was disbanded, noting that “the system has developed very few indicators for the public to judge its performance” (p.4).²⁹

The disillusionment with traditional program evaluation and its inability to reflect organisational performance was not confined to NSW. There is an extensive literature on the utility of evaluation (or lack of it) (see for example Cousins and Leithwood, 1986; Scriven, 1993). Writers such as Weiss (1972) and Chemlinski (1985) identified under-utilisation as one of the foremost problems in evaluation research.

Evaluation research is meant for immediate and direct use in improving the quality of social programming. Yet a review of evaluation experience suggests that evaluation results have not exerted significant influence on program decisions (Weiss, 1972).

Mangano (1992) also notes that too often, evaluators failed to target issues facing decision makers, were slow in responding when they did, and were ineffective in communicating their results. What was offered to decision makers was not enough of what they wanted or valued. Decision makers sought information they could use right

²⁹ A similar conclusion was reached by the Quality in Education Review Committee report (1985); that there was simply no evidence that would allow them to state that education standards had fallen or risen.
away, but evaluators, with some notable exceptions, continued to focus exclusively on long-term analysis.

School inspection

Other means of evaluating the effectiveness of educational services have also been tried, and found wanting in various respects. The Inspectorate provided the system's primary means of quality control for more than a century until disbanded in the early 1990s. The problems with school inspection as a means for evaluating the performance of the Inner London Education Authority (ILEA) noted by Hargreaves, (1988) also apply to inspection in NSW.

We had the tradition going back for most of the century of full inspections, which gave a very detailed report of the work of the school. . . If it was a favourable inspection report you would find that principals had them at the entrance to the school, available for any casual visitor. If the report was not so favourable they are usually kept in a cupboard in the back of the school. . . The chances of a school being inspected under the full inspection model are about 1 in 50 in each year, hardly enough to strike terror or even apprehension in any principal. A second disadvantage was that this model is not as helpful for developmental purposes as we would have wished. It took about a year to get the full school inspection report printed and made public, by which time conditions in the school had most likely changed. Of course we gave the schools oral feedback as soon as the inspection was completed but the gap of a year to complete the written documents was too great. There is a third problem, and a very important one. This system of sampling does not identify your weakest schools, the schools in difficulty, the schools that are being poorly led, or the schools where there is serious under achievement. If you are taking monitoring seriously, samples will not do. You need to identify every school that needs help. This kind of approach using full inspections, though it has some value was not cost effective. . . (Hargreaves, 1988 p.127).

In addition to these problems, the functions of the Inspectorate as an evaluative body in NSW were subverted by their increased role in personnel appraisals and low-level administrative tasks.
School self-evaluation

This lack of external scrutiny contributed, in part, to the growing popularity of self-evaluation at the school level, which was supported by several pilot programs (Osborne, 1989) and resource kits. Many of these self-evaluation projects were based on the action research model popularised by Kemmis and others at Deakin University (see for example Carr and Kemmis, 1983). While school self-evaluation never attained the same degree of formality in NSW as it did in Victoria (Marshall, 1988) or overseas systems, such as the Inner London Education Authority (Simons, 1985), few principals or teachers could have avoided at least some exposure to the concept of evaluation. Despite the groundswell of interest in school self-evaluation from the mid-seventies, it has failed to gain a more permanent base in education systems.

School self-evaluation served several purposes and took a variety of forms. The basic motivation lay in persuading schools to be more accountable and/or more managerially efficient. Early forms of school self-evaluation comprised comprehensive checklists of questions for schools to keep themselves under review. The deficiencies of this approach soon became apparent and local authorities turned to more complex models.

The research literature on school self-evaluation indicates that it has proved to be ineffective (Simons, 1992). The checklist approach has been criticised for being too prescriptive in intent, piecemeal, unconnected to curriculum development and failing to embody a strategy of change which involved the whole school, and the externally imposed schemes failed to match the value systems of school staff (for a more extensive critique see Shipman, 1983). Hargreaves found that schools tended to write long, heavily descriptive and defensive accounts of what they had been doing. Teachers were not good at analysing what they were doing. School self-evaluations were extremely weak at

---

examining what went on in classrooms. They were very weak at analysing pupil performance and were not very good at applying the evaluation results to forward planning.

In NSW, the void created by the disbanding of the existing structures, created the possibility of a fundamentally different way in which evaluation could be conceived and practiced. The view emerged that evaluation was the responsibility of every manager (whether at school or system level) rather than the preserve of the initiated few, and the expectation arose that evaluation should be more firmly embedded and an integral part of the planning and budgeting cycle rather than an "add-on" at the end.

**Performance indicators**

During the mid to late 1980s the concept of performance indicators provided an approach to information management and reporting which captured a great deal of attention in education systems both in Australia and overseas (see Wyatt, 1992 and 1994 for a comprehensive discussion of this movement). The approach offered several advantages:

- Monitoring indicators takes less time than one-off, in-depth program evaluations. Furthermore, once developed, future indicator use is relatively inexpensive.
- Measuring indicators over appropriate time intervals allows an evolving picture to be obtained.
- It encourages managers to focus on outcomes.
- The development and monitoring of indicators can and should be conducted by program managers rather than specialist staff (see Lenne, 1986).

Indicator systems have several features that could allow them to contribute positively to comprehensive school improvement efforts, and indeed have been a driving force in many examples of systemic reform (eg Fetler, 1986). They can provide benchmarks for comparing educational progress over time and place. They can help to
describe changes in key features of the education system and direct attention to present or potential problems and inform policy decisions (Oakes, 1986).

However, the implementation of performance indicator systems as tools to inform evaluation has been largely unsuccessful to date (Winston, 1994). This is not to say that there has not been a great deal of activity in this area, but most school systems have employed a limited range of easily measured variables, and tend to focus excessively on test scores rather than broader measures of education outcomes. Most do not provide any coherent model of a school’s functioning, and so are of limited value for either describing schools or diagnosing their problems. Finally, although some of the more successful accountability systems have been designed and operated at the local school level (see Wyatt, 1992), those managed at the state level are unlikely to promote the kind of systematic, cooperative effort among teachers and principals that is needed to effect significant school change because they are often perceived as imposing irrelevant, additional demands on schools (David, 1987).

Within the NSW Department of School Education, most planning documents, whether at school or system level contain a column headed “performance indicators”. But on closer examination, while the name is present, what are proposed as indicators rarely allow judgement about achievements, and rarer still contribute a dynamic monitoring and decision-making process.

Many proposed education indicators take the form of measures of short–term outputs, such as whether a particular document has been produced or not. These are not good performance indicators for several reasons – they say nothing about the quality of what was done, and also say nothing about what happened as a result of producing the document. It is arguable whether they even fit the basic definition of what an indicator is, in that it does not provide comparative information. For a school executive to know and be able to report whether a document has been produced or not may be important information, but it is an indication of whether a milestone has been reached, not an indication of outcome.
If it is any consolation to educators, this confusion between outputs and outcomes as measures of performance appears to be widespread throughout the public sector. The Department of Finance (1989: 29) noted that:

"...many published indicators are of workload. Workload indicators are about levels of input or activities, not performance."

None of these reasons are sufficient in themselves to suggest that performance indicators have no place to play in the strategic management of education systems. However, it is also obvious that neither the currently available forms of indicators or the uses to which they are currently put provide decision makers with the information needed. The best solution, as Porter (1993) suggests, is likely to be a mix of indicators and context sensitive information about school processes.

**Conclusion**

What is needed is a mechanism that brings together the best features of each of the earlier models in an integrated way, and at the same time avoiding their known pitfalls. The abolition of the Program Evaluation Unit in the latest departmental restructuring represents both a great loss of expertise and a new opportunity to re-think the department’s approach to evaluation.

For a time it appeared that the processes for school review established under the Quality Assurance program would meet these needs (see for example Cuttance, 1993 and 1994; Hightett, 1994; Wasson, 1995). The school review process aimed to assess the effectiveness of practices and processes for achieving improved student outcomes in schools. They were intended to contribute to the process through which schools as learning organisations develop a planned approach to constructing their future. Over a four year cycle, the school reviews were a collaborative enterprise, in which review teams worked with school communities to assess:
how the school goes about the task of meeting the community’s needs for education, in the context of addressing the statewide priorities for student outcomes;

- what the school is attempting to achieve for its community;
- how the school knows it is achieving what it set out to do; and
- how the school responds to its achievements.

The Quality Assurance process too, has passed by the wayside, largely as a result of industrial action by the NSW Teachers Federation, but also because it failed, like the ILEA inspection model described above, to identify schools which were underperforming on a timely basis. It can be argued that the worst schools managed to avoid any external accountability by refusing to submit to school reviews until the end, by which time the process had been killed through industrial action. It failed also because there was no commitment achieved within the system as a whole to the notion that there are indeed standards of performance for which schools should be accountable.

Future Directions

For the outcomes of customers of education systems to be better evaluated, there needs to be clearer identification of what outcomes are to be achieved for which clients from which services, and appropriate standards for service delivery established for each. System-wide and systematic measurement schemes need to be put in place for all of the department’s programs, with an appropriate reporting structure also capable of informing improvement strategies at the program level. Procedures for true performance reporting against the department’s corporate plan, and against each school’s plan need also to be established.

All of these activities need to take place within an enhanced management culture, in which evaluation is seen as an integral and essential part of each manager’s job (and school principals are included in this category), and not as an imposed, feared and unwanted “add-on”. To do so will require ongoing professional development in evaluation theory and practice for some managers, so that they will at least be aware of
when they might need to “buy-in” appropriate resources (including expertise) to ensure that their evaluation efforts are of an acceptable standard. None of the above can happen unless there is a clear understanding of the requirement for cultural change, and a commitment from top management to make it happen.

Whether the focus of evaluation efforts is on accountability or program improvement, the evaluation strategies must ensure that the results are available when key decisions are to be made. There is no point in having highly polished voluminous reports that miss key reporting deadlines. Similarly, the evaluation methodologies chosen need to be such that decision-makers have confidence in the results, and indeed, have methodologies that can provide information that decision-makers can use (Robinson, 1993).

The NSW public school sector has until recently been largely immune to the pressure on other government agencies for greater performance reporting. However, changes in government thinking and parental expectations may well force a rapid end to this position. The challenge for the department to implement an effective service evaluation model will assume considerable importance. The experience of the past efforts will provide many valuable lessons for the future, for those who care to heed them.
Bibliography


NSW Department of Education (1986). Program Performance Evaluation in the NSW. Sydney: Department of Education


Chapter 6

Informing School Improvement Efforts: The Legacy of Quality Assurance Reviews

Unpublished paper presented at the University of Western Sydney Postgraduate Student Conference, Werrington, October, 1996.
Informing School Improvement Efforts – The Legacy of Quality Assurance

Introduction

The Quality Assurance (QA) school review program that operated in NSW between 1993 and 1995 has been extensively described in the literature (see Cuttance 1993a; 1993b; 1995: Hightett, 1994) critically evaluated (Henry, 1993; Hopkins, 1995) and in some quarters, enthusiastically acclaimed. It attracted the attention of educators throughout the South-East Asia and Pacific region, the United Kingdom and North America. The program hosted visits by officials from Ministries and Departments of Education of many nations, all of whom were seeking solutions to the same needs addressed by the QA school review program – the requirement for school systems to simultaneously provide mechanisms for school development and accountability.

The Quality Assurance approach gave effect to what Boston (1992) saw as a key challenge facing the school system, to make sure, be sure and to give assurance to others about the quality of our work in education. Quality Assurance was intended to be an integral part of the work of all within the education system. As for all successful enterprises, the application of a systematic approach to continuous improvement must become fundamental to all aspects of the school system. This approach of attempting to build in quality differed in both philosophy and practice to previous models of quality control in education which were based on inspecting-out inferior outcomes after they had been produced (Cuttance, 1993a). It attempted to put systems in place which sought to prevent performance problems in the first place.
It is not necessary to go into detail in this paper about the reasons why the school review program was abolished. It is sufficient to note that the underlying issues are complex, encompassing interpersonal relationships, crass political manoeuvring, logistical and management difficulties and problems with the methodology underpinning the review process itself. A brief description of the program is, however, necessary to understand the context for the work discussed below.

**The context for Quality Assurance**

Quality assurance school reviews aimed to assess the effectiveness of practices and processes for achieving improved student outcomes in schools, and to contribute to the process through which schools planned for their future.

School reviews were intended to be a collaborative enterprise, in which review teams worked with school communities to assess:

- how the school went about the task of meeting the community's needs for education, in the context of statewide priorities;
- what the school was attempting to achieve for its community;
- how the school knew it was achieving what it set out to do;
- how the school responded to what it knew about its achievements.

The quality assurance process was not based on the assumption that there was something wrong or deficient in the operation of schools. Rather, it reflected the need to provide a clear framework for public accountability and provide the information about current performance required as a basis for improving performance (Cuttance, 1993).

**The School Review Process**

Quality Assurance reviews of individual schools focused on the way in which the school was providing for its community's education needs. They were concerned directly with student outcomes and the quality systems in place to bring about their
improvement. The reviews focussed on issues such as effective management and leadership, learning and teaching, staff training and development, student welfare, community involvement and participation within the perspectives and parameters of statewide priorities and resources. In the review process, staff, students and members of the school community work with a review team to assess the current achievements of the school and the initiatives that would, through the planning and development process, contribute most to future achievements.

The primary focus areas for each review were negotiated with the school community following the analyses of school planning documents, student outcomes data and school self-assessments against the framework for outcome statements for the review of educational practice and school functioning. These focus areas were agreed as the initial understanding of the key aspects of the current performance of the school and of the most important factors in its continuing development. As the review developed, additional foci that emerged from the input of parents, students and staff were either integrated into the previously agreed foci or addressed as a separate part of the review.

School reviews were conducted by teams whose size and composition vary according to the enrolment of the school. The team leader was responsible for setting up the review, managing it during the period in the school and producing the report of the findings and recommendations. The host school principal was involved throughout the review, and participated in the presentation of the findings. The membership of most review teams comprised principals, school executive staff or teachers from other schools, Directors of Schools and parents. The school principal and Director of Schools were accountable for ensuring that the recommendations were implemented appropriately following the review.
Methodology of the Reviews

The specific methodology used in any particular review varied according to the context of the school in question. In general, however, information was collected and analysed using what can loosely be described as qualitative inquiry methods. Review teams gathered their information through structured meetings and interviews with a cross-section of students, school staff (both teaching and non-teaching) and parents in the local community. Processes of random selection were used to ensure that a cross-section of each of these groups were involved in the review. In addition, the reviews also analysed student outcomes data and documents, such as school plans and budgets.

Quality assurance was concerned with recognising and building on strengths and achievements. Importantly, it was not a process for simply identifying problems or weaknesses in schools and providing advice on their remediation. One of the most valuable functions that school reviews performed was to provide external validation of school achievements, and an avenue for public recognition of these successes. School reviews provided an opportunity for sharing information about particularly effective practices both within the local school community and across the school system.

School reviews also served an accountability function by providing the community and the Department of School Education with public information about the success of programs and initiatives of schools. Further, they placed on record the key issues that individual schools needed to address in revising their plans for future development (Cuttance, 1995).

Reports were written on the basis of the findings of the review for each individual school reviewed. These reports are public documents that acknowledge the strengths and achievements of the school and provide information and recommendations for the school's future development.
QA school reviews were only part of the entire QA strategy, although the constituted the largest part of the “public face” of the program. Less visible, but equally important were the program evaluation, systemic reporting (Wyatt, 1994), best practice identification (Carroll, 1994) and student outcomes analysis work conducted under the aegis of the program. While the school review program has now been abandoned, it is these less visible activities that have endured, and that provide the opportunity to bring about lasting change in the effectiveness with which schools and the school system deliver outcomes for the students in their care.

The school review program has left behind it some important legacies. The reviews themselves, made thousands of recommendations for the improvement of individual schools, most of which have been implemented to a greater or lesser extent. These ranged from relatively minor recommendations, such as improving home-school communication by introducing a weekly newsletter, to those suggesting a fundamental reappraisal of the school’s mission. In terms of its success in “building in quality”, the school review process provided a model taken up by many schools for gaining self-knowledge about their processes and practices. Not least of the legacies of QA was the recognition by many teachers and principals that parents, and students, have legitimate views on the quality of schooling they are offered (and that even very young students can articulate these views sensibly and maturely) that should be regularly ought in planning and evaluating school programs. In such schools, a “mini-QA” process is conducted internally, as part of the school’s annual planning cycle (n.b., in the original QA model, schools were to be reviewed on a four–yearly cycle).

The review process itself has also directly contributed to the professional development of a large cadre of teachers and principals, through exposure to a body of knowledge about quality principles, evaluation methodology, the development of skills in critical analysis and writing, and observation teaching and management practices in a range of schools. This experience, at a minimum, provides former QA team members
with the opportunity to reflect on their own practices, and hopefully translate the “best practices” and principles they have observed back into their own schools.

Also of importance was the greater focus on student outcomes the QA methodology developed in its latter days (although this was also part of its undoing), which contributed to the development of a greater awareness in some schools that these outcomes provide a more appropriate basis for strategic planning than administrative rules, current work practices and expected levels of resource inputs.

But more tangibly, the school review reports themselves collectively represent a database about the operation of NSW schools unparalleled in its breadth and richness of detail. The database contains reviews from over 2000 schools and reflects the views of tens of thousands of students, teachers and parents. As the review reports are public documents, the database is potentially available to anyone who has the wherewithal to analyse them.

In a recent address, the President of the NSW Board of Studies lamented that:

“...our research and knowledge base of what happens in NSW schools and classrooms is hopelessly inadequate. I would give a great deal to have access to the sort of data that has been collected in recent years from British and Scottish schools by Her majesty’s Inspectors and the results of the National Curriculum Assessment programs” (Weller, 1996, p.2).

The QA review reports provide an opportunity to greatly enhance this knowledge base, and at very little expense. This paper describes a process and a tool developed for extracting this information from the review reports, illustrated by an example of the kind of output that can be achieved.

The potential offered by the school review reports to inform systemic strategic planning was demonstrated by the publication of the first QA Review Report which synthesised the findings of the first 400 school reviews (see Wyatt, 1994). To be a learning organisation (Senge, 1994) an institution must have an efficient and effective
system which allows it to have knowledge about its own performance. The depth and complexity of the information—from—the—classroom contained in the school review reports provides the most comprehensive information the NSW school system has ever been able to draw upon in its quest to become a learning organisation.

Haertel et al. (1989) note that:

"In order to diagnose problems in a school's functioning and at the same time suggest possible avenues and approaches for improvement, a comprehensive evaluation must examine the school's instructional processes, as well as its formal and informal administrative and decision-making processes. It must elicit the views of the school's participants concerning problems and solutions, as well as inferring problems from evidence of poor learning outcomes, negative parent or community sentiment or other sources". (p2)

The same conditions hold for the evaluation of school systems. The school review reports, taken as a whole, provide valid, reliable, high quality, comprehensive and synthesised information about school operation in areas deemed most pertinent by the schools. It is information grounded in the field, in a form that retains much of the field richness but is available for aggregation. When aggregated it has the capacity to provide a picture of systemic operation that can inform decision making about issues of concern to educators at all levels of the education system.

In an era of greater public accountability for school outcomes, the analysis of issues which lie at the heart of school operations is both timely and essential. Improvements in schools and the school system can be brought about where available information, common sense, and practical knowledge are used to identify and define problems and then select courses of action (Levine, Windham & Bathroy, 1990). When aggregated on a statewide basis, the findings and recommendations of the reviews reveal trends and issues emerging in the management and operation of schools which have the potential to significantly influence policy formulation and resource allocation.
While qualitative methods have been an accepted tool in educational research since the late 1970s (e.g., Guba, 1978), practical guidance on the mechanics of qualitative data analysis (as opposed to data collection) is comparatively thin in the literature. The potential for "fuller appreciation of the phenomena of interest" (Eisner, 1979) offered by qualitative methods must be balanced by the need for studies to be conducted in a disciplined manner which can withstand external scrutiny. Any information offered to decision-makers which proposes directions for action requiring significant resource shifts must have a high level of credibility in terms of methodological rigour if it is to have any chance of acceptance.

The intuition of the individual analyst is sometimes described as the richest and primary source of subjective understanding in qualitative research, as the analysts immerse themselves in the rich and thick data resulting from the transcription of field notes, interviews and records of observations. The weakness of the intuitive approach is that because intuition is such a private process, it is difficult to expose the results to external scrutiny or confirmation. In the construction of meaning from information provided during school reviews, a collaborative approach is used to challenge the intuitive interpretations. This approach combines views from both within and outside the specific school culture and represents both a professional and lay interpretation. As such, the reports themselves have credibility, providing confidence that later analysis is based on valid and tested information.

In the analysis of the school review reports described in this paper a more procedural approach, along the lines of that offered by content analysis (see for example Berg, 1989) was used, in which a hierarchy of significance is constructed based on the frequency with which an issue occurs in the data. Content analysis itself is sometimes criticised as approaching the positivistic (the antithesis of the naturalistic philosophy), as it requires pre-determining the concerns or hypotheses that are to be tested against the data. Alternatively, grounded theory (Glasser & Strauss, 1967) provides a strategy for generating hypotheses from systematically acquired data. One element of the
grounded theory strategy – constant comparative analysis – involves ongoing, systematic organisation and classification of the data into various categories. When categories of data or themes begin to emerge, other data that fit the category and other related sub-themes can then be discerned.

One of the main disadvantages of using qualitative data as a means of influencing policy making is the time-consuming and labour intensive effort required to make sense of the volume of material usually collected. In the case of the QA school review reports, over 30,000 pages of text are involved in the analysis. The analysis of the QA review reports has been simplified by the development of a purpose designed software and a system for pre-coding the review findings and recommendations (see Manefield & Wyatt, 1995).

Coding of Review Reports

In the first year of operation, the analysis of the review reports was conducted by a small team of officers, which, while producing a high quality outcome, was labour intensive and only just within the bounds of manageability within the timescale for reporting. As the volume of data grew, the use of commercial software packages designed to assist analysis of qualitative information was trialled. These commercially available packages proved to be of limited value. Because of the specific nature of the material, including the idiosyncrasies in the writing styles of the various authors of the review reports which meant that similar information was often expressed in different ways, and conversely, similar language was used to describe very different things, it was decided that informed analysts were needed to read and interpret each report. These analysts code each paragraph for entry into a purpose designed software package. A group of casual and retired teachers, all experienced in the NSW system, was recruited and trained for this task.

So that information can be accessed in a flexible manner, an electronic version of the review findings and recommendations for each report is prepared and linked through
the school code to the school type (for example secondary or primary, large or small), school characteristics (for example whether the school is selective), the location of the school (for example the school district), the team leader for the review, and the term in which the review was conducted. This electronic file is then linked to the coding information provided by each report analyst.

The first stage in the procedure is to mark each paragraph with a set of structure codes representing the level of text within each report. This range hierarchically from the focus area, to heading (ie school profile, strengths and achievements, current situation, findings, future directions and recommendations), sub and sub-sub headings to text per se. These codes are entered into the electronic file of each report.

The focus areas for review are selected as one primary unit of analysis. Eleven groups of focus areas account for the majority of focus areas encountered in the reports. An “other issues” category to group miscellaneous focus areas has proved sufficiently comprehensive to cover all salient areas of current school operation. The categories are:

- student welfare;
- school community involvement and participation;
- communication and school image;
- training and development and staff welfare;
- policy decision making;
- operational management and planning;
- learning and teaching;
- assessment and reporting;
- special education and integration;
- curriculum; and
- other issues.

Each paragraph is allocated its focus area category. Themes identified within each focus area are loosely and arbitrarily grouped into those related to schools, system perspectives and initiatives, and curriculum areas. There are thirty two themes under the schools grouping, twenty three under the systemic category and twenty two related to
curriculum. Schools themes include student learning processes, school organisation of classes, student learning outcomes, and equity issues. Systemic themes include whole school development, support staff, isolation, public image, outcomes based education and sport. The curriculum themes relate to the primary and secondary key learning areas as well as such areas as vocational education, literacy and extra curriculum activities. Each paragraph is allocated to a theme or, if appropriate to two or more themes. A compromise has to be reached for some conglomerate paragraphs, realising that in the total database of perhaps thirty thousand pages, secondary mentions of themes are not of sufficient significance to warrant further coding.

**Analysis of the Information**

As alluded to above, the database where this coding is linked to the electronic copy of the report allows analysis in several ways. Data can be searched for any of the focus areas, or themes emerging from the reports. It can be searched for any of the dimensions or aspects of the review frameworks. In addition, the database has a key word search facility supported by a software based procedure which enhances the speed of the search.

As the coding of reports continues, key words are allocated to various themes. This makes coding easier and helps the user of the database to delimit the search requirements. The key word searching can be undertaken at any level of the database from the whole to parts. For example, within the secondary school category, in the findings section, in the focus area of student welfare, under the theme of discipline systems, the word “levels” could be entered. Entering the same word at within another focus area might reveal quite different material. Entering it at a broader level might reveal yet more information and a different definition and use of the word.

Output from the database includes tabulations of frequency of occurrence. For example, within a focus area, the frequency of occurrence of the array of themes can be found. While these are generated by term within the categories specified by the search, a
cut-off of frequencies can be entered so that strings of “one-off” themes can be excluded, and a table summarising the most frequently occurring themes within the focus areas can be generated. This automatically provides the most important (if frequency is the surrogate for important) topics within the search, while also providing an indication of the trend in frequency.

The topics can then be entered as the search parameters. The database is able to provide an electronic or printed output of each of the paragraphs from the review reports that have been coded into the chosen search category. Further synthesis is then possible by reading the paragraphs for recurring sub themes and issues which can then be drawn together in systemic reports.

The following example of an analysis of one issue demonstrates the power of this database to provide information that is relevant to the system and to individual schools. Extracts from the original reports (in subset text) are used to illustrate major themes.

**The Relationship between Curriculum, Teaching, Learning, Assessment and Reporting – Findings from School Reviews**

Assessment plays a key role in the teaching and learning process. It is not surprising therefore, that more than three-quarters of the school review reports to date make some mention of assessment or the closely related themes of testing, reporting and student outcomes.

The most frequent observation made by review reports is that teachers and schools use a large variety of methods of assessment at different times and for different purposes. Teachers variously use both qualitative and quantitative assessment strategies, including commercially produced and teacher made tests. The range of practices in place in schools including the following:

- observations of individual students
- annotated work samples
- checklists
- teacher devised tests
- reading interest inventories
- teachers listening to children read
- student learning logs.

Some innovative methods of charting student progress were mentioned in a small number of reports, for example, the use of photographic records and tape recordings of observed student learning behaviour to further enhance the individualised student-centred approach to teaching in evidence throughout the school.

Review reports describe some examples of good practice in assessment and reporting.

The review found that students learn in an environment where their learning needs are clearly identified, where programs are developed and implemented to meet these needs, and when there is close monitoring to support students in achieving improved learning outcomes.

Information is maintained in individual student folders and is used by teachers to determine individual progress and meet individual learning needs. Student folders are passed on to the student's next teacher to support continuity of instruction and analysis of progress.

The school has in place a variety of assessment strategies to provide information on students' achievement and progress and has the intention expressed in its planning documents, to investigate further outcome evaluation and recording. This step would provide a coherent approach and meaningful information to communicate to students and parents about individual student development.

These findings of good practice, however, are few in number compared to the findings of less satisfactory practices. The words of the review reports speak for
themselves. A lack of coordination of assessment and reporting practices between faculties in secondary schools, and between departments and year levels in primary schools is a common finding. So too, is a lack of shared understanding about the purposes and proper application of various assessment strategies.

Professional development has been undertaken within the school to develop and implement current theory and practice in the area of student assessment and evaluation. Some teachers have produced innovative and comprehensive assessment systems and procedures in reporting student outcomes. However, the school would benefit by capitalising on these developments and producing a comprehensive assessment policy to provide a more coordinated and sequenced approach in reporting student outcomes to parents.

A range of opinions exist among teachers about what the actual grading on a report referred to: performance in standardised tests, classroom units of work or a combination of both. Students were not always able to see the relationship between the report grade and their performance in class. Some parents and students also had difficulty in understanding the relationship between gradings for application.

The review team found that the reporting format was not closely linked to the student learning outcomes in the syllabus document. Of most concern were the methods employed in assessing student work. Testing of student recall was seen as over emphasised and a need was identified for more relevant and varied assessment approaches that truly assess the range of desired outcomes.

Some parents questioned the use of normative pen and paper testing to assess students who had been working all year in a concrete, experiential programme.

These findings led to recommendations similar to the following:

That the school further refine the student assessment and profiling process to create more meaningful and regular information on student outcomes.
In more than half of the review reports, the manner in which assessment information is reported to parents failed to meet their expectations. Parents complain that reports are too infrequent, too late to be of use in identifying emerging problems, and contain too little or confusing, jargonistic information.

A significant group of parents felt that the written report came too late in the year and that more information was needed to provide support at home.

Concern was expressed by a number of parents that the satisfactory/unsatisfactory rating on the personal and social development area on the school report did not convey a full and accurate description of the child and some problems were experienced with report interpretation.

Whilst there was satisfaction from most parents regarding the report format, many parents felt that there was a need for some sort of benchmark reporting to take place and that there needed to be greater standardisation and consistency in reporting between classes and grades. It is recommended that the school review its assessment procedure to ensure:

- consistency of approach to K-6
- effective use of available benchmarks and profiles which indicates a students' relative achievement, and
- parental understanding of the school's assessment philosophy or processes.

On the other hand, parents in many schools spoke highly of the degree to which teachers made themselves available to parents, and the how useful they find these opportunities for face to face reporting.

Parents particularly recognised teachers for the regular feedback they were able to give about student progress. Teachers were seen as supportive and approachable and the value of the school report, regular parent meetings were praised.

There are some tensions evident in the kinds of information parents want to receive about the progress of their children. Some want to return to the place in class and percentage marks familiar to them from their own school days. Some schools have not sufficiently well communicated to all parents the advantages of alternative methods of
reporting student progress. Parents seek information in ways that is easily understandable to them and provides an honest and unambiguous description of what their children know and can do, relative to other children of the same age and relative to some "standard" or set of expectations. The following extracts encapsulates the tensions in reporting requirements.

Many parents were happy with comments on the report form, particularly on the infants reports, and some parents were happy to take advantage of the parent-teacher interviews. Some parents, however, felt that there was a necessity to include a percentage score or some reference to position in class.

A committee of parents and teachers have been working upon a new format for the School Report on pupil progress. Some indecision in regard to rankings or gradings as opposed to the use of descriptives or profiles remains even after many meetings and parents and teacher input through surveys.

Summarising the findings of the analysis of the data from the review reports leads to the following conclusions:

• **Student assessment data does not play a significant role in school planning for quality learning outcomes in a majority of schools.**

A majority of schools do not make extensive use of either external or internal student assessment data in planning for quality student outcomes. Many teachers see external data as lacking relevance for individualised instruction.

Assessment data do not significantly or consistently inform teaching practices. Many primary, but particularly secondary teachers do not perceive assessment to be an essential and integral part of sound teaching and learning practice, seeing it as something largely for the purpose of reporting to parents.
Many teachers noted to review teams that they urgently require professional development in assessment technologies. Secondary teachers, in particular, have a limited understanding of the range of purposes for assessment. There is confusion between the terms "assessment" and "evaluation". Secondary teachers rely almost exclusively on summative rather than diagnostic and formative information. Avenues for professional development addressing these challenges need to be explored.

The further use of computer technology will allow teachers to store and manipulate larger quantities of data to inform their teaching practices and accurately report on student outcomes, particularly outcomes assessed using profiles. It would be virtually impossible to handle profiles data by other than electronic means.

- **Assessment in senior secondary schools is driven by HSC requirements.**

Many teachers of students in senior secondary years believe that they are not at liberty to adopt assessment practices that are best suited to individual learners and their educational needs. There is also a 'trickle down' effect to lower levels of schooling. Typically, secondary school assessment policies refer predominantly to requirements of external assessment (School Certificate and Higher School Certificate). Rarely do they refer to assessment practices for other purposes. Many secondary schools do not have assessment policies that are relevant to Years 7–9. As a growing proportion of students in the senior years are not seeking entrance to universities, it is important that assessment practices for these students reflect the skills and knowledge valued in courses other than the traditional pathway. In vocational courses, the kinds of assessment tasks used are frequently not in keeping with the objectives for the course.

- **Assessment practices in primary schools rely heavily on informal, subjective judgments of teachers.**
While the use of informal methods of assessment such as anecdotal records are useful means of assessment, many teachers appear to rely on these almost exclusively. This has lent an overly subjective and imprecise dimension to their reporting of student outcomes. Tangible student outcomes are rarely the basis of reporting to parents in these instances.

In some areas, few teachers use BST data to inform their teaching practice, seeing them as irrelevant to what is taught and what is learnt. Some Year 2 teachers expressed the belief that BST data for tests administered in Year 3 are not relevant to their teaching practices.

There is widespread confusion among primary teachers and school executives about the meaning of “student outcomes” and “profiles.” While many primary school teachers report that they use “profiles”, it is clear that these take a wide variety of forms, some of which are modelled on the current pupil record card. As the term “student profile” has assumed great importance in the assessment and reporting of student outcomes, it is crucial that teachers have an informed understanding of the term and its relevance in assessing student outcomes. Teachers are concerned about their ability to use profile information. There is a clear need for extensive professional development opportunities for principals and teachers to support the use of student outcomes data and student profiles in assessment and reporting to parents.

- **There is little coordination of the use of assessment data between primary and secondary schools.**

Student outcomes information from primary schools is not widely used in secondary schools. Legislative requirements prevent the transmission of information about BST results to the student’s new high school. Pupil record cards are regarded by secondary and primary principals as inadequate for recording student achievements,
secondary schools frequently conducting “grading tests” at the end of year 6 or early in year 7 to make placement and remediation decisions rather than rely on the advice of the primary teachers. This adds to the hiatus between primary and secondary education. A continuous profile of student achievement following each student from primary to secondary school will contribute to greater recognition of learning as a continuous process.

- **Parent and students play little, if any, part in determining how students are assessed and are not adequately informed as to how assessments are arrived at.**

  Assessment practices are determined in most cases by class teachers and faculty heads/grade supervisors. Parents have little input and frequently comment that they are not adequately informed as to what practices are implemented. Teachers and school administrators commonly believe that the determination of assessment practices is a professional concern of teachers rather than other members of the school and broader community. The role of school councils is usually limited to commenting on the format of reports to parents and, less commonly, ratifying broad assessment policy statements. Students generally play little role in determining assessment practices. Moreover, students (with the exception of students in Years 11 and 12) are frequently not informed as to how they are assessed.

  Students report that they receive information about their programs primarily from teachers. This is provided through awards, results of tests, teacher comments and placement in groups. Teachers report that diagnostic tests form an important part of student assessment with a range of other processes in use that are largely teacher-centred.

**Future Directions**

All schools need to reflect on their assessment practices, determine how appropriate they are, and how they might be improved. Schools will need to consider carefully the reasons for assessing students and adopt the most appropriate procedures.
A range of assessment strategies will be required to assess students' knowledge, skills, and attitudes using both qualitative and quantitative information in formative and summative assessments. They will need to consider the extent to which external examinations will continue to influence their assessment strategies for the whole school population.

Teachers will need to become more proficient in maintaining ongoing records of students' achievement related to clearly articulated outcomes to assist them in making professional judgements about levels of student achievement.

In order to ensure that schools have a consistent approach to assessment, the school's assessment procedures need to be carefully aligned with the school's policy on assessment which should be fully documented and communicated in an accessible form to parents and students. Such an initiative is timely, as many schools recognise a need to develop or update assessment policies and they may need advice on this process.

These policies and procedures need to give as much consideration to the way assessment information is used in the school as they do to what is assessed and how assessment is conducted. Assessment information at present is largely the preserve of individual classroom teachers. Better ways will need to be found for sharing appropriate information with parents, other teachers and the broader school community. Assessment information has the potential to inform whole-school planning, as well as planning at the faculty and classroom level to modify and improve teaching and learning in programs and classroom practices. Schools may need support in the use of appropriate computer programs and technology in order to effectively process, store and communicate the assessment information.

A continuous profile of student achievement following each student from primary to high school is one way of reducing the isolation of assessment information, allowing learning to be seen and understood as a continuous process. This, in turn,
would promote the notion of continuity for students as they move from class group to class group and from year to year.

**Conclusion**

The QA school review reports provide a valuable source of knowledge about the current status of schooling in NSW. When used in this way, there is a danger that this material might date quickly (however, over the three year period covered by the reviews, few discernible trends were evident in the issues of enduring concern to educators). What raises the value of the information in the collective review reports above the level of mere historical interest is the ability of the dataset to provide insights into successful and less successful practices in teaching and learning and school management and governance. Analysis of the material allows researchers to identify a range of problems facing schools and teachers, to identify the solutions obtained by others, and more importantly, the reasons underpinning their success or failure. By identifying the underlying principles of best and worst practice, educators have a series of qualitative benchmarks against which they can test their own practice. At the same time, the analyses of the school review reports provide strong messages for the administrators of the school system. In the example of assessment practices illustrated in this paper, there are some key messages that the system ignores at its peril. That such a widespread cross-section of practitioners demonstrate such a poor understanding of one of the fundamental aspects of the teaching and learning process ought to be of deep concern.

There are some who believe that the Quality Assurance school review process was a profligate use of resources.\(^1\) Such a view shows little understanding of the power of the information generated by the program over the three years during which it operated (which in any case accounted for less than 0.1 per cent of the total NSW

education budget for the period), and the value that remains in the information collected. Used well, this information has the ability to promote change well beyond that offered by the program itself.
Bibliography


Chapter 7

What Students Expect of Their Schools:
Findings from Quality Assurance School Reviews

What Students Expect of Their Schools: Findings from Quality Assurance School Reviews

Quality Assurance school reviews provide an opportunity to gain information about the attitudes and perceptions of teachers, parents and students on a scale not commonly available in the past. This paper draws on the findings of reviews of 430 schools in NSW conducted from Term 3 1993 to the end of Term 1 1995, which included interviews with over 15,000 students.

Within each school review, groups of 5–6 randomly selected students from each grade level are usually interviewed. These interviews focus on many diverse topics, which collectively provide insights into many issues including students’ expectations of schooling.

The Learning Environment

Students expect their schools to be well resourced, attractive and well maintained, with pleasant grounds and a wide range of facilities. Students, as do parents, comment very frequently on the efforts of the school to provide surroundings which are aesthetically pleasing, and respond positively to efforts to involve them in this endeavour. Tree plantings, creation of environmental study areas, provision of shaded areas, paved courtyards and assembly areas, and covered walkways are particularly valued.

Students are appreciative of the school grounds and facilities. Classrooms are freshly painted and well maintained, a feature of the school valued by many parents.
Many students reported that their sense of ownership of the school and its environs, developed by physical participation, had contributed to their own sense of self-worth. Examples of participation given by the students were: working on regeneration of a small piece of rain forest; caring for the schools’ guinea pigs and chickens; and planting trees and setting gardens.

Reviews also recognise student needs for access to well maintained playground space, sporting fields and toilet facilities. The extent to which all students have access to high quality facilities varies from school to school. The challenge provided for schools, and the education system, is to establish means by which student needs can be recognised and addressed within the planning and budget cycle.

In addition to their physical surroundings, students equally value the opportunity to learn in what they perceive to be a safe and secure environment. They want their classrooms and playgrounds to be free from bullying and threats of physical violence, and they expect teachers to maintain good order and discipline. The following examples from two secondary school reviews illustrates this theme.

Students have become increasingly vocal about disruptive classroom conduct, where the “few” prevent the lesson proceeding. These same students are especially critical of teachers who do not adequately manage these situations.

A common complaint among some students concerned the behaviour of disruptive students. While some students were particularly sympathetic to the teachers faced with these students, they were scathing in their comments about any teachers who failed to address the situation caused by disruptive students.

The vast majority of schools are commended by the reviews for the caring and supportive environments provided by teachers. Primary schools in particular are often noted as having a “family” atmosphere, and as being “friendly, safe, non-violent and multiculturally harmonious.”

Younger students value the family atmosphere, the small enrolments, and more particularly the sharing and caring attitude of everyone in the school.
Another example from the high school level demonstrates the commitment of many schools to student welfare:

High profile activities such as the Flag Ceremony, Aboriginal Dancers and the building of an Aboriginal Resource Centre have assisted in the raising of awareness of the multicultural make-up of the school and an appreciation of these students’ traditions and current cultures.

Students recognise the efforts made by schools to foster these conditions, and cite Peer Tutoring, Peer Mediation, Buddy Programs, Peer Support and positive reinforcement schemes, such as merit awards as being particularly effective. Systemic initiatives have also been recognised. For example, one review noted:

The development of a safe, secure environment has been enhanced by implementation of the Anti–Racist Policy and Child Protection Program.

**Involvement in Decision Making**

Reviews increasingly note that schools are providing opportunities for students to participate in leadership activities and decision–making, especially through the establishment of Student Representative Councils (SRCs). Students at all levels express strong desires to have input into the school’s planning and decision making process.

During the review, students expressed support for their school. They believe they have a role in decision making through the SRC, peer support and mediator programs. They believe these programs motivate them and develop confidence.

The Students Representative Council is identified by students and staff as being an effective forum through which students can contribute meaningfully to school governance. Regular reporting back to student members by their peers and a structure for students to present their views to class representatives have led to all students feeling that they have a voice in the school and that their views are valued.
Students expect that their views will be taken seriously. They express concern wherever structures are established which convey the impression that their contribution will be tokenistic or subject to staff veto. They expect that action of some kind should result from their discussions. They are also concerned that the role of the SRC should not be limited only to fundraising.

The Student Council is appreciated by students and staff as providing a forum to discuss aspects of school life. Students comment, however, that more action needs to result from Council deliberations.

The review team noted . . . the successful operation of the Students Representative Council which had done much to introduce students to the responsibilities of governance. However students were divided on the impact of this body. In this context the issue of staff wearing hats outdoors was mentioned.

However, schools where large numbers of students have significant input into decision making are in the minority, and many reviews recommend further development in this area. For example, in one review, a significant number of respondents indicated that there was a need for the Students' Representative Council to have a greater involvement in the school's decision making and opportunities to be involved in a wider variety of tasks, to be more representative of them, for the Council to be more accessible to them, and for the student body to be better informed about its work and achievements, leading to a recommendation:

That the school evaluates present structures for student participation so that students are empowered in decision making, implementation and sharing of information.

**Teaching and Learning**

A question commonly asked of students during the reviews is “what helps you to learn best?” The responses of most students demonstrate a high degree of maturity and sophistication across grade levels, and provide strong messages for teachers about their selection of instructional strategies. These responses include the following:
use of a variety of approaches
active student participation in lessons
encouragement of discussion
less teacher talk
praise for a job well done
valuing of individuals
wider application of hands-on experiences and group work
an increase in excursions to support learning experiences in the classroom;
sound discipline, but not necessarily silent classrooms
teachers listen to students and give them time
teachers give satisfactory explanations
fewer assumptions that students all learn the same way
learning in a team situation.

Students value learning which is purposeful, which engages them and caters for their learning styles. Students comment on how demotivating they find learning which is substantially passive and not interactive. Students commented on the importance of developing capacities for creative and original thinking. Students expressed least satisfaction when they found themselves in learning situations that did not take account of their language needs, learning styles or the student-centred nature of the learning environment.

However, many experienced teachers still favour traditional teacher-centred approaches, with some staff showing little or no variation in their teaching practice.

A small number of students indicated they believed some staff were “old fashioned” and “inflexible”.

Students at the secondary level require more attention to be given to such things as study skills, varied teaching strategies, education for employment, the extension of student/teacher relationships, higher teacher expectation of student achievement, student risk taking, problem solving and creative thinking, and independence in learning.
Students want to share more responsibility for their own learning. A few schools have made notable progress in this regard.

Efforts to create an adult learning environment have resulted in the majority of students taking responsibility for their own learning. Those efforts have created a responsive learning environment marked by freedom and flexibility.

There is much that could be done by schools to become more responsive to these demands. Review reports note few instances of students being involved in the evaluation of their own work, and for many students the processes and criteria used to assess them remain a mystery.

Students value highly their relationships with teachers and expect that teachers will in turn value them as individuals who are deserving of respect. Adolescents need to experience a sense of security and “belonging”, which requires schools to have a strong focus on the welfare of its students and support this focus with a wide range of programs.

Students spoke at length regarding the critical importance of positive, constructive relationships with teachers. They consider it to be the single most important ingredient in their education. Students like to see teachers enjoying their teaching.

The excellent rapport between staff and students has engendered cooperation, trust and enthusiasm.

Senior students appreciate the school’s positive tone manifested in harmonious and supportive staff–student relationships; sensible, mature student behaviour and attitude; strong teacher encouragement for student endeavour; and willing, ready access of staff for student assistance.

**Student Welfare**

The review reports suggest that schools place considerable emphasis on establishing strong welfare programs, and many schools have been commended for the success of these efforts.
Students also expect their efforts to be recognised. Many schools operate award schemes to encourage student achievement and behaviour, but review reports note that the biggest criticism of these schemes is the inconsistency with which they are implemented. Students do not want awards trivialised. Both students and parents feel that there could be more consistency in approach by all teachers and more emphasis placed on merit and consistent effort. In particular, they suggest that more attention could be given to recognising the quiet achiever as well as students who excel academically.

Students appreciate displays of their work and acknowledgment of their achievements when publicly displayed, or announced, at school assemblies and through school newsletters to the home and wider community.

More than anything else, students expect that their learning will be important, relevant and meaningful. Particularly at secondary school level, students seek confirmation that their day to day tasks will help to prepare them adequately for the future. This should not be taken to imply that the majority of students have a narrow, instrumentalist view of the value of education, but that they seek validation of their ability to have some control over their own destinies. In this regard, they want regular, purposeful teacher feedback on their performance.

The review reports note that in many schools there is considerable tension between what students consider important (based on their perceptions of the future), their parents' wishes, the expectations of teachers and the demands of a mainly tertiary entrance examination-centred curriculum (the recent introduction of alternative "Pathways" notwithstanding). This tension has been heightened by the rapid increase in senior school retention and the prevailing economic circumstances, and is perhaps most acute in rural areas. This tension manifests itself in several negative ways, including inappropriate subject choice, underachievement and inappropriate behaviour.

Parents, students and staff identify a tendency on the part of some students to under-achieve as a result of pressure from their peers. It was verified from the school records that
there was a disproportionately small number of boys receiving awards, both within the merit system and at presentation night.

The issue of informed, appropriate student subject choice is one which greatly concerns the school. There still exists a resistance to the consideration of alternative course options.

There is a lack of co-ordination for middle school students; the feeling of powerlessness of some students and the provision of inappropriate and restricted 'pathways' alternatives available. There is a need to address the growing mismatch of expectations of staff, students and parents in relation to curriculum directions and possibilities.

Summary and conclusion

The review reports clearly indicate that students better identify with their school when:

- The school has a strong focus on the welfare of its students and that this focus is supported by a wide range of programs.
- The school perseveres and explores as many avenues of assistance and support as can be found in its approaches to management of students who are having behavioural or learning difficulties.
- The school recognises the benefits to students in curriculum and welfare matters of having close ties with primary schools (and the home) and is engaged in a planned program of further development of links with them both.
- The school recognises the importance of and caters for the recognition of the achievements of its students.
- Staff are approachable, caring, and are committed to catering for the needs of individual students.
- The school offers a well-balanced and varied curriculum, enhanced by a broad range of cultural and sporting activities.
- There is a high level of communication throughout the school community, particularly in relation to curriculum and course and subject choice.
- The teaching and learning processes empower students as learners so that they take responsibility and ownership of their own learning.
Chapter 8

The Achievements of Schools as Seen by Parents: Findings From Quality Assurance School Reviews

The Achievements of Schools as Seen by Parents: Findings From Quality Assurance School Reviews

Quality Assurance school reviews provide an opportunity to gain information about the attitudes and perceptions of teachers, parents and students on a scale not commonly available in the past. This paper draws on the findings of reviews of 430 schools in NSW conducted from Term 3 1993 to the end of Term 1 1995, which included interviews with over 10,000 parents.

The school review reports provide details of the many achievements of the individual schools reviewed, based on the observations of the review team and the perceptions of parents, students and staff conveyed to them. Each school has its own particular strengths and areas for development. These achievements are too numerous to list individually. This paper discusses some of the themes recurring in the achievements of schools as seen by parents.

Quality Assurance school reviews indicate that parents are more involved than ever in their children's education, more knowledgeable about what goes on in the classroom and are generally positive about what they find.

Parents are proud of the schools their children attend and cite as a major achievement the schools' ability to provide a wide range of educational opportunities as well as giving priority to the teaching of basic skills. Parents and community members appreciate the dedication, care and commitment and professionalism of the staff in providing an educationally challenging and student-centred learning environment.
Parents also take pride in the physical appearance of the school and its surroundings. In many country areas, the school often provides a primary focus for the community.

Parents valued the delightful school grounds and buildings which promote an atmosphere conducive to learning and which are respected and appreciated by the school community.

'The Open Door'

In the majority of reviews, schools are commended for the open and welcoming image that they present to the community. The phrase used in one school review report, that school is “a place their children like coming to”, is representative of the general parental perception of a successful school. It is also a place that parents like coming to and where they feel welcome and valued.

Many parents commented on the ease of communication between home and school because of the friendliness and the caring and committed nature of the staff. The nature of these relationships between staff, students and parents, were described as both open and close. Many schools and their principals have adopted an “open door policy” which was well known by parents, teachers and students.

Parents praised the open door policy in classrooms and the warm parent/teacher relationships which made parents feel welcome to contribute to class activities.

Most schools are aware of the need to reach out to all sections of the school community and have instituted many strategies to do so. The translation of school newsletters into community languages is common. Some schools also provide large print versions of school documents for parents with impaired vision. Other schools have accepted the need to be more flexible in their arrangements for parent/teacher meetings so they can be more responsive to parent’s needs. Parental involvement in classrooms and other school activities is widespread (at primary school level in particular). One
school has instituted a “buddy system” for parents to link new or less confident parents to the more confident and experienced, in order to encourage parent involvement.

The message parents receive from these efforts by schools is that they are valued members of the whole school community for both social and educational reasons.

The parents overwhelmingly reported that they were happy to have their children attend the school and that they were made to feel very welcome. They understood that their opinions were sought and valued and that information about their children's education and the school's programs and activities were readily available. There was an exceptionally high degree of recognition of these programs and activities by parents which represents a strong endorsement of the effectiveness of the established lines of communication.

Many reviews cite examples of parent involvement in parent tutoring, particularly in reading, sporting activities and excursions as well as other examples of assisting teachers in the classroom. One review noted a school which invites parents to fill in goal sheets for their children. These goal sheets are then taken into account when planning school programs.

Schools advanced the view that the active presence of parents in the school was appreciated by the students and provided for them a sense of security in the learning environment, promotes student self-esteem, breaking down the barriers between school and home.

Parents also see the educational interests of their children as being strengthened by the relationship between parents and teachers. Parents have achieved a new status and role in schools through this partnership between parents and teachers.

Within the culture of the school there is a partnership between parents and teachers in the teaching-learning process. The review team through observation and interview received information on the high degree of parental involvement in the learning process. Every day numerous parents help in the school's classrooms. These parents have undergone school-based training and assist teachers and pupils. Their presence allows teachers to organise many different activities as part of normal practice.
Parents in many schools expressed the belief that student learning is not solely a teacher or school responsibility but also requires their support and involvement.

Parents as Decision Makers

The opportunity to be involved in substantive decision making about the direction of the school is seen as an achievement by some parents. Further development of parent participation in school governance by means of their membership of School Councils has taken place in 1994. Parents are becoming increasingly aware of this development and identify it as a way of enabling them to be involved in the decision making processes of the school. For many parents, this is perceived as a welcome development away from their traditional fund raising role. It must be noted that a significant number of parents explained that time constraints and family and business commitments precluded their involvement in school programs, but this lack of involvement did not mean lack of interest.

Schools are also increasingly inviting parents onto other decision making forums such as Disadvantaged Schools Committees and School Development Day planning and Training and Development Committees. Some parents indicated that they did not want to be personally involved in any decision making processes but all wanted full and detailed information on the results of school based decisions. Overwhelmingly, they seem to be satisfied that this is being done.

Where there was no culture of improvement in the school, parents had little understanding of this aspect of school life.

Building an Understanding of Diversity

Parents from a wide variety of backgrounds perceive the success of schools in fostering acceptance of cultural diversity as an emerging strength.
Schools are developing awareness of the wide cultural backgrounds of their students and the importance of preserving their diverse cultural identities. Aboriginal education programs are strongly supported by staff, students and parents. Some schools are forming Aboriginal Student and Parents Committees to assist the school in further raising the awareness of the school community.

Staff, parents and community members feel that the varied and effective activities of the Aboriginal Programs give it prominence as the best feature of the school. They believe that there are many indicators of successful outcomes for students, the school and the community.

Parents are proud that school is an environment where children are encouraged to become tolerant and sensitive to the differences in society.

The acceptance and celebration of multiculturalism has been thoroughly embedded in the culture of the school. Staff, parents and students mentioned tolerance as a key feature of school ethos. The many opportunities for developing multicultural understandings through social activities provided at school were highly valued by respondents.

Parents see schools as taking a positive approach to catering for a wide range of student abilities. Many schools have developed programs to cater for students with specific learning difficulties and the developmentally delayed. In this context, parents particularly value the growth and acceptance of integration programs.

**Schools as Happy and Safe Environments**

Reviews generally indicate a high level of parental satisfaction with schools as happy and safe environments for their children.

The school is acknowledged as having a friendly, dedicated and professional staff committed to providing a safe, caring and supportive learning environment. Students are enthusiastic and enjoy their classroom and school activities.

The continuing emphasis on student welfare policies and programs by many schools has led to perceptions of:
• the fostering of growth in self esteem and confidence;
• the importance of students taking responsibility for their own behaviour; and
• a lower incidence of bullying behaviour.

An increasing awareness of and willingness to deal with violence in the playground was noted and appreciated by parents, although some parents are still concerned about levels of playground supervision.

Parents have more knowledge of the increasing use of programs such as Peer Support, Buddy Programs, Personal Development Programs, Child Protection Programs and Life Education Centre Programs and consider they also have contributed to the development of safer, friendlier schools.

It is noted that there are still some parents who are not aware of these and other programs which are outside traditional subject areas.

Reviews show that parents approve of the concept that students benefit from mixing across ages and grades. This approach is more common in primary schools, where events, activities and special programs are increasingly run on a K–6 basis.

Schools’ increasing use of behaviour management processes are seen as effective in developing positive behaviour in the large majority of students. Parents report their support of award or merit systems. They report that students are proud of these awards and like to display them at home. However, reviews have indicated some areas of criticism by parents in the operation of merit award systems. For this reason, it is essential that merit systems operate consistently and fairly.

**Student Achievements**

Ultimately, parents judge the success of a school by the achievements of its students. Review reports cite many instances where parents value highly the performance of the school at public examinations, state–wide academic competitions
such as the Mathematics Olympiad, sports, and cultural activities such as the Rock Eisteddfod. Parents also value highly the many programs launched by schools to cater for the learning needs of individuals.

Parents expressed the view that the school provided a happy, caring environment which encouraged learning. Parents perceived that the school was catering for the needs of all students. Many parents were able to cite particular school programs which they feel were especially valuable in meeting students’ needs.

On the whole, the majority of parents interviewed in most schools reviewed are satisfied with the quality of the teaching programs offered and the standards achieved by students. There are, of course, areas for further development both in particular schools and across the education system in general. Parents in primary schools, in particular, report satisfaction with levels of acquisition of basic skills.

Parents indicated that they were generally satisfied with the approach to the teaching of English and that improvements in standards had occurred in reading.

Parents were pleased with the progress their children showed in mathematics. It is very evident that the school is meeting parent expectations regarding acquisition of functional numeracy skills.

A significant number of parents have very high expectations in relation to their children’s achievements. They recognise and appreciate the school’s commitment to meeting these expectations.

At secondary school level, attempts made by the school to cater for the full range of student abilities and the extensive subject offerings were singled out for praise in many reviews.

Parents and students value the school and its diverse curriculum range and are proud of the image it projects to the wider community.

Many interviewees from all sections of the school community expressed general satisfaction with the senior and junior curriculum. In particular, favourable comment was made about the balanced range of academic and vocationally oriented courses. These were
seen as catering for the educational needs of students. This was perceived as having been achieved without the introduction of subjects which devalued senior study.

While many parents express satisfaction with the outcomes achieved by their children, many others express concerns about the teaching and learning practices evident in many schools. They do, however, appreciate the efforts of many teachers to change towards more child-centred practices.

Review reports indicate that these features are an integral part of the culture of many primary schools and Schools for Specific Purposes. They also suggest that there is scope for development of these features in a great many secondary and central schools.

Summary and conclusion

The review reports indicate that parents more often view schools positively when they are:

- made to feel welcome and see that their opinions and contributions are valued and accepted;
- well informed about curriculum and school processes and structures;
- given regular, meaningful information about student progress and are accepted as partners in the education process;
- offered the opportunity to be involved with school planning and decision making processes;
- invited to visit the school, attend meetings/functions and have access to teachers;
- confident that the school encourages cultural diversity and is seen to welcome the contributions from the entire community; and
- satisfied that a secure, caring environment exists in which their children are happy and encouraged to develop to their full potential.
Chapter 9

What Parents Want from Further Participation and Communication: Findings from Quality Assurance School Reviews

What Parents Want from Further Participation and Communication: Findings from Quality Assurance School Reviews

Quality Assurance school reviews provide an opportunity to gain information about the attitudes and perceptions of teachers, parents and students on a scale not commonly available in the past. This paper draws on the findings of reviews of 430 schools in NSW conducted from Term 3 1993 to the end of Term 1 1994, which included interviews with over 10,000 parents. It discusses some of the findings of the reviews concerning what parents want from further participation and communication in relation to curriculum, teaching and learning, and assessment and reporting.

Quality Assurance school reviews indicate that parents are more involved than ever in their children’s education, more knowledgeable about what goes on in the classroom and are generally positive about what they find.

Parent Participation

Many schools already enjoy a considerable amount of community support. The following list synthesised from the review reports indicates the range of activities that benefit from parent involvement.

- construction of school equipment
- school donations – fundraising
- production of uniforms
- transport on school excursions
- working bees
• classroom support, such as reading groups, process writing, and oral history
• finance committee
• selection panels
• making items for school use
• assistance at swimming, athletics and sports carnivals.

These activities are exemplified in the following extracts from school review reports.

The community is involved in all aspects of the school. This has resulted in additional facilities around and within the school.

The P&C is active in fundraising, curriculum decisions and whole school initiatives.

Parents are prepared to participate in classroom activities and are always present when transport is required or additional adult helpers are required.

Primary schools in particular have high levels of parental involvement in the classroom and other school activities. The term “involvement” is used in this sense to mean working at the request of, and/or in conjunction with teachers. In addition, the opportunity for parents to participate in broader and more substantive decision making is already widespread, with nearly two thirds of NSW public schools having school councils. Some parents saw the formation of a school council as an opportunity to:

• provide a higher level of decision making
• facilitate recognition of school achievements
• assist in the formation of school policies
• support harmonious teacher parent relationships.

Where school councils have been established, the review reports indicate that what parents want is not so much more decision making powers at this stage, but greater clarification of roles, particularly the differentiation between the roles of the School Council and the P&C Association.
Many parents were unsure of the purposes and implications of the school council. Some expressed concern that it may lead to an intrusion on curriculum matters. Others were concerned about its effects on the P&C.

Many parents and staff noted that the roles, functions, purposes and operation of the three parent bodies were unclear and not coordinated.

This is not to say that the needs of all parents in all schools to have access to decision making structure have been met, as the following extracts illustrate:

Some parents expressed the desire to more actively participate in the management and planning structures of the school.

Some parents expressed a wish to have a greater involvement in decision making and saw this as being facilitated by the proposed formation of a school council.

The review reports suggest that the absolute number of parents who express a desire to be more actively involved in decision making is relatively small. However, common across all schools is a desire by all parents not to be excluded from the decision making process, should they wish to be involved. Parents want decision making processes to be open, and value the opportunities that they have to contribute in their own way. More than anything else parents wish to be kept fully informed of the decisions that are made.

When not involved in decision making, parents want to be kept informed about school policies on matters such as welfare, finance, management and administrative issues.

Many schools have been very successful in communicating with their communities, some going to great lengths to achieve this. These efforts send powerful messages to parents that they are valued partners in the learning process, and in return, parents in these schools are more likely to hold positive opinions about the school because of their informed understanding.

Even in those schools which do value and promote parent involvement and participation there are many barriers to further increasing the range and level of support
Currently enjoyed. In some schools, participation is limited to only a small section of the school community for a variety of reasons.

Parent’s knowledge of the school’s planning structures was restricted to members of the committee.

Some community members recall their own school experiences as threatening and while noting the welcoming attitude of the school, feel their own lack of confidence inhibits their relationship with the school. A few parents mention that their interactions with the leadership and some staff made them feel intimidated and uncomfortable. As a result of these experiences some parents feel that their point of view is not heard or valued and their potential to contribute to the directions of the school not realised.

Some of these issues can be addressed relatively easily by schools. In one school, for example, parents said that they “tried to help but were unsure of the expectations and role that they can play”, a problem which could simply be addressed through enhanced communication.

Some parents expressed a wish to provide more assistance in other curriculum areas (other than reading).

In other instances, there are deep divisions within the community and clashes of values and cultures which are beyond the ability of schools to solve on their own. The challenge for public schools is to be as inclusive as possible of the entire community.

Barriers to greater participation included a belief by some that their opinions are not valued by other parents.

Some teachers and a large number of parents commented on the difficulty of reflecting the community needs in the school because of diverse values and attitudes.

While acknowledging the efforts of those involved in the parent body, several parents described it as a ‘closed shop’ and that new members were not actively encouraged to join.
Reporting to Parents

Parents’ needs for increased communication from schools continue to be strongly expressed in 1994 school review reports. In responding to the review teams, parents in the majority of schools say that they have little knowledge of the curriculum in most Key Learning Areas (KLAs), at either primary or secondary school level. The following examples are typical of the findings and recommendations in this area:

A majority of parents interviewed expressed a lack of knowledge and understanding of the Science and Technology Curriculum. They wished to have a better understanding of its content and application.

[It is recommended] . . . that the school explore ways to further develop community understanding in mathematics with a particular focus on methods, homework and expected student outcomes, so that parents are able to be more informed and effective partners in their support of student learning.

Parents indicate that importance needs to be given equally to all KLAs. Some parents also indicate the importance they attach to programs such as First Aid/Health, Child Protection and HIV/AIDS awareness. Parents want the development of collaborative and cooperative skills to be taught across all curriculum areas.

Parents seek to be informed about curriculum and other school programs for a variety of reasons:

- they want to know what is happening to their child during the course of the day (the inability of parents to extract this information from their children is legendary);
- they want to know how they can help their child;
- they want reassurance that their child is learning in ways which will prepare them well for the future.

Parents see their involvement as essential in promoting student learning. They believe involvement in the classroom assists the teacher and benefits students. Reports frequently note that parents appreciate the opportunity to gain further information about the school’s programs through working in classrooms.
Parents seek reassurance that what children learn is worthwhile:

Whilst parents staff and students acknowledged basic literacy and numeracy as being the important subjects students should learn they have also identified a large number of other essential skills and attitudes which they felt should underpin curriculum. These included how to learn; social and living skills; communication and public speaking; confidence; self-esteem and assertiveness; tolerance and understanding; decision making; self control; listening and analysis skills, interpersonal skills; honesty; cooperation and respect for others; and computer skills.

Parents indicated that they wished to be informed about new policies, teaching learning strategies and specific subject syllabuses. They also wanted to be informed on ways parents could assist their children.

In many instances, parents seek this reassurance that students are learning in ways which make sense to them and are familiar from their own school days. This is often expressed in terms of requests for children to be drilled in tables and spelling lists, a finding common in small rural schools. Some parents ask for greater emphasis on handwriting skills and book presentation.

Parents request more rigour in the development of literacy skills, specifically the teaching of spelling and grammar.

Parents say that they gain information about student progress through formal and informal meetings with teachers, from school reports, and from examining student work books and homework. Many parents value highly these tangible symbols of learning.

For many parents, the formal school reporting process is their only real contact with the school during the course of the year. School reports (along with school photographs) are important and often treasured records of a substantial part of family life. It is perhaps not surprising therefore, that a great many of the comments of parents in regard to the assessment process point to perceived inadequacies in these reports. The following extracts from a number of school reports capture these sentiments.
A significant number of parents expressed a desire to know more about the process of assessment, how the level of progress was determined and what methods were used.

Parents were unsure about the methods used to assess student progress.

Parents want more frequent reporting about the progress of their children.

Parent were confused about the format and structure of the reports. They describe the current reports as limited in the information they provide. They wanted to see more information in each KLA and reports on the progress of achievement not just effort. Parents wanted an opportunity to provide feedback on the reports.

Some parents believed that there was a need for more information on student progress (relative to other schools). Others would appreciate more information on yearly reports.

Some parents need assistance in interpreting the information contained in school reports.

Reports should cover social as well as academic achievement and information about behaviour. Parents want comparisons of the individual student against previous achievements.

Parents in many schools express a desire for both comparative information about their child as well as a more substantive information about their achievements. They want realistic and meaningful feedback about their children’s performance and behaviour, but some state that they would prefer negative comments to be passed on at interview rather than on paper. Many say that they want earlier notification that their children are experiencing difficulties. Parents in some small schools feel unsure about their children’s academic standards relative to other schools.

Some parents suggested that the school should inform them of areas for concern as soon as a problem arose and that the half yearly report period was too late to give feedback on student progress.

Many parents want more consistency in reporting. In many schools parents spoke of confusion across the school generated by different teachers doing different things. This became most apparent when the reports for different years use different
models, different terminology, and at different times of the year. These parents reported their frustration and lack of understanding regarding the reasons for the absence of consistency. (Note: findings about assessment and reporting are discussed in more detail in a separate report).

Summary and conclusion

In summary, the themes arising from the information conveyed in the school review reports indicate that:

- a change in culture is needed in many schools to accommodate the needs of parents. This cultural change includes a shift in the relationships between schools, teachers, and parents – wherein continuous improvement and responsiveness become paramount.

- Schools have to be prepared to meet these needs of parents for information in ways that they have not done so previously. The passage may not be smooth in many instances, and schools must be prepared for disappointments.

- Many school reviews note that there is a mismatch between demands of some parents for more meetings and more training and development activities and the reality of their willingness and availability to attend.

- Some schools have achieved considerable success using non-traditional approaches. Suggestions made by parents staff and students include more opportunities for social contact between staff and parents, targeting small groups of parents, being more inclusive by including all caregivers in invitations and negotiating avenues of participation with parents.

- Parents recognise and appreciate access to school committees, P&C Association involvement, school development days and staff selection. In the same way, parents express trust and confidence in their schools.

- Individual school review reports provide a valuable resource for obtaining more information on successful approaches currently employed to foster parent participation and understanding.
Chapter 10

The Senior Curriculum and Subject Choice: Findings from Quality Assurance School Reviews

Unpublished paper
The Senior Curriculum and Subject Choice: Findings from Quality Assurance School Reviews

This paper draws on the findings of Quality Assurance School Reviews in 430 schools in NSW conducted from Term 3 1993 to the end of Term 1 1995, which included interviews with over 10,000 parents, 15,000 students and 10,000 teachers.

Curriculum arrangements are an important concern for secondary schools. Not surprisingly, more than 90 percent of reviews in secondary schools to date have included some aspect of the curriculum as a focus for investigation. More than three-quarters of these reviews have addressed issues concerning the curriculum for senior students. In a minority of reviews, the topic of senior curriculum was also discussed as a student welfare issue.

Broadening the Curriculum

A challenge confronting many secondary schools in recent years has been a rapid increase in the retention of students in the post-compulsory years, resulting in many instances in a considerable broadening of the range of abilities, interests and needs of students in the senior school.

The findings in the review reports indicate that many schools have responded to this challenge by broadening their curriculum offerings. In several instances, schools have adjusted their organisational structure to accommodate these changes. The reviews
indicate that in the majority of schools, the expansion of such senior curriculum options as Joint Secondary School/TAFE (JSST) courses, Training in Retail and Commerce (TRAC), Business–Industry Links and the opportunities these provide for work experience and vocational awareness have met with widespread student, parental and community support.

Respondents indicated the benefits of a broad curriculum catering for the wide range of abilities and aspirations of its post–compulsory students. Innovative timetable structures with a “vocational” day for seniors, and additional periods for the senior school added in the morning over the other four days, provide opportunities for all students to maximise their options. Students and parents valued the small, teacher intensive classes in some of the traditional academic subjects, when other students pursued vocational options.

The review notes the high level of success, enjoyment and independence which students are experiencing in the JSST program.

These vocational programs, when well managed, have proved valuable in extending the school’s profile in the community. There is some evidence that suggests many parents perceive the provision of a broad range of subject options is an important criteria in judging what is a “good” school. Community support for alternative pathways is strongest where it is perceived that its expansion has occurred without detracting from the reputation of the school in traditional academic courses. There is also evidence that suggests that student satisfaction with subject choice is a powerful determinant of their satisfaction with schooling in general.

There is significant appreciation of the ways in which the school’s curriculum is enabling students of all abilities to access courses which are relevant and in which they can enjoy success.

All schools seek to provide a curriculum which is as broad as the available resources will allow, attempting to balance academic studies, cultural experiences, sporting endeavours and creative interests.

In the focus area of the curriculum, the review found that the school offers a curriculum which has a wide range of choice based on the needs of students. Parents and the school
community favour the present curriculum because of its emphasis on academic achievement. Recent additions, especially 1 Unit non-matriculation courses, are seen as making the curriculum more responsive to the needs of senior students.

There was overwhelming support for the vocational options offered to seniors by parents and teachers who consider that many students have returned to the senior school because of the course variety offered. The effectiveness of the Training in Retail and Commerce program, which extensively uses local industry as training centres, has been very strongly supported by many respondents. The Joint Schools and TAFE program is also seen as most beneficial by all groups interviewed.

In small high schools and central schools, particularly those in more isolated reas, concern was expressed about the narrowing of subject choice due to low candidate numbers and as a consequence of falling enrolments. Geographic isolation also presents problems for some schools in terms of their ability to participate in Joint Secondary school/TAFE courses. Some arrangements made by schools to overcome some of these problems are discussed later in this paper.

Despite these successes, in many schools, a need was expressed for a continuing diversification of the senior curriculum with more vocational courses. A significant number of respondents across all groups commented on the benefits to students of subject patterns which included post-compulsory vocational options, most of which still retained a Tertiary Entrance Rank credential. The perceived benefits discussed by a range of respondents included:

- vocational courses count towards the Higher School Certificate;
- access to the Open High School to support an even broader subject choice for tertiary oriented students;
- vocational courses allow students to be treated as adults, interact with adults, and provide the additional focus of relevance and purpose;
- practical skills learned by students are seen as beneficial in seeking employment;
- a wide range of working environments provide competencies to benefit students when they leave school; and
- improvement in students’ self-confidence and maturity.
Despite the success enjoyed by a large number of secondary schools in introducing alternative vocational pathways, a number of unresolved tensions continue to be mentioned in the reviews. Across schools, there is a small but significant number of teachers, parents and community members who question the proper role of schools in vocational education. They argue that other institutions are better placed to deliver such services. For example, one review report notes:

A small number of respondents across all groups thought that non-academic students should enrol in TAFE.

In some schools, some parents and students have negative attitudes to non-traditional courses, arguing that they devalue the efforts of those enrolled in traditional courses. Though senior students have been provided with more flexible learning arrangements and options for study through vocational and life-oriented course offerings, entrenched views based on the traditional academic Higher School certificate pathway still persist.

Many teachers also recognise it as vital to change parental and student perceptions on the importance and relevance of an academic course culminating in a Tertiary Entrance Rank for many of the senior students.

Negative opinions about non-matriculation courses appear strongest in schools which have not traditionally been highly successful in preparing students for tertiary entrance. Many secondary school review reports raise the concern that the Tertiary Entrance Rank (TER) continues to exert undue pressure on subject selection in the senior years. Some students continue to select inappropriate subjects given their previous academic record.

Unrealistic student expectations and mistaken perceptions by parents were identified as impeding appropriate course selection.

There is need to address the growing mismatch of expectations of staff, students and parents in relation to curriculum directions and possibilities.
The fact remains, however, that the issue of informed, appropriate student subject choice is one which greatly concerns schools. Parental and community attitudes exacerbate the aim for Tertiary Entrance Rank scores, even though wider career and vocational opportunities may come from a non-matriculation Higher School Certificate. There still exists a resistance to the consideration of alternative course options. Acceptance of the concept of pathways by the community appears to be very gradual in some places. The majority of schools recognise the need for further communication with the school community and are making active efforts to more fully involve them in the course selection process.

All members of the school community acknowledged that the school has recognised the changing needs of the non-academic students in the senior school and has implemented some structures to address these needs more effectively. At the same time a significant number of parents, staff and students indicated that further restructuring of the senior curriculum was essential to provide a more vocationally oriented program for these students. Many teachers also recognised that it was vital to change parental and student perceptions of the importance and relevance of an academic course culminating in a Tertiary Entrance Rank for many senior students.

A range of respondents suggested that the communication of the dimensions and advantages of the post-compulsory program needs to be reinforced to staff, parents, employers and especially students in the junior school.

Many teachers also recognise the need to expand student's expectations and vocational options through the provision of a more proactive careers program.

Changes to school organisational structures

To accommodate the diverse needs of students in terms of the academic and non-matriculation pathways offered, some schools have made organisational changes. These have included:

- extensions to the school day by increasing the period offering from 8 x 40 minute periods to nine or ten such periods, with earlier starting and later closing times.
• varied school attendance times for junior and senior students to enable
greater use of specialist rooms.

• variations to lesson times with many schools looking to reorganise the
school day to 4 x 75 minute periods, minimising movement between
lessons and providing a lesson framework better suited to practical and
vocational course offerings.

• some alignment of senior timetables between adjacent high schools to
better resource low candidature subject offerings and facilitate shared use
of school facilities.

• The introduction of vertical integration into the senior timetable across
Years 11 and 12, to enhance student choice especially in 1 unit subjects
which may often only attract small numbers (one high school has
successfully offered a choice of 46 options to Years 11/12 employing this
strategy).

In several instances, nearby schools have co-operated to bring about broader
subject offerings and to share resources. One review, for example noted:

Four local high schools have joined to develop a 2 unit Work Skills Course for non-
matriculation HSC candidates. This will involve one day per week placement and the
development of a work competency log.

Some concerns about timetabling structures are raised in a small number of
reviews, which note that the subject choice of students is limited by timetabling
structures. One review recommends:

That the school carry out an evaluation of the structure of the school day, the current
timetables, and the curriculum structure and offerings to ensure that these are meeting the
needs of the changing school population.

The changes outlined above do not suit all schools and the circumstances of all
students. For example, where students have long distances to travel, extended hours are
not always an appropriate response. Students comment that JSST attendance should be
organised so that they do not have to catch up work in school based subjects. Some
reviews highlight the undesirability of having to timetable 3/4 Unit classes before and
after school hours. Other reviews note that respondents express a desire for there to be
more out of school classes to allow wider subject choice and to facilitate re-entry student enrolment.

**Support for subject selection**

Review reports reveal consistent comments by teachers, parents and students about the importance of sound subject choices and careers advice for students about to enter the senior years. Many schools have recognised the need for more extensive and more individualised counselling for both parents and students than in the past. For example, a few schools have included home visits as part of their program in an attempt to accommodate parents who commute long distances to work or shift workers. Other examples are illustrated in the following extracts:

Considerable planning and energy has been expended on the production of several documents compiled specifically to assist students in the subject selection process and enable them to make informed decisions. Some students commented positively on this curriculum informational material and instanced particularly the Senior Curriculum Handbook 1994. The majority of students acknowledged and genuinely appreciated the supportive role of teachers in providing course advice.

In addition to written material on subject selection, the *Way To Go* program has operated for two years. This program aims to provide additional assistance and guidance for Year 10 students before they proceed into Year 11. Parent information evenings, directed at parents of students in Year 10, have been held in conjunction with this program. Most parents and teachers interviewed were of the view that the school has made very considerable efforts to inform parents of the subject choices in the senior school, and the options which are available to Higher School Certificate students. From comments made by staff during interviews it was apparent that some of them believed that despite the school’s efforts, prevailing community attitudes made some parents and students less receptive to alternative, non-traditional courses.

These extracts contrast with other findings in a few schools where it is recognised that Year 10 students would benefit from more information and experiential exposure to the range of senior curriculum courses to assist them in making appropriate career based subject choices.
Year 10 seminars were conducted to explain the Higher School Certificate procedures. Attendance by students and parents was limited. There was concern among some staff that some senior students are inappropriately placed in courses. Some students wanted more advice as to their abilities and directions.

Welfare needs of senior students

The expansion of post-compulsory education has highlighted a number of social issues which impact upon the school and its students. Increasing numbers of students are leaving the family home to complete their schooling and live independently with friends in shared accommodation relying on part-time work and AUSTUDY for support. This trend intensifies the poor motivation of many non-academic students with low self expectation already having trouble adjusting to individual learning and release from class time. Students from South East Asian countries often have families in Australia which are not complete, with important members still located in their country of origin. Given these factors and the on-going problems of class disruption from poor conduct and failure to complete work, schools have extended their welfare networks. There is now widespread support for tightening the procedures for monitoring student attendance and progress in the senior years. Some schools have extended their welfare programs into the community.

A stated priority of the school is to establish itself as a centre of learning for the whole community. This has been pursued through the encouragement of community tutors in the school and by facilitating the re-entry of mature-age students so that they can further their education at the school.

Against this background, secondary schools have intensified welfare programs for living skills, study skills, community liaison with support networks to raise student self-esteem, expectations and goals to enhance student career options. The senior 25 hour PD/Health/PE course was seen to be a valuable course by both students and staff.

Many Year 11/12 students have trouble adjusting to individual learning and managing their study when released from class time. Whilst on the one hand they
demand to be respected and treated as mature individuals, many are poorly motivated to the extent that one school review commented that:

A significant number of those interviewed indicated that post-compulsory students were not encouraged to develop personal and moral autonomy as essential characters of maturation.

Some senior students in some schools still exhibit poor attendance, lack of motivation and under-achievement. Review reports in a number of schools highlight these problems in relation to boys in particular, in association with a culture of peer pressure to not achieve.

**Summary and conclusion**

- Secondary schools have responded to increased enrolments in the post compulsory years through curriculum and organisational change including introduction of vocationally oriented courses, semesterisation, longer school days, extended teaching periods, flexible timetabling, reduced face-to-face teaching time and the establishment of independent learning centres.

- These initiatives are particularly directed at meeting the needs of the increasing number of students who are not seeking tertiary entrance as a result of their studies.

- Community perceptions of the value of the Higher School Certificate have not kept pace with the rate of change in the schools. Unrealistic student/parent expectations for Tertiary Entrance Ranking have generated real problems for some students because of inappropriate subject selection and failure to select appropriate and attainable career goals.

- Counselling and welfare support networks are placed under increased pressure to effectively address the social needs of post-compulsory students. There is a growing awareness of the importance of high quality career advice and guidance prior to Year 11, and for ongoing provision of guidance for students in Years 11 and 12.

- Some senior students in some schools still exhibit poor attendance, lack of motivation and under-achievement. Review reports in a number of schools highlight these problems in relation to boys in particular, in association with a culture of peer pressure to not achieve. This issue is discussed in more detail in a separate paper in this series.
Chapter 11

School effectiveness research: dead end, damp squib or smouldering fuse?

School effectiveness research: dead end, damp squib or smouldering fuse?

This paper begins by asking what we have learned from two decades of studying school effectiveness, then goes on to define a research agenda in the NSW context that can begin to answer some of the questions raised about the basic parameters of "effectiveness". Several intersecting lines of evidence suggest that the current paradigm of research into school effects may have reached a dead end. This evidence includes the results of evaluation studies demonstrating the failure of most projects based on the so-called correlates of school effectiveness; the repeated finding that only a small proportion of variance in student performance can be explained by schools; and more recent research pointing to the greater importance of class membership in explaining student performance.

Recent initiatives by the NSW government to implement performance-based funding and to provide the community with more extensive information about school performance raises the stakes in identifying effective schools to new heights. The task of identifying effective schools is not an easy one, either conceptually, technically or politically. Handled sensitively, the use of school performance information has the potential to contribute considerably to the improvement of schooling outcomes for students. Handled ineptly, the contribution of school effectiveness research will either be irrelevant, or create a conflagration all would rather have avoided.

Introduction

After more than two decades of research into school effectiveness, it is proper to question what it has achieved in terms of improving outcomes for students. Several intersecting lines of evidence appear to suggest that the current paradigm of research into school effects might have reached a dead end. The first of these lines of evidence draws on the conclusion of several large-scale evaluations of school improvement projects based on the correlates of presumed effective schools – that despite massive funding,
and zealous efforts by legislators, administrators and teachers, these projects have not generally led to any significant and sustainable improvements in student outcomes.

A second line of evidence is the continued finding that the size of the effect of schools on student learning is disappointingly small, with a recent meta-analysis of over 80 British and Dutch studies suggesting that, on average, only around 9 per cent of the total variance in student performance can be explained by the effects of attending different schools (Bosker and Witziers, 1995). Further, recent sophisticated multi-level research by Hill and Rowe (1994) and other suggests that most of even this small amount of variance is actually explained by differences attributable to membership of individual classes, not schools. They suggest that teachers, not schools, “make the difference” in student learning.

Underpinning both of these arguments has been the inability of research to produce consistent and unambiguous findings of school effectiveness across different domains of performance or over time. Despite the voluminous literature on school effectiveness, we are not much further advanced from the state of affairs described in Ralph and Fennessy’s (1983) critique: much of the literature takes the form of reviews of reviews, with only a small number of highly influential empirical studies providing the “evidence” cited in paper after paper. Reynolds (1992) concludes that in many ways we have come full circle back to the interpretation of the 1966 Coleman report – that schools don’t make a difference.

So, has school effectiveness research fizzled out? An overall evaluation of the available data on the size and stability of school effects leads to the conclusion (in line with Scheerens, 1992) that school effectiveness models are not as shaky as certain critics would have it, but at the same time not established as firmly as some enthusiastic school improvers would treat them. Recent initiatives by the NSW government to implement performance-based funding, to provide the community with “fairer school information”, and to establish “charter schools” (those identified as failing) raises the stakes in identifying effective schools to new heights. Governments elsewhere also continue to
legislate the publication of more and varied information about school performance. Given the consequences for a school of being designated effective or ineffective, it is fair to ask what is meant by those terms. How might we know if a school is effective? Do schools remain effective over longer or shorter periods of time? Are they equally effective in all areas? Are they equally effective for all their students? There would seem to be little point in attempting to identify schools at one point in time as "effective" in a global sense, if in fact such a notion is a false one.

The task of identifying school effectiveness is not an easy one, either conceptually, technically or politically. Handled sensitively, the use of school performance information has the potential to contribute considerably to the improvement of schooling outcomes for students. Handled ineptly, the contribution of school effects research will either be irrelevant, or create a conflagration all would have rather avoided.

The context for school effectiveness research

The history of the research into school effectiveness has been extensively described elsewhere (Reynolds, 1992; Brophy and Good, 1986; Scheerens, 1992; Creemers, 1994) and it is not necessary to go into it in any great depth here. It is sufficient to note that the major impetus for developments in North America and British research came about as a reaction to deterministic interpretations of findings by Coleman et al. (1966); Jencks et al. (1972); Plowden, (1967) and others that family and neighbourhood characteristics have a greater impact on student performance than individual schools. The popular (if incorrect) interpretation was that schools do not make a difference. Subsequent research (Rutter et al., 1979; Mortimore et al. 1988; Mortimore 1993), examining the relative progress made by students concluded that while background variables are important, schools can have a significant impact. Scheerens (1992) notes that many school effectiveness studies carried out prior to the mid 1980s were hampered by the limited statistical techniques available to them. The
development of more sophisticated techniques and software that allows greater separation of the effects of students and schools (Goldstein, 1987) has led to an explosion in the number of studies, conducted in a variety of contexts, on different age groups, and different countries which confirms the existence of both statistically and educationally significant differences between schools in students' achievements.32

However, the picture is not nearly as clear cut as this impressive body of evidence would indicate at first glance. Very recent research suggests that the notion that schools can be placed on a continuum from effective to ineffective may be inappropriate (see Hargreaves, 1995), and indeed that effectiveness itself may not be a unitary concept. There are questions as to whether schools are differentially effective for all of their students, whether they are equally effective across all curriculum areas, and whether they remain effective over time. To those closely involved with schools this would seem to be merely common sense, but empirical demonstration is another matter. The following sections of this paper examine the evidence relating to these issues in more detail.

Perhaps more importantly, recent work by Hill and his colleagues with data from Victorian primary schools suggests that when information about class membership is considered, the additional proportion of variance accounted for by the school is very small. There has not yet been widespread replication of this finding, other than a limited demonstration of greater within school effects than between school effects in an international test (Scheerens, Vermeulen and Pelgrum, 1989), although the possibility of greater class effects was raised earlier by Fredrickson (n.d). Despite the lack of

---

32 There have been several dozen, if not hundreds of individual school effectiveness studies conducted in the past decade. Many of these have had quite small samples or are limited to a particular education authority. Some of the more widely cited studies include the following: Reynolds (1976, 1982); Gray (1981); Edmonds (1979); Brookover, et al., (1979); Cuffanc (1987); Smith and Tomlinson (1989); Williams and Raudenbush (1989); Nuttwill, et al., (1989); Gray Jesson and Sime (1990); Daly (1991); FitzGibbon (1991); Jesson and Gray (1991); Stringfield et al., (1992); Goldstein et al., (1993); Sammons et al., (1994a, 1994b, 1994c); Thomas and Mortimore (1994); Thomas, Sammons and Mortimore (1994); Bondi (1994); and Hill et al., (1994).
replication, it is clear that the implication of this work is that we are faced by the same conclusions as researchers in the 1960s, that schools do not make (much of) a difference. This may be overly pessimistic, as Reynolds (1992) notes, we are considerably wiser for the journey, and it is clear that individual teachers can make a difference to student’s learning outcomes. Indeed there is cause for hope, because of all the variables that have been associated with effective schools, the quality of teaching has both the most consistently demonstrated impact on student learning and is within the power of schools to do something about.

It would seem necessary for future research to pay closer attention to the issue of teacher effectiveness for there to be significant advances in our understanding of what makes schools effective. Hargreaves (1995) discusses the need to consider the cultural dimension to school improvement. The question to be answered is this: is an effective school more than a collection of effective classrooms, or is there some cultural influence operating over and above the contribution of individuals. While resolution of this issue is beyond the scope of this paper, interested readers are referred to reviews of teacher effectiveness in Witrock (1986); Walberg, (1986); and Hopkins, Ainscow and West, (1994).

A further question to be asked is whether the findings of school effects research in other countries have any relevance to the Australian and in particular the NSW context. The aims, organisation content and delivery of schooling vary greatly from country to country. Few countries have the extensive private school sector that operates in Australia, or academically selective schools within the general education stream. What effects might these have for the interpretation of results? It has been suggested that the focus of much of the U.S. work, in particular, which has focussed on attempts to identify correlates of highly effective inner city, high minority enrolment schools, has very little relevance to the majority of schools in countries such as Australia and the Netherlands.
Research by McGaw et al. (1992) has suggested that Australian school communities have little appetite for the narrow focus on multiple choice tests of literacy and numeracy that is the standard fare in many North American studies. Instead, the outcomes most highly valued by Australian communities included less tangible attributes such as the development of a positive self-concept, a sense of self-discipline and self-worth, becoming a productive and confident member of the adult world and the development of appropriate value system.

Whether this reflects a more sophisticated understanding of the purposes of schooling or simply the lack of a tradition of testing is hard to determine. Surprisingly little is known about the parameters of performance of Australian schools and school systems. There has been a long history of opposition to standardised testing and the public comparison of schools, witnessed by the opposition to the ASSP project in the 1970s (Keeves and Bourke, 1976; Bourke et al., 1981) and the continued opposition of the NSW Teachers Federation to the Basic Skills Tests (Byrne, 1997). While public examinations have been part of the educational landscape in NSW since the establishment of public schools more than a century ago, only the most cursory information has been publicly available about the performance of schools on these examinations (an exception is Williams and Carpenter, 1987).

The most widely known information has been the number of students in the top 1000 places in terms of the Tertiary Entrance Rank, a very imperfect indicator at best. The current NSW government promised as part of its election platform that it would publish “fairer school information” (Carr, 1995) but while prototype models of school annual reports have been disseminated for discussion (DSE, 1996), there has been no final agreement on what form these will take, and at the time of writing are the subject of a ban by the NSW Teachers Federation. If little is known about school performance in absolute terms, almost nothing is known about the comparative value added by schools of various kinds and in various locations over and above what might be expected given the intake characteristics of their students.
So, if effectiveness is to prove to be an elusive concept, is there any point in continuing? Has school effectiveness fizzled out or has it reached a dead end from which we need to retrace our steps for a while and try a new direction?

Schools directly or indirectly touch on the life of almost everyone in modern society, whether as student, parent, teacher, employer or consumer of the goods and services produced by school leavers. Education is a major undertaking of governments around the world. Schools account for a substantial proportion of public and private expenditure, averaging around 4 percent of GDP in OECD countries. The NSW Department of School Education is one of the largest employers in Australia, and has a budget of well over $3 billion annually. In return for this investment, high hopes are held for education as an instrument of social and economic policy for the betterment of individual, community and national well-being. It therefore should come as no surprise that there is intense interest in knowing whether schools are delivering value for money – how effective schooling is and how it can be improved (Hill, Rowe and Holmes-Smith, 1995).

The re-emergence of belief in “market forces” as the dominant economic model in countries including the U.K., USA, Australia and New Zealand in the 1980s has also forced schools and school systems, for better or for worse, to operate in a quasi-market environment in which they must actively seek to satisfy client expectations and compete for student enrolments and thus resources. They must not only be effective, but they must also market themselves on the basis of what they do especially well (Hill, 1995a).

Such interest is not new, but may have become more intense, as moves to “reconstruct” poor performing schools takes hold (see for example, North Carolina State Board of Education, 1997). The education indicators movement of the late 1980s (see OECD, 1994; Smith, 1988; Wyatt and Ruby, 1988) refocussed attention on the need for both educational accountability and improvement to be based on accurate, reliable and defensible collection, dissemination and utilisation of information. The measurement of
student outcomes as a reflection of school effectiveness is an essential and integral part of such information systems. Hill (1995a) notes that the need for reliable information and measurement has been understood for some time by those in industry and business, and the message is becoming increasingly clear within education. This is not to say that it is yet universally accepted across the education sector.

Reynolds and Packer (1992) argue that several converging forces make it likely that the need for research and development in the general area of school effectiveness and school improvement will be even greater in the 1990s than it has been thus far. They hypothesise that in addition to the increased pressure for education systems to demonstrate results, school systems are likely to become more heterogeneous in quality, encouraged by policies that promote greater school differentiation (such as greater school choice and the establishment of more specialist and selective schools). Policies encouraging greater self-management will also mean that schools are more dependent on their own resources, which are not equally distributed. They believe that the likely result of these policies, in the short term, will be a substantial variation in the quality of schools, since the common factors provided by districts and local education authorities to all schools are simply being removed.

In addition, the nature of the current school population is changing, firstly as a consequence of policies of “mainstreaming” children with special educational needs (either as a result of physical, behavioural or learning problems), and also as a result of the retention into senior secondary schooling of a large number of students not necessarily interested in or inclined towards the traditional academic offerings. As these young people are highly sensitive to the quality of what they are offered within their educational setting, the influence of schools is likely to increase (Graham, 1988). Those that adapt best to working with a diverse range of students will be demonstrably more effective than those that cling to the ways of the past.

Slightly further into the future, demographic changes mean that governments in all major industrialised nations will be faced with greatly reduced cohorts of young
people leaving school as a consequence of dramatically reducing birth rates in the 1960s and 1970s. In the United States, each retired person is currently supported by 17 workers. By 2010, if current trends continue, the proportion will be reduced to 1:4, and one of these four will be Hispanic or black, groups not traditionally well served by the schools of today. Assuming the demand for labour remains constant, no society in the future will be able to tolerate the large number of students who “drop out” or leave schools without formal qualifications as at present.  

In these circumstances, the need for governments to be able to identify effective schools and to take active steps to encourage greater effectiveness becomes self evident.

**What have we learnt from two decades of school effectiveness research?**

Perhaps the most important outcome of research into school effectiveness has not been the finding that schools can influence their students’ learning, or that some school related factors seem to lead to better student outcomes than others (Mortimore et al., 1988; Bosker and Scheerens, 1992). Murphy (1992, pp. 94–96) identifies four aspects that he believes are the real legacy of the effective school movement:

- **The educability of learners.** At the heart of the effective schools movement is an attack on the prevailing notion of the distribution of achievement according to a normal curve. There is a clear demonstration that all students can learn.

- **A focus on outcomes.** For a variety of reasons, educators have tended to avoid serious inspection of the educational process. Effective schools advocates, however, argued persuasively that rigorous assessments of schooling were needed and that one could judge the quality of education only by examining student outcomes. Equally important, they defined success not only in absolute terms, but as the value added to what students bring to the educational setting.

- **Taking responsibility for students.** The third major contribution of the effective schools movement is an attack on the practice of blaming the

---

33 Similar sentiments are expressed in the Commonwealth Government’s major education policy document *Strengthening Australia’s Schools* (Dawkins, 1988).
victim for the shortcomings of the school itself. It means an end to the philosophy of “I taught them but they didn’t learn”. The movement has been insistent that the school community takes a fair share of the responsibility for what happens to the youth in its care.

- **Attention to consistency throughout the school.** One of the most powerful and enduring lessons from all the research on effective schools is that the better schools are more tightly linked – structurally, symbolically and culturally – than the less effective ones. They operate more as an organic whole and less as a loose collection of sub-systems. An overarching sense of consistency and coordination is a key element that cuts across the effectiveness correlates and permeates our better schools.

The legacy of the effective schools movement as outlined by Murphy (1992) leads into the related subject of school improvement. School effectiveness research may have intrinsic interest for some, but is ultimately of little value unless it produces something of policy relevance that can help to make schools better. Likewise, Murphy contends, school improvement efforts that lack substantive content, or focus on single curriculum initiatives or isolated teaching practices rather than whole school development are doomed to failure.

For some time, research into school improvement was seen as a somewhat different discipline to school effectiveness, but of late there has been a drawing together of the two traditions (Reid, Hopkins and Holly, 1987; Reynolds and Creemers, 1990; Levine, 1992). Many school effectiveness researchers are profoundly concerned with the implications of their research for policy makers, schools and students, indeed many are employees of school systems (e.g. Nuttall 1989; Fetler, 1989; Webster et al. 1994) or have been commissioned by school authorities to carry out work on their behalf (e.g. Schagen, 1994; Thomas et al., 1994). The work of the ALIS centre at the University of Newcastle (U.K.) is an example of how the results of methodological and theoretical advances have been put into practice for the purposes of informing school improvement (Tymms, 1995).

The mechanisms by which the results of school effectiveness research are put into practice have also been the subject of considerable research (Fullan, 1991; Louis and
Miles, 1991). There are some important lessons to be learnt from these reviews. While an extended discussion of these is beyond the scope of this paper, it is important to note that there is little support for those who would apply the findings of school effectiveness research mechanismically, without reference to a school’s history and context. Rather, the approach advocated by Sammons et al. (1995) and adopted by educational authorities in Scotland, the U.K. and NSW, is that they can be a useful starting point for school self-evaluation and review. While many examples of the deleterious effects of testing programs based on minimum competency testing have been identified (eg. Madaus 1988), it would also appear that the worst fears of those who vehemently oppose the publication of student results seldom come to pass.

A large number of initiatives around the world have attempted to put into practice the findings of the school effectiveness research. Mann (1992) summarises this work as follows.

Since the early 1970s, school practitioners and education researchers in Great Britain and the United States started to document the ‘within-school’ characteristics that could ‘add value’, that is, help children achieve over and above what would be predicted given their family backgrounds, chiefly social and economic status.

Effective schools advocates believed that enough was known about the best teaching practices to help children from low-income families learn the same basic things as other, more privileged students. The basic assertion of effective schools advocates was that ‘compensatory education’ was possible, that there was a set of practices that did not depend on extra money or new grants of authority. If those factors were maximised, simultaneously and persistently, then the neediest children would be able to acquire the same basic skills (literacy, numeracy) as their more privileged peers. Various researchers have created different lists of these factors. Most can be resolved into the five used by the late Ron Edmonds:

1. Strong leadership at the building level.


3. An organizational climate that supports good work by teachers.
4. Curriculum that fosters an "instructional emphasis" or an "academic press."

5. A pupil progress measurement system that is geared more to the next lesson's teaching than the next grade's promotion.

That model continues to be widely used (and widely resisted) in state and local school reform movements around the U.S. and abroad. Some advocates (including this author) have argued that the preferable but unused knowledge base created a moral imperative that should compel teachers to change their teaching. The rhetoric was close to both the goals of a democratic public school and a profession of teaching. The goals lost (p. 224).

The outcome of these efforts, the Hudson Institute concludes, has been a "$500 billion flop". In the U.S., the average reading achievements of 9, 13 and 17 years old students tested in the National Assessment of Educational Progress have not increased over the last 17 years; the national trends appear as flat lines. As Mann describes it:

If those lines were an electrocardiogram, the doctor would go talk to the family. Since they measure education, school people blithely substitute excuses for action: the tests count the wrong things, those are someone else's kids, and so on.

Less emotively, the 1985 Commonwealth Quality of Education Review Committee (QERC) in Australia concluded that assessing effectiveness of schools was constrained by the absence of unanimity of what students should achieve, the lack of effective measures of achievement across the spectrum of educational objectives and the difficulty of separating the effects of schooling from those of the complexity of social processes experienced by learners. QERC noted that physical provision for schooling was qualitatively better than before, and that the qualifications of teachers were higher, but it was unable to provide evidence that the cognitive outcomes of schooling had become better or worse.

Chapman (1988) describes the outcomes of a large scale school improvement effort in Victoria. By the end of 1986, approximately half of Victorian Government schools had joined the School Improvement Plan (SIP). A review conducted in that year found that the overwhelming majority of schools (93%) were satisfied with their overall
progress in school improvement. However, the researchers undertaking the study (Robinson and Rocher, 1987 in Chapman, 1988) concluded from the evidence available “the link between school improvement and student outcomes appears rather weak”. The State Board of Education concluded that the SIP was limited because it paid relatively slight attention to questions of comparisons of schools performance and to thinking about questions of efficiency, effectiveness and cost effectiveness. Nor did it meet the problems of detecting and dealing with schools performing unsatisfactorily.

Not surprisingly, the school effectiveness movement and school effectiveness research has come under considerable criticism from Australian academics and policy analysts. Questionable methodological procedures, narrow concepts of effectiveness, the emphasis on standardised achievement, the danger of recreating the dream of the efficient one-best system of instruction and the conservative and simplistic prescriptions for effectiveness, improved standards and excellence are identified as contributing to a movement which is socially conservative and educationally regressive (Angus, 1986).

Measuring school effectiveness

How the effectiveness of schools can best be measured and reported is a question that perhaps has no single correct answer. One difficulty has been that there is little agreement on what is meant by effectiveness. Reid, Hopkins and Holly (1987) concluded that while all reviews assume that effective schools can be differentiated from ineffective ones, there is no consensus yet on just what constitutes an effective school.

Most current definitions have in common a focus on student outcomes, and in particular the concept of the value-added by the school (McPherson, 1992). This focus implies that a school’s performance is to be judged not on results alone but on the school’s contribution to these results. The definition adopted by an international study of the quality in schooling by the OECD (see Chapman, 1991) encapsulates these elements:
An effective school is one that promotes the progress of its students in a broad range of intellectual, social and emotional outcomes, taking into account socio-economic status, family background and prior learning (p.1).

A study of the perceptions of what constitutes an effective school by the Australian Council for Educational Research (ACER) found that rather than the narrow concentration on test results in literacy and numeracy commonly found in overseas studies, Australian school communities valued most highly the following:

- positive relationship with learning
- development of a positive self-concept
- sense of self-discipline and self-worth
- student’s living skills
- becoming a productive and confident member of the adult world in time.
- the development of appropriate value systems; and
- the preparation of the student for the next stage of learning (McGaw et al., 1992).

However, the bulk of current school effects research accepts an operational definition of an effective school as one in which students progress further than might be expected from consideration of its intake (Mortimore, 1991). In this paper the term value-added is used to mean the extent to which students may, over a given period, have exceeded or fallen below the expected progression from a given starting point. A value-added measure is one which attempts to describe the educational value that the school adds over and above that which would have been predicted given the backgrounds and prior attainments of the students within the school (Hill, 1995a).

Wyatt and Ruby (1988) reviewed a number of different strategies used to construct indicators of school effectiveness, including comparisons against standards, comparison of actual against expected scores, data envelope analysis, gain scores and cluster analysis. Cuttance (1987) and Gray et al. (1990) have also reviewed a number of models for reporting the value-added by schools. Of these, models based on the statistical technique of regression have been the most widely used. There is a general
consensus in the academic literature that multi-level regression models (that is, those that take into account the hierarchically nested nature of the data) provide the best estimates of the sources of variation in performance between individual students, class, schools or higher units of aggregation (see for example, Goldstein, Bryk and Raudenbush, 1992; Hill et al., 1994).

However, education authorities have, in the main, tended to opt for the use of simpler linear regression models. The reasons for this are not clear. It may be that the spread of the knowledge and software needed to operate multi-level models in practice has lagged well behind their theoretical development. But considerations for the need to present information in more accessible form to audiences broader than highly trained specialists and academics may also play a role. The U.K. Schools and Curriculum Authority (SCAA) Report on Value-Added argues that the models used must be as simple and straightforward as possible while still maintaining accuracy. They recommend that, in the first instance, a simple linear regression of student intake scores on student outcome scores on two public examinations be used to derive an effectiveness score. Thomas and Goldstein (1995) on the other hand, argue that empirical interpretation of value-added is anything but simple and straightforward. An example of the presentation of value-added results which avoids the league table approach is shown in Figure 1 below.

Presentation of value-added results

Bosker and Witziers (1995) identify three kinds of value-added measures. The first, Hill (1995) terms "Unpredicted Achievement", described as the level of attainment of students adjusted for family background characteristics, such as socio-economic status or ethnicity, and student general ability (such as measured by IQ tests). The

---

34 See Salganik (1994) for a discussion of the methods used to identify effective schools in ten US state accountability schemes.

35 General ability measures and standardised test instruments have been found to be poorer predictors
second value-added measure might be termed “Learning Gain”, which is defined as the level of achievement of students adjusted for their prior levels of achievement. The third is “Net Progress”, namely, level of student achievement adjusted for family background, ability and initial level of achievement.

Of these options, the Net Progress or fully specified model is generally accepted to provide the most accurate measure of value-added (see Sammons et al. 1994b). However, even within these fully specified models, prior achievement has been consistently identified to be the most important determinant of later achievement. Gray et al. (1990) reviewed a large number of value-added studies and found that correlations between examination outcomes and student’s social backgrounds were typically around 0.35 whereas the correlations between examination results and finely differentiated prior attainment was typically about 0.7.

**Figure 1: Value added results compared to raw results**

![Image of a scatter plot showing the relationship between value-added scores and raw scores. The plot illustrates the distribution of data points across different levels of raw scores and value-added scores.]

of later achievement than curriculum relevant tests (see Madaus, Kelleghan, Rakow and King, 1979).
The SCAA review identified potential input variables and listed them in order of usefulness in predicting later achievement.

1. The most useful are finely differentiated measures of prior achievement.

2. Grouped measures of prior achievement.

3. Social background variables for each individual student.

4. Average social or academic background of the school population.

5. Social and other characteristics of the neighbourhood in which the school is situated or the catchment area from which its students are drawn.

Thomas and Goldstein (1995) argue that while value-added measures are clearly preferable to raw results as indicators of the progress pupils make in schools, league tables or rankings of value-added scores are no better than rankings of raw results. They claim that it would be irresponsible to publish league tables without taking account of the accuracy of the data. Goldstein et al., (1993) and Goldstein and Healy (1995) demonstrate that about two-thirds of all the value-added comparisons between pairs of schools are too imprecise to provide a fine separation of institutions in terms of GCSE results. This is illustrated in Figure 2. Schools can be judged as statistically different (at the 5% level) only when the bars for a pair of schools do not overlap. Goldstein argues that it only makes sense to identify schools at the extreme. This in fact has been standard practice in many systems around the world – schools are typically banded together in broad groupings, although the basis for the grouping differs from place to place. The SCAA report in the U.K. for example, recommends reporting school scores in terms of placement within quartiles, while in California (Fetler, 1989) a comparison group is formed for each school on the basis of the 10 percent of schools immediately above and ten percent of schools immediately below the school ranked according to a composite index of student achievement and family background factors.
In practice, there are several reasons why single measures of value added in themselves are no more satisfactory than raw measures in representing school effectiveness. A good deal of recent research suggests that schools are differentially effective (Nuttall, et al., 1989; Sammons et al., 1994) that is, they enhance the performance of certain kinds of students, say those from high SES backgrounds, but not others. There is also evidence that schools may be differentially effective in some subject areas but not others (Nuttall et al., 1989; Sammons et al., 1993). Further, school effects may not be stable over even relatively short periods of time, that is, they may be effective one year, but not the next (see Thomas, Sammons and Mortimore, 1995; Teddlie and Stringfield, 1993). Tabberer (1994) concludes that the evidence of differential effectiveness brings the consideration of single-feature measures, such as those that underpin league tables, even further into question. The issue of differential school effectiveness is considered in greater detail below.

Stability of school effects over time

It would seem obvious that one of the fundamental requirements for a school to be judged "effective" would be that the outcomes achieved by its students did not
fluctuate greatly from year to year, but relatively little research has been conducted to examine the stability of individual school effects over time. Virtually all of the published research has been cross-sectional in nature. A few studies have examined student growth in academic achievement (Hoffer, Greeley and Coleman, 1985; Willms, 1984), but few have examined year to year variations in school performance.

The studies which have reported in this area have produced inconsistent findings. Early research suggested considerable stability over a period of years (Rutter, et al., 1979), but in more recent work it appears that schools can vary quite markedly in their performance even over a period as short as 2 to 3 years (Goldstein, 1987; Nuttall, et al., 1989). Nuttall et al. (1989) conclude that these results give rise to a note of caution about any study of school effectiveness that relies on measures of outcome in just a single year, or a single cohort of students.

Bosker and Scherens (1992) reviewed a number of studies, and found that estimates of the stability of school effects over time (as measured by between year correlations) were between .35 and .65 for primary schools and .70 and .95 for secondary schools. Some studies in the U.S. (Mandeville and Anderson, 1986) have found even lower correlations (around .10) for elementary school mathematics and language.

Gray et al. (1995) identified eleven studies which had collected data on more than one cohort of students. Four of these looked at British secondary schools. Of these, three reported inter-year correlations that were greater than 0.9 (Willms & Raudenbush, 1989; Nuttall et al., 1989; Sime and Gray, 1990), and one reported ‘middling’ correlation (between 0.5 and 0.9). One study of Dutch secondary schools (Bosker & Guldemond, 1991) also estimated high correlations while another (Roeleveld, de Jong & Koopman, 1990) reported middling ones. Correlations in primary schools studies appear to be lower. Mandeville (1988) and Rowan and Denk (1982) in the USA report low to middling correlations as did Blok and Hoeksma (1993) in the Netherlands.
The widely cited study by Teddlie and Stringfield in Louisiana (1993) followed up eight pairs of outlier schools over a number of years. It is not possible to describe their data in the same way as the other studies. They claim, however, the over the period of their study, four of the schools were stable and effective, four were stable and ineffective, five were improving and three were declining.

It may be that the inadequacy of the statistical models used in many of these studies may have led to an under-estimation of the stability of effects, since they did not separate out sample variance from true parameter variance. Willms and Raudenbush (1988) found that measurement noise can account for up to 80% of the observed year variance.

Instability of estimates of effectiveness over time may have several causes. First, the instruments used to measure student outcomes may not test the same thing in successive years or may not adequately measure what was intended. Second the statistical methods used may produce spurious results. Thirdly, the world may be inherently unstable, and the school effectiveness scores simply reflect this.

Willms and Raudenbush (1988) hypothesise that the explanation for the finding of instability may lie in the manner in which school organisational structure and classroom practices contribute to student performance. These factors can change over time as teachers and principals transfer from one school to another. The impact of staff turnover may be felt more keenly in the generally smaller primary schools than secondary schools, explaining some of the discrepancy in the estimates for the two sectors. Year to year performance may also vary depending on the success of local improvement initiatives or the manner and extent to which wider reforms are implemented. The extent to which school performance is influenced by wider social, economic and political factors is largely unknown. These factors may also vary across communities and change over time.
On the other hand, schools which have developed a positive school climate and in which a tradition of high achievement has been enculturated may experience greater stability than others. The question of whether school effects are stable has implications for accountability schemes based on the presumed ability to identify school effectiveness. Without an understanding of how much the variation between schools is attributable to school practice versus social factors and how much is simply random fluctuation, there is potential for misuse of indicators based on single year measures. Goldstein et al. (1988) point to the need for caution in interpreting school effects at any one point in time:

> It is clear that the uncertainty attached to individual school estimates, at least based on single years data, is such that fine distinctions and detailed rank orderings are statistically invalid.

Gray et al. (1995) propose that a new way of looking at the data is needed. They note that when school effectiveness researchers have been in a position to collect a second years’ data they have sought to replicate their previous findings. A lack of stability between years was interpreted as a threat to the validity of their findings. This orientation may have inadvertently inhibited development in this area. School improvement is essentially about creating change in levels of school performance. If we continue to view instability negatively we may be missing an important point.

For satisfactory studies of changes in school performance over time, Gray et al. (1995) argue that a number of factors must be brought together, including:

- measures of outcome and prior attainment on individual students;
- data on a minimum of three cohorts and preferably more;
- multi-level statistical analysis; and
- an orientation towards examining the data for systematic changes in school performance over time.

To test this framework, Grey et al. analysed data from three successive cohorts in 30 English secondary schools. They found that only a small proportion of the schools in the study (between a fifth and a quarter) were improving or deteriorating in terms of
their effectiveness. A particularly striking finding was that whilst several schools improved in effectiveness, only one initially ineffective school did so consistently. As with any pioneering work, the study is not without its critics. Tymms (1995) queries several methodological aspects, which require further exploration. A later section of this paper suggests some areas for research in the NSW context that can help to advance knowledge in this area.

**Are schools equally effective across the curriculum?**

As with the work on the stability of school effects, the research on differential school effects is both limited and contradictory. The early work of Rutter, *et al.* (1979), and Reynolds (1976) reported high intercorrelations between school’s academic effects and their social effects. Mandeville and Anderson (1986) investigated the intercorrelation between primary school effect indices for different subject areas (maths and reading) and found indices near .70. Brandsma and Knuver (1989) arrived at much the same figure (.72) in Dutch elementary schools. Cuttance (1987) reports correlations for secondary schools in Scotland of .47 and .74 for English and arithmetic respectively.

More recent work suggests schools may be differentially effective in different areas. Mortimore, *et al.* (1988) found substantial variations between schools’ effectiveness on one academic outcome, such as oracy (heavily school influenced), and another, such as reading (less heavily influenced). Smith and Tomlinson (1989) also report substantial variation in the departmental success rates of different schools at public examinations, with these differences being more than simply a function of the school’s overall effectiveness. FitzGibbon *et al.* (1990) report similar findings between English and Mathematics departments.

Nuttall, *et al.* (1993) using data from the ILEA Junior School Project found that some schools were more effective in raising pupil’s performance in one cognitive area rather than another. Only a few schools in their sample had a marked positive or negative effect on both reading and mathematics. The project’s findings, they conclude,
is evidence that no simplistic division of schools into good or bad is possible, even on the basis of results in basic subjects such as reading and mathematics.

Hill, (1995) reports the only Australian data in this area. He found from the Victorian Quality Schools Project that the correlation between value-added measure for 51 primary schools in English and mathematics to be 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 are also differentially effective across the curriculum. Figure 3 illustrates the relationship between performance in English and mathematics. The possibility of replicating this work with the NSW Basic Skills Test is readily apparent.

**Figure 3: Relationship Between Value-added Measures of English and Mathematics Achievement**

Analyses by the Victorian Board of Studies using 1994 VCE data reveal the same phenomenon of differential effectiveness across the curriculum. Of the total of 190 schools, only 22 are in the top 10% for four or more subjects. Intercorrelations between subjects ranged from 0.32 to 0.71.
Figure 4 below illustrates how the relative performance of various curriculum areas within a school can be compared, when both the level of difficulty of the examination and the quality of the candidature for each course has been taken into account.

Only a small number of studies have examined whether schools are equally effective at promoting both social and academic outcomes. Reynolds (1976) shows small academic, but large behavioural and attitudinal differences in the same school. The ILEA study (Mortimore et al. 1988) showed that schools can be differentially effective with respect to their student’s academic and social outcomes. Given the emphasis many schools place on the social development of their students, it would seem to be an appropriate area for further investigation.

Figure 4: Relative Effectiveness across Curriculum Areas

Do schools have the same effect on all of their students?

Once again, the findings in this area are unclear. There is some evidence in the international literature that schools are not equally effective for different groups of students within the same school, such as students from different ethnic groups, ability ranges and socio-economic status. Nuttall, et al. (1993) using data from ILEA primary
schools found that effective schools appear to raise their reading attainment scores for all pupils irrespective of initial attainment level. Conversely, less effective schools seem to depress later attainment scores for all. Gray, et al. (1990) also found little evidence of differential effectiveness in their study of schools in a wide range of LEAs. This effect is shown in the figure below.

Figure 5: Plot of School Slopes Showing Predicted Year 5 Reading Scores

However, Aitken and Longford (1986) found that schools can differ in their regression slopes, suggesting that some may be more effective for pupils of a certain ability level that others, a finding supported by many others (Nuttall, et al. (1989); McPherson and Willms (1987); Willms and Cuttance (1985) Smith and Tomlinson (1989). In the Victorian Quality Schools Project, Hill found that in the case of primary school English attainment, girls make greater progress than boys, students from high socio-economic backgrounds make greater progress than students from low socio-economic status, and classes with a high proportion of non-English speaking (NESB) students make less progress than classes with low proportions of NESB students. For mathematics, the significant factors are also gender, student SES and NESB, but in this case girls make rather less progress than boys. To sum up these confusing findings, the following table drawn from Bosker and Scheerens (1992, p.749) is helpful.
Table 1: Range of Stability Estimates for School Effects

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th>Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across years</td>
<td>.35–.65</td>
<td>.70–.95</td>
</tr>
<tr>
<td>Across grades</td>
<td>.10–.65</td>
<td>.25–.90</td>
</tr>
<tr>
<td>Across classes</td>
<td>.45–1.00</td>
<td>–</td>
</tr>
<tr>
<td>Across subjects</td>
<td>.70–.75</td>
<td>.45–.75</td>
</tr>
<tr>
<td>Across criteria</td>
<td>.00–.05</td>
<td>.35–.70</td>
</tr>
</tbody>
</table>

What might explain these conflicting findings?

Levine and Lezotte reach the following conclusions about why there is so much variability in the school effects research.

Different achievement criteria and different data analysis methods frequently lead to conflicting conclusions about whether a school is effective or ineffective. Judgements depend highly on the grade and subject tested, and whether the test is norm referenced or criterion referenced.

The tendency to draw conclusions about school effectiveness from data derived from low-level learning tasks that are the most easily taught and measured by teachers and are the most easily memorised and regurgitated by students is one of the most distressing of these methodological uncertainties, according to Levine. This tendency is particularly evident in the many studies which depend mainly on outcome criteria such as primary grade test scores, which are generally limited in the extent to which they assess independent learning, conceptual application, thinking and higher order tasks and skills. Also, when scores are aggregated across grade levels, poor performance on higher order tasks in later grades can be masked by acceptable performance in lower grades. Similarly, when scores are aggregated across tests and sub-tests, mechanical skills such as punctuation, spelling and mathematical computation can disguise unacceptable performance on fundamental reading and maths skills such as comprehension and problem solving (Levine and Stephenson, 1987).
Identification of a school as effective is highly dependent on methodological variations, such as whether or not, and in what way socio-economic status measures are used to control for student's background. The regression models used are highly sensitive to the way data is aggregated and disaggregated. Judgements about effectiveness are sometimes made on the basis of only slight achievement differentials. When large sample numbers are involved, trivial differences in achievement can be statistically significant.

The criteria and analytic methods for classifying schools as effective or ineffective have been topics of considerable debate. Many studies contrast the differences obtained when value-added measures are used in place of raw results, but only a few studies have directly contrasted the implications of using alternative methodologies for calculating value-added. Marko (1974) was one of the first to contrast different methods of aggregating data, and found little difference in his sample. However, this research was limited because analysis using multi-level models had not been developed at that time.

The research of Goldstein, (1987) is considered to be the authoritative work in this area, but there is still scope for considerable work on how the complex results of multi-level models can be presented in an easily understandable form by lay audiences. The following figure illustrates how two different methods of calculating and reporting value-added can give very different impressions of the state of affairs in a school.

Figure 6 (a) shows the difference between expected and actual scores based on performance at the mean, whereas Figure 6 (b) shows the same value-added scores for students in the bottom quarter, middle half and top quarter. Figure 6 (a) suggests that the school is doing slightly better than would be expected, but when disaggregated as in Figure 6 (b) it is apparent that the school is doing much worse than expected for its lower ability students, but much better for its more able students.
An agenda for school effectiveness research

Several suggestions for further research have been mentioned above (see also Reynolds and Packer 1992). The databases relating to student achievement held by various school authorities within NSW offer a significant opportunity to contribute to the international knowledge base on effective schooling. There are a number of technical problems that need to be resolved before this potential can be realised, not least being the task of integrating disparate data sets to construct longitudinal records. There are also a number of legal, industrial and ethical issues that also must be addressed before such work can be undertaken. The agenda for further research sketched below assumes that such issues will be resolved in time.

Several recent initiatives, if they are successfully implemented, are of particular interest in terms of providing something new. Recording of information of student outcomes against curriculum profiles may provide much more information about student progress in a wider variety of subject areas and in greater detail than has typically been studied in the past. They also provide the possibility of measurement of the extent to which students can demonstrate mastery of a variety of skills rather than subject matter
knowledge, and thus better reflect the multiple goals of schooling expressed in school and systemic mission statements.

There is also much that more traditional analysis of Basic Skills Test and public examinations such as the School Certificate and Higher School Certificate can reveal about some of the enduring questions about a school effectiveness. The size of the samples, the breadth of curriculum tested, and the relatively long periods of time over which data have been collected are all superior to the bulk of studies reported in the literature, and provide the potential if not for definitive answers, at least answers with a high degree of confidence within the NSW context. Issues concerning the size of school effects, their consistency over time, their consistency across different kind of school output, their consistency for different types of pupil and the applicability of findings across international settings fall into this category.

What needs to be done? One of the shortcomings of the British and North American literature arises from its historical focus on identifying schools that have been effective in teaching disadvantaged youth. Most of these are inner city schools, which have no counterpart in many countries. We need research undertaken in more typical samples of schools.

Larger sample sizes are also needed. More studies of secondary school effectiveness are also needed, particularly since the literature on school effectiveness from the Netherlands and the United States concentrate heavily upon research in elementary schools.

Some British studies have been highly defective in their measurements of pupil intakes into schools, which may have led to invalid assumptions being made about schools or education systems being more effective simply because full allowance had not been made for the intake quality of their pupils. What is needed in the future is multiple indicators of intake, covering a range of pupil academic and social factors, as in the study by Mortimore et al. (1988).
The methodology of measuring ‘value-added’ also needs to be further explored. The early studies using ‘means-on-means’ analyses, where school averages for all pupils are used, as in Reynolds’ (1979, 1982) work, make it impossible to analyse the school experience of different groups of pupil and also have lower explanatory variance. Individual pupil level data rather than group data are now widely agreed to be necessary, both on intake and at outcome (Aitken and Longford, 1986), to permit the appropriate use of multi-level techniques of analysis, which can nest pupils within classrooms and classrooms within schools and the schools within the context of outside-school factors.

There is a need also to broaden the investigation of effectiveness to include social outcomes from schools, which may be independent of academic outcomes. Only a small number of studies thus far have included even very limited measures of behaviour and attendance. Not only do we not know what the parameters of effectiveness are in this domain are, we do not know how it interacts with other domains. Some outcomes may partially determine, as well as being partially determined by, the academic outcomes of schooling. It seems strange that almost every education system around the world places great emphasis on the importance of non-cognitive outcomes for students, but few appear to have ever attempted to determine how effective they are in this area.

Reynolds (1992) outlines some of this further work required. He notes that we are still not completely sure which processes are associated with effectiveness, and also how the school organisational factors have their effects – through their effects upon pupil self-concepts or by direct modelling, for example? We need to know what creates the organisational factors, which may require a degree of historical study since there are those who insist that what makes an effective school is in part the history of being an effective school. There is a need also, he argues to lift the level of abstraction from mere empiricism to a more conceptual level.

This list of research topics is not an exhaustive one. There is a great deal of qualitative investigation needed over and above the statistical analysis referred to above. The scope of the research agenda outlined above suggests that far from being a dead
issue, there is considerable work yet to be done in school effectiveness. The challenge is to ensure that this work leads to a better understanding of school performance and how it can be improved.
Bibliography


Byrne, R. (1997). Why We Should Ban the BST. *Education (Journal of the NSW Teachers Federation)*. Feb 17, p.15.


243


Chapter 12

Understanding Value Added

Forthcoming as No.1 in Corporate Performance Directorate Occasional Paper series
Understanding Value Added

Introduction

From the very beginnings of universal and compulsory education, one of the enduring concerns of educators, bureaucrats, politicians and the community has been how to best develop an understanding of the performance of schools. This concern has been felt world-wide, and in part underlies the establishment of systems of school inspection, adoption of common curricula and examinations, performance indicators, and more recently, quality assurance mechanisms. In the past 30 years or so, the focus of this understanding has shifted from considerations of the quantity of schooling provision to the quality of schooling outcomes.

Capturing the essence of these outcomes of schooling in a form which is fair, unambiguous and easily understood by all has proved to be a far from simple matter. By far the most common, but also the most frequently criticised, measure of school performance to date has been school average performance on an external test or examination. Sometimes what is reported is the percentage of students reaching a pre-determined criterion level of performance (such as percentage of students within a skill band or percent reaching a grade level norm).

There is widespread agreement in the academic research literature that the use of raw results of students alone is inadequate for reporting school performance and for making judgements about school effectiveness. Perhaps more importantly, such results provide little information to a school as to how it can improve its performance. Good examination or test performance may indicate more about the quality of the student intake of a school than about the effectiveness of the teaching and learning process.
The use of "league tables" which rank schools on the basis of their raw results has drawn extensive criticism, especially in the United Kingdom where this practice has been a central plank in successive Conservative government education policies (and is likely to be retained under a Labour government). This criticism has focused on both the fairness or statistical validity of drawing conclusions from these rankings, and on the "market forces" ideology underlying league tables as a means for spurring school improvements.iii

So, if raw performance figures and league tables are as inadequate, misleading or even potentially destructive as measures of school performance as their detractors maintain, does this mean that we should abandon the quest for knowledge about how schools perform, and what contributes to this level of performance? Should we allow a continuation of what Professor David Reynolds describes as "hiding behind our shields of ignorance", where poor performance is passed off by schools as "just a bad year" or perhaps more insidiously, "what you would expect from kids in this neighbourhood"? Is it impossible to find a fair, and statistically valid and reliable means by which school performance can be measured and reported?

The answer to all three questions is a resounding "no"! What we are really concerned with when we speak about school performance is school effectiveness, since our major concern is with identifying what the school has contributed to student achievement, over and above what the student already knows or can do as a result of their innate talents and abilities, or the opportunities for learning provided by the student's family, peer group or community. A considerable literature exists which points the way forward, documenting thousands of studies into school effectiveness, school improvement and educational change.iv This literature demonstrates conclusively that schools within any given education system do differ in the degree to which they contribute to the outcomes achieved by their students, and that there are certain practices which are more effective than others in promoting these outcomes. In other
words, some schools are more effective than others, and within individual schools, some classrooms are more effective than others.

The starting point for identifying the contribution of schools to student performance thus derives from the definition of an effective school. While there are many such definitions, common to most is the belief that an effective school is “one in which students make greater progress than might be expected from a consideration of its intake” (Mortimore, 1991).\(^\text{36}\) The extent to which aggregate student attainment differs from what is expected given consideration of its intake is often referred to as the “value added” by schools.

While the term “value added” has readily slipped into the lexicon of educators, and heads will nod knowingly at its mention, few really understand how it is derived and why this derivation is valid. This paper attempts to explain the concept of “value added” in simple terms. The paper concludes with some comments about the legitimate use of value added measures in school self-evaluation.

The concept of value added

Professor Peter Hill, in his inaugural lecture on school effectiveness and improvement acknowledges that “... methodologically speaking, measuring school effectiveness is immensely difficult” (Hill, 1995). Schools are often large, complex organisations. As the British Government’s Task Group on Assessment and Testing notes, “a school’s performance may only be judged by taking account of many aspects of its work, and many factors outside its control that affect this work”. Attempts to describe in simple terms the concepts and methodology behind value added run the risk of appearing not to understand this inherent complexity within schools. A common

\(^{36}\) More correctly, an effective school is one in which students make greater progress than might be expected from consideration of its intake, given what we know about the relationship between student intake and later performance for students in all schools in the population of interest.
complaint about any attempt to describe schooling outcomes by way of indicators of performance is that they concentrate too heavily on academic outcomes, and further, that what is actually reflected in these academic outcomes is an even more narrow range of academic outcomes. It is undeniable that the goals of schooling, particularly in NSW, as reflected in syllabus and curriculum documents, Departmental policies, and school mission statements, are far broader than the acquisition of subject knowledge alone. Most teachers would also acknowledge that the assessments leading to the award of grades at School Certificate and Higher School Certificate level do in fact reflect more than subject knowledge.

No sensible discussion of school performance would conclude that one single measure (whether it is of value added or otherwise) can adequately sum up all there is to know about a school. A basket of measures (or a set of indicators) is obviously called for, and the real question is what measures should be included in this basket. All things can be quantified somehow, if we so desire, but not all things should be, and the non-cognitive achievements of students may fall into this category. The achievements of sporting teams, musical groups and so on are more easily described in words than a contrived numerical indicator. On the other hand, the acquisition of knowledge skills and attitudes, which are important school outcomes, can be measured by well constructed tests and can be quantified and reporting accordingly."

The School Curriculum and Assessment Authority (SCAA) in the U.K. has reached a similar conclusion, and it is worth reproducing this conclusion at length to reinforce the point made above:

"... no study involving value added measures can ever give a full description of the complex situation which it is attempting to describe. Such measures should therefore always be considered in conjunction with other kinds of evidence concerning the performance of the school. Moreover, it is also important to recognise that value added analyses, as is true of all quantitative analyses, can only be as good as the quality of the data on which they are based.

It also needs to be recognised that there is much more to education than the examination and test results on which value added measures are based. Nevertheless, examination and test results are important. They matter to pupils as measures of educational attainment, and in the case of public examination results, as passports to chosen occupations and careers. They matter to schools and society because they can provide impartial, external
and independent indicators of educational standards both for individual pupils and for schools. Moreover, examination and test results are compatible with and can facilitate other important and worthwhile educational goals. There is therefore an important place in the public debate about education for serious analyses of examination and test results for schools and the performance indicators which can be derived from such results. (SCAA, 1994, p.14).

Measuring value added

From the definition of effective schools given earlier, there are several possible approaches to measuring value added, each attempting to answer slightly different questions, such as:

- "How far did students in this school progress, in comparison to similar students in other schools?"; and
- "What was the unique contribution of the school to the level of attainments reached by its students (i.e., once other factors affecting student attainment have been taken into account)?"

The basic methodology for answering both these questions is essentially similar. Hill (1995) refers to these two measures as "learning gain", and "unpredicted achievement" respectively. He describes a third measure as "net progress", which is the level of attainment adjusted for family background, ability and initial level of attainment. vi

All value added measures attempt to build a model which allows us to both explain and predict, student performance. We can anticipate that many factors may influence student performance, including:

- *those associated with the student*: their interests, motivation, preparedness, general ability and so on;
- *those associated with the student’s family*: their encouragement of the student, the resources they have to support him/her (such as number of books in the home), their knowledge of how the system works, and so on;
- *those associated with the school*: its level of resources, the breadth and quality of the curriculum offered, the school ethos and culture, the focus on achievement, and so on;
- *those associated with the student’s classroom*: the quality of the learning experiences offered, the classroom climate, and so on;

254
• those associated with the student's community: the kinds of resources available to support learning, the regard held for education, the opportunities for students to benefit from achievement and so on.

These are only some examples. We can also speculate that many other factors, such as peer groups, churches, the media and the wider education system may also influence student learning. Many studies have attempted to measure and explain the influence of all of these variables, to help explain why certain performance levels have been achieved and also to point the way towards policies that will encourage greater performance. The specific outcomes of these studies are not considered here but the methodology behind their conclusions is.

The statistical technique known as regression analysis provides the basis for estimating the relative contribution of each of these factors with student achievement. In the simplest case, we can use one factor, say a measure of student interest in a subject, and examine the relationship of this factor with achievement in this subject. This relationship can be shown in a scatterplot, such as the following hypothetical example shown in Figure 1. In the scatterplot, the results of each student in the subject is shown on the vertical axis, and the corresponding measure of interest on the horizontal axis.

**Figure 1: Simple Regression Plot**

![Simple Regression Plot](image)
What the technique of regression allows is a test of the strength of the relationship between the two variables of interest (sometimes called the input and outcome variables). Regression techniques also allow a "line of best fit" to be drawn, which visually summarises this relationship. The equation which describes this line also provides the basis for the estimation of an expected value for achievement (outcome) for any given value of input. The extent to which any point in the scatterplot differs from the expected value given by the "line of best fit" (more correctly by the regression equation) is regarded as being that portion of performance which is not explained by the input measure. This is shown in the following diagram.

**Figure 2: Calculating Regression Residuals**

The distance A–X is that part of student achievement not predicted by student interest, that is, performance which is attributable to some other factor(s). Distance A–X is called the residual value in the statistical literature. In this case, the residual has a positive value indicating that the student’s actual performance exceeds their predicted performance.

The example used in this illustration has considered only one input variable (in this case, student interest). Modern statistical packages allow very easy computation of multiple input factors (in the current example, three input measures may be considered, say student interest, hours of homework completed each week, and student’s socio-
economic status). We can then assess whether including additional factors, or combinations of factors provides a better explanation of the outcome measure (that is, by increasing the precision with which the input factors can predict a given outcome measure). The more precisely we can predict the outcome, the better equipped we are to explore the factors contributing to that outcome.

When computing regression models, there are several important points to note. First, the relationship between input and outcome factors need not be linear (represented by a straight line of best fit as in the above example). In real world situations, and in education in particular, few relationships are linear, as demonstrated in the following example.

**Figure 3: Curvilinear Relationships**

If we are measuring, say, proficiency on a test of physical activity, it is likely that performance levels will increase with greater practice prior to the test, but there is a point at which more practice does not help, but actually begins to detract from performance as fatigue sets in. The fact that the relationship is curved (curvilinear) and not linear, makes no difference though, to our ability to predict performance on the basis of knowledge of the input factor (practice in this case). Simply, the equation for generating the predicted values is more complex.
Second, it should be noted that the input and output factors in the regression equation do not have to be measured on the same scale. In fact, they do not even have to be measures of the same thing. (Remember, what we are trying to do is to examine the relationship between variables and use this knowledge to explain performance). This point will be reiterated again later. Value added does not necessarily measure student growth or progress directly.

Third, while regression may be used to explain performance, it does not imply that one factor causes an effect found in another factor, even when there is a strong or positive relationship between the factors. It may be that another factor, connected with one of the factors, but not measured or included in the model is the real cause of the effect. There may be an apparently high correlation between hours of instruction and outcome, for example, but it may be the degree of engagement in learning rather than the passage of time that is the important factor.

Fourth, as with all measurement, the estimation of the expected value for the regression equation and thus the residual distances has a certain degree of error associated with it. This degree of potential error can also be accurately estimated. While it is usual to speak of residuals as a single value, strictly speaking what is implied is that the true value for the residual is likely to be somewhere in a zone defined by the estimated residual plus or minus a margin for error.

The practical significance of considerations of standard errors comes when there is a need to make judgements between two or more observations. Unless this estimate of the amount of error is taken into account, we cannot have absolute confidence that two apparently different scores are not in fact essentially similar. It is partly for this reason that Professor Harvey Goldstein, one of the world’s leading authorities on the measurement of school effects, argues that single measures of value-added, and league tables of value added scores are just as useless as league tables of raw scores.\textsuperscript{16}
How can this technique be used to estimate the value added by schools?

The hypothetical examples discussed above have attempted to explain some of the fundamentals of how regression techniques can be used to predict and explain individual students’ results. The same principle can be applied to predict and explain a school’s results.

In essence, the “value-added” attributed to a school is calculated in the same way, that is, it is the difference between the predicted score for the school (based on the relationship between inputs and outcomes for each individual student in all schools) and its actual score. For convenience, it is usual to calculate these scores at the midpoint for schools, but there is no reason why this cannot be done at any other point in the distribution. (In fact, in the models proposed for NSW, it is recommended that value added scores be calculated at three points representing low, average and high performance: the 15th percentile, the 50th percentile and the 85th percentile, in recognition that schools can have different effects on students of different abilities). These scores can be displayed graphically in the form of column charts for ease of interpretation, as in Figure 4 below.

**Figure 4:** Sample Value Added Display
The most recent advances in statistical methodology for examining the relationships between variables, called multilevel modelling or hierarchical linear modelling, allows even more precise estimation of the unique contribution of schools to student achievement. The mathematics behind multilevel modelling is complex, and interested readers should consult Bryk and Raudenbush (1992), and Goldstein (1987) for a technical discussion. It is sufficient to note here that multilevel models take account of the complex nature of the organisation of education, where students are located within classes, classes within schools, schools within districts and so on. Multilevel models allow the unique influence of each of these levels to be identified. The calculation of multilevel models is complex, but the interpretation of the results is no different in principle to that of the simple model proposed for NSW.

Developing appropriate models of value added in NSW

A vast number of studies have attempted to identify and quantify the factors which contribute to student and school achievement. The results of these studies can be summarised as follows.

- The most powerful predictor of student performance is a finely differentiated measure of prior performance (such as a scale from 0–100).
- Curriculum based measures provide better predictors of school achievement than standardised tests or IQ tests.
- Broadly based, aggregate measures (ie those based on results in a number of subject areas) produce more stable results than do single-subject measures;
- Grouped measures of prior attainment (such as grades A–E) are the next best input variable, if finely differentiated measures are unobtainable;
- Social background characteristics at the individual student level are weaker predictors of outcomes than prior performance measures, but can be used when prior performance is not available;
- Average social or academic background of the school is far less preferable than individual student level input data (DSP status would be an example of this); and
- The least helpful type of input variable is the social or other characteristics of the neighbourhood in which the school is situated or of the catchment area from which its students are drawn (such as census based data).
The studies reviewed typically found correlations of about 0.35 between examination outcomes and social background variables, whereas correlations between finely differentiated prior attainment measures and examination outcomes were typically about 0.70.

These findings have been tested against data available in NSW (the results of some of these analyses are reported in the School Accountability and Improvement Technical Paper). As an illustration of the kind of relationship between prior performance and later performance in NSW schools, Figure 5 below shows the relationship between a sample of school mean aggregate scores at Year 10 and Year 12 level for 1994. The correlation between these two measures is 0.78. The correlation between the scores of the 1994 School Certificate cohort and their Year 6 BST aggregate was 0.74, and for the Year 6 cohort and their Year 3 performance was 0.80.

The results of these analyses are consistent with those found overseas. When more complex models were tested, including examining the interaction of academic and social background factors (such as gender, NESB status, and socio-economic status), it was found that the addition of these other variables adds only minuscule amounts of additional explanatory power to the model. One reason why this might be so is that whatever influence these other factors may have on achievement has already been included in the prior performance measure. For example, gender and SES influences prior attainment and because of this it is inherent in the more simple models.

In the interests of keeping the model as simple as possible while maintaining statistical validity, these findings suggest that the best model to apply to NSW secondary school data is one in which the value added by schools is defined by the relative outcomes achieved by students given their performance on a prior measure of achievement. The NSW Department of School Education does not have comprehensive and reliable data on the socio-economic status (SES) of individual students, and to collect this data would be expensive and intrusive. Given that when good measures of prior attainment are available, very little is added by inclusion of social background
factors such as SES, there seems little justification for collecting this information for this purpose. What we do have are very good measures of the academic performance of entire cohorts of students at different points in their school careers. Few education systems around the world are in this fortunate position.

**Figure 5:** 1992–94 Cohort School Mean Scores

**Summary and conclusion**

This paper has attempted to present a simple explanation of the derivation of one class of measures of school effectiveness – those attempting to capture the value added to student learning by schools. The extensive literature on this topic has been distilled with a view to identifying what might be an appropriate direction to be taken in NSW. It has been argued that a model which uses finely differentiated measures of prior achievement to explain later performance is the preferred option. As described in the School Accountability and Improvement Model *Technical Paper*, this can be operationalised using a scale derived from the School Certificate reference tests to create the prior performance measure, and using a scale derived from the Higher School Certificate course scores as the outcome measure. In doing so, it is important to note that what is produced is not a measure of the growth in students between Year 10 and
To understand student growth requires both the later and prior measures of performance to be on a common scale.

Value added is about *relative* performance, once other background factors have been taken into account. The measurement of student growth in learning in a particular school is about quantifying how much the students have increased their mastery of certain bodies of knowledge and skills. The value added by schools is about the relative amount of progress that students have made compared with what students of the same backgrounds have achieved.

Value added measures have an important place in school self evaluation, when considered in conjunction with a range of other data. As with any other kind of performance data, they should never be used on their own to form a judgement about a school. Value added scores should be used as indicators of whether students in a school are broadly performing above, below or close to the level that would be expected, rather than as precise measures for which specific numerical improvement targets can be written.

They are most useful when they are used to contrast differences in performance between sub-groups within a school, such as between students with differing levels of initial ability and gender. Examination of the stability or change in value added scores over time can also provide valuable information about the kind and scale of change necessary to bring about radically improved outcomes for students.
Bibliography


Chapter 13

Recent Developments in Annual School Performance Reporting in New South Wales

Recent Developments in Annual School Performance Reporting in New South Wales

Introduction

Annual school performance reporting is a relatively new experience in Australia, but has a longer history in other countries such as the U.K., France and the U.S. (see for example, Fitz-Gibbon, 1992; Jaeger, et al., 1993; Le Guen, 1994). Public reporting of school outcomes is intended to serve several purposes, including providing public accountability, informing parental choice, and informing school improvement efforts. However, the extent to which school reporting has, or can, produce substantial changes in outcomes for students has rarely been clearly demonstrated (see for example Tymms, 1995).

This paper describes the school accountability and improvement model that has recently been established in New South Wales (NSW), including the annual school reporting program, and discusses the reasons why it is believed that this model will be more successful than previous efforts.

The context for accountability and improvement in NSW

The NSW public school system is the largest in Australia, and one of the largest in the world. It has over 750,000 primary and secondary students, employs 66,000 teachers in more than 2,200 schools. The education budget in 1996 was in the order of $4 billion, or more than 20% of the entire state budget. The state education sector is the
largest employer in the country. One in eight NSW residents spends his/her day in a public school. Students in NSW public schools come from a wide range of socio-economic, language and national backgrounds. Around 76% of students complete the full 13 years of schooling and qualify for the Higher School Certificate, the primary entrance requirement for tertiary study.

Like many public and private organisations in Australia and around the world, the past decade has been one of constant reorganisation and restructuring for the NSW Department of School Education. Early restructuring efforts in the late 1980s saw the virtual destruction of the strong, centralised bureaucracy which had prevailed for more than a century, in favour of a “school-centred” system, guided and supported by 10 regional offices. In practice, the regional offices were themselves abolished in 1995. A newly empowered central office (although a much slimmer and geographically dispersed version of its predecessor) has been created, with 40 district offices established to provide curriculum, technology and administrative support directly to schools. It was intended that attached to each of these District Officers would be a School Improvement Officer, an appointment at a very senior level, who would have responsibility for implementing the Accountability and Improvement model in the District’s schools.

The latest restructuring aligns with current management thinking on organisational structures. Charles Handy, in his book *The Empty Raincoat*, sums up the role of the Head Office of organisations, as follows:

Small the centre should be, and partially dispersed, but it must be strong and well informed. The centre, after all, carries the ultimate responsibility for the whole. Its reserve powers typically include ‘new money’, ie the choice of strategic investments, ‘new people’, ie the right to make the key strategic personnel decisions in the group; the design and management of the information system, which is the artery of the organisation; and most controversially, the right of invasion when things go wrong. Only those at the centre have the view of the whole. They cannot run it, and should be too few in number to be tempted, but they can nudge, influence, and if they have to, interfere. The centre’s principle task is to be trustee of the future, but it needs to be sure that the present does not run out before the future arrives (1994, pp 118-119).
Accountability and improvement structures figure prominently in the new configuration, including the establishment of a Corporate Performance Directorate tasked with the implementation of school and systemic reporting mechanisms, development of the school and program evaluation schedules, and facilitation of research across the department. These tasks align well with, and provide vital support for the centre in Handy's conceptualisation.

In its 150 year history, the NSW state Department of Education (and its predecessor the Department of Public Instruction) has adopted several different approaches to accountability. A brief history of the various approaches, including the Inspectorate, school self-evaluation, performance indicators, and Quality Assurance school reviews, is discussed in Wyatt (in press). None of these, on their own, have proved satisfactory as systemic accountability measures, and all have eventually been abandoned for this purpose (although it must be noted that the inspectorate was in existence for more than a century before its ultimate demise). No government can tolerate an education system which does not have a strong accountability and improvement process if the system is to be counted as world class.

School Accountability and Improvement: A rationaleXIII, Xiv

Governments everywhere are increasingly seeking reassurance that the school systems they are paying for are delivering the results they want (however these are defined). Moreover, they are also concerned that there needs to be continual improvement in the standards of provision and attainment if students are to be equipped to fully participate in and contribute to an increasingly complex and changing society. In knowledge based economies, human rather than natural resources will be the key to increased economic prosperity. In many western countries, the 1980s and 1990s have seen strong pressure from governments to strengthen educational performance across the board.
On the whole, some researchers conclude that these reform efforts have had little success. (see Mann 1992, for example). The outcome of these efforts, the Hudson Institute concludes, has been “a $500 billion flop”. They note that in the U.S., the average reading achievements of students at the ages of 9, 13 and 17 years old have not increased over the past 17 years; the national trend lines appear flat.

These conclusions have, of course, been hotly debated, but it is worth questioning why we would expect that simply putting information in the public domain would bring about improvement. The notion of “holding people accountable” in the sense of requiring the resignation of those unable to meet performance expectations (in other words, performance contracting), has little currency in Australia. Neither has the idea that schools can be “shamed” into improvement been acceptable, as the largely negative reaction to one Sydney newspaper’s attempt to do so has proven (“The Class We Failed”, Sydney Daily Telegraph, January, 1997). The “market forces” so beloved by Thatcherite politicians, which are supposed to bring about improvement by forcing the closure of unpopular schools, have yet to exert themselves in this way in the decade since policies allowing school choice were introduced.

A fundamental premise underlying the model to be described here is that neither accountability nor school improvement efforts, on their own, will be sufficient to produce improved student outcomes. A successful process of school evaluation will include a balance between accountability and development and improvement processes. This dual focus recognises the importance of both pressure (through accountability processes) and support (through development strategies) in achieving change. The importance of this combination of pressure and support for effective school improvement is highlighted by Fullan (1991):

Pressure without support leads to resistance and alienation; support without pressure leads to drift or waste of resources. (p. 91).
A close examination of the literature on school reform and school improvement efforts suggests that the majority of the schemes described have attempted to achieve change through a concentration on one or the other of these two aspects. Few of the reforms of the 1970s and 1980s have attempted to combine the two. Other education systems have also adopted this change perspective, including New York State’s School Quality Review Process, with its commitment to “top–down strategies for bottom–up reform” (Olson, 1994).

A second premise underpinning the approach adopted in NSW is that successful school improvement will not derive from centrally prescribed mandates, but that successful change can only come from working with schools, rather than doing things to them. This premise also derives from a strong research perspective and a philosophy of empowerment (see for example Hargreaves and Hopkins). MacBeath (1996) concludes, from the perspective of the Scottish education system, that:

.... The way in which we foster quality and performance from within and support it from without is the real test of the education system’s integrity. A rigorous, realistic national framework of internal/external evaluation can achieve that in a way which gives to all stakeholders – school staff, parent, pupils, governing bodies, local authorities – a place in the process. (p.5).

In an era in which schools have been given greater autonomy and freedom to manage their own affairs, more sophisticated processes are needed to ensure that system standards and public expectations are being met. The trend is to set explicit targets for schools, but to allow them greater freedom, within budgetary and other constraints, as to how they achieve them. The monitoring and evaluation of performance of the individual schools against those targets is becoming more widespread and accepted in schools. Schools are now taking more responsibility for their own performance, are subjected to closer public scrutiny and are finding new ways of improving student outcomes in a world of more demanding standards.
Measures of both cognitive and non-cognitive student learning outcomes are needed to inform a comprehensive whole school evaluation. These outcomes need to reflect the complexity and breadth of the educational process and include student achievement of syllabus outcomes, performance in tests and examinations, attendance rates, students' post-school destinations, and students perceptions and attitudes. In recognition of the importance of within school as opposed to between school differences in performance, student outcome indicators are needed which identify the variations of performance of different groups within the school. This approach allows a more detailed analysis of performance within and between key learning areas and by different subgroups of students. A focus only on average school performance may mask important issues, particularly equity issues.

Key principles

The new emphasis on accountability demands that there is good quality, reliable data on schools. NSW has a considerable advantage over many other systems in this respect. Comprehensive and comparable student outcomes data is available over a long period of time from public examinations, at two points in secondary school, and statewide tests of literacy and numeracy at two points in primary schools. In spite of the well known inadequacy of examinations in summing up the real learning achievements of a student, or the performance of a school, they are currently the best performance indicator available particularly as they also act as passports to higher education and credentials for employment and are therefore seen as high-stakes by the students (OECD, 1995).

In addition, the planned introduction of an outcomes approach based on syllabus outcomes in a staged framework will in due course provide another means by which schools can assess the performance of students against benchmarks of what can be expected of students at a particular stage of their learning. This will provide for even
more comprehensive indicators of student and school achievement, particularly at primary school level.

The model described here is based on the belief that even sophisticated indicators and analyses of student outcomes only provide information about what is happening in a school, not why things are this way, and more importantly, what to do to bring about improvement. Wilson, (1995) recognised the same problem in relation to the measurement of student and school performance in the United States.

Reliance on testing student performance to assess schools and reform progress appeared badly flawed. Standardised tests were limited. Test score information provides little wisdom or guidance either to teachers about how to teach better or to policy makers about what policies were effective in supporting schools (p.91).

Oakes (1989) and Porter (1991) and others have argued persuasively for the need to identify school process indicators to inform policy and decision making. Vastly different strategies can be required to bring about improvements in schools with apparently similar levels of performance, depending on their context (see Hopkins et al., 1994). Measuring performance and identifying strengths and weaknesses are not enough — although they can raise expectations in the short term, a process of providing support and resources to meet the needs identified by the accountability process is the key component in any model which aims to produce improvement in schools performance. Without a framework within which to act on the analysis and reporting of outcomes data, including agreed action plans, follow up advice and monitoring to help a school to improve, a sound program of teacher development and more willingness to put resources into schools with problems, post-evaluation improvement in many schools is likely to be short-term and limited.

The NSW Accountability and Improvement Process

The accountability and improvement process for NSW government schools contains four elements. These are:
• an annual school self-evaluation, which is externally validated
• annual school reporting
• school reviews
• systemic monitoring and reporting.

Each of these four elements is discussed briefly below.

1. Annual School Self-Evaluation

All schools will undertake an annual self evaluation, assisted by a school improvement officer. This process will comprise three parts:

• evaluation of a broad range of student learning outcomes
• evaluation of specific areas of school practice
• evaluation of the school’s quality systems.

Evaluation of student learning outcomes

Many schools already use learning outcomes data in the routine monitoring of their effectiveness. The proposed evaluation of student outcomes is therefore not a new concept for schools. What is new is that evaluation will be systematic across the government school system and the relevant results of the evaluation process will be reported publicly.

The basic purpose of using student learning outcomes is to make them transparent so that all relevant stakeholders can analyse them and use them to improve opportunities and ultimately outcomes for students. To use student learning outcomes in a positive and developmental way requires a multi-dimensional perspective of a schools performance. Outcomes information that will be considered for secondary schools includes:

• attendance rates over time
• retention rates
• post-school destinations
- relative outcomes for different types of students, comparison of a school’s performance in different curriculum areas
- results in public examinations
- measures of relative student progress or the “value-added” by the school.

This broad range of indicators recognises the complexity of schools and schooling and the range of students’ educational outcomes.

**Evaluation of specific areas of school practice**

The evaluation of areas of school practice complements the analysis of outcomes evaluation by tapping into areas of a school’s operation such as climate, ethos and management practices, which are valued by the school staff and community but not necessarily accessed through an outcomes evaluation.

The evaluation of specific areas of school’s practice will be undertaken by schools using software supplied and developed by the Department of School Education. A feature of the process is the involvement of students in the self evaluation and the subsequent comparison of the students’ responses with the responses of staff. The areas of school practice which are the subject of this type of evaluation are Learning and Teaching, School Management and Development Planning, and Leadership and Culture. The school self evaluation framework statements have been developed through a rigorous and widely consultative process undertaken over the last three years. The self-evaluation framework and supporting software are described more extensively in Hightett, (1994); Carroll, (1994); and Wasson, 1995).

**Evaluation of the school’s quality system**

The quality systems in a school include school planning and development structures and processes, processes for curriculum delivery, staff training and development, communication and self improvement processes. The information about
quality systems is accessed through the responses given in the self-evaluation framework instruments.

The quality systems evaluation is a major step in diagnosing a school’s developmental capacity and identifying system requirements for support to improve performance. It identifies the internal structures and processes used by a school to monitor and evaluate its own performance. This information will be essential, if Hargreaves’ (1995) conclusion is correct and transferable from the English to the Australian context, that less than a third of schools have the capacity to improve themselves without outside direction and support.

2. Annual School Reporting

The three processes described above together lead to information to be included in the annual report for each school. Annual school reporting is the major strategy for implementing the government’s policy on providing the public with fair, reliable and objective information about the performance of schools. The public report assists schools in identifying priorities and directing their attention and resources to specific aspects of their performance, and results in families being better informed about quality and performance of individual schools. This enables them to be more meaningfully involved in their child’s education, to ask more relevant questions about the performance of their local school, and to choose schools, where alternatives exist, on a more informed basis than at present.

This approach to reporting has several strengths:

- contextual information is provided as a background to performance;
- qualitative as well as quantitative information is included;
- a range of indicators is provided to ensure that performance is not judged on a single indicator;
- the indicators recognize the complexity of the work of the school;
- simplistic league tables are avoided; and
the recurrent nature of the process and identification and follow-up of specific targets encourages a focus on development.

Specific indicators to be reported are described in greater detail below.

3. School Reviews

The annual assisted self-evaluation process will identify some schools as performing satisfactorily and therefore not requiring a more detailed review. Most school reviews in the new model will be focused on schools which are identified as under-performing, although there will be some reviews of high performing schools with a view to sharing good practice across the system. School reviews will be conducted by teams of external personnel with special expertise and lead by a senior departmental officer.

The rationale for moving away from a program of routine, scheduled reviews, as in the old Quality Assurance model, to one of strategically identified and focused reviews is to more effectively and efficiently use limited resources. Other systems have also found that routinely scheduled inspections and reviews have several major shortcomings (Hargreaves, 1988; Burchill, 1995). As a result of these reviews of under-performing schools, the Department’s support services can be directed to those schools which are most in need of them. In particular, District Superintendents will be able to target a proportion of their district’s resources to provide comprehensive support for improving performance in schools that are under-performing.

4. Systemic Reporting

Information about the annual self evaluations of schools is consolidated to provide reports on performance and areas for further improvement at district and system levels. One of the purposes of using such information is to assess whether there is improvement in the ways the system provides support to its schools. In particular, the analysis of information, notably the disaggregation of learning outcomes data according to student subgroups, allows for a more informed evaluation of how the
system is meeting the equity needs of all subgroups in the student population. This in turn, is a valuable basis for systemic decision making.

The Annual School Reports

The annual school reporting process attempts to implement some of the central principles of good practice in the use of performance indicators, as enunciated by Oakes (1987), in that they present summary information in graphical form, which:

- consists of statistically valid information;
- provides benchmarks for measuring progress or regression over time, or differences across geographical areas, such that substantive interpretations can be drawn from presentation of the data;
- represent important policy issues or aspects of education that might be altered by policy decisions;
- can be easily understood by a broad range of citizens concerned with education; and
- are feasible in terms of time, cost and expertise required for production.

The array of indicators attempts to capture three important qualities about school performance: gross productivity, net productivity, and equity (Willms and Kerckhoff, 1995). Gross productivity is defined in terms of the absolute or unadjusted performance of students at the public examinations. Net productivity is defined in terms of the level of performance of students adjusted for their characteristics on entry. Equity is defined in terms of the extent to which student performance is associated with personal characteristics, such as gender. NSW schools have often individually (if ineffectually) reported on their raw test performances in the past. However, this is the first time that they have been required to do so systematically and publicly, revealing their full range of performances, and not just the triumphs of the best students. This is also the first time that there has been public reporting of measures which truly reflect on school performance in terms of the value-added to student learning. The methods for establishing value added are based on the effective schools research (summarised in Sammons et al., 1994; Rowe, Hill and Holmes-Smith, 1995), and draw on multi-level

The following section summarises the data which will be reported to parents and the community for the first time. Much more detailed data will be made available to the school for the purposes of informing their self-evaluation, including longer time series data for the indicators shown below. Also provided will be indicators which show, for example, the relationship between raw and adjusted exam performance in the school, and performance at the course level (students may currently choose from over 100 courses for the HSC). The reports will also contain sections which describe the schools context, including staffing and financial position, and conclude with a section detailing the specific improvement targets for the following year, and progress made towards current year targets. These sections, while seen as critical elements of the reporting framework, are not discussed further here.

1. **School Certificate Results**

The aim of this indicator is to provide information on how the distribution of student performance in the core School Certificate subjects within a school compares to the distribution for the state as a whole. Core subjects (at present, English, Mathematics and Science) are moderated on a state-wide basis. For a significant number of students, the School Certificate is the only credential they will obtain before leaving school and for others it reflects the skills needed in later years of study.

The indicator in Figure 1 is presented in the form of column graphs, showing the percentage of the school’s candidature awarded grades A to E in the year prior to the compilation of the Annual School Report. “N” grades (no award) are excluded for the purposes of calculation of this indicator.
2. Higher School Certificate Results

2.1 Aggregate results

The HSC results indicator provides insight into the level of overall student performance in the school. The measure used is the Tertiary Entrance Rank (TER) – which is commonly recognised by the community as an index of overall HSC performance. For those aspiring to higher education, it is an important outcome of their schooling.

The aim of this indicator is to provide information about the distribution of HSC performance for successive cohorts within a school over time. It allows parents to judge whether the absolute levels of performance in terms of attainment of a Tertiary Entrance Rank are increasing, decreasing or remaining stable over time. The indicator also provides information about the proportion of the cohort not opting for a TER qualifying program. The TER is, as its name implies, a rank, and it is therefore not proper to report school performance in terms of a mean score. It is, however, legitimate to describe the distribution of the ranked scores within a school and the externally validated TER provides the clearest available indicator of relative merit for students in the school.
The indicator is shown in the form of a grouped column graph showing the percentage of the school's student cohort attaining TER scores in the 0–25, 26–50, 51–75, 76–100 range and the non–TER eligible category, as in Figure 2. Percentages of students in each cohort are used rather than the number of students to allow easier comparisons over time. The number of students in each year cohort is provided in the legend. Provision of data over three years is the preferred option, to give sufficient information to discern the beginnings of possible trends, while at the same time avoiding a cluttered and potentially confounding presentation. The graph will not be produced for cohorts of less than 20 students when student scores are spread across the TER categories. These schools will describe student performance in the accompanying text.

Figure 2: Higher School Certificate Results

As noted earlier, a full picture of school performance can only be gleaned by appreciation of both the levels of absolute performance of students and the amount of progress that they make. While the results of the thirteen or so years of schooling most students receive cannot fully be summed up by a single number, overall performance at the Higher School Certificate does provide a means of recognising the outcomes achieved in the senior high school years. This indicator therefore needs to be read in the context of
the value added to students shown in the relative comparisons of school performance indicators.

2.2. HSC Performance in Different Curriculum Areas

One of the known difficulties encountered whenever statistics such as averages or medians are used to describe performance is that they can disguise as much as they reveal. In terms of student achievements, mean HSC results do not provide any information about the strengths or weaknesses of a school in particular curriculum areas. Some schools may have a tradition of excellence in some areas (for example science), but not achieve notably higher results over all KLAs and courses. An increasing body of international literature suggests that not all schools are equally as effective at teaching all subjects (see for example Rowe et al., 1995), although arguably, this should be the goal of comprehensive schools aiming to provide an all-round education for their students. With the establishment of high schools specialising in particular curriculum areas (such as agriculture, sport, performing arts, languages and technology), it could be expected that the additional resources and emphasis devoted to the area of specialisation should mean that, other things being equal, performance in that area is relatively better than in other areas.

The smallest unit on which it is possible to report on student performance at the HSC on a comparable basis is at the course level. However, for public reporting purposes, it is less desirable because in any one year it has the potential to be heavily influenced by the performance of individual students. Uncritical examination of course results may be unfairly used to apportion praise or blame on individual teachers, since many schools will offer only a single class for each course.

The Key Learning Areas (KLA) and KLA groups provide a convenient aggregation of courses that allows confidentiality to be preserved in the main, but at the same time to highlight areas of particular strength or weakness. The majority of Higher School Certificate students study courses in English, Mathematics, Human Society and
its Environment and to a lesser extent Science. Other KLAs are more specialised and less widely supported in terms of student numbers, making it more difficult to preserve student and teacher anonymity in public reports. For this reason they have been grouped into a single category including the Languages Other Than English, Technology and Applied Studies, Creative Arts and Physical Education KLAs. The aggregation within this diverse group of courses clearly represents a compromise between the desire to make comprehensive comparisons and the need to preserve anonymity. It is acknowledged that in this group of courses there may not be an equally meaningful construct within particular schools. An outstanding performance in one course could be masked by weak performances in other courses, and the overall effect drawn towards a median result, for example. Nevertheless, effects outside more traditional core curriculum areas can legitimately, in many schools, be summarised within one KLA group. Schools have the opportunity to elaborate on outcomes for specific KLAs, subjects and courses in the commentary accompanying the charts in this section.

The School v State (Figure 3) aims to answer the simple question: How does the performance of this school’s candidature in a particular KLA (or KLA group) compare to the performance of all candidates in the KLA (or KLA group) across the state?

**Figure 3: School v State Comparison**

![Graph showing State Average scores for various KLAs](image)

The aim of the school v state comparison is to determine how well results in KLA or KLA groups in a particular school compare to average levels in the same KLA
achieved across the state, using the common metric of the transformed unit scores. This index itself is not computationally complex. In the first place, the large sample size in each KLA or KLA group makes it possible to deem the state score a fixed point estimate. However, to make a “fair” comparison, the interpretation of differences between school and state results needs to consider the variability in KLA scores and sample sizes within each school and KLA. For this reason small differences in average scores should not be judged as significantly different from the state mean since small differences may have little substantive meaning to schools.

2.3. Relative School Performance Indicators

The inclusion of indicators of relative performance is predicated on the understanding that measures of absolute (or unadjusted) performance of students do not necessarily reflect equally on the effectiveness with which schools produce these results or on the relative efforts that they must make to produce equal levels of outcomes. In recognition of the fact that students do not begin any period of schooling with equal degrees of preparation, motivation or resources, relative school performance indicators use valid “value-added” concepts to provide a comparison of student performance when all are placed on a “level playing field”. They compare “like with like”.

The indicators included in this section of the reports produce information about the amount of value added by the school relative to expectations based on performance throughout the state. The high school report provides this indicator for students with low, average and high prior levels of achievement, relative to similar students in other schools in the state. The high school version of the indicator draws on research that suggests that schools are not equally effective for all their students and that these differential effects can be obscured if only school mean results are reported.

There are several alternative methods of calculating the value-added by schools to student performance. (see, for example; Hill & Rowe, 1996, Sammons et al., 1994).
In this case, the term value-added refers to the amount of progress students in a particular school make relative to what they could have been expected to make, given what is known about their prior performance and the relationship between prior performance and later outcomes for the state as a whole. Research in this area has demonstrated that the attainment of the children when they first enter the school is the single most important determinant of subsequent achievement, and a growing literature exists both documenting this and discussing how fair and valid comparisons can be made by taking it into account (Goldstein, 1988:197)

The research literature on value-added approaches suggests that the best predictor of later academic performance is finely differentiated measures of prior student achievement The recent School Curriculum and Assessment Authority (SCAA) report in the U.K. recommends that an indicator of value-added can be provided by a simple linear model in which the outcome measure is plotted along the vertical axis against the input measure plotted on the horizontal axis. This allows the estimation of the proportion of students in a given jurisdiction (such as a school or district) achieving results which are better or worse than expected. Other more complex models can include data about the socio-cultural context of schools and their students to provide a more comprehensive view of relative school effectiveness. However, by far the most important correlate of current performance is prior performance, and since prior performance is itself influenced by demographic, background and socio-cultural factors, the need for other adjustment to the statistical models is much reduced.

Relative school performance in high schools compares school and state performance at three points commensurate with lower, medium and higher achieving students as in Figure 4. Three sets of comparisons are provided in the form of school-level residuals summarising the effects associated with students falling into different categories of prior achievement.

Two relative performance models are presented for each high school, one related to School Certificate outcomes and the other to Higher School Certificate outcomes. In
each case the intake-adjusted regression models are centred at the 15th, 50th and 85th percentile points, and residuals for schools calculated at each point. No effects are reported for schools with small numbers of students falling within the range of 15 percentile points above or below the 15th and 85th percentile reference points (clearly, residuals reported in such cases extrapolate beyond the scope of the school's data and have little practical meaning).

The use of 15th, 50th and 85th percentiles as reference points provides comparisons of substantive interest. The 85th percentile reference point, for example, provides a measure appropriate to high performing students in comprehensive schools. Similarly, the lower achieving reference point is of interest in monitoring and planning remedial programs. Considerable strength is added to this approach through fitting multilevel models which can account for the inherently hierarchically nested nature of school outcomes data.

Annual School Reports for secondary schools will report two forms of relative school performance information, one relating to the value-added from year 6 to year 10 and one relating to the value added from year 10 to year 12. The graphical presentation for this indicator is shown in Figure 4.

**Figure 4: Value Added from Years 10 to 12**
3. **Attendance Rates**

Attendance rates are of interest as educational indicators as both a measure of school outcome (in that it can be considered to be a reflection of school climate), and as an important determinant of later academic achievement. The school effectiveness research literature consistently identifies engaged learning time (sometimes called ‘time on task’) as one of the strongest predictors of student academic performance – of which school attendance is an important element (see, for example, Levine and Lezotte, 1990). Several British studies have also examined attendance (and more specifically truancy) as a social/affective outcome of schooling. Rutter *et al.* (1979), for example, found that some urban high schools were better than others in promoting students’ academic and social success. In terms of attendance, academic achievement and desirable classroom behaviour by pupils, it was evident that most schools that tended to be high on any one of these measures also tended to be high on the other two measures. Similarly, schools that ranked low on one outcome also ranked low on the other two outcomes. The design of these studies was such that they were unable to prove any causal link between school ethos and attendance, but the variation in outcomes was not able to be explained by intake factors alone. It can be inferred in these cases, that factors within the school were affecting attendance rates.

For example, poor attendance may result in:

- non-progression leaving gaps in student learning and limiting further progress for the group,
- developmental sequence in programs being interrupted by the need to repeat key sections of work for absent students; and
- the development of casual attitudes and diminished work ethics.

The aim of this indicator is to provide a comparison between average attendance rates for the school, the district and the state, over a four year period. The indicator quite deliberately focuses on attendance (as a reflection of school achievement) rather than absence, which is more properly the focus of information for school management. Attendance data will be reported in narrative form.
4. **Post-School Destinations for Secondary Students**

An indicator of post-school destinations will be included in each secondary school’s report. This indicator recognises that some of the outcomes of schooling, as opposed to programmatic outputs, are complex, long-term and diffuse. At the present time, the major purpose of schooling in NSW is seen as providing an all-round education that “aims to produce citizens who are capable of functioning productively in our present society and who are equipped confidently to meet the challenges of the future” (Carrick, 1989:39). Within the context of this broad general education, students should develop skills which will allow them maximum flexibility and adaptability in their future employment and the capacity to respond with maximum vocational mobility to the challenges provided by technological and social change. This vocational preparation, if not the acquisition of specific vocational skills, is thus an important outcome to be expected from schooling. At the same time, senior secondary schooling has been viewed traditionally as preparation for continuing education, and continues to be valued for this purpose, with the external examinations leading to the award of the Higher School Certificate (HSC) remaining as the primary means of selection to most tertiary institutions.

The extent to which students are able to access continuing education and employment can thus be viewed in part as a reflection of the extent to which their schooling has prepared them for these futures. Many complex and varied factors impinge on the post-school destinations of individuals. In addition to local, state and national conditions, other factors including the expectations of the local community and family aspirations and traditions also come into play. While not attempting to establish any causal relationships in this area, this indicator encourages schools to reflect on the extent to which their goals and methods are congruent with the needs, wishes and outcomes of their clients.

The aim of this indicator is to provide information about the post-school destinations for students who leave school after completing the School Certificate or
2. Higher School Certificate Results

2.1 Aggregate results

The HSC results indicator provides insight into the level of overall student performance in the school. The measure used is the Tertiary Entrance Rank (TER) – which is commonly recognised by the community as an index of overall HSC performance. For those aspiring to higher education, it is an important outcome of their schooling.

The aim of this indicator is to provide information about the distribution of HSC performance for successive cohorts within a school over time. It allows parents to judge whether the absolute levels of performance in terms of attainment of a Tertiary Entrance Rank are increasing, decreasing or remaining stable over time. The indicator also provides information about the proportion of the cohort not opting for a TER qualifying program. The TER is, as its name implies, a rank, and it is therefore not proper to report school performance in terms of a mean score. It is, however, legitimate to describe the distribution of the ranked scores within a school and the externally validated TER provides the clearest available indicator of relative merit for students in the school.
after completing the HSC, or who leave during years 11 or 12 before the HSC – in their first year after leaving school. The cohorts reported on are (in 1997) those who left school at any time during 1996.

The data is presented as in Figure 5, showing the percentage of students in various categories of post-school destination. Percentage figures are preferred in this graph over simple numbers of students in each category to provide a common basis for comparison. Readers can be provided with the total numbers of students in each cohort in the text of the report. Specific groups of school leavers can be noted in the text if appropriate.

The graph would not be produced when the number of leavers in a cohort is less than ten. Where this is the case, the school may prefer to include a descriptive statement about the destinations of students.

These indicators are being constantly refined, and may be added to, modified or deleted as time goes on.

**Figure 5  Post-School Destinations**

<table>
<thead>
<tr>
<th>Destination</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAFE</td>
<td>15%</td>
</tr>
<tr>
<td>Employment</td>
<td>45%</td>
</tr>
<tr>
<td>Unemployment</td>
<td>16%</td>
</tr>
<tr>
<td>Other Schools</td>
<td>4%</td>
</tr>
<tr>
<td>Unknown</td>
<td>2%</td>
</tr>
<tr>
<td>Tertiary</td>
<td>18%</td>
</tr>
</tbody>
</table>

**Conclusion**

The school accountability and improvement process incorporates the key aspects of an effective system for raising the quality of NSW government schools which
have been highlighted in the educational research on school effectiveness and improvement over the past two decades.

Schools in NSW operate in a system which has devolved much authority and responsibility to the individual school and its community. The four elements of the School Accountability Process provide one mechanism for ensuring the accountability of government schools.

In addition to the accountability function, the process maintains a strong emphasis on school development and improvement. It seeks to support the establishment of quality practices at all stages of the work of schools and it aims to ensure schools are highly effective in achieving the best possible outcomes for students.

The process integrates self-evaluation, the provision of fair, reliable and objective information to parents, school reviews, systemic reporting and District and State offices in a way which strengthens and improves the performance of individual schools and the system as a whole. It provides an opportunity to improve the performance of schools because it is flexible and directs resources to support those schools which have the greatest developmental needs.

Evidence from the NSW Quality Assurance school review program over the past few years indicates that the emphasis on external review has focused the attention of many schools on their own internal review processes. This development will be encouraged in the new process. It is appropriate for such a process to monitor school performance in a way which obviates the need for comprehensive reviews of schools which are performing well on their own. This frees up resources for other priorities such as targeting support to schools needing to improve their internal development capacity.

Overall, this process has much potential, but its impact will need to be evaluated in due course. Many school reporting programs have failed to deliver on their promise, and it remains to be seen whether this comprehensive approach fares any better. We believe it will.
Bibliography


NSW Department of School Education (1996). School Accountability and Improvement in NSW public schools: Technical Paper, Sydney


---

i See the OECD (1994) publication Making Education Count: Developing and Using International Indicators, particularly the chapter by Ruby on Indicators, Reporting and Rationality, and OECD (1995) Schools Under Scrutiny for an extensive discussion of the historical, philosophical and empirical underpinnings of this phenomenon.

ii Attempts by the Sydney Morning Herald to depict school quality by ranking the percentage of students in the TER top 5000 is an example of this method.


As McGaw (1997) points out in his review of the NSW HSC, the interpretation of these quantified results is not as simple as it might appear.

See also Bosker and Witziers (1995) for a discussion of different models of school effectiveness.

See Levine (1992) and Walberg (1986) for reviews of this literature.

See Goldstein and Thomas (1995) and Goldstein and Spiegelhalter (1996) for further discussion on the importance of considering standard errors in value added research.


As theory in this field develops and more data becomes available, different models should continue to be tested. This recommendation is made for senior secondary schooling only; other models may be more appropriate for primary school grades.

This same rationale supports the proposed use of aggregate measures derived from the Basic Skills Tests in Year 5 or 6 to predict performance on aggregate measures at the Year 10 level, and at the HSC for appropriately matched cohorts.

I would like to acknowledge the contribution of colleagues Bill Clarke, Patrick Dias, Neville Higlett, Eric Jamieson, Max Smith and Dave Wasson to the development of the NSW School Accountability and Improvement model.

This section draws on material published by the NSW Department of School Education (1996) in a pamphlet titled School Accountability and Improvement in NSW Public Schools: a rationale.