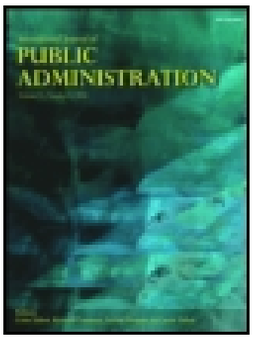

International Journal of Public Administration

ISSN: 0190-0692 (Print) 1532-4265 (Online) Journal homepage: www.tandfonline.com/journals/lpad20

On the unintended consequences of publishing performance data in the public sector

Peter Smith

To cite this article: Peter Smith (1995) On the unintended consequences of publishing performance data in the public sector, International Journal of Public Administration, 18:2-3, 277-310, DOI: 10.1080/01900699508525011

To link to this article: https://doi.org/10.1080/01900699508525011

INT'L J. OF PUB. ADMIN., 18(2&3), 277-310 (1995)

ON THE UNINTENDED CONSEQUENCES OF PUBLISHING PERFORMANCE DATA IN THE PUBLIC SECTOR

Peter Smith
Department of Economics and Related Studies
University of York
York YO1 5DD
United Kingdom

ABSTRACT

Most performance indicator schemes in the public sector have been implemented on the assumption that they will yield benefits in terms of efficiency and equity. Less attention has been paid to the potential costs of such schemes. Drawing on experience from a range of sources, this paper identifies eight consequences of publishing performance data that are not necessarily intended, and which are likely to be dysfunctional. The paper gives examples of such phenomena from the UK public sector, and suggests ways in which they can be mitigated. While not challenging the desirability of publishing performance data, the paper concludes that the performance indicator philosophy is based on inadequate models of production and control. Most performance indicator schemes will therefore fail unless serious consideration is given to the deficiencies described in this paper.

INTRODUCTION

The definition of performance in a public sector organization is often elusive. This has not prevented governments throughout the world from publishing an increasing volume of data pertaining to various aspects of public sector performance.[1] In different contexts, such data have become known as performance measures, performance indicators, or reports of service efforts and accomplishments. This paper makes no attempt to distinguish between possible nuances of meaning attached by some commentators to these terms. Instead, it seeks to describe in very general terms some of the unintended consequences that might arise when performance information of any sort is made available.

The arguments for publishing performance data are diverse. However, chief amongst them is the desire to secure ***control*** of the public sector organization.[2] Such control might be external, of the sort that legitimate stakeholders - such as voters, service users, elected representatives and taxpayers - might wish to exercise. Or the control might be internal, as sought by managers within the organization. Hofstede[3] calls the former ***political*** control. The latter sort of control might be thought of as ***managerial*** control.

Between these two extremes, in many countries an increasing number of hitherto unitary public sector organizations are being divided into quasi-autonomous business units, typically resulting in the creation of a ***purchaser*** organization (the principal) and a ***provider*** organization (the agent). For example, under the UK government's "Next Steps" initiative, the previously monolithic central government Department of Social Security has been split into a small central policy unit and a number of agencies, of which the most important is the Benefits Agency, responsible for administering a wide range of social security benefits.[4] The relationship between principal and agent, previously internalized, must now be conducted through the medium of a formal contract, and performance data usually play an important part in securing

conformity of the agent to the requirements of the principal. The model of control this arrangement gives rise to might be termed *agency* control.

Our emphasis on control appears to contradict the common assertion that the principal objective of performance indicator (PI) systems is to enhance various notions of economic efficiency.[5] Certainly, efficiency and effectiveness may be important objectives of many of the stakeholders in a public sector organization. However, the pursuit of economic efficiency is only one aspect of control. In particular, an important consideration in many public sector organizations is that users of the service should be treated equitably. For example, underlying the UK National Health Service (NHS) is the notion that patients in equal "need" should have equal access to health care facilities, regardless of their personal circumstances or where they live. The massive set of PIs developed for the NHS therefore includes a large amount of data on waiting times for elective surgery, reflecting the interest of the central government controller in the equitable treatment of patients.[6]

Another interest of some stakeholders is the preservation of financial control, and adherence to cash limits. Thain and Wright[7] argue that the UK central government sees this consideration as the dominant criterion of control, often at the expense of allocative efficiency, and PI schemes are seen as an important element in securing financial control. Thus, although it embraces the pursuit of efficiency and effectiveness, the concept of control is broader, incorporating other legitimate interests of stakeholders, such as equity, probity and financial management.

The dominant philosophy informing the use of performance indicators is the notion of managerial cybernetics.[8] In its crudest form this envisages the managerial process as follows. Organizational objectives are identified. Performance indicators are developed to reflect these objectives. Targets are set in terms

of the performance indicators. Management then chooses action and effort intended to achieve the targets. Progress towards targets is monitored using the PIs, and - if there is a divergence from targets - new targets are set and appropriate remedial action is taken. So the process continues.[9]

Thus central to the cybernetic model is the notion of ***feedback***, as serviced by the system of PIs. The model is clearly based on phenomena in the physical and biological sciences in which systems are kept in control via negative feedback mechanisms.[10] However, "organizational control is a complex and ill-understood activity precisely because it involves an attempt to control a complex network of self-controlling human beings".[11] When dealing with human organizations, the system one is trying to control is intelligent. The humans within the organization can anticipate the actions of the controller and - if it is to their advantage - may take action to frustrate the wishes of the controller.

It is therefore the thesis of this paper that the implementation of a system of performance indicators is in most public sector organizations likely to have unintended consequences, in addition to - or in place of - the intended consequence of securing control. It is most emphatically not intended to challenge the desirability of publishing more data on public sector activity and achievements. Rather, the purpose of the paper is to warn of the unintended and possibly dysfunctional consequences of careless use of such data, and to suggest ways in which the adverse consequences of publication can be minimized. The ultimate criterion for judging the usefulness of a PI scheme is the magnitude of its benefits in relation to its costs. The paper suggests ways in which some of the unintended costs of such a scheme can be mitigated.

The emphasis throughout this paper is on the impact of PIs on the ***internal*** management of public sector organizations. Many PI schemes are ostensibly aimed at external stakeholders, such as

electors. Smith[12] argues that, although these schemes appear to be different in nature to those intended as internal control mechanisms, in practice they are likely to have a profound effect on the internal management of the organization. Similarly, internal management in organizations subject to agency control is likely to be strongly influenced by the terms of the agency contract. The findings of the paper are therefore likely to be relevant to any situation in which performance data - whether directed at political, agency or managerial control - play a significant part in guiding the activities of the organization.

The paper is organized as follows. The next Section introduces the framework of the discussion. Eight types of unintended consequence that might arise when implementing a PI system are then described, with examples from the UK public sector. The concluding section suggests mechanisms for mitigating dysfunctional consequences, and draws some general conclusions.

THE SOURCES

Three types of organization and their associated literatures yield experience relevant to the study of the behavioural impact of performance data. The first is the firm. Economists have advanced several theories explaining the existence of firms, of which Williamson's transactions costs theory is one of the most celebrated.[13] The important feature of the firm is that - to a greater or lesser extent - it relies on bureaucratic processes rather than markets to allocate resources internally, a function for which the provision and processing of performance data is vital.[14] Management accountants are centrally concerned with such activities, and the management control literature is replete with experience from the corporate sector.[15]

The second type of relevant organization is the centrally planned economy, in which a deliberate decision is made to seek

to plan the allocation of resources to all activities in the economy, and to repudiate the use of markets. In particular, the experience (and subsequent collapse) of the Soviet system offers valuable lessons for those interested in the dysfunctional consequences of an excessive reliance on performance data.[16] Thirdly, the direct experience of governments seeking to control their bureaucracies has been studied extensively by political scientists, and there is therefore a useful literature of existing experience with PI schemes.[17]

The characteristic common to all three types of organization is a reliance on direct bureaucratic control. In these circumstances, performance information is required for a number of purposes. Most fundamentally, it is required to gain an understanding of the production function of the organization, and to assess the performance of the organization relative to that production function. In practice, of course, except for the simplest tasks, it is rarely possible to know the precise form of the production function. Instead, it must be inferred, either by examining activities amongst a range of comparable organizations, or by observing the activities of a single organization over time. The performance of organizations relative to the production function must then be assessed using the same data. Thus, performance data serve a dual purpose: first, in assessing performance relative to the efficient production function; and second in determining desired progress towards efficient production, through the setting of targets.

It is this duality which gives rise to many of the difficulties described in this paper. In order to make the control scheme meaningful, agents must be rewarded, either implicitly or explicitly, for progress towards targets. This leads to two problems. First, to a greater or lesser extent, agents can anticipate how their targets will be set. And second, the link between efforts and performance is often very poorly understood. Therefore, unless the incentives inherent in the reward scheme are directly

compatible with the performance targets - which are in turn perfect representations of organizational objectives - the behavioural responses of agents are likely to be imperfectly matched to the requirements of the principal. The principal:agent literature suggests that any divergence of objectives between the two parties compromises the ability of the principal to secure an appropriate level of effort from the agent.[18]

UNINTENDED CONSEQUENCES

A scrutiny of the three branches of literature described in the preceding Section brings to light a huge number of instances of unintended behavioural consequences of the publication of performance data. This paper distils this extensive experience into eight distinct types of phenomenon, as follows:

(1) tunnel vision;
(2) suboptimization;
(3) myopia;
(4) measure fixation;
(5) misrepresentation;
(6) misinterpretation;
(7) gaming;
(8) ossification.

All eight are the result of a lack of congruence between the goals of the agent - as moderated by the implicit reward scheme - and the actual goals of the principal. In the first three the lack of congruence arises because of a divergence between organizational objectives and the measurement scheme. Numbers (4) and (5) result from an inability to measure complex phenomena with precision or fidelity. Problems (6) and (7) reflect an inability on the part of the controller to process performance data correctly, while (8) indicates an inability to respond to new circumstances. These unintended circumstances are now discussed in detail.

Tunnel vision

Tunnel vision can be defined as an emphasis by management on phenomena that are quantified in the performance measurement scheme, at the expense of unquantified aspects of performance. Thus, for example, maternity service managers in the UK National Health Service are increasingly being held to account by a single performance indicator: the perinatal mortality rate.[19] This indicator certainly reflects a very important aspect of the effectiveness of maternity services. However, most commentators would accept that there are other important unquantifiable objectives which maternity service managers should pursue, such as the minimization of handicaps in surviving infants, and the enhancement of the entire experience of pregnancy and childbirth for healthy mothers. There is nevertheless clear evidence that the emphasis on the quantifiable perinatal mortality rate is distorting the nature of maternity services, to the detriment of the non-quantifiable objectives.[20] Indeed, a report by the House of Commons Health Committee[21] makes just this point, and recommends a more balanced approach to the organization of maternity services.

The problem of unquantified objectives is particularly acute in the public sector for three reasons. First, the various stakeholders usually hold a large number of diverse objectives with respect to a public sector organization. It is often impractical or impossible to identify all of these objectives. It is for example instructive to note that - in developing a set of hundreds of performance indicators for local governments - the Audit Commission[22] was criticized for a lack of completeness! In any case, specification of more than a handful of objectives is almost certainly meaningless, because it is impossible to devise a managerial reward scheme that satisfactorily reflects achievement in more than three or four dimensions.[23] A decision therefore has to be made on which objectives to emphasize in the performance measurement scheme, possibly resulting in the omission of considerations which are known to be important.

Second, many important objectives in the public sector simply defy adequate quantification. The maternity service example above illustrates this point, and the attendant danger that unquantifiable objectives will be ignored in the scheme. The principal difficulty which distinguishes the public sector from private enterprise is the absence of the unifying concept of revenues as a measure of the benefits of the organization's activities.

And third, it is a common feature of public sector services that the ramifications of their activities extend well beyond the immediate target of service delivery. Thus, for example, the activities of schools have implications for truancy, and therefore for a whole range of broader social issues. In principle, therefore, it is desirable constantly to monitor the environment to detected unanticipated consequences of an agency's activities. Clearly, almost by definition, such externalities cannot be captured in a performance measurement scheme. Yet managers should be responsive to them, and stakeholders should be able to hold managers to account if adverse externalities arise. Undue emphasis on quantified performance may preclude such flexibility.

The problems of measuring all salient features of a public sector organisation's activity are therefore considerable. However, failure to capture important dimensions of performance in a measurement scheme is highly likely to distort behaviour away from that required by at least some stakeholders. In line with practice from the corporate sector[24], it is therefore desirable to seek to quantify every objective, however crudely, if it is intended to use the measurement scheme as a pivotal instrument of control.

An alternative method of overcoming the problems arising from the quantification problem is effectively to downgrade the importance of the measurement scheme, and instead nurture a sense of shared values with staff, so that their objectives are congruent with those of the organization. Ouchi[25] argues that "clan" control - whereby the centre relies on a sense of shared

objectives with staff rather than a formal control mechanism - may be appropriate in situations where there is poor ability to measure outputs and imperfect knowledge of the production process. In such circumstances, endemic to the public sector, Ouchi suggests that market or bureaucratic forms of control may be dysfunctional. For many years, the NHS relied on a form of clan control to motivate staff, and secure the health aims of society.

Finally, the above discussion indicates that no measurement scheme can hope to capture all the consequences of a public sector organization's activity, and that it is therefore important constantly to scan the environment to detect unanticipated externalities. This principle is central to the "mixed scanning" approach to planning advocated by Etzioni.[26] Any performance measurement scheme should therefore be embedded in a broader environmental monitoring system, which might incorporate controls such as peer review and accreditation..

Suboptimization

Suboptimization is the pursuit of narrow local objectives by managers, at the expense of the objectives of the organization as a whole. This phenomenon is potentially endemic to any hierarchical organization in which control is secured by means of explicit performance criteria. For example, the UK Department of Social Security is seeking to implement a simplified system of collecting data on personal circumstances from benefit claimants, who are expected to gain from a simpler process for claiming. However, this initiative has met resistance from managers in local benefit offices because the retraining required to implement the new systems threatens achievement of the targets they have been set, expressed in terms of rapid settlement of benefit claims.[27] Thus pursuit of a key organizational objective - the implementation of new technology - is compromised by the lack of congruence with targets being set for individual benefit offices. Great care

should be taken, therefore, to maximize the congruence between high-level objectives and the objectives set for devolved management.

Suboptimization is a particularly serious danger when managerial discretion is strictly devolved, as in the agency model of management, in which control is secured by means of formal contract. Many of the outputs of the public sector are the ***joint*** outputs of several agencies, and the fragmentation of services may compromise the inter-agency cooperation necessary to pursue the stakeholders' objectives. It is therefore often extraordinarily difficult in the public sector to mesh the targets set for the devolved units with the top level objectives of the organization as a whole. Where this is important, a more flexible style of control may be appropriate, in which pursuit of high-level objectives by devolved agencies can be rewarded outside the scope of the formal control scheme. However, care must be taken that this stratagem itself does not lead to dysfunctional consequences - for example if employees perceive that it results in inequitable treatment.

Suboptimization is also potentially important at the level of the individual. Most outputs of the public sector are the result of team rather than individual efforts. As a result, if the implicit reward scheme is directed at individuals, suboptimization can arise. For example, managers may find it difficult to motivate members of their team whose rewards are not perfectly congruent with those of the manager. In any case, even if perfect congruence can be secured, there arises the problem of "free-riding", whereby individual team members rely on the efforts of their colleagues to secure targets, judging that their own lack of effort may not compromise team performance.

The problem of suboptimization is not necessarily soluble. Instead, the controller needs to recognize that there is a trade-off between the beneficial incentive effects of a formal control mechanism and the dysfunctional consequences of suboptimization.

The optimal form of organizational design depends heavily on the nature of the activity undertaken by the organization. However, for many public sector activities, it is likely that the rigid model of control implicit in market or bureaucratic designs is likely to be inadequate.

Myopia

Many of the activities of the public sector yield rewards over very long time periods. Indeed, it could be argued that an important reason why markets fail to deliver an optimal level of public goods is the systematic inability of citizens to perceive their long term benefits. Such goods may therefore be more appropriately provided by the public sector. Performance indicators offer a snapshot of organizational activities. Where such activities have long term consequences, therefore, they are likely to be imperfect measures of organizational performance in two senses. First, the current PI is likely to indicate the results of managerial endeavour over many years. And second, the current PI cannot reflect the future consequences of current managerial actions. The PI therefore offers an imperfect reflection of the efficacy of current management. Of course, the current PI may offer some help in judging the performance of *previous* management, but in many circumstances such information is of limited usefulness.

Thus PIs can induce managerial ***myopia***, the pursuit of short term targets at the expense of legitimate long term objectives. For example, successive UK national governments have grappled with the problem of devising a suitable performance measurement and reward scheme for police services, and an influential report by an industrialist has proposed a much stronger emphasis on measurable performance, such as the proportion of reported crime that results in a prosecution.[28] Yet it can be argued that many of the measurable outcomes of police activity are the results of years of

endeavour, and that many objectives of the police service may be best served by the long-term pursuit of strategies such as improved relations with the communities served. Such strategies may divert effort away from activities yielding more immediate success, and do not necessarily bear fruit in the short term.

The problem of myopia is exacerbated by the short term career perspectives of many workers, an attitude encouraged by the use of fixed term contracts of employment. If a manager understands that the long term consequences of his or her actions will eventually be reflected in PIs, the long term objectives of the organization are more likely to be respected if the manager expects to be in post to bear the consequences of those actions. Thus Sheehy's advocacy of fixed term contracts for police officers[29] would appear to run the risk of accentuating the bias towards short term objectives intrinsic to any performance measurement and reward scheme.

Myopia is a constant danger with all performance measurement schemes, although its importance depends on the nature of the service under consideration. Thus relatively routine functions, such as refuse collection or administration of welfare payments, may not be affected to any great extent. However, services with more elusive objectives, such as education and police, may be profoundly susceptible to myopia induced by PI schemes, and great attention should be paid to minimizing any dysfunctional consequences.

Strategies to counter myopia include measuring *processes* addressing long term issues, rather than seeking to measure outcomes. For example, it may be more appropriate to hold a public health department to account with measures of vaccination rates rather than outcome in terms of the incidence of disease. Of course, in having recourse to process variables, the performance analyst must be aware of possibly dysfunctional consequences other than myopia. The above analysis also suggests

that the extent of myopia may also be reduced by encouraging workers to adopt longer term career perspectives. However, it must be recognized that performance measurement schemes are intrinsically myopic, and that long term issues can usually never be satisfactorily captured by PIs. It may therefore be necessary to encourage managers to pursue long term objectives be means other than formal PI schemes.

Measure fixation

Even if clear objectives can be identified for a public sector organization, devising comprehensive measures of performance to reflect those objectives can often be far from trivial. If a measure does not fully capture all dimensions of the associated objective, managers may be encouraged to pursue strategies which enhance the reported measure rather than further the associated objective.[30] Such ***measure fixation*** can be defined as an emphasis on measures of success rather than the underlying objective.

For example, there has been great concern in the UK at the length of time that NHS patients have had to wait for elective surgery. The UK government therefore has an objective of reducing such waiting times. Yet converting this apparently straightforward objective into operational measures of performance and associated targets is fraught with difficulty. The first attempt was reflected in a "Patient's Charter", which stated that no patient should wait more than two years for surgery.[31] This requirement does appear to have borne fruit, in the sense that health authorities have almost universally eliminated waits in excess of two years.

However, at least two (presumably) unintended consequences have arisen. First, there has been an increase in the number of patients waiting more than one year: indeed there is evidence to suggest that average waiting times have increased.[32] Further, early reports suggest that the patients who are benefiting from the

waiting time initiative are those requiring relatively minor procedures, at the expense of patients awaiting operations for more serious conditions. Of course it is possible that this consequence was foreseen and intended. However, it is more likely that the initiative is proving dysfunctional, in the sense that it is diverting care away from those in most need.

The second consequence of emphasizing long waiting times has been a reported change in behaviour on the part of consultant surgeons. In the NHS, patients are referred by general practitioners to such consultants, and the waiting time for surgery is measured from the day the consultant decides to place the patient on a waiting list. Surgeons can therefore improve their performance, as measured by waiting times, first by delaying the preliminary consultation with the patient; and second by delaying placing the patient on the waiting list after that consultation. Both of these actions may have adverse health consequences for the patients affected, without any compensating gains for other patients in the NHS. One proposal to overcome the first problem is that the waiting time should be measured from the date of referral by the general practitioner. However, the second problem is more difficult to address, because the consultant surgeon is the "gatekeeper" to surgical treatment, and is therefore able to control patients' entry onto the waiting list without any significant audit of outcome in terms of health.

Clearly one means of reducing the prevalence of measure fixation is to increase the number of measures by which a manger might be held to account. However, an increase in the number of measures may reduce the focus of managers, and leading to demoralization and a reduction in managerial effort.

A better approach is to recognize that most measures are proxies for output, and that the ultimate arbiters of the quality of that output are the organization's clients. In measuring the performance of the organization, it therefore seems appropriate to

seek the opinions of the clients.[33] A robust and acceptable client satisfaction measure should, in principle, obviate the need for other performance measures. However, in making this strategy operational, two problems arise: identifying the "clients"; and eliciting their views. In the corporate sector, there are two easily identified clients: owners and customers. Owners can express their satisfaction through their dealings on capital markets, and using their influence as shareholders. Customers can indicate their satisfaction through the market mechanism and their willingness to pay. In the public sector, the notion of the client is much more elusive. It is certainly not just the user. If, for example, the satisfaction of claimants were the only criterion of the UK Benefits Agency, there might be little incentive for the Agency to pursue objectives relevant to other stakeholders, such as the minimization of fraudulent claims. There are moreover usually only crude mechanisms for expressing satisfaction, such as elections and migration. Thus, great care must be exercised in defining the ambit of any client satisfaction measurement scheme.

Misrepresentation

If, in seeking to exercise control over a public sector organization, excessive reliance is placed on PIs, there is clearly an incentive for managers to manipulate data under their control to show their organization's performance in the most advantageous light. ***Misrepresentation*** is the deliberate manipulation of data so that reported behaviour differs from actual behaviour. The scope for misrepresentation is particularly broad in the public sector because many of the data used to measure performance and hold staff to account are under the direct control of those staff.[34]

Misrepresentation can take two forms: "creative" reporting and fraud. Creative reporting is possible when there is some discretion about how an event can be recorded, and can take many forms. In the corporate sector, much attention has been paid to

accounting choices, which can profoundly affect reported performance.[35] In the public sector, creative reporting is especially likely if some element of professional judgement is needed to describe an event. For example, if a doctor has a choice of two diagnosis related groups to which to allocate a particular patient, she has a strong incentive to choose the group which maximizes her apparent workload.

Clearly the possibility of fraud is also always present, particularly when the scope for external audit is limited. The incentive for fraud is likely to increase in proportion to the reliance placed on the PI scheme in determining managerial rewards.

Misrepresentation is potentially dysfunctional because it misleads the controller about the activities of the organization. This may lead to misallocation of resources, and inequitable treatment of staff and clients. The perception that a control mechanism is unfair may be a serious disincentive to goal congruent behaviour on the part of staff.[36] Clearly misrepresentation can be reduced by increased audit effort, and by threatening serious sanctions if misrepresentation is detected. Controllers must determine the optimal balance between the dysfunctional consequences of misrepresentation and the costs of monitoring.

In addition, it is desirable to give front-line staff the incentive to record data consistently and truthfully. This may be best done by integrating the performance data collection system into administrative information systems which staff use to run their own units. By being given "ownership" of the data in this way, front-line staff may find it in their own interests to ensure that the data are of high quality, as they facilitate efficient operation of their unit.

Misinterpretation

In practice, many of the production processes in the public sector are immensely complex, and building a realistic model of them is severely taxing. This problem is intensified by the fact that most public sector organizations must continue to operate however difficult the local environment, in contrast to firms, which can choose to cease trading in adverse conditions. Full account must therefore be taken of the external environment in which public sector organizations are operating when inferring the nature of the production function and assessing performance. Thus, even if the available data were a perfect representation of reality, the problem of interpreting the signals emerging from the data is often extremely complex. In other words, although in possession of all the facts, bounded rationality might cause the controller systematically to ***misinterpret*** them, and to send the wrong policy signals to the agent.

Although an organization's PIs can usually be interpreted only in relation to performance in other jurisdictions, judging performance on a truly comparable basis is often immensely complex. There are five reasons why there might be differences in the reported performance of two organizations:

(1) they might be pursuing different objectives;
(2) they might face different operating environments;
(3) they might have different costs of inputs;
(4) they might use different measurement methods;
(5) they might have different levels of efficiency.

Clearly disentangling these five determinants of variability is a major task.[37]

A particularly prevalent form of misinterpretation arises from the politician's preoccupation with issues of equity. Indeed the quest for equity is one of the most important features

distinguishing political from corporate systems of control. Yet, in spite of the importance of the concept, it is often difficult to translate the goal of equity into operational targets. In practice, because of the difficulties of interpretation noted above, it is often the case that crude measures of *equality* are used as proxies for equity objectives. In the absence of skilled interpretation, therefore, local political pressure is often for apparently poor performers simply to devote more resources to reducing the performance disparity from other organizations.[38]

Thus there is likely to be intense pressure in most public sector organizations for a reduction in the variation of published performance from that of other apparently similar organizations. The important point to note is that the pursuit of such equality may not be desirable on either efficiency or equity criteria. It may result in organizations diverting resources towards activities which do not yield as many benefits as might have been secured elsewhere. Convergence - the tendency for public sector organizations to pursue equality of reported performance - is therefore a potentially dysfunctional phenomenon.

A particularly serious potential for dysfunctional misinterpretation is occurring in the UK education system. The central government has decreed that the results of public examinations in all state schools must be published in the form of "league tables". This initiative is intended to enable parents to make more informed choices about the choice of school for their children.[39] However, the requirement to publish results in such a form has been strenuously opposed by many teachers and parents on the grounds that interpretation of raw performance data is complex, and that misleading inferences about school performance might therefore be made.

An example of the pressures that are likely to arise if such publication takes place is given by the UK government's presentation of the percentage of seven year olds securing

satisfactory results in national tests in different local governments.[40] In presenting these results, the Secretary of State claimed that "there are unacceptable discrepancies between the performance of different [jurisdictions]". However, no attempt was made to seek to understand why the variability arose. The implication was that local governments at the lower end of the performance range must seek vigorously to improve, regardless of whether efforts to that end are an efficient use of resources.

Because interpretation of public sector performance data is often beyond the scope of lay commentators, it becomes the domain of experts. In the corporate sector, investors can perceive an obvious personal advantage in securing accurate and timely advice about performance, and so are willing to pay analysts to interpret financial statements and other sources of performance data. In the public sector, however, there is no equivalent market for independent intermediaries to fulfil this role. Expert analysis of public sector performance data therefore has the characteristics of a public good: society might benefit from such analysis, but individuals are not prepared to pay for it. Because of this, many governments have set up public sector audit organizations to examine performance on behalf of various stakeholders. In the UK this role is fulfilled by the National Audit Office for central government, and the Audit Commission for local government and the NHS.

Notwithstanding the existence of audit offices, in many cases the only apparent source of analytic experts is the very organization the performance of which is being analyzed. For example, health professionals have an opportunity to lead the debate about the interpretation of their own PIs, and clearly have an incentive to bias the analysis to their own advantage. Smith (forthcoming) calls this phenomenon "interpretive" misrepresentation.

Clearly improvements in the supply of expert advice and in the awareness of ordinary citizens are therefore potential means of

reducing the extent of misinterpretation. In particular, it is important to acknowledge that variations in published performance will always exist, yet may signify very little more than stochastic variability. Perhaps the most important requirement is to seek to understand the *reasons* for variations in performance, a task for which the models of management accounting and operational research have been designed. It is often possible to formulate policy prescriptions sensitive to an organization's objectives only after carefully modelling the system under examination. The discussion of convergence moreover suggests that policy makers should have a very clear idea about the efficiency and equity objectives they are pursuing. Otherwise - even with correct interpretation of performance data - inappropriate policy signals might be sent.

Gaming

Fundamental to the difficulties encountered in managing the Soviet economy was the problem of offering managers appropriate incentives to improve productivity.[41] To be sure, rewards were offered for reaching targets. However, those targets were typically set with reference to previous behaviour.[42] Therefore, although managers had a modest incentive to fulfil targets, they had a much stronger incentive to keep future targets modest by chronically underperforming. They could blame external factors for any apparent failure, and - without market disciplines - could continue to function at modest levels of productivity. The phenomenon whereby targets are irreversibly raised in response to one year's productivity improvement is known in the Soviet literature as the "ratchet" effect.

The analogy with many public sector organizations is clear. As noted above, performance appraisal is highly complex, and it is often difficult to establish any benchmark for performance which is independent of the organization's own past performance.

Managers therefore have an incentive to minimize the apparent scope for productivity improvements, as any reported improvement in one year will result in increased expectations (and targets) for future years. Such behaviour is known as ***gaming***, which can be defined as the deliberate manipulation of behaviour to secure strategic advantage.[43] Thus, while misrepresentation leads to distortions in ***reported*** behaviour, gaming is the equivalent manipulation of ***actual*** behaviour. It is therefore potentially severely dysfunctional, as the Soviet experience suggests.

To illustrate the importance of the concept, consider the use of cost per pupil as a performance measure in schools. Incentives might be given to head teachers to reduce this PI. However, in doing so, the heads might make their own future targets tougher, if these are based on this year's costs. They therefore have to balance the rewards for achieving this year's target against the penalty suffered by the implicit reductions in rewards that might arise in future years. They can blame any failure to reduce costs on local socio-economic circumstances, and it is often difficult for the controller to produce directly comparable evidence to gainsay such arguments. Thus the school system controlled in such a manner may be chronically inefficient.

A number of strategies exist for the controller to minimize gaming. The first is to use a range of PIs. It is much easier to game only one PI. The complication of balancing several PIs makes successful gaming much more difficult. It may also be appropriate for the controller to maintain some uncertainty about which indicator is to be used to influence rewards. More generally, the reward structure should be flexible enough to respond to managerial efforts which contribute towards improvements in future targets as well as merely the achievement of current targets.[44]

A further strategy to minimize the gains from gaming is to develop benchmarks of performance which are independent of the

organization's past behaviour, perhaps using techniques such as data envelopment analysis.[45] Of course, as noted above, the use of comparative data to this end is complex, but if managers understand that their future targets will not be influenced to any great extent by current performance, then they are far more likely to seek out opportunities for improved productivity. In this context, the comparative data can be thought of as introducing a pseudo-market into the public sector, and fostering inter-organizational competition.[46]

Finally, it should be noted that managers are most likely to indulge in gaming when they expect to be in post for a long time, as the rewards of a modest target regime accrue over a number of years. Thus, in contradiction to the policy prescription for myopia, one strategy for reducing the importance of gaming is to encourage short term career perspectives amongst managers, perhaps leading them to expect a change of job every two or three years. In this way, the benefits to the manager of achieving this year's target appear large in relation to the benefit of underperforming as a means of maintaining easier targets in the future.

Ossification

One of the most important virtues claimed for competitive markets is that they continuously offer entrepreneurs incentives to seek out new opportunities. By contrast, the discussion above suggests that bureaucratic performance measurement schemes can inhibit innovation, and lead to ***ossification***: organizational paralysis brought about by an excessively rigid system of performance evaluation. The need to choose performance measures and set targets in advance means that new threats and opportunities may be ignored by managers. Moreover, the discussion of gaming suggests that a predictable system of performance evaluation may offer considerable scope for manipulation on the part of management.

The danger of ossification arises because of the inevitable delay in designing and putting in place a performance evaluation scheme, and the effort required to change it subsequently. Yet the discussion above suggests that any scheme is likely to be deficient to a greater or lesser extent, and so will need to be regularly reviewed and updated. Hofstede[47] notes that the simple closed loop cybernetic system may lead to "dynamic conservatism", and advocates a "second-order loop ... which periodically adjusts the standards of the first-order cycle". Thus there is a need to have a system to control the content of the PI scheme: in short, a control over the control mechanism. Now clearly the second order control cannot be explicit and rational, otherwise it becomes indistinguishable from the original PI scheme. Instead it must be messy and political, responding to new circumstances, the nature of which - almost by definition - is likely to be unforeseeable.

The difficulty the UK government has had in implementing its Citizen's Charter[48] illustrates the problems involved in seeking to infuse an entire administration with the PI philosophy. Many government departments were lukewarm in adopting the Charter's principles, and Hofstede's analysis suggests why this might be. The stark clarity and explicitness of PIs is antipathetic to the political process of responding to new developments. As a result, the PI philosophy is best suited to situations in which there is reasonable agreement on objectives, and relatively little danger of unexpected challenges, so the need for second order control is minimal. However, where the second order, political aspect of control is dominant, the PI philosophy may be inappropriate. In such circumstances, the need for regular political intervention in control may render the PI scheme redundant.

Ossification is therefore minimized by ensuring that managers are rewarded for behaviour that anticipates new challenges, even if the activity does not yield improvements in existing PIs. In addition, it is imperative to ensure that the PI scheme is kept under constant (probably annual) review. Clearly, constantly changing

the measurement scheme may adversely affect managers' motivation, and there is a need to involve all levels of management in any review. However, the inability to respond to new circumstances is intrinsic to most bureaucratic systems of control, and so mechanisms to avoid such inertia should be built into the design of any PI scheme.

DISCUSSION

The information technology revolution has brought about a massive reduction in the costs of providing performance data in the public sector, leading to the belief that more such data should be provided. It is not the purpose of this paper directly to challenge the desirability of publishing performance information. Rather, it is intended to emphasize that such dissemination introduces risks and side-effects that may not be anticipated. The paper categorizes the sorts of unintended consequences that may arise, and suggests means of minimizing adverse repercussions.

Economists would argue that the decision about which data to provide and how they should be used must be viewed within the context of a cost-benefit model. In principle, data should be provided up to the point where their marginal benefits equal their marginal costs. Benefits clearly include improvements in allocative and technical efficiency resulting from better informed decision making and goal congruent incentives built into the performance measurement scheme. Costs include the direct costs of collecting and processing data. And we have also shown that there exist a plethora of less direct potential costs of PI schemes. These can be mitigated using the suggested strategies. However, the strategies themselves impose costs. So the wise controller must balance a large number of conflicting pressures to determine how to implement a system of performance measurement.

Table 1 summarizes the strategies that have been suggested for minimizing dysfunctional consequences, and the specific problems

TABLE 1

Summary of Strategies for Minimizing Unintended Consequences of Performance Measurement

Strategy	Tunnel vision	Suboptimization	Myopia	Measure fixation
1. Involve staff at all levels	+	+	+	+
2. Retain flexibility in use of PIs	+	+	+	+
3. Quantify every objective	+		+	+
4. Keep system under constant review	+	+		+
5. Measure client satisfaction				+
6. Seek expert interpretation of PIs				+
7. Maintain careful audit of data				
8. Nurture long term career perspective			+	
9. Keep the number of indicators small	--			--
10. Develop independent benchmarks				

TABLE 1 Continued

Summary of Strategies for Minimizing Unintended Consequences of Performance Measurement

Strategy	Misrepre-senting	Misinter-preting	Gaming	Ossifi-cation
1. Involve staff at all levels	+	+	+	+
2. Retain flexibility in use of PIs	+	+	+	+
3. Quantify every objective			+	+
4. Keep system under constant review	+		+	+
5. Measure client satisfaction	+	+	+	
6. Seek expert interpretation of PIs	+	+	+	
7. Maintain careful audit of data	+	+	+	
8. Nurture long term career perspective			--	
9. Keep the number of indicators small		+	--	--
10. Develop independent benchmarks			+	

they are designed to mitigate. The symbol "+" indicates a putative beneficial effect, the symbol "-" a potentially detrimental effect. Note that four strategies address a large number of problems, and so are likely to be applicable in most situations:

(1) Involving staff at all levels in the development and implementation of PI schemes;
(2) Retaining flexibility in the use of PIs, and not relying on them exclusively for control purposes;
(3) Seeking to quantify every objective, however elusive;
(4) Keeping the PI system under constant review.

The importance of the next three strategies is more dependent on the type of organization being controlled, being most relevant when objectives are poorly defined and measurement of output problematic:

(5) Measuring client satisfaction;
(6) Seeking expert interpretation of the PI scheme;
(7) Maintaining careful audit of the data.

The final three strategies are designed to address specific difficulties - myopia, misinterpretation and gaming - and so should be considered when any of these is especially important. Note however that they may have negative effects relevant to other dysfunctional phenomena.

(8) Nurturing long term career perspectives amongst staff;
(9) Keeping the number of indicators small;
(10) Developing performance benchmarks independent of past activity.

The naive model of production on which the PI philosophy is based is the source of many of the difficulties noted in this paper. It is assumed that units of production are isolated entities which consume inputs and produce well-defined outputs in a static

framework. Indeed, the typical PI is the simple ratio of an output to an input. However, in practice, many of the outputs of the public sector are difficult to identify and measure; many are produced jointly with other organizations; and many unfold over a long period.

The designer's objective is to find a the pattern of control over the organization which secures the objectives of society. However, the optimal design of a PI scheme is inextricably linked to the nature of the organization's activities, and the mechanisms whereby it is controlled. A performance measurement scheme which fails to recognize that the control of public sector organizations is often necessarily a complex, ill-defined, dynamic process is doomed to fail. Certainly the provision of a great deal of data relating to performance may in many situations lead to a greater potential for control congruent with the objectives of stakeholders. However, this paper suggests that, unless that control is exercised with care and discretion, severely dysfunctional consequences may arise.

ACKNOWLEDGEMENTS

This paper has been presented to the UK Local Education Authorities Research Group, the Operational Research Society, the UK Department of Social Security and the Universities of Leeds and York. Thanks are due to Alec Nove, Tony Culyer, Alan Williams, seminar participants, the journal's referees and others for numerous helpful comments.

REFERENCES

1. Fédération des Experts Comptables Européens *Performance Measurement in Public Sector Management*, Chartered Institute of Public Finance and Accountancy, London, 1991.

2. Stewart, J. D. "The Role of Information in Public Accountability" in A. Hopwood and C. Tomkins (eds) ***Issues in Public Sector Accounting***, Phillip Allan, Oxford, 1984, pp. 13-34.

3. Hofstede, G. "Management Control of Public and Not-for-profit Activities." ***Accounting, Organizations and Society*** 6 (1981): 193-211.

4. Carter, N. and Greer, P. "Evaluating Agencies: Next Steps and Performance Indicators." ***Public Administration*** 71 (1993): 407-416.

5. Tomkins, C. ***Achieving Economy, Efficiency and Effectiveness in The Public Sector***, Kogan Page, London, 1987.

6. Trumpington, Baroness "OR in Government." ***Journal of the Operational Research Society*** 38 (1987): 907-912.

7. Thain, C. and Wright, M. "Planning and Controlling Public Expenditure in The UK, Part II: The Effects and Effectiveness of The Survey." ***Public Administration*** 70 (1992): 193-224.

8. Beer, S. ***Decision and Control***, Wiley, Chichester, 1966.

9. Anthony, R. N. and Young, D. W. ***Management Control in Nonprofit Organizations***, third edition, Richard D. Irwin, Homewood, 1984.

10. Forrester, J. W. ***Principles of Systems: Text and Workbook***, MIT Press, Cambridge, 1968.

11. Emmanuel, C., Otley, D. and Merchant, K. ***Accounting for***

Management Control, Second edition, Chapman and Hall, London, 1990, p. 12.

12. Smith, P. "Outcome-related Performance Indicators and Organizational Control in The Public Sector." ***British Journal of Management*** 4 (1993): 135-151.

13. Williamson, O. A. ***Markets and Hierarchies***, Free Press, New York, 1975.

14. Tomkins, C. ***Corporate Resource Allocation***, Blackwell, Oxford, 1991.

15. Emmanuel, Otley and Merchant, ***Op. Cit.***

16. Nove, A. ***The Soviet Economic System***, Allen and Unwin, London, 1980.

17. Carter, N., Klein, R. E. and Day. P. ***How Organizations Measure Success: The Use of Performance Indicators in Government***, Routledge, London, 1992.

18. Baiman, S. "Agency Research in Managerial Accounting: A Second Look." ***Accounting, Organizations and Society*** 15 (1990): 341-371.

19. Committee of Public Accounts ***Maternity Services, Thirty-fifth Report, Session 1989-90***, HMSO, London, 1990.

20. Smith, ***Op. Cit.***

21. Health Committee ***Maternity Services, Second Report, Session 1991-92***, HMSO, London, 1992.

22. Audit Commission ***Citizen's Charter Indicators: Charting a Course***, HMSO, London, 1992.

23. Hopwood, A. G. *Accounting and Human Behaviour*, Prentice/Hall, Englewood Cliffs, 1974.

24. Kaplan, R. S. and Norton, D. F. "The Balanced Scorecard - Measures That Drive Performance." *Harvard Business Review* 70 (1992): 71-79.

25. Ouchi, W. G. "A Conceptual Framework for The Design of Organizational Control Mechanisms." *Management Science* 25 (1979): 833-848.

26. Etzioni, A. *The Active Society: A Theory of Societal and Political Processes*, The Free Press, New York, 1968.

27. Benefits Agency *Business Plan, 1993/1994*, Department of Social Security, London, 1993.

28. Sheehy, P. *Inquiry Into Police Responsibilities and Rewards, Volume 1: Report*, HMSO, London, 1993.

29. *Ibid.*

30. Kerr, S. "On The Folly of Rewarding A While Hoping for B." *Academy of Management Journal* (1975): 769-783.

31. Department for Education *Choice and Diversity: A New Framework for Schools*, HMSO, London, 1992.

32. Guardian, "The Big Rise in Queues for Hospital Treatment." *The Guardian*, 13 May 1992.

33. Pollitt, C. "Bringing Consumers into Performance Measurement: Concepts, Consequences and Constraints." *Policy and Politics* 16 (1988): 77-87.

34. Flynn, N. "Performance Measurement in Public Sector Services." *Policy and Politics* 14 (1986): 389-404.

35. Watts, R. L. and Zimmerman, J. L. *Positive Accounting Theory*, Prentice/Hall, Englewood Cliffs, 1986.

36. Jones, C. S. "Universities, on Becoming What They Are Not." *Financial Accountability and Management* 2 (1986): 107-119.

37. Smith, P. "The Use of Performance Indicators in The Public Sector." *Journal of the Royal Statistical Society, Series A* 153 (1990): 53-72.

38. Smith, P. "Negative Political Feedback: An Examination of The Problem of Modelling Political Responses in Public Sector Effectiveness Auditing." *Accounting, Auditing and Accountability Journal* 5 (1992): 5-20.

39. United Kingdom Government *The Citizen's Charter*, HMSO, London, 1991.

40. Department of Education and Science "Tests of 7 Year Olds Reveal Unacceptable Variations," *DES Press Release 436/91*, The Department, London, 1991.

41. Kornai, J. *The Socialist System: The Political Economy of Communism*, Clarendon Press, Oxford, 1992.

42. Birman, I "From the Achieved Level," *Soviet Studies* 30 (1978): 153-172.

43. Jaworski, B. J. and Young, S. M. "Dysfunctional Behavior and Management Control: An Empirical Study of Marketing Managers." *Accounting, Organizations and Society* 17 (1992): 17-35.

44. Weitzman, M. "The New Soviet Incentive System." ***Bell Journal of Economics and Management Science*** 7 (1976): 251-257.

45. Ganley, J. A. and Cubbin, S. J. ***Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis***, North-Holland, Amsterdam, 1992.

46. Smith, P. "Assessing Competition Among Local Authorities in England and Wales." ***Financial Accountability and Management*** 4 (1988): 235-251.

47. Hofstede, ***Op. Cit.***

48. Smith, "The Use of Performance Indicators in The Public Sector," ***Op. Cit.***