

“What Gets Measured Gets Done”: Headteachers’ Responses to the English Secondary School Performance Management System

Deborah Wilson¹
Bronwyn Croxson²
and
Adele Atkinson³

¹*CMPO, University of Bristol*

²*New Zealand Treasury*

³*Personal Finance Research Centre, University of Bristol*

August 2004

Abstract

English secondary schools operate within a performance management system, which includes league tables reporting school performance across a number of indicators. This paper reports the results of an interview-based study, showing that head teachers care about their school’s place in the league tables, and that they believe this system affects behaviour. The effects they identify include some unintended consequences, not necessarily related to improved overall school performance, including focusing on borderline students who can boost a pivotal indicator: the number of students gaining five A*-Cs at GCSE. This behaviour reflects, in part, the dual role played by headteachers: they are both educationalists (serving the interests of all pupils); and school marketers, concerned with promoting the school to existing and prospective parents. The behaviour is also consistent with economic theory, which predicts a focus on that which is measured, potentially at the expense of that which is important, in sectors characterised by incomplete measurement, by multiple stakeholders and containing workers with diverse objectives. We conclude that, given that performance indicators do affect behaviour, it is important to minimise unintended consequences, and we suggest the use of value-added indicators of student performance.

Keywords: education, performance measures

JEL Classification: I2

Acknowledgements

Thanks to the Leverhulme Trust for funding this work through the Leverhulme Centre for Market and Public Organisation, and to Simon Burgess, Carol Propper and Marisa Ratto for comments on previous drafts. Particular thanks to our respondents for agreeing to take part in this study. The views expressed in this paper are our own.

Address for Correspondence

Deborah Wilson
CMPO
School of Economics, Finance and Management
University of Bristol
12 Priory Road
Bristol BS8 1TN
D.Wilson@bristol.ac.uk

1 Introduction

Performance management in government has received increasing interest since the 're-inventing government' movement of the late 1980s (Osbourne and Gaebler 1992). In the US, performance measures have been employed in many federal and state programmes dealing with, for example, job training, welfare to work, and child support enforcement, as well as in the health care sector (Marshall et al 2000), while the No Child Left Behind Act put the use of such accountability systems based on student test performance into federal law in January 2002. In the UK, the Financial Management Initiative, introduced in the early 1980s, embodied performance management but was assessed as being unsuccessful in influencing public sector resource allocation or in increasing the degree of public sector accountability (Osbourne et al 1995; Sharifi and Bovaird 1995). The use of performance measures has increased through the previous Conservative and current Labour administrations. Performance targets, their publication, and the linking of such targets to the resources allocated by Treasury to government departments is now widespread in the UK public sector. In addition, individual providers of public services now may get specific rewards linked to their performance, either directly – the top performing hospitals gain 'earned autonomy', for example; or indirectly – schools that perform well attract more students and hence more resources.

The public sector has specific features which complicate the causal link between the measurement of performance and its improvement (Behn 2003). First, bureaucrats often serve multiple stakeholders, including service users, taxpayers and politicians (Dixit 2002). Second, partly a consequence of the first, public sector agencies often have to work towards several goals, some of which may be in conflict, and each of which may be complex and/or difficult to define. The multiple and sometimes vague goals of the public sector mean that performance relative to these goals is difficult to measure. At best, a performance measurement system will provide an imperfect picture of a complex process, which leaves scope for the performance measure to be 'gamed', i.e. it leaves scope for unintended responses (Smith 1995).

A third feature is that public sector workers may exhibit 'public sector' or 'intrinsic' motivation, which may cause them to respond in particular ways to the incentives embodied in a performance measurement system. Intrinsic motivation has been defined as occurring when an individual is "motivated to perform an activity when [he/she] receives no apparent reward except the activity itself" (Deci 1971, page 105, cited in Jones and Cullis 2003). Crewson (1997) distinguishes intrinsic rewards from extrinsic rewards (such as a pay rises or promotion). The distinction matters since it may affect the impact of rewards on behaviour. Extrinsic rewards may not work: they may 'crowd out' workers' intrinsic motivation and so have a negative impact on public sector performance (Frey 2000).

Le Grand (2003) argues that the move towards more market based systems of public service delivery, often incorporating performance measurement, represents a change in belief regarding motivation. Specifically, he argues that there has been a shift from believing public sector workers are 'knights', i.e. predominantly public spirited or

¹ According to Behn (2003, page 599), "what gets measured gets done" is perhaps the most famous aphorism of performance measurement.

altruistic, to thinking of them as ‘knaves’, i.e. motivated primarily by their own self interest. Le Grand argues that public sector incentive structures should be robust, in the sense that they should not rely on any particular assumption regarding motivation, but should appeal to both ‘knight’ and ‘knave’. In this way, provider incentives may be aligned in such a way as to help achieve the objectives of both service user and government as two key stakeholders.

This paper investigates the extent to which the performance management system currently employed in the English secondary school sector aligns schools’ and stakeholders’ objectives. In the English education sector, parents are active agents (‘queens’ rather than ‘pawns’ in Le Grand’s terminology), who are able to choose to which school they would prefer to send their child. Headteachers are also active, since parental choice between schools is informed by the annual publication of school performance tables, giving headteachers an incentive to improve (at least measured) performance in order to attract more students. We consider both the impact of what has become the key target performance indicator – the percentage of a school’s students gaining at least five GCSE passes at grade C or above (%5A*-C) – and the recently introduced value-added performance measures (PMs).

In the study reported in this paper we analyse headteachers’ perceptions of the incentives created by these alternative summary PMs. Our analysis is based on 21 semi-structured telephone interviews with headteachers of secondary schools in England, conducted between December 2002 and May 2003. This coincided with the publication of the first value-added performance measures, in January 2003, as part of the 2002 secondary school league tables². The interviews formed part of a wider project on the introduction of performance related pay for teachers (Burgess and Crosson 2001). Here, we report the results from one section of the interview, in which we asked our respondents questions on three related topics: (a) how they viewed the current league tables and their perceptions of other stakeholders’ views; (b) their responses to the key target indicator, %5A*-C; and (c) the likely impact of the new value-added performance measure.

We find that headteachers perceive parents as active users in the English education quasi-market, and that they do respond to the competitive pressure created by parental choice. There is evidence of ‘knavish’ as well as ‘knightly’ responses, including strategies intended to boost the %5A*-C indicator, regardless of the impact on overall student performance. While the introduction of a value-added PM was cautiously welcomed as a more accurate basis for measuring student progress and hence school performance, the majority of respondents did not expect it to lead to a change in their behaviour, primarily due to the current dominance of the %5A*-C measure. On the basis of these findings, we suggest ways in which the school performance tables may be altered so that organisational incentives are better aligned to both headteacher motivation and stakeholder objectives.

The rest of the paper is structured as follows. In the next section we describe and discuss the use of performance measures in the education sector, culminating in four specific hypotheses regarding how we might expect headteachers to respond to the English secondary school performance tables. We briefly outline our methodology

² <http://www.dfes.gov.uk/performance/tables/index.shtml>

and describe our sample in Section 3, before reporting our findings in Section 4. Section 5 concludes.

2 Performance measures in the English education sector

The education sector in England has been subject to relatively high levels of public monitoring since the implementation of the 1988 Education Reform Act and the creation of the education quasi-market. The reforms introduced open enrolment, overlapping catchment areas and devolved budgets based on per capita funding, with money following students. Parental choice is informed by two forms of publicly available performance measures: Ofsted reports; and the annual publication of summary PMs in school performance tables, commonly referred to as the league tables. Here, we focus on the latter. In this system, schools face an implicit financial incentive to improve their performance, if by doing so they attract more students and hence more funding³. Le Grand (2003) argues that this incentive structure is robust to both ‘knavish’ and ‘knightly’ motivations. Headteachers are the relevant decision makers and have the incentive to maintain the financial health of the school, either through a self interested desire to keep their own jobs, or through the more altruistic aim of creating a greater surplus with which to provide better services for students. As we will show, the design of the performance measurement system affects the precise balance between such motivations, as well as how they are manifest in headteachers’ behaviour.

The performance management system in the English education sector uses three types of summary PM to try and capture movement towards the goal of improving student performance (Kane and Staiger 2002). *Levels* are the raw output scores of a cohort of students at a specific point in time, often reported as the percentage of that cohort achieving a particular target. One key PM in England, for example, is the percentage of a school’s students who gain at least five GCSE passes or equivalent at grade C or above (%5A*-C)⁴. *Changes* aim to capture the improvement of successive cohorts at the same grade in the same school across time (such as the improvement measures currently published in England), while *gains* provide a measure of the progress of one cohort between two points in time. The value-added PMs, published for the first time in the 2002 English school performance tables, provide an example of a *gain* PM. The aim of using a value-added PM is to better isolate the impact school environment has on student progress between two points in time. It does this by incorporating prior attainment, which helps to account for factors beyond the school’s control, such as family background and other personal characteristics.

In the 2002 and 2003 English secondary school league tables the following PMs were published for each school: eight *level* PMs relating to GCSE exam scores and to Key Stage 3 test scores; two improvement or *change* measures; and indicators of student absence, both authorised and unauthorised. In addition, two value-added (*gain*) PMs

³ This contrasts with an explicit financial incentive scheme, in which a direct financial reward is made to the organisation or individuals within it – essentially pay for performance (Burgess et al 2002).

⁴ The National Curriculum in England comprises four Key Stages. Pupils take tests in English, maths and science at the end of Key Stage 1, 2, and 3 (at ages 7, 11 and 14 respectively). The end of Key Stage 4 marks the end of compulsory schooling at age 16, at which point pupils take a range of exams known as GCSEs and GNVQs.

were published for each secondary school: one provided an indicator of the average value-added by the school between the ages of 11 and 14 (Key Stage 2 to 3), and the other between the ages 14 and 16 (Key Stage 3 to 4)⁵.

The production of education is a complex process, so any one PM will at best be an imperfect measure of the multiple tasks undertaken by the school; indeed, some of these tasks may be inherently unmeasurable (Dixit 2002). One aim of publishing multiple PMs is to provide more information on this range of tasks. There is, however, a trade off between comprehensiveness and transparency: there is a danger that multiple indicators are unwieldy and hard for parents to evaluate. In practice, parents and other users often manage the complexity by focusing on one indicator – as has occurred in the English secondary education market where the headline PM has become the percentage of a school’s students gaining at least five GCSE passes at grade C or above (%5A*-C).

Koretz (2002) identifies three problems with inferring educators’ performance from that of their students: the limitations of measures employed; the difficulties in drawing inferences regarding gains in student performance; and the perverse incentives that may be created. Propper and Wilson (2003) discuss the first two in some detail; here we focus on the last, the implications of alternative PMs for the responses elicited from schools, given that the aims of the government may not coincide with those of schools and headteachers (who may have ‘knavish’ as well as ‘knightly’ motivations). Different PMs create different incentives for schools. As the government tries to improve student outcomes by implementing performance measurement, schools may respond with improved outcomes as well as other, unintended and less desired behaviour⁶. Here we highlight three manifestations of such behaviour in the education context.

The first follows from the fact that any summary PM is an incomplete measure of a complex process. This creates an incentive to concentrate on those parts of the process which are included in the summary measure, possibly to the detriment of other, less quantifiable, tasks. This type of behaviour includes, for example, the ‘teaching to the test’ strategies observed by Wiggins and Tymms (2002), Dære and Strayer (2001) and Jacob (2002). The second arises when a PM does not control for heterogeneity in the quality of the inputs. This occurs if schools are able to achieve better relative positions in *levels*-based ranking exercises by virtue of having more able students, rather than because they necessarily deliver better quality education. A PM based on raw test scores therefore gives schools (and headteachers) an incentive to select those students most likely to improve its ranking. There is evidence of schools responding to such PMs by engaging in ‘creaming’ strategies, either when admitting students (Gerwitz et al 1995; Whitty et al 1998), or when deciding which types of student should take which tests (Figlio and Getzler 2002; *Times Educational Supplement* 2002). One way to reduce the incentive for such behaviour is for the PM to account directly for differences across the student population, an argument for the use of value-added measures of school performance.

⁵ Details on how the latter value-added PM is calculated are given in Wilson (2004). Performance tables for each school can be viewed at <http://www.dfes.gov.uk/performance/tables/index.shtml>

⁶ The possibility that the publication of summary PMs may cause public agencies to exhibit dysfunctional behaviour is well documented in Smith (1995).

Finally, the design of PMs affects the incentives schools face when allocating resources across any given student population. As stated above, *levels*-based PMs are often reported in terms of the percentage of a school's population attaining a specific target. They therefore focus schools' attention on borderline students (Fitz-Gibbon and Tymms 2002), and create an incentive to shift their activities or to target their resources to students who might otherwise just miss the target. This may be to the detriment of the other students, and may or may not be welfare improving. Wiggins and Tymms (2002) and Deere and Strayer (2001) both provide evidence of such behaviour. A particular issue highlighted in the English secondary education sector is the incentive to focus on students who are perceived to be on the borderline between gaining a C and a D at GCSE – so-called C/D borderline students. Additional resources targeted towards this group may boost a school's ranking with respect to the %5A*-C indicator, but the impact on other students in the school is unclear (Mansell and Wright 2004).

In summary, the production of education is a complex process, with multiple stakeholders including the government and parents. It is difficult to capture the objectives of any one of these stakeholders using a set of summary PMs. Measurement difficulties are compounded if the objectives of stakeholders conflict with one another. But, as Behn (2003) argues, even good measurement is not enough: the aim of any performance measurement system is to *improve* performance. And here the role of the headteacher, and specifically his/her responses to the incentives created by the PMs employed, is central. The following quote from Schofield (2001, page 83) captures the complexity of the headteacher's role:⁷

“As a consequence of policy initiatives, managers of [education] institutions have had to learn to become ‘business’ managers, entrepreneurial contractors with state purchasers and public strategists, all within an accountability system built around a managed market, but still with central state funding.”

The motivation of headteachers determines how they respond to the incentives within the performance management system, while the design of the system determines the incentives created. In the English education system, the quasi-market provides an implicit financial incentive to improve (measured) performance, if parents respond by choosing to send their child to the school. There are also non-pecuniary rewards and penalties attached to the prestige associated with a good – or poor – league table position.

So (how) do we expect headteachers to respond? Given the above discussion, we put forward the following hypotheses:

- headteachers will respond to the incentives created by the annual publication of summary performance measures – the precise way in which they do so will depend both on the specific PM being considered and on their individual motivation;

⁷ The context of Schofield's study is the durability of bureaucracy amid 'new public management' reform in the National Health Service. The quote, however, applies equally well to managers of healthcare and education institutions.

- we expect headteachers both to try to improve performance and to employ strategies to game the system;
- the introduction of a value-added PM should mitigate the incentive for some of the dysfunctional behaviour caused by the focus on the %5A*-C indicator;
- the implicit financial incentive system, with money following the student, should mean that headteachers take parents' views seriously.

Before investigating these hypotheses, we briefly outline our methodology and provide some details of our sample.

3 Methodology and sample

The interviews reported in this paper were conducted as part of a wider project on the impact of the introduction of performance related pay for teachers and the impact of choice on educational outcomes. The sampling strategy and more details on the background to the larger project are outlined in Burgess and Croxson (2001). Here we report the results of just one section of the interview, which concentrated on eliciting headteachers' views on the impact of league tables and on the introduction of value-added performance measures.

We carried out 22 semi-structured telephone interviews, which were recorded and subsequently transcribed, between December 2002 and May 2003⁸. All but one of these interviews was with the school's headteacher; on one occasion the deputy headteacher was interviewed. We interviewed by telephone, rather than face to face, since our sample is geographically dispersed and comprises senior people, likely to have to rearrange times at short notice. The topics covered in the interviews were unlikely to be sufficiently personal to heads to need the sensitivity afforded by face-to-face interviews.

We analysed transcripts for emergent themes, which we then checked against each answer to particular questions. Where appropriate, we categorised and tabulated the answers to specific questions. Our subjects were guaranteed anonymity, and the results have been reported accordingly. We have included quotations to illustrate the different types of views expressed.

The schools in our sample comprised seven foundation, ten community and five voluntary aided schools. We interviewed three headteachers from grammar schools, none from secondary moderns, and 19 from comprehensives. 17 were co-educational, with two all girls schools and three all boys schools. There were three Roman Catholic and one Church of England school in our sample, and the rest were non-denominational. The number of students in each school ranged from around 200 to approximately 2,000, and 11, i.e. half, had a sixth form. Our respondents also had a wide range of other characteristics: they ranged from 13.6% to 99.2% of students gaining at least 5 GCSEs at grade C or above; from 0.0% to 29.3% of students with special educational needs (without a statement); and from 1.4% to 51.4% of students

⁸ One interview was cut short before the questions relevant to the results reported here, hence we report the results of 21 respondents in this paper. Note that we did not find any systematic differences between the responses of those interviewed before the publication of the first secondary school value added performance measures, in January 2003, and those interviewed after that date.

eligible for free school meals. Finally, there was also a fair geographical spread within our sample: from Cornwall to Sunderland, Dover to Liverpool, and including schools in major conurbations such as Birmingham, Manchester and Greater London.

4 Results

We asked our respondents questions on three related topics: (a) how they viewed the current league tables and their perceptions of other stakeholders' views; (b) their responses to the key target indicator, %5A*-C; and (c) how they viewed the likely impact of the new value-added performance measure.

(a) The current league tables

We began by asking an open question: **Do you know your school's position in the league tables?** We were interested in finding out with which schools headteachers compared their own performance, both in terms of national/local comparisons and in terms of school type, and also on which PMs they focussed. The majority were interested in their ranking relative to neighbouring schools, usually within the same Local Educational Authority (LEA), and within the diocese for one Roman Catholic comprehensive. Five headteachers said that they also compared their performance nationally, but only with schools of the same type. These five were the grammar schools and/or single sex schools from our sample, plus the smallest mixed comprehensive. As the headteacher of a girls grammar school said:

“So I usually know where I am; I will rigorously buy all the papers and look and chart up results of my school against the top 20 girls' grammar schools every year But I don't look down the table. I look at what I would call a benchmarking group of the top 20 or so girls' schools and the top 30 or so state selective schools, because I know it's not fair to look outside that group because they're not selective as we are.” (*grammar, girls*)

When asked which PMs the headteachers focused on, 18 out of the 21 respondents identified the %5A*-C target indicator. Two of the grammar school heads said that this indicator wasn't an issue for them as they would expect to achieve 100%. Each indicator was mentioned at least once: %5A*-G; %1A*-G; average point score; value-added; and both authorised and unauthorised absences. Only one head took an interest in this last indicator, and she/he did so “only to see who's telling the biggest lies”⁹.

So why the focus on %5A*-C? As one headteacher put it:

“Although it's the stupidest measure For one reason because it's raw data, and the other reason is of course because it's incredibly dependent on a small number of children at the threshold of performance. So it's the silliest measure anybody ever came up with, but it's totally engrained on the public psyche, isn't it?” (*comprehensive, mixed*)

⁹ Fitz-Gibbon (1996) discusses the massaging of truancy rates in response to their inclusion in the league tables.

Another head said:

“It’s only important to me in so far as I think it’s very important to my students because it gets them into certain places [6th form colleges] The world seems to have decided that five A to Cs is an indicator of something for our students it’s an important measure. They like the idea of coming from a school where you do well. There’s no doubt about that, and so do parents We’re now stuck with it still unless the government are going to do something really radical and do something with the exam system.” (*comprehensive, mixed*)

Out of all the indicators published, %5A*-C is always the headline figure, hence headteachers have to know about it, even if they do not believe it is useful.

What came across very clearly in the interviews was the wide range of indicators and measures employed within schools to monitor their own performance. Respondents discussed various ways in which they collected baseline data for students entering the school and how targets were subsequently set for each individual student in order to monitor individual progress. A wide range of value-added measures were discussed, including the ALIS, YELLIS and MidYIS systems administered by the CEM Centre at Durham University¹⁰. Additionally, several respondents continually monitored current school performance relative to its own past performance. In effect, therefore, both *change* and *gain* PMs are increasingly, and routinely, being employed as internal performance management tools, while one *level* PM – %5A*-C – continues to dominate the published league tables.

We then asked: **Who (else) do you think takes an interest in the league tables?**¹¹ This brought a mixed response. All but one headteacher said parents, some distinguishing between current and prospective parents. Governors and the LEA were the next two most common responses. Three heads mentioned the teaching staff and the impact of the school’s league table position on morale. There was a range of opinion on the importance these parties attached to the %5A*-C ranking position. For example, with regard to parents’ views, heads commented:

“Parents have a very mixed view of them [the league tables], some of them take no notice at all, others pour over them. There’s an element of society that looks on them in the same way as they look at the premier league table in the Sunday paper, you know they’re like a sort of football competition.” (*grammar, boys*)

“I think the parents’ view of the league tables in some cases has changed. I think that middle class parents are still quite hung up on the idea...very interested and very concerned, and feel that percentage point leagues have great significance. But I don’t think they’re as unsophisticated in their interpretation as they were five or ten years ago. I think they’re much more sophisticated now. And I think also they’re much more prepared and

¹⁰ <http://www.cemcentre.org>.

¹¹ Headteachers were able to give multiple responses to this question.

willing to look at other aspects of the life of the school. They're looking now at the position in league table as one factor only. They're concerned about the pastoral care of the school, the way in which the schools treat children." (*comprehensive, mixed*)

There was also a range of views regarding the interest of LEAs in league table positions, which may reflect the practice of different LEAs:

"The LEA are obsessed with league tables because they have their own league tables to perform in." (*comprehensive, mixed*)

"I think certainly people within local education authorities and even the government have now a better understanding of how schools are performing against the kinds of pupils that we have in our school. And I certainly think our local authority is less interested in the performance table position. They take a lot of trouble to get the kind of data which would enable them to see how the school has actually performed rather than through league tables." (*comprehensive, mixed*)

There was more consensus when we asked: **Whose view is most important to you?** 14 respondents said parents¹²; three said the school's governors, two didn't specify and one said that his/her own view of the school's performance was the most important to him/her. The most common reason why parents' views were held in such high regard by the respondents can be illustrated by the following response:

"The parents. Because the school actually exists, as you know, it's funded in the main by the number of kids who come through the door." (*comprehensive, mixed*)

Most of the headteachers we interviewed were very aware of the role they played in maintaining or increasing school rolls. Several spoke of the importance of marketing, including the importance of maintaining good links with local media in order that their school be reported favourably. The %5A*-C indicator was seen as central to marketing the school, although the respondents were well aware of the flaws of this indicator as a means of comparing the performance of different schools:

"No I don't think it's fair at all I mean we're a school as you know in a very deprived area with over 50% free school meals when two thirds of children come to our school with reading ages one year and many of them two years below the chronological reading ages, and we're expected to get those children up to the level of the grammar schools, because we're directly compared with them, in five years. That's absolutely ridiculous." (*comprehensive, mixed*)

"So you're not comparing like with like. And it's really getting that message over that just as I said earlier, you know, a school could be doing a tremendous job with its kids but you know if you are working with

¹² Two of these said parents and children, or the community in a wider sense.

plywood there's no way you're going to make boats which require oak.”
(*comprehensive, mixed*)

In summary, the headteachers we interviewed employ a range of *change* and *gain* PMs within the school to help improve student outcomes. They additionally find it necessary to focus on one published *level* PM, the percentage of students gaining at least five ‘good’ GCSE passes, largely because it is the primary focus of parental, media and, to a lesser extent, LEA attention. Out of this group, our respondents view parents as the most important stakeholder, reflecting their view of themselves as their school’s marketer.

(b) Responses to league tables and the %5A*-C indicator

We explored headteachers’ views about league tables in more detail, in order to gain insights into whether league tables in general, and target indicators such as %5A*-C in particular, created the kind of perverse incentives discussed in Section 2. We therefore asked heads: **Have you ever targeted resources at particular areas in order to try and improve the school’s position in the league tables? If so, how?**¹³

Responses suggest that headteachers have two (not mutually exclusive) types of strategy: strategies that target underachieving students generally, regardless of the existence of league tables; and strategies developed in response to league tables.

Ten respondents discussed the strategies they had in place to target underachievement across the full ability distribution, regardless of their league table position. Strategies include the use of mentors, improving the information and communication technology (ICT) facilities available to students, introducing more vocational courses, and providing ‘catch up’ classes in literacy and numeracy.

With respect to the second type of strategy, there was only one headteacher who discussed anything other than targeting students at the C/D border, which in itself is a reflection of the importance of the %5A*-C target indicator¹⁴. As one respondent said:

“By choosing five or more A to Cs and making such a big issue out of it, it’s no surprise that most schools put a huge amount of energy and resources into those students who are on the C/D borderline.” (*comprehensive, mixed*)

Eight headteachers stated that they currently target resources at C/D borderline students in order to improve their league table position, a further two said they had done so in the past, and one said that s/he had tried, but failed, to avoid doing this. Conversely, two said they simply didn’t have any spare resources to target in such a way, and six stated that they deliberately didn’t follow such a strategy. Within this latter group, we got some very definite responses:

“We deliberately don’t do that. I know many schools do, but we deliberately don’t We target resources to those that are under performing in general rather than on the C/D boundary only.” (*comprehensive, mixed*)

¹³ Headteachers were able to give multiple responses to this question.

¹⁴ One respondent was additionally interested in G/U borderline pupils (those on borderline between the lowest pass grade and an unclassified or fail), although to a lesser extent.

“And I haven’t jumped on the bandwagon of introducing vocational courses in order to colour the results.” (*comprehensive, boys*)

This second respondent was referring to the practice of putting students in for GNVQs, which can be worth up to four GCSEs at grade C, in order that only one additional GCSE (at grade C or above) is required for a student to hit the target. The ICT GNVQ introduced and now sold by Thomas Telford school was mentioned by several respondents¹⁵. As the respondent who said s/he tried to avoid such targeting explained:

“And of course when I first heard about it [the four GCSE equivalent GNVQ in ICT organised by Thomas Telford school] I stupidly thought ‘Oh, it’s all right, the government are reviewing GNVQs, that’s a scam, they’re not going to allow it to carry on.’ But of course, I wasn’t cynical enough, I should have recognised that actually that particular scam suits the government, because it helps schools to meet the unrealistic targets they’re setting.” (*comprehensive, mixed*)

The following quote illustrates this practice:

“...we started last year, we introduced a GNVQ course which is a double award and we deliberately targeted that at the middle of the road pupils, those pupils who might get 4 A to C passes or who might get 6 A to C passes if we put them into this course.” (*comprehensive, girls*)

In addition to entering students in general, and C/D borderline students in particular, for GNVQ exams, a wide range of strategies to target resources at C/D borderline students were discussed by our respondents. These included mentoring, Saturday revision classes and homework clubs, and more intensive monitoring of such students’ performance. The following quotes illustrate some specific examples of strategies to try and boost the key target indicator. These examples show that these targets have a substantial impact in schools:

“Well, we give them the ‘Gee up Neddy, what are you going to do about this?’ sort of exhortation advice. ... So they have Gee up Neddy 1 in September [of Year 11 (age 15/16)]. ...[then in January, after mock exams have been marked] we bash them over the head with Gee up Neddy 2 in which we discuss their predicted grades, whether things are on course or not. And obviously at that point we get heavy with the kids who’ve got four C grades and tantalising numbers of Ds on their list.” (*comprehensive, mixed*)

“A lot of strategies that we have been working on this term have been targeted at the C/D borderline student. ... You know, the C/D borderline people do make a difference. We’ve also targeted for instance a group of Year 8 [age 12/13] students. We were looking at this type of youngster, the C/D borderline youngster in Year 8, because obviously it’s not good enough just to target

¹⁵ According to the Times Educational Supplement, Thomas Telford school made £7 million pounds in two years selling its online ICT GNVQ courses (Mansell and Wright 2004).

them in Year 11 ... but it's this type of youngster that is on this borderline that could go either way that ... always makes in a small school such a difference to us. And what we're doing with these ten youngsters in Year 8, we've considered them to be the cultural architects, the ones that can influence others in their peer groups. And so they're fairly cool customers, you know well-regarded by their peer group, as well as being in this little middle cohort, as identified by their CATs or by their subject teachers. And we're putting in some extra self esteem raising workshops it's being closely monitored and then obviously continue to mentor them and then compare their progress with a group who haven't been so supported, a group of another ten youngsters, you know as a control group." (*comprehensive, mixed*)

"... but identifying your youngsters who are going to get Ds, who with a bit of a push might get Cs. I refer to them as the fragiles. And the fragiles we put into a separate form. So I've got one form group of 20 youngsters who are all fragiles. And we make sure that the tutor there spends a lot of time mentoring these youngsters and we put lots of time and effort into them. Because ... there's only 100 or so children in the year group, so each child is worth a percent. So if I get those 20 through along with my sort of 30 that were predicted to anyway, that would be very nice, thank you." (*comprehensive, mixed*)

The wider potential impact of targeting resources at the C/D borderline student was recognised by this last respondent. When we asked him/her whether his/hers was a successful strategy to get these borderline students from D to C, s/he replied:

"It is successful; as an educationalist I'm not entirely happy with it because I'm still thinking we're putting all these extra resources in these youngsters to keep A to C figures up, but maybe there's a youngster who's going to get an E which is really good for them and who might end up getting an F because we're not spending as much time with them. ... The bright kids still prosper I don't think they miss out at all. But I think the lower ability ones potentially do." (*comprehensive mixed*)

Headteachers are engaged in activities aimed at boosting the performance of students that have been identified as underachieving. While this is partly going on independently of the publication of the league tables, the centrality of the %5A*-C indicator seems to dictate what is perceived as 'underachieving', leading to a focus on those students on the C/D border. Such a focus may create a conflict between the role of the headteacher as an educationalist (serving the interests of all students) and his/her role as the marketer of the school.

So, headteachers recognise the flaws in the current key PM, and react to it in ways that may not be welfare enhancing. The next section discusses the extent to which this might be alleviated by the publication of the new value-added performance measures.

(c) The impact of value-added performance measures

As we discussed above, many schools already use variations of value-added measures internally as a performance management tool to measure individual student progress.

We asked: Do you think that the new [DfES] value-added indicators give a more accurate reflection of school performance than PMs that measure raw outcomes?

18 respondents answered this question. Out of these, only two were definitively against value-added PMs, and three were not sure whether they would be more accurate. Seven headteachers were positive about their introduction, while the remaining six were generally in favour of the use of value-added PMs in theory, but had reservations about the particular form currently being published in the DfES league tables and/or the methodology underlying its calculation¹⁶.

Their reservations included whether one PM should measure value-added from age 11 to age 16 (KS2-KS4) or, as is currently the case, whether there should be two measures: one covering ages 11 to 14 (KS2-KS3); and one covering ages 14 to 16 (KS3-KS4). The range of views expressed on this issue is reflected in the following comments:

“So I’m not really bothered whether it’s Key Stage 2 to Key Stage 4, or Key Stage 3 to Key Stage 4, or 2 to 3. I think what I am more interested in is: are those tests a fair indicator of what a child has learnt and are they measuring the things that we think are important to measure.” (*comprehensive, mixed*).

“I think that [publishing KS2-KS3 and KS3-KS4] is really rather stupid because of course they’ve been able to do that for the last two years where we’ve got Key Stage 2 to GCSE value-added, and we’ve plotted that in school. That looks very healthy for us. I like that one.” (*grammar, girls*).

“... Key Stage 2 is no reflection whatsoever of the children’s ability. Because you coach them. I mean you can’t blame them [primary schools] for doing it because they have league tables as well. I certainly think perhaps Key Stage 3 to 4. Because we’ve got control of that, it’s all in our school, it’s all with our kids. So I think that’s probably fairer” (*comprehensive, mixed*).

Other aspects of the current methodology that were questioned by our respondents included the fact that the government doesn’t publish uncertainty (or confidence) intervals for each school’s value-added PMs, which makes it difficult to judge the extent to which differences between scores are significant; the imposition of the cap at a student’s eight best GCSEs; and the oversimplification involved with using only one, school average PM, since it doesn’t take account of differences in subject difficulty¹⁷. One headteacher of an all boys comprehensive school would also prefer gender specific value-added PMs, as “we honestly believe there is a lowering in the rate of maturity of boys once they’re in the second phase” and hence that whole cohort value-added measures don’t accurately reflect the performance of single sex boys schools. Given these caveats and reservations, however, there was a general consensus amongst our respondents that the move towards using value-added PMs as the basis for measuring and comparing school performance was, in principle, a positive step forward:

¹⁶ Five of these six were interviewed after the 2002 league tables were published (i.e. after January 2003). Out of the seven who were most positive, three were interviewed after publication and four prior to the league tables being released.

¹⁷ Propper and Wilson (2003) discuss all these points in more detail.

“But having said all that, I think the value-added has to be enormously better than raw results.” (*comprehensive, mixed*)

“The idea is an excellent one and I’m pleased they’re doing it and that’s really how everyone should be judged. How it’s going to work out in practice, I reserve judgement.” (*comprehensive, mixed*)

“I’d be happier to be measured against value-added than I would against some other things, do you know what I mean?” (*comprehensive, mixed*)

We then asked: **Do you expect the publication of the new value-added indicators to have an impact on you as headteacher?** Six respondents said it would have an impact, primarily because it would more accurately reflect their school’s performance. Six were not sure or wanted to reserve judgement. Eight said that the publication of value-added PMs would not have an impact on them as headteachers. A range of reasons why not were expressed:

“...at the moment I don’t understand or trust it.” (*comprehensive, mixed*)

“It [the publication of VA] won’t affect me in the sense that we already spend a huge amount of time analysing that. It will only affect me if I’m not able to demonstrate that we’re making better progress than other schools round about.” (*comprehensive, mixed*)

“... if it comes out as I expect, it will make me quite happy and proud. You know be able to put my nose in the air, hopefully against some of those schools who’ve thought they were superior to us just because they have a better intake all the time.” (*comprehensive, mixed*)

There was more consensus when we asked: **Do you think the publication of value-added information would have an impact on other parties in the education system?** Several respondents suggested that school governors and LEAs would be interested, but that there was already a lot of value-added information being shared between these parties and the school via individual school performance management systems and the annual PANDA returns¹⁸. Several also made the same point in relation to their teaching staff. One headteacher thought that it would impact on teachers by giving them a more realistic idea of their own performance. Eighteen respondents discussed whether or not the value-added PMs would have an impact on parents. Of these, only one thought parents definitely wanted the new performance measures; one said that it depended on how they were marketed; and one did not know if or how parents would react. One talked of the need for schools to engage in a learning process with parents to show them what value-added scores mean. Two noted that different parents would react in different ways. As one headteacher put it:

“.... some parents are very involved and follow these things really carefully. Others leave them [their children] at the gate and say ‘you look after them’. So

¹⁸ PANDA returns are Ofsted’s Performance and Assessment Reports, made available annually to all schools as part of the Autumn Package. See <http://www.teachernet.gov/management/tools/panda> for more information.

there's going to be a range of responses. But on balance I think parents will be pleased to see progress." (*comprehensive, mixed*)

Over half the respondents to this question, however, did not think parents would engage with the new value-added performance measures, at least in the short to medium term. Two main reasons emerged for this. First, the complexity of these value-added PMs and the subsequent lack of understanding by the majority of parents. Second, the fact that this was yet another PM to add to the wide range of data already published on different aspects of school performance, and that, out of these, the %5A*-C target indicator was already engrained on the public psyche. The following quotes illustrate these points:

"What concern I have is that parents might get confused. But as parents are becoming more sophisticated looking at performance information I really don't know." (*comprehensive, mixed*)

"I don't think the public understand where these [value-added] indices come from ... and actually now I think people are quite blasé about the league tables. I think they've been satiated with them. And the more complicated that the government makes the statistics, the more it will confuse and put off the parents from even bothering to look because they're old hat." (*comprehensive, mixed*)

"I think the problem is that the population as a whole and newspapers and the media and everyone else will still go on the raw results." (*comprehensive, mixed*)

"I think it's so difficult for people. You know they're so wedded to the idea of good school equals good results I think they think value-added is just a more complicated way of coming up with the same thing. And so really it's going to be a very long time if ever before they're taken seriously by people out there in the media and so on. I just don't think they understand them." (*comprehensive, girls*)

So while our respondents were generally – if cautiously – positive about the move towards value-added PMs as the basis for measuring and comparing school performance, they expressed doubts that their introduction would have a substantial impact on behaviour. There seem to be two reasons for this. First, different types of value-added information are already used within the school and shared with stakeholders such as school governors and the LEA. Second, the new value-added PMs are unlikely to have an impact on parents, given their perceived complexity and the fact that they are additions to an already crowded league table dominated by the %5A*-C indicator.

5 Conclusion

The production of education involves multiple stakeholders and complex objectives, which a performance management system will only partially capture. There is therefore scope for undesired as well as desired responses. Headteachers play a central

role in the link between measurement of and improvement in school (and hence student) performance. How they respond in practice depends both on the incentives created by the particular performance management scheme employed, and on their individual motivation. In this paper we analyse 21 interviews with headteachers to investigate their views about, and strategies in response to, the publication of specific PMs used in the English secondary education sector.

Our findings suggest that headteachers are generally aware of their school's position in league tables, in particular with regard to the %5A*-C target indicator. While conscious of this PM's flaws as a measurement tool, the fact that it is the 'headline figure' used by other stakeholders means that they have to pay some attention to improving their school's performance relative to it. This may involve strategies to 'game' the indicator, often by re-allocating resources to those students targeted at the C/D borderline. While all respondents were aware of this strategy, not all followed it. The introduction of a value-added PM was generally, if cautiously, welcomed. The majority of respondents did not expect it to change their behaviour, however, given that it is perceived as a complex addition to an already crowded league table. Throughout the interviews it was parents who emerged as being the most important stakeholder for headteachers, who believed that parents' behaviour was unlikely to be affected by the publication of value-added PMs, at least in the short term.

We can therefore accept all but one of the hypotheses listed in Section 2. Headteachers do respond to the incentives created by the league tables, and we find evidence both of genuine effort to improve student performance and of some gaming of the system. Parents' views are taken particularly seriously by headteachers. The introduction of the value-added PM does not, however, seem likely to reduce the incentive for some of the dysfunctional behaviour caused by the focus on %5A*-C. The reasons for this seem to be particular to the performance management system employed in the English education sector, rather than to be related to the use of value-added as a measurement tool *per se*. Indeed, the use of value-added as an internal performance management tool within schools was widespread amongst our respondents.

The design of the performance management system is crucial here: it determines the balance between measured and unmeasured performance, between extrinsic and intrinsic reward. The intended purpose for each PM should determine both its form and the decision whether or not to publicly disclose the resulting information. Value-added measures are already widely used by schools as internal performance management tools. Given that value-added is a more accurate basis for measuring and comparing school performance, and so should be incorporated in the published league tables, our analysis has the following policy implications.

The specific form of the value-added PM needs to be considered, in particular with regard to its applicability to specific stakeholder(s) (Propper and Wilson 2003). The current value-added PMs provide measures of school performance which do not distinguish between the effects of resource levels, of teachers, or of peer groups. The indicators do not, therefore, accurately measure one important government objective: raising standards for a given resource base. However, they do provide an indication of the impact of the whole package, something of direct interest to parents when choosing schools. The importance our respondents placed on parents as stakeholders

suggests that such a value-added PM has the potential to facilitate an effective implicit financial incentive scheme in the education quasi-market, in which money follows students. In order for this to be realised, however, parents need to be informed. There hence needs to be a shift in focus, away from the %5A*-C indicator and towards value-added performance measures. This is likely to require both simplification of the existing English secondary school performance tables and a process of learning for parents regarding how to interpret the results. Consultation with headteachers regarding these changes may further help to ensure that provider incentives and stakeholder objectives are better aligned. In this way, the use of value-added PMs may improve both what gets measured and what gets done in secondary schools.

References

- Behn, R.D. (2003), 'Why measure performance? Different purposes require different measures', *Public Administration Review*, 63:5, 586-606.
- Burgess, S. and B. Croxson (2001), *The impact of teacher pay reform on pupil attainment: an outline of the CMPO research project on the Performance Threshold*, CMPO, University of Bristol, Working Paper 01/036.
- Burgess, S., C. Propper and D. Wilson (2002), *Does performance monitoring work? A review of the evidence from the UK public sector, excluding health care*, CMPO, University of Bristol, Working Paper 02/049.
- Crewson, P.E. (1997), 'Public service motivation: building empirical evidence of incidence and effect', *Journal of Public Administration Research and Theory*, 7:4, 499-518.
- Deci, E.L. (1971), 'Effects of externally mediated rewards on intrinsic motivation', *Journal of Personality and Social Psychology*, 18:2, 105-115.
- Deere, D. and W. Strayer (2001), *Putting schools to the test: school accountability, incentives and behaviour*, Department of Economics, Texas A&M University (March).
- Dixit, A. (2002), 'Incentives and organizations in the public sector: an interpretive review', *Journal of Human Resources*, 37(4), 696-727.
- Figlio, D.N. and L.S. Getzler (2002), *Accountability, ability and disability: gaming the system?*, http://bear.cba.ufl.edu/figlio/fig_getz.pdf (April), accessed 14/10/02.
- Fitz-Gibbon, C.T. (1996), *Monitoring Education: Indicators, Quality and Effectiveness*, London: Cassell.
- Fitz-Gibbon, C.T. (1997), *The Value-Added National Project: Final Report: Feasibility Studies for a National System of Value-Added Indicators*, London: School Curriculum and Assessment Authority.
- Fitz-Gibbon, C.T. and P. Tymms (2002), 'Technical and ethical issues in indicator systems: doing things right and doing wrong things', *Education Policy Analysis Archives*, 10:6, <http://epaa.asu.edu/epaa/v10n6>, accessed 12/02/02.
- Frey, B. (2000), 'Motivation and human behaviour', in P. Taylor-Gooby, *Risk, Trust and Welfare*, Basingstoke: MacMillan, 31-50.
- Gerwitz, S., S.J. Ball and R. Bowe (1995), *Markets, Choice and Equity*, Milton Keynes: Open University Press.
- Jacob, B.A. (2002), *Accountability, incentives and behaviour: the impact of high-stakes testing in the Chicago public schools*, NBER Working Paper 8968, NBER: Cambridge MA.

Jones, P. and J. Cullis (2003), 'Key parameters in policy design: the case of intrinsic motivation', *Journal of Social Policy*, 32:4, 527-547.

Kane, T.J. and D.O. Staiger (2002), 'The promise and pitfalls of using imprecise school accountability measures', *Journal of Economic Perspectives*, 16:4, 91-114.

Koretz, D.M. (2002), 'Limitations in the use of achievement tests as measures of educators' productivity', *Journal of Human Resources*, 37:4, 752-777.

Ladd, H.F. (1999), 'The Dallas school accountability and incentive program: an evaluation of its impacts on student outcomes', *Economics of Education Review*, 18:1, 1-16.

Le Grand, J. (2003), *Motivation, Agency and Public Policy: Of Knights and Knaves, Pawns and Queens*, Oxford: Oxford University Press.

Mansell, W. and G. Wright (2004), *Easy way to the top of the league?* Times Educational Supplement, page 8, 16/01/04.

Marshall, M., P. Shekelle, R. Brook and S. Leatherman (2000), *Dying to Know: Public Release of Information about Quality of Health Care*, London: Nuffield Trust.

Osbourne, D. and T. Gaebler (1992), *Reinventing Government*, Lexington MA: Addison-Wesley.

Osbourne S., T. Bovaird, S. Martin, M. Tricker and P. Waterson (1995), 'Performance management and accountability in complex public programmes', *Financial Accountability and Management*, 11: 19-37.

Propper, C. and D. Wilson (2003), 'The use and usefulness of performance measures in the public sector', *Oxford Review of Economic Policy*, 19:2, 250-267.

Schofield, J. (2001), 'The old ways are the best? The durability and usefulness of bureaucracy in public sector management', *Organization*, 8:1, 77-96

Sharifi, S. and T. Bovaird (1995), 'The financial management initiative in the UK public sector: the symbolic role of performance reporting', *International Journal of Public Administration*, 18: 467-90.

Smith, P. (1995), 'On the unintended consequences of publishing performance data in the public sector', *International Journal of Public Administration*, 18: 2/3, 277-310.

Times Educational Supplement (2002), *League table bonus attracts schools to vocational option*, 23 August.

Whitty, G., S. Power and D. Halpin (1998), *Devolution and Choice in Education: the School, the State and the Market*, Milton Keynes: Open University Press.

Wiggins, A. and P. Tymms (2002), 'Dysfunctional effects of league tables: a comparison between English and Scottish primary schools', *Public Money and Management*, 22:1, 43-48.

Wilson, D. (2004), 'Which ranking? The impact of a 'value-added' measure of secondary school performance', *Public Money and Management*, 24:1, 37-46.