

# **Evaluating the work of teachers in Australian schools**

## **Vision and reality**

**Elizabeth Kleinhenz ACER**

**Lawrence Ingvarson ACER**

**Rod Chadbourne Edith Cowan University**

**Paper presented at the AARE annual conference Brisbane**  
**2002**

[kleinhenz@acer.edu.au](mailto:kleinhenz@acer.edu.au)

ingvarson@acer.edu.au

r.chadbourne@ecu.edu.au

## **Evaluating the work of teachers in Australian schools**

### **Vision and reality**

**Paper for presentation at the AARE Conference Brisbane December 2002**

**Elizabeth Kleinhenz ACER**

**Lawrence Ingvarson ACER**

**Rod Chadbourne Edith Cowan University WA**

#### **Abstract**

Given the importance of effective teacher evaluation to successful teaching and learning, it is surprising that many of the research findings in this area are being ignored. Discredited procedures continue to survive and are even born again within reinvented systems of bureaucratic and managerial control

This paper argues for an approach to teacher evaluation that reflects the findings of contemporary research. It falls into three parts. The first part reviews the relevant literature and distils some features and principles that underpin successful policy and practice. The second part describes a case study that was part of a research project, begun in 2000 that is investigating teacher evaluation in Australian state education systems. The final part of the paper assesses the 'reality' of the situation observed in the case study against the 'vision' provided by identified features of good practice. In this way, an attempt is made to discover how well Australia is being served by methods currently used to assess the work of its teachers. Finally, some suggestions are proposed for the future.

#### **Introduction**

Gaps between the ways in which teacher evaluation systems are currently being implemented in schools, and accounts of effective principles, programs and practices found in the teacher evaluation literature are, increasingly, apparent. This paper examines some of those differences and suggests directions for improving teacher evaluation in Australian schools and education systems.

The paper falls into three parts. The first part uses a review of relevant literature to discuss:

- the purposes of teacher evaluation,
- the use of standards in teacher evaluation
- the roles of participants in teacher evaluation processes

The second part of the paper, in light of the first, describes a case study that was recently carried out in a Melbourne school that is implementing the system of teacher evaluation known as 'staff performance and development.' The discussion centres on the arrangements

made in this school to implement Victorian state government policy in the area of school based teacher evaluation.

The third and final part of the paper summarises the findings of the case study. It then draws some conclusions about the extent to which teacher evaluation in this school reflects or diverges from conclusions drawn about effective practice in the first part of the paper.

### *Background considerations*

As part of an ARC funded research study, we have mapped current teacher evaluation programs and processes currently undertaken in schools and education systems in all Australian states. A number of findings and observations provide a backdrop for this paper.

Two main trends are apparent from the study. The first trend is towards teacher evaluation programs developed in schools in accordance with state government mandates and centrally developed performance management systems. In these programs, teacher evaluation is the responsibility of school principals. The processes are usually hierarchical: principals' performance is 'reviewed' by a system level manager, the principal 'reviews' the performance of the assistant principal and senior teachers; the assistant principal and senior teachers 'review' the performance of other teachers. These programs are now common in government and non government schools in all Australian states.

The second trend is towards teacher evaluation for provisional and full teacher registration. This is now carried out, or about to be carried out, by teacher registration bodies in those states where teacher registration bodies are established, or being established. (New South Wales, Queensland, South Australia, Tasmania, Victoria).

Virtually all teacher evaluation practices in Australia, whether the responsibility of school principals or teacher registration bodies, are at an embryonic stage of development. Teacher evaluation programs that have been developed as systems of teacher performance review tend to follow models derived from the business world. They do not adequately reflect the complexities of teaching as a professional occupation and do not sufficiently address the 'technical core' of teachers' work. (Kleinhenz, Ingvarson & Chadbourne 2001). Teacher registration bodies have yet to develop their own sophisticated means of empowering teachers, whom they represent, to evaluate their own teaching practice and provide guarantees of quality.

Australian teachers' experience of evaluation is different from that of teachers in other countries, e.g. the US and the UK. It is now approximately thirty years since the systems of inspection used to evaluate teachers during the preceding hundred years, were jettisoned. In those thirty years, most Australian teachers have had little or no experience of formal scrutiny of their work. Cultural norms of privacy and individualism remain strong, especially in some secondary schools. This contrasts with the situation in most states of the US where teachers have long been accustomed to pencil and paper tests of their competence and (unpopular and ineffective) classroom visitations by principals. It also contrasts with conditions in the UK, where a form of external inspection survives in government schools to this day.

Australia, then, is something of a *tabula rasa* with regard to teacher evaluation. Performance management systems are still relatively new and their impact not yet strong. Stringent evaluation for certification or registration purposes is a foreign concept to Australian teachers. The debate about evaluation, if there is to be one, is yet to take place.

## **PART ONE**

### **Purposes, standards and participant roles in teacher evaluation.**

#### *Purposes*

Writers on teacher evaluation agree that the overarching purpose of teacher evaluation is to ensure that children are well taught. This purpose embodies two generally recognized strands or categories: to improve teachers' work and help them to keep pace with new knowledge about effective teaching, and to guarantee teaching quality. (Stronge, 1995; Duke, 1990; McLaughlin & Pfeiffer 1994; Shinkfield & Stufflebeam, 1995) (Ingvarson 2001 ).

Stufflebeam (1988) points out that meeting the educational needs of students and the community is the primary, overarching purpose of teacher evaluation. He stresses the need for all participants in teacher evaluation systems to work towards this purpose:

Educational systems exist to meet the needs of students and the community; so all elements of those systems, including personnel evaluation, should be directed toward achieving that purpose. (Stufflebeam, 1988, p. 22).

Improving teacher quality is not the only motive a school or education system may have when it conducts teacher evaluation. Schools and systems also seek quality assurance through using evaluation to assess basic competence, dismiss teachers for incompetence and ensure that the 'right' people are promoted to key positions. This creates a kind of tension. Teachers who feel under pressure to justify their work practices to their assessors for fear of dismissal or failure to gain a promotion will be unwilling to expose weaknesses, even though identification of weaknesses is a necessary first step towards improvement.

One of the dilemmas facing teachers then, and now, is the belief that, on the one hand, the evaluation function should lead to professional growth while, on the other hand, it provides a ready weapon for manipulation by administrators. What potentially should be good may be seen as functionally insidious. (Shinkfield & Stufflebeam, 1995, p. 16).

Historically, teacher evaluation systems have failed to solve this dilemma. They have also failed to provide genuine guarantees of quality, although the appearance of such guarantees may have been created.

Seventy years of empirical research on teacher evaluation shows that current practices do not improve teachers or accurately tell what happens in classrooms. Administrator reports do not increase good teachers' confidence or reassure the public about teacher quality. Teacher evaluation as currently practiced does not identify innovative teaching so that it can be adopted by other teachers and used in teacher education programs. Finally, current procedures do not reward exemplary teachers. (Peterson, 2000).

A major problem is that teachers are suspicious of the evaluators' motives and dislike the methods used to assess their work. McLaughlin & Pfeiffer see these 'organizational issues' as 'the primary obstacle to initiating and carrying out meaningful teacher evaluation' (McLaughlin & Pfeiffer p.5). Winning the trust of teachers, they say, is a necessary precondition for any successful evaluation system:

The most difficult problem of teacher evaluation, then, is not only to develop a better instrument. It involves organizational questions of getting started - how

to overcome the resistance and negative attitudes that exist about teacher evaluation. (McLaughlin & Pfeiffer, 1998, p.5).

'Getting started' entails recognising that purposes need to be clearly understood and shared by all stakeholders. (Darling Hammond, Wise & Pease, 1983; Stiggans & Duke, 1990). It also means commitment to the idea of *improvement* as a major purpose of evaluation. Without such shared understandings and commitment teacher evaluation descends into a pointless and, in many ways, hypocritical game in which teachers and administrators, in their own ways, tacitly conspire, to protect their own and each others' territory. This usually means assiduous avoidance of serious engagement with the 'hard' issues involved in bringing about real improvement in teaching quality.

It is easy to say that teachers and administrators alike need to look beyond the narrowness of their 'own' interests to the broader purposes of teacher evaluation, especially the purpose of improving teaching quality. However, if this were as simple as it sounds, it would have happened years ago. That it has not, indicates that arriving at common understandings of purpose is a major and difficult undertaking. There are serious problems and dilemmas that need to be worked through. Schools need help to uncover and tackle the complexities and uncertainties that impede clarity and prevent teacher evaluation from fulfilling its potential to be a useful and practical tool for improvement.

Schools cannot 'implement' teacher evaluation systems if they do not know why they are doing so. Finding answers to the 'why' - answers that are acceptable to all stakeholders, will require solid investment of time and energy. Nonetheless, reassessment and clarification of the purposes of teacher evaluation, in relation to various interests, is an essential first step for any school or system that is struggling to make teacher evaluation 'work'.

### **The role of teaching standards in teacher evaluation**

Teaching standards are essential to a professional model of teaching and provide a sound basis for teachers' learning. (Darling Hammond, (1988; 1999). As such they assist the teaching profession to evaluate its own practice and provide its own guarantees of quality. Standards serve several purposes in teacher evaluation. They explicate what teachers can be expected to know and do on the basis of research and best practice, they describe desirable levels of performance and they clarify for teachers the elements of what they can expect to get better at over the long term. Good sets of standards legitimate teachers' claims to the possession of special knowledge of a kind that lends itself to professional rather than bureaucratic evaluation methods. As Elmore (1996), points out:

The existence of external norms is important because it institutionalizes the idea that professionals are responsible for looking outward at challenging conceptions of practice in addition to looking inward at their values and competencies. (...)Without some kind of external normative structure, teachers have no incentive to think of their practice as anything other than a bunch of traits. The existence of strong external norms also has the effect of legitimating the proportion of teachers in any system who draw their ideas about teaching from a professional community and who compare themselves against a standard external to their school or community. External norms give visibility and status to those who exemplify them. (Elmore 1996..)

It is difficult to conceive of effective teacher evaluation in the absence of coherent, valid and reliable standards:

Without a framework, the structure is reduced to something the mentor, coach, or supervisor has in her head, and thus reflects the personal beliefs that individual holds about teaching, regardless of whether these have ever been made explicit. (Danielson, 1996 p. 7).

Ingvarson (2002 p.6) points out that the nature and content of teaching standards should vary according to their purpose. Although 'core principles' do not change, different sets of standards are suitable for different purposes such as accrediting teachers, job selection, and certification of accomplished practice.

Because the primary purposes for teacher evaluation, quality assurance and professional learning, are interdependent, teaching standards must be able to serve these two purposes, as well as the more specific purposes of particular evaluations. In particular, professional learning systems for teachers need to be underpinned by good sets of professional teaching standards.

Many sets of standards have been developed in recent years, but relatively few have tackled the issue of how they are to be *used* to achieve the primary and specific purposes of particular evaluations. Most have not been designed with assessment in mind, an essential feature if their aims are to be achieved.

Arguably, the most useful and effective sets of teaching standards emerge when they are designed with assessment in mind. Confronting though the idea may initially appear, the most valid and coherent sets of teaching standards are those that are designed as tools for the assessment of teacher performance. This is because:

A standard is a *tool* for rendering appropriately *precise* the making of *judgements* and *decisions* in a context of *shared meanings and values*. Sykes and Plastrik (1993).

Teachers themselves understand these things when they design student activities and methods for assessing student learning. Quality student assessments are designed to provide useful feedback and to provide avenues for learning. The same applies to teacher learning. New methods for assessing teacher performance against standards are emerging that are in themselves excellent vehicles for professional learning. Examples include the standards of the National Board for Professional Teaching Standards (NBPTS) (1989), the Interstate New Teacher Assessment and Support Consortium (INTASC) (1992) and PraxisIII/ (1994).

In our view, these are powerful examples of how to integrate standards into systems for teacher evaluation and professional learning. The NBPTS standards are for the advanced certification of teachers of accomplished practice by the Board, which is an independent teacher professional body. The Praxis III/Pathwise and INTASC standards were developed for state and local agencies to use in making teacher licensing decisions. Most states in the US now base their local standards for teacher licensure on the INTASC or Praxis models. These standards are part of a comprehensive set of guidelines, task standards and performance standards with supporting resources, for determining whether teachers have met the standards. Standards are research based and were extensively field-tested before being adopted. Teachers are required to demonstrate, through a range of task-based evidence, that they meet the standards. The tasks they perform are examples of normal aspects of teachers' work, not artificial 'add ons.' As such, they provide a "natural harvest" of evidence such as students' work samples and teaching artefacts, that can be used for assessing teacher performance.

### *Links between standards and teachers' professional learning*

According to Elmore, (Elmore 2000) teachers learn best when their learning is focused on the elements of the 'technical core' of teaching and learning. The technical core comprises the central features of teaching and learning, such as knowledge of subjects, knowledge of students, knowledge of how to facilitate students' learning - features that are articulated in good sets of standards for teachers' practice. There is thus no need to 'make' links between teachers' professional learning and standards of practice if the standards adequately, coherently and comprehensively express this technical core of what teachers know and are able to do.

With a framework of professional practice in hand,... participants can conduct conversations about where to focus improvement efforts, within the context of shared definitions and values. These conversations can focus on means, not ends, and they can be conducted in an environment of professional respect. (Elmore, 2000)

There is no shortage of examples, in Australia and overseas, of standards that were not developed with assessment in mind and that do not form part of an integrated professional assessment and teacher learning system. Such sets of standards court the obvious danger of falling into disuse, or of never being used in the first place. Even when co-opted to serve as criteria for various teacher assessment purposes and projects these standards appear to be of limited use.

### **Participants in the evaluation process**

The participants in all teacher evaluation processes are the teachers (evaluatees) and the people who evaluate their work (evaluators). (In processes of self evaluation, teachers are both evaluators and evaluatees.) There are two major types of assessors: principals/line managers operating in some form of hierarchical structure, and peers with advanced levels of expertise.

### ***Evaluation by principals/ line managers***

Historically, evaluation has been something done *to* teachers by people like inspectors and principals. Teacher evaluation as performed by these people may well have served quality assurance functions in minimalist ways. In some instances it may even have led to improvements in the work of some teachers, but there is little evidence of it being effective on a large scale. On the contrary:

Seventy years of research on principal ratings of teachers shows that they do not work well. Well designed empirical studies depict principals as inaccurate raters both of individual teacher performance behaviours and of overall teacher merit. These problems go beyond simply developing a better rating form, informing teachers of the items on the rating forms, or improving the training of principals as raters, although these strategies are offered by many who are interested in preserving current practice. (Peterson, 2000, p. 19).

Other research shows a high degree of disrespect for teacher evaluation conducted by principals and administrators:

Almost all respondents (to a survey of 32 district school offices) ...felt that principals lacked sufficient resolve and competence to evaluate accurately. (Wise, Darling-Hammond, McLaughlin & Bernstein, 1984, p.2).

Twelve studies from 1921 to 1959 reached the same conclusion: that the correlations between the average principal's rating of teacher performance and direct measures of teacher effectiveness were near zero. (Medley & Coker, 1987, p. 242.)

Teachers interviewed for this study roundly criticized formal supervision and evaluation practices...(saying) that they are...not effective for improvement...Administrators are rarely prepared to offer...useful advice...virtually never...providing an opportunity for learning...Very good teachers...regard the practice as an institutional obligation to be endured rather than an opportunity to be seized. (Johnson 1990, p. 266.)

(These quotations are three of many, quoted by Peterson (2000), to show the ineffectuality of the type of teacher evaluation systems carried out in the US in which principals and district administrators are responsible for managing bureaucratic, hierarchical processes).

These findings, from research done in the US tally with similar research findings in the UK. A research team from Exeter University (Chamberlin, Wragg et.al. 2001) found, for example, that there is considerable teacher and head teacher dissatisfaction with the evaluation methods prescribed by the British Department for Education and Training (DFET) for the 'Threshold' position. The purpose of this evaluation is to decide which teachers should be eligible to 'cross the threshold' between competent teaching and highly effective practice, thereby gaining a substantial pay increase. Implementing this process is the responsibility of Head Teachers and highly paid External Reviewers. In the 1000 schools of their study sample the researchers found that the success rate of applicants was 97%, (the same as the national success rate). This figure alone raises questions about the effectiveness and validity of the evaluation. It also raises the obvious question of whether the evaluation was necessary in the first place. Surely, commented many observers, it would have been easier and cheaper to have simply given these teachers a pay rise.

The Exeter research team found that head teachers themselves were "vitriolic" in their criticism of training to conduct the assessments:

If you are selling 'double glazing' that hardly anyone wants, that has been badly designed, that is incomplete and that does not really work, then you are probably on a loser from the start - especially if you only heard about it the day before and don't really understand it yourself.' (Chamberlin, Wragg et. al 2000 (a), p. 2).

While most teachers were not overly critical of the role played by principals in their evaluations, some, especially the few who failed, were highly critical and claimed that they had been victimized. Only seven of the 174 unsuccessful teachers felt that the judgment of their case was justified, The other 167 were "shocked" "furious" or "demoralized" For example, with respect to some of their principals, some unsuccessful teachers said:

'I have had many confrontations as a Union rep and she bears grudges'

'He is extremely unsupportive because I have not always agreed with him. He has blocked my provisional development since his arrival three years ago.'  
(Chamberlin Wragg et al 2000 (b))

The most significant disadvantages of the processes that made individual school principals responsible for teacher Threshold evaluations arise in the following areas:

- The differences in the way in which the Threshold Assessment procedure is carried out in different schools
- The problems relating to the reliability and comparability of performance data
- The basis on which some head teachers made judgements about an individual's performance
- Possible bias and victimization by head teachers
- The potentially divisive nature of the procedure

As for the role of the External Reviewers, their judgements differed from those of Head Teachers in less than 0.4% of cases. Not surprisingly, one secondary teacher summed up an external reviewer as:

'Nice man, served no useful purpose.' (Haynes, Wragg et al 2000, p. 14).

### ***Evaluation by peers***

An alternative to teacher evaluation by principals and external administrators is peer appraisal. The most effective models of teacher evaluation using peer appraisal are to be found in the USA. Three examples are

#### *The National Board for Professional Teaching Standards.*

The work of the NBPTS is now familiar to all who have an interest in teacher evaluation. The Board has drawn heavily on the best brains in the educational measurement world and conducted more research than any other body in developing its certification system for teachers of accomplished practice. Although not without its critics, NBPTS certification is a system that continues to grow and gain credibility and recognition from all sectors from teacher unions and state governments to business and parent organizations.

Teachers who apply for National Board certification as highly accomplished teachers are asked to complete ten separate assessment tasks (six portfolio entries and four performance centre tasks). These tasks aim to represent the range of abilities in the content standards.

Assessments are carried out by specially trained peer assessors who are themselves Board certified teachers of accomplished practice. Each task provides independent evidence of performance on several standards. Every standard is assessed in several ways.

#### *The Interstate new Teacher Assessment and Support Consortium (INTASC).*

INTASC is a major national program that facilitates collaboration among states in the development of standards for licensing *new* teachers. It articulates ten 'common core' principles that set out to define the knowledge, dispositions and performances necessary for

a learner centred approach to teaching. It has established sub-committees to translate these core principles into standards and performance based assessments across the curriculum.

For example the mathematics sub-committee, consisting of highly regarded teachers of mathematics, teacher educators and researchers from across the country, translated the ten core principles into 'Standards for Beginning Teachers of Mathematics'.

Assessors are exemplary teachers who receive similar training to that of the NBPTS assessors. Assessments broadly follow the National Board practice of portfolios and assessment centre tasks, but differ in their levels of demand and complexity.

### *Praxis III*

Praxis III is a framework of standards for guiding teacher learning during induction and assessing the performance of beginning teachers. It was developed for the use of local state agencies that are empowered to license new teachers.

Unlike the NBPTS standards and the INTASC standards, the Praxis standards are generic rather than subject or stage-of-schooling specific. However they share a number of similarities with the NBPTS and INTASC standards and assessments. These include a shared philosophy, a similar standards 'architecture' and the use of trained peer assessors.

The most significant difference between the NBPTS and INTASC evaluation systems and Praxis III is in the way the standards are operationalised in assessments. While the first two use portfolios and assessment centre exercises, the Praxis III peer assessors use classroom observations. Teachers are assessed on criteria that are part of the standards framework. The assessors use prepared materials including "Questions for assessors' reflection." Several levels of performance are identified for each 'Domain' of the standards.

### *The Level 3 Classroom Teacher position Western Australia*

While many schools in Australian government and non government systems of education have instituted their own peer appraisal schemes, we found in our study only one example of teacher evaluation that used non school based peer appraisal within a standards based assessment that was built into the career progression of teachers. This was the evaluation for the Level 3 Classroom Teacher position in Western Australia.

This position gives promotion and a pay rise of up to \$5000.00 for teachers who are able to demonstrate outstanding teaching performance. The evaluation is carried out quite separately from an applicant's school. It is based on portfolios of evidence and a review in which teachers make a presentation to other teachers who are being assessed. Two trained peer reviewers, who are themselves classroom teachers, assess both portfolios and the presentation.

### *Advantages of evaluation by peers*

The use of peer assessors for teacher evaluation has many advantages and has the support of most commentators on teacher evaluation. The main advantages have been summarized by Peterson (2000 p. 122) as:

- Balancing expert knowledge of curriculum with the practical experience of "teaching immature, non voluntary persons in a group setting".

- Applying subject specific knowledge to recognise subject specific "subtleties", recognise effective strategies and give useful feedback
- Applying their own subject knowledge to judge the subject knowledge, and the ability to organize subject knowledge for teaching purposes, of the person being evaluated
- Including and explaining student achievement data in teacher evaluation.

Peterson comments that evaluating the work of their peers is also a valuable professionalising activity for teachers:

A sense of professionalism is strengthened with the idea of shared craft knowledge. Peer review makes it possible to exchange information and techniques. It makes exemplary practice available for others to follow. Finally, the professional and political standing of teachers in the society is enhanced with the self-regulation of peer review. (Peterson , 2000, p.123).

### **Part 1: summary and conclusions**

The purposes of teacher evaluation, under the overall 'umbrella' of serving the educational interests of students and the community fall into two main categories: teacher development and quality assurance. While there are some inherent dilemmas and even apparent contradictions between these purposes, they are interdependent and complementary. Effective teacher evaluation, even when it is concerned with specific purposes such as competitive selection for teaching jobs, and promotion addresses elements of both teacher development and quality assurance.

Standards explicate what teachers know and can do. They provide a basis for effective teacher evaluation and for teachers' professional learning. Good sets of standards are developed with assessment in mind and are embedded in programs of professional learning for teachers. Assessment of teachers' work has similarities with assessment of students' work in that it provides feedback for reflection and further development on the basis of criteria derived from the standards.

There is a great deal of evidence to suggest that teacher evaluation that is done *to* teachers *by* principals and administrators has never been effective, especially in terms of improving the quality of teaching. Evaluation carried out by knowledgeable peers, on the other hand, has been shown to meet teachers' own expectations of effective processes that understand and recognise the complexities of their work. Teachers learn from this kind of evaluation experience and welcome the professional feedback from their peers.

For teacher evaluation to be effective, therefore, all participants need to understand why the evaluation is occurring and share common understanding of the desired outcomes. Standards need not only to be in place, but to be embedded in professional learning systems that help teachers to know how to improve their practice and when and how learning takes place. A way of achieving this is to place the onus on teachers to provide evidence that they meet the standards. This can be done by, for example, asking teachers to complete tasks that necessarily engage teachers in analyzing and evaluating their work with colleagues.

## **PART TWO CASE STUDY**

This section of the paper describes a case study in teacher evaluation that was undertaken in a Victorian government school in 2002. It aims to show the extent to which teacher evaluation in this school reflects or diverges from the conclusions about effective teacher evaluation described, in relation to the literature, in the first part of the paper.

Teacher performance review was introduced into Victorian government schools in the mid 1990s. The present system was agreed upon between the Victorian Department of Education and Training (DE&T) and the Victorian branch of the Australian Education Union (AEU) in 2001. It is now fully operational under the direction and management of principals in all government schools in Victoria.

The objectives of this system are described in terms of ensuring the quality of the education system and schools and improving of teaching quality. (DE&T,2001, p.6). Teachers must complete a twelve-month 'performance cycle' during which they are required to demonstrate their competence against the professional standards for their level. All must attend three meetings with a reviewer, (the principal or a delegate of the principal, at the start, middle and end of the cycle).

School principals are responsible for all evaluations. 'Failure' results in the withholding of salary increments for those teachers who are still progressing on a salary scale. Further disciplinary action, including instigating processes for dismissal may be taken if a teachers' performance is unsatisfactory. The system of staff performance and development was agreed in the 2001 industrial agreement between DE&T and the Australian Education Union, (AEU) Victorian branch.

### **The case study**

This section of the paper describes the findings of case study that was undertaken in a secondary school in an outer Melbourne suburb in 2002. It draws some conclusions about the purposes, use of teaching standards and participant roles in the teacher evaluation processes observed at this school. As a case study of only one school, it is open to the criticism that its findings cannot be generalised to other schools. Indeed it teacher evaluation processes are carried out quite differently in some other schools. (This is one of the major problems of teacher evaluation under the performance management model as it operates in Victorian schools, especially in terms of its validity, reliability and credibility). Lessons can be learned, however, even from such a limited case, and a degree of generalisation is possible, not to other case studies but to 'theory, analogous to the way a scientist generalises from experimental results to theory'. (Yin, 1984).

### **Selection of the school**

The school was selected for study on the basis of a recommendation from a Department of Education Assistant Regional Director, who believed that the school had developed an exemplary model of staff performance and development. The pseudonym 'Sundale Secondary College' has been given to the school in order to maintain confidentiality. Pseudonyms have also been used to ensure the anonymity of the respondents.

## **Selection of participants**

'Bob' the principal, in consultation with the chief researcher, chose which teachers would be interviewed. Bob described the performance management program at his school as 'a sensitive issue'. He was honest in saying that he would not recommend interviewing certain disaffected members of staff because this might create a negative impression of the processes. These teachers, he said, were in a small minority; the great majority of teachers supported the program.

Seven people were interviewed: the principal, an assistant principal who was a reviewer, a leading teacher who was a reviewer, and four teachers. The principal and assistant principal were male, the two leading-teacher-reviewers were female, three teachers were female and one was male. There were age differences among the participants, but these were not seen to be of significance for the study.

## **Collection and analysis of data**

Interviews with teachers and the school principal provide most of the data for the study. These interviews were tape recorded and transcribed. Corroborating evidence was obtained from school documents and from relevant state Department of Education & Training documents. Data was analysed as it 'converged' on the three questions that led the investigation: These were:

1. What are the purposes of teacher evaluation in this school? Are they the same as the purposes of the education system of which the school is a part? To what extent do they meet the school/system purposes?
2. What standards are being used for teacher evaluation in this school? How are they perceived and used by participants in the evaluation processes?
3. Who are the participants in the evaluation processes in this school, and what are the key characteristics of their roles?

## **Performance review processes at Sundale Secondary College**

The performance of every staff member at Sundale, including the principal, is monitored and assessed by a person at a higher level of the hierarchy. The Regional General Manager reviews the principal's performance. The principal reviews the performance of the assistant principals and leading teachers. Leading teachers review the performance of classroom teachers. The principal participates in review processes for all teachers and is responsible for the final assessment.

Processes are carried out in accordance with Department of Education and Training (DE&T) directions (DE&T 2001). The school administrators inform staff about cyclical review processes in a variety of ways. These include reminders at staff meetings, news updates on school bulletins, staff memos and leadership team memos.

Including the principal, nine people act as reviewers at Sundale. The assistant principals and leading teachers review the performance of sixty teachers. Each reviewer 'manages' the performance of eight or nine teachers.

At the start of the process, teachers choose a reviewer and their choice is considered by the leadership team who make the final allocation. The principal reported that every effort is made to ensure that reviewers are familiar with the work of teachers, e.g. teachers who teach mainly VCE are reviewed by the VCE coordinator.

Processes follow the "cycle" as prescribed by DE&T. The first meeting takes the form of a discussion between the teacher and the reviewer during which the review processes, the teaching standards, and expectations of performance are explained. At the second (mid-cycle) meeting, the reviewer and teacher discuss key elements of the teacher's performance in relation to the standards.

Before the mid-cycle meeting, teachers prepare a 'statement of performance' on a pro-forma that allows a brief, (about ten lines) statement in relation to each standard. The statement of performance is expected to include professional development learning outcomes.

At the end of the review period, the reviewer makes a summative assessment of each teacher against each standard. All standards must be met for the outcome to be successful. An end of cycle assessment pro-forma is signed by the principal and the teacher.

The Sundale principal meets with every teacher at the end of the cycle, discusses the reviewer's report, informs them of the result of their assessment, recognizes their achievements and encourages their efforts.

### **Perceptions and understandings of purpose**

#### *The principal*

The principal saw the purpose of annual performance review as (a) to ensure that teachers were teaching effectively (b) to help teachers improve their work (c) to recognise teachers' achievements. He believed that the processes strengthened his authority to deal with teachers who were seriously under performing. He said he was glad about that because it empowered him to do what he felt was right for the students in his school:

Kids shouldn't have to put up with disgraceful teachers. And the system shouldn't be supporting them. I've known teachers who couldn't even be put in front of a class on their own, they've had to be put in the library or somewhere, but the system kept them on.

The performance management processes at Sundale were first implemented in the mid nineties as part of the now discontinued 'Professional Recognition Program' (PRP). During the past two years, with the appointment of a new principal, and in consideration of changes that followed the discontinuing of the PRP and the 2000 Industrial Agreement, processes have been refined and strengthened.

The principal said that these processes helped to identify three teachers whose work was not satisfactory. When one of these teachers learnt that the outcome of his review was likely to be unsatisfactory he resigned. The other two were dismissed after lengthy and difficult procedures involving the DE&T and the Australian Education Union. (AEU). All other teachers so far have been successful in their reviews and no pay increments have been withheld.

The principal saw the importance of all staff understanding the purposes and process of review. This, he said, should happen at the first meeting in the cycle:

That first meeting is crucial. You have to get it right at the start. The crucial thing is to have shared understandings from the beginning.

The principal appeared to be more focused on the quality assurance purposes of the review and on the administrative details of the procedure than on details of how individual teachers' performance might be improved in the course of and as a result of the review.

### *The reviewers*

The reviewers said that helping individual teachers to improve their work was their major focus. Quality assurance purposes, they said, were a 'given' in that the school was obliged to carry out the requirements of the Department. Both reviewers interviewed said that they were vehemently opposed to the idea of their role as being to 'weed out' unsatisfactory teachers, saying that there were other processes to deal with such people.

One reviewer described the accountability factor as a 'prod' that led teachers to reflect on their performance and seek improvement.

Teachers have to do it and that's it, we don't have to think about that. And that's good because it gets them going. What we're doing here is helping them to reflect on their strengths and weaknesses, identify areas of improvement, set goals and help them to reach them (Sam, reviewer).

The other reviewer interviewed said that for her, the purpose of review tended to vary with the person whose performance she was reviewing. As she saw it, there were *three groups* of teachers. The first group was made up of teachers who were apathetic and even hostile to the processes. It included a disproportionate number of older teachers. The purpose of review for these teachers, she believed, was mainly to ensure that they were meeting the standards to a satisfactory level. Since these people were resistant to change, it was difficult or impossible to talk with them about improvement. The review processes alone could not alter this situation.

The second group of teachers, in the view of this reviewer, comprised those who were 'anxious to do the right thing' but unsure of how they should go about it. She saw the process as 'helpful' in encouraging such teachers to identify their strengths and weaknesses and seek out appropriate professional development that would help them to improve. For these people the purpose was reassurance and 'pointing them in the right PD directions'.

The third group of teachers, she said, was made up of teachers who were 'doing great things'. Their work was improving anyway and would have done regardless of review processes. Some of these teachers were 'career minded'. They welcomed the review as a way of showing senior people what they were doing and they appreciated the recognition. This reviewer saw the purpose of annual review for these people as being to ensure that their energies were channeled towards the school curriculum priorities and to recognise superior performance.

Without the review we mightn't even know what these people were doing. It's amazing, some of the things....Because we know we can give them the recognition they deserve.

### *The teachers*

All four teachers interviewed saw the purpose of the review as being more about meeting accountability requirements than improving their performance. Three of the four teachers

were anxious about 'passing'. One teacher, 'Elsie' said that she felt seriously overworked and that the review processes had just added to an already unbearable load.

I am a meticulous person. I put together a mound of stuff in a portfolio for my review. I put in the professional reading I'd done, and the PD courses and stuff from the internet, course planning, assignments, games and classroom activities. The principal has high expectations and this places more pressure on us. But that's not his fault, it's the government's.

Elsie could see that a purpose of performance review might well be to help her to improve her teaching. But although she had experienced preparing for her interviews as a useful reflective exercise, it had been more of a burden than a way of helping her to improve. She also felt that because most of the work had been done at home, it was somehow disconnected from the workplace.

'Ross', taught ICT. He said that he 'loved' teaching and enjoyed his classes. In contrast to Elsie, had spent less than one hour on preparing for his first review meeting. He too saw the main purpose of review as accountability (which he 'totally' supported), but was confident that he could more than satisfy requirements. He was not planning to supply written evidence for his mid cycle review, but was prepared to do so if he was asked. He spoke enthusiastically and at length about his subject, ICT. He said he was continually trying to improve, but improvement came more from activities he was pursuing in and outside the school on his own initiative than from the review processes. He agreed that the review provided recognition of his work, but said that the enthusiasm of his students and other intrinsic rewards of the job were much more important to him.

'Lily' also focused on accountability as the main purpose of annual review, saying that she approved of this purpose:

Older colleagues used to talk about inspectors. I thought it wasn't such a bad idea. There is a desperate need for accountability in the state system. I respect the process, we need to be able to justify our salaries.

She felt that the 'bureaucratic' nature of the processes made it difficult to see annual review as a vehicle for improvement. (By 'bureaucratic' she was referring to the form filling and the schedule of interviews). Her reviewer had pointed out to her the advisability of undertaking some professional development in the use of ICT. She had found this helpful and had followed the advice. But she said there had been no opportunity in her interview to discuss, for example, an assignment she had prepared for her students using Bloom's Taxonomy and Multiple Intelligence theory.

I wanted to talk about that because I thought it was a very good assignment. But there was no discussion of teaching content. This was strictly a matter of time. If it was part of our ongoing work it would make more sense. At the moment it feels like it's something we're doing to please the Department.

Of the four teachers, 'Marita' appeared to be the most focused on improvement as a purpose of the review processes.

It's not just a check up. It's to make sure I'm on track and to look at ways I may improve.

She saw the review as a means of recognizing her strong points and helping her to identify areas for improvement. For example, she explained:

Marita: I do heaps of extra things with the students. I love being with the students - camps, the school production, talent quests, song fests. I love that, it gets me closer to the students. But I worry about some things. At uni we didn't do anything about planning CSF units. How do you do it?

Int: Well, that's a very big question. Has the review process helped you with that?

Marita: Ye-es - but there wasn't really time. It just helped me to know that's one area where I need to improve.

## Standards

The standards developed by the DE&T for schools to use in annual staff performance and development reviews are generic. There are different sets of standards for teachers at different career stages.

These standards take into account the complexities of teaching practice and school culture, and the developmental processes that characterize all teachers' profession from a beginning teacher to a highly skilled professional. (DE& T 2001).

The standards that applied to the teachers interviewed at Sundale were those for 'experienced' and 'beginning' teachers.

The standards are set out in the performance review 'pro-formas.' Teachers must satisfy reviewers of satisfactory performance in relation to all standards. The following example of a pro-forma that was completed by a teacher is for Standard 1 of the eight *Experienced Teacher* standards.

### Experienced teacher performance and development

Professional standards	Statement of performance (teacher completes with help of reviewer)
Demonstrate a high level of knowledge of relevant curriculum areas, student learning processes and resources and can apply it in implementing programs which enhance student learning	<ul style="list-style-type: none"> <li>• <i>Prepare units of work that reflect CSF outcomes.</i></li> <li>• <i>Prepare an assignment for students that incorporates Information Technology.</i></li> <li>• <i>Develop a resource bank of print, film and internet materials/web sites.</i></li> </ul>

Most interviewees at Sundale, including the reviewers, disliked what they called the 'bureaucratic' language of the standards and felt no sense of 'owning' them. It was hard to discover what they meant by 'bureaucratic', but, clearly, they did not feel that the language of the documents was 'their' language. The standards were seen as having been imposed on

teachers by the employer and using the employer's language. The following comment was typical:

It's the unfriendly jargon of the standards. Teachers see it as more appropriate for businesses than schools. It takes quite a bit of time to work through that and by the time you do, some teachers are so turned off that, you, know, they don't want to know about it. (Sam, reviewer).

However, no respondent had a serious quarrel with the content of the standards. The principal, teachers and reviewers believed that that they were useful for delineating the *areas* in which teachers' work could be reflected upon for purposes of improvement. Teachers agreed that reflecting on their performance as it related to each of the standards (a) increased their confidence that they were 'covering' all areas and (b) caused them to think more systematically about their teaching in relation to each area. Having to make a statement of performance in relation to each standard had, for some, caused them to try something new and to reflect upon it.

I developed major assignments in English and French that incorporated Gardner's multiple intelligences. I don't know whether I would have done that without the Review - maybe not. I certainly learnt from that and having to meet that standard made me focus on that area.

Teachers also saw the usefulness of linking the statewide standards with the school priorities of literacy, numeracy and ICT. As part of their performance review all teachers had to prepare a unit of work that incorporated some aspect of ICT. This task was linked to Standard 2 - 'Demonstrate high-quality classroom teaching skills and successfully employ flexible and adaptive approaches and constructive strategies to allow students to reach their full potential.'

### Participants' roles

As Table 1 shows, in addition to the teachers whose performance is being reviewed, there are three main participant roles in performance review processes at Sundale High School.

Role in review	Position in school	Review duties
Chief Reviewer and person responsible to DE&T for review outcomes.	Principal	Directs processes  Delegates review process management duties to assistant principal

		<p>Delegates review duties to assistant principals and leading teachers.</p> <p>Makes final summative judgement on basis of reviewers' reports.</p> <p>Conducts end of cycle interview with every teacher and informs them of result of process.</p>
Manager of Annual review processes	Assistant principal	<p>Manages processes</p> <p>Informs staff</p> <p>Liases with reviewers</p> <p>Reviewer</p>
Reviewers	Assistant principals and leading teachers	<p>Conduct three formal interviews with teachers in the review "cycle". Act as mentors and advise teachers of possible amendments to their statements of performance and pd plans.</p>

### ***The principal***

As well as having chief responsibility for staff performance and development, Bob, the principal of Sundale High School was personally involved in the review processes in several ways: he reviewed the performance of the assistant principals and leading teachers, liased with the assistant principal who managed the processes for teachers' reviews, made the final summative judgement of every teacher's performance on the basis of reviewer's reports and met with every teacher individually to inform them of the result and discuss it with them.

Bob said that he used 'both formal and informal' evidence when making his judgements of teachers' performance.

Bob: I usually accept the reports (of the reviewers) but I interview every teacher myself. They appreciate that feedback. But I use my own observations as well.

Int: How do you mean?

Bob: I make sure I move around the school and talk to teachers about their work and their students. You get a very good idea talking to teachers on yard duty and so on. And there's evidence like VCE results.

Int: Do you ever go into classrooms to observe teachers' work?

Bob: No I don't do that. Don't need to. I have a pretty good idea.

### ***The Reviewers***

The reviewers saw themselves as mentors. They were very anxious that teachers would not think of them as judging or criticizing their performance. Although the actual interviews were short, the reviewers were prepared to spend extra time helping teachers to prepare their statements of performance and PD plans. The bulk of advice was in the area of PD. The school subscribed to a PD system that allowed teachers to choose courses and sessions from a range of advertised 'products'. Several teachers and reviewers cited ICT as an example of an area of personal 'weakness' that was picked up by the review processes. Reviewers helped these teachers to select a PD program suited to their interests and levels of expertise. Reviewers were thus, in one sense, 'brokers' between the teachers and the PD programs.

Every teacher interviewed said that he or she chose their reviewer because they 'felt comfortable' with that person. The reviewers stressed that they were trying to put teachers at ease and the teachers said that they felt at ease with their reviewers. While the teachers were glad of 'tips' to improve their PD plans and statements of performance, they did not believe that the reviewers had substantially helped them to improve their teaching. This was largely because, in most cases, the reviewers taught in a different subject area.

The teachers appeared surprised by the idea that a reviewer might become involved in a specific aspect of their teaching or work with them to help improve student learning outcomes in a particular Key Learning Area. Ross, the ICT teacher, for example said that there was no other teacher in the school with his expertise in ICT. He said that while it might have been 'nice' to have had his review with a person of equal or higher expertise, this was not possible in the school setting. In any case, he said, the 'comfort' factor was probably more important. He felt very 'comfortable' with his reviewer

You don't use jargon, you don't get caught up in your subject. You know there are limits.

Ross' first interview had taken ten minutes:

(The reviewer) corrected my English, made some changes, de-jargonised it, made it more user-friendly.

Ross' reviewer was the person in charge of student discipline. As Ross had few or no discipline problems with students, he had little to do with his reviewer in the normal course of his work.

Elsie said that she had great respect for her reviewer, a person with more than forty years teaching experience who was highly respected in the school.

Elsie: She was really fantastic, really supportive and she made some useful suggestions and gave me a focus. The structure helps her to support people and she's very good at that. Everyone says so.

Int: Did it (review meeting) help you to become a better teacher?

Elsie: (Surprised by the question) I don't think so. The meeting didn't focus on my actual teaching. It mainly affirmed what I was already doing. (The reviewer) teaches in a different area and the meeting only took about twenty minutes.

Int.: Do you have ongoing contact with (reviewer)?

Elsie: Well, yes, she's a sub school leader, but not really about the review or teaching.

Marita had the same reviewer as Ross. When asked how she prepared for her review, she said:

I started by just jotting down a few ideas before the meeting. I discussed them first with (the reviewer) and he was very helpful. I also talked to some other members of staff. Some of them were really good but some, especially the older ones, were really apathetic.

Marita did not think that the reviewer had sufficient time or evidence to judge her performance as a teacher. She was grateful for his reassurance but would have welcomed more specific feedback:

Marita: The only way he could have really known was to come into my classroom.

Int: Would you have been happy with that?

Marita: Of course, it's the only way. Or ask my students. I'd have no problem with that. They're the best people to ask and I know they like me.

Lily's reviewer was the person responsible for the daily operations in the school.

She said there had not been a 'wide choice' of reviewers and she had chosen that

person because she 'felt comfortable' (that word 'comfortable kept re-emerging) with him. Lily taught in three areas, the main one was LOTE. She said her reviewer was a maths teacher. Maths was not one of her own subjects.

The interviewer was a former LOTE teacher and a short conversation about teaching LOTE occurred, serendipitously, during one part of the interview.

Lily: That's really interesting. It's so good to talk with someone about that. I always feel I need more help with those bits (of the LOTE CSF). Thank you.

Int: So that's not the kind of conversation you'd have as part of your performance review?

Lily: Oh no, (the reviewer) is a maths teacher. He wouldn't know about LOTE.

Int: Would it be more helpful if he was a LOTE teacher?

Lily: . Yes, I suppose it would be. He can't really be a mentor. Yes, It would be very good feedback.

### **PART THREE**

#### **SUMMARY OF FINDINGS AND CONCLUSIONS IN RELATION TO THE CASE STUDY**

This section of the paper summarises the findings of the case study research and assesses the extent to which the system of performance review and development at Sundale High School reflect or diverge from conclusions drawn about effective practice in the first part of the paper. It also assesses the extent to which the staff performance and development system established in this school meets government defined objectives

#### **Perceptions of the *purposes* of the review**

There appeared to be tacit agreement and acceptance among the principal, the reviewers and the teachers that accountability as an aspect of quality assurance, was the major purpose of the review processes. The principal was interested in the notion of staff development as a purpose of teacher evaluation and concerned to ensure that staff development was occurring. But he did not give evidence of having specifically considered *how* teacher evaluation processes might be linked to improving teachers' practice, beyond referring them to appropriate professional development outside the school.

The reviewers said that they saw staff development as the main purpose of the review processes, although quality assurance was also a necessary and defensible purpose. However their main intentions seemed to be to make teachers feel 'comfortable' and to give them useful tips and hints about improving their practice. The generally limited amounts of time they were able to give to the processes meant that their suggestions for improvement did not go far beyond helping teachers to identify strengths and weaknesses and directing them to appropriate (out of school) professional development. They had used the standards *as categories* in which to make comments but they had not even started to think that the standards might be used to indicate *levels* of performance

The teachers said that they saw the main purpose of the review as being 'a check' to ensure that their practice met basic requirements. They were appreciative of the reviewers attempts to make them feel 'comfortable' and to act as their mentors. They saw this mentoring as helpful but generally did not expect that it would lead them to change their practice in the classroom or to significant professional learning.

The teacher evaluation literature reviewed in part 1 of this paper suggested

- a. That effective teacher evaluation systems enjoy the trust and confidence of all participants
- b. That all participants should have shared understandings of the evaluations' purposes
- c. That teachers are more supportive of teacher evaluation systems that are designed to serve purposes of teacher development than of systems designed for accountability purposes
- d. That in effective teacher evaluation systems, the purposes of quality assurance and teacher development are complementary.

There was considerable goodwill in respect of review processes at Sundale High School. The principal, the reviewers and the teachers trusted that all participants would play their role in an honourable manner that would disadvantage no one. This was no mean achievement.

Participants shared an understanding that the main purpose of the review was to ensure teaching of acceptable quality throughout the school. Contrary to some of the observations reported in the literature, teachers at Sundale did not appear concerned that performance *improvement* fell behind accountability as the main purpose of review.

There seemed to be tacit understanding among all participants that the processes should be as undemanding as the meeting of basic accountability requirements would allow. The processes were simple and time spent on them was minimal. 'Comfortable' was the adjective most commonly applied to the relationship between reviewers and teachers. The mechanics of achieving teacher performance improvement targets were neatly siphoned off in a process of identifying teacher strengths and weaknesses, then directing them to professional development programs outside the school.

The teacher evaluation processes at this school were, therefore, observed to be achieving *some kind* of quality assurance purpose. The question is, however 'What *sort*, or what *degree* of quality?' In the absence of more rigorous processes for defining and assessing quality it was difficult to define, let alone assess it. The lack of a systematic focus on improving the quality of teachers' work was both a result and a cause of this lack of definition. How can teachers improve the quality of their work if improvement is not a clearly shared purpose of review? How can they learn to perform work of higher quality if there are no shared understandings about quality and no shared perceptions of how it might be achieved?

Ensuring that improvement and quality assurance were equally important and complementary purposes of the review system, as the literature suggests, would have meant more work for the already overloaded administrators and teachers at Sundale. However, in view of the high level of goodwill and respect between participants in review processes, it may be possible in the future to focus more clearly on improvement as a goal. This would entail identifying the conditions that need to be in place to ensure the effectiveness and productivity of the evaluation. This would involve such changes as building professional learning teams with shared goals that align with professional standards and curriculum standards. It would include placing a responsibility on teachers to provide evidence of meeting the standards through, for example, engaging collaboratively in teaching tasks that encouraged reflection and self-assessment.

### **Teaching standards**

The (statewide) criteria used for the review processes were derived from the content standards developed for teacher performance review in Victoria. They were seen to be useful in that they indicated the *categories* in which teachers could be expected to perform - knowledge of content, knowledge of practice, etc. Teachers found reflecting on their performance and preparing their statement of performance against the standards a useful exercise which provided a sound foundation and structure for discussions with their reviewers.

There were, however, no mechanisms for assessing *how well* teachers performed in relation to the standards. There was little to suggest what sort and how much *evidence* teachers should present in order to be assessed against the standards. In this respect, the assessment was hardly a measured assessment at all, but rather a subjective judgement based on the reviewers' own tacit knowledge, in relation to each criterion, about the performance of the teachers whose performance was being judged

In the teacher evaluation systems discussed in the first part of this paper (NBPTS, INTASC, Praxis III WA Level 3) standards were embedded within professional learning systems for teachers. By systematically assessing teachers' performance using rubrics that differentiated between levels of performance, and precise evidence requirements, teachers and assessors were able to use the standards as practical 'hands on' tools for improvement. Quality assurance purposes were automatically taken care of within these processes, almost certainly more rigorously than in the processes currently being followed in Victorian government schools of which Sundale High School, was one example.

### **Roles of participants**

The literature clearly shows that improvement in teacher performance does not occur when systems of teacher evaluation are hierarchically managed and implemented. This was confirmed in the case study. The processes did not allow reviewers to engage with teachers in a disciplined, rigorous process that was directly related to teachers' daily problems of practice. The reviewers were therefore unable to influence teachers' professional learning beyond giving helpful advice, and suggestions for professional development. The relative superficiality of the process was underlined by the fact that the reviewers taught different subjects from those taught by the teachers whose performance they were reviewing.

It is not inconceivable that the Sundale teachers may have improved their practice as an indirect result of the review processes, by undertaking the recommended professional development. However this is a much less powerful means of improvement than the type of teacher evaluation, (NBPTS, INTASC, Praxis III, WA Level 3) in which assessment is by peers and teachers engage deeply in processes of reflection, assessment and development that directly relate to the technical core of their work.

The findings of the case study suggest that the system of staff performance and development instituted by the Victorian Department of Education and Training (DE&T) will only 'work' if school educational leaders and teachers are able to increase their capacity, in the form of requisite knowledge and skill, to make it work. Present indications are that schools are prepared to be accountable and that increased accountability will improve the quality of teaching in schools. Teachers *can* learn to move beyond outmoded notions of

bureaucratic accountability towards taking responsibility for their own professional learning and improvement, but it is unfair to ask them to do this if they are unclear about *why* they are doing it, *how* to do it and what are the roles and responsibilities of the participants in the various processes.

Arriving at an understanding of these issues is a major undertaking, and it is only a first step towards establishing an effective system of teacher evaluation. Putting an accountability system in place and leaving schools to implement it is, obviously, not enough. Without the requisite knowledge, skills and understanding, desired improvements will not eventuate. As Elmore points out, it is unlikely that systems will achieve the goal of improving the quality of teaching until they are prepared to invest in building the capacity of educational leaders and teachers to undertake the necessary tasks.

The imperative here is for professionals, policy-makers and the public at large to recognise that performance-based accountability, if it is to do what it was intended to do - improve the quality of the educational experience for all students and increase the performance of schools - requires a strategy for investing in the knowledge and skills of educators.....If the public and policymakers want increased attention to academic quality and performance, the quid pro quo is investing in the knowledge and skill necessary to produce it. (Elmore, 2002, p.5).

## References

Chamberlin R. P. Wragg E.C. Haynes, G. Wragg, C.M. (2001) (a). *Performance-related pay and the teaching profession: A review of the literature*. Background paper not presented at BERA Annual conference Leeds 14 September, 2001.

Chamberlin R. P. Wragg E.C. Haynes, G. Wragg, C.M. (2001) (b) *Head Teachers' Perceptions of Threshold Assessment*. . Paper presented at BERA Annual conference Leeds 14 September, 2001.

Darling-Hammond, L. Wise, A. E. & Pease, S.R. (1983) *Teacher evaluation in the organizational context: a review of the literature*. Review of Educational Research , 53, 3.

Danielson C. (1996), *Enhancing professional practice; A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling Hammond, (1999) *Reshaping Teaching Policy, Preparation and Practice, Influences of the National Board for Professional Teaching Standards*. American Association of Colleges for Teacher Education: New York.

Department of Education Employment & Training (DEET) (2001). *Victorian Government Schools. Staff Performance and Development Handbook*. Melbourne, Victoria.

Communications Division for Employee Relations Department of Education, Employment and Training.

Duke, D.L (1995). *The move to reform teacher evaluation*. In Duke (Ed.), *Teacher evaluation policy: From accountability to professional development*. (pp.1-12). Albany, NY: State University of New York Press.

Dwyer C.A. *Praxis III: Classroom Performance Assessments Assessment Criteria*. Princeton New Jersey Educational Testing Service.

Elmore, R. (1996). *Getting to scale with educational reform*. In S.H. Fuhrman & J.A. O'Day (Eds.), *Rewards and Reform: Creating Educational Incentives that Work*. Jossey Bass.

Elmore, R. (2002), *Bridging the Gap Between Standards and Achievement. The Imperative for Professional Development in Education*. Washington D.C. Albert Shanker Institute.

Elmore, (2000), *Building a New Structure for School Leadership*. Washington D.C. Albert Shanker Institute.

Haynes, G.S. Wragg, E.C. Wragg, C.M. Chamberlin, R. P. (2001). *Threshold Assessment. The Experiences and Views of Teachers*. School of Education and Lifelong Learning University of Exeter. Paper presented at BERA Annual conference Leeds 14 September, 2001.

Interstate New Teacher Assessment and Support Consortium (1992). *Model Standards for Beginning Teacher Licensing and Development: A Resource for State Dialogue*

Ingvarson, L.C. and Chadbourne, R. (1997). *Self-Managing Schools and Professional Community: The Professional Recognition Program in Victoria, Australia*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, March, 1997.

Ingvarson, L.C. (2001). *Strengthening the Profession: A Comparison of Recent Reforms in the USA and the UK*. Australian College of Education Seminar Series, ACE: Canberra.

Ingvarson L.C. (2001) *Development of a National Standards Framework for the Teaching Profession* An issues paper prepared for the MCEETYA Taskforce on Teacher Quality and Educational Leadership, ACER Melbourne.

Ingvarson L.C. (2002) *Building a Learning Profession*, Australian College of Educators Deakin West ACT

Kleinhenz, E. Ingvarson, L Chadbourne R. (2001) *Teacher Evaluation Uncoupled: a Discussion of Teacher Evaluation Policies and practices in Australian States and their Relation to Quality Teaching and Learning*

Paper presented at the AARE conference, Fremantle 2001.

Johnson S.M (1990) *Teachers at work. Achieving success in our schools*. New York: Basic Books.

McLaughlin M & Pfeiffer, R.S. 1988. *Teacher Evaluation, Improvement, Accountability, and Effective Learning* New York Teachers College Press

Medley D.M. & Coke, H. (1987, *The accuracy of principals' judgements of teacher performance*. Journal of Educational Research , 80, 242-247.

Peterson, K.D. (2000). *Teacher evaluation: a comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press Inc.

National Board for Professional Teaching Standards (1989) *Towards high and rigorous standards for the teaching profession*. Detroit. NBPTS.

Stronge, J.H. (1995), *Balancing individual and institutional goals in educational personnel evaluation: A conceptual framework*. Journal of Personnel Evaluation in Education, 21, 131-151.

Stufflebeam, D. (1988), *The Personnel Evaluation Standards* California, Sage publications

Shinkfield, A.J. & Stufflebeam, (1995), *Teacher evaluation: Guide to effective practice*. Boston Kluwer Academic Publishers.

Stiggans, R.J., & Duke, (1988), *The case for commitment to teacher growth: Research on teacher evaluation*. Albany, N.Y.: State University of New York Press.

Sykes G. & Plastrik, P. (1993), *Standard setting as educational reform*. Washington DC: American Association of Colleges of Teacher Education

Wilson, B. & Wood, J.A.(1996) Teacher evaluation: a national dilemma. Journal of Personnel Evaluation in Education 10, 1, 75-82.

Wise, A.E., Darling-Hammond, L., McLaughlin, M.W., & Bernstein H.T., (1984). *Teacher evaluation, a study of effective practices*. Santa Monica, CA: RAND

Yin, R.K. (1984), *Case study research: Design and methods* (Applied Social Research Methods Series, Vol 5). Beverly Hills, CA: Sage.