TEACHER EVALUATION UNCOUPLED

A DISCUSSION OF TEACHER EVALUATION POLICIES AND PRACTICES IN AUSTRALIAN STATES AND THEIR RELATION TO QUALITY TEACHING AND LEARNING

Elizabeth Kleinhenz ACER
Lawrence Ingvarson ACER
Rod Chadbourne Edith Cowan University

Abstract

There is now broad acknowledgement that student learning is directly linked to the knowledge and skills of teachers. It is a short step from this understanding to recognising the importance of teacher evaluation in schools. The education systems in all Australian states now have teacher evaluation policies and practices in place, but it is as yet unclear whether the directions that are currently being taken will result in the desired improvements in teacher performance or even whether basic goals of accountability will be met. ACER and Edith Cowan University are currently engaged in the second year of a three year research project that maps teacher evaluation practices across all Australian states and territories. This paper reports on some of the findings of the project. A critical discussion of the "top down" performance management systems that are linked to pay increases in some states is offered and three specific and relatively recent initiatives in teacher evaluation at the "accomplished teacher" level are discussed in some detail. These are the evaluations for the Victorian Experienced Teacher with Responsibility (ETWR) and Western Australian Level 3 Classroom Teacher positions, and the UK "Threshold" classification. It is argued throughout the paper that progress will only be made when the limitations imposed upon teacher evaluation by "loose coupling" are overcome and teachers are encouraged to participate in the development of evaluation systems that respect to the complexity and depth of their professional knowledge and practice.

Introduction

For a very long time, most teachers in Australian schools have tended to resist the notion of opening up their work to professional or other scrutiny. The reasons are not hard to find. They are grounded in a culture forged by historical resentment of attempts to "inspect" the work of teachers. Yet teachers do understand that effective evaluation is necessary for improvement. They know, for example, that sound methods of assessment are essential for improving the educational performance of their own students. And they are starting to see the inevitability and reasonableness of becoming more accountable and of improving their teaching practice on the basis of reflection and review.

Research continues to tell us that quality of teaching is the most important factor in successful student learning and indeed it only makes sense, as Elmore remarks (2000, p.9) that students will learn if they are effectively taught. Clearly, if the goal of education is to improve student learning, and if teacher performance is the most important factor in this, there is an urgent need to seriously and effectively evaluate the work of teachers for the twin purposes of public accountability and improvement.
This paper is based on some of the findings of a research project, begun in 2000 and due for completion in 2002, that set out to discover how well Australia is being served by current approaches to teacher evaluation. The first aim of the project was to document existing teacher evaluation practices in Australian schools and school systems, the second to identify the quality of some actual teacher evaluation practices in terms of a range of criteria and the third to investigate the relationship between some current teacher evaluation practices and the quality of teaching and learning by means of case studies in a small number of schools from different school systems.

The paper presents an overview of project findings to date and argues for an approach to teacher evaluation that closely involves teachers themselves and draws on their knowledge and skills as a professional group, as opposed to being based on a set of bureaucratic procedures that are done to or imposed upon teachers. The voice of teachers through relevant professional bodies, it is argued, needs to be heard in teacher evaluation processes that occur at every stage of a teacher's career, from pre service through induction and career progression to the phase of highly accomplished teaching. The paper examines and compares in some detail the evaluation procedures for the Victorian Experienced Teacher With Responsibility (ETWR) and the Western Australian Level 3 Classroom Teacher, (L3) classifications, which, in some respects, exemplify bureaucratic, (ETWR) and professional (L3) models. It also discusses some of the findings of recent research (Chamberlin et. al.2001), on the teacher evaluation procedures, (bureaucratic), used to assess applicants for the "Threshold" classification in the UK, arguing that shortcomings of the Threshold evaluation procedures are the result of "loose coupling" between management and the "technical core" of teaching as described by Elmore (2000). The exposed flaws of the British system serve as a warning against the use of similar teacher evaluation methods in Australia, some of which also appear to operating in an educational context that is characterised by loose coupling.

"Loose coupling"

Elmore's (2000) discussion of loose coupling distinguishes between the "technical core" of education systems and the "administrative superstructure." The core includes the essential features of teaching and learning: making decisions about what should be taught and how, grouping students, planning learning schedules and evaluating student achievement. According to loose coupling theory, this technical core, says Elmore, is surrounded and sheltered by an "administrative superstructure" which comprises management and administration at the school and higher levels. The crux of the loose coupling argument as it relates to education is that because the technical core of education systems is weak, disordered and uncertain, largely because of doubts about the status of teachers' professional knowledge, (Bidwell, 1965, Lortie, 1975,) the surrounding administrative arrangements prefer to ignore it thereby acting mainly to shield the fragile core from external scrutiny and criticism.

Elmore draws a distinction between management of the structures and processes that surround the technical core of education and management of the core itself. Although it may seem that educational management is about managing the processes of teaching and learning, (the technical core,) it is, in reality, nothing of the sort. Educational management, according to loose coupling theory, manages things like tracking systems and athletic programs and numbers of students in schools, but leaves the technical core of teaching and learning to the idiosyncratic practice of individual teachers who are "buffered" from outside interference by the very structures which, ostensibly, were set up to manage their work. This buffering has the effect of falsely reassuring the public that all is well at the technical core.
Buffering consists of creating structures and procedures around the technical core of teaching that, at the same time, (1) protect teachers from outside intrusions in their highly uncertain and murky work and (2) create the appearance of rational management of the technical core, so as to allay the uncertainties of the public about the actual quality or legitimacy of what is happening in the technical core. This buffering creates what institutional theorists call a "logic of confidence" between public schools and their constituents. (Elmore 2000 p.6.)

The theory of loose coupling, Elmore continues, explains why most innovation occurs at the system level rather than in the classroom. Such innovation is often at best insensitive to the technical core of teaching and learning and at worst may do damage, e.g. alienating students in over large schools, or establishing grouping practices that exclude some students from learning, or setting up athletic programs that prevent students from engaging in other forms of learning. The most telling example of all is perhaps that of site based management structures that discuss "everything about everything except the conditions of teaching and learning." (Elmore 2000 p.6)

**Loose coupling and teacher evaluation**

A logical extension of these arguments is that where teacher evaluation procedures are designed and implemented at the level of the administrative superstructure rather than at the technical core, the core itself is buffered from the kind of scrutiny that may contribute to its improvement, and is therefore likely to remain weak. Ingvarson, (1994), has noted two "imperatives" for teacher evaluation: the need to safeguard the educational interests and welfare of all students (public accountability); and the need to ensure that teachers continually review and improve their practices in the light of contemporary research and professional standards. (Professional accountability.) It can readily be seen, in the light of loose coupling theory, that administratively conceived evaluation procedures that do not address and include the core of teaching practice - do not involve teachers in the development of standards of practice for example - will satisfy neither or these two imperatives. In fact they are likely to have the opposite effect, in line with Elmore's arguments, of protecting the core from examination and criticism, i.e. they will fail to provide public guarantees of quality, and will discourage reflection and review among teachers themselves, thereby allowing, and even by default causing, teachers' knowledge and skills to remain weak and uncertain.

**Research project findings**

The mapping phase of the study has documented teacher evaluation practices in all Australian states at four or five stages of a teacher's career.

1. The pre service or initial teacher education phase
2. The phase of gaining first employment as a teacher, (effective provisional entry to teaching)
3. The induction stage, (effective full entry to teaching)
4. The career progression stage (the stage of annual progression through the incremental salary scale, usually 10-12 years.)
5. The highly accomplished teacher phase. (This is not applicable in all states.)

**Phase 1 the pre service phase:**

The purpose of this evaluation or series of evaluations - the first of a teacher's career- is to gain a tertiary qualification that is acceptable for employment as a teacher. The expectation
is that the evaluation will test the professional knowledge and skills of the applicant on the basis of sound evidence and that the successful graduate will be equipped with the necessary knowledge and skills to work as a professional.

In all Australian states the education faculties of universities decide their courses of study and the evaluation tasks and criteria for their teaching degrees and diplomas. Our study found that various statutory committees, whose reporting lines are not clear, accredit university courses. At present, although practising teachers may be members of these committees, they are generally not a majority. This contrasts with accreditation arrangements for other professions, where professional bodies exercise course accreditation functions. In states that have Teacher Registration Boards, those Boards may exercise a course accreditation function.

How teacher preparation courses are to be evaluated and accredited, and who should be the body or bodies responsible, is a central issue for this phase. If teacher education course accreditation structures are only loosely coupled with the realities of the classroom - the technical core - there is little guarantee that the courses will reflect these instructional realities or that that evaluations of new teachers will be conducted on the basis of what really counts as knowledge about teaching. A vital issue for this phase is therefore to ensure that the various course accrediting structures are properly equipped to perform their function. Professional bodies which, like the proposed Victorian Institute of Teaching, comprise a majority of teachers, and teacher registration boards can play an important role in ensuring the quality of teacher preparation courses and the reliability of the credentials they issue.

**Phase 2: First employment as a teacher.**

There is a fine line at this stage between "selection" and "evaluation": applicants are obviously being selected for a job, but they are also in effect going through an evaluation process that is crucial because it **gives them provisional entry to the teaching profession.** In the course of our investigations it was found that, especially in the larger and more centralised state education systems, this evaluation tended to be perfunctory and hardly thought of as an evaluation by the employer. Understandably perhaps, given the large numbers of teachers and vacancies, emphasis was on satisfying administrative needs such as matching teachers with subject and stage of schooling positions rather than on making judgements of teacher quality, the latter having been assumed on the basis of tertiary results.

Standards used as the basis for making employment decisions in all states were, for the most part in the form of brief and generic criteria. These standards were used more as administrative check lists than as attempts to probe applicant ability. "Passing" the evaluation meant getting a job and this was often governed more by job availability than merit e.g. it was easier to get a job in a hard to staff subject area like Maths or LOTE or in a remote geographical area. There was an obvious cross over between employability and quality, with employability rather than quality being the main criterion for provisional entry to the profession, a situation that is unlikely to improve if teacher shortages worsen.

In Victoria, where, as part of the devolution of school management to school site level, this evaluation is conducted by principals, usually with the assistance of a panel, at the school level, it is probable that the evaluation touches more closely on such features of the technical core of teaching as the applicant's relationships with students, knowledge of curriculum planning and assessment, and knowledge of the subject to be taught. Schools, after all have a strong vested interest in getting the "best" teachers. (One Victorian bureaucrat interviewed said that such evaluation was the most "rigorous" of all, for that reason.) But in individual schools as in whole education systems issues of employability may
take precedence over quality, the criteria are likely to be imprecise and principals as evaluators generally lack suitable training in evaluation methods.

**Phase 3 Induction**

All state education systems recommend that schools carry out formative appraisal of teachers on "probation" and most employers require a summative evaluation to be conducted by the principal after one or two years of teaching to end the probation period. Requirements for evidence vary, but it is generally recommended that teachers provide material evidence of effective teaching such as lesson plans and samples of assessed student work. Classroom visitations as part of evaluation are not the norm, but do happen in some schools and systems.

These requirements for formative and summative teacher evaluation during the induction phase are generally set out in the form of brief "guidelines" for principals, who, in all states, have the final responsibility for induction and end of induction summative evaluations. Improvements in the printing and graphic design industries have served the administrative superstructure well by setting out these guidelines in visually attractive form, but the smart packaging does not disguise the generally superficial nature of assessment criteria which in every case fail to sufficiently address the complexities and idiosyncracies of the technical core - what teachers should be expected to know and be able to do at this stage.

It is one thing for education departments to issue guidelines and quite another for the guidelines to be followed in schools. There is evidence of great variation between schools in the way induction is carried out, from virtually no induction at all through to demanding and elaborate processes in some schools. It was acknowledged by all interviewees in our project that while some principals were meticulous in setting up induction processes and processes for formative and summative evaluation, others simply "signed off" the teacher at the end of the probation period. Several interviewees doubted whether some teachers would even know that they were being evaluated at this stage. There are figures to show that in all states very very few teachers fail the summative evaluation at the end of the probation period. There are thus absolutely no guarantees that in all Australian schools teacher evaluation procedures at this phase meet public accountability or quality assurance standards. Nor can it be said that the existing procedures are helping new teachers to review and improve their performance, except in those schools, (probably a minority), that have made a conscious effort to develop their own appraisal systems. Common sense backed up by loose coupling theory tells us that neither public nor professional accountability requirements can be met in all schools as long as these processes fail to reach the technical core of teaching practice and are left to be implemented by poorly supported, untrained (in evaluation methods), and overworked school staff.

**Phase 4 Career progression**

There is an expectation in all Australian state government schools that teachers will go through processes of regular performance appraisal which are the responsibility of the principal. Like the guidelines for induction, (which may also be part of overall performance review systems), the details for these systems of teacher performance management, including standards for performance, may be attractively presented in impressive publications, but this does not disguise their failure to adequately address the real issues and complexities of teaching and learning.
As with processes for the induction phase, formative and summative assessment is generally recommended for teacher performance appraisal together with the recommendation that teachers be required to present written evidence of competence, e.g. unit and lesson plans, evidence of achievement of agreed goals or of professional development undertaken, at one or more formal interviews with the principal and/or line manager. Standards for performance review have been developed by most state Education Departments and in some states there has been agreement with the AEU to use these standards for the purposes of Annual Review linked to performance pay. Such measures represent interesting attempts to assess and improve teachers’ performance by linking pay to improvements in teaching, but measuring improvement is only as good as the standards and methods used to do so. All standards used in Australian teacher performance management systems are generic, and their quality, as measured against commonly accepted standards for evaluation in education. (E.g. Joint Committee on Standards for Educational Evaluation, 1988.) is questionable. Certainly the evaluation implementation methods show little in the way of consistency or expertise in evaluation procedures.

A teacher’s failure to demonstrate satisfactory performance can lead to dismissal, but it was found that the number of such dismissals in all states is very low and is usually not the result of the appraisal process itself, but a response to concerns raised by, e.g. colleagues or parents. There appears also to be wide variation, not only among state systems but also among individual schools in the ways performance review is carried out. The AEU in Victoria, for example, where salary increments are linked to performance, suggests to its members that full review processes, including an interview, should only happen where there is doubt about a teacher’s performance, yet many schools choose to maintain comprehensive Review processes linked, in some cases, to teachers’ professional development plans. In some schools teachers themselves have taken the initiative to develop appraisal systems whose chief aim is to reflect on and improve teaching practice. We will investigate some of these teacher developed appraisal programs in the third stage of our study, but it is expected that they will not be easy to find since such programs have no official status in any state education system.

The incentive of annual salary increments linked to performance review does not seem to be particularly effective. The majority of teachers in all states are already at the top of the incremental salary scale, (this takes about 10-12 years,) so this incentive does not apply to them. Of those teachers who are eligible, very very few fail to access annual increments. This suggests a parallel with the British Threshold evaluation, (see below), where, because almost 100% of teachers "passed" it has been queried (Chamberlin et.al 2001) whether the results justified their trouble and expense.

This phase covers a substantial period of a teacher’s career. Yet, our research so far is suggesting that education systems in all Australian states would be hard put to demonstrate that the work of teachers in this period is being evaluated in ways that address the core of teaching practice. This is partly because of the great variation in teacher appraisal practices among schools. But it is also largely because of the absence of standards that adequately explicate the real work of teaching - what it is that teachers can be expected to know and be able to do in specific domains of practice. It is now widely accepted that comprehensive, congruent, domain specific standards provide the only credible basis for making useful judgements of teacher competence. (Ingverson 2001, Berliner, 1992, Shulman 1987, Brophy, 1991). Without such standards it is unsurprising that the technical core of teachers’ knowledge and practice remains weak.
**Phase 5 Highly Accomplished Practice.**

In the years after the 10 or 12 year career progression stage some teachers are promoted to administrative positions up to and including principal positions. Evaluation for these better paid and higher status positions is largely on the basis of administrative competence, (which raises interesting questions about how higher rewards escape classroom teachers. If systems reward what they most value, what does this say about teaching?) Of those teachers who choose to simply continue teaching, little in the way of evaluation seems to be demanded by most schools and systems. Not a lot is known about the work of these shadowy creatures who occupy the netherworld of the classroom, and their practical knowledge and skills are seldom recognised, let alone reflected in standards for teacher evaluation.

In some states there is a teacher classification between the highest point on the automatic salary scale and administrative classifications. This type of position exists to reward teachers who choose to remain in the classroom rather than move into administration. Examples include the Level 3 Classroom Teacher position which was introduced in WA in 1997 and the Experienced Teacher with Responsibility position, brought in in Victoria in 2001. There are also positions at this level in Queensland and South Australia, (Advanced Skills Teacher) and the Northern Territory, (Teachers of Exemplary Practice.) In order to explore some of the features of evaluations for positions at this level for this paper it was decided to "zoom in" on those for the Western Australian Level 3 Classroom Teacher and for the Victorian Experienced Teacher With Responsibility (ETWR) classifications.

**Comparison between evaluations for the Level 3 Classroom Teacher Classification (WA) and the Experienced Teacher with Responsibility (ETWR) (Vic)**

The Final Report on the first (trial round) of teacher evaluations for the Level 3 position, prepared by Jasman and Barrera, (1998,) notes that the developers of the Level 3 teacher classification in WA addressed two suggestions of Ingvarson and Chadbourne, (1994) that recognise the limitations of an application/interview process for teacher assessment. These are

1. that evaluation should be summative, criterion based and use multiple sources of data to demonstrate that the evaluatee has attained a particular standard of professional knowledge and skill, and
2. that teachers should be assessed by a "college of specialists." Wallace, Wildy and Louden, (Unicorn May 1999,) note that this is consonant with current research and writing on teacher career models.

The evaluation was conducted in two stages. In the first stage, applicants were required to present a portfolio of evidence to be assessed by teachers who, in the first round of appointments, (1997) were selected on the basis of recognised ability and in the second, (1999), had demonstrated their competence by succeeding in the L3 evaluation themselves, (i.e. they were now Level 3 teachers.)

In the second stage of the evaluation, applicants who had succeeded in the portfolio stage were assessed further in a half hour presentation in a group situation, thereby providing a further source of evidence.

Evaluation processes for the ETWR were on the basis of an application to the principal addressing statewide criteria and, in most schools, an interview with the principal or nominee and a panel. There are other significant differences between the evaluations for the two positions which are discussed under the headings below:
1. Visions of educational leadership

Elmore draws a clear distinction between the respective spheres of influence of administrative and educational leadership. In the first, the administrative sphere, the work of leaders is confined to the "ritualistic" tasks of management - The second sphere is the mysterious world of teaching and learning "where teachers, working in isolated classrooms, under highly uncertain conditions, manage the technical core."

Selection procedures for the ETWR position do not reflect an understanding of the notion of educational leadership and give scant recognition of the existence of a technical core of teaching practice. This is clearly seen in the suggestions for the extra ETWR "jobs" in the school which all successful applicants are required to negotiate with their principals after the selection processes have been completed. The "key role" of the ETWR was described in the guidelines for the position as "an excellent classroom teacher and...a role model and mentor for other teachers in the school" (DEET 2001 p.2) but the suggested extra ETWR jobs include managing first aid, sport, camps and fundraising! (DEET 2001(a) p.3.) Such suggestions that these senior teachers, as a result of promotion, should be required to perform jobs that could by done by any reasonably competent clerk not only demean teachers and the occupation of teaching, they demonstrate a classic misunderstanding of the role of an educational leader in a school.

Conceived at administrative levels and implemented by principals, the ETWR evaluation is something done to teachers - they had no say in establishing the criteria for assessment, they are constrained by the criteria ( and by a page limit) in their application and there is insufficient time in one interview, (in some schools the principal must interview 30 or more teachers within a limited time so the interviews are necessarily short and confined to basics,) - to present evidence of what they are trying to achieve, how they are achieving it and how they might share their expertise with other teachers..

In contrast, the W.A. L3 concept and related evaluation processes provided an example - certainly the best discovered in Australia so far in the course of the project - of a sustained attempt to discover leaders capable of mentoring other teachers and innovating at the educational core. The clear aim was to set in place an educational leadership position that would bring about changes in teachers' ways of working in order to improve the core functions of teaching and learning.

Our interviews with a senior Education Department of Western Australia administrator and the president of the Western Australian branch of the Australian Education Union and discussions with L3 teachers showed that these key stakeholders in the L3 classification shared a remarkably clear recognition of the essential differences between administrative leadership and educational leadership and a determination that the L3 position should be about the latter. It was clearly understood that successful applicants would be called upon to set examples of excellent teaching, provide other teachers with advice and mentoring, conduct programs of professional development and allow other teachers into the classroom to observe their work.

While it was expected that L3 teachers would negotiate an extra leadership role with their principals, it was spelt out very clearly that this role was to be in the area of mentoring and leading other teachers to reflect on and improve their teaching knowledge and practice and not in any administrative area. This clear goal was reflected in the evaluation processes which sought evidence of teaching excellence and of the capacity to share demonstrated knowledge and skills.
2. Evaluation methods and evidence required

It was the stated intention of the Victorian Education Department that the rewards attached to the ETWR position were to be in recognition of teaching excellence, i.e. excellence in those attributes that make up the technical core of teachers' knowledge and practice. (DEET 2001.) It therefore seems reasonable to expect that evaluation processes for the position should have taken the form of rigorous explication and investigation of those attributes that belong to the core, together with requirements for multiple sources of evidence of competence in relation to those attributes. Such cannot said to have been the case for the ETWR evaluation where the application/interview process used was almost identical to the processes used to employ a secretary or cleaner! The only evidence required, apart from the teacher's own declarations of competence in relation to the criteria, was in the form of (usually verbal), reports from referees.

Evaluation processes and evidence required for the L3 position were more rigorous and certainly more focussed on the core of teaching and learning, although they lacked the sophistication of, for example, the methods used in the assessments of the US National Board for Professional Teaching Standards. Completing a portfolio with a statement backed up with the teacher's own choice of evidence for each generic competency, though much less random than the ETWR methods, is still somewhat haphazard, as the applicants are given no indication of best practice in the relevant domains.

The reflective review stage of the L3 evaluation drew considerable criticism, especially from unsuccessful applicants. A common complaint, even from successful applicants, was that it advantaged teachers who were good at "selling themselves", which was not one of the criteria. It needs to be noted, however, that this stage of the evaluation does provide an additional and qualitatively different source of evidence. While the evidence for this and the portfolio stage may be both qualitatively and quantitatively insufficient to enable evaluators to make well grounded judgements, it nonetheless represents a serious attempt at requiring applicants to present more than a single kind of evidence.

3. Assessors

The person with responsibility for the summative evaluation for the ETWR position, upon which selection depends, is the school principal who may choose to delegate some of the work to a nominee and/or to make use of selection panels. Having the principal and/or other teachers at the school as the assessors has some advantages, chief of which is that these people see the applicant's work in its actual context, making it easier to identify those contextual features that could influence the applicants performance. (Joint Committee on Standards for Educational Evaluation, 1988 p.90.) Against this must be weighed the disadvantages that arise from the adverse effects on evaluations of complex micro political issues in schools. Available data suggests that the ETWR success rate in some schools is much higher than it is in others. This suggests difficulties that may well originate in the inclinations and personalities of principals.

Even if micro-political and personality factors could be dealt with, it is difficult to see how principals could be expected to carry out a thorough assessment of every applicant for this classification. The vast majority of principals have not been trained in evaluation methods, nor do they have the comprehensive knowledge of method and content in relevant areas of teaching practice that is required for effective teacher evaluation.
In contrast, the assessors for the L3 are now teachers who have successfully completed the evaluation themselves. They are generally not personally known to the teachers who are being evaluated, so that it is arguably less likely that bias will occur. They are given two days training in evaluation methods, but the assessors we spoke with believed this to be insufficient. Even so, their two day training is two days more than most Victorian principals have had!

Because they have been evaluated themselves under these processes and because they are currently working as L3 teachers, it is likely that the L3 evaluators will have a far deeper understanding of the complex teaching and educational leadership qualities that are being assessed than school principals under the ETWR arrangements. The L3 evaluation is thus very much about the core of teaching and is carried out collegiately by people who themselves work at the core.

In neither system do the assessors necessarily have experience or knowledge in the same subject areas or levels of schooling as the applicants. For the L3 evaluation, it is quite possible, for example, for a teacher of Year 1 or Beginners to be assessed by a teacher of Year 12 Physics, or vice versa. This is not in accord with the findings of current writing and research on teacher evaluation, which recommend domain specificity for both standards and evaluators. It takes little imagination to see that such domain specificity has a far better chance of probing the technical core of teaching than the application of generic principles across many disciplines and year levels.

4. Standards

The standards (called "competencies" for the L3 and "selection criteria" for the ETWR) that are used for evaluating applicants for the two positions are similar. They are:

**ETWR Selection criteria**

1. Demonstrated high level of knowledge of relevant curriculum areas, high level classroom teaching skills and a proven capacity to improve student learning
2. Significant contribution to the development, implementation and evaluation of curriculum programs and policy in accordance with the key goals and priorities in the School Charter and proven capacity to respond at the school level to initiatives that enhance student learning
3. Demonstrated commitment to ongoing professional learning and proven capacity to model excellent teaching skills within the school
4. High level communication and inter-personal skills when relating to students, parents and work colleagues
5. Demonstrated productive contribution to the school program

**Level 3 Classroom Teacher competencies**

1. Utilise innovative and/or exemplary teaching strategies and techniques in order to more effectively meet the learning needs of individual students, groups and/or classes of students. (Core idea: teaching.)
2. Employ consistent exemplary practice in developing and implementing assessment and reporting processes. (Core idea: Assessment.)
3. Engage in a variety of self-development activities, including a consistently high level of critical reflection on one’s own teaching practice and teacher leadership, to sustain a high level of ongoing professional growth.(Core idea: reflection.)
4. Enhance teachers’ professional knowledge and skills through employing effective development strategies. (Core idea: Working with Others)

5. Provide high level leadership in the school community through assuming a key role in school development processes including curriculum planning and management, and school policy formulation. (Core idea: School development.)

The ETWR criteria and L3 competencies reflect a view of teaching as both an individual and collective enterprise. (Sykes 1990.) Many WA teachers saw the competencies as "skewed away from classroom activities and towards school wide responsibilities" (Wallace Wildy and Louden 1999, p. 8.) It is certainly clear that both sets of "standards" are not confined to classroom practice, but, given the defensibility of the view that teaching is a collective rather than individual endeavour, the main shortcoming of the standards is not so much this as their lack of coherence and failure to capture the essences of what good teachers should know and be able to do in specific domains of practice. What a good teacher knows and is able to do in order to teach a young child to use multiplication tables is different from what a good teacher knows and is able to do in order to lead a discussion on the French Revolution.

A simple comparison between the above standards and the following example from the NBPTS Adolescence and Young Adulthood/Mathematics standards, (National Board for Professional Teaching Standards 1996.) illustrates this point:

111. Knowledge of Mathematics

Accomplished mathematics teachers have a broad and deep knowledge of the concepts, principles, techniques and reasoning methods of mathematics that they use to set curricular goals and shape their teaching. They understand significant connections among mathematical ideas and the applications of these ideas to problem-solving in mathematics, in other disciplines and in the world outside school. (NBPTS 1996, p.13.)

It should also be noted that unlike the ETWR and L3 criteria and competencies, NBPTS standards (and standards developed on the NBPTS model) are "cast in terms of actions that teachers take to advance students outcomes (and )incorporate the essential knowledge, skills, dispositions and commitments that allow teachers to to practice at a high level." (National Board for Professional Teaching Standards, 1996, p.1.). When judged against such sets of standards, it can be seen that neither the ETWR criteria nor the L3 competencies adequately reflect what accomplished teachers should know about helping students to learn. They are open to the same criticism as that which has been directed at the criteria developed by the private consultancy firm of Hay /McBer for the British Threshold classification.

… the Hay McBer research does not reflect well what highly accomplished teachers know about how to help students learn what they are teaching. Its criteria do not identify what effective English teachers know about how to help students write better or what effective Science teachers know about how to probe a student’s initial beliefs about a concept in science, and how to use that knowledge to anticipate and deal with possible misconceptions. These generic characteristics of effective teaching lend themselves to the development of observational checklists and other evaluation methods for managers to use who may not have expertise in the field of teaching they are assessing. (Ingvarson 2001. P.6).
Overall, the evaluation processes for the L3 position were seen to have the following advantages over those for the ETWR:

- The L3 processes showed a much clearer vision of educational leadership than those for the ETWR, and a built in expectation that this vision would be reflected in the evaluation for and the implementation of the role.

- The evaluation methods for the L3 position called for a greater variety and quantity of evidence than the ETWR processes and employed dual assessment methods, thereby providing a broader base upon which judgements could be made.

- The fact of the L3 evaluation being held at a neutral venue with assessors who were not known to the applicants, (or if known were required to state this), reduced problems arising from conflict of interest or school micro politics and increased the probability that processes would be fair across the system. There appear to be serious issues of fairness and equity in relation to the ETWR processes.

- The use of "trained" peer assessors in the L3 evaluation processes probably provided more likelihood of sound judgements than the principal/school based panel ETWR processes. As teachers who had been successful in the evaluation themselves, the peer assessors could be expected to have a better knowledge and understanding of the competencies, the processes and evidence required than principals and panels in the ETWR processes.

Comparison of the L3 and ETWR with evaluation for the British "Threshold" classification.

There are striking similarities between the L3 and ETWR evaluation procedures and those of the British Threshold, and, especially, between procedures for The Threshold and the ETWR. All three positions are designed to reward with a substantial salary increase the same category of teachers, - those who have reached the top of the level of salary progression that precedes administrative levels. The intentions of all three classifications are similar - to reward teaching excellence, to encourage good teachers to remain teaching, and to promote educational leadership. For both the Threshold and the ETWR the process is essentially one of application and interview on the basis of brief generic "criteria."

The results of research at Exeter University (Chamberlin et. al 2001), reveal precisely the kinds of flaws that can be expected of teacher evaluation practices that are conceived and executed at an administrative level Among the issues perceived as "of concern" by the Exeter researchers are:

- The suitability of a form filling exercise for assessing whether teachers have met the Threshold "standards"
- The unrealistically high (97%) "success" rate
- The different ways in which the Threshold Assessment procedure is carried out in different schools
- The need for better training on how to complete the application form and more in-house advice and support for this task
- The basis on which some head teachers make judgements about an individual's performance
o Possible bias and victimisation by head teachers
o The overwhelmingly negative impact of the processes on unsuccessful teachers' motivation, self-esteem and, in some cases, general health

The researchers noted that "to tie pay to performance raises an expectation that some changes in practice should occur". (p.6) Yet only a small minority of teachers in the sample believed that the Threshold Assessment had caused any change to their practice and of these the "vast majority referred to increased record keeping: A significant 40% of unsuccessful teachers said that the procedure had had a "detrimental" effect on their performance.

As an initiative designed to boost staff retention and recruitment, the Threshold assessment appears to have been a failure. Sixty eight per cent of teachers surveyed in 2000 said that the assessment had had a quite" or "very" negative impact on their sense of well being. Head teachers also reported damaging effects on staff morale.

Reports that are now coming in on the ETWR (AEU News 2000) are similar to those documented by Chamberlin et. al on the Threshold: Eighty five per cent of ETWR applicants were successful in the first round, but this does not include over 500 which are still in the Appeals process. There are claims of discrimination, e.g. at one High School the principal is said to have changed the selection panel's recommendations to falsely suggest that the panel had failed the application of an AEU sub-branch representative. Some principals, it is claimed, discouraged certain teachers from applying. The AEU received complaints of some applicants being required to undergo psychological tests and some being required to give powerpoint presentations at their interviews. In some schools, the principal did not bother with interviews at all. There were huge inconsistencies in the results, with some schools "passing" every applicant and others awarding success to only a few. In true Australian sport loving fashion, one wit circulated a scorecard that compared the ETWR results of his and a neighbouring school: "The ETWR Cup: Kurnai 26 def. Lowanna 12. Do we blame the players or the coach?" (AEU, 2001, p.3.)

Conclusion

Despite broad acknowledgement that successful student learning is primarily dependent on the knowledge and skills of teachers - the core of teaching practice - the results so far of the study described in this paper indicate that there is much room for improvement in the evaluation methods used to evaluate teachers' performance as teachers in Australian states. Loose coupling theory tells us that it is unlikely if not impossible that teachers' work will improve in the absence of the kind of evaluation that explores and critically assesses performance on the basis of the reality of what teachers know and are able to do – the "technical core" of practice. The same theory lends itself to the suggestion that bureaucratically conceived and executed systems of teacher evaluation - systems that fail to reach the technical core of teaching - will not truly fulfill their ostensible purpose of meeting public accountability requirements. While they may purport to provide the public with assurances of teacher quality, such assurances, because they are based on inadequate processes and information are certain to be suspect.

The technical core of teaching will not be reached for evaluation or any other purpose without the full participation and co-operation of teachers. This is true today even more than in the past. As curricula become more demanding, goals for student learning more ambitious and the social contexts of classrooms more uncertain, the work of modern teachers lends itself less and less to evaluation by traditional bureaucratic means. Assessment that is
done to teachers by bureaucrats or administrators is a waste of time. The results of teacher evaluation for the British Threshold, the Victorian ETWR and for most performance management systems are now telling us that this kind of assessment verges on farce because it is based on superficial standards, because it is carried out by untrained and overworked managers who lack the necessary knowledge and skills and because it relies on ill formulated evidence which lacks credibility. These systems are proving to be incapable of providing guarantees of teacher quality. Because they alienate and annoy teachers and, in many cases, principals, they do not lift morale or encourage teachers to improve their performance.

The best example of teacher evaluation found so far in this project has been in the methods used for the WA Level 3 classroom teacher position. The major strength of these evaluation methods was seen to be the contribution and participation, at most stages of the process, of teachers with proven high levels of teaching knowledge and skills. Even this evaluation, however, fell short of involving teachers in the central task of developing professional standards that could be relied upon to show teachers what it is that they need to know and to be able to do in order to become better teachers. This was understandable, if not inevitable, since there is not, as yet, in Western Australia or any Australian state a professional body which could take up this responsibility, "an institution that will enable the teaching profession and education authorities to talk with each other on equal terms and to exercise their shared responsibility for the quality of teaching and learning in schools." (Ingvarson 2001.p.1).

The relationship between building a responsible teaching profession and establishing more effective systems of teacher evaluation is becoming clearer. This relationship is gaining the recognition of policy makers, as evidenced by current initiatives to set up Colleges or Institutes of Teaching in several Australian states, and in the recommendations of various reports, including the NSW Ramsey Report (2000). In seeking a way forward we will need to abandon the ultimately futile bureaucratic teacher evaluation methods of which UK Threshold provides such damning evidence of failure. The theory of loose coupling helps us to understand how such evaluation methods fail because they all but ignore the core business of instruction, with all its idiosyncracy and complexity. Only practising teachers know what happens at the core of teaching and learning and only practising teachers, as Elmore says, are currently managing the core, whatever their "managers" may think.

The way forward, therefore, is to start not at the level of the administrative superstructure but at the instructional core, by encouraging teachers themselves to build a critical, knowledgeable and dependable profession that will assume responsibility for developing standards of practice as the basis for evaluating the work of its own members. By these means, the fragile core will be strengthened, teachers will, at last, know what it is that they should be expected to know and be able to do in order to improve and they will receive collegiate support to do so. Only then will it be possible to provide the public with genuine guarantees of quality in the area where it matters most - the area of teachers' professional knowledge and skills.
References

Australian Education Union (Victorian Branch) Advice to AEU Members Applying for Experienced Teacher With Responsibility Classification

Australian Education Union News November 2001


Department of Education Employment and Training 2001(a) Implementation Kit

Victorian Government Schools Agreement 2001

Department of Education Employment and Training 2001(b) Victorian Government Schools Teacher Class Handbook


National Board for Professional Teaching Standards (1996) Adolescence and Young Adulthood/Mathematics Standards for National board Certification


Wallace Wildy & Louden *Selecting Teachers for Advanced Certification: The Western Australian Level 3 Experience* Unicorn May 1999