

Performance-related Pay: the Views and Experiences of 1,000 Primary and Secondary Head Teachers

Ted Wragg, Gill Haynes, Rosemary Chamberlin, Caroline Wragg

Dr Gill Haynes, Dr Rosemary Chamberlin and Dr Caroline Wragg are Research Fellows at the School of Education and Lifelong Learning at the University of Exeter. They are members of the Teachers' Incentive Payment Project (TIPP) research team, directed by Professor Ted Wragg, investigating the introduction of performance-related pay in the teaching profession in the United Kingdom.

CORRESPONDENCE

Professor Ted Wragg, School of Education and Lifelong Learning, University of Exeter, Heavitree Road, Exeter EX1 2LU

Tel: +44 (0)1392 264826; Email: E.C.Wragg@ex.ac.uk

ABSTRACT

This is the first of two papers describing a study of the introduction of performance-related pay into the teaching profession in the United Kingdom. It reports the views and experiences of a national random sample of 1,000 primary and secondary head teachers in over 150 local education authorities in England who were responsible for implementing one strand of the government's performance-related pay scheme, Threshold Assessment. The second paper describes the views and experiences of teachers who were unsuccessful in crossing the threshold and therefore did not obtain a pay increment.

Head teachers did not find it difficult to assess the five standards that teachers had to meet in order to receive their £2,000 additional performance payment, but they were very critical of the training they received, the amount of time they had to spend, and the changing ground rules. The success rate was 86% of all teachers eligible, but 97% of those who actually applied were awarded the additional payment.

Most heads dealt with the applications entirely on their own, though one in six, mainly in the secondary sector, shared the task with senior colleagues. Unsuccessful candidates were few in number, but most were deemed to be failing on more than one aspect of their teaching. While those who were successful in crossing the threshold were pleased and relieved, unsuccessful applicants were said to be bitter, threatening action, in several cases leaving the school.

External Threshold Assessors had to visit every school. In only 71 cases out of 19,183 applicants in our sample of schools was there disagreement. Three quarters of heads felt Threshold Assessment had made a little or no difference to what teachers did in the classroom. This is confirmed by our other

studies, which suggest that teachers simply keep more careful records, rather than change how they teach. Some 60% of heads were opposed to performance-related pay, but 39% were in favour of it in principle, though most of these were unhappy about the way it had been put into practice.

Keywords: performance-related pay; teachers; threshold assessment.

INTRODUCTION

Advocates of performance-related pay claim that its primary purpose in any organisation is to recruit, retain and motivate the workforce. It can also help to focus employees' minds on particular targets or goals (Protsik, 1966); communicate to employees an organisation's core values, and change the culture of that organisation (Kessler and Purcell, 1991). Other benefits cited by the OECD's study of performance-related pay in the public sector (1993) include: saving an organisation money by reducing automatic increments; encouraging greater accountability; and enhancing job satisfaction.

A number of empirical studies have shown that performance-related pay can have a motivational effect (Lazear, 1999; Fernie and Metcalf, 1996). However, there is evidence that performance-related pay works best where there are clearly measurable outcomes (Murnane and Cohen, 1986). A study of civil servants in the Inland Revenue by Marsden and Richardson (1994) found that the introduction of performance-related pay did not result in staff in general feeling more motivated, indeed there was some evidence of demotivation. Other studies undertaken in the public sector have reported similar findings (Marsden and French 1998, Richardson 1999).

While performance-related pay for teachers was only introduced in England in the year 2000, the United States has witnessed many schemes. A survey of plans undertaken by the Education Research

Service in 1978 (Johnson, 1984) found that most ‘merit plans’ had been discontinued for a wide range of reasons including problems in judging teachers’ performance; failure to apply criteria fairly; teacher and union opposition; falling morale; division amongst staff; cost. Despite this, at the end of the 20th century, merit schemes were still emerging in a number of states (Johnson, 2000; Janofsky, 1999). A full review of the research literature on performance-related pay undertaken by the authors of this article was published in an earlier issue of *Research Papers in Education* (Chamberlin et al, 2002).

The Teachers’ Incentive Pay Project (TIPP) is a three year research project, funded by the Leverhulme Trust at the University of Exeter. It is an independent investigation into the impact on classrooms and schools of performance-related payments to teachers, which were introduced by the government on a national scale during the 2000-01 school year.

The research involves four linked studies of Threshold Assessment (the awarding of a pay increase of £2,000 to eligible teachers whose applications are successful) and Performance Management (improving the quality of teaching). The four studies are:

Study 1 Two national surveys of over 1,000 head teachers in primary and secondary schools in England on (a) Threshold Assessment and (b) Performance Management. The findings of the first of these is reported in this paper.

Study 2 Intensive case studies of primary and secondary schools in the Midlands and South of England, with a particular focus on observing classrooms in order to elicit what, if anything, has changed as a result of performance-related pay.

Study 3 Studies of teachers who (a) applied and were successful, (b) applied but were unsuccessful (the subject matter of the second of our two papers), (c) were eligible but did not apply.

Study 4 Interviews with key people such as external Threshold Assessors, agencies providing training for heads and assessors, head teacher and teacher union officers.

This paper reports the analysis of the first of the national surveys in Study 1, describing the experiences and views of 1,000 head teachers in primary and secondary schools. Head teachers were charged with running the Threshold Assessment exercise in their schools, subject to inspection by external Threshold Assessors. The purpose was to determine which of the teachers who were eligible, that is at or above point 9 in the pay scale, and who applied (for some did not), should be awarded a salary increase of £2,000. Their experiences and views about the assessment, therefore, are of considerable importance.

At the beginning of the project extensive interviews with 31 primary and secondary head teachers were carried out. A questionnaire was then constructed to elicit the views and experiences of a large national sample of head teachers on several of the major features of Threshold Assessment, such as the training they had received, applications, time spent, the role of Threshold Assessors, success rates, reactions of those involved. There are many more primary schools than secondary schools, so it was sent to a random sample of one in five primary and middle school heads and one in two secondary heads in schools in England.

Mailed questionnaires typically receive about a ten per cent return, and it might have been lower, since many head teachers were boycotting paperwork at that time. However, the return rate was exceptionally high, some 53% (1,225 out of 2,325) responding within four weeks, from schools in over 150 different local education authorities.

This paper reports the first 1,000 questionnaires analysed (52% primary, 48% secondary) on all the quantified data and these feature in several tables below. In addition there were several sections where heads could respond in their own words. Qualitative analysis of people's spontaneous comments is extremely time-consuming, so we performed an intensive analysis of a random sample of half the freehand responses, 500 questionnaires.

Training for head teachers

Heads were vitriolic in their condemnation of the two training days they had received from the private companies charged with carrying it out, only one in eight describing it as 'good'. Some even walked out and several refused to return for the second day. A massive 57% said they were very dissatisfied and a further 20% said that their training was merely 'adequate'.

Their comments were scathing, many describing it as the worst training they had ever attended on any subject. The trainers themselves appeared ill prepared. Heads said some had confessed openly that they knew little about school management, mechanically putting on numerous overhead transparencies for a few seconds, often with little understanding of the actual content:

“It was the worst experience of my professional career – insulting, disorganised. The linesman at (the football ground where the training took place) was quite interesting – the training was not!”

“Unutterably boring. No unscheduled questions could be answered and these were what I went for, since I can read the information as well as the trainer.”

“If you are selling ‘double glazing’ that hardly anyone wants, that has been badly designed, that is incomplete and does not really work, then you are probably on a loser from the start – especially if you only heard about it the day before and still don’t understand it yourself!”

A small number of heads (about seven per cent) felt that there had been an improvement by the second training day:

“The first was poor. At the second, clearly, lessons had been learned.”

“First session dreadful (muddled, unclear); second session better – system and thoughts had been ironed out.”

“First day extremely controlled with no opportunity for professional discussion. Second half day much better – informative. Treated as a human being.”

Many heads commented that they would have appreciated being briefed by fellow professionals who had themselves been thoroughly trained first. They felt that the conventions were being invented as time elapsed, rather than being thoroughly thought out in the first place. This led to confusion and different interpretations of what they should do, especially about the more complicated cases:

“Even after the second session it was clear that head teachers had differing attitudes towards the process and that uniformity/consistency was unobtainable. One key point was the issue of a significantly inadequate application from a good candidate, perhaps an over modest or self-deprecating person (“George” in the exemplar materials if I recall correctly). From my group it was clear that some head teachers intended to treat such applications as ‘not yet met’ whereas others were prepared to ‘fill in the gaps’.”

Shifting ground rules

Policy seemed to be changing during the training period and this generated considerable uncertainty. Trainers themselves appeared unsure about their central messages and this transmitted itself to the head teachers, adding to their concern. There was considerable confusion about the processes and conventions to be applied.

The ground rules appeared to be shifting constantly, even between the two training days, and on such fundamental issues as how many teachers should progress through the threshold. This rapid change in policy was confirmed by assessors themselves when we interviewed them in *Study 4* of this research project, which will be reported more fully in a later paper.

Trainers were given fresh instructions as training progressed, to reflect what appeared to be changes in policy. This affected head teachers' confidence to proceed, especially when they were denied the opportunity to ask questions, or when the answers seemed uncertain and imprecise. Many said they had initially been given the impression that relatively few applications would be successful, but then the message from trainers switched dramatically, implying that most teachers would get the £2,000 pay award:

“There were far too many questions which the leaders/advisors were unable to answer. Some answers had even changed overnight since their briefing meeting!”

“At the first whole day training we were led to believe that only ‘super’ teachers were eligible for threshold payments, but at the second half day it was ‘satisfactory and above’. Confusing!”

The standards of teaching

Teachers' applications were judged on eight standards under five main headings:

1. Knowledge and understanding
2. Teaching and assessment
3. Pupil progress
4. Wider professional effectiveness
5. Professional characteristics

Some of the standards were more easily judged than others and there were several differences between primary and secondary heads' responses. When heads were asked to rate, on a four point scale, the ease or difficulty they felt when making judgments, most reported it to have been easy rather than difficult. Table 1 shows their perceptions of the assessments.

Table 1: Percentage of 1,000 head teachers saying how easy/difficult it was to assess each of the five standards

	very easy	quite easy	quite difficult	very difficult
1. Knowledge and understanding	33	54	12	1
2. Teaching and assessment	29	56	14	1
3. Pupil progress	22	43	29	6
4. Wider professional effectiveness	26	51	20	3
5. Professional characteristics	26	49	22	3

Table 1 shows that fewer than a quarter reported difficulty on four out of the five areas. It seemed a little harder to judge *Pupil Progress*, however, with over a third saying it was 'quite difficult' or 'very difficult'. This was for a variety of reasons, sometimes because more than one teacher taught the same class, as frequently happened in secondary schools, so picking out the contribution of just one of them was not always straightforward. Children with Special Educational Needs usually learn more slowly than other pupils and some heads found it difficult to make a fair assessment of what might reasonably be expected. One primary head pointed out: "Pupil progress was sometimes in part a result of other interventions in the school e.g. booster classes." A secondary head was concerned at some teachers selecting those groups with the best evidence of progress: "What about other classes?!"

Primary head comments on the pupil progress standard sometimes reflected a more holistic view of progress, less influenced by numerical data. This may, in part, be due to a lack of statistical information available to primary teachers in years when there were no national tests:

“There was insufficient data for individual teachers (e.g. Year 4 and Year 5) to support their applications regarding pupil progress. This is now being addressed so that future applications will be easier to complete.”

Several heads referred to ambiguities in the evidence they should expect from teachers or were allowed to use and would have liked further guidance:

“*Evidence* could be very selective. Therefore a teacher chose what to put down – did you judge only that or on everything you knew about them? In other words, were you judging the application or the applicant?” (Secondary)

Some head teachers said judging the standards had been easy because of the way in which their teachers completed the application form:

“My colleagues spent much of their summer half term completing their forms. I was very impressed with the detail and thoroughness.” (Primary)

“Most of my staff worked hard to produce strong applications which were easy to assess.” (Secondary)

These were generally schools where the head had also indicated that teachers had been given comprehensive training and support on how to complete their applications. This was, apparently, not the case in all schools:

“Teachers do not have much guidance for completing the application forms. There is a need to make the criteria against which assessment is made very clear to teachers.”

This may have been the reason why some teachers submitted what one head described as evidence which was “sketchy or unclear”. Several head teachers pointed to very capable teachers who submitted poor applications and less effective teachers who submitted good ones:

“Some of our most successful (as judged by OFSTED, me, advisers, parents, pupils) did not do so well with the application form as some of our merely competent teachers.” (Secondary)

Time

The amount of time heads said they spent on each application ranged from under an hour, in 11% of cases, to over four hours, in 8% of schools. Table 2 shows the range. Secondary school heads averaged about one and three quarter hours per application, primary heads nearer two hours. The burden was taken almost entirely by head teachers themselves, only one in six (17%) saying that it was shared with a senior colleague.

Table 2: Time spent on each application by 1,000 head teachers

Time spent on each application	Percentage
Under 30 minutes	2
30-60 minutes	9
1 hour	35
2 hours	31
3 hours	15
4 or more hours	8

Many heads, both primary and secondary, resented the amount of time and bureaucracy involved with the Threshold Assessment procedure:

“No one made the days longer so that I could cope with 32 forms. The three who I judged ‘not yet met’ took a long time to consider, consult and complete. I had to get them right and they probably took 3 hours each at least.” (Secondary)

“It was more a question of the process being very time consuming rather than easy/difficult.”
(Primary)

Heads often preferred to use their knowledge of the staff in their school and felt that the bureaucracy was there to impress the assessors, though they were not convinced that the latter would be any more enlightened about what were sometimes seen as vague and highly subjective guidelines:

“I am accustomed to judging the work of teachers by many informal observations and by seeing them at work. The judgement I had to make about the evidence was whether it would be strong enough to convince an outsider who had not observed their work.” (Primary)

“With the exception of teaching and assessment, they were not standards but guidelines, leaving a lot to judgement. I do not object to this as long as my judgement is respected. The standards on pupil progress were particularly vague.” (Secondary)

“The problem was not in understanding the evidence but in understanding the standard. Teacher effectiveness is on a continuum and exactly where the threshold line is, is not clear even now. I only hope the assessors have a clear idea.” (Secondary)

Sharing the task

Heads often did not delegate any of the assessment to a senior colleague because “We were told not to”, or because they thought it not fair on the overworked deputies or on the staff who were being assessed. One said:

“Senior colleagues applied. I did not consider it advisable to delegate this work.” (Primary)

Another stated:

“I did all the assessment and verification. I spent a lot of time on the latter because I wanted to be able to show beyond doubt the accuracy/rigour of the process.” (Secondary)

Of the one in six heads (17%) who did share the task, three quarters were from the secondary sector:

“I dealt with half of the applicants and my deputies dealt with a quarter each. We all discussed all the applicants and I reviewed all the assessments made.” (Secondary)

Although most heads did not necessarily *delegate* the assessment of applications to anyone, they did, however, detail a variety of ways in which they used their deputies, senior and some middle managers. Heads sought assistance for a variety of tasks. ‘Verifying evidence’ was the most frequently mentioned followed by ‘helping staff’. Colleagues were asked to give second opinions and also to undertake data analysis, lesson observation and to chase evidence:

“Deputies and two senior teachers were involved in the initial assessment of those they line-managed.” (Secondary)

“I asked a deputy to comment on my provisional judgements where I felt the standards may not have been met.” (Secondary)

“I regularly discussed the process with the core leadership team but I made the judgements myself.” (Primary)

The teachers who applied

Most eligible teachers (on point 9, or beyond, of the pay scale) applied for the £2,000 payment in the 1,000 schools studied. Table 3 shows that there were 21,749 teachers eligible to apply in our sample of 1,000 schools, and 88% of those actually put in an application. In three quarters (76%) of schools, however, at least one teacher who was eligible did not apply. Our other studies in this research project, to be published later, will show that people did not apply for different reasons, some refusing on principle, others being discouraged by the head or a senior member of staff.

'Success' rates

There were 18,684 successful and 499 unsuccessful applications. In nearly three quarters of schools (72%) all teachers who applied were successful. These figures are not as straightforward as they look and are open to various interpretations. They can be presented in different percentage forms. The success rate in terms of *all eligible candidates* was 86%. Taken as a *percentage of those who actually applied*, however, the success rate was 97%.

Table 3: Teachers in the 1,000 schools who applied to cross the pay threshold

Number of teachers who were eligible to apply	21,749
Number of teachers who actually did apply	19,183 (88% of teachers eligible)
Number of teachers who were successful	18,684 (86% of teachers eligible, 97% of all teachers who applied)
Number of teachers who did not apply	2,566 (12% of teachers eligible)
Number of teachers who applied without success	499 (3% of all teachers who applied)
Schools where every applicant was successful	72%
Schools where not every eligible teacher applied	76%

It is difficult to suggest what figure in between 86% and 97% should be seen as a 'true' success rate, given that many teachers thought unlikely to succeed were discouraged from even applying, and in this

particular study we do not know all the relevant details. More information on this aspect is given in the second of these two papers. Whichever way the figures are interpreted, it is certainly the case that the vast majority of eligible teachers did succeed in their applications.

Advising teachers

Some heads felt strongly it was not their role to advise teachers on whether they should apply. Indeed, they had been told they were not to intervene and only a few (7%) said they had advised people not to apply:

“I am amazed you ask! It would be totally improper.” (Secondary)

“I did not feel this was a suitable action to take.” (Primary)

“It would have been wholly inappropriate to do so.” (Secondary)

Other heads said they had been warned by their union not to advise staff at all in this respect:

“Even had I thought it necessary, I was warned off by my professional association because of legal consequences.” (Secondary)

Nonetheless the seven per cent who did intervene, and this was twice as likely to occur in a secondary school as in a primary, said they had made sure that certain teachers were told, directly or indirectly, that their applications would *not* be supported:

“The deputy head who was given responsibility for doing preparation work with eligible staff had a number of conversations with several who did not choose subsequently to apply.”
(Secondary)

This was done through intermediaries. Given the fact that heads were not supposed to talk direct to applicants or potential applicants this process had a slightly surreal feel to it. Nevertheless two were actively discouraged. (Secondary)

“I explained [to the teacher] that there were deficiencies in one area and that I could not support her application.” (Primary)

Other heads indicated more subtly to teachers their applications would not be supported:

“Not directly, but in training sessions to staff I did make criteria clear and discuss scenarios that may make it difficult to cross the threshold (e.g. recent disciplinary interview).” (Secondary)

“The three who did not apply were all receiving support in order to improve their performance. Although no direct advice not to apply was given, they were aware of the management’s view of their shortcomings.” (Secondary)

“Training alluded to standards. I gave the message not to bother if they did not meet the standards.” (Secondary)

Other heads encouraged *all* the teaching staff to apply:

“I advised all eligible to apply.” (Secondary)

“I insisted all should apply and that I would support them.” (Primary)

“I advised all of those eligible to apply. I told them I considered them all to be worthy of success and that they should support one another by entering the fray together.” (Primary)

As we found with the interviews we carried out with head teachers before the process began, some of the most effective members of staff had been initially very reluctant to apply and had had to be cajoled by their head into submitting an application:

“I did tell one very good teacher who is fearful of failure that he had to apply!” (Secondary)

“I persuaded two to apply who were not going to apply.” (Primary)

“I had to work hard to persuade some very deserving colleagues to make an application.”
(Secondary)

“I had to encourage some to apply. They did not think they would good enough or that they would get it – often they were the *best* teachers!” (Secondary)

Unsuccessful applicants

The vantage point of teachers deemed ‘unsuccessful’ or ‘not yet met the standards’ is described in the second of our two papers (Haynes et al, 2002) so in this paper we only give the head teachers’ perceptions. Since only 3% of applicants were unsuccessful the numbers involved in the 28% of schools where someone was turned down were bound to be small, in three quarters of cases it was either one or two teachers. Heads reported that it was extremely rare for a teacher to have failed under only one of the five ‘standards’. Most teachers were judged ‘not yet met’ on at least two or more:

“Known problems observed by me in class management; lack of colleagues’ confidence in team leader; questionable pupil progress.” (Secondary)

“Poor planning; constant parental complaints; lack of support for other colleagues; negativity of attitude.” (Primary)

The most commonly mentioned standards were *Teaching and Assessment* and *Pupil Progress*. Within *Teaching and Assessment*, poor class management, poor teaching quality and lack of marking were mentioned specifically.

Some heads who had judged the applications of *supply* teachers in their school pointed to the problems that this special group may face in providing evidence:

“Supply teacher - who had only worked occasionally in the school and had not been formally monitored.” (Primary)

“Two [who didn’t cross the threshold] were supply teachers who had great difficulty accessing evidence. Equal opportunities issue here, since most supply teachers are female.” (Secondary)

There were also a number of cases where heads felt unable to support teachers’ applications because they believed that the evidence cited on their forms could not be substantiated and ran counter to the head’s own knowledge of the teacher:

“Scrutiny of planning and work samples, and lesson observations, indicated that claims made in the application were not correct.” (Primary)

A number of heads mentioned the *quality* of the application form itself:

“Muddled application” (Primary)

“Woefully inadequate form. Teacher had spent 30 minutes on it.” (Secondary)

The Threshold Assessor

The role of the external Threshold Assessor was potentially extremely sensitive. A number of heads had themselves trained as assessors, not always with the intention of going round other schools, checking their procedure, sometimes more to be able to understand the process from the inside.

Furthermore there had been a great deal of publicity about daily payments in excess of £300 and many assessors were also inspectors with the Office for Standards in Education (OFSTED).

In the event over 90% of heads were satisfied with the actual arrangements for the assessor's visit and Table 4 shows some of the answers to questions about the assessor. They felt well informed about the structure of the visit and what was needed, though fewer than half (46%) the assessors gave any indication in advance to their personal views about the applications. A few heads grumbled about the amount of documentation required and some felt the assessor was behaving too much like an OFSTED inspector, but these were a small minority of cases:

“Quite comfortable. The answer was speedy, efficient and affirmative of the work we had done.” (Secondary)

“I am proud of what we achieved and I don't object to letting people know this or how hard my staff work and how good they are.” (Secondary)

There was a significant minority who were irritated by it and found that it was time consuming, not always relevant or in their opinion they were asked for too much. Some ventured into sarcasm to explain how they felt:

“Over the moon! It is so refreshing having another inspection by someone of doubtful ability checking up on me yet again. Clearly I cannot be trusted and I now understand that.”
(Secondary)

“More bureaucracy with inspectorial overtones.” (Secondary)

Table 4: Responses of 1,000 heads to questions about the Threshold Assessor’s visit

Yes, satisfied with arrangements for assessor’s visit	92%	
Head ‘well informed’ or ‘quite informed’ about the structure of the visit	94%	
Assessor asked for information about school context before visit	92%	
Assessor discussed sample of applicants before visit	86%	
Assessor indicated own view of applications before visit	46%	
Assessor asked for teachers’ evidence to be available on day of visit		90%
Assessor did job ‘very effectively’ or ‘quite effectively’	97%	

Providing evidence for the Threshold Assessor

Assessors had been warned in their own notes of guidance (Cambridge Education Associates Professional Guidance Note No 5 for External Assessors dated 19 February 2001) not to be excessive in any requests for evidence. The emphasis was meant to be on verification:

“Only the focused evidence necessary to enable verification of the selected Standards of those in the sample should be requested ... Underlying all strategy in this regard must be the principle that the purpose of the Threshold Assessment visit is to verify and not to assess. Large volumes of evidence are neither necessary nor desirable in this regard.”

Head teachers were asked what *type* of evidence the teachers in the sample were asked by the Threshold Assessor to provide. Some simply stated that the evidence requested was tailored to support statements on individual teacher’s applications. Others provided a list of different types of evidence. Most were

clearly intended to provide evidence under the standards relating to *Teaching and Assessment* and to Standard 3 – *Pupil Progress*:

“Teachers’ lesson plans; exam data; pupils’ planners; pupils’ work; schemes of work; GCSE specific material; evidence of quality teaching; assessment evidence.” (Secondary)

“Short-term planning showing how different pupil groups have their needs met; evidence of how they mark and assess children’s work: marking records, samples of books, pupil portfolios; evidence of lesson observations by SMT of the teacher’s work.” (Primary)

“Whatever they [the applicants] wanted to provide. Some provided a box full – quite wore the assessor out. It was his first job. He quickly realised he had too much evidence – here all day.” (Secondary)

Lesson observation

This section and the following one are extremely brief, for a very good reason: there is virtually nothing to report. That assessors might enter classrooms to observe lessons was a theoretical possibility, they had the option of so doing, but not a reality. The emphasis was very much on paperwork. The truth of the matter is that, out of 1,000 primary and secondary schools, only one head teacher reported that the assessor had actually observed a lesson. Hundreds of thousands of hours must have been expended nationally by assessors scrutinising paperwork and interviewing people, but virtually no time at all observing anyone teach.

Disagreements between Threshold Assessors and head teachers

This section too is brief. Out of the 19,183 applicants considered by assessors in the 1,000 schools in this sample, there were only 71 reported cases in 63 schools (14 primary, 49 secondary) where head teacher and assessor disagreed. This is about one third of one per cent. The two parties agreed over about 99.6% of teachers. Inside this infinitesimal disagreement there were 44 cases where the assessor thought the teacher a failure, but not the head, and 27 cases the other way round. Some of these cases were of supply teachers, others involved long term ill health and lack of evidence.

Nearly all final decisions had been reached in an amicable manner. In only two cases was there a hint of tension in the relationship:

“I was able to provide enough verbal evidence and reason about why the applicant should cross the threshold and eventually convinced her. (Primary)

“Each time the assessor raised a point of contention, my team and the teacher concerned produced additional evidence to back my judgement. We refused to accept that a member of staff should fail on a technicality. Eventually he gave in.” (Secondary)

It does not seem easy to justify the vast amount of time and effort involved and the huge cost of sending assessors to every school in the country, at expensive daily rates, when there was agreement over nearly a hundred per cent of cases. A spot checking verification system in some schools, or an appeals process might be more cost effective.

Overall rating of Threshold Assessors

Whether or not they are needed, assessors do appear to have been an extremely effective lubricant in what could have been a most difficult assignment. Their high approval rating from head teachers (97% rated 'very effective' or 'quite effective', as shown in Table 4) may in part be explained by the massively high level of agreement between the two parties, but heads' written comments were also mainly positive, with few carping. "Professional," "efficient" and "thorough" were descriptions of the Threshold Assessors, repeated time and again by both primary and secondary head teachers. A few complaints were made about the process being "too OFSTED based", or the individual assessor being "an OFSTED clone", almost all from primary heads, whereas all other comments, positive, negative or ambivalent, came equally from both sectors:

"Excellent - very professional and knowledgeable of the school context." (Secondary)

"He strove hard to make the process positive. He provided valuable comments on his perception of aspects of our policy, based on evidence seen." (Primary)

"An external assessor for all schools should not be necessary. Head teachers should be able to make this decision (if it should be made at all) with monitoring of a few schools to check consistency." (Primary)

"Nice man - served no useful purpose." (Secondary)

Reactions of teachers

Heads informed most teachers (over two thirds) individually of the outcome, though in some schools where all teachers had been successful it was announced to the whole group. Unsuccessful teachers were almost all informed individually, only in four schools was any announcement about failure to cross the threshold made in a group context.

Successful teachers

The same words were used over and over to describe the reactions of successful teachers. Given that 97% of applicants eventually received their £2,000 payment, there was much relief and celebration:

“Pleased/relieved/grateful.”

“Good, positive reaction. They felt they had worked hard to prepare their applications.”

“68 people produce different reactions. All were pleased; most believed it to be their entitlement too. There was considerable relief.”

“Pleased, not surprised. Would have been incandescent had they failed!”

“Very pleased! I got chocolates and thank you cards!”

Some successful schools still experienced resentment of the elaborate processes involved:

“Obvious relief, but irritation at the entire process.”

“They were very pleased and to some extent relieved. There was also a degree of resent about being made to ‘jump through hoops’.”

“Relief and pleasure mixed with antagonism about the whole process. Upset that they had to prove themselves and that my judgement was not sufficient.”

Unsuccessful teachers

Though far fewer in number, the 3% of unsuccessful teachers were mainly bitter. Words which featured most frequently in heads’ comments were, as one might expect, “angry”, “disappointed” and “upset”. Some teachers were described as “accepting” and or clearly determined to improve their performance. A number of heads found the experience deeply stressful and demoralising, especially when their colleague was ill or sought redress:

“Disbelief and anger immediately. Meeting finished in an unsatisfactory way and colleague became absent with stress a few days later.”

“Very aggressive. Is appealing. Seeking alternative employment. Bringing unions/legal advice in.”

“Very upset. Has requested a review – I feel we should have been trained for this.”

“Disappointed – keen to know why – what aspect of work had failed; anxious to put weaknesses right and what could they and the school do to support them?”

The effect of Threshold Assessment on teachers and teaching

The impact of Threshold Assessment on classroom practice appears to have been minimal, in the eyes of head teachers. Very few (2%) thought it had improved classroom practice ‘a lot’. Only one in five (19%) thought it had had ‘some’ impact. Over half (52%) felt it had had no influence at all and a further quarter (24%) said it had ‘a little’ effect. There was virtually no difference when heads’ assessment of the impact of Threshold Assessment on successful and unsuccessful applicants were analysed separately. Over three quarters still said it had had little or no effect. *Study 2*, our intensive year long scrutiny of classrooms in individual schools, confirms that the main influence of Threshold Assessment has been to persuade teachers to keep more detailed records of children’s work so they will have more written evidence on a future occasion, rather than change the way they teach.

When performance-related pay was first announced there was a fear that it might have a divisive effect in staffrooms, but that was before anyone knew how many teachers would be successful. Most heads felt the process had had little positive or negative effect on staff relations, mainly because most teachers had been successful. A tiny minority said that it had led to difficult staff relations, often affecting their relationship with an unsuccessful candidate. In a few cases sympathy had been expressed by the staff to the unsuccessful candidates and in one case the head had said that this had caused problems for them, but in the main the process was not seen as unduly influential on the relationships among staff.

Heads' overall appraisal of performance-related pay (PRP)

Of the sample of 500 questionnaires analysed intensively for qualitative responses, 461 head teachers (256 secondary, 205 primary) had responded freehand to an invitation to comment on the whole issue of performance-related pay (PRP). Each head teacher's response was coded and placed under one of three main categories: Positive, Negative, and Mixed. Sub-categories were then created under each main heading. The data were also analysed separately for primary and secondary head teachers, to elicit any differences between the two.

Performance-related pay can be seen in two parts: Threshold Assessment, which was the subject of this research study, and Performance Management, the continuing process whereby teachers are encouraged and supported to improve their skill. It should be mentioned that, although the question had asked about performance-related pay *in general* for teachers, it was clear that many of the head teachers' comments related only to the Threshold Assessment part of it, and their experiences of this process may have skewed their responses. Indeed, some heads felt that the process had been introduced in the wrong order, in that Performance Management should have been the starting point followed by Threshold Assessment and the rewarding of those thought to be doing well:

“It seems to have been a missed opportunity to have the first cohort judged before a year of agreed performance standards! I would have preferred standards described, then a year of performance, then threshold assessment.” (Primary)

Overall, 60% of head teachers indicated that they were against PRP, in principle. Feelings sometimes ran high on this topic:

“I am firmly opposed to it. It is a cumbersome, wasteful and degrading process.” (Primary)

“Totally opposed. We work in a ‘difficult’ area, our success comes from mutual support and high levels of collegiality. Many of the most valuable things we do for our pupils do not fit onto performance rates, including values, morals, extra-curricular experiences. Petty bonuses are divisive and attempts to prove worth are a distraction. It is a process long discredited and discarded by industry and in disrepute in NHS Trusts.” (Secondary)

Only 1% expressed ‘mixed’ feelings, while 39% said they thought PRP was a good idea *in principle*, though many had reservations about its current practice. Several of these spoke with as much feeling in favour of the principle, as those who had railed against it:

“I think it is definitely the way forward. It motivates teachers to improve and reflect on their practice. I have never seen so much interest expressed in the . . . evaluation of school and class data!” (Primary)

“I am very pleased with the principle – as a head teacher I have gained from such a system – quality teachers should also gain.” (Secondary)

However, of those in favour of PRP in principle, a large majority expressed concerns about its current implementation. These tended to focus on their experiences with the Threshold Assessment procedure:

“If done properly and fairly I agree with the best teachers getting additional rewards. I believe the system devised offered an opportunity to allow this to occur. I do not believe the practice has realised the opportunity.” (Secondary)

“In principle I feel that good teachers should be rewarded and those who are ineffective should not. It’s the volume of work collecting evidence on both groups which concerns me.” (Primary)

Further analysis was undertaken to identify any differences in attitude between primary and secondary heads. Table 5 reveals that primary heads are much more opposed to PRP than their secondary counterparts:

Table 5: Percentage of primary and secondary head teachers’ (in 500 schools) comments about the merits of performance-related pay *in principle* in three categories

Phase	Positive	Negative	Mixed
Primary	27%	71%	1%
Secondary	48%	50%	1%

One reason may be that secondary schools have for much longer had statistical data on pupil achievement and progress, and therefore this type of process is less alien to them. Another reason for this disparity lies in the sometimes different cultures of primary and secondary schools. Primary school staff see themselves more as a single team, secondary schools with their different subject departments and faculties appear more fragmented. Certainly, amongst ‘negative’ primary heads, by far the largest sub category was that relating to the potentially divisive nature of the PRP. While just over a fifth of secondary heads expressed concern about this, the figure was a third amongst the primary heads:

“I have no problem in setting targets and objectives, these are vital for development and improvement. I object to the role I had to play. I feel it drove a wedge between myself and staff – which was unnecessary and unproductive. Fortunately all applicants were successful. I dread to think of the consequences if any had been unsuccessful.” (Primary)

Other reasons for opposition to PRP cited by both primary and secondary heads were worries about standardisation within and between schools, and philosophical and practical concerns about measuring teachers’ performance:

“I do not agree with it, as inner city schools/different cohorts/children with special needs/different ethos in each school make decisions subjective, to a degree, and can be unfair.” (Primary)

“I am very uncertain about relating remuneration to a crude set of performance criteria for teachers. I do not believe it practicable to reduce the job of teachers to some kind of ‘productivity’ measure.” (Primary)

“The introduction of PRP is problematic unless a fair and transparent set of criteria can be applied. We all know some teachers are more effective than others. However, how do you quantify the difference? (Secondary)

Those in favour of PRP clearly did not see these issues as problematic and were enthusiastic about the opportunity to reward good teachers:

“As a former LEA officer who has observed hundreds of lessons, I feel strongly that good teachers should be rewarded.” (Primary)

“Jobs done well should be rewarded. Coasters should not gain additional rewards.”
(Secondary)

“Valuable for individuals to see they are rewarded for doing a good job. Children benefit.”
(Primary)

Perhaps the most striking feature of the comments made by ‘positive’ heads was the rarity of explicit references to PRP’s potential impact on *learning*. The focus was very much on the benefit to teachers rather than their pupils. Apart from the last quote above, only one other primary and one secondary head in this group referred at all to the possible repercussion for pupils:

“If managed well, it can enhance pupil performance.” (Primary)

“Generally in favour of the acknowledging the practice of good teachers. However, I feel the standard or pass mark should be the equivalent of OFSTED grade 3 – i.e. good teachers. This would lead to improved classroom practice.” (Secondary)

Amongst those *opposed* to the principle, a number of heads, particularly secondary, expressed the view that PRP would have *no* beneficial impact on teacher performance or classroom practice but, as with responses to other questions, they were often talking within the context of Threshold Assessment, rather than considering the wider picture including Performance Management:

“Very little positive effect on teaching and learning. Negative effects – even the best teachers were concerned that they may not succeed. Gave the school an air of uncertainty, even panic, on occasions.” (Secondary)

“The Threshold Assessment process has encouraged staff to use the school databases and staff now look far more closely at value added. However, it has not improved classroom practice and performance – teachers are not better teachers as a result of the activity.” (Primary)

For heads opposed to PRP the problems seemed obvious. They saw it as divisive, demotivating, impossible to implement fairly, yet one more bureaucratic burden, not likely to raise standards. The call from this group was to increase the pay of *all* teachers in order to address the current retention and recruitment problems and to deal separately with poorly performing teachers through capability procedures.

Those in favour believed that PRP provides heads with the opportunity to reward their good teachers, but they, too, acknowledged the problems with the policy in practice. Moreover, they were seriously worried about the funding implications when, two years’ later, teachers who have crossed the Threshold would be eligible to apply for an increment to move even higher up the pay scale.

It also became apparent from the analysis that, at that time, there were still head teachers who had not made the link between Threshold Assessment and Performance Management, the two key strands of the government's PRP policy. The comment below seemed to exemplify this:

“It is not ‘performance related pay’. That would involve judging teachers’ performance annually and paying them on the basis of those judgements. A daunting and potentially very adversarial task.” (Secondary).

Some 203 head teachers decided to make additional comments, entirely of their own choosing, not structured or solicited by the questionnaire. The majority of these were overwhelmingly negative about the process and these were primarily focused on the additional workload that this had created for both heads and teachers and that it was seen as a bureaucratic and costly exercise:

“I cannot believe that the DfEE and NUT made it one of the most spectacularly frustrating exercises I have carried out in 18 years of headship.” (Secondary)

“Yet another example of a badly thought through ‘top of the head, I had a dream’ government mentality. More paperwork for heads, more paperwork for teachers.” (Primary)

“Very little was achieved from the many hours of work that went into form completion. I agree that the form provided a point of focus, but was the time spent justified? I would have liked to

have seen a rise for the profession which was open to all and not dependent upon hoop jumping.” (Secondary)

“Crazy and expensive. I (and most heads) have shown my commitment to improving teaching and learning through undertaking, when necessary, a competence procedure, and on other occasions providing focussed training and support. The teaching profession needs improved status, respect and pay (which are undeniably linked). This process has not, to my mind assisted.” (Secondary)

A small number of heads (9) commented that they themselves had received no pay rise for all the work that they had incurred as a result of the process. Some commented on the unfairness for those competent younger staff who had not yet reached the appropriate point on the pay scale and so were not even eligible to apply. There were also, in a few cases, concern about the erosion of differentials between classroom teaching staff and those on leadership/deputy head scales, and also about the effects on recruitment.

“Threshold Assessment for all staff, not just those who are on point 9 [of the pay scale]! This has caused much discussion and heated debate and everyone believes that it should be open to all.” (Primary)

“Until there is a sufficient supply of good teachers, all this jumping through hoops is not helpful. If you have a Physics teacher (with a Physics degree) are you going to upset him with an ‘iffy’ Threshold Assessment? Recruitment is a far more pressing concern.” (Secondary)

CONCLUSION

Head teachers felt poorly trained and poorly prepared for assessing teachers' crossing the pay threshold and receiving their additional payment. Contradictory messages from trainers left them uncertain as to whether only the best teachers should receive the payment, or whether almost all, other than those who were incompetent, should succeed. In the end 97% of teachers who applied were actually successful, although 20% of eligible teachers chose not to apply in the first round

Although head teachers did not find it difficult to assess the standards their teachers had to meet in order to cross the threshold, they found the paperwork involved in assessing teachers' applications burdensome and extremely time-consuming. Heads felt that they should have been trusted to make judgements about their own staff without the intervention of an external Threshold Assessor, although there were very few complaints about the way in which these assessors undertook their role. Indeed, there were only 71 reported cases of disagreement between the head and the assessor of the 19,183 applications reported in this questionnaire survey. In almost every case these were resolved amicably.

Head teachers were divided in their opinion of the use of performance-related pay itself, with 60% against it and 40% for it. Even those who supported it in principle expressed concerns about the Threshold Assessment procedure. Primary heads were much more negative about performance-related pay than were secondary heads.

Since the fieldwork for this research was undertaken, the procedures for Threshold Assessment have been modified. Threshold Assessors no longer have to make visits to every school where there are

applicants to cross the threshold. Instead, the contact with schools must be mainly by telephone, although a random sample of two schools out of every ten receive a visit. The initial blanket cover of schools in the first manifestation of performance-related pay was intended to ensure that head teachers carried out the procedure properly and fairly, but clearly it was difficult to justify the vast amount of time and effort involved.

As our review of the literature revealed, performance-related pay has often been controversial and problematic in countries where it has been tried, sometimes leading to termination because of the difficulties associated with its administration. A great deal of value has emerged from this study, recording a whole school system trying to introduce additional rewards on a mass scale. Whether the scheme is enhanced, survives, or ceases is strongly influenced by political and financial factors, not solely by research findings. What we have described, however, does offer decision makers some fundamental evidence, should they choose to use it.

REFERENCES

CHAMBERLIN, R. P., WRAGG, E. C., HAYNES, G. S. and WRAGG, C. M. (2002). 'Performance-related pay and the teaching profession: a review of the literature', *Research Papers in Education*, 17, 1, 31-49.

FERNIE, S. AND METCALF, D. (1996). 'It's not what you pay, it's the way that you pay it: Jockeys' pay and performance', *CentrePiece Magazine*, 2 June.

HAYNES, G. S., WRAGG, C. M., WRAGG, E. C. and CHAMBERLIN, R.P. (2002). 'Threshold Assessment: The experiences of teachers who were unsuccessful in crossing the threshold', *Research Papers in Education, (this issue)*

JANOFSKY, M. (1999). 'For Denver teachers, a pay-for-performance plan'. *New York Times*, 10 September.

JOHNSON, S. (2000). *Teachers' compensation and school improvement: a review of the literature and a proposal to build capacity*. Paper prepared for the National Education Association.

JOHNSON, S. (1984). 'Merit pay for teachers: a poor prescription for reform'. *Harvard Educational Review*, 54, 175-85.

KESSLER, I. AND PURCELL, J. (1991). *Performance-related pay: theory and practice*, Paper presented at the 10th Colloquium for the European Group of Organisation Studies, Vienna, July 1991, Oxford: Templeton College: The Oxford Centre for Management Studies.

LAZEAR, E. (1999). *Performance Pay and Productivity*. Stanford California: Stanford University Research Paper.

MARSDEN, D. AND FRENCH, S. (1998). *What a performance?* Centre for Economic Performance: London School of Economics.

MARSDEN, D. AND RICHARDSON, R. (1994). 'Performing for pay? The effect of 'merit pay' on motivation in a public service', *British Journal of Industrial Relations*, 32, 2, 243-262.

MURNANE, R. J. AND COHEN, D. K. (1986). 'Merit pay and the evaluation problem: why most merit pay plans fail and a few survive,' *Harvard Educational Review*, 56, 1, 1-17.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD) (1993). *Private pay for public work: performance related pay for public sector managers*. Paris: OECD.

PROTSIK, J. (1996). 'History of Teacher Pay and Incentive Reforms', *Journal of School Leadership*, 6, 3, 265-89.

RICHARDSON, R. (1999) *Performance-related pay in schools: an assessment of the Green Papers*, London: NUT.