Performance-related-pay in the UK public sector

A review of the recent evidence on effectiveness and value for money

Zofia M. Bajorek and Stephen M. Bevan
The Work Foundation, London, UK

Abstract

Purpose – The purpose of this paper is to provide an up-to-date, comprehensive, independent and credible assessment of relevant academic and other literature since 2007, on the effectiveness and value for money of performance-related-pay (PRP) in the public sector.

Design/methodology/approach – PRP was studied using both economics-based literature and literature from the organisational and management field (including human resources, management sociology and psychology). An initial search of databases identified 7,401 documents regarding PRP in the public sector, which was reduced to 57 final papers included in the study (27 in the health sector, 16 in the education sector and 16 in the civil service) after abstract and full paper screening.

Findings – The review found some evidence that PRP schemes can be effective across the three domains of the public sector for which there was evidence available (health, education and the civil service), but findings within and between the sectors are mixed, with scheme effectiveness often dependent on scheme design and organisational context.

Research limitations/implications – The research highlights the importance of considering both economics-based and organisational literature when discussing PRP in the public sector, and the implications for motivation and PRP design.

Practical implications – The results indicated that the design of PRP schemes could influence their effectiveness and outcomes, and the research suggests how the challenges of designing and implementing PRP schemes can be overcome in the public sector.

Social implications – The review highlights that when implementing PRP schemes there may be gender differences in their overall effectiveness (especially in education) and there must be consideration for how fairly the PRP scheme is perceived.

Originality/value – The paper uses literature from economics and behavioural sciences when looking at the motivational implications for PRP in the public sector.

Keywords Motivation, Public sector management, Performance related pay, Rewards system

Paper type Literature review

Introduction

The use of pay incentives to motivate performance has long attracted interest in the academic literature (Marsden, 2009), and in practice 80 per cent of OECD member countries have either implemented performance-related-pay (PRP) schemes or are in the process of doing so (Bregn, 2013). The motivational aspects of PRP have been heavily contested and this paper aims to looks at the implications for motivation from two different areas of literature: economics and behavioural sciences. PRP has long been an important idea for economists, and the principal-agent theory (Prendergast, 1999) has often been discussed by economists when explaining the role of PRP and behavioural

Journal of Organizational Effectiveness: People and Performance
Vol. 2 No. 2, 2015
pp. 94-109

2051-6614
DOI 10.1108/JOEPP-03-2015-0011

The authors would also like to acknowledge Kathryn Ray, Beth Foley, Tiffany Tsang and Damian Walne for their contribution to the original review.

Funding: The research was funded by the Office of Manpower Economics.

outcomes, with the assumption that an incentive or reward related to performance will result in a change in behaviour and outcomes. The principal's interest is that a worker works as hard and efficiently as possible to maximise profit, but the agent aims to minimise their effort to reach the basic standards needed to receive pay. PRP is then used to better align the interests of the principal and the agent.

Organisational literature, however, focuses on broader theories to explain human motivation and work (Bregn, 2013). Reinforcement theory suggests that an individual's behaviour can be encouraged through rewards and praise, and when these are experienced their behaviour will be modified accordingly (Reilly, 2003). Expectancy theory developed by Vroom (1964) and later developed by Porter and Lawler (1968) is the belief that individuals will exert effort in a task and they expect that it will lead to an outcome that they value (Perry *et al.*, 2009). Thus, according to expectancy theory when an organisation introduces a PRP scheme, employees should work harder if they value the monetary reward that is offered, and if they believe that the reward will be achieved by increasing their effort. Organisational literature has also emphasised the role of equity theory (Adams, 1963) in social relations as being important for staff morale. Equity theory argues that individuals seek a balance between their input and their reward outcomes (Reilly, 2003), and a perceived lack of fairness may have considerable detrimental effects on performance (Bregn, 2013).

There is some empirical support for principal-agent theory and its underlying assumptions in the private sector (Frey *et al.*, 2013), e.g. Lazear (2000) reported that the introduction of piecework pay in a glass company raised productivity significantly. However, academics have questioned whether such assumptions transfer to the public sector. Reviews studying the role of PRP in the public sector have been undertaken (e.g. Makinson, 2000) as the public sector had introduced incentive schemes at local and national levels, but there had been limited progress in making a link between pay and performance in most public sector groups. Prentice *et al.* (2007) conducted a review of evidence on PRP in three public sector services (the civil service, healthcare workers and teachers) finding that there is evidence that public sector workers do respond to financial schemes, but the overall benefits of the schemes for society could not be assessed.

Use of PRP in the UK public sector

The use of individual PRP in the UK public sector was first introduced as part of formal pay policy in the 1980s. Although it has taken several forms (from appraisal-driven pay, to skill-based progression to team-based bonuses) and has only been used comprehensively in parts of the public sector (e.g. the civil service), its use has been based on a number of principles. First, that an element of public servant's rewards should contain a "contingent" element which reflected their individual performance. Second, that this contingent element should be used to motivate, incentivise and reward higher performance. Third, that the value this "contingent" element would be variable in line with affordability criteria. Fourth, that using performance-related rewards to differentiate between higher level performers and poor performers would promote feelings of distributive justice in the workforce and lead those whose performance who was poor to strive for improved performance ratings. However, the use of PRP in the public sector has always been controversial. Many public sector trades unions and some academics have argued that PRP is a mechanism borrowed from the commercial sector which is inappropriate in an environment where an "ethos" of public service rather that the so-called "profit motive" dominates the culture (Marsden and French, 1998; Kessler and Purcell, 1992). In addition, others have argued that the use of PRP

was always intended by government to individualise the employment relationship in the public sector as part of an explicit attempt to undermine the role of trades unions in collective approaches to pay bargaining. Other, more operational, concerns about the use of PRP have included its potential to reinforce the gender pay gap and adverse outcomes for women (Bevan and Thompson, 1992), the motivational impact of very low or zero awards in periods of low inflation or economic austerity and whether the direct, indirect and displacement costs of administering these schemes can be justified in terms of the performance "premium" they generate. More recently there have been concerns about the use of "guided distributions" to moderate the allocation of the highest performance ratings, with debate about whether these undermine any motivational impact of PRP.

Conceptualising PRP in the public sector

As economic and organisational theories regarding PRP schemes have developed, academics and practitioners have argued whether the logic of PRP can be applied to all organisations – with challenges particularly marked in the public sector (e.g. Bregn, 2013; Frey *et al.*, 2013). One of the main concerns discussed is the extent to which public and private sector employees differ with respect to their sources of motivation (e.g. Brewer and Selden, 1998; Heath, 1999). Intrinsic motivation (the motivation that causes an individual to perform activities because they like them) is of high importance in the public sector (Larkin *et al.*, 2012). Lindenberg (2001) identifies two types of intrinsic motivation: enjoyment-based intrinsic motivation referring to a satisfying flow of activity, and obligation-based intrinsic motivation, which refers to an activity with the goal to act appropriately (thought to be of particular relevance in the public sector). Frey *et al.* (2013) argued that if an employee experiences obligation-based intrinsic motivation, they will follow the norms for their own sake and will take the well-being of others into account without expecting a reward. Thus public sector workers may derive their motivation from their belief in the intrinsic value of the service they are providing, their commitment to a professional code of behaviour and the need to work collaboratively instead of being in competition for financial rewards (Anderfuhren-Biget *et al.*, 2010; Ashraf *et al.*, 2014; Leigh, 2013; OECD, 2009; Taylor and Taylor, 2011).

By their nature, public services generate a multitude of outcomes, and thus their objectives are not always clear cut. Besanko *et al.* (2006) provided the example of the police force, highlighting that explicit contracting for results is not feasible, and if the service was subjected to PRP, there was little evidence that "good outcomes" (such as reduced crime) occurred. In addition, ultimate outcomes in other sectors (such as educational performance, health improvements) will only be visible in the long term, raising questions about the complexity and ambiguity of outcomes and the feasibility of measuring the impact of a PRP scheme.

Alchian and Demsetz (1972) noted that in the public sector, in contrast with the private sector, there is no principal claimant and there may be a wider variety of stakeholders (e.g. individuals using the service, public sector managers, unions, professional bodies, the Government). All could claim an important stake in public services, but their objectives will differ. Additionally, the objectives of these stakeholders can change over time (OECD, 2009; Burgess *et al.*, 2010). Thus PRP schemes in the public sector need to reconcile outcomes that emerge from these multiple stakeholder interests.

Delivery of tasks in the public sector tend to be as a result of collaboration, based on the combined contributions of a large number of individuals, thus an individual's responsibility

for performance outcomes may be difficult to observe and attribute (Frey *et al.*, 2013). Desired outcomes can be influenced by uncontrollable events, or decisions that are taken outside the individual's control (Gibbons, 1998). If PRP schemes are based upon individual efforts, this may be a barrier towards successful implementation in the public sector and individual incentives can mitigate against any necessary team work (Perry *et al.*, 2006).

The aim of this research was to provide an up-to-date (since 2007) comprehensive, independent and credible assessment of relevant academic and other literature on PRP in the public service. This includes identifying features of public services that make the implementation of PRP different to the public sector and identifying PRP schemes that can be more effective for practitioners to implement.

Methodology

The Prentice *et al.*, (2007) report had reviewed the literature regarding PRP in the public sector up to 2007; however, the report only focused on economics-based impact studies of PRP. As the aim of this review was to provide a post-2007 comprehensive, independent and credible assessment of relevant academic and other literature on PRP in the public services, this study also included relevant literature from the fields of human resources, management sociology and psychology, economics and a review of any relevant "grey literature" (literature that is not published in academic media or is in the process of publication).

An initial search of the literature post 2007 was undertaken using the key search terms and databases as presented in Table I. This initial search identified 7,401 documents and articles. The abstracts of all these articles were reviewed using priority search times and the relevance of the article to the review aims. This initial sift reduced the articles to 277 for the second sift which included reviewing the whole paper where content was appraised in more detail, including making judgements on the sectors and occupations covered by the study, the detail of the PRP scheme that was used, the geographical focus and the study's aims. Where studies were impact assessments, the Maryland Scale for evaluations was applied (a five-point scale designed to classify the strength of evidence based on their use of comparison groups). Other studies (reviews and process studies) were appraised both on the rigour of the research methods used and their relevance to the review questions. This resulted in 59 articles to be included in the study, 27 focusing on PRP in the health sector, 16 in education and 16 in the civil service (including both local and central government) (see Table II). The most recent studies discussing PRP were found in the health service sector, whilst robust impact studies in the civil service were particularly sparse. The review found no new evidence of applications of PRP relevant to the armed forces, policing or prisons and justice.

Results

The search, screening and appraisal process highlighted that there is no shortage of published material referring to PRP in the public sector, however; there is relatively little research providing new evidence about the implications of PRP on individual or organisational outcomes, and its cost-effectiveness or economic impact. There is also limited evidence from within the UK, with a high share of the available research material coming from the USA.

The effects of PRP on incentivised outcomes

There is some evidence of positive effects from PRP schemes across the three public services (education, health and the civil service).

Key search terms	Sources
Set A: Populations public sector; public service*; public organisation*; government; non-profit; social services; teach*; civil serv*; doctor*; physician*; nurs*; health*; polic*; judicia*; prison*; military; armed forces; local authorit* NHS; dentist*; general practitioner*, GP*, consultant*, hospital*, school*	*Academic databases* Business Source Premier ProQuest Business Databases Academic Search Complete Web of Science EconLit
Set B: Interventions Performance-related-pay; pay for performance; performance-based pay; performance-based wages; performance pay; performance targets; performance management; performance measure*; merit pay; bonuses; wage*; pay; compensation; reward*; performance incentives; financial incentives; payment by results; contribution; team-based; variable pay; appraisal-related pay; incentive pay	*Grey literature* Centre for Economic Performance (CEP) CIPD Institute for Employment Studies (IES) National Bureau of Economic Research (NBER) National Center for Performance Incentives (NCPI) National Institute of Economic and Social Research (NIESR) OECD Policy Exchange World Bank
Set C: Study type Evidence; evaluat*; empirical; experiment*; impact; effect*; outcome*; assess*; case stud*; value for money; cost effectiveness; implementation; review; experience; study; survey	
Set D: Outcomes Performance; improve*; motivat*; job satisfaction; quality; service delivery; effectiveness; achievement*; productivity; output; efficien*; behaviour; effort; morale; discretion*; input; outcome*; recruitment; retention; turnover; divisive; discriminat*	

Table I.
The literature search terms and databases used

	Impact studies	Review studies	Process studies
Health	13	9	5
Education	10	5	1
Civil service	1	8	7

Table II.
A breakdown of studies included in the review by sector

In education PRP was seen to result in positive implications for students' performance in standardised tests (e.g. Figlio and Kenny, 2007; Podgursky and Springer, 2007; OECD, 2009) which suggests evidence in favour of PRP in improving student test scores. Woessman's (2011) analysis of PISA tests found that students in countries where teacher's salaries were adjusted for performance scored around 25 per cent of a standard deviation higher on maths tests (after controlling for student, school and country measures). Studies also indicated that PRP could have a positive impact on "value-added" measures. For example, Atkinson *et al.* (2009) reported that introducing a PRP scheme for teachers in the UK added on average nearly 90 per cent of a GCSE grade and 40 per cent of a value-added grade for child per eligible teachers. However, several single-programme impact studies noted no significant effects of the use of PRP on pupil examination performance (e.g. Goodman and Turner, 2009; Springer *et al.*, 2012). Therefore, it cannot be concluded in education that PRP interventions will consistently lead to improvements for the incentivised outcomes, and other factors such as scheme design, teacher support and wider accountability and performance measures should be included.

Positive evidence of the implementation of PRP was found in the health sector, most frequently for quality of care measures for chronic diseases and preventative health measures. Hurley *et al.* (2011) found that incentives led to the increase in the provisions of four of the five preventative services targeted in Ontario, Canada. Van Herck *et al.* (2010) concluded in their systematic review that PRP led to a 5 per cent improvement in incentivised physician performance measures, although variations across studies were reported. Although these positive effects were reported, the evidence suggests that any improvements as a result of PRP were small and short lived (So and Wright, 2012; Houle *et al.*, 2012), either as a result of the modest size of the payment scheme (Chung *et al.*, 2010; Greene, 2013), or because performances in certain measures were already high limiting the room for further improvements (ceiling effects) (e.g. Petersen *et al.*, 2013; Van Herck *et al.*, 2010). There has been limited experience of PRP in UK hospitals, however, Sutton *et al.* (2012) focusing on the implementation of the Advancing Quality pilot in the northwest of England demonstrated a significant reduction in mortality rates as a result of the scheme (mortality rates reduced by 1.5 percentage points more than non-participating hospitals).

There are mixed findings regarding the effectiveness of PRP initiatives in the civil service. Burgess *et al.* (2010) evaluated a PRP pilot programme in the UK tax office, finding that the scheme (using team-based targets) had positive implications for both individual and team performance. However, Binderkrantz and Christensen (2012) reported no evidence for the effects of PRP in the Danish Government and improved public sector management. The results from this review indicate that in the civil service any positive effects of PRP schemes may have been influenced by other factors such as intrinsic motivation or team and management structures.

The effects of PRP on staff attitudes, motivation and behaviour

Understanding the link between PRP and job satisfaction attracted much attention in the literature, and is especially important as a result of the significant link between job satisfaction and work motivation. This is of particular interest in the public sector as it has been suggested that the financial incentives may "crowd out" any intrinsic motivations that public sector workers have (Frey *et al.*, 2013).

In education, the evidence of the implications for PRP on teacher's attitudes and behaviours is mixed, but overall can be viewed as more negative than the findings for student test scores. Lavy (2009) identified an increase in the provision of extra instruction after school hours among teachers involved in the PRP scheme in Israel and Belfield and Heywood (2008) noted that co-operative teacher working relationships were consistently associated with PRP. However, Jones (2013) reported that PRP increased the focus on remuneration for the work that teachers did, but negatively effected their intrinsic motivation for teaching. Gius (2013) found that teachers who worked in PRP districts were less likely to believe that teaching was important, showed less enthusiasm for their job and were more likely to leave for employment that was better paid than those who did not work for merit pay areas.

Several studies have also suggested that responses to PRP schemes may differ according to individual characteristics such as gender and prior experience. Leigh (2013) and Jones (2013) report evidence to indicate that male teachers respond more positively to and support PRP schemes than their female counterparts, and Jones (2013) also highlighted that women are more likely to reduce their hours under PRP than men. Evidence also suggests that teachers with more experience display negative reactions to PRP in comparison to early career teachers (Jones, 2013; Leigh, 2013), although it is

unclear whether this results from hostility to changes in the system, or previous negative experiences to PRP.

Taylor and Taylor (2011) analysed data from the 2005 International Social Survey Programme on public service employees from 15 countries, and found that although both wages and public sector intrinsic motivation improved behaviour, many government workers, especially those at supervisory levels were affected more by public sector motivation, a finding that was mirrored by Anderfuhren-Biget *et al.* (2010) in a Swiss study. Forest (2008) in a review of theoretical and empirical studies about PRP in the public sector in the USA, England and France, concluded that individual remuneration practices, can in the long-term lead to negative effects on intrinsic motivations thought to drive public sector employees. However, Stazyk (2013) investigated managers in US local government jurisdictions and found that employees who worked under a variable pay system, rather than the standard compensation system had higher levels of public service motivation, greater role clarity and reported greater job satisfaction. Consequently, there is mixed evidence in the civil service regarding whether financial motivations as a result of PRP schemes crowd out intrinsic motivation.

The evidence available examining the link between PRP and behaviour, attitudes and motivation of those in the health sector is mixed. Petersen *et al.* (2013) found that in the USA any positive effects of individual financial incentives relating to physician practices to blood pressure control rapidly wore off after the incentives were removed, indicating that physicians were responding directly to financial incentives. However, in the UK after quality and outcomes framework incentives had been removed for specific outcome targets, high-performance levels remained (Kontopantelis *et al.*, 2014). A number of studies from the UK have suggested that PRP have resulted in a loss of autonomy for healthcare providers and also undermined their sense of professionalism; for example Gillam *et al.* (2012) indicated that health professionals felt that care was becoming less patient centred and there was a greater emphasis on protocol-driven care. The results of the review also suggested that the motivational effects of financial incentives are contingent on the degree to which the incentives are consistent with their work-related values. Young *et al.* (2012) reported that physician responses to incentives to promote clinical tests and screening to diabetic patients were stronger among those less concerned about whether the incentives represented a threat to their autonomy and for physicians who believed that the programme's goals aligned to their professional goals.

The effects of PRP on work organisation and team relationships

Economic theory focuses its attention on the effects of PRP on individual motivation and effort, however; the HR literature places an emphasis on wider factors in the organisation that can influence the effectiveness of PRP interventions, including workplace relationships and organisational structures.

In the civil service, for example, Burgess *et al.* (2010) discussed the importance of effective managers contributing to positive outcomes of PRP. In their study out of the two teams who participated in the incentive scheme, even though both were engaged, the team who engaged in task relocations and had managers motivating their employees to be more efficient showed an increased performance. Workplace relationships, especially "perceived fairness", were highlighted as important in the civil service literature. Bregn (2013) showed that differential payment for the same task was associated with reduced effort and motivation for employees who were paid less, the implication being that PRP schemes may have detrimental effects on performance if they appear to be unfair in either their design or their implementation.

There is some evidence to suggest that improvements in the quality of care provided in the health sector as a result of PRP initiatives may have occurred as a result of changes to the organisation of work, as opposed to increased individual effort. Kontopantelis *et al.* (2014) noted that following the withdrawal of incentives for a number of activities in the quality and outcomes framework in the NHS, levels of performance were maintained. It was suggested that this resulted from the introduction of quality improvement infrastructures that maintained high standards of performance, for example, the increased use of computers, decision support, clinician prompts and patient reminders and recalls (Gillam *et al.*, 2012). In the USA, Sautter *et al.* (2007) also found that where incentives were seen to lead to improved performance, the resources that hospitals used to invest in quality improvement (e.g. redesigning clinical processes, improved reporting and feedback) and the organisational context were seen as more valuable than payment incentives. However, McDonald *et al.* (2007) found that the quality and outcomes framework also resulted in changes to the distribution of work, with nurses gaining greater responsibility but not receiving performance payments, causing resentment and adverse team relations.

Designing PRP schemes

Evidence suggests that the inconsistencies seen in the effects of PRP schemes in the public sector could occur as a result of their design. The nature and range of the goals and targets included in PRP schemes is an important consideration. For example Eijkenaar *et al.* (2013), Leigh (2013) and Lundstrom (2012) proposed PRP schemes will be more effective when goals are specific and easy to track and measure, with Coleman (2010) noting that outcomes which are more difficult to measure may be more prone to gaming, although simple targets could lead to a misallocation of effort from individuals participating in the schemes. PRP schemes in the education sector have been primarily focused on simple measures such as students' performance in standardised tests, although Yuan *et al.* (2013) questioned whether test scores adequately captured all the important elements of teaching performance, with Fryer (2011) suggesting outcome measures such as pupil attendance, homework completion, etc., could be additional incentivised outcomes. In the healthcare sector, Van Herck *et al.* (2010) found that selecting goals with greater room for improvement resulted in higher effect sizes than when there was less room, and process indicators (e.g. treatment targets) yielded better response than outcome measures (e.g. hospital readmissions) (So and Wright, 2012; Van Herck *et al.*, 2010). However, Gravelle *et al.* (2010) mentioned that when a wide range of targets are introduced, PRP schemes can become too complex to administer and collect reliable measures.

The modest size of payments in many PRP schemes has been offered as an explanation for inconsistent and sometimes limited effectiveness (e.g. Chung *et al.*, 2010; Greene, 2013; Schmidt *et al.*, 2011). Lester *et al.* (2013) reported that the fairly large incentive in the quality and outcomes framework (if staff gained maximum points this could equate to 20 per cent of their salary) was an important factor in its success. The size of the incentive could also have positive implications for the take-up of voluntary PRP schemes (Van Herck *et al.*, 2010). However, other reviews of PRP in healthcare have failed to find consistent relationships between the size of incentives provided and their success (Frølich *et al.*, 2007; Elovainio, 2010). The relationship between payment frequency and the success of PRP schemes has not been established, with Eijkenaar *et al.* (2013) reporting that timely payments lead to improved outcomes, but Chung *et al.* (2010) noting that frequency had no effect on physician responses to PRP schemes.

How measures of performance are constructed also have an important bearing on behavioural responses to PRP schemes. Absolute measures, although easier to report, many reduce the incentive for low-performers as they would struggle to reach targets, however, relative measures will reward those with lower baseline measures as they have more chance for improvement. In the health sector most PRP schemes target absolute measures as they are easier to track and these measures are more likely to report positive effects (Elovainio, 2010; Eijkenaar *et al.*, 2013).

Rank ordering of staff is another performance-based measure (ranking employees according to performance and then awarding bonuses to a set number of them). This method provides control over expenditure (as a set number of staff will receive a payment in any given round) (OECD, 2009), but can also lead to false distributions, where even those who have achieved their targets are in the bottom 10 per cent penalised for poor performance. Similarly, the proportion of staff that are likely to receive a bonus can have an impact on behavioural outcomes. Figlio and Kenny (2007) reviewed PRP schemes in education, finding that schemes which offered a bonus to a large proportion of employees led to fewer positive behavioural outcomes in comparison to those which were more selectively distributed. Although, Schmidt *et al.* (2011) studying PRP schemes in German public services reported that selective schemes were perceived as unfair and reduced employee motivation.

There are also mixed results for behavioural outcomes based on whether PRP schemes were implemented on a voluntary or a mandatory basis. Greene (2013) reported on the national Practices Incentive Programme in Australia, and found that higher performers were more likely to be participants in the scheme, and consequently there was often limited room for improvement. Van Herck *et al.*'s (2010) review found mixed results for behavioural outcomes when schemes were implemented voluntarily. For example, there was evidence suggesting that voluntary schemes do not result in an over representation of high performers, although a US study on Premier Quality Hospital Incentive Demonstrations did find significant differences in behaviours between participants and non-participants in a voluntary scheme.

The level at which employees are involved in the design of PRP schemes and what the outcome measures are can have an impact on their effectiveness. Schmidt *et al.* (2011) found that more participative schemes in German public services achieved a higher level of acceptance among employees and were more effective in motivating employee effort than non-participative systems. Eijkenaar *et al.* (2013), So and Wright (2012) and Van Herck *et al.* (2010) indicated that in healthcare, programmes designed collaboratively and implemented with effective communication of scheme targets and rewards also achieved better results in comparison to non-collaborative PRP schemes.

Key challenges to PRP in the public sector

Findings from the data also identified a number of challenges to the principles underlying PRP in the public sector. Misallocation of effort occurs through a tendency for performance incentives to focus on tasks or outputs which are easily measurable, encouraging employee emphasis on these to the detriment of others (which may be equally significant to task outcomes). Gaming, the maximising of incentive gains without increasing actual performance or minimising effort, may be seen as an extreme form of misallocation of effort. In the education sector there was evidence to suggest that teachers narrowed their approach, focusing their attention to particular tests or on borderline pupils so that they achieve the threshold target (Neal, 2011). This "teaching to the test" approach has been viewed by many as a misallocation of effort. In healthcare, there was

some evidence of reallocation of effort towards incentivised tasks (Gillam *et al.*, 2012) and low-level gaming of the system through exception reporting practices (Gravelle *et al.*, 2010). However, the introduction of PRP schemes also resulted in positive spill overs to non-incentivised aspects of care, making a significant contribution to the overall value of the scheme (Sutton *et al.*, 2012), as a result of promoting more structured team-based care and positive changes to the organisation of the workplace (Gillam *et al.*, 2012).

A characteristic of public sector activity is its reliance on team work and collaboration to achieve positive performance outcomes. The review found evidence that PRP schemes resulted in perceptions of "unfairness" in pay, leading to detrimental effects on employee effort (e.g. Schmidt *et al.*, 2011). Bregn (2013) reported how in the civil service differential pay for the same tasks led to reduced effort from the employee who was paid less as a result of the poor design or implementation of a PRP scheme. In the education sector male teachers responded more positively and supported more strongly PRP systems (Jones, 2013; Leigh, 2013) with evidence from the wider literature suggesting that as women have more pressures on their time outside of work (usually care-related activities), they may be less able or willing to exert effort to gain returns from PRP schemes (Green and Heywood, 2010). However, if the implementation of PRP schemes also results in improved management practices, and task redistribution among teams is perceived as fair and efficient, the team targets (especially in the civil service) can improve (Burgess *et al.*, 2010).

Frey (1997) discussed how the introduction of PRP in the public sector may "crowd out" or impair intrinsic public sector motivations. Evidence of this is suggestive, but inconclusive in the public sector, partly due to difficulties in defining and measuring intrinsic motivation (e.g. Anderfuhren-Biget *et al.*, 2010; Stazyk, 2013). Suggestive evidence includes that of Jones (2013) who reported how teachers displayed a reduction in unpaid co-operative activities in areas where PRP schemes had been introduced, while levels of paid activities remained the same and Weibel *et al.* (2009) who found a negative effect on the completion of more complex tasks in favour for simple tasks in the civil service. However, in the UK health sector the evidence (e.g. McDonald *et al.*, 2007; Lester *et al.*, 2013) suggests that general practitioners thought that the quality and outcomes framework acted as a further incentive to provide what they themselves regard as good practice in clinical care, indicating that the framework did not act as a threat to crowding out intrinsic motivations of healthcare staff.

As well as challenges in the design of PRP schemes in the public sector there are also related difficulties in the measurement of outputs. Different ways of measuring of and rewarding performance can have implications for participant behaviour, for example relative measures that reward improvement are more likely to stimulate improvement among poorer performers. Thus matching measures to overall performance goals sought is important. A final challenge also revolved around the outcomes of PRP and occurs when incentivised outputs do not satisfactorily represent valued or desired outcomes from the service. This is a specific challenge in the public sector because of the complexity of public service "good", and the variety of stakeholders with an interest on service outcomes.

Discussion

The review did find some evidence that PRP schemes can be effective across the three domains of the public sector for which there was evidence available (health, education and the civil service). However, the findings within and between sectors are mixed, with scheme effectiveness often dependent upon the occupational and organisational context. Importantly, the design of the PRP scheme had an important bearing on the

results obtained, although this review is unable to provide with certainty which type of scheme will be best overall, due to the methodological limitations of the evidence base (studies have rarely tested the overall impact of the scheme design, and evaluation designs have tended to inhibit any opportunities to undertake a meta-analysis (Flodgren *et al.*, 2011; Houle *et al.*, 2012, Scott, 2009), and because effectiveness of PRP schemes are also contingent upon the public sector service context and the specific policy aims within these contexts.

The weight of evidence between the sectors studied varied across the services. More robust evidence came from the health and education sector. For example, in education, positive effects (and design) tended to be concentrated on students' performances in standardised tests, whereas in health, process measures (e.g. treatment targets) in primary care and preventative services have shown the most improvement as a result of PRP. However, in all sectors where any positive effects were found, the effect sizes were often small and short lived. This emphasises the need and the value of undertaking longer-term follow-up evaluations of PRP schemes.

Practitioner advice for PRP design options

The results indicated that the design of PRP schemes, especially in the public sector could influence scheme effectiveness and outcomes. However, there are ways in which these challenges can be overcome for the future design and operation of PRP schemes.

When considering the nature and range of goals and targets that PRP the review suggested that a broader range of targets helped to prevent a misallocation of effort, but simple, specific targets were clearer and easier to understand for staff and could enhance scheme effects. Selecting goals where there is greater room for improvement is likely to yield higher rates for improvement due to "ceiling effects", and ways of preventing misallocation of effort and gaming centre around devising schemes with a broader suite of targets that reflect the full range of valued outcomes. Practitioners should combine "hard" measures of performance with "softer" forms of performance management such as appraisals, so that less easily measured goals and outcomes are still captured. Similarly, when addressing whether absolute or relative measures of performance are used, although absolute measures can reduce the incentive for low-level performers, relative measures can reward the lowest providers the most. Thus, a combination of relative and absolute measures may be optimal.

The results indicated that the public sector is characterised by collaborative activity, suggesting that team targets could be more effective than individual goals, however; questions remained regarding how targets set at organisational levels can be conveyed to all staff, and how transparent the measurements were. Evidence across the three public sectors suggests that individual or small group targets may be more effective than institutional targets, and could allow for effective peer monitoring of behaviour. Another concern over collaborative activity in the public sector was the difficulty in attributing outcomes to individual effort which could result in perceived unfairness and detrimental implications for productivity and performance. Potential methods of addressing these concerns include making the targets as transparent as possible and designing PRP schemes in consultation with staff, allowing their input on goals, targets and means of measuring performance. Finally when discussing designing PRP schemes with concerns over team collaboration, although rank order tournaments can control expenditure, false distribution or false ranking can become a disincentive. Thus, practitioners should use this scheme with caution, and if implemented ensure appropriate appraisals and feedback are provided so individuals are able to understand how their performance compared with others.

There was no robust relationship established between size and frequency of payments and PRP scheme effect size, and large payments can raise concerns about the cost-effectiveness of the schemes for organisations. Although reward size may increase participation in voluntary PRP schemes, there are concerns in the public sector that PRP schemes can crowd-out or impair intrinsic motivations. Methods to address this may include designing collaborative PRP systems to ensure that targets align with employee professional goals, to consider other non-financial forms of quality improvement, and use other methods of motivating performance, e.g. strengthening professionalism.

Performance measurement problems were also identified as challenges to PRP in the public sector particularly associated with the negative effects of particular performance measures and managerial subjectivity in assessment. Potential methods of addressing these issues include involving employees in discussions about the most appropriate metrics and performance measures, to clarify objectives and engage employees more directly with the goals of the organisation and to ensure that targets align with the overall improvement goals and combining measures where appropriate (e.g. linear measurements above a minimum threshold, or a combination of attainment and improvement). Ways to minimise situations where PRP outcomes do not represent the desired outcomes from the service include having a wide range of outcome measures that adequately capture the desired outcomes of the service, undertaking longer-term monitoring and evaluation of public sector PRP schemes to ensure that outcomes align with organisational and policy goals, to conduct cost-effectiveness analyses to assess whether PRP schemes are actually cost-effective in meeting the desired outcomes, and to compare different improvement strategies and means of motivating performance to ensure that there is value for money.

Limitations and opportunities for future research

Although a comprehensive review of the literature was undertaken, it is clear from the results that further pilots and evaluations with robust, ideally experimental methods need to be undertaken in this area, which would be beneficial to providing further guidance on aspects of scheme design within specific public policy context. Similar to an observation made by Prentice *et al.* (2007) evidence from the UK public sector is limited, and many conclusions are made from international studies. Consequently, care must be taken when generalising results across countries, especially as samples may not be representative of the UK population, and incentive schemes should reflect the environment in which they operate. Consequently, there is still a need to understand more fully the nature of PRP schemes and their effectiveness in UK public services, prioritising policies and practices used in the UK. A number of UK public sector services are still unrepresented, especially the armed forces and the prison and justice sectors, thus more research regarding the effects of PRP schemes in these areas is needed. Finally, there would be merit in undertaking longer follow-up evaluations of PRP schemes to measure the longevity of PRP schemes in the public sector.

References

Adams, J.S. (1963), "Toward an understanding of inequity", *Journal of Abnormal and Social Psychology*, Vol. 67 No. 5, pp. 422-463.

Alchian, A.A. and Demsetz, H. (1972), "Production, information costs and economic organization", *American Economic Review*, Vol. 62 No. 5, pp. 777-795.

Anderfuhren-Biget, S., Varone, F., Giauque, D. and Ritz, A. (2010), "Motivating employees of the public sector: does public service motivation matter?", *International Public Management Journal*, Vol. 13 No. 3, pp. 213-246.

Ashraf, N., Bandiera, O. and Jack, B. (2014), "No margin, no mission? A field experiment on Incentives for public service delivery", *Journal of Public Economics*, Vol. 123, pp. 1-17, available at: www.sciencedirect.com/science/article/pii/S0047272714001546 (accessed March 2015).

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H. and Wilson, D. (2009), "Evaluating the impact of performance-related pay for teachers in England", *Labour Economics*, Vol. 16 No. 3, pp. 251-261.

Belfield, C. and Heywood, J. (2008), "Performance pay for teachers: determinants and consequences", *Economics of Education Review*, Vol. 27 No. 3, pp. 243-252.

Besanko, D., Dranove, D., Shanley, M. and Schaefer, S. (2006), *Economics of Strategy*, 4th ed., Wiley, Chichester and Hoboken, NJ.

Bevan, S.M. and Thompson, M. (1992), "Merit pay, performance appraisal and attitudes to women's work", Report No. 234, Institute of Manpower Studies, Brighton.

Binderkrantz, A. and Christensen, J. (2012), "Agency performance and executive pay in government: an empirical test", *Journal of Public Administration Research & Theory*, Vol. 22 No. 1, pp. 31-54.

Bregn, K. (2013), "Detrimental effects of performance-related pay in the public sector? On the need for a broader theoretical perspective", *Public Organization Review*, Vol. 13 No. 1, pp. 21-35.

Brewer, G.A. and Selden, S.C. (1998), "Whistle blowers in the federal civil service: new evidence of the public service ethic", *Journal of Public Administration Research and Theory*, Vol. 8 No. 3, pp. 413-440.

Burgess, S., Propper, C., Ratto M., Von Hinke Kessler Scholder, S. and Tominey, E. (2010), "Smarter task assignment or greater effort: the impact of incentives on team performance", *Economic Journal*, Vol. 120 No. 547, pp. 968-989.

Chung, S., Palaniappan, L., Wong, E., Rubin, H. and Luft, H. (2010), "Does the frequency of pay-for-performance payment matter? – Experience from a randomized trial", *Health Services Research*, Vol. 45 No. 2, pp. 553-564.

Coleman, T. (2010), "Do financial incentives for delivering health promotion counselling work? Analysis of smoking cessation activities stimulated by the quality and outcomes framework", *BMC Public Health*, Vol. 10, pp. 167-172, available at: www.biomedcentral.com/1471-2458/10/167

Eijkenaar, F., Emmert, M., Scheppach, M. and Schöffski, O. (2013), "Effects of pay for performance in health care: a systematic review of systematic reviews", *Health Policy*, Vol. 110 Nos 2/3, pp. 115-130.

Elovainio, R. (2010), "Performance incentives for health in high-income countries: key issues and lessons learned", Background Paper No. 32, World Health Report, Geneva.

Figlio, D. and Kenny, L. (2007), "Individual teacher incentives and student performance", *Journal of Public Economics*, Vol. 91 Nos 5/6, pp. 901-914.

Flodgren, G., Eccles, M., Shepperd, S., Scott, A., Parmelli, E. and Beyer, F. (2011), "An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes", *Cochrane Database of Systematic Reviews*, No. 7, doi:10.1002/14651858.cd009255, available at: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009255/pdf

Forest, V. (2008), "Performance-related pay and work motivation: theoretical and empirical perspectives for the French civil service", *International Review of Administrative Sciences*, Vol. 74 No. 2, pp. 325-339.

Frey, B. (1997), *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar, Cheltenham and Brookfield, WI.

Frey, B., Homberg, S. and Osterloh, M. (2013), "Organizational control systems and pay-for-performance in the public service", *Organization Studies*, Vol. 34 No. 7, pp. 949-972.

Frølich, A., Talavera, J., Broadhead, P. and Dudley, R. (2007), "A behavioral model of clinician responses to incentives to improve quality", *Health Policy*, Vol. 80 No. 1, pp. 179-193.

Fryer, R. (2011), "Financial incentives and student achievement: evidence from randomised trials", *The Quarterly Journal of Economics*, Vol. 126 No. 4, pp. 1755-1798.

Gibbons, R. (1998), "Incentives in organizations", *Journal of Economic Perspectives*, Vol. 12 No. 4, pp. 115-132.

Gillam, S., Siriwardena, A. and Steel, N. (2012), "Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework – a systematic review", *Annals of Family Medicine*, Vol. 10 No. 5, pp. 461-468.

Gius, M. (2013), "The effects of merit pay on teacher job satisfaction", *Applied Economics*, Vol. 45 No. 31, pp. 4443-4451.

Goodman, S. and Turner, L. (2009), "Group incentives for teachers: the impact of the NYC school-wide bonus program on educational outcomes, Discussion Paper No. 0910-05, Department of Economics, Columbia University, New York.

Gravelle, H., Sutton, M. and Ma, A. (2010), "Doctor behaviour under a pay for performance contract: treating, cheating and case finding?", *Economic Journal*, Vol. 120 No. 542, pp. F129-F156.

Green, C. and Heywood, J. (2010), "Profit sharing and the quality of relations with the boss", *Labour Economics*, Vol. 17, pp. 859-867.

Greene, J. (2013), "An examination of pay-for-performance in general practice in Australia", *Health Services Research*, Vol. 48 No. 4, pp. 1415-1432.

Heath, C. (1999), "On the social psychology of agency relationships: lay theories of motivation overemphasize extrinsic incentives", *Organizational Behavior and Human Decision Processes*, Vol. 78 No. 1, pp. 25-62.

Houle, S., Mcalister, F., Jackevicius, C., Chuck, A. and Tsuyuki, R. (2012), "Does performance-based remuneration for individual health care practitioners affect patient care?", *Annals of Internal Medicine*, Vol. 157 No. 12, pp. 889-899.

Hurley, J., Decicca, P., Li, J. and Buckley, G. (2011), "The response of Ontario Primary Care physicians to pay-for-performance incentives", Working Paper Series No. 2011-02, Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, available at: http://chepa.org/docs/working-papers/11-02.pdf (accessed March 2015).

Jones, M. (2013), "Teacher behavior under performance pay incentives", *Economics of Education Review*, Vol. 37 No. C, pp. 148-164.

Kessler, I. and Purcell, J. (1992), "Performance-related pay: objectives and application", *Human Resource Management Journal*, Vol. 2 No. 3, pp. 16-33.

Kontopantelis, E., Springate, D., Reeves, D., Ashcroft, D., Valderas, J. and Doran, T. (2014), "Withdrawing performance indicators: retrospective analysis of general practice performance under UK quality and outcomes framework", *British Medical Journal*, Vol. 348 No. 330, pp. 1-11.

Larkin, I., Pierce, L. and Gino, F. (2012), "The psychological costs of pay-for-performance: implications for the strategic compensation of employees", *Strategic Management Journal*, Vol. 33 No. 10, pp. 1194-1214.

Lavy, V. (2009), "Performance pay and teachers' effort, productivity, and grading ethics", *American Economic Review*, Vol. 99 No. 5, pp. 1979-2011.

Lazear, E.P. (2000), "Performance pay and productivity", *American Economic Review*, Vol. 90 No. 5, pp. 1346-1361.

Leigh, A. (2013), "The economics and politics of teacher merit pay", *CESifo Economic Studies*, Vol. 59 No. 1, pp. 1-33.

Lester, H., Matharu, T., Mohammed, M., Lester, D. and Foskett-Tharby, R. (2013), "Implementation of pay for performance in primary care: a qualitative study 8 years after introduction", *British Journal of General Practice*, Vol. 63 No. 611, pp. E408-E415.

Lindenberg, S. (2001), "Intrinsic motivation in a new light", *Kyklos*, Vol. 54 Nos 2/3, pp. 317-342.

Lundstrom, U. (2012), "Teachers' perceptions of individual performance-related pay in practice: a picture of a counterproductive pay system", *Educational Management Administration & Leadership*, Vol. 40 No. 3, pp. 376-391.

McDonald, R., Harrison, S., Checkland, K., Campbell, S. and Roland, M. (2007), "Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study", *British Medical Journal (International Edition)*, Vol. 334 No. 7608, pp. 1357-1359.

Makinson, J. (Chair) (2000), *Incentives for Change: Rewarding Performance in National Government Networks*, Public Services Productivity Panel, HM Treasury, London.

Marsden, D. (2009), "The paradox of performance related pay systems: why do we keep adopting them in the face of evidence that they fail to motivate? Discussion Paper No. 946, Centre for Economic Performance (CEP), available at: http://cep.lse.ac.uk/pubs/download/dp0946.pdf (accessed March 2015).

Marsden, D. and French, S. (1998), "What a performance: performance-related pay in the public services", Special Paper No. CEPSP10, Centre for Economic Performance, London School of Economics and Political Science, London.

Neal, D. (2011), "The design of performance pay in education", Working Paper No. 16710, NEBR, Cambridge, MA.

OECD (2009), *Evaluating and Rewarding the Quality of Teachers: International Practices*, OECD, Paris.

Perry, J., Engbers, T. and Yun Jun, S. (2009), "Back to the future? Performance-related pay, empirical research and the perils of persistence", *Public Administration Review*, Vol. 69 No. 1, pp. 39-51.

Perry, J.L., Mesch, D. and Paarlberg, L. (2006), "Motivating employees in an new governance era: the performance paradigm revisited", *Public Administration Review*, Vol. 66 No. 4, pp. 505-514.

Petersen, L., Simpson, K., Pietz, K., Urech, T., Hysong, S., Profit, J., Conrad, D., Dudley, R. and Woodard, L. (2013), "Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial", *Journal of the American Medical Association*, Vol. 310 No. 10, pp. 1042-1050.

Podgursky, M. and Springer, M. (2007), "Teacher performance pay: a review", *Journal of Policy Analysis and Management*, Vol. 26 No. 4, pp. 909-949.

Porter, L.W. and Lawler, E.E. (1968), *Managerial Attitudes and Performance*, R.D.Irwin, Homewood, IL.

Prendergast, C. (1999), "The provision of incentives in firms", *Journal of Economic Literature*, Vol. 37 No. 1, pp. 7-63.

Prentice, G., Burgess, S. and Propper, C. (2007), *Performance-Pay in the Public Sector: A Review of the Issues and Evidence*, Office of Manpower Economics, London, available at: http://webarchive.nationalarchives.gov.uk/20130106083031/http://www.ome.uk.com/downloads/Performance%20pay%20in%20the%20Public%20Sector.%20A%20review%20of%20the%20issues%20and%20evidence.pdf (accessed March 2015).

Reilly, P. (2003), *The Link between Pay and Performance*, Institute for Employment Studies, Brighton.

Sautter, K., Bokhour, B., White, B., Young, G. and Burgess, J. (2007), "The early experience of a hospital-based pay-for-performance program", *Journal of Healthcare Management*, Vol. 52 No. 2, pp. 95-107.

Schmidt, W., Trittel, N. and Muller, A. (2011), "Performance-related pay in German public services", *Employee Relations*, Vol. 33 No. 2, pp. 140-158.

Scott, I. (2009), "What are the most effective strategies for improving quality and safety of health care?", *Internal Medicine Journal*, Vol. 39 No. 6, pp. 389-400.

So, J. and Wright, J. (2012), "The use of three strategies to improve quality of care at a national level", *Clinical Orthopaedics & Related Research*, Vol. 470 No. 4, pp. 1006-1016.

Springer, M., Hamilton L. and Mccaffrey, D. (2012), "Final Report: Experimental Evidence from the Project on Incentives in Teaching (POINT)", National Center on Performance Incentives, Vanderbilt University, Nashville, TN.

Stazyk, E. (2013), "Crowding out public service motivation? Comparing theoretical expectations with empirical findings on the influence of performance-related pay", *Review of Public Personnel Administration*, Vol. 33 No. 3, pp. 252-274.

Sutton, M., Nikolova, S., Boaden, R., Lester, H., McDonald, R. and Roland, M. (2012), "Reduced mortality with hospital pay for performance in England", *New England Journal of Medicine*, Vol. 367 No. 19, pp. 1821-1828.

Taylor, J. and Taylor, R. (2011), "Working hard for more money or working hard to make a difference? Efficiency wages, public service motivation and effort", *Review of Public Personnel Administration*, Vol. 31 No. 1, pp. 67-86.

Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. and Sermeus, W. (2010), "Systematic review: effects, design choices, and context of pay-for-performance in health care", *BMC Health Services Research*, Vol. 10, pp. 247-259, available at: www.biomedcentral.com/1472-6963/10/247

Vroom, V.H. (1964), *Work and Motivation*, Wiley, New York, NY.

Weibel, A., Rost, K. and Osterloh, M. (2009), "Pay for performance in the public sector – benefits and (hidden) costs", *Journal of Public Administration Research and Theory*, Vol. 20 No. 2, pp. 387-412.

Woessman, L. (2011), "Cross-country evidence on teacher performance pay", *Economics of Education Review*, Vol. 30 No. 3, pp. 404-418.

Young, G., Beckman, H. and Baker, E. (2012), "Financial incentives, professional values and performance: a study of pay-for-performance in a professional organization", *Journal of Organizational Behavior*, Vol. 33 No. 7, pp. 964-983.

Yuan, K., Le, V., Mccaffrey, D., Marsh, J., Hamilton, L., Stecher, B. and Springer, M. (2013), "Incentive pay programs do not affect teacher motivation or reported practices: results from three randomized studies", *Educational Evaluation and Policy Analysis*, Vol. 35 No. 1, pp. 3-22.

Corresponding author
Dr Zofia M. Bajorek is the corresponding author and can be contacted at: zbajorek@theworkfoundation.com