

---

# Teacher incentives in South Africa: a theoretical investigation of the possibilities

PAULA ARMSTRONG

---

Stellenbosch Economic Working Papers: 07/14

KEYWORDS: INCENTIVES, TEACHERS  
JEL: I2; J5

PAULA ARMSTRONG  
DEPARTMENT OF ECONOMICS  
UNIVERSITY OF STELLENBOSCH  
PRIVATE BAG X1, 7602  
MATIELAND, SOUTH AFRICA  
E-MAIL: PAULA@SUN.AC.ZA



UNIVERSITEIT  
STELLENBOSCH  
UNIVERSITY



A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE  
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

---

# Teacher incentives in South Africa: a theoretical investigation of the possibilities

PAULA ARMSTRONG

---

## ABSTRACT

---

This paper investigates different theoretical models of incentives for teachers in education. It highlights key characteristics likely to render incentives successful in encouraging productive behaviour, provides evidence of where these systems have been successfully and unsuccessfully implemented internationally and the likelihood of successful implementation of teacher incentive programmes in South Africa.

Keywords: incentives, teachers

JEL codes: I2; J5

## 1. INTRODUCTION: TEACHER QUALITY IN SOUTH AFRICA AND THE POSSIBILITY OF INCENTIVES

For the country's level of development and the proportion of the budget that is spent on education, South Africa perform substantially below what is expected of it in terms of educational performance (National Education Evaluation and Development Unit, 2012: 13). Performance amongst learners in this country is markedly worse than amongst learners in countries which are poorer than South Africa. In the third round of international tests conducted by the Southern African Consortium for the Measurement of Educational Quality (SACMEQ) in 2007, South Africa performed worse than most countries (a total of 14) in both mathematics and language (Spaull, 2011: 43), managing to outperform Lesotho, Mozambique, Uganda, Malawi and Zambia, leaving South Africa lying 10<sup>th</sup> out of a total of 15 countries). Educational quality in South Africa therefore leaves much to be desired. .

Teacher quality and low levels of teacher effort are often cited as major drivers of South Africa's education crisis. The question then is why South African teachers are performing as poorly as they do. NEEDU's National Report of 2012 suggests that teachers are unable to ensure high quality education for students either because they *won't* or because they *can't* (NEEDU, 2012: 20). In the case that teachers and schools *won't* provide quality education, poor performance is the result of lack of discipline amongst staff member and any remedial action should focus on changing behaviour. In the case that teachers *can't* provide quality education, then intervention should focus on improving and enhancing the knowledge base of teachers to equip them with the skills necessary to ensuring that quality teaching and learning takes place in classrooms (NEEDU, 2012: 20). This paper focuses on the potential to change teacher behaviour in a way that enhances student performance.

Many argue that the lack of effort amongst teachers is the greatest hindrance to the development of student skills in South Africa, suggesting that the appropriate policy response should be directed towards designing attractive incentives for teachers. High levels of absenteeism from school and from classrooms, poor lesson preparation and lack of interest in the progress of learners are key signs that teacher effort is critically low in South Africa, and it is often reported that such low levels of effort result from weak incentive systems.

This problem is not particular to South African. International research findings indicate that an important source of service delivery failure in education is the teacher. The principal-agent problem in education is complicated to a large extent because of the nature of service

provision in education (Bruns, Filmer and Patrinos, 2011: 10-11). The interaction between teachers and learners is discretionary (in that it is the teacher's own judgement that will determine what is taught and how it is taught), variable (in that a teacher is required to tailor their teachings style to a diverse group of students) and requires repeated interaction. It is exceptionally difficult to define beforehand what type of behaviour and actions teachers must take. It is challenging therefore to specify a contract according to which they are expected to conduct themselves as well as to monitor their behaviour (Green, 2011).

Many advocate the implementation of incentives for teachers as a possible channel through which to enhance student ability and skills development. An important motivation for implementing incentives for teachers is that they explicitly state the results that are valued by the education authority and therefore of teachers, thereby enhancing their accountability. The agreement between authorities and schools is enhanced because expectations with regard to service provision are explicitly specified. Bruns et. al. (2011: 19-20) discuss pay-for-performance as a possible avenue to incentivise teachers and ensure accountability. Pay-for-performance largely leaves the teacher salary scale untouched while creating incentives at the margin. Unlike salaries, this kind of incentive pay rewards teachers for what they actually do or achieve during a specified period, rather than for what their qualifications and training suggest they are capable of doing. The correlation between the observable characteristics of teachers and the level of achievement attained by their students is weak. (Bruns et. al., 2011: 20).

This paper examines the theoretical characteristics of incentive systems, detailing the potential for distortion as well as the conditions under which peer pressure is effective as an accountability- and performance-enhancing force. Section 2 provides a brief overview of incentives in economic theory and specifically in education. Section 3 presents theoretical models of incentives systems and investigates the characteristics of such systems while section 4 draws up a brief framework from the theoretical models presented in section 3 according to which international incentive systems are evaluated in section 5. Section 6 analyses the potential for the introduction of incentive systems in South Africa and explores incentives inherent to the South African teaching profession.

## 2. ECONOMIC THEORY AND INCENTIVES

This section is as an introduction to thinking around incentives in an economic context with specific reference to the case of education. The objective is to sketch the broad framework in which the theoretical models of section 2 are situated.

Research on the use of incentives in organizations generally makes use of an economic framework which analyses the difference in the objectives of different individuals who make up an economic organization. Specifically, the objective of the organization as a whole (which is characterized by the objective of the owner or “principal” of the organization) is contrasted with the objectives of workers within the organization, i.e. the objectives of the “agent”. Differences in these objectives imply that workers will not necessarily behave in a way that maximizes the objective of the organization as a whole or with the principal’s objectives for the organization, therefore rendering it less productive and worsening the situation for workers in the long run by diminishing either productivity or employment. Incentives may therefore be used to encourage workers to work towards the objectives of the organization as a whole.

A classic example of such incentives is what is referred to as a “piece rate” – this is a rate paid to workers based on their level of productivity or the number of “pieces” they produce (hence the name “piece” rate) as opposed to a salary paid based on hours. Increased profits will result from workers producing more per hour, so incentivizing this production directly may have a greater positive impact on productivity than would be the case if worker were paid by the hour (Hout and Elliott, 2011: 21).

The efficiency of incentives is however very much dependent on the social relations that evolve around “piece work”. Complexities beyond simply paying for productivity instead of paying hourly arise when dealing with incentives, and these require some understanding of how incentives work in different contexts and for different people. Five complexities requiring attention are

1. Finding performance measures
2. Different incentive effects on different people
3. Uncertainty and control
4. Effects of working in groups
5. Weighing the costs and benefits of incentives.

These are each considered in turn below.

## **2.1 Finding Performance Measures to Use with Incentives**

In most jobs, output cannot be counted in any meaningful way, making it difficult to measure the contribution of each worker. Often the qualitative aspects of the job performed are more important than the aspects that can be quantified. The difficulties inherent in quantifying incentives are a major constraint to providing them, and the gap between the measures available for the measurement of incentives and the actual value of the output has important implications for the operation of incentives. “Objective” measures of performance obviously focus on the quantifiable aspect of the job at the expense of qualitative aspects, given the difficulty in measuring these. It is not surprising then that when incentives are attached to the quantifiable performance measures, workers focus on these aspects of the job and neglect those that do not affect the performance measure. This is problematic when performance measures are not closely aligned with the true value of the work being done. The performance measures are said to be “distorted” when they result in behaviour that is detrimental to the true value of the organization or simply fail to enhance the value to the firm (Hout and Elliott, 2011: 22-23).

Defining the underlying goals that performance measures should reflect can be highly problematic. In the case of education, schools are required to ensure that students meet some minimal standard of academic performance, at the same time as being responsible for developing cognitive skills, physical and emotional development, preparing students to enter the workplace and society and students’ health and safety. Whilst these goals are not inconsistent they compete for limited resources, forcing schools to make difficult trade-offs. Ideally each goal would have one performance measure, but this is simply not practical to implement and so further trade-offs in the selection of performance criteria are needed (Dixit, 2002: 712). Once performance measures have been selected, further decisions need to be made regarding how each performance measure will be weighted in the overall incentive scheme. Agreeing on a performance measure to be implemented is therefore a considerable challenge.

A theoretical analysis (Baker, 2002 discussed later) shows that more important than correlation between the performance measure and the value of the organization is whether behaviour that improves the performance measure also enhances genuine value; in the case of education, this would be genuine learning amongst students. This distinction is important

because a performance measure may be correlated with a wide range of outcomes absent incentives (i.e. high levels of the performance measure is correlated with high overall performance). However, once incentives are attached to the performance measure, behaviour that increases the performance measure at the margin may not enhance overall performance at all. This is known as “gaming” the system (Koretz, 2008: 24).

In the field of education, the strength of incentives currently in place is not straight forward. Studies of cheating by teachers show that even when incentives are small, some teachers react quite strongly in a distorting way while others genuinely increase their level of effort (Hout and Elliot, 2011: 26). As mentioned before and as discussed later, the behavioural impact of a given incentive or set of incentives will cause different people to behave in different ways and will depend largely on the context in which individuals work.

Incentives systems therefore inevitably have some level of distortion in them and so the objective of evaluating incentive systems is not to ascertain whether incentives exist, but rather whether the costs involved in minimizing the distortion are outweighed by the benefits of implementing the incentive system.

## **2.2 Different Effects of Incentives on Different People**

Economic theory postulates that different people are likely to respond differently to incentive structures. Amongst people for whom the target is easily attainable, incentives are likely to result in greater levels of effort relative to those for whom the target is difficult to achieve (Lazear, 2000: 186). This means that people who are able to reach the target are likely to be attracted to and remain at the organization while those who are likely to be unsuccessful will become discouraged and leave. This may enhance productivity since low productivity workers will leave the organization and be replaced by more productive workers. An application of this to the area of teaching by Lazear (2000) applies a model in which teachers have different abilities to raise student test scores and produces the result that some teachers increase their effort while others leave the teaching profession. The fact that the teachers that left the profession are the teachers who are less able to raise test scores would be an enhancement in the overall effectiveness of teaching over time.

A variation of this model is one in which teachers respond to incentives by either increasing effort or by increasing effort in test preparation (assuming test scores are the performance measures upon which the incentives are based). The performance for teachers engaging in

both forms of behaviour will increase student test scores, but actual learning will only take place amongst students taught by teachers in the first group. This type of distortion is discussed in detail in section 2.

### **2.3 Effects of Uncertainty and Control in Providing Incentives**

In almost all jobs, a worker's actual performance and the performance measure may be affected by factors outside of the worker's control. In the case of pay-for-performance incentives, it is likely that payoffs received by workers as part of the incentive scheme will depend both on the workers' effort and factors beyond their control. As a result, in situations where their pay is going to be affected by factors beyond their control, they will require compensation for this uncertainty in the form of higher levels of pay (Hout and Elliott, 2011: 27).

Theoretical analysis shows that optimal incentive schemes will allocate less weight to measures that are more dependent on factors beyond the worker's control (Baker, 2002: 3) since these incentive schemes require greater average pay levels in order to compensate workers for the greater level of uncertainty in their pay relative to what they might have received in another job (without incentives). Ultimately then, although productivity is enhanced by the use of incentives, the higher level of average pay that workers require is costly (Baker, 2002: 4-5).

There are numerous factors beyond the control of schools and teachers that impact on student learning. In particular, characteristics of students' home environments result substantial variation in the performance between students and over time – one of the primary reason for the strong opposition to incentives based on student performance amongst teachers.

### **2.4 Effects of Groups in Providing Incentives**

Trade-offs in performance available performance measures are inevitable in environments where the value of work done depends on cooperation amongst workers. Performance measures that consider only individual worker production will neglect the contribution made by individual workers to the productivity of the entire team, while measures of the productivity of the team will provide a vague and possibly inaccurate indication of individual workers performance because they are determined by the contributions of the entire team. Whether team-level or individual-level measures of performance should be incentivized is



dependent on the relative importance of cooperation and on the uncertainty inherent in such performance measures (Hout and Elliot, 2011: 30).

Education is a prime example of a field requiring cooperation amongst workers. Research in fields other than economics points out the importance of understanding how schools function as organizations i.e. understanding the extent of cooperation required amongst teachers. For example, sociological research looks at understanding incentives in a setting where consequences are not explicitly defined, investigating how incentives are communicated in an informal way between workers (Meyer and Rowan in Hout and Elliott, 2011: 30). The combination of economic and other research indicates the importance of considering the impact of incentives beyond those directly relevant to the individuals within an organization. Furthermore, the degree to which work is done jointly and the degree to which the direct effect of incentives will be transmitted to other members of the organization must be considered (Hout and Elliott, 2011: 31).

An organizational structure as complicated as a school requires one to consider the different roles and interactions between individuals working at the school. There may well be value in incentivizing individuals relatively high up in the hierarchy who have the capability to transmit the incentives in ways that result in cooperative behaviour (or at least encourage such behaviour). Importantly, the extent to which the diffusion of incentives is imperfect requires that the behaviour which one is trying to adjust, as well as the ability of members in different roles within the organization to adjust the behaviour, be considered carefully.

## **2.5 Weighing the Costs and Benefits of Incentives**

The fact that distortion occurs with performance incentives does not mean that they should not be used at all. Distortion will mean that it is more difficult to measure the benefits associated with incentives and that some parts of the organization may function less efficiently than they might have in the absence of incentives. It may however still be true that the benefits resulting from the implementation of incentives are greater than the costs. The improvement in cognitive skills is associated with improvements in income distribution, individual earnings and economic growth (Hanushek and Woessman, 2008) and an incentive system that produces true learning gains may well produce a sizeable net benefit. However, unless one accounts for the level of distortion in such test-based incentives the size of the benefit is likely to be grossly exaggerated.

Incentives can therefore prove useful in the context of education, despite the problems inherent in their implementation. Section 3 provides a theoretical analysis of incentives in the teaching profession.

### **3. THEORETICAL MODELS OF TEACHER INCENTIVES**

Section 2 placed incentives in the context of Economic theory and discussed some of the complexities involved in making use of incentives. This section looks at particular models of incentives in the context of teaching.

The payment of incentives on the basis of inputs or outputs is a central question in the literature on teacher incentives. The risk inherent in rewarding workers on the basis of what they produce means that output-based incentives are not always fair. As discussed above, this characteristic of output-based pay is particularly relevant in education. In some cases therefore, it may be necessary and preferable to reward workers on the basis of input. Lazear (2003) explains input- and output-based pay.

#### **3.1.1 Output-based Pay**

The most important objective in an education system is the education of the population so as to ensure productivity, allowing individuals to generate skills and thereby an earning to sustain themselves and in turn generate economic growth. It may therefore be said that generating and developing earnings capacity is the central objective of education (Lazear, 2003: 183).

Earnings, however, are only observed sometime after individuals have left the schooling system rendering it impossible to determine teacher pay on this basis. As a result, student achievement test scores are usually used as proxies.

When test scores are used as the basis for teacher pay, it is difficult to distinguish the increases in test scores resulting from teacher efforts and those resulting from the activities of others. Furthermore, improved test scores may result from distortive behaviour on the part of teachers as opposed to genuine enhancement in effort levels.

#### **3.1.2 Input-based Pay**

The advantage of paying workers based on inputs is that it removes the risk inherent in output-based pay, therefore discouraging teachers from focussing exclusively on the performance metric according to which they are paid.

Inputs-based pay may be seen as a solution to the problem of “teaching to the test”, or distortion in the context of the incentives literature (Baker, 2000; Holmstrom and Milgrom, 1991). In the extreme case, a perfect measure of the disutility of working<sup>1</sup> would be used as the measure according to which workers were paid. In this case, there would be no incentive to focus on one area of the curriculum at the expense of others. Teachers would receive full compensation for their efforts and would likely be willing to do what is in the best interest of their students, regardless of the level of effort required. Using disutility as the basis for compensation would ensure that teachers remained indifferent to which areas of the curriculum they emphasise and hopefully inducing them to make sure that all areas are covered. This results from the fact that providing payment on the basis of disutility ensures that they receive enough to compensate them for teaching the “disagreeable” subjects.

Inputs-based pay works well insofar as hours worked are a proxy for the disutility of teaching. However, when teachers care about what activities they engage in while teaching, inputs-based pay will no longer be effective at eliciting sufficient effort from teachers. The non-teaching labour market deals with this problem to some extent by compensating different occupations differently. Lazear (2003: 195) explains that “[institutions are prevented from] hiring professors of accounting as the same wage than can attract professors of organisational behaviour” due to the fact that “[the wage] difference is compensation for perhaps less pleasant or more difficult work”.

### **3.1.3 What works better?**

The literature on teacher incentives favours output-based pay for two reasons. The first advantage (the “informational” part) is that output-based incentives clearly signal to teachers what is valued and required. The second advantage, (the “alignment” part) ensures that the objectives of teachers closely parallel those of society as a whole. It may well be the case that teachers work hard, but there may be disagreement about which areas of the curriculum are important. By tying teacher compensation to an agreed-upon metric, it is possible to ensure that the education provided by teachers results in the accumulation of agreed-upon skills

---

<sup>1</sup> It is assumed that different aspects of teaching or teaching different parts of the curriculum have different disutilities.

amongst learners (Lazear, 2003: 209). Finally, there may be some divergence between the preferences of teachers and the best interests of their students. Lazear (2003: 182-183) gives the example of teachers deciding against giving assignments because of burden of grading, even though they are fully aware that their students would benefit from completing the assignments. Therefore compensation based on students' performance on the assignment may induce teachers to behave in a way that is beneficial to their students.

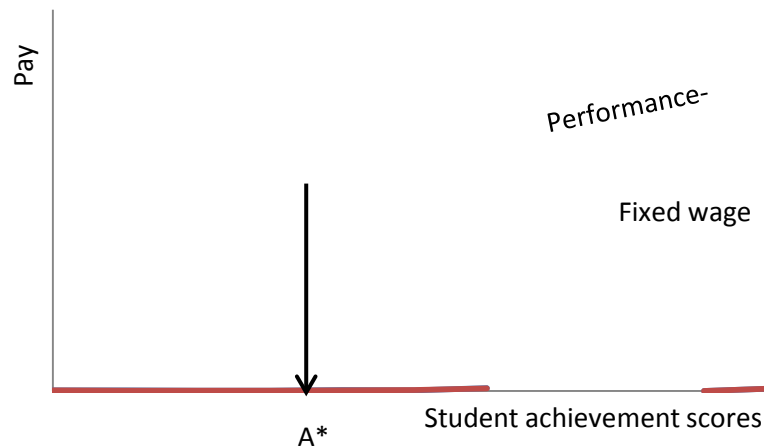
If incentivising teacher performance is thought to be effective in enhancing the quality of education, then it is important to understand the mechanism through which the improvement is likely to occur. Section 2.2 explains two possible avenues of influence through which teacher incentives are likely to impact on student performance – through an “incentive effect” and a “sorting effect”.

### **3.2 Sorting versus Incentive Effects and the Likelihood of Success**

The “selection” aspect of compensation is arguably the most important aspect in the context of education. The individuals attracted by various compensation schemes determine the quality of individuals within a profession. It is widely accepted that teacher quality is critical to quality education and so the question of who is likely to enter the profession is crucial.

When someone is faced with an incentive based on student performance, they must improve student performance in order to be rewarded. Sorting refers to using incentives in order to attract individuals who are better able to improve test scores to the teaching profession, resulting in a teaching force that is better able to improve student performance. Payment based on output therefore attracts individuals who are best able to enhance student performance and may discourage weak teachers from continuing in the profession given the relatively lower levels of pay they will receive as a result of their inability to improve student performance. Sorting therefore results in better teachers replacing less able teachers. This differs from the effect of incentives; in the case of incentives the individual is not replaced but rather adjusts their behaviour (Lazear, 2003: 187).

**FIGURE 1: Incentive and sorting effects**



Source: Lazear, 2003

Figure 1 illustrates the difference between the incentive and sorting effects of compensation schemes. The first pay schedule (fixed wage) is one according to which teachers are paid a fixed salary regardless of the level of performance of their students. The second (performance-based pay) increases teacher wages based on student performance. Teachers able to raise student performance above  $A^*$  therefore do better on a performance-based schedule while those unable to do so will do better on a fixed wage schedule. Pay based on student performance therefore favours teachers who are able to increase student performance relative to those who are unable to do so.

So is teacher quality and teacher productivity driven primarily by teachers' innate ability or does the level of effort put into teaching influence the level of productivity of teachers? Neal (2011) discusses the importance of distinguishing whether teacher quality is dependent on teacher effort, teacher ability or a combination of both. He maintains that it is only if teacher quality is a function (at least in part) of teacher effort that incentives are likely to influence student performance. Ignoring teacher effort as a contributing factor in educational quality counters much of the empirical literature on the subject of teacher productivity; if poor teachers are teachers who are unable to master the skill of teaching well, then incentive provision is unlikely to improve student performance in any significant way. However, if poor teachers are unmotivated and "lazy" teachers, then the introduction of well-designed incentives may well contribute to improved student performance. The extent to which teacher productivity is driven by effort determines to some degree the likelihood of success in the implementation of incentives (Neal, 2011: 8).

It is undeniable that teachers vary substantially in their levels of productivity, in South Africa and internationally. The fact that differences in productivity exist says nothing about whether teachers provide efficient effort given their level of talent. Teachers may well operate according to different “effort norms”. It is therefore necessary to ascertain whether or not productivity is likely to change if effort is incentivized, and this means ascertaining whether or not productivity depends on teacher effort.

### **3.3 Moral Hazard and the Risk of Distortion**

*“I come to the following pessimistic laws. . .: The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor”* (Campbell, 1976). This section discusses the risk involved in incentivising teacher performance. It discusses the risk of encouraging sub-optimal behaviour and how this comes about with the introduction of incentive schemes.

#### **3.3.1 Multitasking and the Risk of Distortion**

A criticism often levelled at performance-based incentives is that if incentives are used in order to encourage teachers to provide higher quality of education to their students, teachers may improve test scores (according to which payments are made) without actually enhancing learning – behaviour known as distortion. Distortion may also occur when the sorting effect is at play. If we believe that education is valuable to individuals because of its impact on future earnings and we are using student test performance as a measure of learning, it may well happen that teachers who are able to increase student test scores without influencing their learning are drawn into the profession (Lazear, 2003: 187).

Neal (2011: 10) uses a special case of the multi-tasking model of Holstrom and Milgrom (1991) is often used to explain the behaviour of teachers in various merit pay schemes.

Assume that in an education system, teachers allocate their effort between two tasks,  $t_1$  and  $t_2$ , i.e. the amount of time the teacher allocates to task 1 and task 2, respectively. The human capital production function is given by

$$(1) \quad h = f_1 t_1 + f_2 t_2 + e$$

where human capital acquired by the student as a result of teacher effort is  $(h - e)$  (Neal, 2011: 10)  $h$  denotes additional student skill and it is measured in dollars.  $f_1$  and  $f_2$  are

constants and  $e$  is a random error term that captures factors affecting a student's rate of learning beyond the teacher's control.  $h$ ,  $t_1$  and  $t_2$  are not observable by the authority, but the authority is able to observe a statistical measure of teacher performance  $p$

$$(2) \quad p = g_1 t_1 + g_2 t_2 + v$$

where  $g_1$  and  $g_2$  are constants and  $v$  is a random error influencing measured performance (Neal, 2011: 10-11).  $v$  and  $e$  are shocks and are independently drawn with mean zero. They are also assumed to be independent of  $t_1$  and  $t_2$ . The teacher's utility function is

$$(3) \quad U = X - C(t_1, t_2)$$

where the teacher's expected income is given by  $X$  and the cost associated with any pair  $(t_1, t_2)$  is given by  $C(t_1, t_2)$ . An optimal compensation contract is designed by the education authority, given by

$$(4) \quad w = s + bp$$

in which  $s$  is the base salary and the bonus measure  $b$  is paid according to the measure of performance  $p$  (Neal, 2011: 11). Given  $b$ , a salary  $s$  can be chosen to result in a given level of teacher effort based on some utility option  $U_0$ .

The optimal bonus rate solves

$$(5) \quad \max_b f_1 t_1(b) + f_2 t_2(b) - C(t_1(b), t_2(b))$$

subject to

$$[t_1(b), t_2(b)] = \arg \max_{t_1, t_2} s + b(g_1 t_1 + g_2 t_2) - C(t_1, t_2)$$

The optimal bonus rate therefore maximizes the difference between the human capital that results from the teacher's actions and the cost to the teacher of those actions (Neal, 2011: 11). Teachers will respond to any bonus rate  $b$  by choosing actions that maximize utility given  $b$ . The cost function for teacher effort is given by

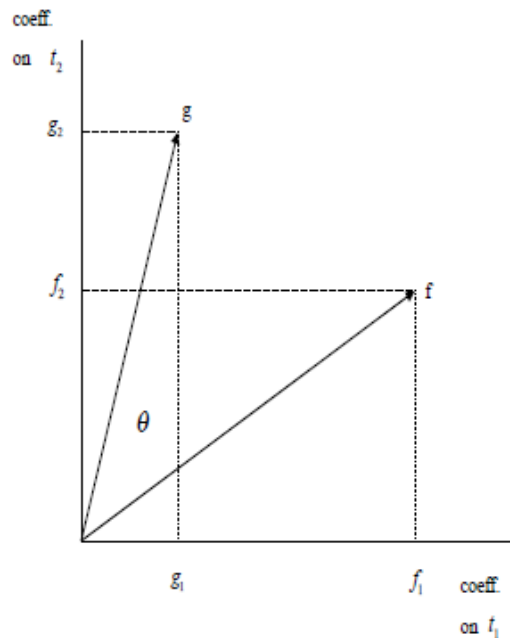
$$(6) \quad C(t_1, t_2) = 0.5(t_1 - \bar{t}_1)^2 + 0.5(t_2)^2$$

where  $\bar{t}_1$  is a norm for the amount of effective instruction. This is considered fixed and not affected by the incentive system and so is taken as given in the calculation of the optimal incentive structure. Importantly, it is stipulated by the education authority and is therefore assumed to be observable by them. This assumption is relaxed shortly. From the cost function, it can be shown that the optimal bonus rate is

$$(7) \quad b^* = \frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2} = \frac{\|f\|}{\|g\|} \cos \theta$$

with  $\theta$  the angle between vectors  $(f_1, f_2)$  and  $(g_1, g_2)$  (Neal, 2011: 12).

**FIGURE 2: Angle between vectors  $(f_1, f_2)$  and  $(g_1, g_2)$**



The formula for  $b$  indicates that the alignment factor  $\theta$  is important for the optimal bonus rate. If the vectors are orthogonal, such that  $(f_1 = 0, f_2 > 0)$  and  $(g_1 > 0, g_2 = 0)$ , then  $\cos \theta = \cos 90 = 0$  and  $b^* = 0$ . In the case of perfect alignment,  $\cos \theta = \cos 0 = 1$ .

### 3.3.2 The Efficiency of Incentive Pay in Education

Neal (2011: 13) discusses whether the presence of at least some incentive pay is optimal in this model, i.e. whether  $b^* > 0$  is optimal. The version of the model presented above indicates that incentive pay is optimal as long as  $f_1 g_1 + f_2 g_2 \neq 0$ . In cases where  $b^* < 0$ , it is possible for the authority to institute  $b^* > 0$  by simply specifying a new performance measure  $p' = -p$ .



As long as all of the constants  $(f_1, f_2, g_1, g_2)$  are non-negative and at least three are strictly positive, then the condition that  $f_1g_1 + f_2g_2 > 0$  will hold. In other words, as long as

- i) one of the teacher's activities contributes to output as well as the performance measure,
- ii) the other task contributes to output or the performance measure or both, and
- iii) neither task is detrimental to either the performance measure or real output,

then  $b^* > 0$  is optimal (Neal, 2011: 13). If  $t_1$  are activities that generate genuine increases in human capital, and  $t_2$  are activities that may be considered "gaming the system", such as teaching to the test or changing the answers of students before the assessment is graded, then it is widely assumed that  $f_1 > 0, g_1 > 0, g_2 > 0$  and if gaming activities do not diminish human capital amongst students,  $f_2 \geq 0$ . Optimal policy should always then include  $b^* > 0$ . The separability of the cost function is an important assumption for this framework.  $b^*$  is only optimal to the extent that teachers can consciously and willingly decide on the combination on  $t_1$  and  $t_2$ , implying that they are able to distinguish between which of their actions genuinely enhance the skills of students and which of them simply enhance their performance measure. Without this assumption, the optimal policy of  $b^* > 0$  is not a robust feature of the multi-tasking model (Neal, 2011: 13).

Given that education requires time of both students and teachers, and given the restricted nature of students' attention and energy, it may be worth considering cost functions that take the form

$$(8) \quad C(t_1, t_2) = 0.5(t_1 + t_2 - \bar{t})^2$$

In this cost function,  $t_1$  and  $t_2$  are perfectly substitutable and  $\bar{t}$  is a total effort norm that impacts on teacher costs. It is assumed that teachers choose  $t_1 = \bar{t}$  and that  $t_2 = 0$  when no incentives exist. Given this setting, if the education system chooses  $b > 0$ , then teachers will choose  $t_1 = 0$  as long as  $g_2 > g_1$  and as long as there are many combinations of  $\bar{t}, f_1, f_2$  and  $g_2$  that would result in  $b = 0$ . So when  $f_1 > f_2$  and  $f_1\bar{t}$  is baseline output, an incentive scheme that results in teachers substituting a small amount of  $t_2$  for  $t_1$  will lower human capital gains to students without affecting teacher remuneration costs (Neal, 2011: 13-14).

However, for  $\bar{t}$  low enough, this may still result in increased total surplus. Since  $t_1$  can never be negative, the diminished output associated with the loss of  $t_1 = \bar{t}$  may be compensated by

the benefits associated with increasing  $t_2$  far beyond  $\bar{t}$ . Therefore, whether or not an optimal bonus rate ( $b^*$ ) exists depends on what type of instruction happens in the classroom (effective teaching or “gaming” behaviour) and on the norm  $\bar{t}$  within the education system (Neal, 2011: 14).

The nature of teacher activities denoted by  $t_2$  that result from the introduction of incentive schemes is important to consider, as well as the relative values of  $f_2$  and  $g_2$  given the assumption that  $g_2 > g_1$ . Equally important, however, is to consider whether increased  $t_2$  activities result in increased teacher effort or whether it is simply substitution away from effective teacher behaviour  $t_1$ . If teacher effort is directed away from effective teaching practices towards gaming behaviour, the effect this will have on learning depends on the level of productive effort initially. In education systems where very little effort was directed towards effective teaching practices, an increase in gaming practices will increase the overall performance of the system because of the overall increase in teacher effort. In education systems in which the level of effort devoted to  $t_1$  is high, incentives systems which result in less effective teaching practices and more gaming practices will result in a decline in the overall output of the system (Neal, 2011: 14). For example, it may be argued that in a system where teacher effort is extremely low, a change from no teaching to “teaching to the test” still represents an increase in the amount of teaching and learning compared to what was happening before, albeit only in areas which are likely to be tested and therefore impact on the performance measure. Therefore the overall output of the education system has increased. However, in the case where effective teaching and learning are taking place in which students’ skill bases are being expanded in all areas and not just those related to material likely to be tested, a movement towards “teaching to the test” and away from genuine skills-enhancing teaching will diminish the output of the education system. Furthermore, although it is possible to say that enhancing  $t_2$  in systems characterised by extremely low levels of  $\bar{t}$  may still enhance overall output of the education system, the long-run implications of encouraging such behaviour are significant and the development of hard-working teaching force which contributes in a meaningful way to the skills development of students becomes less likely.

The question is therefore whether incentive systems lead to improvements in student performance that result from increases in student skills, or whether the improvement in performance is the result of gaming behaviour which improves the measure of student

performance without enhancing students' skills to the same extent that quality skills-enhancing teaching would.

### **3.3.3 Contamination and Hidden Actions**

Coaching is not an optimal allocation of teacher effort, but some positive skills enhancement may result from certain forms of coaching. Indeed if coaching reflects a reduction in leisure time on the part of teachers rather than a reduction in effective teaching time, then teacher performance under an incentive scheme may well be enhanced relative to what it would have been under an accountability program alone. However, teachers respond in ways other than coaching that constitute what Koretz (2002) refers to as cheating and which are unequivocally wasteful from the perspective of public welfare.

Evidence from Jacob and Levitt (2003) on the behaviour of Chicago teachers clearly constitutes cheating, when teachers altered the answers of the students on high-stakes assessment. This type of cheating is usually quite easy to detect because the performance of students being taught by cheating teachers is significantly out of line with their performance in other areas. In addition, the performance of these students reflected significant increases from the previous year, but only small increases (if any at all) in the year following the cheating.

Figlio and Winicki (2005) present evidence from Virginia that on the day that an assessment took place, the sugar content of the meals given to students was increased. It appears that school officials were responding to literature on the positive correlation between academic performance and glucose intake. This is a prime example of the behaviour taken by agents to enhance their performance measure without actually having a real impact on skills enhancement.

The multi-tasking model underlines the possibility that inefficient teaching behaviour may result from the implementation of incentive schemes. However, it has also draws attention to the fact that in situations where the level of teacher effort is very low, even teaching practices that enhance only the performance measure may result in an overall increase in teacher effort and so there is an overall increase in welfare within the education system. In terms of the long term objectives of education (the improvement of students' skills), it is clear that teaching behaviour that enhances the performance measure without actually altering the level of human capital generated amongst learners is undesirable. This emphasizes a problem with

using measures of student performance as the basis upon which incentives are paid. The fact that it is susceptible to “gaming” highlights the difficulties faced by the education authority with regards to monitoring teacher responses to incentive programs. A possible remedy to the issue of monitoring is the use of partnerships and peer pressure amongst teachers to illicit appropriate teaching behaviour in the presence of an incentive program. Subsection 2.4 analyses the channels through which peer pressure and partnerships are likely to influence behaviour under incentive schemes.

### 3.4 Peer Pressure and Partnerships

The creation of partnerships and profit-sharing arrangements stems from the need for internal motivation which arises mainly because of the inhibitive cost of monitoring employees, or even when monitoring is not expensive, the accuracy of observable performance measures may come into question. By rewarding and punishing workers as a team, sufficient incentive may be provided for workers to supply an adequate level of effort. The idea is that team members are in a better position to control and discipline each other at the same time as having a real incentive to do so, given that the overall payoff and therefore their individual payoff is contingent on producing a level of output that is only attainable with a certain amount of effort from each individual worker (Kandel and Lazear, 1992: 802).

This subsection explores the conditions under which peer pressure operates.

#### 3.4.1 Free-Rider Effects and Peer Pressure

Suppose that a group of identical workers produces output  $f(\mathbf{e})$ , so that output is a function of the individual effort of each worker  $e_i$ . Therefore,  $\mathbf{e}$  is an N-dimensional vector of the levels of workers’ effort with N workers.  $f(\mathbf{e})$  is assumed to be non-separable in  $e_i$  ensuring a reason for partnerships. This would be tantamount to saying that pupil education depends on the effort of all of their teachers and that overall effort is non-separable (Kandel and Lazear, 1992: 803).

The cost of effort is denoted  $C(e_i)$  with  $C' > 0$  and  $C'' > 0$ . The worker seeks to maximize

$$(9) \quad \max_{e_i} \frac{f(\mathbf{e})}{N} - C(e_i)$$

with first-order conditions

$$(10) \quad \frac{f_i(\mathbf{e})}{N} - C'(e_i) = 0.$$

Efficient production requires the maximization of total surplus

$$(11) \quad \max_{e_1, e_2, \dots, e_N} f(\mathbf{e}) - \sum_{i=1}^N C(e_i)$$

with first-order conditions

$$(12) \quad f_i(\mathbf{e}) - C'(e_i) = 0 \quad \forall i.$$

Given that  $C'' > 0$ ,  $\mathbf{e}^*$ , the solution to (12) is thus larger than  $\mathbf{e}$  the solution to (10) for all  $N > 1$ . The level of effort chosen in a partnership is below the efficient level (Kandel and Lazear, 1992: 804).

When effort is observable, the best solution would be to pay workers  $a + be$  where  $b = f_i(\mathbf{e}^*)$ . However, the root of the problem is the observability of effort. Payment on the basis of effort is therefore ruled out and we investigate the effect of peer pressure on efforts.

The “peer pressure” function is written as

$$(13) \quad \text{peer pressure} = P(e_i; e_j, \dots, e_N, a_i, \dots, a_N)$$

The pressure felt by worker  $i$  is dependent on their own efforts ( $e_i$ ), the effort of their peers ( $e_j, \dots, e_N$ ) and other actions taken by their peers ( $a_i, a_j, \dots, a_N$ ). The actions of workers<sup>2</sup> have no effect on output, but they do involve some cost to the workers. Cost is therefore redefined as  $C(e_i, a_i)$ . The general maximization problem for worker  $i$  then becomes

$$(14) \quad \max_{e_i, a_i} \frac{f(\mathbf{e})}{N} - C(e_i, a_i) - P(e_i; e_j, \dots, e_N, a_i, \dots, a_N)$$

---

<sup>2</sup> These are discussed in more detail later.

Peer pressure differs from the cost of effort in the sense that it is social and dependent on the effort and actions of others, so it is to some extent external to the worker. That is,  $P(\cdot)$  is subject to manipulation by the group in which the worker works.  $C(e)$  is not. We may therefore think about  $C(e)$  being the exogenous part of the utility of effort (in the sense that it is determined independently of the effect that peer pressure may have on utility) and  $P(\cdot)$  being the part that is endogenous and cultural (Kandel and Lazear, 1992: 804). Introducing the peer pressure function is an attempt to explain differences in preferences and work ethics amongst workers. Making explicit assumptions about  $P(\cdot)$  allows us to make statements about the tastes that drive particular behaviour.

The Cournot-Nash assumption that the actions of all other workers are taken as given is used. Each worker receives  $\frac{f(e)}{N}$  in a pure partnership of size  $N$ . The actions of other workers (the  $a$ 's) are assumed to have no effect on  $P$  and are therefore set to zero. The worker's problem becomes

$$(15) \quad \max_{e_i, a_i} \frac{f(e)}{N} - C(e_i, a_i) - P(e_i, \dots)$$

with first-order conditions

$$(16) \quad \frac{\partial f / \partial e_i}{N} - C_1 - \frac{\partial P}{\partial e_i} = 0$$

given that  $a_i$  is set to zero.

Peer pressure implies that  $\partial P / \partial e_i < 0$  (the disutility associated with peer pressure diminishes as workers exert more effort), which means that the level of effort that solves (16) is greater than the level of effort that solves (10). That is, effort is higher with peer pressure than without peer pressure. The  $P(\cdot)$  function implies that workers get utility from effort. This has only been assumed, however. Workers working in an environment in which peer pressure exists may be worse off because of it, however. Peer pressure may increase effort, but it does not necessarily mean higher utility because pressure itself is a cost that is borne by all firm-members. Despite the higher output that results from peer pressure, workers may not enjoy working in a high-pressured environment (Kandel and Lazear, 1992: 805).

### 3.4.2 Creating Peer Pressure

Considering the actions of organizations in the peer pressure function allows us to analyse how peer pressure is created and how partnerships result in higher effort levels.

The effectiveness of peer pressure as a motivator is conditional on two things: first of all, the effort of member  $i$  must affect the well-being of other team members in order for them to have incentive to exert pressure on him. Secondly, team members must be able to affect the choices of member  $i$ . Profit-sharing in some form (or sharing of the incentive in the case of teachers) is required in order for the first component to hold. If workers are paid straight salaries, their level of effort does not impact on the salaries received by any of their peers, removing the incentive for workers to exert pressure on their peers since they are not concerned with the level of effort they choose. Profit-sharing is therefore necessary for peer pressure to provide motivation for workers to influence their peers' behaviour. However, it is not sufficient. Even if workers have reason to influence the behaviour of their peers, they need to be able to exert pressure for peer pressure to actually provide incentives. Both components are assumed to be created and manipulated to some extent by the education authority (Kandel and Lazear, 1992: 806).

Pressure can be classified as either internal or external: internal pressure exists when workers feel disutility from hurting others, regardless of whether others are able to identify the offender or not. Sociologists call this "guilt". External pressure or "shame" occurs when disutility is dependent on others being able to identify the worker hurting others. The lack of observability implies only guilt or internal pressure will be an effective form of pressure. Shame or external pressure requires that workers are able to observe each other's efforts. If  $a_j$  denotes the monitoring ability of workers, shame requires  $a_j > 0$ , while guilt would create pressure effectively for  $a_j = 0$ . Therefore,

$$\frac{\partial P(e_i; e_j, \dots, e_N, a_i, a_j, \dots, a_N)}{\partial e_i} < 0$$

holds only when  $a_j > 0$  in the case of shame, but holds for  $a_j = 0$  in the case of guilt. The implication is that investment may be required to create guilt (Kandel and Lazear, 1992: 806 – 807).

Another aspect contributing towards the effects of incentives is empathy; if workers are monitored by people who are not members of the profit-sharing group, it is unlikely that workers will be motivated to behave efficiently.

$N^*$  denotes the number of profit sharers that the individual cares about. The peer pressure function is defined  $P(e_i, \dots, 0) = 0$ , indicating that if the individual worker does not care about any peers, there is no relevant pressure. The maximization problem then becomes

$$(17) \quad \max_{e_i} \frac{f(\mathbf{e})}{N} - C(e_1, \dots) - P(e_i, \dots, N^*)$$

where  $N$  is the number of individuals sharing the reward.  $N$  may consist only of members of the education authority. If workers do not empathise with them, then  $N^* = 0$ , in which case  $P = 0$  and peer pressure would be useless as a motivating force. Allowing only the individuals about whom workers care to share in the rewards results in maximum motivation.  $P(\cdot)$  therefore becomes endogenous in the sense that it becomes dependent on  $N^*$ , and it may be worthwhile to invest some resources in altering  $P(\cdot)$  in order to improve incentives faced by teachers. The environment in which initial investments in loyalty are likely to be most effective and necessary is characterized by two features, namely the unobservability of workers' efforts and the complementarity of production (Kandel and Lazear, 1992: 807 - 808).

### 3.4.3 Mutual Monitoring

In addition to exerting effort, workers are able to monitor each other and are able to penalize other workers caught shirking. For now, the punishment is assumed to be nonpecuniary (so it takes the form of physical or mental harassment) (Kandel and Lazear, 1992: 811). In this case,  $a$  can be thought of as peer monitoring. The expected penalty of being caught shirking now becomes

$$P(e_i; a_j, \dots, a_n, N).$$

All workers are considered to be identical and so the monitoring decision of worker  $k$  is identical to that of worker  $j$ , which means the penalty may be written

$$P(e_i, (N - 1)a_j).$$



The assumption of identical workers in the output function implies that  $i$ 's maximization problem becomes

$$(18) \quad \max \frac{f(e)}{N} - C(e_i, a_i) - P(e_i, (N-1)a_j)$$

with first-order conditions

$$(19a) \quad \frac{f_i(e)}{N} - C_1 - P_1 = 0$$

and

$$(19b) \quad \frac{N-1}{N} f_i \frac{\partial e_j}{\partial a_i} - C_2 = 0.$$

The choice of monitoring level  $a$  for each worker must satisfy (19b), according to which other workers respond to worker  $i$ 's choice of  $a$ . Workers generally believe that their co-workers will respond to their increased monitoring effort. Differentiating (19a) with respect to  $a_j$  will show the response of worker  $j$ 's level of effort to worker  $i$ 's choice of  $a_i$ , given that the problem is symmetrical across workers:

$$(20) \quad \left. \frac{\partial e_j}{\partial a_j} \right|_{(19a)} = \frac{-P_{12}}{\left(\frac{f_{ii}}{N}\right) - C_{11} - P_{11}}.$$

The denominator of (20) is unambiguously negative which means that the sign of the expression will be the opposite of the sign of the expected punishment,  $P_{12}$ . The expected punishment is related to the accuracy of detection – an increase in monitoring will presumably increase the accuracy of the measurement of co-worker effort, in which case those workers' levels of effort will increase, resulting in  $P_{12} < 0$ .  $P_2$  does not enter, indicating that simply increasing monitoring efforts has no effect on the level of punishment. It is the *interaction* of increased monitoring and increased effort that accounts for  $P_{12} < 0$ , and simply increasing monitoring without allowing for this interaction (i.e. the worker's ability to avoid the higher punishment level by increasing their work efforts when co-workers increase their monitoring efforts) will have no effect. If we believe that increasing monitoring will have no impact on worker efforts, then peer pressure loses its value as an incentive (Kandel and Lazear, 1992: 812).

The success in using peer pressure to incentivise teachers depends on how successful the education authority is in creating “guilt” amongst teachers. As the model shows, the number of workers that the individual cares about or who will share in the profits will influence the level of effort contributed by the individual. In the context of education, we may understand this to mean that when an incentive is provided for a whole school and not just individual teachers, it is likely that the monitoring activities of teachers will be effective in enhancing the effort levels of their peers. Important to acknowledge when peer pressure creates incentives for workers is that it is only likely to result in increased effort levels if joint performance evaluation (as opposed to relative performance evaluation) is used to measure performance (Che and Yoo, 2001: 526). Under relative performance evaluation workers are compared to their colleagues and are relatively worse off when their colleagues perform well given that relative performance evaluation rewards the highest performing workers within the organization. It therefore compares co-workers and so increased effort from one worker will “disadvantage” their colleagues. Relative performance evaluation therefore penalizes workers when their colleagues succeed (Che and Yoo, 2011: 529).

In settings where workers are likely to have “repeated interactions”, or settings in which long term contracts are in place, the implicit incentives created by joint performance evaluation may prove useful. The implicit incentives arise as a result of the fact that workers are rewarded on the basis of both their own performance and that of their colleagues (Che and Yoo, 2001: 529), so hardworking co-workers increase the likelihood of reward which is not the case under relative performance evaluation. Furthermore, the possibility that shirking in the current period may be “punished” with shirking by co-workers in the next period or at a later stage is “self-enforcing” in the sense that workers “punishing” shirking co-workers in subsequent periods is not stipulated in any incentive contract and therefore does not need to be enforced by the education authority. Che and Yoo (2001: 231) explain that this is an endogenous characteristic of the relationship in a setting of joint performance evaluation and is likely to ensure that workers actually work (as opposed to shirk).

### **3.5. Return to Distortion**

The model discussed above indicates that introducing monitoring amongst teachers will theoretically enhance their level of effort. However, it is not clear whether the expected increased earnings will result in an increase in human capital-enhancing teaching behaviour or behaviour that may be considered distortion. Therefore, peer pressure does not necessarily

result in improved long-term outcomes for students – the ultimate objective of implementing teacher incentives. Although peer pressure as an incentivising instrument may enhance the level of effort in a setting in which observability is limited, it does not remedy the problem of distortion in teaching behaviour.

So what features of the theoretical models discussed above should be considered most important to successfully incentivise teachers to enhance the performance of their students? Section 4 puts together a brief framework of the characteristics that enable incentive systems for teachers to result in improved performance for their students. Section 5 then presents international examples of teacher incentive programs and analyses them according to the characteristics mentioned in section 4.

#### **4. CHARACTERISTICS OF SUCCESSFUL INCENTIVE PROGRAMS**

Three aspects of incentive programs need to be investigated when evaluating the likelihood of success: whether or not incentive programs are likely to improve student outcomes through sorting or through enhanced effort, the risk of distortion that arises with the introduction of incentive programs and in the case of low observability of worker effort (which is certainly the case in the teaching profession), whether internal source of performance monitoring and accountability can be utilised to provide incentives (i.e. peer pressure). Each of these is discussed briefly below.

##### **4.1 Sorting versus incentives**

Whether or not incentive programs are likely to result in improved performance depends on whether the program result in higher levels of effort amongst teachers (assuming that increased teacher effort will result in improved student performance) or whether individuals better able to enhance student performance will be drawn into the teaching profession. Whether incentive programs are likely to result in sorting or enhanced effort is an important question to consider because it is possible that the introduction of incentive programs may have negative implications for poor performing schools. Clotfelter, Ladd, Vigdor and Diaz (2004) examine this question in detail using data from North Carolina. This is discussed in more detail in section 5.

## **4.2 Potential for distortion**

Another aspect to consider is the likelihood of distortion. The extent to which the introduction of incentives results in genuine improvements in learning as opposed to improvements in performance measures as well as the long term effects that incentives will have on student learning are important considerations. Neal (2011: 14) points out that even distortive behaviour may be a net gain in the case of extremely low productivity, e.g that teaching to the test may be better than no teaching at all. . The risk for gaming behaviour as well as potentially short-term benefits should be investigated.

## **4.3 Possibility for peer pressure**

Finally, the likely success of using peer pressure to incentivise teachers is dependent on teachers being able to monitor each other as well as being able to impose a penalty on workers who are found shirking. The combination of monitoring and penalising is important for this method of incentivising to be effective. Importantly, it requires that incentives be provided at the level of the school and not for individual teachers.

According to the features highlighted in this section, how do incentive systems measure up internationally? Do they result in either sorting or enhanced teacher effort, and does this in turn result in improved student performance? Where improvement occurs, is this genuine human capital development or is it the result of gaming behaviour by teachers? Do the incentive systems in place enable the use of peer pressure as an incentivising mechanism? Section 5 discusses incentive systems from India, Israel, Kenya, Brazil, Chile, USA and Finland, analysing whether or not they are likely to fulfil the criteria outlined in the section 4.

# **5. INTERNATIONAL EXAMPLES OF INCENTIVE PROGRAMS**

As mentioned above, this section discusses incentive systems implemented internationally. The last 3 incentive systems discussed in this section – those of the USA, Chile and Finland - are not analysed in terms of the framework laid out in section 4, but rather serve as examples of the risks associated with implementing incentive programs (USA), the success with which pragmatic incentive programs can be implemented in a developing country context (Chile) and the potential benefit that comes from incentivising the teaching profession through factors such as selectivity, prestige and relatively attractive compensation (Finland).

## **5.1 Andhra Pradesh, India (2005 – 2007)**

An example of performance pay on an individual level based on student learning outcomes can be found in a random control trial (RCT) conducted in India. In this study, individual teacher bonuses were awarded in 100 schools, bonuses were awarded to groups of teachers in 100 schools, an extra contract teacher was provided in 100 schools and a school grant was awarded to 100 schools. A further 100 schools were also included as a comparison group. The study was conducted over 2 years, and performance bonuses were promised to teachers at the beginning of the following school year (Muralidharan and Sundararaman, 2011). Bonuses were awarded to any teacher or school managing to increase student test scores by at least 5 percentage points, with higher increases being awarded with larger bonuses. Individual and group-level bonuses were paid at the beginning of the next school year, and block grants and extra contract teachers were provided unconditionally at the beginning of the school year.

At the end of the 2 year programme, significant differences existed between individual- and group-level bonuses, as well as between schools receiving teacher bonuses and those receiving either block grants or extra contracted teachers. Individual incentives increase student performance on tests by 0.27 standard deviations (roughly 9 percentage points) in comparison with an increase of 0.16 (roughly 5 percentage points) for group incentives. Input strategies (block grants and extra contracted teachers) also improved student performance, but by a substantially smaller amount – 0.08 standard deviations.

During the study, teachers were monitored by observers through unannounced classroom visits as well as interviews (Muralidharan and Sundararaman, 2011: 67). Teachers were monitored through 20 to 30 minute classroom observations in which enumerators coded whether or not certain actions took place in the classroom from a position at the back of the class, without interfering with proceedings. Teachers were then also interviewed about their teaching methods and practices. The interviews took place at the end of the school year after testing had taken place but before results were available. Teachers were asked open-ended questions about how their teaching practice differed over the school year (Muralidharan and Sundararaman, 2011: 67 – 68). Although differences in classroom observations were not significant between treatment and control schools, teachers teaching in schools which received the incentives were significantly more likely to have given extra classes after school hours, to have focussed additional attention on weaker students, to have assigned more class work and more homework and to have given more tests as practice for exams. It is true that self-reported behaviour is less credible than classroom observations. However, the authors

find a positive and significant relationship between teachers' self-reported behaviour and the performance of their students, suggesting that the teachers' reports are likely credible (Muralidharan and Sundararaman, 2011: 68). They conclude that although there was no difference in the proportion of teachers captured as "actively teaching" by classroom observers, it is likely that teachers increased the intensity of their teaching efforts (Muralidharan and Sundararaman, 2011: 69).

This study therefore provides evidence that teacher effort was enhanced by the introduction of incentives. The brief framework constructed in section 4 lists 3 criteria according to which incentive systems can be evaluated; whether incentive systems are likely to improve student performance through sorting in the teaching profession or enhanced teacher effort or both, whether incentive systems are likely to introduced distortion into teacher behaviour and whether peer pressure is a likely channel through which incentives can be implemented. Evaluating the incentive system introduced in Andra Pradesh according to this framework reveals that the incentives were effective in enhancing student performance through enhanced teacher effort. From the reported results it is difficult to see whether the performance gains were long-lasting and reflecting genuine skills development, or short term and reflecting an improvement in the performance measure. Although the gains were observed at the end of both years in the program, students were not tested after the incentive program had ended so whether or not the effects of the program outlasted its duration is unclear. The fact that teachers reported higher levels of practice tests and more extra tuition suggests that "coaching" may have been responsible for enhanced student performance. However, an increased focus on weaker students indicates higher levels of effort directed towards improving student human capital. In terms of whether group incentives are likely to render peer pressure a viable channel through which incentives can be implemented, it is not clear whether teachers in Andra Pradesh used any kind of pressure or punishment to achieve increased effort levels. What we can see is that individual level incentives improved student performance by a greater margin than group level incentives, indicating that individual level incentives were more effective, although the group level incentives did also result in improvements in test scores.

## **5.2 Israel (2001)**

Lavy (2009) presents evidence from a tournament-type bonus programme introduced in Israel in 2000. In 18 schools, teachers were ranked within schools based on their value-added

contributions to predicted matriculation marks of their students after controlling for socioeconomic characteristics, grade level, their level of study and school-level fixed effects. A control group of 18 schools was also included as (Lavy, 2009: 1980). Teachers were ranked according to the deviation of the mean residual of the pass rate achieved by the students in their class, as well as the mean residual of the score that students achieved in various subjects (Lavy, 2009: 1983). . Students who did not take the examination were assigned a score of zero but not excluded from the sample. The top four ranked teachers in each subject (English and Maths were the core subjects, with awards being offered in other optional subjects) received an award that amount to a substantial proportion of their salary (25% for teachers ranked first, 19.2% for teachers ranked second, 11.7% for teachers ranked third and 5.8% for teachers ranked forth) (Lavy, 2009: 1983). Significant positive effects on student achievement were observed, with increases being observed in test-taking amongst high school seniors, average scores and average pass rates in both mathematics and English. Although test-taking was optional, students who appear on enrolment lists but who do not take the test are given a score of zero but not excluded from the sample, therefore diminishing the incentive to discourage weak students from taking the exam (Lavy, 2009: 1983). Furthermore, given that the incentive program was only revealed to teachers in the middle of the year (December 2000, with exams being written in June 2001), it is unlikely that teachers were able to influence the composition of their class in order to ensure a stronger group of students (Lavy, 2009: 1982 – 1983). Teachers in school for which the incentives were offered reported interesting modifications in their behaviour compared to teachers in control schools. Teachers in treatment schools were significantly more likely to track students by ability in the classroom (Lavy, 2009: 2004), offer extra classes after school (Lavy, 2009: 2004) and to adjust their methods of instruction to the individual needs of students (Lavy 2009: 2004). As a result, an increased proportion of students took the mathematics exam in schools in which the incentive was offered relative to those in the control group.

Interestingly, it was observed that teacher effectiveness (as measured by their success in obtaining the bonus on offer) was uncorrelated with their observable characteristics (such as education level, gender, age, years of experience or certification level) but was correlated rather with the calibre of university that the teacher attended (Lavy, 2009: 2004-2005). A significantly higher level of effectiveness was observed amongst teachers who graduated

from top-ranked Israeli universities in comparison to those who attended teacher colleges or less-prestigious universities.

This incentive program is interesting to analyse in terms of section 4's framework. The results indicate that student performance did improve. Lavy (2009: 2004) reports that ability tracking, extra classes and focus on the specific needs of students was observed amongst teachers in both control and treatment schools, but that this behaviour was more prevalent amongst teachers who were offered the incentive. This suggests that the incentive program encouraged increases in genuine effort and not merely "gaming behaviour". It is therefore unlikely that improvements in student performance resulted from gaming behaviour (Lavy, 2009: 2004). The observation that teachers who qualified at more prestigious universities performed better is an interesting and potentially important one. If individuals with higher ability to perform well in terms of student performance were the ones who attended prestigious universities, this may well indicate that there may be a potential to draw high ability individuals into the teaching profession through substantial monetary rewards attached to performance, as they are most likely to benefit from such a program. This provides some support to the hypothesis that teacher incentives may enhance student performance through a process of sorting.

### **5.3 Kenya (1997)**

Glewwe, Ilias and Kremer (2003) present evidence from an incentive program run in 50 rural schools (with a control group of 50 schools) in Kenya awarded in-kind prizes (such as bicycles) which were valued at a significant proportion of a typical fourth- to- eighth grade teacher salaries for improvements in average student performance over 2 years (Glewwe et al., 2003: 14). The prizes were awarded as group incentives, and performance was measured as improvement from baseline test scores obtained on Kenya's district-wide government exams. Prizes were awarded for "top-performing" schools and for "most-improved" schools, with schools being eligible for awards from only one of these two categories. Three prizes each were awarded for first, second, third and fourth place, resulting in 24 out of 50 schools receiving prizes over the 2 year time period for which the program was in place. Teachers therefore felt that the reward was obtainable (Glewwe et al., 2003: 14).

A higher proportion of students in schools for which rewards were available compared to control schools achieved gains in test scores. By the second year, an average gain of 1.4 standard deviations was observed in treatment schools, with the largest effects



being observed in geography, religion and history (roughly 0.34 standard deviations in the first year, and 0.20 standard deviations in the second year of the program), followed by mathematics and science (with improvements of 0.20 and 0.15 standard deviations, respectively) (Glewwe et al., 2003: 29). However, these improvements did not persist. Differences in test performance had disappeared a year after the program had ended. Glewwe, Ilias and Kremer (2003: 33) speculate that the introduction of rewards resulted in a short-run focus for improving test performance. They suspect, for example, that teachers may have focused more on short-run approaches such as increased coaching in test-taking techniques rather than an increased focus on pedagogical adjustments that may have resulted in longer-term learning gains. No evidence of behavioural changes was observed, with teacher absenteeism failing to decline over the period for which the rewards were offered (Glewwe, 2003: 20). There was also no evidence that more homework was assigned relative to the baseline year (Glewwe et al., 2003: 21). However, during the second year of the program, schools that were eligible for rewards were more likely (by 7.4 percentage points) to conduct extra exam preparation classes (Glewwe et al., 2003: 22). Interestingly, when researchers changed the format of the exam written by students from the format in which the government exams were presented, there was no difference in the performance of schools who were eligible for performance bonuses and control schools, suggesting that the benefits of increased exam preparation classes was limited to performance in the government exam, i.e. the target that teachers were aiming at (Glewwe et al., 2003: 30). It thus did not extend to more general learning. The authors report that there was no evidence of outright cheating amongst schools who received the incentive (Glewwe et al., 2003: 26). However, the fact that there was no significant difference between the performance of students in schools who received the incentive and those who did not when students wrote a different exam (one for which incentives were not provided) provides strong evidence that the improved student performance amongst students in the incentivised schools was unlikely the result of genuine human capital development. Furthermore, the fact that improved performance did not continue after the termination of the incentive program is a clear indication of the short-run focus of teacher effort (Glewwe et al., 2003: 29). There is thus strong evidence of distortion in this incentive program. The incentive program did achieve improvements in student performance *on the performance measure*, but not on any other measure. Increased efforts amongst teachers were therefore directed towards gaming behaviour.

#### **5.4 Pernambuco, Brazil (2008 onwards)**

The Brazilian government's establishment in 2007 of the Index of Basic Education development (IDEB) is central to Brazil's incentive structure that awards schools for improvements in student performance as well as other characteristics. IDEB captures school performance on Prova Brasil test results (national assessment conducted every two years for all fifth, ninth and twelfth grade students in maths and language) with administrative data on enrolment, repetition and grade promotion (Fernandes in Bruns, Adams and Luque, 2012: 9). Importantly, IDEB results are reported widely in the Brazilian media and targets for each school within the 26 state and 5564 municipal school systems have been established by the federal government.

The state of Pernambuco implemented an incentive system in 2008 which rewarded school staff for the attainment of school improvement targets (Bruns, Filmer and Patrinos, 2011: 169). All teachers in schools achieving at least 50 percent of the target set by the federal government received bonuses proportional to their school's level of achievement. The size of the bonus is substantial by international standards since state education departments budget an additional month's payroll for the program each year, and so if less than 100 percent of schools achieve the bonus, the mean bonus for those who received it will be greater than an additional month's salary. Schools achieving less than 50 percent of their target receive no bonus. School principals have no say in the distribution of the bonus, with each teacher in the school receiving equal percentage bonuses on their monthly salary (Parandekàr, Amorim and Welsh, 2008: 2)).

The initial targets are established according to which quartile of the performance distribution schools fall into, with performance targets being more or less ambitious according to the quartile. The differentiation of targets allows for an analysis of how achieving targets in one year impacts on the likelihood of achievement in subsequent years, as well as of how the achievement of targets and receiving a bonus is likely to impact on teacher behaviour (Parandekàr et al., 2008: 2)).

The program was widely accepted by schools in Pernambuco, where 64 percent of school principals indicated that the program was appropriate and 66 percent indicated that they experienced the policy as having a positive impact on their schools, regardless of whether or not they received the bonus. Furthermore, schools for whom targets were more ambitious achieved greater student progress than those with less ambitious targets. Indeed, learning

levels across the state increased substantially, with language score improving for the eighth and eleventh grade by 0.44 and 0.57 standard deviations over the period of a year, respectively. As the program was applicable across the entire state, these gains are raw score gains and not gains relative to any comparison group. Schools that narrowly missed achieving the bonus in 2008 improved more in 2009 than did schools who barely achieved it. It therefore appears that not receiving the bonus improved school motivation and performance. Finally, schools in which teachers spent a larger proportion of time on instruction had a much greater likelihood of achieving the bonus (Bruns et al., 2011: 172).

Overall, teachers in schools achieving bonuses spent considerably less time on activities other than teaching and were also observed (in unannounced visits to the school) to make greater use of classroom resources (Bruns et al., 2012: 166). However, because of a lack of a “control group”, the causality of “better” teacher behaviour cannot be inferred from the analysis conducted on the schools in Pernambuco. It is not clear whether the change in teacher behaviour reflects greater incentive to perform well or whether students in schools achieving the bonus are better students and easier to teach and manage. However, the fact that bonus-achievers came from all parts of the performance distribution, including a substantial number of low-performing and low-income schools, suggests that the performance bonus may well induce improved teacher behaviour (Bruns et al., 2011: 174).

Evaluating the incentive scheme in place in Pernambuco in terms of the framework introduced in section 4, it does not appear that sorting takes place. It is impossible to analyse whether individuals with higher ability entered the teaching profession. However, improvements were observed across the distribution of schools. This is likely the result of individualised targets for each school – a feature that should seriously be considered in the case of South Africa. In terms of distortion, although it is possible that increased teacher effort may be directed towards ensuring improvements on the performance measure according to which teachers are evaluated, the fact that the performance measure is not limited to student test performance eliminates the possibility of distortive behaviour to some extent. It is possible that principals were dishonest in the reporting of enrolment and repetition rates, but the fact that these elements are included in the IDEB performance measure mitigates the risk of distortion to the extent that performance is not measured exclusively by student performance on the Prova Brasil test results. The potential to make use of peer pressure as an incentivising force is heightened by the fact that incentives are awarded at the level of the school. However, whether or not teachers are realistically able to monitor

each other's behaviour and enforce punishment in the case of shirking is unclear from the results reported above. Group incentives do however heighten the probability that this will occur.

### **5.5 Chile (since 1991)**

The System for Measuring the Quality of Education (SIMCE) was introduced in 1990 (Gustafsson, 2006). The objective of SIMCE is the identification of schools in special need of interventions. It is conducted yearly and involves the testing of an entire grade (approximately 300 000 learners), in either grade 4, 8 or 10 (alternated cyclically), ensuring that each grade is tested every three years (Gustafsson, 2006). Importantly, the tests are marked at a single national centre and reports on the performance of individual schools are made available to the public (Delannoy, 2000). Schools are compared within their region and within their socioeconomic category. It is argued that this encourages school principals to over-state the school's poverty level, therefore enabling the school to compete against weaker performing schools. Furthermore, weak performing learners may be discouraged from attending school on the day that the testing takes place (Delannoy, 2000). Teacher incentives in the Chilean education system have been introduced in phases. The first phase was introduced in 1991 with the Teacher Statute. It introduced a system designed to reward continued service as a teacher. Part of the monetary incentives introduced in Chile in 1991 were those directed towards teachers in difficult-to-teach urban schools and remote rural areas (Gustafsson, 2006). A 2003 regulation of the 1991 Teacher Statute details a 12 point index of school remoteness or difficulty, attaching various values to each indicator in order to ascertain the size of the incentive necessary to compensate for the remoteness and difficulty of the school. Provincial authorities then distribute earmarked funds as incentives through teacher salaries to those teaching in the most remote, rural and otherwise difficult schools (Gustafsson, 2006).

The National Performance Evaluation System (SNED) was introduced in 1996, according to which all teachers in a well-performing school are rewarded. SNED is conducted biannually and is heavily reliant on learner performance data provided by SIMCE. Schools are compared within regions (of which there are 13) and within socioeconomic groups (of which there are 5), resulting in 65 groups in which SNED comparisons are made. The top 20 percent (approximately) of schools in each group are considered to be outstanding performers and receive additional funding for 2 years, 90 percent of which is paid to teachers as a monetary

incentive, and 10 percent of which is allocated to schools and which may be spent according to the schools' own development plan. The money rewarded to teachers in outstanding schools amounts to approximately a month's salary, effectively providing a 13<sup>th</sup> pay cheque for them (McMeekin, 2000).

SNED is based on a number of indicators and not entirely on the learner performance data gained from SIMCE. These included value added (improvement in SIMCE scores since the last evaluation), school governance features and learner retention and graduation rates (Gustafsson, 2006). Roughly half of schools have received SNED incentives at some stage which means that the perceived likelihood of receiving an award is quite high.

The Teacher Evaluation System was put in place in 2000. Evaluation happens every 4 years, and involves 4 items: documentation from the teacher related to assessment of learners, a series of the teacher's lesson plans, reflective notes produced by the teachers and a 40 minute video recording of a lesson. University-based evaluators then assess the material and classify teachers as excellent, competent, basic or unsatisfactory (Gustafsson, 2006). A number of non-monetary incentives exist for good performance in the Teacher Evaluation System, namely the eligibility of good performers to participate in the Ministry's overseas experiential learning programme.

The Pedagogical Excellence Award (AEP) was also introduced during this phase. This award is a monetary incentive and it is relatively independent of the Teacher Evaluation System. Teachers are divided into four segments according to their years of teaching experience, with teachers in the first segment having between 2 and 12 years of teaching experience, teachers in the second segment having between 12 and 21 years of experience, and so on. A quota of AEP is established for each region, and the competition involves teachers submitting a portfolio of their methodology and writing a test covering both subject content knowledge and methodology. Winners receive a monetary award roughly equal to a month's salary for all the years in which they remain in their segment (Gustafsson, 2006). Only a small group of teachers actually receive the award – 1500 received it in 2004 and 722 teachers received it in 2005. It is clear, therefore, that the AEP has a very small reach. Teachers do not seem to consider it worthwhile to enter the competition. In addition to its role as an incentive program, the AEP serves as a selection system into the Network for Teachers' Mentors – a remunerated programme. AEP recipients may therefore be seen to have a double monetary incentive, with the financial award received for good performance and with the opportunity to

earn additional money as part of the Network for Teachers' Mentors (Gustafsson, 2006; Mizala and Romaguera, 2004).

As from 2005, teachers who achieved an "excellent" or "competent" rating on the Teacher Evaluation System are able to take a test administered on one day of the year at a national level (held on the same day as the AEP evaluations are held), according to which the Variable Allocation for Individual Performance (AVDI) is awarded for good performance. The AVDI amounts to between approximately 15 and 25 percent of the Minimum Basic National Pay. Teachers who are not yet eligible to be evaluated for AVDI may still apply for AEP (Gustafsson, 2006).

Finally, a very interesting and controversial incentive in place in Chile is the Demerit List system. According to this system, a teacher receiving unsatisfactory ratings in the Teacher Evaluation System for three consecutive years, despite receiving rigorous professional support and the assistance of an assigned tutor, will be dismissed and will receive a dismissal package. Importantly, teacher unions have agreed to this (Gustafsson, 2006).

### **5.5 Incentive Systems: What Does the Evidence Say?**

The incentive systems presented above all provide evidence that introducing teacher incentives results in improvements in the measured performance of students. In the case of Andhra Pradesh, although it is not possible to see whether the improved performance lasted beyond the duration of the incentive program, important adjustments in teacher effort were observed. In particular, an increased focus on weaker students is an important result. In the case of Israel, teacher efforts were also found to increase. Interestingly, correlation between the positive effect of the incentive program and student performance was observed most strongly for teachers from elite universities, suggesting the potential for more able individuals to benefit most from such incentive systems. In the case of Kenya, although improvements in the performance measure were observed with the introduction of teacher incentives, the improvement was not generalizable to tests that were not included as part of the performance measure according to which teachers were rewarded. Furthermore, the improvement in test performance disappeared once the incentive system had been removed. Both of these observations suggest that gaming behaviour in the form of "teaching to the test" took place, rather than genuine skills development. The cases of Pernambuco and Chile illustrate the value of creating individualised incentives for schools. The observation that schools achieved improvements across the socioeconomic status distribution indicates the potential for system-

wide improvement when the socioeconomic context of schools is taken into account. The Chilean example illustrates the possibility of introducing an element of fairness to incentive programs by comparing schools within their socioeconomic status and within their region. In addition to ensuring fair competition, this also ensures that a large number of schools receive the award, therefore increasing the probability of receiving it. The fact the AEP and the AVDI do not include measures of student performance may be perceived by teachers to increase the fairness of performance pay since their achievement is not affected by student performance which is widely known to be affected by numerous factors outside the teacher's control. However, while this eliminates the possibility of teachers behaving in a distortive way in order to qualify for rewards, it also means that teachers are rewarded on the basis of inputs rather than outputs. It is therefore possible that despite the fact that teachers perform well on these measures, they may not be able to enhance student performance. The AEP and AVDI are nevertheless useful programs to enhance teacher pedagogical and content knowledge.

The next 2 incentive programs (from North Carolina in the USA and Finland) present examples of the potential risks and benefits associated with different incentive systems. The final example discussed – that of Finland – is presented as an example of an education system in which the absence of incentive pay does not compromise the performance of the Finnish education system.

### **5.6 North Carolina, USA**

Clotfelter, Ladd, Vigdor and Diaz (2004) present evidence from North Carolina in the United States of America. The purpose of analysing the accountability system introduced in North Carolina is to highlight some of the risks associated with the introduction of incentive systems. North Carolina introduced an accountability system in the 1996 – 1997 academic year entitled ABS (A for accountability, B for basic skills and C for local control) (Clotfelter et al., 2004: 254). The accountability program involves evaluating the gains in Maths and Reading scores from year to year, as well as the proportion of students performing at or above the grade appropriate level (Clotfelter et al., 2004: 255). Making use of the gain scores mitigates the disadvantage experienced by schools catering for students from lower socioeconomic status because it does not consider the level of test scores. Rewards are distributed at the level of the school, and schools' performance is evaluated relative to their expected gain for each year. A school's expected gain is calculated using the state average

which is adjusted for the initial level of proficiency of the school's students as well as for mean reversion (Clotfelter et al., 2004: 255). A school is classified as having met its expected growth if the mean score of the students attending the school is at least as large as the calculated expected gain. Schools in which the mean student score is at 10 percent higher than the calculated expected growth are classified as "exemplary", and schools not reaching their expected gain are classified as either "no recognition" or "low-performing" schools. "Low-performing" schools differ from "no recognition" schools because in the former, less than 50 percent of students performed at the appropriate grade level, whereas in "no recognition" schools at least 50% of student performed at the appropriate grade level (Clotfelter et al., 2004: 255). Growth standards are therefore school-specific and low-performing schools are those that reached neither their school-specific growth standard, nor the 50 percent of grade appropriate performance standard. Financial bonuses of \$1500 were awarded to teachers in "exemplary" schools (Clotfelter et al., 2004: 255). Schools results are made publically available on the ABC's website.

As mentioned earlier, this study provides interesting evidence on the potential negative implications of introducing incentive programs. Clotfelter et al. (2004: 256) report that schools labelled as "low-performing" experienced higher teacher turnover in the years following the introduction of the accountability system and subsequent classification. Interestingly, in weak performing schools in which less than 50 percent of students performed at the appropriate grade level but in which student test performance had met the expected test performance (and who were therefore not labelled as "low-performing"), teacher turnover did not change (Clotfelter et al., 2004: 258).

This example illustrates that labelling low-performing schools as such imposes additional "costs" on them in the form of higher teacher turnover. Higher teacher turnover makes it difficult to create continuity and momentum in reform efforts. This represents a challenge associated with rewarding some schools and not others: it may incentivise teachers to migrate away from weak performing schools towards better performing schools in which performance-related rewards are more likely. The extent to which this is a real risk depends on the ease with which teachers are able to move in and out of schools, and the extent to which teaching posts are available in schools classified as exemplary.



## 5.7 Finland

The Finnish education system is characterised by a very high level of equality, with schools performing in the lowest decile achieving average marks higher than the OECD median. It is therefore clear that educational quality is high throughout the Finnish education system (August, Kihn and Miller, 2010: 18).

Selection into the teaching profession in Finland is highly competitive, and once candidates have been selected to enter the profession, they are required to obtain a master's degree in a five-year program. Students must fall within the top 20 percent of their secondary school academic cohort. Students qualifying to apply to teach are examined in the first round of screening, after which only the top performers are invited to write an exam based on education literature. This is a further round of selection, after which top performers on the second exam are interviewed and screened on "softer" skills in order to ascertain whether or not they are likely to excel in the teaching profession. This third round of screening also includes a "micro-teaching exam", in which students are evaluated in a classroom-like setting so that examiners are able to observe whether or not students work well with children (August, Kihn and Miller, 2010: 19).

Compensation for Finnish teachers is surprisingly modest, with teachers earning approximately 81 percent of per capita GDP (August, Kihn and Miller, 2010: 19). Performance pay and bonuses are not given to teachers. Graduate level training for teachers is paid for by the Finnish government, and students receive a living stipend. Interestingly, the complete absence of union politics within the Finnish teaching profession differentiates it quite substantially from the profession in other countries (Simola, 2005: 460).

Prestige is widely believed to account for the popularity of the teaching profession amongst top-performing students. Indeed, the traditionally favoured professions like lawyers, psychologists, physicians, engineers and journalists all trail teaching in terms of the number of applications at Finnish universities (Simola, 2005: 459). It is said that "people know that if you've been trained as a teacher you must be something really special" (Pasi Sahlberg in August, Kihn and Miller, 2010: 19). As a result of this signal of high quality emanating from the teaching profession, teachers have a substantial amount of autonomy in their work and are well-trusted by the public and the political and economic elite (Simola, 2005: 460). Teachers have a significant amount of authority in school policy and school management, textbook selection, course content, student assessment and budget allocations within schools and

importantly, are left to teach the prescribed curriculum in the way that they see fit (August, Kihn and Miller, 2010: 19-20).

### **5.8 Potential Costs of Incentive Systems and Alternative Solutions**

The last 2 examples of international incentive system (or in the case of Finland, a description of the factors that contribute towards the attractiveness of the teaching profession) have pointed out the possible risk associated with the introduction of incentive programs to the teaching profession. The higher teacher turnover that resulted from the labelling of schools as low-performing schools with the introduction of the accountability program had a negative effect that may not have been anticipated by the education authority. High teacher turnover makes it difficult to achieve any continuity and gather any momentum in schools reform, therefore hindering the opportunity for these low-performing schools to improve their performance. This is a useful demonstration of the possibility that negative externalities may be associated with the introduction of incentives. In the case of North Carolina, “punishing” poor-performing schools may worsen their situation.

The Finnish example illustrates the possibility of ensuring higher quality teaching without offering pay-for-performance type incentives. The prestige of the teaching profession in Finland ensures that it remains a highly selective profession, admitting only the high performing individuals into teacher training courses. Possibly as a result of this selectivity, teachers enjoy a high degree of trust and respect and have a large degree of autonomy in how the curriculum is taught. The superior performance of the Finnish education system is a testament to the high quality of teachers in Finland. However, the degree of selectivity and prestige observed in the Finnish teaching profession may be what is required for the achievement of the performance standards observed in that education system. It is debatable whether this is replicable in a developing country context.

Section 6 discusses incentives systems from around the world, from both developed and developing countries. The first 5 examples of specific incentive systems provide evidence on the impact on student performance and teacher behaviour, while the last 2 provide examples of potential costs associated with implementing incentive systems as well as an example of how the inherent incentives in the teaching profession (as opposed to those introduced through the implementation of an incentive scheme). These inherent incentives provide an

example of sorting into the teaching profession. Are incentives likely to be effective in improving educational quality? What are lessons for South Africa? What can we learn from international experience?

## **6. SOUTH AFRICA: WHERE DO WE STAND?**

This section explores the lessons that South Africa should take from international experience and analyses the incentives inherent in the teaching profession in this country. It looks at the prospects for professionalising the teaching profession as a means of enhancing accountability within the profession and finally looks at the performance monitoring system currently in place in the South African education system – the Integrated Quality Management System (IQMS).

### **6.1 Lessons from International Experience**

The first 4 incentive systems analysed in section 4 (those of Andhra Pradesh, Israel, Kenya and Pernambuco) illustrates the possibility for improvements in student performance through the implementation of incentive systems. In all cases, measured student performance improved with the introduction of the incentive system and the authors of the various studies (with the exception of the Kenyan case) indicated that the improvements appeared to result from genuine increases in teacher effort. However the only study in which results are reported when the incentive system is no longer running (the Kenyan example) showed that improvements in performance did not persist beyond the time period of the incentive system. Whether or not improvements in test scores were the result of gaming behaviour or genuine increases in teacher effort is therefore debatable. It still appears as if there is something to be said for incentivising teachers based on student performance.

An interesting aspect of incentive systems in place in Brazil and Chile, and one that is pertinent to the South African context is dividing the education system into subsections in the setting of incentives. In the case of Chile, schools compete within their regional socioeconomic category. In the case of Pernambuco in Brazil, school targets are set according to where they perform on the performance distribution. Given the extent of inequality in South Africa, comparing schools across socioeconomic quintiles would be grossly unfair. The educational and socioeconomic background of both students and teachers renders comparison within socioeconomic groups a much fairer format in which incentives may be introduced in South Africa. An appealing feature of the IDEB targets in place in Pernambuco

is that they account to some extent for school specific characteristics by setting targets at the level of the school. Consideration for the difference circumstances across schools is crucial in the context of South Africa and should be replicated where possible.

The discussion of the externalities resulting from labelling poor performing schools as such is illustrated in the example of North Carolina's ABC accountability framework. Attracting teachers to teach in undesirable locations is already a problem in South Africa. Labelling schools as under-performing is likely to exacerbate the problem, particularly if rewards are promised to teachers in schools that are publically recognised as performing well.

The Finnish example of illustrates the possibility of ensuring a high quality teaching force without explicitly introducing an incentive system. One of the key elements in the success of the Finnish education system is the prestige associated with the teaching profession. Section 4 details the competitiveness of teacher training courses, the high quality teachers that result from these courses and as a result, the high degree of trust and autonomy enjoyed by Finnish teachers. Perhaps the most startling feature of the teaching profession in Finland is the modesty of teacher wages relative to per capita GDP. This highlights the importance of the prestige associated with the teaching profession in ensuring a high-quality teaching force and brings to the light the "inherent" incentives in the Finnish teaching profession. Admittedly, it is unlikely that high ability individuals are likely to be drawn into the teaching profession without the promise of financial reward in a developing country context. The following section analyses the consequences of the South African teaching profession. Ironically, whereas the prestige of the teaching profession in Finland results in the profession's attractiveness despite the relatively low levels of remuneration, the relatively low levels of remuneration in the teaching profession render it a relatively less prestigious profession in the South African context.

## **6.2 Incentives Inherent in the South African Teaching Profession**

Although this paper deals predominantly with explicit incentive programs implemented to enhance student performance, it is important to recognise the incentives implicit in the salary structure of the teaching profession. Earlier sections discussed the possibility of attracting high-ability individuals to the profession through the implementation of incentives as part of an explanation for the "sorting" effect of incentives. Equally, if not more, important for the possibility of attracting high-ability individuals to the profession is their earnings potential over the entire span of their career. An exploration of teacher incentives should therefore

include a discussion of the incentives inherent in the profession, independent of those introduced with the express purpose of enhancing teacher effort.

An additional year of service in the teaching profession in South Africa is associated with approximately 1% increase in remuneration (Education Labour Relations Council (ELRC), 2011: 14). This appear to be the case “across the board”, with an additional year of service. This is a point of contention amongst stakeholders. Each salary notch is 1% higher than the previous one. An REQV level 14 qualified teacher is employed at notch 85. It is possible for that individual to progress at a rate quicker than 1%, but if an REQV level 14 qualified educator remains a classroom based educator (post level 1) will only have progressed to notch 125 after 40 years of service (ELRC, 2011: 14). Despite improvements in notch level progression<sup>3</sup>, the slow rate at which teachers progress up the salary scale is problematic and likely to continue to cause consternation amongst educators.

The implication of this slow progression up the salary scale plays out in the type of individuals attracted to the teaching profession. Lortie (1975) explains the phenomenon of staging in remuneration for different professions. Staging refers to individuals within a profession receiving different levels of remuneration at different stages of their career. He explains that fields in which an individual’s income increases substantially from one stage to the next usually reflect a significant change in status between different stages of the career. In contrast, in fields in which remuneration does not change significantly from one stage or phase to the next, differences in the status of individuals at different stages in their careers are less prominent and in many cases for all practical purposes, absent altogether (Lortie, 1975). Teaching (particularly in the South African case) is significantly un-staged. This has particular effects on the occupation as a whole.

Remuneration for teachers may be considered “front-loaded” in the sense that very little progression from the initial salary notch happens over time. The salary level at which teachers begin is therefore high relative to their eventual earnings potential. In comparison to many other professions requiring a degree for entry (such as law, accounting, engineering amongst others), teaching can be described as being relatively “career-less” given that there is

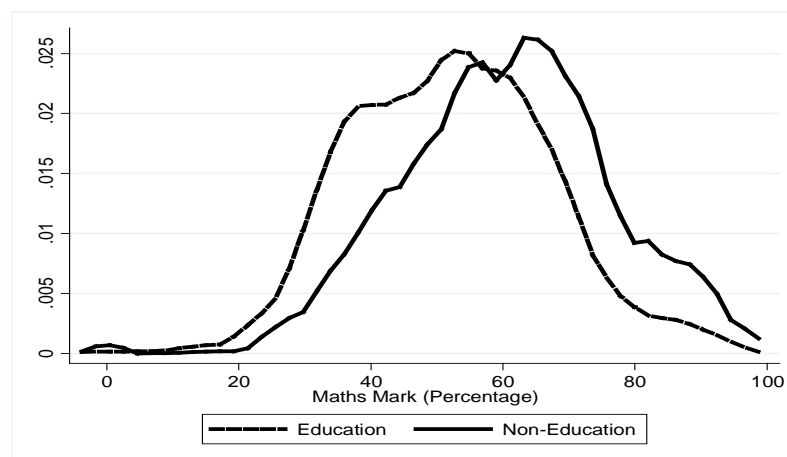
---

<sup>3</sup> Collective Agreement number 4 of 2009 allowed for notch progression based on years of service which saw an improvement in the salary progression of teachers who had been teaching for longer. For example, in 2007 teachers who had been teaching for 40 years with an REQV level 14 qualification earned just 29% more than entry-level teachers. The agreement saw these teachers earning 62% more than entry-level teachers in 2010. Similarly, educators with 11 to 15 years of teaching experience who were earning just 8% than entry-level teachers in 2007 earned 26% more than this group in 2010 (ELRC, 2011).

relatively little opportunity for upward mobility in terms of building a career and the status of a young entry level teacher is not markedly different from that of a teacher with some experience in the profession (ELRC, 2011: 15).

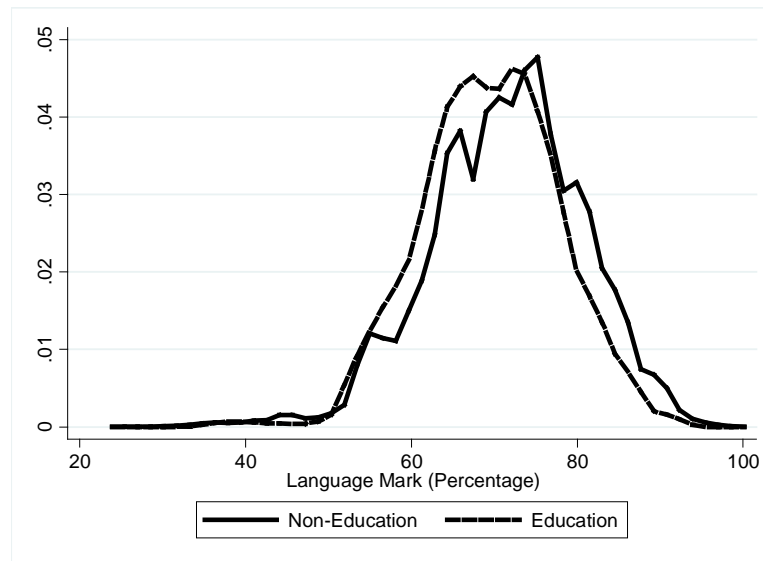
Overall then the structure of teacher salaries (i.e. the lack of staging or marked progression up the salary scale for South African teachers) has consequences for the type of individuals who enter the teaching force. If staging remuneration within professions does indeed play the roles described above, then it becomes important to understand the likely impact that this will have on the profession as a whole. Evidence from the first year student enrolled in different faculties at the University of Stellenbosch is provided below.

**FIGURE 3: Distribution of Matric Mathematics Marks for First Years Enrolments, 2005 – 2009**



Source: Data on first year enrolment at the University of Stellenbosch, 2005 - 2009

**FIGURE 4: Distribution of Matric Language Marks for First Years Enrolments, 2005 – 2009**



Source: Data on first year enrolment at the University of Stellenbosch, 2005 - 2009

Figures 3 and 4 respectively show the performance in matric Mathematics and Language for students enrolled in the first year of university studies in the Education faculty and in the other faculties between 2005 and 2009. The figures show that the distribution of marks for students enrolled in the education faculty lie to the left of those of students enrolled in other faculties, indicating weaker performance in both mathematics and language amongst students enrolled for education training.

The inherent incentive in the salary structure of the teaching profession in South Africa does little to attract high ability individuals to the profession. The slow rate of progression through the salary scale provides little motivation for teachers to remain in the profession, particularly those with high earnings potential in non-teaching professions. In terms of sorting, therefore, it is unlikely that high-ability workers will be attracted to the teaching profession in South Africa.

### **6.3 Prospects for Peer Pressure as an Incentivising Force: Professionalising Teaching**

In terms of using peer pressure to incentivise teachers into exerting an acceptable amount of effort, section 3 referred to the possibility of professionalism within teaching as a vehicle through which “guilt” can be created. This subsection examines professionalization in the teaching profession, with a specific focus on the context of South African teachers.

The ELRC's *Revised Salary Structure Proposal of 2011* indicates that teachers feel that their social and economic status have been eroded (ELRC, 2011: 28). There is a strong need amongst teachers to be recognised as professionals. There is a strong belief that the knowledge and skills requirements for teachers are equal to what is required in other professions, and that this is likely to be the case increasingly given that teaching now officially requires a degree for entry into the profession (ELRC, 2011: 28).

Pratte and Rury (1991: 64) describe professionalism as “an ideal to which individuals and occupational groups aspire, in order to distinguish themselves from other workers”. The characteristics of a profession from which professionals derive their prestige are a) mastery of a distinctive body of knowledge, b) the control of membership of the profession and c) a commitment to the well-being of their client.

Professionals are expected to have expert knowledge in their field and the profession is largely based around the identification with a distinctive body of knowledge. Organisations employing professionals function largely as communities of associates as opposed to being based on supervisory authority. The expertise of the professionals is the basis for their professional autonomy and authoritative power (Ambroise and Harley, 1988).

In terms of controlled access to the profession, the creation, diffusion and implementation of the professional standards of practice are assumed collectively by members of the profession. The licensing and to some extent the education of members is controlled by members of the profession and entrance into the profession (typically graduate level university study) is rationed to those achieving a minimum level of competency in their education up to that point. Having completed the academic program, candidates are required to pass demanding tests of their theoretical knowledge as well as complete a kind of internship during which they are continuously evaluated. The function of controlling the membership of the profession therefore serves the function both of controlling induction into the profession as well as the function of ensuring that the acquisition of knowledge is standardised across the profession (Pratte and Rury, 1991: 66).

Finally, professional practitioners commit to the pursuit of the welfare of their clients, usually through the acceptance of codes of ethics established by professional peers. Client dissatisfaction with the service of professionals is therefore dealt with through the initiation of a legal process rather than by reporting to the supervisor of the professional (Ambroise and Harley, 1988).



Lortie (1975) explains that the educational preparation of teachers is relatively general by comparison to that required by individuals entering into what are typically regarded as professions in the labour market. Teacher education does not require the same degree of intellectual rigour required in professions such as law, medicine or engineering. Education lacks the scientific roots or scholarly development that characterise these professions (Lortie, 1975), and the fact that teacher education and training takes place largely in the format of lecture and discussion, whereas traditional professions often require variations in settings such as laboratories, in order to master skills necessary to practice the profession. Teaching therefore does not require mastery of distinctive body of knowledge to the same extent that other widely-recognised professions do.

Controlling access to the teaching profession in South Africa is not characteristic of “an ideal to which individuals and occupational groups aspire, in order to distinguish themselves from other workers” (Pratte and Rury, 1991: 72). Charles Simkins (2010: 11) reports that some 45% of new teacher registrations with the South African Council of Educators (SACE) in 2009/10 were provisional registrations given to teachers with less than the required qualifications. Therefore, a very large portion of individuals entering the occupation and being allowed to practise as teachers legally did not have the required qualifications (Simkins, 2010: 11). This stands in stark contrast to the pursuit of prestige and exclusivity inherent in controlling access to a profession.

South African teachers are exposed to approximately 150 to 160 practice teaching days (in the case of a 3 year teaching qualification), 200 to 220 practice teaching days (in the case of a 4 year teaching qualification) or 50 to 55 days (in the case of a 1 year postgraduate teaching qualification) over the duration of their teacher education (University of Stellenbosch, 2012). This is justification for Lortie’s (1975) observation that “one of the striking features of teaching is the abruptness with which full responsibility is assumed.” By comparison to professions requiring extensive residency or clerical work once formal education has been completed, the “internship” required by teachers is significantly less rigorous.

It seems then that teaching in South Africa (as in many other countries) is ill-fitted to professionalization. The nature of the knowledge required to enter the occupation and the absence of the rigorous entry procedures that exist for other professions make it difficult to justify a call for the professionalization of teaching in its current state. The teaching profession would have to be “recast” as one requiring rigorous preparation and one which

holds a fair amount of prestige. Countries in which teaching is considered a prestigious profession have however achieved phenomenal educational success.

In the absence of strong professionalism, peer pressure may incentivise teachers if they are able to monitor each other's behaviour, if they are able to exert pressure on teachers who do not exert the required amount of effort, and if this pressure is likely to result in a change in behaviour. Section 5.4 explores the monitoring system currently in place amongst South African teachers and considers the probability that such monitoring would result in increased effort amongst teachers.

#### **6.4 Mutual Monitoring: The Integrated Quality Management System**

In the absence of professionalization, peer pressure is created through mutual monitoring of workers. The use of peer pressure to enhance effort levels requires that monitoring by peers actually result in enhanced effort, which in turn results from the possibility of imposing punishment on team members who do not pull their weight.

Currently, the system whereby teacher performance is measured and recorded in South Africa is the Integrated Quality Management System (IQMS). The IQMS handles teacher evaluation through self-evaluation by teachers and through a development support group (DSG) (ELRC, 2003). The self-evaluation and the evaluation conducted by the development support group use the same instrument so that teachers are familiar with the criteria according to which they are to be judged. The DSG is comprised of the teacher's immediate senior (i.e. head of department, or deputy principal in the case of a head of department) are one peer in their field of specialisation. They may choose which of their peers they would like to be part of the DSG. Teachers are graded according to 12 performance standards and both the self-evaluation and the evaluation conducted by the development support group are considered in the overall evaluation. The evaluations happen once every year with primary objective of helping teachers develop growth plans (ELRC, 2003). Teachers are therefore involved directly in monitoring the performance of their peers. However, this is intended to have a very nurturing and developmental role and is most definitely not designed to enable teachers to apply any kind of pressure or disciplinary actions against their peers. Furthermore, the fact that teachers are allowed to choose the peer who will form part of the DSG provides the opportunity for teachers to influence the type of assessment they may receive from the group members. Although it is stipulated that educators must appoint peers from their field of specialisation, there is no other requirements or restrictions on who teachers may choose to

be part of the group. It is therefore possible that teachers will choose peers most likely to provide them with a better assessment. In this way, distortion may occur insofar as a teacher is understood to influence the outcome of the performance measure.

Peer pressure as a mechanism for incentivising teachers is unlikely to prove effective. The fact that it is a once-yearly exercise, combined with the ability of teachers to choose their assessors means that it is highly unlikely that teachers will be able to exert any kind of pressure on their co-workers, and further that increased monitoring will result in higher effort levels.

## **REFERENCES**

Auguste, B, Kinh, P, and Miller, M. 2010. Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching. A publication by McKinsey and Company.

Baker, G. 2002. "Distortion and Risk in Optimal Incentive Contracts" in *The Journal of Human Resources*. 37(4). 4 – 24.

Bruns, B, Evans, D and Luque, J. 2012. *Achieving World-Class Education in Brazil*. Washington, DC: The World Bank.

Bruns, B, Filmer, D and Patrinos, H. 2011. *Making Schools Work: New Evidence on Accountability Reforms*. Washington, DC: The World Bank.

Che, Y and Yoo, S. 2001. "Optimal Incentives for Teams" in *American Economic Review*. 91(3). 525 – 541.

Clotfelter, C, Ladd, H, Vigdor, J and Diaz, R. 2004. "Do School Accountability Systems Make it More Difficult for Low-Performing to Attract and Retain High-Quality Teachers?" in *Journal of Policy Analysis and Management*. 23(2). 251 – 271.

Clotfelter, C, Ladd, H and Vigdor, J. "Who teaches whom? Race and the distribution of novice teachers" in *Economics of Education Review*. 24 (2005). 377 – 392.

Delannoy, F. 2000. "Education Reforms in Chile, 1980 – 1998: A Lesson in Pragmatism" in *Country Studies*. 1(1). 1 – 94.

Green, J. 2011. *Education, professionalism and the quest for accountability: hitting the target but missing the point*. New York: Routledge.

Dixit, A. 2002. "Incentives and Accountability in the Public Sector: An Interpretive Review" in *The Journal of Human Resources*. 37(1). 696 – 727.

Education Labour Relations Council (ELRC), 2003. *Integrated Quality Management System (IQMS) for School-Based Educators*. Pretoria: ELRC.

Education Labour Relations Council (ELRC). 2011. *Revised Salary Structure Proposals*. Pretoria: ELRC.

Figlio, D and Winicki, J 2005. "The Effect of School Accountability Plans on School Nutrition" in *The Journal of Public Economics*. 89(2). 381 – 394.

Glewwe, P, Illias, N and Kremer, M. 2003. "Teacher Incentives". NBER Working Paper number 9671.

Gustafsson, 2006. "Learning from Chile's Teacher Incentives Experience". Unpublished paper. University of Stellenbosch.

Hanushek, E and Woessman, L. 2008. "The Role of Cognitive Skills in Economic Development" in *The Journal of Economic Literature*. 607 – 668.

Homstrom, B and Milgrom, P. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design" in *The Journal of Law, Economics and Organization*. 7. 24 – 52.

Hout, M and Elliot, S (eds). 2011. *Incentives and Test-Based Accountability in Education*. Washington, DC: The National Academies Press.

Jacob, B and Levitt, S. 2003. "Rotten apples: An investigation of the prevalence and predictors of teacher cheating" in *Quarterly Journal of Economics*. 118(3). 843 – 877.

Kandel, P and Lazear, E. 1992. "Peer Pressure and Partnership" in *The Journal of Political Economy*. 100(4). 801 – 817.

Koretz, D. 2008. "A Measured Approach: Value Added Models are a Promising Improvement, but No One Can Evaluate Teacher Performance" in *American Educator*. Fall 2008. 18 – 39.

- Lavy, V. 2009. "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics" in *American Economic Review*. 99(5). 1979 – 2011.
- Lazear, P. 2003. "Teacher Incentives" in *Swedish Economic Policy Review*. 23 (10). 179 – 214.
- Lortie, D. 1975. *Schoolteacher: A Sociological Study*. Chicago: University of Chicago Press.
- McMeekin, R. (2000). "Implementing school-based merit awards: Chile's experience. Technical notes". Country studies: Education reform and management publication series. 3(1). Washington, DC: World Bank.
- Mizala, A and Romaguera, P. 2004. "Teachers' Salary Structure and Incentives in Chile". Working Paper for the Department of Economics at Universidad de Chile. Number 193.
- Moloi, M. and Strauss, J. 2005. *The SACMEQ II Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education*. Harare: SACMEQ. [online]: <http://www.sacmeq.org/reports.htm>. Accessed: 24 October 2012.
- Muralidharan, K and Sundararaman, V. 2001. "Teacher Performance Pay: Experimental Evidence from India" in *Journal of Political Economy*. 119(1). 38 – 77.
- Neal, D. 2011. "The Design of Performance Pay in Education". NBER Working paper number 16710.
- National Education Evaluation and Development Unit, 2013. *National Report 2012: The State of Literacy Teaching and Learning in the Foundation Phase*. Pretoria: National Education Evaluations and Development Unit.
- Parandekar, S, Amorim, E and Welsh, A. 2008. "Prova Brasil: Building a framework to promote measurable progress in learning outcomes" in *En Breve*. 121. 1 – 4.
- Pratte, R and Rury, J. 1991. "Teacher, professionalism and craft" in *Teachers College Records*. 93. 59 – 72.
- Simkins, C. 2010. "The maths and science performance of South Africa's public schools". Working paper for the Centre for Development Enterprise. Number 1, September 2010. Available from [www.cde.org.za/attachment\\_view.php?aa\\_id=333](http://www.cde.org.za/attachment_view.php?aa_id=333). Retrieved: July 2011.

Simola, H. 2005. "The Finnish miracle of PISA: historical and sociological remainkrs on teaching and teacher education" in *Comparative Education*. 41(4). 455 – 470.

Spaull, N. 2011. "Primary school performance in Botswana, Mozambique, Namibia and South Africa: A comparative analysis of SACMEQ III. SACMEQ III Working paper No. 8. [online]: [http://www.sacmeq.org/reports.htm#working\\_papers](http://www.sacmeq.org/reports.htm#working_papers). Accessed: 7 April 2013