Chapter 1

Introduction

Background to the Study

I was an army officer for twenty years before becoming a teacher. The annual "PR19" performance report was a critical factor in determining my potential for promotion. On the positive side, my superiors told me what I was doing well, and advised where I could improve, but they also had the power to stop my career dead with unfavourable ratings in critical areas. The expectations of this appraisal were underpinned by a series of generalist and employment specific courses that clearly outlined the requirements of each rank level. My peers and I charted our progress up the promotion chain by way of our reports and courses. My teaching experience was different. There was no clear career or promotion chain. Appraisals focussed on workplace procedural conformity, verifying the prescribed number of professional development days completed, fulfilment of co-curricular responsibility, and acknowledgement of school policies and ethos. Length of service and connections seemed to be the key to advancement.

Within my own teacher appraisal experience, I felt that the real business of teaching and learning came secondary to organisational conformity. The student classroom surveys I had been doing for years at my own instigation, and in pursuit of self-knowledge and better teaching practice, were irrelevant to the appraisal process because there was no procedure for their effective inclusion and analysis. More detailed evaluations for tenure or promotion were more of the same. They included classroom visitations during which my students behaved unnaturally well, probably due to them being disposed positively to me already as well as awareness of another teacher (and a senior one at that) sitting behind them in the classroom. I felt dissatisfied and unfulfilled because I cynically saw my appraisals yielding a well-rehearsed interview technique instead of assisting me to improve my pedagogy. Colleagues experiencing the same appraisal system were apprehensive before, and angry or dismissive after.

Whilst my students seemed to like both my teaching and my subject area, I felt isolated from even my colleague next door. Short of their classroom anecdotes, I actually had almost no idea of how my fellow teachers taught. Meanwhile student talk provided glimpses of their realities and perceptions. They referred to injustices and inflexibility, to teachers setting work and sneaking out down the road for a cigarette, to being called *little bastards*, and to being taught exactly the same lesson two days in a row. After a couple of years one suspects which teacher

is "worth their salt", and who is merely going through the motions. In some cases, teacher failings became common knowledge among old scholars, parents, and colleagues beyond their own school. Rarely did anything seem to be done to help such teachers overcome their shortfalls. More often than not, retirement made the problem go away.

To my mind, if teachers were to improve their pedagogy they needed some mechanism to identify what they were doing right or wrong in the classroom. Anecdotal comments about teacher competence made by student and parents, and indeed by teacher colleagues, left no doubt that both good and inadequate teaching existed in classrooms around Tasmania. The appraisal system in my school did not appear to be addressing the requirement, prompting me to ask:

- *Does there exist in Tasmania any method of evaluating the teaching act in an open, honest, non-threatening, and pedagogically constructive fashion?*

From the first, I felt that the answer to this question would intrinsically involve teacher learning because shortfalls in pedagogy would generate training requirements by either the school or the individual. Conversely, evaluation of teaching is the mechanism by which the effectiveness of training and learning can be measured. In current parlance the term *professional development*, not training, is used about teaching. The problem with this term is that it includes a wide range of activities including formal courses, lectures, workshops, conferences, and meetings. The effectiveness of these was often open to question from my own observations of teacher responses. Like my classroom students, teachers exhibited both enthusiastic on-task, and dismissive off-task, behaviours; some actively engaged in the specific event, whilst others used the time to catch up on marking, or amused themselves passing notes. Attendance clearly did not equate to learning. I was led to enquire:

- *How could the professional development managers claim success, and how could they justify the costs in time and money?*

Aim of the Study

The experience of evaluation systems in two careers, supported by over ten hugely enjoyable and pedagogically rewarding years in the classroom, prompted me to question:

- *Is there any evidence in my own state of Tasmania of evaluation of pedagogy and is this connected to professional development requirements?*

It appears logical that one first should be able to identify individual teacher's strengths and weaknesses before solutions can be sought or developed. Conversely, the effectiveness of any enhancement or remedial measures cannot be gauged without an evaluative mechanism. In other words, how can one identify what areas of a teacher's practice require improvement, development, or reinforcement, without an examination of their current practice and the application of their previous professional learning, and how can the effectiveness of new or remedial learning be verified without some form of validation? If such a nexus does not exist on a broad scale, then the effectiveness of policies for change are open to question as the mechanisms for developing and monitoring changes in teaching practices cannot be monitored. On an individual level, the question became:

- *Is there a process to help me identify both the good and poor aspects of my classroom practice so that I can build on strengths, and identify and rectify personal weaknesses?*

The results of this research had to be more than a simple positive or negative. If there was evidence of evaluative practices, then these could be reviewed to determine if they met teacher needs and the requirements for enhancing pedagogy in accordance with the research. If evaluation was absent, then it was open to question why not even appraisal systems remain in widespread use, particularly after the theoretical discussion of such evaluation in Australia over the past fifteen years. Factors working against evaluation could then be examined and a direction proposed to help teachers find a way to review and enhance their pedagogy in a meaningful fashion.

Context of the Study

This research was conducted at a time of significant educational development in Tasmania. The Department of Education was introducing a major curricular change entitled the Essential Learnings (sic). This sought to establish cross-curricular themes rather than subjects as the foundation for student learning. At the same time, the Teachers Registration Board Tasmania had implemented a registration system for teachers, and was developing standards of practice for teachers in the state along with other stakeholders. Concurrently, the Commonwealth Government was becoming increasingly active in the area of teaching standards and accountability through the establishment of the National Institute for Quality Teaching and School Leadership, and introducing regulations on national benchmark testing as a precondition for federal assistance. Teachers were being required to teach differently and appeared to be facing increasing accountability pressures. It appeared a strong possibility that public discussion would soon shift to teacher quality and competence, and the means of

assessing both. Expanded national benchmark literacy testing and moves to publicly rank school outcomes raised an even more sinister spectre of teachers being held accountable for other social shortfalls as happens in the United States. Objective evidence for criticism of, or even support for, teachers in such an environment did not appear to me to be possible without a mechanism to monitor and rectify pedagogy.

My perspective on this research context was strongly influenced by the work of Linda Darling-Hammond (1997) who has argued that the current structure of schools perpetuates industrial revolution practices that sought to produce obedient skilled factory workers. She asserts that progress has rendered this structure less relevant to a society where initiative and creativity are required to meet the challenges of changing work and lifestyle structures. In response to perceived shortfalls, educational bureaucracies are increasingly prescribing not just what is to be taught, but also how. Consequently, teachers are losing their independence to pursue learning and understanding in their students by being forced to conform to behavioural objectives through curricular direction. Teachers may thus increasingly become *apparatchiks* executing a set program on an increasingly rigid outmoded educational assembly line. If Darling-Hammond is correct, then this research was undertaken at a time when Tasmanian teachers were in danger of losing professional independence. Training and learning would be directed on a collective scale to the implementation of the curricular mandate and not to the enhancement of individual pedagogical needs. Furthermore, there was the imminent danger that the standards being developed by other than teachers in Tasmania could further erode teacher professional independence by sanctioning what was "correct" in the new and more controlled educational environment.

Background to Evaluation

Before I could research Tasmania, I had to answer a simple question:

• *What does evaluation actually mean in the context of teaching?*

Teachers are not army officers, and students are not soldiers. Neither are teachers floor managers, and students employees. According to the research literature, suspicion and fear have characterised the history of teacher evaluation from old style school inspection in the United Kingdom through to more bureaucratic appraisal models directed to social and financial accountability objectives in the United States.

In reviewing this history one very quickly gets the feeling that, more often than not, grand ideas in the pursuit of teacher quality have culminated in unfulfilled promises. In Australia,

terms such as *appraisal* have been more common, generally focussed on superiors identifying and rewarding performance using procedures transferred from the commercial sector. However, widespread implementation of any system appears to have foundered frequently due to teachers retreating to a collective industrial stance, and the absence of any formal and publicly available standards. However, teacher quality and evaluation has not died as an issue, and indeed appeared to be resurfacing. I needed to determine:

- *What are the roles played by teachers and school in evaluation?*
- *What factors work against teacher evaluation and can these be used to determine a list of ideal requirements?*
- *What lessons can be drawn from evaluation systems elsewhere, and what are the research outcomes flowing from them?*

To determine the roles of schools and teachers, I felt it necessary to analyse evaluation within the context of the school as a unique social organisation charged with mass education of the young. Not only does Weber's theory of bureaucracy define the structure of most schools, but associated organisation theory, particularly the works of authors such as Bidwell (1965) and Pusey (1976), reveal that the affective relationship developed in classrooms by teachers as part of the very act of teaching and learning establishes unique institutional relationships. The work of these theorists suggests that the hierarchical school structure operates on a formal bureaucratic level different from the affective relationship between teacher and students in the classroom. Appraisal systems based on commercial or administrative models and instituted by management are incapable of penetrating the teaching act because the administrative and classroom educational purposes operate within different paradigms. Indeed, Pusey suggests that teachers actively work to exclude a perceived hostile formal bureaucratic level of the school from their classroom reality.

The difficulties, primarily beyond Australia, in establishing evaluation are examined in light of this organisational limitation. The work of Darling Hammond (1997) is drawn on to suggest that in some cases evaluation is utilised to mask systemic failures, so the real issue of pedagogical quality is missed. Meanwhile, others point to the superficial and ritualistic evaluations that fail due to the absence of any measure of teacher competence that is authoritative and objective. Schalock (1987), and Fullan and Hargreaves (1991) question the ability of any outsider to quantify teacher effectiveness and productivity due to the many factors that combine uniquely to generate student learning.

- *If the organisational theorists are correct and evaluation has been so elusive and divisive, do the motives behind the evaluation provide some key lessons?*

Simply put, if we cannot identify clearly why evaluation is conducted, then one cannot determine who should do it, or whether it was successful. Is the evaluation aimed at accountability (summative) or improving teaching (formative)? If improved teaching is the aim then evaluation ought to encourage teachers to be open and collegial and to admit shortfalls with the intention of remediation. However, school based appraisal processes do not appear often to have done so. The research of Chadbourne and Ingvarson (1994) in particular is drawn on to identify an alternative direction for teacher evaluation centred on the National Board for Professional Teaching Standards (NBPTS) in the United States. National Board for Professional Teaching Standards certification is based on individuals formally reflecting on their practice, and subjecting themselves to centralised assessment by teaching peers. Other forms of assessment, such as the duties based model of Scriven (1994) and the hybrid form of Danielson (1996) are introduced as alternatives that direct teachers to enhance their performance. Whether summative or formative in structure, these theoretical perspectives direct the argument increasingly towards the teacher as holistic practitioner and away from their assessment as mere organisational functionaries.

The common element of these researchers is an attempt to define teacher standards. Kleinhenz, Ingvarson and Chadbourne (2002) state that standards

> 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 (Kleinhenz et al. 2002, p.5).

The absence of any comprehensive, professionally agreed, and implemented code of standards in Australia suggests that teachers often "feel their way" after graduation from teaching courses. Judgements of their effectiveness or quality also cannot be justified because there is no objective professional measure to which they need to conform, as compared, for example, with nurses, as noted by Ramsey (2000). The requirements of the particular school or particular sector, such as those conducting a group of religion-based private schools, apparently become the de facto measure in the absence of anything else.

However, it is incomplete to attribute responsibility to school management alone; teachers are the other half of the process. This leads to the question:

- *What role should individual teachers play in evaluating and developing their pedagogy?*

The work of Schon (1983) identifies the elements of reflective practice where teachers evaluate their work to rectify shortfalls as a matter of course in their daily practice. Such an active posture makes evaluation and its associated learning the province of the teacher, not the school or teaching system, and establishes true professional consciousness. However, if reflective teaching is not the norm, then the professionalism of teaching is open to question. Chapter 2 identifies the characteristics of a profession and reveals that Ramsey (2000) and others find teaching seriously wanting in a number of critical areas, such as self-regulation against an established code of practice or standards. If this is the case, then it is suggested that teachers are ill equipped to meet the challenges of scrutiny. This also explains why the initiatives for evaluation in Australia have come from others than teachers themselves.

Educational change is considered in Chapter 2, because evaluation includes identification of areas for pedagogical improvement, and thus changes for the better, in individual practice. Furthermore broad educational change such as has occurred in Australia, and more particularly Tasmania, in recent years demands different teaching approaches and thus a change in how teachers operate.

- *If there is no way of knowing what teachers did, are doing now, and should do in the future, then how can we substantiate reports of improved or deteriorated teaching practice?*

Furthermore, if curricular reform is a key political agenda and teachers feel destabilised by it or threatened by any bureaucratic evaluation mechanism, then superficial conformity may be all that eventuates. What appears to be change may end up as no change at all.

Whilst Chapter 2 concludes with a suggested list of design elements for effective evaluation, these may not be exhaustive and the relationship between them is not quantified. Rather, they flow from the analysis of the factors that have in the past undermined or negated the value of evaluation of teachers' work, and could do so in the future. They also provide a theoretical best practice outline against which data collection could be devised and judgements made. As a list of positive elements, they may become guides to avoid the pitfalls of the past, and engender confidence rather than fear about evaluation of teaching.

As a researcher into my own teaching practice as well as that of colleagues, Chapter 2 already provided me with explanations for my personal dissatisfaction with my own school appraisals

and my frustration at not readily identifying a mechanism to help improve my teaching. For the reader of this paper it should convey an understanding of the complexity of evaluation as a key issue for the profession and reinforce that more is required than just devising a system and telling teachers and schools to implement it.

Professional Development in Perspective

The fundamental question addressed in this section of the research is:

- *How effective is professional development, and is it connected to evaluation?*

This chapter commences with a historical overview of the various inquiries in the 1970's and 1980's that recommended the establishment of ongoing teacher education as part of normal practice, and then connected teacher learning with the national training agenda as part of economic reconstruction. There is repeated criticism of professional development management and accountability in these studies. The continuation of problems was highlighted in the landmark report by Ramsey (2000) which suggested that questions of effectiveness remained, despite professional development becoming an established part of the teaching landscape over the intervening twenty years. One of the ongoing issues in these reports is the need for teachers to accept responsibility for professional development beyond that directed or mandated by employers.

Whilst professional development may be less contentious than evaluation, this does not mean that it is easier to get right. Guskey and Sparks (1996) and others reveal that proponents of change often fail to appreciate the differences between schools and teachers and the appropriate developmental needs of each to achieve success. Hargreaves (1997) and Darling-Hammond (1998) are but two of the theorists who identify the need for schools to prepare an appropriate professional development framework that allows change to progress in a truly collegial environment where teachers feel free to experiment. There appears to be common feeling amongst researchers that mandated changes are often delivered through directed professional development sessions that actually disempower teachers in the change process. This raises the question:

- *Can professional development be directed at teachers as a generic whole, or are there identifiable differences between teachers that affect professional development effectiveness?*

This question directed the study towards an examination of stage theory. Theorists such as Oja (1991), Hubermann (1995) and Joyce and McKibben (1982) reveal that teachers not only are different people, but also exist at different stages of their professional lives. Consequently, professional development aimed at a whole staff is in fact a contradiction in terms, as no such homogeneity exists. At an individual level, stage theory is shown to be a valuable tool for meeting each teacher's particular learning needs and structuring appropriate learning environments. Nonetheless, Fessler (1995) points out that stage theory is not immutable, and teacher stages cannot be arbitrarily assigned. Rather the stages should be broad categories that guide trainers to helping each teacher develop effectively. Such direction may seem easy, but, if we refer back to the studies at the end of last century, the responsibility cannot be placed exclusively on school administrators.

- *What is the role of the individual teacher in professional development?*

There is a strong theoretical support base for the teacher to be an active and independent participant in professional development; meeting individual learning needs through reflection or collegial interaction based on real classroom challenges. Ingvarson and Chadbourne (1998a, 1998b, 2002), two highly respected theorists in this area in Australia, repeatedly stress the importance of teachers engaging with their own learning needs. Their advocacy of the National Board for Professional Teaching Standards process is supported indirectly by others, such as Asayesh (1993) who criticises the banking or "shovelling in" approach to professional development, and Viadero (2002), an advocate of action research.

Whatever the level of professional development, systemic, school, or individual…

- *How do we know that professional development activities have been successful?*

There is strong advocacy in the literature for validation of teacher learning. Guskey and Sparks (1996), in particular, support others in the assertion that professional development is unlikely to achieve its aims unless all aspects of the learning and support structures are subjected to scrutiny to determine effectiveness which is then looped back to enhance programs. Again Ramsey (2000) supports this view and places it into an Australian context. Researchers argue that in the absence of validation, there is no basis for determining future learning needs. Attendance at courses or presentations does not of itself represent learning for adults any more than it does for children sitting in the classroom. Understanding and internalisation must be verified and, in turn, prompt remedial action where necessary. The literature suggests that it makes no sense for teachers and schools to validate the learning of students through testing

and culminating performances whilst failing to apply the same basic requirement to their teachers.

The organisational and individual elements of Chapter 3 are drawn together to show that there is broad support in the literature for effective professional development to be not just the responsibility of the school or the teachers alone, but also rather an interdependent process of school support and encouragement alongside teachers' individual motivation and commitment. Furthermore, it is widely stressed that professional development should not be an occasional thing but be embedded in the daily life of the school. Individual and collective needs should be monitored, supported, and coordinated if school and teacher effectiveness are to be encouraged. In this interdependent relationship, if either side gets it wrong the system can break down and generate the criticisms that have been evident in the examined reports into teacher learning.

As with the chapter on evaluation, Chapter 3 extracts what appear to be the key elements of effective professional development from the research and reports examined. This provides both a foundation for the derivation of research questions and a touchstone for research results. Matched up with the evaluation elements from Chapter 2, this chapter balances out the evaluation and learning nexus by completing a summary of best practice from research and theory.

Research Questions

In a 2002 conference paper, Kleinhenz, Ingvarson and Chadbourne (2002) stated that in the thirty years since the abandonment of the school inspection system in Australia 'most Australian teachers have had little or no experience of formal scrutiny of their work…Australia, then, is something of a *tabula rasa* with regard to teacher evaluation' (Kleinhenz et al. 2002, p.3). If such experienced and respected researchers had identified a paucity of evaluation then it was unlikely that this research would identify reflective evaluation practised by professionally minded teachers in Tasmania. Both as teacher and researcher this left me with dilemmas and concerns. My study was partly a personal journey to identify ways of defining the good parts of my teaching and finding ways to rectify my failings. The literature suggested that there was a way ahead; the elements of evaluation could be identified. However, the mechanisms to do so just weren't in place in Australia.

The principal question concerning this researcher was to determine:

What evaluation of pedagogy occurs in Tasmanian high schools, how is this connected to professional development, and how might the relationship between the two be facilitated in order to continually enhance the quality of teacher classroom practice and student learning?

Given that works on appraisal in Australia were published over fifteen years ago and my own experience proved that some appraisal by superiors remained in place, it seemed more productive to transfer the research focus to this form of evaluation. Appraisal results could be used to assess how this type of evaluation generates reflection on pedagogy, defines areas for professional development, and, in turn, verifies the outcomes of such learning. Sensitive and well-conducted appraisal did not have to be totally threatening and might help teachers focus on their pedagogy. If the teachers were not evaluating themselves, but, rather, the schools were conducting some form of evaluation for the purposes of developing pedagogy, there would at least be some level of systemic or school quality control in teaching. The slate may not be completely blank as suggested above.

I wished to examine the perceptions of the purposes of evaluation and the emotional responses of teachers who had been part of an evaluation process. If the aims of evaluation were seen as other than the improvement of teaching, and if negative emotions were preponderant, then the environment for further developments in evaluation in any other form may well be negative. Conversely, if teachers welcomed the experience then the cause of professional reflective practices might have been advanced. Emotional responses would be backed up by examining teachers' actions subsequent to the evaluation, and perceptions of their own level of influence within the school. Even in the case of appraisal by a superior, teachers could assume a level of ownership and demonstrate independent professional rather than subservient behaviours. I also sought to determine evidence of reflective behaviours amongst all teachers regardless of evaluation experiences. High levels of mutual observation of teaching could signify reduced levels of classroom isolation and greater professional interest in sharing actual classroom practices. Frequent participation in regular meetings to discuss curriculum and teaching approaches both within and across learning areas might also indicate an ongoing dialogue about teaching and learning beyond mere anecdotal discussions in the staffroom.

This translated into the following detailed research questions:

• *How often are teachers evaluated by superiors and for what purpose?*

- *What are teacher opinions of, and reactions to, such evaluations and how do they respond afterwards?*
- *Is this form of evaluation seen as affecting teaching practices positively?*
- *Regardless of the level of appraisal being practised, do teachers manifest reflective behaviours such as mutual observation and the discussion of pedagogy within and beyond their learning areas?*
- *Do students see evidence of teacher reflective behaviours and changes in teaching in response to evaluation activities or practices?*

I did not expect that participation rates would reveal much about professional development due to it being an accepted part of modern teaching practice; it is something everyone does. My questions were rather more directed to the purposes of the professional development itself: teacher responses to various professional development events, and teacher feelings of ownership or influence. I sought to identify whether teachers seriously thought about their professional learning needs, discussed them with their supervisors and then made appropriate choices to modify and enhance their pedagogy. More positive reasons for professional development choices would suggest reflection on pedagogy and personal learning needs. Validation of the outcomes in some coherent form either individually or through the school professional development system would demonstrate that there is a cyclical process in place to continually review teachers' professional needs. Of course, evaluations would be part of such validation and would illustrate the relationship between the two. Where evaluation processes were not in place, school based validation of professional development may well achieve a similar outcome.

The detailed research questions examined are:
- *What is the basis for teacher selection of professional development, particularly in the absence of evaluation mechanisms?*
- *How do teachers perceive professional development management and effectiveness?*
- *Are professional development activities validated and related back to classroom teaching and learning?*
- *How do teachers perceive both professional development overall and whole school professional development events in particular?*

The best outcome for the profession in Tasmania from the perspective of the literature review would be for teachers at least to manifest the behaviours in both collegial interaction and

professional development that would illustrate characteristics of individual reflective responsibility, even if formal evaluative processes were not in place. The less attractive option for the future would be for teachers to be passive workers attending but not taking responsibility for the quality of their work and learning, and for them to see themselves as powerless to affect their current situation. Whatever option prevailed, the outcomes would dictate possible solutions to meet the ideal requirements outlined in the initial chapters of this study.

These research questions are not reported on as discrete entities because of the relationship within an ideal evaluation/ professional development/ validation cycle. For example, teacher suspicions of evaluation can be underpinned by the presence or absence of validation of professional development, and inadequate or poorly targeted professional development may fail to yield any classroom outcomes and thus contribute to the absence of any evaluative motivations. Furthermore, this paper considers issues of standards as a focus for evaluation, but it does not seek to examine the contents of particular standards systems in detail. Rather, whilst standards are desirable for profession-wide reflective evaluation, their promulgation itself may not necessarily yield that outcome. Rather, the intimate involvement of teachers and the development of reflective professional identities individually and collectively are seen as critical preconditions.

Research Methodology

This research was initially intended to be a qualitative study of case study schools. However, this presented the researcher with a problem. A qualitative approach might have illustrated issues in the particular schools studied, and even identified commonalities across the case studies, but it would have had to rely totally on other research statements about the absence of evaluative practices generally for any attempt to extend lessons from the qualitative research across Tasmanian high schools and teachers. In the absence of any empirical evidence of evaluation practices in Tasmanian schools, the results of this study would have been open to serious criticism. Furthermore, an exclusive focus on quantitative methodology might have provided further data, but would at the same time have removed the "faces" from the individual teachers, knowledge of whom the researcher felt was essential if the study was to detect evidence of individual evaluative practices or help our understanding of why the profession had not moved towards monitoring and enhancement of its own quality.

Furthermore, because the literature led increasingly towards evaluation as a means of individual professional empowerment for teachers, I was concerned that quantitative research methodology alone would be unfair to teachers by categorising them as a homogeneous group. So often in public discussion, such as by politicians and union representatives, teachers are spoken of as a collective type, despite the arguments of stage theory and the obvious diversity of school environments. I felt that statistics and figures alone would mask the very colleague teachers whom theorists were expecting to assume reflective evaluation practices individually in their classrooms. In this case I felt a resonance with researchers like Viadero (2002) who could rejoice at the work done by identifiable individual fellow classroom teachers working to change pedagogy at an intimate level based on their own understanding and initiative rather than as the result of imposed grand systemic plans. I also felt that it would be hypocritical to look at teachers at a detached quantitative level at a time of major curricular change in Tasmania where colleagues as a whole were being directed how to teach regardless of their individual, school, and classroom circumstances.

I was fortunate to be invited to insert questions pertinent to this topic into a survey being conducted by the Faculty of Education of the University of Tasmania, entitled *School Governance, Instructional Practices and Student Outcomes in Tasmanian Schools*. The initial concept was for this research on evaluation to be incorporated into the governance study, but this did not eventuate due to administrative difficulties that changed that study and resulted in the departure of key researchers, one of them being my supervisor. Nonetheless, this opportunity relatively early in the study allowed me to utilise the quantitative paradigm to establish a macro framework of teacher participation in evaluation and professional development, and reflective and collegial practices in Tasmanian high schools into which the case studies could be positioned. Fortunately, the governance questions also provided additional data on influence levels that allowed me to investigate and illustrate manifestations of organisation theory applied to the research topic. From the perspective of topic, by the end of the quantitative phase of the research I was able to establish not just that evaluation was a minority activity in Tasmanian high schools, but also that teacher practices were not conducive to the existence of reflective professionalism. The qualitative phase could thus advance focussed on the micro factors and issues impacting on the level of individual schools, classrooms, and teachers beneath empirical data that had not existed previously.

A number of schools were approached to participate in the study, but principals seemed reluctant to participate in this research, perhaps because of the strains of a state-wide

curriculum change underway at the time. Nevertheless, the lack of enthusiasm of school administrations was both notable and disappointing. Whilst one must accept the right of Principals to determine access to the organisations they manage, there appeared to be a disinclination to expose the absence of evaluative procedures to scrutiny. In one case, a school that declined participation was in fact at the time considering development of an evaluation model based on the work of Charlotte Danielson (1996). Involvement in this research might well have benefited that process. Even within the four schools that were ultimately studied, the low rate of response in some cases may itself be a finding. Teachers may have been overloaded by directions for change, or they may be apathetic, or reluctant to engage in exactly the sort of reflection that is sought in the literature. It may also be that, in teachers often unrealistically busy working lives, completion of the survey was perceived as yet one more task that was not compulsory and would not deliver obvious direct benefit to the participant. Perhaps it was naive of the researcher to expect that practitioners would leap at the opportunity to participate in a survey about teaching and the evaluation of the quality of teaching.

Nonetheless, four high schools were selected for the qualitative phase of this study. All were of reasonable size and represented the independent, state, and systemic sectors of the educational community. Schools are not further identified, as urban or rural, single sex or co-educational, and so on. Such discrimination in an environment as small as Tasmania threatens the prospect of school identification that would not be the case in larger cities or states in Australia. To ensure the anonymity of the schools and all participants, and to ensure maintenance of ethics guidelines, this researcher felt he had to take care not to just to avoid identification, but also to avoid any categorisations that might exclude, and thus threaten identification by default. It should also be noted that issues of school structure and student demographic are not seen as impacting on the issues of evaluation and professional development directly in this research study, although local conditions would need to be taken into consideration in the application of evaluations as is mentioned in the literature review.

The case studies focussed on analysis of school evaluation and professional development procedural documentation, teacher survey responses, and student focus group discussions. In the case of the teacher surveys, these were not sent to all teachers but delivered to the schools and left for individuals to collect after a verbal briefing about the nature of the study and the ethical requirements for complete anonymity of both individual responses and school identity. Through this approach the researcher sought to establish some level of collegial empathy towards the study being undertaken by a practising colleague rather than making the surveys

appear a purely academic or organisational activity from beyond the school environment. Whilst the response rate in each school could indicate the level of interest in a matter that ought to be a significant professional attention, it might just as well reflect staff morale or governance issues beyond the scope of this research. Rather than focus on the missing element, the study concentrates on the responses themselves as providing evidence of attitudes, habits, prejudices, fears, and perceptions of exactly the individuals who would need to be empowered to pursue a course of reflective professional evaluation and consequent effective and validated professional development. If there would be some difficulty convincing and empowering those who responded, then how much more effort might it take to include those teachers who may not have been sufficiently motivated to participate in an activity directed towards their professional enhancement?

The focus group discussions were with student representative council members, thus providing a cross-grade sample from the schools. The researcher facilitated the discussions during school time and in the absence of any other member of staff. To comply with ethical requirements, all participants had parental permission, were told that the discussion would be taped and offered the opportunity to withdraw. Complete anonymity was assured and students were very supportive in maintaining non-specific identifiers; in only one case was there a need to retape to erase a name. After some initial reluctance, students warmed to the discussion. The list of questions developed for the student discussion were not adhered to slavishly but amplified or explained as required, or omitted if answered in previous discussion. The aim was to ensure that discussion flowed and that students were engaged in the topic, and to avoid the possibility of a "boring" judgement that might have hampered the extent of responses. This approach worked well, with student engagement being lively and enthusiastic. Although some individuals stood out as having strong views and verbal capabilities, the researcher sought to be as inclusive as possible with statements or opinions often redirected to others or the group as a whole for verification. The analysis of the discussion was reduced to key headings outlined in Chapter 5. It should be noted that the researcher had to assume student honesty in their responses at the time of the discussion and rely on the development of empathy with them as an objective researcher rather than another teacher. Where parental responses are referred to these must be accepted as second-hand accounts. However, given that the students in each school had no previous indication of the nature or content of the focus group discussions, they verified the anonymity of their statements in a couple of cases before responding, and that discussion tended to flow very quickly in the time available, there would be little reason for students to have answered in any other than an honest fashion. Furthermore, commonalities in

responses across schools in a number of areas supported the validity of student perceptions on the research topic.

Each school is allocated a chapter in the study so that it can be analysed as a unique entity. Not only do the results provide an insight into the particular research aspects, but they also indirectly demonstrate whether each school is focussed on the enhancement of teacher quality or is directed towards other agendas. In the case of School 4, which had an appraisal system in place and revised the process during the course of the study, the potential outcomes of the revised documentation are particularly interesting in the light of the teacher responses to the surveys. Unfortunately this school did not provide access to students for focus group discussions.

During the course of this study, in 2004, a group formed primarily of bureaucrats and school representatives and established under the Australian Government Quality Teacher Programme, started to devise standards for teachers in Tasmania. Whilst a meeting of teachers was held and about fifteen apparently exceptional teachers were recruited to "unpack" the practices that made them so, the actual involvement of Tasmanian teachers as a whole in the developmental process was negligible. There was no documentary or bibliographic evidence of a rigorous examination of research on either teacher standards or evaluation practices in the literature circulated by this body, and the methodology was not defined in detail: no web site was established, nor information pack circulated. The progress of the work of this group was followed throughout the next two years to determine its impact in the light of both the readings and my research results. Such a development was fortuitous given increasing interest in teacher quality and the potential for the Teachers Registration Board Tasmania to assume a quality assurance function in its future re-registration procedures.

The Tasmanian standards examination is deliberately located at Chapter 10 to reflect the chronological occurrence of this development within the progress of my research. Whilst over the preceding years this study conducted a detailed and referenced examination of the literature, considered the mutual dependence between evaluation and professional development, and sought to establish empirical evidence of the situation in Tasmanian high schools and the attitudes of teachers, the Tasmanian standards process provided no public evidence of similar rigour over the two years from its inception through to the issue of its initial standards publication, despite the apparently impressive composition of the group. Anecdotal evidence exists that academic review was invited only once the decisions had been

made and the brochure was in its final draft stage, suggesting a desire to improve the participation list rather than extensively involve the educational academic community, teachers included. The circulation of the *Tasmanian Professional Teaching Standards Framework* (undated) document in late 2006 also coincided with the finalisation of the conclusions of this paper. Quite clearly that document and the conclusions of this research are headed in different directions and the criticism of the failure of that process to establish a more solid intellectual foundation is intended. If teachers were already concerned and suspicious of evaluation as illustrated in this research, then the Tasmanian standards implementation, conducted in response to national policy and without examination of prevailing responses to evaluation and the realities of professional development participation, has the potential to alienate teachers to the extent that the Tasmanian environment would no longer be blank, as suggested by Kleinhenz, Ingvarson, and Chadbourne (2002), but become demonstrably hostile, to the detriment of the improvement of classroom pedagogy.

Research Conclusions

Whilst each section of the quantitative and qualitative research contains separate findings, these are consolidated within Chapter 11, where the evaluation environment, professional development practices, and the standards developments are combined to illustrate whether indeed the nexus between evaluation of teaching and the consequent professional development requirements exist in Tasmanian schools. However, the study does not stop at the point of judgement. Rather it moves through these results to suggest a new professional environment where teachers can assume control of their pedagogy which appears to be in the process of being supplanted by increasing bureaucratic direction. In this respect, this study is influenced by the work of Linda Darling-Hammond (1997) who argues that the modern age requires a re-conceptualisation of schooling out of the current hierarchical industrial revolution model where the agenda was to train loyal factory workers rather than thinkers. This paper is not so arrogant as to suggest a new way of schooling. However, it does submit a proposition that teachers can reinvent themselves into the type of reflective professionals advocated by so much eminent research. In doing so, teachers would evaluate their own teaching and meet their professional development needs by working with colleague teachers, rather than so-called experts.

Whilst the establishment of professional standards is clearly very important, it may be more useful for teachers initially to utilise an extant model as a focus for development of professional evaluative practices rather than waiting for a consensus on standards. Such an

adopted standards structure would undoubtedly not be perfect and require adjustment to satisfy local conditions. Nonetheless, the clear professional focus of any associated certification process and experience in its utilisation and modification could result in the almost immediate empowerment of teachers. The significance of this for the Tasmanian context is that this Australian state is sufficiently small for many teachers to know each other within the one regional education system. State wide moderation of senior secondary teachers already occurs, providing perhaps a laboratory for teachers together to chart a course towards the enhancement of their pedagogy as independent professionals, evaluating their classroom practice as a matter of course and directing their learning needs to meet their own truly professional requirements.