

Performance-Based Rewards for Teachers: A Literature Review

Owen Harvey-Beavis*

For distribution at the 3rd Workshop of Participating Countries on OECD's Activity *Attracting, Developing and Retaining Effective Teachers*
4-5 June 2003, Athens, Greece

February 2003

* This paper is provided for the information of participants. It was prepared by Mr Owen Harvey-Beavis, a graduate student at the University of Melbourne, Australia, during an internship at the Education and Training Policy Division, Directorate for Education, OECD for the period January-February 2003.

The opinions expressed in this paper are the sole responsibility of the authors and do not necessarily reflect those of the OECD or the governments of its Member countries.

TABLE OF CONTENTS

1	INTRODUCTION	3
2	TYPES OF PERFORMANCE-BASED REWARD SYSTEMS.....	3
3	ARGUMENTS SUPPORTING PERFORMANCE-BASED REWARDS	6
4	ARGUMENTS OPPOSING PERFORMANCE-BASED REWARDS.....	11
5	DIFFICULTIES IN IMPLEMENTATION	15
6	EVIDENCE OF THE EFFECTS ON EDUCATIONAL OUTCOMES.....	19
7	CONCLUSION	24
	REFERENCES	26
	ANNEX 1: SYSTEMS OF PERFORMANCE-BASED PAY IN DIFFERENT COUNTRIES	30
	ANNEX 2: FEATURES OF PERFORMANCE-BASED SYSTEMS.	36
	ANNEX 3: THE METHODOLOGY AND CONCLUSIONS OF THE KENTUCKY AND CHARLOTTE-MECKLENBURG STUDIES OF TEACHER ATTITUDES (1998-2002)	37
	ANNEX 4: BALLOU AND PODGURSKY'S STUDY OF TEACHER ATTITUDE (1993)	39
	ANNEX 5: THE DALLAS SCHOOL ACCOUNTABILITY AND INCENTIVE PROGRAMME (CLOTFELTER AND LADD, 1996; LADD, 1999).....	41

1 INTRODUCTION

1. This paper examines the current academic and policy literature surrounding performance-based reward programmes for primary and secondary school teachers. It is a working paper for the activity “Attracting, Developing and Retaining Effective Teachers” being conducted by the Organisation for Economic Co-operation and Development (OECD).

2. Performance-based rewards have a long history in education, particularly in the United States. In the last ten years, a number of countries have adopted pay-for-performance strategies to modify the traditional salary scales (Refer to Annex 1). The distinguishing feature of a performance-based scheme is that it rewards or sanctions teachers based upon some form of performance evaluation (Chamberlin, et al, 2002). Distinctions in performance-based reward programmes are found in the skills assessed and the rewards provided. Most individually-based programmes have used pecuniary rewards for high levels of performance, usually defined in terms of student outcomes or teacher skills and knowledge. More recently, some analysts have proposed that intrinsic rewards, such as seeing students improve in performance, and increased feelings of well-being are better motivators of teachers. Other rewards include increased holiday time and professional development courses.

3. Many of the earlier programmes tended to focus on individual performance, in particular merit pay (Richardson, 1999), with recent debates more likely to consider group-based reward programmes, or knowledge and skill based rewards (Odden, 2000a; Odden and Kelley, 2002).

4. One limitation of this paper should be noted: the academic and policy literature examined is exclusively from English-speaking countries, and especially the United States. All reports considering the evidence of the effectiveness of performance-based rewards examine United States programmes.

5. This paper has five further sections. Section two examines different types of performance-based reward programmes. Section three examines the arguments that are put forward in their support. Section four examines the arguments that reject them. Section five considers the reasons why performance-based reward programmes are apparently difficult to implement. And section six summarises current evidence on the effects of performance-based reward programmes.

2 TYPES OF PERFORMANCE-BASED REWARD SYSTEMS

6. This section summarises the types of performance-based reward systems that are found in the literature and the education systems of OECD countries.

7. One of the prominent features of performance-based reward systems is the number of different varieties of programmes. It is necessary to be aware of the particular characteristics of the schemes when examining the research and policy literature, as distinctions between specific models are often made (see, for example, Odden and Kelley, 2002). Some key dimensions along which performance-based rewards differ include the following:

- Whether the programme focuses on individual teacher performance or school-based performance;
- Whether the compensation is pecuniary or non-pecuniary, and whether sanctions exist for poor performance;
- The duration of the reward, and in particular, whether the reward is given once only, for a limited duration, or permanently;
- The reward levels, and in particular, whether there are ascending rewards for increased teacher or school performance, or whether the performance evaluation allows teachers to progress to a new salary scale;
- What is evaluated, and in particular, whether the evaluation is made on the basis of observation, a portfolio, acquired qualifications or student performance;
- Who evaluates the teacher, and in particular, whether the evaluation is completed by the principal of the school, an external review, or peer review;
- The scope of the reward, and in particular, whether all teachers who fulfil criteria are rewarded, or just a specific quota; and
- Whether the performance-based rewards supplement or replace the existing salary scale system.

8. Annex 2 contains a chart that outlines the different dimensions of performance-based rewards in the form of a diagram.

9. Despite this great variety, there are three main models of performance-based reward programmes that are commonly examined in the literature and are found in education systems. The first model is ‘merit-pay’, which generally involves individual pecuniary awards based on student performance, and classroom observation (McCollum, 2001). The second model is ‘knowledge and skill-based’ compensation, which generally involves individual pecuniary rewards for acquired qualifications and demonstrated knowledge and skills, which are believed to increase student performance (Odden, 2000b). Knowledge and skill-based pay differs from merit-pay because it provides clear guidelines on what is being evaluated (Odden and Kelley, 2002). The knowledge and skills evaluated are, it is argued, linked to teacher proficiency, meaning knowledge and skill-based pay increases teachers’ ability (Odden and Kelley, 2002). The third model is school-based compensation, which generally involves group-based pecuniary rewards, typically based on student performance (Odden and Kelley, 2002). Table 1 elaborates the distinctions between these models.

Table 1. Models of Performance-Based Reward

Characteristic	Knowledge and Skill-Based programmes	Merit-Pay programmes	School-Based programmes
----------------	--------------------------------------	----------------------	-------------------------

Recipient of compensation	Individual teachers	Individual teachers	Schools, who often have discretion to distribute rewards to staff.
Scope of compensation	All teachers who can exhibit skills and knowledge are rewarded.	Mixed, with some programmes providing universal rewards, while others are limited by quotas.	Mixed, with some programmes providing universal rewards, while others are limited by quotas.
Type of compensation	Primarily financial. Some analysts argue that intrinsic rewards, such as enjoyment from increasing student performance, form part of this system.	Typically financial.	Primarily financial. Some analysts argue that intrinsic rewards, such as enjoyment from increasing student performance, form part of this system.
Areas evaluated	Assessed on the basis of demonstrated skills and knowledge which are thought to be linked to increased teacher performance. Often comes in the form of additional qualifications.	A range of areas are assessed, including: a portfolio of teacher accomplishments, class room observation and student performance.	Typically student performance is used to evaluate schools. This can occur through 'added value' gains to student scores, or absolute student achievement.
Who evaluates	Typically, will be completed by external review.	A range of evaluators, including peer review, the principal and external review.	Typically, will be completed by an external review.
Duration of compensation	Typically short-term, generally requiring periodic evidence that teachers maintain skills and knowledge.	The duration of the reward is varied, but most often annual.	Typically an annual bonus.
Relationship to existing compensation system.	Typically would replace the salary scale, in full or in part.	Typically supplements, but can replace the salary scale.	Typically supplements the salary scale
Compensation level	Typically ascending levels of rewards tied to increasing levels of skills and knowledge acquired.	Mixed, depending on the relationship to existing compensation systems. If the salary scale is used, there is a greater propensity to have a single compensation	Typically one level of reward.

level. If it is not used, there is a greater inclination to have multiple rewards.

3 ARGUMENTS SUPPORTING PERFORMANCE-BASED REWARDS

10. This section describes the arguments made by a range of analysts who are concerned to support the view that net benefits can be expected from using performance-based rewards programmes for teachers. The arguments in support of, and in opposition to, performance-based rewards are summarised in Table 2.

Table 2. A Summary of the Arguments in the Literature

Arguments In Support of Performance-Based Rewards	Arguments Against Performance-Based Rewards
The current system is unfair and rewards experience and formal qualifications instead of performance.	Fair and accurate evaluation is difficult because performance cannot be determined objectively.
Performance-based rewards improve the governance of schools by increasing the efficiency of resource allocation.	School administration becomes hierarchical and co-operation between school management and staff is strained.
Performance-based pay motivates teachers to perform at their best.	Performance-based financial incentives do not provide incentives for teachers to improve.
There is increased collegiality between teachers and administration.	There is reduced co-operation between teachers.
Student performance is increased, and teacher quality improves.	A range of perverse outcomes occur because of teacher 'game playing' and a narrowing of the curriculum.
The market provides the best approach for efficient allocation of resources, and this model can be applied to teaching.	The market has no place in education.
Relative to other education reforms, performance-based rewards provide a relatively cost-efficient solution.	To implement a performance-based scheme is expensive and time consuming.

Performance-based schemes increase political and public support of education systems.

3.1 The current system is unfair and rewards experience and formal qualifications instead of performance

11. Under most current systems of a salary scale, teachers are rewarded for the number of years spent teaching and the number of tertiary degrees, rather than their performance (Odden, 2000a). For this reason, many analysts believe the salary scale system determines teacher compensation on incomplete criteria. For example, Hoerr (1998) argues that any non-merit-based system is unfair for exceptional teachers because they are judged on inefficient criteria. This will cause, it is argued, talented teachers to leave the education system because excellence is not fairly rewarded (Odden, 2001). Only when performance is rewarded and teachers command salaries equal to the private sector without having to progress up an arbitrary salary scale, will the best talent be attracted and retained (Solomon and Podgursky, 2001).

12. Proponents point out that research has found no consistent links between education credits or degrees and student performance, and only modest links between experience and student performance (Heneman and Milanowski 1999; Hoerr, 1998; Tomlinson, 2000). The existing salary scales are thus at best only loosely related to the expertise and skills needed in the classroom (Mohrman, Mohrman and Odden, 1996). If the pay structure is based on this formula, it inevitably produces unsatisfactory outcomes as it is not well aligned to education output (Odden, 2000a). Thus, a substantial body of literature argues performance-based reward systems are an improvement on the efficiency of salary scales.

3.2 Performance-based rewards improve the governance of schools

13. Several analysts have argued that performance-based pay schemes improve the administration of schools. Under a performance-based pay scheme, principals must know the quality of teachers in all classrooms (Hoerr, 1998). This type of evaluation, it is argued, means principals must summatively evaluate teachers, rather than formatively evaluate, and so more objective decisions about teacher quality are made. Research showing that in performance-based systems, many principals report they evaluated teachers more harshly than they would have in a non-performance-based system (Murnane and Cohen 1986, 9) is used to support this argument. As a safety precaution, Solomon and Podgursky (2001) advocate principals becoming recipients of school wide performance-based rewards, to ensure they remain objective in their evaluation.

14. It is also argued that a movement to school-based rewards can increase the precision of resource allocation by encouraging resource alignment from top down, by setting organisational goals, and from the bottom up, as teachers are gaining feedback, and benefiting from better resource allocation and policy coherence (Kelley, 1999). This can occur because school goals are clarified in a performance-based reward system, and teachers have an increased incentive to share information with administrators since they benefit from improved outcomes.

3.3 Provides motivation to teachers

15. One of the largest benefits reported by proponents of performance-based rewards is an increase in the motivation of teachers. It is argued that performance-based pay will increase teacher motivation by adequately rewarding productivity gains. This perspective links the attitude of teachers to student

outcomes, by arguing that once the motivation and skill of the teacher determine salaries, teacher quality will be improved. Within the literature, Tomlinson (2000) argues that performance-based pay is about motivating people, and developing performance-oriented cultures. Teachers who are not motivated by financial rewards, can be encouraged with non-financial rewards (Odden, 2000a). These rewards can include, for example: satisfaction from high student achievement, recognition, influence, learning new skills, and personal growth (Tomlinson, 2000; Odden 2000b). As Odden and Kelley (2002; Kelley, 1999) argue school-based rewards are a means of providing motivation by introducing clear goals to the whole school, and facilitating student achievement.

16. While it is argued that teachers are not motivated by money (see, for example, Firestone and Pennell, 1993), financial reward must have some influence on career choices for at least some teachers (Richardson, 1999). Some point out that past research suggests money has an influence on teachers' motivation (Refer to Annex 3), and others argue money is one motivator among many (Odden and Kelley, 2002). Hence, it is argued a performance-based policy which involves a monetary component would attract teaching talent by providing rewards that motivate a large range of people. A further benefit may occur through a rise in the socio-economic status of teachers, which should also attract and motivate talent (Solomon and Podgursky, 2001). However, for this to be feasible, more revenue would be required for teacher salaries.

17. Solomon and Podgursky (2001) argue that when teaching is rewarded based on outcomes, quality teachers can be moved to areas of low socio-economic status since these areas can be specifically rewarded. Different criteria can be used to determine rewards for different areas based on the socio-economic, racial and gender demographics of the student population.

3.4 Increased collegiality

18. Earlier merit-pay models were criticised for adversely affecting collaboration between teachers (see, for example, the American Federation of Teachers (AFT), 2001). In response, a large body of literature argues that performance-based reward systems can increase collegiality by rewarding co-operation between teachers (Solomon and Podgursky, 2001; Cohn, 1996), especially through administering group-based pay (Mohrman, Mohrman, and Odden, 1996; McCollum, 2001). This kind of management technique can redesign the work of teachers so they are interdependent, and acknowledge their interdependence (Mohrman, Mohrman and Odden, 1996). Even some opponents of performance-based rewards argue there is some evidence of increased collegiality when group performance rewards are employed (See, for example, Firestone and Pennell, 1993).

3.5 Student outcomes improve

19. According to a range of analysts, the most fundamental goal of performance-based rewards is to increase student performance. For example, Odden (2000b) argues there is a causal link between the quality of teaching and the level of student outcomes, meaning any method that increases the quality of teachers should improve student outcomes. By introducing objective standards which can be used to determine whether teachers have skills to increase the performance of students, the quality of teachers would be established, and also improved (Mohrman, Mohrman and Odden, 1996). Some argue this occurs when evaluation focuses on the knowledge and skills of teachers, which provides an incentive for all teachers to improve, and also an intrinsic reward through professional development (Solomon and Podgursky, 2001). Moreover, performance-based pay can target educators to key objectives and important subjects as a means of increasing student performance (Mohrman, Mohrman and Odden, 1996; Odden,

2001). Proponents argue that teachers may actually gain freedom to innovate, since they no longer have to focus on process, but rather student outcomes (Solomon and Podgursky, 2001).

20. Furthermore, it is argued there will be a greater consistency in teaching standards across school jurisdiction since the best teachers would not be grouped in the highest achieving, lowest disadvantaged and racially homogenous areas (Tomlinson, 2000). This would occur when objective performance rewards create a market where movement between schools would become easy, and the true value of teachers is established. Teachers would not be locked into a district based on their seniority and qualifications, but would have adequate opportunity to move to jurisdictions where their talent is most highly valued (Solomon and Podgursky, 2001). Conversely, poorly performing teachers would be sanctioned by the market, and command a reduced wage. If retention of teachers is affected by the opportunity cost of staying in the profession, this policy would attract the most capable teachers and discourage the least capable teachers.

21. Under a policy of performance-based rewards, the 'best' possible graduates can be recruited by guaranteeing a competitive market based salary. This would give teachers the capability to move beyond the starting salary and be paid at a comparable level to the private sector workforces (Mohrman, Mohrman, and, Odden, 1996; Odden and Kelley, 2002).

3.6 Increased political and public support of education systems

22. A theme in the literature is that performance-related pay increases the support of education by politicians and the public (Solomon and Podgursky, 2001). Reportedly, the public feels that current teacher compensation rewards mediocrity (Tomlinson, 2000). Therefore, it is argued, by providing performance-based rewards, political support of the education system can be generated. Odden (2000b) outlines a plan that successfully garnered educator, union and policymaker support, in Vaughn Next Century Learning Centre in Los Angeles, as evidence these groups can come to a consensus on the implementation and design of these programmes.

3.7 A financial investment

23. Some analysts have argued that the introduction of performance-based rewards can be revenue neutral as the existing salary schedules, which reward seniority and academic qualifications can be flattened, and the revenue gained from this reform can be targeted at rewarding teacher performance (Solomon and Podgursky, 2001). However, this appears to be inconsistent with these authors' previous advocacy for a system of increased teacher salaries. Previous programmes that attempted to provide revenue-neutral performance-based systems have been unsuccessful due to a lack of funds and teacher opposition. In contrast, Mohrman, Mohrman and Odden (1996) argue the private sector model shows that costs can be kept down because the workforce becomes flexible and versatile, in particular teachers will need to have and use a range of pedagogical techniques, which suggests the revenue required to implement this strategy would be relatively low. However, the private sector model may have limited relevance to the public sector, as resources are finite, and schools do not generate additional financial resources with increased productivity (Milanowski, 2003). One possibility is for average class size to increase, which allows teachers to be paid more, without increases in education funding.

3.8 The market provides the best approach for efficient allocation of resources

24. The intellectual foundations of performance-based rewards are found in private sector models. Because the private sector requires productive workers to compete against other agencies, they have

developed policies that seek to maximise output from a set input, or minimise input for a set output. Advances in efficiency, it is argued, can be made in the public sector by observing and adapting private sector worker motivational techniques (Odden and Kelley, 2002). Large firms with complex organisational structures that change their workplace practices to increase productivity and quality can be used as a model. Proponents argue these organisations provide a benchmark for teaching because they have very similar environments to schools, and often use performance-based methods of remuneration (Mohrman, Mohrman and Odden, 1996; Odden, 2000a; Ballou and Podgursky, 2001). Any advances in reward strategies for knowledge and skill-based pay in the private sectors thus provide a blueprint for educational salary schedules (Odden, 2000a). Models are also evident in the government and non-profit organisations, such as the higher education model, which suggests performance-based reward programmes are not mutually exclusive with the public sector (Solomon and Podgursky, 2001).

25. With the introduction of new evaluation systems, such as knowledge and skill-based pay, evaluation of person-based human resources systems can occur. Significant educational bodies including the National Commission on Teaching (U.S.) are accepting this method, and the benefit from using benchmarks, it is argued, is an improved education system (Bainbridge, 2000). This is not to suggest that competency models are inevitably going to work, as these programmes need to be carefully organised to ensure that the goals, culture and political realities of the organisation align (Heneman and Ledford, 1998). This is particularly important, because ‘recalcitrant’ teachers who believe the evaluation process is unfair (Murnane and Cohen, 1986) can undermine the adoption of private sector models.

26. Ballou (2001) argues that if teaching were special, it would not be expected to find performance-based reward systems operating in private schools. Since private schools exhibit a much greater frequency of performance-based rewards, and have much greater bonuses when they do use these schemes, it appears education should not be separated from market logic (Ballou, 2001). While private schools still do not use these techniques all the time, suggesting there are some costs associated with implementing performance-based programmes, it shows teaching is not inherently unsuited to evaluative systems of remuneration (Ballou, 2001).

27. In summary the main arguments in favour of performance-based rewards are:

- The current system is unfair and rewards experience instead of performance;
- School administration would improve, especially when school-based compensation programmes are implemented;
- Teacher motivation would improve, with an emphasis on knowledge and skill and school-based reward models in the literature;
- Teacher co-operation would improve, which is presented as an argument in support of school-based reward programmes. There is some concern about the effect merit-pay systems have on teacher co-operation;
- Student outcomes would improve;
- Political and public support of the education system would improve, which is presented as an argument specifically in support of merit-pay, but can be used in support of all systems of performance-based rewards;
- These programmes represent a relatively cheap financial investment in education; and

- The market provides the best model for efficient resource allocation, which is predominantly used to support knowledge and skills and school-based systems but can be presented as an argument in favour of all models of performance-based reward programmes.

28. In general, most arguments principally support knowledge and skills and school-based rewards, which shows a movement away from the support of merit pay in recent literature.

4 ARGUMENTS OPPOSING PERFORMANCE-BASED REWARDS

29. This section describes the arguments made by a range of analysts who are concerned to oppose the view that net benefits can be expected from using performance-based rewards programmes for teachers.

4.1 Fair and accurate evaluation is difficult

30. A wide body of literature criticises the evaluation procedures of performance-based rewards. In this literature it is argued that goals are hard or impossible to establish in teaching because key education outcomes have not been identified, and this necessarily reduces goal clarity (Storey, 2000). One problem evident, it is argued, is the complexity of designing a programme that balances clarity of goals and diverse evaluation criteria, since clear criteria are required to measure productivity gains. This problem is compounded since evaluation is often done through proxies, such as self-report surveys that ask teachers about the motivational impact of the programme, which are at best indirect measures (Richardson, 1999). Rather, it is argued, teacher commitment and knowledge is often a better guide for good instruction than observing and assessing their performance (Firestone and Pennell, 1993).

31. Some analysts argue the performance of a student is beyond the control of a teacher. Rather than viewing the teacher as a single actor, the vital roles played by the school, the principal, and the family should be acknowledged (Holt, 2001). This means the ‘cause’ of educational achievement is difficult to establish, and includes numerous actors, not simply teachers (Evans, 2001). Confounding this problem, it is argued, is that the best teachers are often given classes that perform lowest academically, and may therefore be punished under a performance-based payment system (Evans, 2001). Even the recent efforts to establish ‘value-added’ evaluation criteria are considered problematic because they are in the embryonic stages of development, and there are clear socio-economic and racial biases in these systems (Clotfelter and Ladd, 1996).¹

32. Erroneously rewarding teachers is considered a problem with performance-based programmes (Cutler and Waine, 2000). How do you adequately evaluate a teacher based on student outcomes when previous teachers may have taught superior learning techniques (Cited in Solomon and Podgursky, 2001)?

¹ Clotfelter and Ladd (1996) argue that school systems have a clear choice when designing systems whether to control for socio-economic, racial and gender characteristics. They argue there is a trade-off between adjusting for differences in schools, and the possibility of sending undesirable messages to the community that a school system has a reduced expectation of some students’ outcomes. They report systematic differences in student progress which can be attributed to socio-economic, racial and gender characteristics.

While group-based rewards attempt to overcome this problem by evaluating teacher performance as a whole, questions remain about the equitable division of rewards given the complex relationships that exist between teachers and student outcomes. This questions whether schools are much too complex organisationally for accurate evaluation to occur (Cited in Storey 2000).

4.2 School administration becomes hierarchical

33. It is argued that proper employee evaluation requires an equal participation and relationship between the key participants. When pay is linked to performance, any equality is undermined because there is inevitably a judgmental aspect that makes this equal relationship obsolete (Cutler and Waine, 2000). Teachers, on the one hand, use evaluation as a formative process, allowing them to see how they are performing, and how they can improve. Administrators, on the other hand, use evaluation for summation, which considers evaluation as a process used to gauge teachers worth (Barber and Klein, 1983). This is supported by Murnane and Cohen (1986) who argue principals in the 1980s United States school system were found to prefer giving better evaluations than the teachers actually deserved to build trust between the administrators and the teaching staff, and also as a form of formative evaluation. Thus, it is argued that a functioning professional relationship between the principal and the teachers would be undermined by the use of performance-based rewards.

34. It is also argued morale can be reduced because merit pay creates unfair competition between teachers (AFT, 2001). Teachers who have not been rewarded can question the fairness of evaluation, as there are frequently no transparent criteria. Even if the evaluation process is completed accurately and fairly, teachers may still feel aggrieved if they are not considered competent (Ramirez, 2001) and new hierarchies can be evident in administrators who now have power over teachers and the curriculum (Holt, 2001).

4.3 The incentive systems do not motivate teachers

35. Another common criticism is that teachers are not particularly motivated by pecuniary reward so they will not respond to financial incentives. If money is a relatively small motivator for teachers, attempts to focus on monetary-reward systems can have the consequence of increasing resentment towards management, and reducing employee loyalty, resulting in a reduction in productivity (Ramirez, 2001). This is supported by numerous surveys that suggest intrinsic rewards are very important to teachers (Firestone and Pennell, 1993). Firestone and Pennell (1993) argue that evaluation can undermine the intrinsic rewards for teachers, as the “feedback in the form of performance evaluation undermines intrinsic motivation, *even when the evaluation is positive*” (emphasis in original). It is argued that non-monetary rewards may be better motivators, such as extra holidays. This has been observed in Canada, where many teachers take up the opportunity for unpaid leave. This raises the question of whether the current models of performance-based rewards are flawed because they fail to recognise actual teacher motivations (Chamberlin, et al, 2002). However, Odden (2001) argues that while research has shown current teachers to be motivated by intrinsic rewards, this does not mean potential teachers would not be motivated by financial rewards. These potential teachers could well be talented, but have hitherto been employed within the private sector because of inadequate financial rewards available for teachers.

4.4 Reduced co-operation between teachers

36. The literature cites reduced collegiality between teachers as a major problem with performance-based reward programmes. Even proponents argue that many of the early systems of performance-based rewards had a problem with encouraging co-operation, as systems of merit-based pay are considered at

odds with the team-based nature of teaching (Odden, 2000a). Hoerr's (1998) argument that programmes need to be carefully designed or competition between staff members may reduce collegiality among teaching colleagues echoes these sentiments. This, Hoerr (1998) and Odden (2000a) argue, is a function of poor programme design, rather than an inherent characteristic of performance-based rewards.

37. Nevertheless, a large body of literature argues these programmes have a negative effect on teacher collegiality. For example, Chamberlin, et al, (2002) argues that competition amongst teachers, in a profession where co-operation is essential, undermines any attempt to introduce performance-based rewards. The American Federation of Teachers (AFT, 2001), a United States teacher union, argues that previous programmes created divisions between teachers, as they were classified as either 'winners' or 'losers' (see also, Storey, 2000). It is argued that even when a school-based system is used, collegiality is adversely affected, sometimes because limited funding means the average reward is often so small it is meaningless (Malen, 1999), sometimes because of the 'free rider' problem. The 'free rider' problem occurs when some teachers who are not contributing to the outcomes of students are rewarded because of others' actions (Cutler and Waine, 2000).

4.5 Unwanted outcomes

38. Opponents of performance-based reward systems argue there can be significant problems with the outcomes of these systems. The American Federation of Teachers (2001) argues performance-based reward programmes can create a system where the curriculum is narrowed and a 'teaching to the test' mentality becomes evident, which restricts the advancement of students in areas not tested. This occurs when only specific skills or outcomes are measured and rewarded (Chamberlin, et al, 2002). The result is a narrowed education, with an under-emphasis on subjects which are hard to evaluate, meaning the breadth of intellectual activities in schools is narrowed (Holt, 2001; Ramirez, 2001). A typical question asked by critics is: how would a performance-based system reward characteristics such as honesty, civic responsibility, etc (Evans, 2001)? Further problems could become apparent if teachers 'game play', and develop responses that generate rewards against the spirit of teaching (Malen, 1999). These concerns are relevant for group-based programmes because the unwanted outcomes can occur on a school-wide, rather than individual basis. This can cause institutional limitations of the curriculum and a downgrading in importance of certain subjects that are not measured (Chamberlin, et al, 2002). In other words, by measuring student output, perverse rewards can be encouraged.

39. Poorly performing students may suffer under a performance-based pay system because they may require significant tuition to improve. Teachers would focus a disproportionately large amount of their time on the students most likely to gain from their tuition to maximise the benefit derived, generally argued to be the middle band of students (Murnane and Cohen, 1986). Evans (2001) questions how this would affect schools in low socio-economic areas, since the time needed for improved student outcomes may be substantial. While a school-based reward strategy provides an incentive for the most poorly performing students to be encouraged and improved, teachers may still concentrate their efforts on those students who are most likely to cross a threshold. The highest and lowest performing students may be neglected because they do not represent a quality investment of teachers' time (Chamberlin, et al, 2002). In the same manner, if poorly performing schools are under funded, a school-based strategy will not work until additional funds and expertise are provided (Malen, 1999).

4.6 Financially irresponsible

40. The literature argues performance-based reward schemes require significant performance-related supplements in salary if they are to be implemented successfully. On these arguments, increased salaries

would require increased education revenue, which may be politically difficult (Hoerr, 1998; Holt, 2001; Chamberlin, et al, 2002). Furthermore, if evaluation and reward is expensive, any attempt to level the salary schedule and supplement rewards is ignoring past failed attempts at performance-related pay (Barber and Klein, 1983).

41. Even some proponents of performance-based rewards acknowledge that administering such a system would also require an extensive bureaucracy. For example, Odden (2000) argues that it would be expensive to adequately evaluate every teacher, and would require considerable resources if this evaluation were to be completed regularly. Furthermore, the time needed to administer this kind of a system would have severe budgetary implications (Cutler and Waine, 2000).

4.7 The market approach is inadequate

42. Numerous analysts question the application of market ideas to teaching. This body of literature argues education is a public good, and should not be analysed within a market framework. For example, Richardson (1999) questions the success of individual performance-based reward systems in the public sector in comparison to the private sector. Their lack of success, he argues, means that these private sector models are ill suited to the public sector. Other analysts point out that teachers work with human beings, and not robots or inert objects. In this way, teaching is different from the private sector precisely because education fashions and works with human beings (Cited in Solomon and Podgursky, 2001). Teachers are not permitted to discard any of their “products”, and must consider a wide range of student outcomes, including reading, computation, inferential reasoning and critical analysis, creative expression, handwriting, exposition, social adjustment and more (Chamberlin, et al, 2002). Thus, it is argued, schools are not factories, and you cannot translate the systems of factories into schools and education institutions successfully. Closely related to this argument, Firestone and Pennell (1993) assert there is evidence that teacher commitment is positively correlated to reading and language arts achievement, meaning policies that damage teacher commitment would damage these student outcomes. So when teachers and the public believe that formal education is important to society and has important effects on individual life outcomes, any policies that have the *potential* to undermine teacher commitment should be rejected because the high stakes involved (Firestone and Pennell, 1993).

43. It is further argued that schools cannot operate in a purely ‘rational’ manner because they are not purely technocratic, nor are they apolitical (Malen, 1999). Management techniques based on the private sector are thus bound to fail when the work involves deliberative judgement rather than procedures. For example, merit pay is often used in workplaces where there is a visible output which can be measured, and employee practices and outcome can be easily identified, such as in a clothing factory. In contrast, teachers must use different practices based on individual student characteristics, which are difficult to identify. This means the market has no capacity to increase productivity in these workplaces, because the factors that increase student achievement are difficult to identify and define (Holt, 2001).

44. There are no universally accepted characteristics of a good teacher, so it is distinct from other services where output is easily measured, and techniques for improving productivity can be easily identified (Murnane and Cohen, 1986). There are numerous actors who have a stake in educational outcomes, including children, parents, taxpayers, potential employees, teachers and the government, which is separate from the private sector where the number of principals is limited (Burgess et al, 2001).

45. Most market-based group reward systems do not have a predetermined amount of revenue available, but will distribute a portion of profits from the additional benefit derived from increased productivity. This is not a possibility for public education since resources are fixed, and do not vary with changes in productivity (Mohrman, Mohrman, and Odden, 1996). Teachers rarely have control over

school resources, meaning extra salaries or bonuses can be difficult to fund (Mohrman, Mohrman, and Odden, 1996). This contrasts with the private sector, where increased productivity will generate increased profits, decreased outlays, or costs passed onto the consumer (Chamberlin, et al, 2002). This occurs because the product of labour is easily identified in the private sector, while the product of teachers' labour is not easily identified, nor rewarded (Mohrman, Mohrman, and Odden, 1996). This implies that individual merit pay will be difficult to administer in education because individual teacher quality is hard to measure on the basis of student outcome.

46. In summary the main arguments in opposition to performance-based rewards are:

- Objective evaluation of teachers is difficult,
- It would create hierarchies within school administration which would detrimentally affect student outcomes, which is particularly the case for individual forms of performance-based rewards;
- The incentive system would not motivate teachers;
- There would be reduced co-operation between teachers, which is presented as an argument primarily in opposition to merit-pay;
- A range of unwanted and perverse outcomes would be promoted, which is presented as an argument against using student outcomes as a measure of teacher performance;
- It would be an expensive programme, which is presented as an argument against all systems of performance-based rewards that offer a significant financial reward; and
- The market is an inadequate model for the public sector, which is used as an argument against any model of performance-based rewards.

47. In general, merit-pay is the most contentious system of performance-based rewards, but there is also concern with skill and knowledge and school-based models.

5 DIFFICULTIES IN IMPLEMENTATION

48. This section describes the difficulties in the implementation of performance-based reward programmes observed by a range of analysts.

5.1 Widespread unionisation of teachers

49. The literature consistently argues that one of the major difficulties in the implementation of performance-based reward programmes has been the existence of teacher unions who have been strong opponents of these programmes (Ballou and Podgursky, 1993; McCollum, 2001). Schools are typically highly unionised workplaces, and teacher unions have traditionally rejected movements towards merit pay (Tomlinson, 2000; AFT, 2001). Wage differentiations on the basis of subject taught, and any sort of

subjective evaluation of teachers for rewards has been rejected outright, possibly because of existing collective bargaining strategies (Ballou and Podgursky, 2001). Typically, unions employ a range of arguments to reject attempts to introduce performance-based rewards, particularly focusing on doubts about accurate evaluation of teachers.

50. By lobbying legislatures against merit pay, unions have frequently changed the shape of systems or reduced the number and frequency of performance-based reward programmes (Ballou and Podgursky, 1997). Ballou (2001) reported that a common feature of schools with performance-based reward systems were the lack strong unions, which suggests that teacher unions can exert strong influences on school reform. This means radical reforms can be difficult to implement where union presence exists.

51. Contemporary efforts to introduce performance-based rewards therefore have to consider unions before implementation. However, this has been possible, as there are a group of teacher unions in the United States who now support the Consortium for Research and Policy in Education's (CRPE) efforts to introduce knowledge and skills based pay (Odden, 2001b).

5.2 Teacher opposition

52. Another reported reason for the failure of performance-based reward programmes is the apparent opposition of teachers. Ballou and Podgursky (1993) argue teachers have been opponents of performance-based pay. Explanations for this opposition vary widely, with some attributing this opposition to the reduction of autonomy of teachers because of constraints on their teaching style and outputs (Firestone and Pennell, 1993). When teachers' autonomy is threatened, they are likely to respond negatively which may impact on student outcomes (Firestone and Pennell, 1993). Furthermore, Malen (1999) argues there is a fundamental tension between the policy makers and the public, and teachers, since the most attractive component of performance-based pay with policy makers and the public has been the individual and differentiated selection criteria, whereas teachers often have deep-seated concern about the fairness of individual evaluation. This is also one of the most common concerns cited within the literature, which suggests that there is a conflict between past programmes of individual performance-based rewards, and teacher motivation (Firestone and Pennell, 1993).

53. Highly politicised and sanctioning programmes can increase the stress levels of teachers which can cause further teacher opposition. For example, the Kentucky School-Based Performance Award (SBPA) had statistically significant less anticipation of positive outcomes than the Charlotte–Mecklenburg SBPA and a distinguishing feature between the systems was the existence of sanctions for poorly performing schools in Kentucky (Kelley, Heneman and Milanowski, 2002). When these programmes become politicised, there appears to be a greater likelihood of teacher opposition. Other analysts argue staff room culture is inimical to a form of performance pay system. Hence, staff room culture must be changed before any performance-based systems of reward can be implemented successfully (Storey, 2000). This may be overcome relatively easily by including teacher input in the design and implementation of performance-based reward programmes (Firestone and Pennell, 1993).

54. A study of teachers' attitudes towards performance-based rewards was conducted by Ballou and Podgursky (1993) (Refer to Annex 4 for methodology and discussion of this study). They found that most teachers surveyed were in favour of additional pay for additional duties, and as part of a career ladder where performance dictated the speed of advancement (Ballou and Podgursky, 1993). However, there was some concern that the evaluation process could be seen as unfair or inadequate. This means performance-based rewards, in particular pay, is considered to be difficult to administer objectively and fairly (Ballou and Podgursky, 1993). Unsurprisingly, performance-based rewards are reported to be more popular when it is viewed as supplementing, rather than replacing, other forms of salary (Ballou and Podgursky, 1993).

55. The level of pay in a school district appears to have no influence on teachers' attitude towards merit pay, yet it was more likely to be supported by teachers with low salaries and by ethnic minorities such as black and Hispanic educators (Ballou and Podgursky, 1993). Attitudes towards merit pay were found to be independent of the number of students eligible for free lunches, suggesting the socio-economic status of the students does not affect teacher views in the United States. Ballou and Podgursky (1993) reported a distinction between private and public school teachers, with private school teachers being more in favour of performance-based pay. This research suggests that teacher attitudes are more malleable than is argued by some analysts, since this research points towards different teacher attitudes depending on programme design.

5.3 Political opposition

56. Traditionally a wide range of political groups have been involved in the organisation and promotion of performance-based reward programmes. Implementation can be difficult because any one of a number of bodies can discontinue programmes. For example, Ballou (2001) argues legislators, school superintendents and school boards all have the power to discontinue performance-based reward programmes in the United States. As supporting legislators leave office, the political will to continue what can be a costly enterprise can disappear, particularly in times of economic recession (Ballou and Podgursky, 1997; McCollum, 2001). As Cohn (1996) argues, in times of economic recession it can be difficult to implement new performance-based strategies, and existing programmes come under political attack. One possible explanation is the dollar costs of these programmes are more easily measured than the more vague benefits in student outcomes, so a cost-benefit analysis cannot be completed easily by policymakers (Chamberlin, et al, 2002).

5.4 Poor design and implementation:

57. Poor design and planning in the past has created difficulties in implementing new performance-based pay systems. This sets up the expectation that because it hasn't worked in the past, it will not work in the future (McCollum, 2001). This is one of the few areas in the literature where a consensus is evident. Analysts, both proponents and opponents of performance-based rewards argue that previous attempts had poor design and implementation (Mohrman, Mohrman and Odden, 1996; Ramirez, 2001). Problems in developing fair and reliable indicators and the training of evaluators to fairly apply these indicators undermine any attempt to implement programmes (Storey, 2000).

58. One problem identified is poor goal clarity because of a large number of criteria, which restricts teachers' understanding of the programme and makes implementation difficult (Richardson, 1999). Explanations of how, and on what criteria teachers are assessed may be difficult to articulate. When this occurs, it is almost impossible to give valuable feedback and maintain teacher support for the programme (Chamberlin et al, 2002). If administrators cannot tell workers why one worker got a bonus, while another did not, the programme would face severe pressures (Murnane and Cohen, 1986). Stress levels may also be increased when teachers are expected to work harder towards multiple goals (Kelley, 1999).

59. Several proponents of performance-based reward systems argue that previous systems have been simplistic in their design and implementation. Successful strategies are needed to expand professional development so teachers can learn the new knowledge and skills that are required for skill and knowledge based pay (Odden, 2000b). As performance-based curriculum requires deep conceptual understanding of curricula content, and an array of pedagogical strategies, a great deal of strain is placed upon teachers (Mohrman, Mohrman and Odden, 1996).

60. One example of a recent attempt to overcome this problem is the Consortium for Policy Research in Education's (CPRE) work on sophisticated performance-indicators for teachers. They argue these tests can be applied for accurate and objective evaluation (Odden, 2000a) of core teacher skills to be completed easily and consistently both across and within school jurisdictions (Odden, 2000b). These tests control for a number of social factors such as socio-economic differences, racial differences and previous student outcomes by providing bonuses tied to school performance, which are weighted according to these factors (Odden, 2000a). Similarly, Cohn (1996) advocates the use of evaluation by arguing student test scores measures the most fundamental student achievement.

61. Another technique was developed by Solomon and Podgursky (2001) who use regression analysis techniques based on student results to show the effectiveness of teachers. Student scores before the start of an academic year were compared to their end of year scores, with various factors such as socio-economic indicators controlled for, to provide an evaluation of teachers (Solomon and Podgursky, 2001). Teachers can thus be assessed on how much they have added value to student outcomes, which can be considered an accurate tool for evaluation. Therefore, it is argued, evaluations can be made with minimal error, and teacher effectiveness objectively established (Solomon and Podgursky, 2001). In fact, Solomon and Podgursky (2001) argue "schools are probably more amenable to monitoring individual performance than are most private goods or service-producing firms", because of the ease of measuring the 'added-value' of education. Furthermore, because these evaluations can be measured externally to the schools, political bias in teacher promotion is reduced (Solomon and Podgursky, 2001).

62. However, it has also been argued that previous financial bonuses have been comparatively small, which undermine the motivational value of the programmes. A great deal of literature has noted that the rewards offered have not been enough of an incentive to change teacher behaviour (Malen, 1999). The money rewarded has been limited and this has meant that arbitrary quotas were often established which provided only small incentives to a majority of practitioners (Chamberlin, et al, 2002). Further problems can occur when there is a belief that teachers will not get rewards even for increased performance (Richardson, 1999). This problem has been highlighted in several studies, including the Kentucky and Charlotte-Mecklenburg programmes, with scepticism about future reward bonuses evident in even well established programmes (see Kelley, Heneman and Milanowski, 2002).

5.5 The entrenched ideas of the public school system

63. Publicly administered schools are typically the most common type of educational institution, and their management system can become the norm. Consequently, private schools are more likely to follow these management norms and employ the single salary scale used by the publicly administered schools, rather than radical performance-based reward programmes. This would have the effect of reducing the incidence of performance-related pay by making it harder for private schools to administer these compensation packages (Ballou, 2001). Moreover, public school systems are more likely to be composed of large schools with large numbers of teachers, meaning significant changes in the payment structure become more difficult. Smaller school systems such as private schools are much better equipped to make radical changes to the reward structure but even these schools face problems because of the entrenched ideas of the public school system (Ballou and Podgursky, 2001).

64. In summary the explanations for difficulties in implementing performance-based reward programmes for teachers are:

- The opposition from teacher unions, particularly in relation to merit-pay models;
- The opposition from teachers, particularly related to concern about unfair evaluation;

- Political opposition, especially during times of economic recession;
- The poor design and implementation of previous programmes, especially in relation to merit-pay models; and
- The existence of an entrenched public school management system which opposes market-based models.

65. In general, most of the difficulties in implementing performance-based rewards occurred in earlier merit-pay systems.

6 EVIDENCE OF THE EFFECTS ON EDUCATIONAL OUTCOMES

66. Considering the breadth of argument surrounding performance-based reward programmes, there is a surprising lack of empirical evidence evaluating their effect. Table 3 summarises the potential effects of performance-based reward programmes on teacher, student, classroom, school, school system and societal levels. This section reviews the research that has been completed on performance-based reward programmes.

Table 3. The Potential Effects of Performance-Based Rewards

Level	Potential Area of Effect
Teacher	<ul style="list-style-type: none"> • Teacher motivation and effort • Teacher recruitment and attrition. • Teacher knowledge and skills • Teacher autonomy
Student	<ul style="list-style-type: none"> • Student performance • Student truancy • Student drop-out
Classroom	<ul style="list-style-type: none"> • Pedagogical techniques • Teacher focus on specific students
School	<ul style="list-style-type: none"> • Collegiality between teachers • The efficiency of resource allocation in schools. • The relationship between teachers and school management • School organisational goals
System	<ul style="list-style-type: none"> • The revenue required for teacher salaries and the education system • The culture of educators • The form and content of the curriculum
Societal	<ul style="list-style-type: none"> • Public and political support, particularly the publics' perception of the teaching profession.

6.1 Kentucky and Charlotte-Mecklenburg Group-Based Performance Reward Programmes

67. One of the most commonly cited studies considering the outcome of group-based performance rewards are Kelley, Heneman and Milanowski's (Kelley, 1999; Heneman and Milanowski, 1999; Kelley, Heneman and Milanowski, 2002) studies of the Kentucky and Charlotte-Mecklenburg's School-Based Reward Programmes (See Annex 3 for the methodology and conclusions of Kelley, Heneman and Milanowski (2002)). It is argued school-based reward programmes are beneficial because they motivate teachers, and this improved motivation increases student performance, which has a positive overall effect on student outcomes. The purpose of the study was twofold. First, it assessed how School-Based Reward Programmes affected teacher motivation.² Second, it attempted to find out how teacher motivation affected student outcome. They concluded that school-based reward programmes motivated teachers to perform better. They also found that more highly motivated teachers were more likely to be teaching in areas of high student outcomes. They did not consider the size of student gains.

68. The authors acknowledge there are several methodological problems with their study. The lack of a control group to test the motivational effect of salary scales was one such problem. There were no comparisons made with other similar jurisdictions that did not use performance-based pay. This meant that there may have been confounding variables that explain increased teacher motivation or student outcomes the effect of which could not be controlled for under this research design, so Kelley, Heneman and Milanowski's analysis of the motivational impact of SBPA programmes vis-à-vis the salary scale is thus problematic. Even though the authors acknowledge this problem, they do still attempt to address these issues.

69. The authors collected qualitative data and conducted a literature review to select variables that were used in the regression which tested the relationship between teacher motivation and student outcomes. There is evidence that the variables tested were not comprehensive because the correlation coefficient squared was 0.345 in the Kentucky sample, and 0.337 in the Charlotte-Mecklenburg study, meaning a large proportion of variance was not explained by this model. A more theoretical approach could have produced links between the concepts, and thus, a way of explaining the observed effects. This may have occurred because Charlotte-Mecklenburg undertook a number of reforms to education around a similar time, which could explain some of the variance in student outcomes.

70. Furthermore, student test performance was used to measure the outcome of the programme, so the measure of student success was a function of the programme they were embedded within. This means a school was considered successful if it raised student test scores, and by no other measure. This raises the question of whether the positive programme outcomes were limited to academic improvement. However, if the intention of the programme was only to improve student academic performance, this suggests there was a clear connection between programme goals, and teacher understanding of these programme goals. Moreover, the extent these results can be generalised to other systems is uncertain, as Kentucky and Charlotte-Mecklenburg had different effects on teacher motivation. One possible explanation for this difference was the existence of sanctions in the form of being labelling as a 'school in crisis', or a 'school in decline' in the Kentucky group. This may have reduced teacher motivation because of increased levels of stress.

71. Because this research uses teacher motivation as a proxy for increased student outcomes, it is difficult to fully integrate these findings into an evaluation of performance-based reward systems, because not all variations can be explained in their model. In other words, Kelley, Heneman and Milanowski's (2002) work is problematic because their conclusion requires a missing conceptual step to link programme

² Teacher motivation was measured by a self report survey administered to a random sample of teachers from Kentucky and Charlotte-Mecklenburg.

influence to student outcome. As well, while there may be evidence of increased teacher motivation, the extent to which there is also teacher de-motivation needs to be considered at length.

72. Smith and Mickelson (2000) evaluated the outcomes of Charlotte-Mecklenburg by contrasting them to other urban school districts in North Carolina. Their conclusion that there is no benefit to the range of reform policies introduced in the early 1990s contrasts strongly with the previous analysis by Kelley, Heneman and Milanowski (2002). Smith and Mickelson (2000) examined student outcomes, defined as SAT scores, student proficiency scores, and drop out rates for a range of age levels. They compared progress on these criteria against state-wide averages. Their statistical model did not control for the effect of teacher characteristics, student characteristics or school characteristics (other than using dichotomised 'black' and 'white' schools). This limits the study because there may be differences between the control models used to explain the slow student improvements exhibited in Charlotte-Mecklenburg relative to other school districts. Further difficulties are evident in the interpretation of this study because the Charlotte-Mecklenburg undertook a complex reform programme that involved many other policy developments and not just performance-based rewards. This makes it difficult to identify the unique effect of these programmes. Thus, there are a number of concerns about this study and so what it contributes to understanding on the impact of performance-based reward programmes is limited.

6.2 South Carolina Individual and Group-Based Reward Programmes

73. Cohn and Teel (1992) made an early evaluation of the South Carolina teacher incentive programme. This programme used both school-based performance rewards and individual performance-based rewards as a method for increasing school accountability. Cohn and Teel (1992) conducted a quantitative study of the impact this programme had on student gain scores. A random sample of participating districts was taken, including both group and individual reward programmes. The results of these were compared to the state averages. A multiple regression analysis was conducted, which controlled for a number of variables, including teacher experience, student gender, teacher gender, teacher race, student race, teacher education and student disadvantage. The results showed that the teachers who gained awards in individual programmes had a higher gain score than non-participating teachers. The authors conclude this means the programme was distributing the rewards to the 'right' teachers. However, this study is limited by statistically weak results. Frequently, the regression and cross tabulation models failed to give significant results, even at the 10% level. This undermines the usefulness of the study. Even if their results had been more persuasive, it is interesting to note that Clotfelter and Ladd (1996) argue the South Carolina reforms are too broad to make conclusions regarding the effect of teacher incentive rewards. In other words, it would be very difficult to establish, given the complexity of the programme, which characteristics actually affected measured changes in teaching behaviour.

6.3 The Texas Education Agency Study of the Link Between Teacher Salary and Student Outcome

74. Hanushek, Kain and Rivkin (1999) conducted an important study that examined the relationship between teacher salaries and student outcomes in Texas. Data developed by the Texas School Project in the United States was used, and a multivariable analysis was undertaken to attempt to find a causal link between teacher pay and student outcomes. Their data was broken into teacher experience, teacher gender, and teacher movement and the effects of increases in salaries were examined. While this study does not consider performance-based rewards, it is useful because it links teacher salaries to student performance. They found that increased salaries are correlated with improved student outcomes. The largest salary effects were found in schools with no staff turnover and no probationary teachers. This challenges the

traditional assumption that increased salaries (including performance-based rewards) increase student outcomes because of new talent being drawn into the teaching profession (Refer to sections 3.3 and 3.6). The conclusion, when taken literally, is that salary increases motivate experienced teachers to work harder, rather than by raising the quality of new graduates. The implications of this study for performance-based reward programmes are considerable, since this research seems to support the argument that teacher performance can be increased through extrinsic rewards.

6.4 The Dallas Group-Based Performance Reward Programme

75. Ladd (1999) examined the Dallas school accountability and incentive programme, which consisted of an elaborately designed group-based performance reward strategy (Refer to Annex 5 for the methodology and conclusions of this study). Every year about 20% of schools awarded bonuses of \$1000 U.S. for each teacher and principal, and \$500 U.S. for non-teaching staff, and a further 30% of schools are rewarded a bonus of \$425 U.S. per teacher. Ladd (1999) reports there were a range of perspectives of this programme, with many considering it a progressive attempt to improve student outcomes, and many a narrow use of performance-measures to deflect attention from the poor state of the school system. The study showed an increase in the number of grade seven students who passed the Texas Assessment of Academic Skills (TAAS) every year from 1992-1995 relative to the baseline year, 1991, and other schools in Texas. While Hispanic and white students had a significant and positive increase in their performance in math and reading, there was no significant effect on black students' results.

76. The programme was designed to increase the level of accountability of the district as a means of improving student outcomes. Personnel were awarded \$2.5 million, with half being paid by local businesses, to the most effective 20 percent of schools. This means that schools could not accurately gauge their likelihood of obtaining performance-based reward bonuses during the school year. However, because the system used a complex 'added value' formula to establish student increases in test scores, it was able to discount racial, gender and socio-economic factors from the measures, thereby enhancing the overall validity of the evaluation process. An earlier study by Clotfelter and Ladd (1996) concluded the Dallas programme was able to successfully strip all socio-economic factors from test scores. This may not necessarily be considered beneficial as there is a trade-off between socio-economic bias, and perpetuating the idea that some social groups cannot achieve consistently. On the one hand, there is the danger that if socio-economic, racial and gender factors are not controlled for, schools from racially homogenous and affluent areas may consistently be rewarded. Clotfelter and Ladd (1996) reported that in the South Carolina performance-based programme, schools of low socio-economic status rarely won awards, which seems to support the perspective that the inclusion of socio-economic indicators is important in an analysis of student improvement. On the other hand, if these indicators are controlled for, society can be sent a message that the education system has lower expectations of some students. This raises the possibility that students from lower socio-economic indicators are expected to perform poorly in education on an institutional basis. The optimum level to control for these biases while minimising the negative message sent to society about decreased expectations of certain students is not clear and presents an important issue in programme design.³

³ South Carolina responded to the problem of socio-economic bias in the distribution of rewards by dividing schools into five groups based on socio-economic characteristics and rewarding the top performers in each category. This turned out to be problematic because some of the lower performing schools in the upper socio-economic bands felt they were treated unfairly. This mechanism also encouraged a perverse response from principals who realised their ability to gain an award was based on the socio-economic category they were placed in (Clotfelter and Ladd, 1996).

77. Ladd's (1999) study is a sophisticated attempt to measure the effect of group-based performance rewards on student outcomes. By using controls of several similar jurisdictions and a baseline of results before the programme was initiated, Ladd (1999) was able to reduce the chance of confounding variables influencing the outcome of the results. However, a concern when evaluating this study was the unexplained increase in student results the year before the programme started. Ladd (1999) considers two possibilities: first, it represents a movement back towards the mean state result since results in Dallas were well below the state average, or second, it was the result of positive publicity of the project before it was undertaken. This casts some doubt on the cause of increases in student outcomes. Yet Ladd (1999) argues that student results would have been positive even if 1992 was taken as a baseline year since results were consistently positive from 1992 to 1995. To test this hypothesis, it would be useful to have standardised students' results before 1991 to see the long-term fluctuations in student pass rates. This may give a clearer indication of whether the rise in student pass rates was caused by the adoption of a group-based performance reward programme, or whether it reflected a more general movement towards state means. Without this information, it is difficult to estimate the effect on student outcomes.

78. However, because the average pass rate was used as a measure of programme success, this study does not consider whether teachers focused on moderately performing students. This is an important question since it has been argued that these students represent the best allocation of teacher time for average student performance outcomes (See section 4.5). A more thorough analysis would have been produced if Ladd (1999) measured in what ranges changes in test score occurred, and in particular, whether there was any evidence of systematic teaching bias towards students within a given range. This limits the ability to generalise these results as no conclusions can be drawn on the average increase in student performance, but only the average increase in the number of students who pass the TAAS. While Ladd (1999) argues that the average passing rates were a reasonably accurate method of measuring student achievement because of the relatively low passing rate of students in Dallas, it can be argued that this is a poor indicator of actual student increases, particularly in higher achieving areas.

79. A survey of younger students could not be completed by 1996, meaning the study is somewhat limited in the breadth of students examined. Would the results apply only to grade seven students, or all students in the school? Are the gains that were found simply accelerated gains that would normally have occurred in grade eight, or nine, or ten? The answers to these questions are unclear from these data, but may have important implications for the interpretation of these results. It should also be noted that the collection of data from 1994 and 1995 cohorts occurred at grade six level in the spring, instead of grade seven level in the fall (Ladd, 1999). Ladd (1999) is concerned with the impact on outcomes, as students tested in spring may have been more likely to have performed at a lower level.

80. Alternative measures suggest the programme had beneficial outcomes. The number of grade seven students who dropped out was reduced by a statistically significant amount in comparison to the baseline year and other jurisdictions. This suggests that there are benefits to the educational system. Another interesting, but unexpected, outcome was an increased turnover in principals after the accountability programme was introduced. Ladd (1999) interprets this as positive outcome, since it suggests a greater accountability of principals, and may represent a more rigorous selection of competent principals. However, it is unclear what the actual optimal level of principal turnover is, so the merits of this outcome are not known.

81. Other, non-empirical measures support the argument that the Dallas programme produced positive outcomes. Clotfelter and Ladd (1996) argued that there was a much better chance of accurate administrative alignment within schools as it reduced the barriers between teachers and administrators. To put it another way, the administration of schools improved because teachers were more likely to share relevant information with administrators. This is consistent with Kelley's (1999) thesis on resource alignment discussed in section 3.2.

82. This study found a statistically significant increase in student pass rates within schools participating in the programme compared with of state-wide standardised tests relative to the baseline year and relative to other urban centres in Texas. This makes the study a valuable contribution to the literature. The analysis was able to control for a number of variables, meaning the statistical analysis is robust. Therefore, this study provides some evidence that a group-based performance reward system can increase students' ability to pass the TAAS in math and reading. Thus, this study also provides limited evidence that performance-based reward systems can be used to increase the level of achievement, even if the ability to generalise the results is limited by the Dallas programme's characteristics. This is also limited by the lack of data to explain the failure of black students to increase in their pass rates.

6.5 Conclusion

83. Considering the available research, it would appear there is some evidence of improved student performance in group-based performance reward programmes. However, the research is limited to the United States, and there are concerns about the study designs and data. Further problems in generalising these results exist because reform programmes generally encompass a range of initiatives, not simply a performance-based reward system, so working out how each factor contributes to educational outcomes becomes difficult. At the very least, there is no evidence of negative student outcomes with the adoption of the latest performance-based reward programmes. Furthermore, the evidence presented within the literature is almost unanimously positive in their assessment of performance-based reward programmes. To fully understand the effect that performance-based reward has on student outcomes, a wider study is required to consider how increased or decreased rewards affect teacher behaviour. Since most programmes provided bonuses around \$1000 U.S. per year, the relationship between increasing rewards and student outcomes is unknown. From a policy perspective, it would be useful to analyse the effect increased bonuses had on teacher behaviour and student outcome.

7 CONCLUSION

84. In documenting the arguments surrounding performance-based rewards, there has been no attempt to make normative judgements. The first section looked at the different types of performance-based programmes. The next three sections of this paper summarised the major themes and arguments that have been placed within the academic and policy literature. The fifth considered the empirical evidence on actual programme outcomes.

85. What has become evident is there are a wide range of opinions, and little consensus on even the fundamental issues surrounding performance-based rewards. To compound these problems, there is very limited research on the value of performance-based reward programmes. Among the few studies that have been completed, considerable doubts have been raised about the validity of the conclusions reached due to limitations in design. The limited evidence suggests there are some benefits evident within a group-based performance programme pertaining to increased student outcome. There is no evidence regarding individual performance-based programmes.

86. There is a wide consensus that previous attempts at introducing performance-based reward programmes have been poorly designed and implemented. Most concern surrounds merit-pay systems,

which have been heavily criticised in the literature. Given the range of values, principles and ideas that surround performance-based reward programmes, any design and implementation is probably best carried out with intensive consultation and adequate funding. The past has shown that performance-based compensation will inevitably be short lived without careful design and implementation.

REFERENCES

- American Federation of Teachers (2001) 'Merit pay', retrieved 15 January 2003 from:
<http://www.aft.org/issues/meritpay/meritpay.html>.
- Bainbridge, W (2000) 'Why schools should consider marketplace pay', *American School Board Journal*, 187 (7), pp 54-55.
- Ballou, D (2001) 'Pay for performance in public and private schools', *Economics of Education Review*, 20, pp 51-61.
- Ballou, D and Podgursky, M (1993) 'Teachers' Attitudes towards Merit Pay: Examining Conventional Wisdom', *Industrial and Labor Relations Review*, 47 (1), pp 50-61.
- Ballou, D and Podgursky, M (1997) *Teacher Pay and Teacher Quality*, W. E. Upjohn Institute for Employment Research, Michigan.
- Ballou, D and Podgursky, M (2001) 'Defining Merit', *Education Matters*, 1 (1), pp 16-25.
- Barber, L and Klein, K (1983) 'Merit Pay and Teacher Evaluation', *Phi Delta Kappan*, 65 (4), pp 247-51.
- Bergem, T (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Norway', *Eurydice*, retrieved 29 January 2003 from
<http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Burgess, S; Croxson, B; Gregg, P and Propper, C (2001) *The Intricacies of the Relationship Between Pay and Performance for Teachers: Do teachers respond to Performance Related Pay schemes?* CMPO Working Paper Series No. 01/35, Bristol.
- Chamberlin, R; Wragg, T; Haynes, G and Wragg, C (2002) 'Performance-related pay and the teaching profession: a review of the literature', *Research Papers in Education*, 17 (1), pp 31-49.
- Clotfelter, C and Ladd, H (1996) 'Recognizing and Rewarding Success in Public Schools'. In Ladd, H (ed), *Holding Schools Accountable: Performance-Based Reform in Education*, The Brookings Institution, Washington DC.
- Cohn, E (1996) 'Methods of Teacher Remuneration: Merit Pay and Career Ladders', in Becker, W and Baumol W (eds), *Assessing Educational Practices: The Contribution of Economics*, Russell Sage Foundation, New York.
- Cohn, E and Teel, S (1992) *Participation in a teacher incentive program and student achievement in reading and math, 1991 proceedings of the business and economic statistics section*, American Statistical Association, Alexandria, VA.

- Coolahan, J (2002) *Draft Report: Ireland*, OECD Background report: attracting, developing and retaining effective teachers.
- Cutler, T and Waine, B (2000) 'Mutual Benefits or managerial control? The role of appraisal in performance related pay for teachers', *British Journal of Educational Studies*, 48 (2), pp 170-182.
- Department of Education, Employment and Training (DEET), Victoria (2001) *Victorian Government Schools: Staff Performance Handbook*, Department of Education, Employment and Training, Melbourne.
- Devos, G and Vanderheyden, K (2002) *Thematic Review- Flanders, OECD Background report: Attracting, developing and retaining effective teachers*, retrieved 29 January 2003 from <http://www.oecd.org/EN/document/0,,EN-document-609-5-no-21-36392-0,00.html#title1>.
- Doukas, C and Smyrniotopoulou, A (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Greece', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Eicher, J and Chevailler, T (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- France', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Esteve, J (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Spain', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Evans, D (2001) 'No merit in merit pay', *American School Board Journal*, 188 (1), pp 48-50.
- Farnsworth, B, Debenham, J and Smith, G (1991) 'Designing and Implementing A Successful Merit Pay Program for Teachers', *Phi Delta Kappan*, 73 (4), pp 320-25.
- Firestone, W and Pennell, J (1993) 'Teacher Commitment, Working Conditions, and Differential Incentive Policies', *Review of Educational Research*, 63(4), pp 489-525.
- Hanushek, E; Kain, J and Rivkin, S (1999) *Do Higher Salaries Buy Better Teachers?*, Working Paper 7082, National Bureau of Economic Research, Cambridge.
- Held, F (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Denmark', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Heneman, H and Milanowski, A (1999) 'Teachers Attitudes about Teacher Bonuses Under School-Based Performance Award Programs', *Journal of Personnel Evaluation in Education*, 12 (4), pp 327-41.
- Heneman, R and Ledford, G (1998) 'Competency Pay for Professional and Managers in Business: A Review and Implications for Teachers', *Journal of Personnel Evaluation in Education*, 12 (2), pp 103-121.
- Hoerr, T (1998) 'A case for merit pay', *Phi Delta Kappan*, 80 (4), pp 326-27.
- Holt, M (2001) 'Performance Pay for Teachers: The Standards Movement's Last Stand?', *Phil Delta Kappan*, 83 (4) pp 321-317/

- Ingvarson, L and Chadbourne, R (1997) 'Reforming Teachers' Pay Systems: The Advanced Skills Teacher in Australia', *Journal of Personnel Evaluation in Education*, 11 (1), pp 7-30.
- Jeuthe, E (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Germany', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Kelley, C (1999) 'The Motivational Impact of School-Based Performance Awards', *Journal of Personnel Evaluation in Education*, 12 (4), pp 309-26.
- Kelley, C; Heneman, H and Milanowski, A (2002) 'Teacher Motivation and School-Based Performance Awards', *Education Administration Quarterly*, 38 (3), pp 372-401.
- Kim, E and Han, Y (2002) *Background Report: Korea*, OECD Background report: Attracting, Developing and retaining effective teachers, Korean Educational Development Institute.
- Kotásek, J; Valenta, J and Brožová, S (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Czech Republic', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.htm>.
- Ladd, H (1999) 'The Dallas school accountability and incentive program: an evaluation of its impacts on student outcomes', *Economics of Education Review*, 18 (1), pp 1-16.
- Liang, X (1999) *Teacher Pay in 12 Latin American Countries: How does teacher pay compare to other professions, what determines teacher pay, and who are the teachers?* The World Bank, Latin America and the Caribbean Regional Office, Human Development Department, LCSHD Paper Series No 49.
- Luukkainen, O and Eurydice Finland (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Finland', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Malen, B (1999) 'On Rewards, Punishments, and Possibilities: Teacher Compensation as an Instrument for Education Reform', *Journal of Personnel Evaluation in Education*, 12(4), 387-94.
- McCullum, S (2001) 'How merit pay improves education', *Educational Leadership*, 58 (5), pp 21-24.
- Milanowski, A (2003) 'The varieties of knowledge and skill-based pay design: A comparison of seven new pay systems for K-12 teachers', *Education Policy Analysis Archives*, 11 (\$), Retrieved 30 January 2003, from <http://epaa.asu.edu/epaa/v11n4>.
- Mohrman, A; Mohrman, S and Odden, A (1996) 'Aligning Teacher Compensation with Systemic School Reform: Skill-Based Pay and Group-Based Performance Awards', *Educational Evaluation and Policy Analysis*, 18 (1), pp 51-71.
- Murnane, R and Cohen, D (1986) 'Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and a Few Survive', *Harvard Educational Review*, 56 (1), pp 1-17.
- Odden, A (2000a) 'New and better forms of teacher compensation are possible', *Phi Delta Kappan*, 81 (5), pp 361-66.
- Odden, A (2000b) 'Paying Teachers for Performance', *School Business Affairs*, June, pp 28-31.

- Odden, A (2001) 'Defining Merit', *Education Matters*, 1 (1), pp 16-24.
- Odden, A and Kelley, C (2002) *Paying Teachers For What They Know And Do: New And Smarter Compensation Strategies To Improve Schools*, 2nd Edition, Corwin Press, California.
- OECD (2002) *Education at a Glance: OECD Indicators 2002*, Paris
- Ramirez, A (2001) 'How merit pay undermines education', *Educational Leadership*, 58 (5), pp 16-20.
- Richardson, R (1999) *Performance Related Pay in Schools : An Assessment of the Green Papers, A report prepared for the National Union of Teachers*, The London School of Economics and Political Science.
- Schratz, M and Resinger, P (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- Austria', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Smith, S and Mickelson, R (2000) 'All that Glitters is Not Gold: School Reform in Charlotte-Mecklenburg', *Education Evaluation and Policy Analysis*, 22(2), pp 101-127.
- Solomon, L and Podgursky, M (2001) *The Pros and Cons of Performance-Based Compensation*, Milken Family Foundation, Pascadena.
- Storey, A (2000) 'A leap of faith? Performance pay for teachers', *Journal of Education Policy*, 15 (5), pp 509-23.
- Tomlinson, H (2000) 'Proposals for Performance Related Pay in English Schools', *School Leadership and Management*, 20(3), pp 281-298.
- Vossensteyn, H (2001) 'National summary tables on the reforms of the training and conditions of service of teachers in lower secondary education- The Netherlands', *Eurydice*, retrieved 29 January 2003 from <http://www.eurydice.org/Documents/KT3tables/en/FrameSet.html>.
- Waterreus, J (2001) *Incentives in Secondary Education*, Max Goote, Amsterdam.

ANNEX 1: SYSTEMS OF PERFORMANCE-BASED PAY IN DIFFERENT COUNTRIES

87. This annex examines systems of performance-based compensation for teachers found in OECD countries. The characteristics of these systems will be outlined, and will then be categorised based on the list of features developed in section 2. This section is limited because the literature is drawn from English language sources, meaning non-English speaking OECD countries are somewhat underrepresented.

Table 4. Summary of Performance-Based Reward Programmes in OECD Countries

Country or State	Type of Teacher Remuneration/Evaluation	
	Outline of the System	Classification of the System
Australia	Teachers who have reached the highest level in the salary scale can apply to become an Advanced Skills Teacher (AST). Teachers are evaluated on performance-based criteria to determine whether they can pass to a new salary scale with three levels or rewards. It was designed to reward experienced classroom teachers, and to discourage them from moving into administrative positions (Ingvarson and Chadbourne, 1997). While several states had abolished the AST system by 2001, it still exists in many Catholic schools, and in the Northern Territory (Waterreus, 2001). However, many states now use performance-based evaluation for movement up the salary scale. For example, in the state of Victoria, progression up the salary scale for government school teachers is dependent on successful performance evaluation. The principal evaluates the teacher based on a pro-forma and decides whether they progress to the next salary scale increment (DEET, 2001).	<p><i>The AST System</i></p> <ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; • Mixed duration depending on the state jurisdiction; • Evaluated on a range of criteria; • Evaluated by classroom observation and peer review; • Mixed availability, as some states introduced a quota, while others did not; and • Created an additional salary scale <p><i>The Victorian System</i></p> <ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; • Permanent reward on progression to the next salary scale increment; • Evaluated on the basis of a pro-forma; • Evaluated by their principal, • All teachers are entitled to apply; and • Supplements the salary scale.
Austria	Although Austria has no system of performance-based rewards, this is currently a topic of debate (Schratz and Resinger, 2001). Teacher salary is currently determined by: the type of employment	Does not contain characteristics of performance-based reward systems.

	contract, the formal teacher workload and experience.	
Belgium (Flemish Community)	Every teacher must be evaluated once every three years on criteria established by their job description. This can be used to reduce pay, or for dismissal.(Devos and Vanderheyden, 2002).	Evaluated individually; Pecuniary sanctions; and All teachers are evaluated.
Czech Republic	No evidence of performance-based reward programmes (Kotásek, Valenta and Brožová, 2001).	Does not display characteristics of performance-based reward systems.
Denmark	Since 1999, teachers' wages have been determined by four factors: a basic wage, a function wage, a qualification wage, and a results wage. The results wage is based on the attainment of quantitative and qualitative results, and operates as an individual performance-based reward (Held, 2001).	<ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; and • Evaluated by multiple criteria.
Finland	No evidence of performance-based incentive programmes (Luukkainen and Eurydice Finland, 2001).	Does not display characteristics of performance-based reward systems.
France	Teacher compensation is dependent on experience and performance-based evaluation. The 'head of the school' (40%) and an outside 'inspectorate' (60%) award the evaluation component. The 'head of school' mark is awarded annually based on the principal's evaluation, but the 'inspectorate' mark for pedagogical competency occurs infrequently, with often ten years between evaluations (Waterreus, 2001). This system is currently under review, and the Monteil Report makes a number of suggestions for future systems of teacher evaluation. These include; <ul style="list-style-type: none"> • "A change in the role and methods of supervision • Evaluation based on an activity report, itself produced periodically by teachers; • Closer coordination of evaluation with in-service teacher training; • The supervision and support of new teachers by a tutor" (Eicher and Chevailler, 2001). 	<ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; • Ascending rewards based on individual performance and experience; • Mixed evaluation methods; • Evaluated by an outside reviewer and a by the principal; • Available to all teachers who successfully complete the evaluation; and • Supplements the existing salary scale.
Germany	The <i>Bundesbesoldungsgesetz</i> creates a link between progression up the salary scale and teacher performance (Jeuthe, 2001). Performance as well as seniority is considered before teachers progress to the next increment Waterreus, 2001).	<ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; • Permanent rewards; and • Is available to all teachers who fulfil the criteria.
Greece	Greece has no formal performance-based compensation system. A system was introduced in 1997 to evaluate teachers' performance, the performance of school units and the effectiveness of the school system. This evaluation only affects initial appointment, and not salary (Doukas and Smyrniotoulou, 2001).	Does not display characteristics of performance-based reward systems.

Ireland	Teachers share a common pay scale, which is not influenced by performance or merit (Coolahan, 2002)	Does not display characteristics of performance-based reward systems.
Korea	Korea has merit-pay for ‘teachers of excellent educational activities’, which provides a hypothetical performance bonus in education. In practice only 10% of the bonus is paid differentially based on teacher evaluation (Kim and Han, 2002), meaning performance evaluation does not affect teacher compensation considerably.	The small portion of revenue that is distributed differentially can be categorised as: <ul style="list-style-type: none"> • Individually evaluated; • Pecuniary rewards; • Used as a bonus; and • Supplements the salary scale
Mexico	Mexico’s Carrera Magisterial Programme is a voluntary, individually-based performance pay system. Participating teachers from primary and secondary schools are subject to an annual global evaluation, with salary increases linked to this evaluation. By 1997, 50% of teachers were participating in this programme (Liang, 1999).	<ul style="list-style-type: none"> • Evaluated individually; • Pecuniary rewards; and • Have annual performance-evaluations.
New Zealand	A system of individual performance-based pay is used. Teachers progress along the salary scale based on performance, experience and formal qualifications. In practice, most teachers will pass the performance criteria. If the budget for teacher salaries is not limited, teachers receive automatic promotions (Waterreus, 2001). Broad performance criteria are determined by the education department, but the application of specific performance indicators is left for each school to determine. Despite this, the assessment process has to include a range of evaluation methods, including classroom observation, self appraisal, and an annual interview (Waterreus, 2001).	<ul style="list-style-type: none"> • Evaluated individually • Pecuniary rewards; • Permanent rewards; • Ascending levels of reward based on experience, qualifications and performance; • Evaluated through a range of techniques including observation, portfolios and self report; • Evaluators include the individual teacher, and the principal; and • Provides al reward to all teachers who fulfil the criteria.
Norway	New evaluation procedures are being considered, along with salary increases. These policy initiatives are not necessarily connected (Bergem, 2001).	Does not display characteristics of performance-based reward systems.
Spain	No evidence of performance-based reward programmes (Esteve, 2001).	Does not display characteristics of performance-based reward systems.
Sweden	Sweden has a small component of informal, individual performance-based evaluation. For the first five years of service, teachers have a yearly pay increase, regardless of performance. Any further pay increases are centrally determined, and schools are free to give individual pay increases, on government and school-based criteria. This potentially constitutes a limited, but internal performance-evaluation. There is little incentive for schools to deny salary increases (Waterreus, 2001). In practice, this means that after teachers progress to the top of the centrally determined salary scale in Sweden, the	Classification under the categories outlined in section 2 is difficult because of the limited component of performance-based rewards in the formal system. However, some characteristics are: <ul style="list-style-type: none"> • Individually evaluated; • Pecuniary rewards; • Permanent rewards; and • Supplements the existing salary scale. Despite this classification, it is

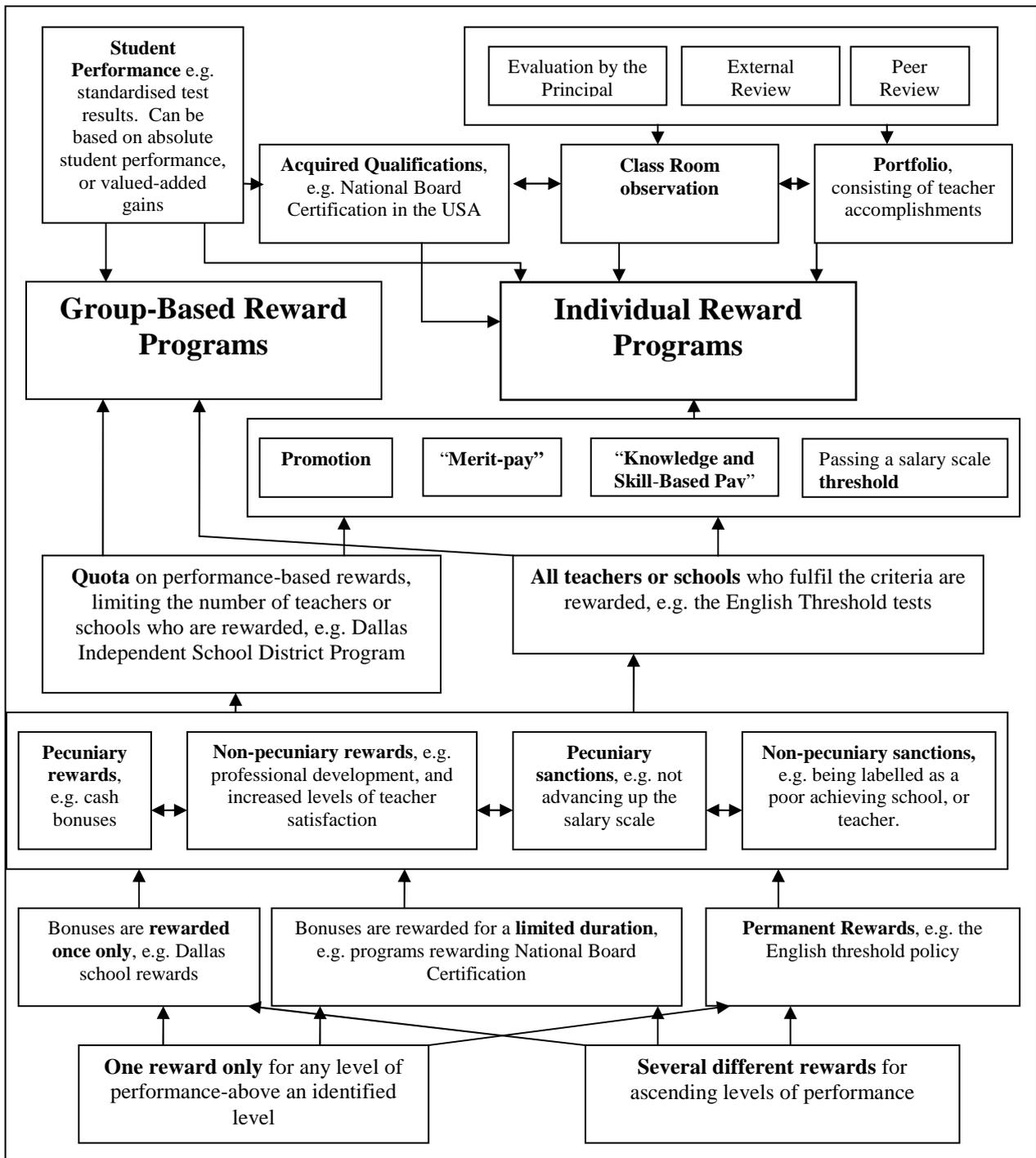
	school determines any further salary increases based on individually established criteria (OECD, 2002), which can incorporate performance-based measures.	unclear how this system can be categorised, given that a range of criteria can be used to measure performance. Potentially, the Swedish system can operate as a merit-pay or knowledge and skill pay system.
The Netherlands	No evidence of performance-based reward programmes (Vossensteyn, 2001).	Does not display characteristics of performance-based reward systems.
United Kingdom (England and Wales)	<i>The Threshold Assessment.</i> Once a teacher progresses to the top of the salary scale, they can take a performance-based test to advance to a new salary scale. There are sixteen criteria that need to be successfully met for transition beyond the threshold (Tomlinson, 2000; Cutler and Waine, 2000).	<ul style="list-style-type: none"> • Individually evaluated; • Pecuniary rewards; • Permanent rewards; • A single reward level allowing transition to a new salary scale; • A large range of criteria evaluated; • Evaluated by external and internal review; and • Is available to a quota of about 25,000 teachers per year.
	<i>Fast tracking.</i> This allows talented teachers to progress quickly through the salary scale. On successful completion of performance-based criteria, fast track teachers can progress two salary scale points per year, instead of the traditional one. This has a significant implication for teacher salary and teacher progress, as time taken to reach the threshold can be hypothetically halved (Tomlinson, 2000). Selected candidates are given a bursary of £5,000, with an expectation that their subsequent performance will increase substantially.	<ul style="list-style-type: none"> • Individually evaluated; • Pecuniary and intrinsic rewards; • Limited duration; • Ascending levels of rewards, and • Supplements the existing salary scale.
	<i>Advanced Schools Teachers.</i> These teachers have a special salary spine, and are required to exhibit outstanding skills based on excellent student outcomes, excellent subject knowledge, excellent ability to plan, excellent ability to assess, and excellent ability to support other teachers (Tomlinson, 2000).	<ul style="list-style-type: none"> • Individually evaluated; • Pecuniary rewards; • Centrally determined, wide ranging criteria based on demonstrated skills and knowledge; • Ascending levels of reward; and • Has a new salary scale.
United States- Douglas County, Colorado	Teacher salary consists of a base pay, and several bonuses: <ul style="list-style-type: none"> • Knowledge based pay; • Performance-based pay measured on experience and evaluation; • Outstanding teacher awards; • Skill blocks; • Group incentive pay, and • Site based responsibility pay. (Tomlinson, 2000, 	This pay structure is varied and encompasses several performance-based reward strategies. This system has components of merit-pay, school-based pay and knowledge and skill pay. There are, multiple reward levels, multiple criteria for evaluation, and multiple evaluators.

	295-96).	
United States- Kentucky	<p>Kentucky has a system of school based performance awards. Teachers are provided with salary bonuses based on student performance (Tomlinson, 2000). Schools are measured on an index of student assessment scores covering seven academic areas (reading, writing, math, science, social studies, arts/humanities, and vocational/ practical living) and school-level indicators including drop-out rates, school attendance and transition to a successful adult life. Poorly performing schools are allocated additional funding, labelled as a 'school in decline' or a 'school in crisis', and have a 'Distinguished Educator' assigned to improve student outcomes (Kelley, Heneman and Milanowski, 2002). Kentucky also has a performance-based teacher-licensing system (Odden, 2000b)</p>	<ul style="list-style-type: none"> • Group-based evaluation; • Pecuniary rewards, and non-pecuniary sanctions; • Rewards allocated every two years as a bonus; • Several levels of rewards and sanctions based upon whether schools reach their performance goals; • Evaluated on the basis of 'added value' to student test scores; • Evaluation carried out by external review; • All schools who reach their performance target are rewarded; and • Supplements the salary scale.
United States- North Carolina	<p>The district of Charlotte-Mecklenburg has a school based performance award programme which evaluates student achievement in nine areas: reading, writing, math, social studies, primary grade readiness, higher level course enrolment, end-of-course subject mastery, attendance and drop-out rates (Heneman and Milanowski, 1999). High achieving schools are given maintenance goals; other schools are given improvement goals. Teachers earn a bonus of between \$750 and \$1000 (Kelley, Heneman and Milanowski, 2002).</p>	<ul style="list-style-type: none"> • Group-based evaluation; • Pecuniary rewards; • Rewards distributed annually as a bonus; • Single level of reward; • Evaluated on the basis of 'added-value' to student performance; • Evaluated by an external review; • Rewarded to all teachers who fulfil the criteria; and • Supplements the existing salary scale.
United States- Ohio Cincinnati	<p>Performance-based teacher-licensing systems (Odden, 2000b). Cincinnati is introducing a plan that will include knowledge and skills based salary bonus, and a school-wide bonus for student outcomes (Odden, 2001).</p>	
United States- South Carolina	<p>South Carolina has a school-based performance award programme (Heneman and Milanowski, 1999). Student academic performance, taking into account past academic results, is used to determine school effectiveness. Schools are placed in one of four categories depending on socio-economic indicators. Within each group the top 25% of schools and the top 25% of all schools who 'exceeded expectations' get a bonus that distributed to the staff (Clotfelter and Ladd, 1996).</p>	<ul style="list-style-type: none"> • Group-based evaluation,; • Pecuniary rewards; • Annually rewarded bonuses; • Single reward level; • Evaluated on the basis of 'added value' to student outcomes; • Evaluated by external review; • A quota of schools are rewarded; and

		<ul style="list-style-type: none"> • Supplements the salary scale.
United States-Texas	<p>The Dallas Independent School District uses school-based performance awards. Success is determined by a complex student ‘added value’ test scores, taking into account the racial and socio-economic status of students. The top 20% of schools are awarded bonuses of \$1000 US per teacher per year (Waterreus, 2001), and the next 30% of schools are given \$425 US per teacher (Clotfelter and Ladd, 1996).</p>	<ul style="list-style-type: none"> • Group-based evaluation; • Pecuniary rewards; • Annual bonus; • Two reward levels based upon the magnitude of the mean ‘added value’ of student test scores; • Determined by external review; • Rewards a quota of schools; and • Supplements the salary scale.

88. In summary, many of the OECD countries examined have some components of performance-based rewards for teachers. The most common programmes are variations of individual merit-pay and group-based rewards, with financial rewards being used to supplement an existing salary scale. No system examined based teacher salary completely on performance evaluation, with experience and formal qualifications remaining strong determinants of teacher compensation. This suggests that most of the OECD countries who have adopted performance-based compensation programmes have introduced these measures incrementally and conservatively.

ANNEX 2: FEATURES OF PERFORMANCE-BASED SYSTEMS.



ANNEX 3: THE METHODOLOGY AND CONCLUSIONS OF THE KENTUCKY AND CHARLOTTE-MECKLENBURG STUDIES OF TEACHER ATTITUDES (1998-2002)

Methodology

89. Heneman, Milanowski and Kelley conducted three studies on the same data set between 1998 and 2002 which examined teacher attitudes to school-based reward programmes systems in Kentucky and Charlotte-Mecklenburg. They used a mixed method, qualitative and quantitative study. The qualitative results were triangulated with literature on education and motivation to produce the quantitative survey. The qualitative research consisted of semi-structured, open-ended interviews with teachers and principals. Geographic and socio-economic variation was considered in the qualitative surveys. Survey data was collected within one evaluation period, so teachers' expectations could be tested against actual student outcomes

90. The quantitative component consisted of a survey that was tailored to each of the two districts' performance-based pay programmes. These surveys were administered in 1997. The response rate was 39% in Charlotte-Mecklenburg and 30.9% in Kentucky giving a sample size of 1,150 and 1,750 respectively. The Charlotte-Mecklenburg sample somewhat over-represented elementary teachers, but covered geographical and socio-economic areas well. The Kentucky sample was generally representative of the teaching population. By aggregating teacher motivation, they attempted to compare schools composed of motivated teachers to non-motivated teachers as a means of determining a connection between teacher motivation and student outcome. The purpose of the experiment was therefore twofold. First, they were assessing how student-based reward programmes affected teacher motivation. Second, they were attempting to find out how teacher motivation affected student outcome.

Authors' Conclusions

91. This study found that teachers considered 'goal attainment rewards', including payment bonuses for attaining rewards and the public recognition from attaining these rewards, as very important. 'Learning' outcomes were also important for the teachers surveyed (Heneman and Milanowski, 1999), including seeing and being responsible for improvements in student performance and working collaboratively with peers. This may be rhetoric, as the actual evidence supports the hypothesis that teachers are motivated by financial rewards. While teachers in the Kentucky school-based rewards system rated 'school improvements' as more beneficial than extra salaries in the research conducted, when teachers allocated programme rewards, 98% voted to use some or all of the rewards as a salary bonus (Kelley, 1999). There was some concern in relation to goal clarity in school-wide programmes, meaning there would be some concern about the motivational capacity of schools (Kelley, Heneman and Milanowski, 2002). Despite this concern, Kelley, Heneman and Milanowski (2002) concluded that teachers knew about school goals, and were committed to them. Other analysts have taken these results to show that school-based performance awards also support school improvement efforts (see, for example, Tomlinson, 2000), as there are motivation impacts from both positive and negative circumstances.

92. Teachers were concerned that they would not be paid the bonus upon successful completion of school-wide criteria. “A striking finding in both the qualitative and quantitative data from both programme sites was the low perceived probability that the bonus would actually be paid when school goals were met.” (Kelley, Heneman and Milanowski, 2002)

93. In conclusion, Kelley, Heneman and Milanowski (2002) argued, that in all models, the higher the average teacher expectancy of student results, the greater the school-based evaluation outcomes. These analyses, it is argued, shows that between-school differences in teacher expectancy are a predictor of schools’ future outcomes. “Overall, we believe that our result[s] suggest that SBPA programs have the potential to contribute to motivating teachers to improve student achievement. However, the potential was not fully realized in these first-generation programs. The implication is that program designers should set goals that are perceived by teachers as achievable and should develop coherent systematic approaches to support goal attainment...These considerations suggest that providing a bonus may be the least problematic aspect of a SBPA program, whereas setting realistic goals, maximizing perceptions that achieving the goals will lead to positive outcomes, minimizing stress reactions, and providing enabling conditions are where effort and attention need to be focused. Motivational impact is not guaranteed simply by promising a bonus.”

ANNEX 4: BALLOU AND PODGURSKY'S STUDY OF TEACHER ATTITUDE (1993)

Methodology

94. Ballou and Podgursky drew on data from the 1987-88 Schools and Staffing survey (SASS) in an attempt to establish teachers' views of merit pay. This survey randomly sampled 9,300 public schools and 3,500 private schools, with approximately 56,000 public and 11,500 private school teachers from these schools surveyed. The response rate was 83% among public school teachers and 73% among private school teachers. These data distinguished between a wide range of teacher characteristics, training, assignments and attitudes. This study drew on the 'Incentives and Compensation' section, which asked teachers their opinions of particular types of compensation strategies. Ballou and Podgursky (1993) used data measuring teachers' attitude towards pay incentives derived from this survey. The surveys asked teachers to rate on a four-point scale (strongly favor, mildly favor, mildly oppose and strongly oppose) their attitude to a number of incentive pay schemes. For the results of this survey, refer to Table 5. Their analysis distinguished between public, Catholic, other religious and private non-religious schools.

Table 5. Teacher Attitudes Towards Various Types of Incentive Pay

'For each of the following incentives, please indicate whether you favor or oppose the incentive, and whether you now receive the incentive	Strongly Favor (%)	Mildly Favor (%)	Mildly Oppose (%)	Strongly Oppose (%)	Receive Incentive (%)
1. Additional pay for assuming additional responsibilities as a master or mentor teacher (e.g. supervising new teachers)	58.8	28.8	5.7	6.7	9.1
2. Additional pay for teaching in a shortage field (e.g. math, science)	24.5	29.2	20.4	25.9	1.3
3. Additional pay for teaching in a high priority situation (e.g. in an inner-city school)	41.0	36.2	12.0	10.7	1.3
4 Salary increases as part of a career ladder in which teachers progress through several promotional levels based on their performance	40.9	30.4	11.7	16.9	16.3
5 A merit pay bonus for exceptional performance in a given year	28.8	26.4	16.0	28.8	2.7
6 A school wide bonus for all teachers in a school that shows exceptional performance or improvement in a given year	34.7	30.5	14.2	20.6	2.8

(Source: Ballou and Podgursky, 1993).

95. They used a linear model to analyse multiple variables. The variables tested were: teaching experience, whether the teachers surveyed were in a merit-pay system at the time of survey, the level of education (a master's degree or better), gender, race, starting pay, community type, socio-economic status of the student population as measured by the percentage eligible for free lunches, perceived class ability,

the type of school they were from, and some locations of school districts that have used merit-pay systems. The sample size for this analysis was reduced to 33,865 because of incomplete data.

Critique and Discussion

96. The conclusions of Ballou and Podgursky are somewhat limited in their application to current performance-based reward systems. First, their conclusions are based on old data. The extent to which these attitudes can now be generalised is unknown. More recent research would be needed, especially given the growth of performance-related pay as an issue since this data was gathered. Teacher attitudes may have changed significantly since this survey given the recent level of debate surrounding performance-related pay.

97. Second, these conclusions are limited in the extent they can be generalised to other performance-based pay system because the question format of the survey does not reflect the reality of contemporary programmes. The survey asked teachers about their attitudes to 'additional pay' so it is fair to assume surveyed teachers would have concluded pay would supplement current income. This would have the effect of over-emphasising support for merit-pay programmes for teachers, since these programmes rarely supplement overall teacher income. While Ballou and Podgursky (1993) argue their findings would remain robust even given a change of wording or context, this is doubtful given the strong teacher opposition to merit-pay reported elsewhere (see Kelley, 1999). Even though the statistical analysis proved to be appropriate, there is considerable doubt whether these results on teacher attitudes would be replicated if different worded questions were used.

ANNEX 5: THE DALLAS SCHOOL ACCOUNTABILITY AND INCENTIVE PROGRAMME (CLOTFELTER AND LADD, 1996; LADD, 1999).

Background

98. Schools in the Dallas Independent Schools District are evaluated on the basis of their 'added-value' of students' test scores. A two-stage regression is carried out on data considering examining student test scores. The first stage examines individual test scores from the current and prior year, controlling for a host of socio-economic indicators including race, gender and eligibility for free lunches. In the second stage, regressed data from the start of the year was compared to regressed data from the end of year (Ladd, 1999). The 'added-value' of each student was aggregated and averaged across each school to give a mean 'added-value' for all students enrolled in the school during the year. This complex procedure was carried out to attempt to ensure that all schools had an equal chance of winning an award, not just the affluent and culturally homogenous schools. , Dallas administered multiple measures of student outcomes, including two tests given annually, a state test, and a national test, to try to overcome criticisms of a narrowed curriculum (See section 4.5). Student attendance and drop out rates were used to supplement these measures.

Methodology

99. Ladd (1999) used the TAAS to measure students' pass rates between 1990-91 and 1994-95 in reading and mathematics. The 1990-91 data covered the year before the scheme was introduced, and the 1994-95 data covered the most recent data available at the time of writing. The use of pass rates are an inaccurate method, but, it is argued, have some merit because of relatively low pass rates of early tests. Moreover, socio-economic indicators are controlled for, meaning racial, gender and affluence should have limited affect on the test outcomes. Average test score gains are compared to five other large Texas cities to attempt to control for student gains not caused by programme characteristics. The control schools had different racial characteristics, but were all urban schools with the same state-wide test requirements. The third grade students could not be analysed because of incomplete data. Ladd (1999) relied on grade 7 student data.

100. For math and reading, analysis was done for the group as a whole, and by children from economically disadvantaged homes, Hispanics, blacks and whites. Student attendance was examined relative to other Texas cities, as were dropout rates and principal turnover. A complex statistical analysis measured the pass rates of Dallas students, controlling for city and socio-economic indicators. All equations were weighted either by the number of students in the school, or the number of students in the subgroup (Ladd, 1999). A second analysis was completed, controlling for the ratios of pupils to teachers and the number of teachers with more than five years experience.

101. The school principal turnover was also examined, to attempt to examine whether there were administrative changes in schools which used group-based performance reward systems. Clotfelter and

Ladd (1996) conducted an earlier attempt to evaluate the performance of this programme, and their analysis considered the conceptual difficulties in designing, implementing and evaluating these programmes. For this evaluation, only the conclusions of Ladd (1999) will be examined in depth.

The author's conclusions

102. For all four years, pass rates increased more in Dallas than they did in any of the control groups. However, there were gains before the programme was fully implemented. This phenomenon could reflect a movement back to the state mean, or the positive effects of publicity of the programme (Ladd, 1999). One possible positive conclusion is that Dallas schools maintained a consistent level of advancement even when the novelty of the programme diminished, meaning the programme was successful. Further evidence of increased student outcomes was apparent when the other cities were factored into the analysis. Dallas had increased student outcomes in comparison to all the control cities, except for one small district in one city that undertook reform comparable to Dallas. The relative pass rates of students of economic disadvantage, blacks, Hispanics, and students with limited English proficiency is lower as the percentage of each of these categories increase as a proportion of the grade (Ladd, 1999). The less transient the school population, the higher the average increase in pass rates.

103. Results based on race showed that in reading and math, Hispanic and white students in Dallas increased significantly more than Hispanic and white students relative to the baseline year and the control cities. There was no evidence that black students improved (Ladd, 1999). Differences in the number of teachers with at least five years experience, and the average number of pupils per teachers produced no significant results, with the exception of Hispanic students in math, who had increased pass rates the greater the proportion of teachers with over five years experience.

104. These results were complicated by the introduction by the Texas government of an aggressive policy aimed at increasing the resources of low-performing schools during the surveyed period. To overcome this, an analysis was done excluding 25% of schools with the highest proportion of students of economic disadvantage. The results followed the same pattern of the previous data, with significant increases in the scores white and Hispanic, but not black students (Ladd, 1999). The evaluation of the programme depends on the interpretation of the first year results. If all the measured gains were true programme impacts, then increases in student outcomes were in the order of 10 to 20% relative to the state average. If the gain in 1992 were because of a return to the mean, the programme impacts can be viewed less favourably (Ladd, 1999).

105. The dropout rate of high-school student decreased more for Dallas than any other city in comparison across all years, with two years (out of three surveyed) showing significance. Turnover rates for principals jumped dramatically after the first year of the programme. It appears that the school districts were more willing to change principals than in the past. Qualitative research suggests that these principals were either demoted or fired, so this does not represent a shuffling of poorly performing principals (Ladd, 1999). Ladd (1999) concludes that the turnover of principals is desirable as it represents district administration being more proactive in making principals accountable to student results.

106. Overall, Ladd (1999) concludes there is reason to be cautiously optimistic about performance-based reward programmes. The seventh grade results are positive and large, but only significant for Hispanic and white students. Other positive results are the reduction in dropout rates and the high turnover of principals. However, more research needs to be to examine the ranges of student increases in scores. Further research could be conducted to make a more accurate cost-benefit analysis of performance-based reward systems.