

Policy Briefs

Strengthening the profession?

A comparison of recent reforms in the UK and the USA¹

Educational policy makers in many countries recognise the need to focus their policies more directly on factors affecting the quality of teachers. Common to these policies are attempts to reform teachers' pay systems and career paths to place greater value on teachers' work and give stronger incentives for professional development. Investing in effective modes of on going professional learning is regarded increasingly as one of the most effective means of improving student learning outcomes.

This article examines two approaches to reforming the teaching profession, one from the UK, the other in the USA. In the case of the UK, the focus will be on a comprehensive government 'performance management' system for the teaching profession in England and Wales, introduced in 2000. In the USA, the focus will be on 'professional certification'; an emerging system for giving recognition to 'accomplished' teachers provided by an independent professional body, the National Board for Professional Teaching Standards. Both reforms aim to improve the attractiveness

of teaching as a career and to provide teachers with greater rewards for evidence of professional development.

These two approaches to assessing teacher performance for career progression will be compared on a number of criteria, particularly their capacity to engage all teachers in effective forms of professional development and assist them to reach their full potential, the fundamental aspiration of any performance management system. Each depends, of course, on credible methods for assessing teacher performance. One, it will be argued, looks backwards and has little chance of achieving its aim; the other points to a possible future and has the potential to radically change the way we think about professional development and methods for assessing teacher performance. It recognises the power of professional forms of recognition and demonstrates the commitment teachers are prepared to give to the task of developing their own standards and methods for assessing performance.

Dr Lawrence Ingvarson
Head of Teaching and Learning Division

This paper can be downloaded at the ACER website www.acer.edu.au

¹An earlier version of this paper was published in 2001 by the Australian College of Educators as Paper No. 4 in its Quality Teaching Series.

There are valuable lessons for Australians in comparing these reforms, even though we have been attempting our own since the late 1980s. The 1998 Senate Report on the status of teaching in Australia, *A Class Act*, is a recent example, together with Commonwealth Government's initiative 'Teachers for the 21st Century', announced in August 2000. Others include the proposed Institute of Teaching in Victoria and the Ramsey Report in NSW (Ramsey, 2000). All point to the need for better mechanisms for providing good teachers with recognition. They indicate that we are well aware of the need to strengthen the profession and place greater value on good teaching, but perhaps still unsure how best to go about it.

The UK and US approaches to reform will be compared on a number of dimensions. Specifically, the degree to which they involve teachers in the tasks that are central to strengthening the profession:

- ◆ Developing standards that describe what accomplished teachers know and do, and
- ◆ Developing and operating a valid system for assessing whether teachers have attained those standards.

This paper argues that commitment of the profession to reforms such as these will depend on the creation of independent institution through which the profession can exercise a major responsibility for these tasks. A central purpose of such an institution would be to 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.

Recent teaching reforms in the UK and the USA

Many countries are shifting their reform efforts to focus on teacher policy and strategies that relate more directly to the promotion of quality teaching and learning. In 1998 the Secretary of State for Education in the UK released a Green Paper late called *Teachers: Meeting the Challenge of Change* (Secretary of State for Education and Employment, 1998). It foreshadowed a wide range of reforms

designed to 'accord the teaching profession an entirely new status'. The Green Paper sets out the government's vision for 'a modernised teaching profession'.

Teachers deserve rewards for good performance, better career prospects, (and) opportunities to keep their skills and subject matter knowledge up to date . . . We propose two pay ranges for classroom teachers, with a performance threshold giving access to a new, higher range for high performing teachers with a track record of consistently strong performance. Crossing this threshold would be a significant career step. Above the threshold, teachers would continue to focus on classroom teaching but would be expected to take responsibility for making a wider contribution to raising standards in their school. (p. 32)

The Green Paper introduces a new career structure to recruit, retain and reward good teachers in England and Wales. Teachers at the top of the incremental salary scale who demonstrate high and sustained levels of achievement and commitment will be able to cross the 'threshold' (top of the incremental salary scale) to gain a 10% pay increase and access to three further pay steps on the new extended pay scale. The first group of teachers was invited to apply for the threshold assessment in 2000, eighteen months after the Green Paper was published. Over 200,000 teachers did so. Headteachers were given the task of assessing applicants against a brief draft set of standards that teachers received three months before their application forms were due. (A more extensive list of standards was completed later – see below)

In similar vein, the USA has had a series of major national reports since the early 1980s expressing concern, among other things, about the capacity of the profession to compete with other occupations for quality graduates and to retain good teachers in the classroom (National Commission on Excellence in Education, 1982; Carnegie Forum on Education and the Economy, 1986; National Commission on Teaching and America's Future, 1996). One of the most significant recommendations of the Carnegie Forum was that a National Board for Professional Teaching Standards (NBPTS) be established.

The NBPTS was founded in 1987 with a broad base of support from state governors, teacher unions and school board leaders, administrators, college and university officials, business executives, foundations and concerned citizens. It is a non-profit, non-partisan organisation governed by a 63-member board of directors, the majority of whom are teachers (NBPTS, 1989). The National Board core functions are to establish high and rigorous standards for what accomplished teachers should know and be able to do, and to develop and operate a national voluntary system to assess and certify teachers who meet these standards.²

The long-term aim of the NBPTS is to build a national certification system operated by teachers for teachers in all schools. Through its certification process, the Board aims to provide an independent and credible assessment of teacher performance against profession-determined standards. The hope of the Board is that, as its certification gains credibility, education authorities will increasingly give recognition and incentives for teachers who become National Board Certified Teachers. (Most states now give some form of recognition, which may range from a salary increase to accepting certification as a means by which teachers may meet state requirements for license renewal purposes.)

Although the aims of the UK threshold reforms and the National Board are similar, the methods for implementing them are strikingly different. The UK and the US approaches will be compared on the following criteria:

- ◆ Purposes for teacher evaluation
- ◆ Methods used for developing *standards* for accomplished practice;
- ◆ Methods used for *assessing* whether teachers have attained those standards; and
- ◆ The extent to which each is likely to lead to build a more effective *infrastructure* for *professional learning*, related to improved student learning outcomes.

Purposes for teacher evaluation: Performance management and professional certification

Two purposes for teacher evaluation should be distinguished. The first is to safeguard the educational interests and welfare of students and ensure that their teachers are able to fulfil their contractual duties. This purpose is based on the undeniable requirement that teachers be publicly accountable. Standards for this purpose are mainly generic and common to all teachers. Responsibility for this kind of evaluation rests with the state and is exercised through bureaucrats and school managers. 'Performance management' is commonly used to describe this type of system. Legitimate authority for defining teachers' work is seen as resting properly with democratically elected governments and responsibility for implementation is delegated to school management.

The second purpose emphasises the complementary need to ensure that teachers continually review and improve their practices in the light of contemporary research and profession-defined standards. Responsibility for developing standards for high quality practice and 'certifying' those that reach them is usually delegated to professional bodies in most professions. With professional certification, legitimate authority stems from professional expertise and values, and responsibility for defining teachers' work and 'good practice' is delegated to professional bodies.

These two forms of authority are complementary. Performance management systems emphasise that teachers owe a duty of loyalty to the policy goals of democratically elected governments. The professional model emphasises that teachers also owe a duty of loyalty to clients and professional standards. Can these two purposes best be met within the one system of teacher evaluation, or do they point to the need for two separate systems (Ingvarson, 1994)?

²The NBPTS is a private organisation – government did not set it up – but teachers form two thirds of its membership and dominate its policy making and its operations. Approximately 50% of its funding over the past thirteen years has come from untied Federal Government grants. The rest has come from a wide variety of bodies such as the Carnegie Foundation, and the National Science Foundation. In 1997 President Clinton pledged to support the NBPTS in its goal of certifying 105,000 teachers by 2005 with grants amounting to nearly \$20m per year.

Performance management versus professional certification

The UK government has established a performance management system that combines these two purposes. Under the Green Paper reforms, annual appraisals and threshold assessments are two parts of the same performance management system. There is no clear distinction between them. Responsibility for developing standards and teacher evaluation policy is in the hands of government. Implementation is in the hands of school managers. There is no role here for an independent professional body. Annual appraisals of teachers are conducted by senior school managers and assessments of teachers at the performance threshold are conducted by headteachers (monitored by an external assessor drawn from a pool of 'nationally trained experts').

In contrast, the NBPTS operates independently of employing authorities. Professional standards, by definition, are not regarded as employer or school specific. Procedures for appraisal or performance management are regarded as the responsibility of school managers within local school districts, or part of state requirements for renewal of state licences to teach. This responsibility is quite distinct from the certification function that the NBPTS provides.

The NBPTS, not employers, awards certification as an independent professional body. However, employing authorities, unions and other stakeholders are active partners in the development and operation of the Board's certification system. Board certification provides an endorsement that a teacher has attained performance standards set by the profession. Certification belongs to the individual. It is portable – it is not tied to a particular position, or job, or role within any specific organisation, nor is it an academic qualification, or an accumulation of academic credits. Specially trained peers from outside their school system assess the evidence that teachers provide for Board certification, for example, portfolio entries, although workplace colleagues and managers may be called upon to validate evidence provided by candidates.

Assessors must be currently teaching in the same field as the candidate. They can not assess teachers known personally to them.

Though the two purposes for teacher evaluation outlined above overlap, they help to distinguish two spheres of responsibility – one quite properly that of government, the other perhaps best delegated to the profession. While the first purpose reflects teacher accountability to management and to the public, the second points to the need for the development of strong normative structures for accomplished teaching with which teachers identify strongly.

Methods for developing standards

Procedural validity: who develops the standards and how?

Fair assessment of teacher performance relies upon valid and clear standards. Standards aim to define teacher's work and what is to be assessed. There are two aspects to validity here. The first concerns the process by which the standards are defined and who is involved, or *procedural* validity. The second refers to whether teachers who meet the standards are more likely to provide higher quality learning opportunities to learn than are those who do not. In measurement terms, standards aim to define the domain of what is to be assessed. In other words they should also have *content* validity. Content validity also relates to the match between the assessment tasks and the construct of interest – in the case of this chapter, accomplished teaching at the threshold in the UK, or standards for National Board certification in the USA.

Several texts on educational measurement describe these validity procedures, including Messick (1992) who introduces the concept of consequential validity. Valid standards and assessments should, for example, promote professional development. Standards should provide teachers with a vision of highly accomplished practice. Assessment tasks themselves should be a vehicle for learning.

Hattie (forthcoming) draws on the work of the National Commission of Certifying Agencies in the USA to identify criteria for assessing *procedural validity* when specifying a set of standards for any profession. The process by

which a set of standards is developed will be a critical issue, not only for the validity of the subsequent operationalisation of assessment procedures, but also for their legal defensibility. In summary, these criteria are designed to ensure:

- ◆ the integrity and independence of the body responsible for developing the standards;
- ◆ that the standards developing body is composed primarily of those who are already highly accomplished practitioners;
- ◆ that the diversity of perspectives in the profession is represented;
- ◆ that the process of defining the standards is developed on a sound scientific basis and that the process of developing the standards be formally documented; and
- ◆ that a wide sampling of agreement is sought for the standards from the major professional groups regarding the appropriateness and level of standards.

Procedural validity and the development standards in the UK and the USA

The procedures used to develop standards for threshold assessments in the UK and for National Board certification can be compared using these criteria. In the case of the threshold assessment used in 2000, it is difficult to find evidence that those involved gave any attention to either the procedural or content validity of their standards before implementation.

Standards for the threshold assessment in the UK appeared early in 1999, shortly after the Green Paper. A final one-page version of the threshold standards appeared in March 2000. According to a DfEE official, 'they were developed by DfEE officials in conjunction with education consultants and other interested parties. In finalising the standards, the Department also took account of comments made by members of the teaching profession who were consulted about the draft threshold standards published in February 1999'³.

In 1999 the Department for Education and Employment in England decided to fund a major project to develop a framework describing effective teaching – teaching standards in effect – with a mind to using the findings to support the Green Paper reforms at a later date. Instead of enlisting existing national teacher/subject associations it commissioned a private consulting firm, Hay/McBer, to undertake the task, missing perhaps a major opportunity to promote the development of the profession. The UK approach to developing standards cannot be rated highly in terms of procedural validity.

The NBPTS provides detailed evidence of the procedures and the research it undertakes to ensure the validity of each set of standards (Moss, forthcoming). When the Board decides to develop standards for a particular certification field such as high school science, or generalist primary teaching, it takes the following steps. It appoints a standards committee, the majority of whom must be distinguished teachers currently practising in that field. Teachers usually chair the committees. Other committee members may include experts in child development, teacher education and the relevant academic discipline. Committee appointments last three years, the period it usually takes to complete the cycle of developing the standards.

The Committees are charged to develop standards that identify what accomplished teachers in that field know and do. They do this within a framework of core principles for accomplished teaching developed by the Board. While these principles are 'generic', the challenging task for members of the Board's standards committees is to work through what they mean for particular curriculum areas and levels of schooling.

The NBPTS reviews the draft standards before a public review process. It works closely with professional teaching associations to establish advanced standards of knowledge and practice in their respective fields. Many teachers from these organisations serve as board members and standards committee members.

³Personal communication

Content validity: Reflecting the diversity and complexity of teachers' knowledge and skill

Hay/McBer produced a major report in June 2000, *A Model of Teacher Effectiveness*, based on its findings (DfEE, 2000b). The approach used by Hay/McBer is circa 1970s process-product research on teacher effectiveness, which seeks to find correlates between generic teacher behaviours and student outcomes. There is little to reflect the major paradigm shift that has taken place over the past twenty years in research on teaching, based on extensive evidence that what expert teachers know and do is in significant respects subject and level specific (Berliner, 1992; Shulman, 1987; Brophy, 1991).

Consequently, 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.

In contrast, the NBPTS is developing advanced standards in more than 30 *certification fields*, working with teacher associations, educational researchers and unions. The National Board's certification fields are structured around student developmental levels (For example, early childhood, primary, early adolescence, adolescence and young adulthood) as well as by subject area. Standards have already been developed in 21 fields and a typical set of National Board standards is thirty to forty pages long.

Before the NBPTS was under way, many teacher associations in the USA had already demonstrated their ability to write convincing and challenging teaching standards. The first, and one of the best, was

produced by the National Council for the Teaching of Mathematics in the USA (NCTM, 1991), based on a clear vision of quality learning in mathematics, and what teachers need to know and be able to do to implement that vision. National subject associations in Australia for mathematics, science and English teachers are doing the same (Ingvarson, 1999a). Needless to say, teachers develop a powerful sense of ownership for standards developed in this way.

The National Board has built on this work of teacher associations and unions. The Board's standards for each certification field are embedded in particular subject areas and teaching levels, consistent with recent research on the domain-specific nature of expertise. They might describe, for example, what accomplished teachers of science know and do to engage their students in scientific inquiry, or what English teachers know about learning to write. They acknowledge that what an accomplished teacher of art knows about how help students progress in that field is different from what an accomplished early primary teacher knows about how to promote development in numeracy, and NBPTS standards aim to reflect that knowledge (Brophy, 1991).

The NBPTS case is an illustration of how the development of standards can provide an excellent opportunity to place greater responsibility for the professional development agenda in the hands of teachers, with beneficial consequences for all. That opportunity to strengthen the profession has yet to be taken up by policy makers responsible for the Green Paper reforms in the UK.

How are teachers assessed against the standards?

There is a striking contrast between the methods developed for the threshold assessment in England and Wales and those developed by the National Board for certification. Once again there are fundamental differences in the role that the teachers play at all stages of the assessment process and the degree of attention given to research issues before implementation.

Assessment at the threshold in the UK

In the case of the threshold assessment in England, teachers and their organisations played no role in the development of the methods for assessing teacher performance. Nor were they involved in carrying out the assessments. The assessments were implemented before a trial period and without any research into their validity or feasibility.

The method used for assessing teacher performance at the threshold is almost breathtakingly crude. Teachers are asked to complete a form that they submit to their headteacher. The form contains pages with boxes on each page for teachers to fill in for each standard. In relation to the threshold standard above, for example, it is suggested that teachers summarise evidence in a box less than two inches deep, that they consistently and effectively:

- ◆ Use their knowledge of pupil's learning needs to plan lessons and sequences of lessons, to target individuals and groups effectively and to ensure good year-on-year progression;
- ◆ Communicate learning objectives clearly to pupils;
- ◆ Make effective use of homework and other opportunities for learning outside the classroom.

As sources of evidence, teachers are told they can 'use feedback from classroom observation, evaluation of performance through a school's monitoring system or from OfSTED inspections. Other sources could be teaching materials, record books, pupil's work and marking of homework.' How headteachers are to ensure some reliability in their own assessments, or consistency with other heads, in interpreting this type of unstructured evidence is not clear. There is also a risk of unfairness here as teachers will make different interpretations of how much and what type of evidence to present.

Although there is a long history of research on teacher evaluation (Millman & Darling-Hammond, 1990), there is little evidence that this was drawn on. There is also considerable

research on the need for training to control bias (Scriven, 1994), and problems with cronyism in the micro-politics of school life (Blase, 1990) that might have been relevant. In the opinion of many headteachers, the one brief day of training they received was unlikely to equip them sufficiently to carry out these evaluations as well as they would wish (TES, March 31, 2000; TES June 9, 2000).

Similarly, there is little evidence of research into fundamental issues such as the reliability of the methods used to assess or score teacher performance or the comparability of assessments from headteacher to headteacher. It is one thing to write a list of standards, which include statements such as 'Teachers should demonstrate that they consistently and effectively plan lessons and sequences of lessons to meet pupil's individual learning needs' (Threshold Standard 2) and quite another to establish clear guidelines, or rubrics, for assessors as to what counts as meeting that standard. Little attention seems to have been given to *setting* standards – how good is good enough to meet the standard – and ensuring there is similar interpretation of evidence from school to school.

One headteacher wrote to the TES (June 30, 2000) describing the heavy workload that threshold assessments caused, but concluded 'My thanks will be seeing my good teachers rewarded – like most people in the profession, they are a great bunch!' This attitude may be commendable, but there is plenty of research evidence to suggest that it is unwise to expect principals to make discriminating evaluations of the performance of their own staff when they have to live with the consequences (Bridges, 1990).

My prediction is that the threshold pass rate in the UK will be around 90%, as it turned out to be for the Advanced Skills Teacher in Australia some years ago. We know what the consequence of that was. The concept of a career path based on evidence of professional development was debased. The assessment had little credibility, the position had little status and the salary progression quickly became automatic, which may have been the intention in the first place.

Assessment for National Board Certification

Teachers applying for Board certification undertake two forms of assessment, a portfolio and an assessment centre.⁴

The portfolio

For the portfolio, teachers prepare six 'entries' of three types, each providing evidence relevant to several standards. Two entries are based primarily on *student work samples*, two are based on videotape clips of class discussion and the last two are based on documentation of professional accomplishments outside the classroom. Each entry is like a whole 'piece' of a teacher's work, and each entry represents a different facet of that work (eg evidence of ability to sustain a quality whole class discussion; evidence of ability to monitor student progress.)

Some titles give the some idea of the wholistic nature of the portfolio entries: *Teaching a Major Idea Over Time*; *Making Real World Connections*; *Probing Student Understanding*. Each entry takes teachers about 20–30 hours to prepare and is about 12 pages long. Teachers are provided with a clear structure for the portfolio and guidelines for preparing their entry. Entries also provide teachers with guidelines on how their entry will be scored.

As an example, one of the portfolio entries for high school English teachers, *Analysing Student Writing*, asks them to analyse a piece of writing from each of three students, in the context of the student as a developing writer and their approach to teaching writing in the class. Teachers are advised to collect samples of work over time from a larger number of students and to select later those they will use in the actual portfolio entry. Candidates must include the prompt or assignment that occasioned the writing, all the drafts, other student work that shows the writing process that the student used, peer or teacher conference notes, and any written feedback the teacher provided.

The next two entries are based primarily on *videotape clips*. Candidates are advised to make several videotapes, and from different classes, from which they can select one later for their entry, and to involve other teachers and students in making and analysing them. One entry asks teachers to demonstrate the strategies they used for *small group teaching* (the other focuses on whole class discussion). They are asked to present a 20-minute videotape (uninterrupted and unedited) in which students work purposefully in small groups and a ten page written commentary about their teaching as seen on that tape. The focus of this entry is on the development of students' ability to engage with the teacher and with each other in meaningful discourse as they work in small groups on an important topic in language arts and on their integration of teaching. Teachers are provided with detailed guidelines about the questions they should address in their written commentary, similar to those above for the student writing entry.

The final two entries ask teachers to document their accomplishments in areas outside the classroom, such as leadership in curriculum and professional development within their school, and wider contributions to educational policy and practices through work in professional organisations and other settings. Teachers are asked to provide evidence of the significance and impact of their accomplishments, not just descriptions, and letters of verification must come from someone who is personally knowledgeable about the accomplishments they are describing.⁵

The Assessment Centre

One of the National Board's guiding principles from its inception has been that highly accomplished teachers should have a rich understanding of the subjects they teach and appreciate how knowledge in their subject is created, organised, linked to other disciplines and applied in real world settings. Evidence of this type of subject specific pedagogical knowledge can be achieved in part through

⁴More details of these assessment methods can be found in Ingvarson, 1999b).

⁵As part of the Australian Science Teachers Association Standards Project, teams of science teachers in four states evaluated the relevance and suitability of NBPTS portfolio entries for Australian conditions during 2000. Their reports strongly endorse the validity of this approach to assessing their teaching and its value for professional development.

portfolio entries. However, the Board believes supplementary methods are needed, so teachers attend an 'assessment centre' for one full day on any day that suits them over a six-week period near the end of the school year. The day usually consists of four 90-minute sessions. Candidates are asked to respond to four specific prompts, most of which assess the currency of their knowledge of relevant subject matter and research on teaching that subject matter. Some are based on stimulus materials sent to candidates well in advance of the assessment centre date. Here is one example.

Teaching and Writing: Candidates are sent three professional articles about teaching and learning writing. They are told to read the articles in preparation for an exercise in which they will be asked to use information from the article as a basis for talking about teaching writing. At the assessment centre, a written scenario is presented. Candidates are asked to construct an argument on a topic presented by the scenario, using the stimulus articles as a support.

It is expected that teachers will take most of a school year to complete their portfolios. The 'pass rate' is around forty percent. Teachers who miss out can reapply and 'bank' portfolio entries on which they did well enough to meet the Board's standards.⁶

Assessing teacher performance for NBPTS certification

The National Board has taken many years to develop its assessment methods. Teams of teachers are involved at all stages of the development and trialing process. Many of the most highly regarded academics in the field of educational measurement have conducted research on the validity, reliability and generalisability of the assessment methods. Pearlman, (forthcoming) documents this work thoroughly. It is important to note that many promising ideas for assessment tasks have been tried and found wanting. The portfolio entries may sound simple, but they

have been through a lengthy period of development and testing by teachers before being considered ready for formal use. Teachers must see them as feasible and 'authentic' teaching tasks – tasks they regard as a normal part of their work. Unlike the UK Threshold assessment, each task is designed as a coherent piece of teaching and to provide evidence of multiple standards. The tasks must clearly link back to the standards and assessors must be able to score the entries reliably.

These assessments call for assessors who have deep knowledge of the relevant teaching field. School managers may have been successful teachers, but they are unlikely to have the expertise to assess across all the certification fields, not to mention the many potential sources of bias that may come in when assessing their own staff.

Last year (2000) nearly 10,000 teachers applied for Board certification. Hundreds of teachers are assembled each summer at several centres across the USA and trained to carry out the assessments. Most have never been assessors before, but after four days training they reach high levels of reliability in their assessments (Gitomer, forthcoming). They can not assess teachers known to them already. Teachers are trained to assess only one portfolio entry or one assessment centre exercise. Assessments are constantly monitored to check on reliability. Two teachers working independently assess each entry and exercise. This means that, across the six entries and four exercises, twenty teachers may be involved in the assessment of one teachers' evidence. (After two to three weeks examining the work of many candidates it is understandable that these teacher-assessors come to value this as one of the most valuable professional development experiences they have ever had.)⁷

⁶These requirements for NBPTS certification will change from 2002. Candidates will be required to complete four entries instead of six and the assessment centre is reduced to half a day and six 30 minute exercises. Board research indicates that this will not diminish the reliability of NBPTS certification decisions.

⁷A party of ten Australian teachers observed the NBPTS assessor training and assessment process last year at one of the sites. Their reports following the visit indicate that they were all highly impressed with the rigour of the process.

How rigorous are the two systems for assessing teacher performance?

No research has been reported as yet for the threshold assessment in England. It was implemented without any research on its psychometric qualities. However, one of the cardinal rules for reliability in assessment is that the process should involve multiple assessors and multiple forms of evidence. With only the headteacher as the assessor, and the type of application form that teachers currently complete, it is unlikely that the threshold assessment in England will meet this criterion. (External assessors are used, but their main role is only to verify, through sampling, that a headteacher has applied the performance threshold standards correctly (DfEE, 2000).

The National Board actively seeks rigorous evaluations of its certification system. It has conducted many research studies on the reliability of its assessments. Annual 'Technical Analysis Reports' from independent expert bodies are produced for the Board containing audits on the psychometric characteristics of the Board's assessments. Reliability, of course, does not guarantee validity – that the Board's certification process is effective at identifying teachers who actually are highly accomplished. This is not a simple problem to research. However, the NBPTS has commissioned several research studies to investigate the validity of its assessments. One was reported recently (Bond, Smith, Baker & Hattie, 2000) and can be accessed at the National Board's website (www.nbpts.org). It provides independent confirmation that successful candidates differed significantly from those who did not gain certification on a range of measures of the quality of their classroom teaching practices.

The National Board recently invited universities and research agencies to submit ideas for research on the the Board's work. Silver (et al 2002) reported a study based on an intensive, independent reanalysis of portfolios submitted by mathematics teachers, which indicated that candidates who gained certification made significantly higher cognitive demands on their students than teachers who did not. The Board commissioned

Bill Sanders, creator of the Tennessee Value Added Assessment System, to undertake a major study of the impact of its certification on student learning outcomes.

No research of this quality or rigour has been conducted for the Threshold assessment in England. However, unlike the NBPTS approach, one of the Threshold standards asks applicants are to provide evidence that, as a result of their teaching, 'their pupils achieve well relative to the pupils' prior achievement, making progress as good or better than similar pupils nationally.' Teachers are told that this evidence 'should be shown in marks or grades in any relevant national test or examinations, or school based assessment for pupils where national tests are not taken.' This value-added approach to assessing teacher performance sounds reasonable, but it places a great deal of faith in the validity of the national tests, as measures of what the teachers are teaching, and without mechanisms for isolating teacher effects from a range of other factors affecting student test scores.

It remains to be seen whether research in the UK will provide evidence to confirm interpretations of teacher effectiveness based on these test scores. But this does point to another difference between the Board and the threshold assessments. The Board's portfolio entries ask teachers to provide evidence that they can provide high quality *conditions* for learning, consistent with current research and professional judgment and with professional models of accountability, as set out in the standards. Portfolios provide direct evidence of student work and classroom activity – the 'outcomes' of the conditions for learning established by the teacher – whereas the threshold assessment attempts to hold teachers accountable for standardised 'outputs'.

More research will be needed before we have a better understanding of which approach is the more valid basis for teacher assessment, which is fairer and which has more consequential validity. In this context, consequential validity refers to the relative capacity of the performance management system in the UK and the certification process in the US to have wider benefits for professional development and student learning.

Linking the assessment process to professional development and improved student learning

As mentioned earlier, both the teacher reforms in England and Board certification in the USA aim to provide stronger incentives and recognition for professional development. Each reform can, therefore, be examined in terms of its potential to engage most teachers in more effective methods of professional development, as well as effects on self-esteem and professional relationships in the schools. Each reform can also be analysed in terms of its capacity to place greater responsibility for the professional development system in the hands of teachers and their organisations.

National Board assessments are designed to be a vehicle for learning. There is considerable evidence that teachers who have been through the NBPTS certification system regard it as one of the most powerful professional development experiences they have ever had (Tracz et al. 1995; Heller, 2002). The NBPTS conducted a survey in 2001 of 10,000 teachers who had been through its process of certification. In summary, teachers reported that:

- ◆ the certification process had made them better teachers (92%),
- ◆ was an effective professional development experience (96%),
- ◆ equipped them to create better curricula (89%),
- ◆ improved their ability to evaluate student learning (89%),
- ◆ enhanced their interaction with students (82%), parents (82%) and colleagues (80%).

Some representative comments include:

“The National Board Certification process was by far the best professional development I have been involved in. I did not realise how much I still needed to learn about impacting student learning. I learned so much through hours of analysing and reflecting.”

“I gained valuable insight of myself as a teacher. The process helped me to assess my teaching abilities as no administrator could have. Most importantly, my students benefit from my self-improvement.”

“Working with other teachers in my school who were also working on certification was rewarding”

“It was the hardest thing I have ever done and it is something I am so glad that I tried. I am immensely proud of the work I turned in – even if I did not make the needed grade. It has made me a better teacher and colleague.”

The certification process has the effect of engaging many teachers in forms of professional learning that are consistent with research on the characteristics of effective professional development (Little, 1993; Wilson, 1999; Ball & Cohen, 1999). Teachers regularly claim that they have become better teachers *as a result* of the certification process (Ingvarson, 1999b; Wolf, forthcoming).

One reason appears to be that each portfolio entry must contain evidence of what the students are doing (video), or evidence from student work samples – evidence which can be directly linked to what the teacher is doing, together with analysis and reflection by the teacher on that evidence. Completion of each entry necessarily engages a teacher in close analysis of, and reflection on, their teaching and its effects on students' learning, tied to concrete examples of student work. Candidates often join networks of other candidates or form their own to help them do this. They often make many videos for their portfolios and use other teachers and students to help them with their analysis, a process strongly encouraged by the NBPTS. The standards provide significant reference points for this process, consistent with the idea of schools and networks of teachers as professional learning communities.

According to Berliner (1992), opportunities for reflected-upon classroom experience with colleagues help to explain the differences between novices and expert teachers. The research on teacher change indicates that this kind of interaction with colleagues about the details of student learning and de-privatisation of practice is one of the defining characteristics of good professional development.

Little research has been completed on the effects of undertaking the threshold assessment in the UK, although it is possible to say that the process is very different.

Reports so far indicate few teachers gain professional development from completing the threshold application form Haynes et al.(2001) surveyed English teachers who had prepared for the threshold promotion and found that 1% reported that the experience had had a positive experience on their practice. 98% said it had had detrimental effect on their morale, almost the reverse of teachers' response to National Board certification. Avenues for collegial analysis and reflection through preparation for the threshold seem unlikely to develop as well, unlike those created by or taken up rapidly by Board candidates. Marsden (2001) surveyed teachers a year after they had applied for the Threshold. 92% agreed with the statement that the 'Threshold has had no effect on the quality of my work because it is already at the appropriate standard'. A cause for concern was that the majority of the teachers believed that school managers would use the threshold process to reward their favourites, a situation that can not occur with the NBPTS certification as assessors can not assess teachers whom they know personally.

UK teachers are not asked to provide 'whole' examples of their work, as teachers are for NBPTS portfolios, nor are they asked to provide analysis and reflection on actual 'cases' of teaching, making use of videos or student work. It seems unlikely, therefore, that the threshold process will engage teachers in similar kinds of analyses to those for National Board candidates. UK teachers can provide evidence for the threshold assessment such as 'schemes of work, lesson plans, feedback from observation (by a school manager or OfSTED inspector), but these tend to be isolated and unrelated pieces of information.

The Threshold guidelines advise teachers, Don't say "I make sure my teaching is appropriate to each child's needs". Do say "Feedback from observation/OfSTED praised the way..." This advice does not seem to place

much value on self-analysis and reflection. Unlike the Board's assessment tasks, the threshold form provides little structure or guidance for analysis and reflection on teaching, or about how the evidence teachers provide will be 'scored'.⁸

Conclusion

This article has compared two strategies for reforming the teaching profession, one from the UK, the other from the USA. The focus was on the capacity of these reforms to strengthen professional responsibility and engage all teachers in effective forms of professional development.

In England, the 1998 Green Paper *Teachers: meeting the challenge of change* 'sets out the Government's vision for a modernised teaching profession'. The Minister's aim is to 'strengthen school leadership, to provide incentives for excellence, to engender a strong culture of professional development, to offer better support to teachers to focus on teaching in the classroom, and to improve the image, morale and status of the profession.'

This could have been a vision to which the profession was invited to contribute, for example, by engagement in developing standards and assessments, but that was not the direction in which the government chose to proceed. The policymakers in the UK who devised the Green Paper have not brought the profession along with them in these reforms. Nor have they capitalised on the many opportunities these reforms might have provided to enhance the responsibility the profession undertakes for the development of standards, assessments and professional development. Professional involvement in its operation has been minimal. (Mahoney et al 2002)

As the Green Paper indicates, the threshold is part of a unified system for the management of teacher performance, not a system for professional development and recognition.

⁸I have just worked with a group of teachers at Monash University who have completed NBPTS portfolios together over several months and, despite the work involved, there is no doubt they enjoyed the experience. It is hard to imagine that teachers would enjoy filling in the form for the threshold assessment. The Board's approach appeals to teachers' professionalism and imagination – 'show us an example of how you engage your students in scientific inquiry', the other seems to threaten, and deaden – show us your 'books marked in line with school policy'.

Managers manage teacher performance and teachers teach. The threshold reform treats assessment as an event not a process for learning. As a consequence perhaps, the performance assessment process for the threshold in its current form does not look as if it will stand up well to scrutiny against standards for the evaluation of systems for evaluating teachers (e.g. Joint Committee on Standards for Educational Evaluation, 1988).

The need for a better balance between political and professional authority is not recognised in the Green Paper reforms. In fact there is little evidence in official documents that teaching standards and assessments might be even considered to be an area where the profession could have expertise. While the Hay McBer model emphasises that effective teachers 'create trust', provide 'challenge and support', and build 'confidence' and 'respect for others', these standards do not appear to have been applied by the government to the way it works with the teaching profession in England. As good teachers develop, they slowly learn how to 'let go' control. Maybe governments have to learn to do the same, if they are to promote the capacity and commitment in the teaching profession on which the success of their reforms will depend.

In the current context in the US, the National Board for Professional Teaching Standards plays a more limited role in reform. It is not the kind of all-embracing government initiated reform we see with the UK Green Paper. Such an approach would not be possible anyway in the federal US context. The National Board has progressed slowly and steadily because it has worked on carrying out one core function well – to provide a national voluntary system to assess and certify teachers who reach high standards.

As NBPTS certification gains credibility, governments and education authorities are creating a market for 'National Board Certified Teachers'. Forty-four states now give tangible forms of recognition, such as salary increases. Many are reorganising their professional development resources to build an infrastructure of resource centres, teacher

networks, university-school partnerships, and so on, to support teachers preparing for NBPTS certification and reallocating PD funds to cover costs of certification. Certification is also redefining the nature of university masters courses that US teachers routinely take for salary progression.

The independence of the Board helps to ensure this job is carried out in the most credible way possible. As the Boards' certification gains respect, indications are that the effects of this function will be far reaching.

The National Board is an example of how the profession can build its own infrastructure for defining teaching standards, promoting development towards those standards and providing recognition for those who reach them – a standards-based professional development system (Ingvarson, 1998). It is a system for which teachers, individually and collectively, can feel responsible. As a professional development system, it is complementary to, not a replacement for, the in-service education that employers should provide to support the implementation of changes and reforms they have initiated.

A professional certification system is a broad reform strategy for the collective advancement of the profession, one that does not rely primarily on government action or the imperfect working of the market (Sykes, forthcoming). The NBPTS, for example, was forged from the mutual interests of teacher organisations, politicians and other stakeholders in quality teaching, who found they were readily able to build a common vision for the teaching profession. Although they had their differences in other arenas, when they came together, in the arena that the NBPTS provided, to talk only about good teaching and how to promote it, those differences faded into the background. The discussion had been moved outside the industrial agenda. To attend a meeting of the NBPTS is to witness something of a miracle for eyes accustomed to Australian industrial relations. Yet industrial relations between unions and employers in the US have been as fraught as in Australia, or the UK, perhaps more so (Kerchner, Koppich, & Weeres, 1997).

What is clearer now is the necessary relationship between the development of teaching as a profession and the development of more effective systems for teacher evaluation and professional development based on profession-defined standards. As we contemplate strategies for revitalising the teaching profession and promoting quality professional development in Australia, the message from these two reforms is that the strategy of establishing an independent national body with a clearly defined certification function such as the NBPTS has much more potential than performance management system such as that in the UK.

References

- Ball, D.L. & Cohen, D.K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds), *Teaching as the learning profession: Handbook of policy and practice*. San Francisco: Jossey-Bass.
- Berliner, D.C. (1992). The nature of expertise in teaching. In E.K. Oser, A. Dick & J-L Patry (Eds.), *Effective and responsible teaching: the new synthesis* (pp. 227–248). San Francisco: Jossey-Bass.
- Blase, J. (1990). *The politics of life in schools: Power, conflict and cooperation*. Newbury Park, CA: Sage Publications.
- Bond, L., Smith, T., Baker, W., and Hattie, J. (2000). *The certification system of the National Board for Professional Teaching Standards: A construct and consequential validation study*. Greensboro, NC: Center for Educational Research and Evaluation, The University of North Carolina at Greensboro.
- Bridges, D. (1990). Evaluation for tenure and dismissal. In J. Millman, & L. Darling-Hammond, L. (Eds), *The new handbook of teacher evaluation* (pp. 35–45). Beverley Hills: Sage Publications.
- Brophy, J. (Ed.), (1991). *Advances in Research on Teaching, Volume 2. Teachers' Knowledge of Subject Matter As it Relates To Their Teaching Practice*. Greenwich, CT: JAI Press.
- Carnegie Forum on Education and the Economy, (1986). *A nation prepared: Teachers for the 21st Century*. Washington, D.C. Carnegie Forum on Education and the Economy.
- Department for Education and Employment (2000a). *Teachers: Meeting the challenge of change. Technical consultation document on pay and performance management*. London: Department for Education and Employment. (www.dfes.gov.uk/teachers).
- Department for Education and Employment (2000b). *A model of teacher effectiveness report by Hay McBer to the Department for Education and Employment*. London: Department for Education and Employment.
- Gitomer, D. (Forthcoming). Reliability and NBPTS assessments. In L.C. Ingvarson (Ed.), (Forthcoming) *Assessing Teachers for Professional Certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier.
- Hattie, J. (Forthcoming). Validating the specification of standards for teaching: Application to the National Board for Professional Teaching Standards. In L.C. Ingvarson (Ed.), (Forthcoming) *Assessing Teachers for Professional Certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier.
- Heller, J.I. & Gordon, A. (2002). Impact of the National Board Certification Process on Teachers' Pedagogical Perception and Teaching Practice. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 2002.
- Ingvarson, L.C. (1994). Setting standards for teaching: an agenda for the profession. In F. Crowther, B. Caldwell, J. Chapman, G. Lakomski, & D. Ogilvie (Eds). *The workplace in education: Australian perspectives*. Australian Council for Educational Administration 1994 Yearbook. Rydalmere, NSW: Edward Arnold.
- Ingvarson, L.C. (1998). Professional development as the pursuit of professional standards: The standards-based professional development system. *Teaching and Teacher Education*, 14(1), 127–140.

- Ingvarson, L.C. (1999a). Science teachers are developing their own standards. *Australian Science Teachers Journal*, 45(4), 27–34.
- Ingvarson L.C. (1999b). *The power of professional certification*. Unicorn, 25(2), 52–70.
- Joint Committee on Standards for Educational Evaluation, (1988). *The Personnel Evaluation Standards: How to Assess Systems for Evaluating Educators*. Newbury Park: Sage Publications.
- Kerchner, CT., Koppich, J.E. & Weeres, J.G. (1997). *United mind workers: Unions and teaching in the knowledge society*. San Francisco: Jossey-Bass.
- Little, J.W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15(2), 129–151.
- Mahony, P., Hextall, I., & Mentor, I. (2002). *Managing teachers' performance? A critical policy study of Threshold assessment in England*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 2002.
- Millman, J. & Darling-Hammond, L. (1990). *The new handbook of teacher evaluation*. Beverley Hills: Sage Publications.
- Messick, S. (1992). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd Edition., pp.13–103). New York: Macmillan.
- Moss, P.A. (Forthcoming). A critical review of the validity research agenda of the National Board for Professional Teaching Standards. In L.C. Ingvarson (Ed.), (Forthcoming), *Assessing teachers for professional certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier
- Murnane, R. & Cohen, D.K. (1986). Merit pay and the evaluation problem. *Harvard Educational Review*, 56(1), 1–17.
- National Board for Professional Teaching Standards (1989). *Toward high and rigorous standards for the teaching profession*. Detroit: NBPTS.
- National Commission on Excellence in Education (1983). *A nation at risk: The imperative of Educational reform*. Washington, D.C.: U.S. Department of Education.
- National Commission on Teaching and America's Future (1996). *What matters most: Teaching for America's future*. New York: National Commission on Teaching and America's Future.
- National Council of Teachers of Mathematics (1991). *Professional Standards for the Teaching of Mathematics*. Reston, Virginia: NCTM.
- Pearlman, M. (Forthcoming). The Design Architecture of NBPTS Certification Assessments. In L.C. Ingvarson (Ed.), (Forthcoming) *Assessing Teachers for Professional Certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier
- Ramsey, G. (2000). *Quality matters. Revitalising teaching: Critical times, critical choices*. (Report of the Review of Teacher Education.) Sydney: NSW Department of Education
- Secretary of State for Education in England & Wales. (1998). *Teachers: Meeting the Challenge of Change*. London: Department for Education and Employment.
- Scriven, M. (1994). Using the duties-based approach to teacher evaluation. In L.C. Ingvarson and R. Chadbourne (Eds). *Valuing teachers' work: new directions in teacher appraisal*. Melbourne: Australian Council for Educational Research.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Education Review*, 57(1), 1–22.
- Silver, E., Mesa, V., Benken, B.M., & Mairs, A. (2002). *Putting it together: What do the NBPTS data tell us about teaching and assessing mathematics for understanding in the middle grades?* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 2002.

Sykes, G. (Forthcoming). Educational reform and the National Board for Professional Teaching Standards. In L.C. Ingvarson (Ed.), (Forthcoming) *Assessing Teachers for Professional Certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier.

Tracz, S. & Associates (1995). Improvement in teaching skills: perspectives from National Board for Professional Teaching Standards field test network candidates. Paper presented at the Annual Meeting of the American Educational Research Association, April 18–22, San Francisco.

Wilson, S. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. In A. Iran-Nejad & D.P. Pearson (Eds), *Review of Research in Education*, Volume 24. Washington: American Educational Research Association.

Wolf, K. & Taylor, G. (Forthcoming). Effects of National Board for Professional Teaching Standards Certification on teachers' Perspectives and practices. In L.C. Ingvarson (Ed.), (Forthcoming) *Assessing Teachers for Professional Certification: The National Board for Professional Teaching Standards*. Amsterdam: Elsevier