

ARTICLES

WHAT IS PUBLIC SERVICE IMPROVEMENT?

GEORGE A. BOYNE

Five conceptual models of public service improvement are derived from the large literature on organizational effectiveness. These are the goal, systems-resource, internal process, competing values and multiple constituency models. The strengths and weaknesses of each of these models is evaluated and a working definition of improvement is proposed. This emphasizes that concepts and measures of public service improvement are political rather than technical, and contingent rather than universal. Conclusions are drawn on the implications for academic research and policy development.

In the last two decades there has been a pandemic of public sector reforms (Pollitt and Bouckaert 2000). Governments across the globe have reorganized and restructured public organizations in an effort to produce better services (Boyne *et al.* 2003). The current Labour government in the UK is engaged, as are others, in this quest for public service improvement. Indeed, the Blair government has stated that delivery on its pledge to raise service standards is the single most important criterion for judging the success or failure of its second term of office. This has triggered a vigorous policy debate on the most appropriate path towards higher performance – and especially on the potential contribution of the private sector. Although an analysis of the relative merits of alternative policies and strategies is clearly important, there is a prior question that has, so far, received little attention: what is public service improvement? The aim of this paper is to evaluate the validity of a range of potential answers to this fundamental question.

A new policy community of practitioner organizations has quickly swarmed around the theme of service improvement. This includes elements of the

George A. Boyne is Professor of Public Sector Management, Cardiff Business School, Cardiff University.

Public Administration Vol. 81 No. 2, 2003 (211–227)

© Blackwell Publishing Ltd. 2003, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Cabinet Office (e.g. the Performance and Innovation Unit), the Prime Minister's Office (e.g. the Delivery Unit), the Treasury (e.g. the Public Service Productivity Panel), the Audit Commission (and especially the Best Value Inspectorate), the Commission for Health Improvement, and the Improvement and Development Agency (the 'self-help' arm of the Local Government Association). None of these bodies has produced a comprehensive definition of the concept that is supposedly central to their missions. Indeed, discussions of what public service improvement actually means are usually missing from their documents. For example, the corporate plan of the Commission for Health Improvement (2001) contains a glossary that defines 50 terms that are important to its remit, but public service improvement is not one of them. Similarly the 'Wales Programme for Improvement' (National Assembly for Wales 2001) has much to say about local authority plans and processes, but is silent on improvement itself.

Academic contributions to the definition of improvement are also, as yet, extremely limited. This is at least partly because published literature has tended to lag behind policy debates. A bibliographic search of electronic databases (e.g. Web of Science, BIDS) failed to identify any academic publications on public service improvement. Nor is there any strong tradition of work on this in the main academic disciplines that have contributed to the study of public administration (e.g. economics, politics, social policy). There is, however, a large literature on the concept of organizational effectiveness that is directly relevant to issues concerning the meaning of 'improvement'. This is not to claim that an answer to the question 'what is public service improvement?' is readily available in the effectiveness literature. Indeed, as will be shown below, a central message of prior research in this area is that universal criteria of effectiveness (and, by implication, improvement) do not exist. As Heffron (1989, p. 323) argues, although effectiveness can be defined broadly as 'producing the intended or expected result, there is no agreement on what the intended result of organisations is or should be'. Nevertheless, the effectiveness literature contains a variety of conceptual models that can contribute towards a definition of improvement.

In the first part of this paper the origins and nature of the literature on organizational effectiveness are outlined and differences between the concepts of effectiveness and improvement are identified. Following this, five main schools of thought on organizational effectiveness are summarized, and their strengths and weaknesses as a basis for defining public service improvement are analysed. Finally, a working definition of public service improvement is provided and conclusions are drawn for academic research and policy development.

EFFECTIVENESS AND IMPROVEMENT

The concept of effectiveness has been central to organizational theory for over a century. Indeed, it can be argued that the main purpose of theorists such as Weber, Taylor and Barnard was to develop models of effective

organizations. Rainey (1997, p. 125) notes that 'virtually all of management and organization theory concerns performance and effectiveness, at least implicitly' (see also Hall 1987; Heffron 1989). However, it was not until the late 1950s and early 1960s that sustained attempts were made to explicate and operationalize alternative definitions of organizational success (Steers 1975). This literature flourished in the 1960s and 1970s, especially in sociology and management journals in the USA. Many of the key papers were published in journals such as *Administrative Science Quarterly* and *Academy of Management Review*. By the early 1980s it seemed that most of the conceptual terrain had been mapped and effectiveness research dwindled from a torrent to a trickle. Indeed, the term effectiveness began to be replaced by narrower concepts such as 'quality'. However, the new literature on service quality largely reworked issues that had already been covered by writers on effectiveness (see Cameron and Whetten 1996).

Different interpretations of effectiveness are analysed below. For the present, it is important to note the following three general characteristics of the literature, and their implications for an analysis of public service improvement:

- (i) Conceptual models of effectiveness tend to concentrate on the success or failure of single organizations. By contrast, public service improvement refers to the achievements of sets of organizations (for example, all hospitals, housing associations, local authorities or schools). In other words, the focus is on raising service standards across an 'industry' rather than in an individual agency. This raises questions about whether an increase in the mean performance of service providers is enough to show that improvement has occurred, regardless of the variation around the mean. Can a change be labelled 'improvement' if the standards achieved by some organizations have declined and if the population in some areas receives poorer services? If the mean level of performance rises, but the gap between the best and worst also grows, is this improvement? Some government departments have recently introduced 'floor targets' that are intended to raise service standards fastest in the poorest local communities (Wintour 2002). This poses fundamental questions about the role of equity in judgements about public service improvement, and the potential trade-off between equity and other dimensions of performance (Keeley 1978; Le Grand 1991; Meier *et al.* 1999).
- (ii) Most of the empirical research on effectiveness consists of comparisons of the relative achievements of different organizations in a single time-period. In other words, the focus is on the 'comparative statics' of effectiveness. By contrast, improvement is clearly a *dynamic* phenomenon that concerns temporal shifts in service standards in relation to a baseline. This implies that it may be possible to claim improvement (or deterioration) in public services through the selection of a benchmark year

in which performance was weak (or strong). Alternatively, if the time frame for comparison is fixed, the desired impression may be created by manipulating the set of improvement criteria.

- (iii) The effectiveness literature concentrates on the success of *organizations*, whereas current policy debates are about *services*. This difference is important because public services are often (and increasingly) provided by networks of organizations (Rhodes 1996). Thus organizational effectiveness and service effectiveness are conceptually distinct (Provan and Milward 1995, 2001). Some organizations may become more effective in their contribution to service provision while the service as a whole worsens (because the effectiveness of other agencies declines, or because the activities of different agencies are poorly co-ordinated). Service improvement, in other words, may require more than higher effectiveness by individual organizations, or even by all organizations that are involved in providing a service. Improvement may not occur unless whole service delivery systems (including public, private and voluntary organizations) get better.

In sum, public service improvement is a dynamic phenomenon, inherently raises questions of equity, and is usually concerned with the performance of multi-organizational networks rather than the achievements of single organizations. The five models that are discussed below emphasise different criteria of effectiveness. Nevertheless, each of them can, in principle, accommodate these conceptual and organizational complexities of public service improvement.

MODELS OF IMPROVEMENT

The goal model

This is the oldest and, on the surface at least, the simplest model of organizational effectiveness (Maynard-Moody and McLintock 1987). It is based on the view that formal organizations are established for a purpose – to achieve objectives that individuals, families or communities cannot achieve on their own. Thus, at the time of their establishment, organizations have a major goal or set of goals. Over time, these goals may change or be embellished, but they continue to guide the strategic direction of an organization. Moreover, the extent to which such goals are attained can be used to assess the degree of success or failure (Mohr 1973; Simon 1964).

In principle, this model has direct relevance to the meaning and measurement of public service improvement. A major goal of the national health service, for example, may be to raise the standards of the population's physical wellbeing. Whether improvement has occurred can in turn be assessed through indicators of mortality and morbidity. If readings on these indicators decline, then the health of the population is higher and it may be inferred that public service improvement has occurred. Research on organizational effectiveness in the public sector has usually adopted a version of

the goal model. For example, Rainey and Steinbauer (1999, p. 13) argue that a public agency is effective if it 'achieves the mission as conceived by the organization and its stakeholders, or pursues achievement of it in an evidently successful way'.

In practice, however, the goal model is not so easily applied. First, public service organizations may not have formal goals that are clearly expressed in legislation or other documents. This partly reflects the political constraints on the policy process: different interests can support the same policy for different reasons. Therefore, 'It is not irrational for a policy maker to defend a policy as good without being able to specify what it is good for' (Lindblom 1959, p. 84). Indeed, goal ambiguity is likely to be an advantage when wide political support is necessary in order to get a policy adopted. Secondly, the formal goals that are explicit are likely to be broad 'mission statements' rather than concrete objectives that can be used to evaluate whether improvement has occurred. For example, the goal of the NHS may be described as the promotion of 'better health', but this quickly dissolves under closer examination: whose health? better in what respects? physical or mental health? as judged by whom? Questions such as these quickly turn consensus on service improvement into conflict on resource allocation. Thirdly, public organizations often have a multiplicity of goals. Even on a random basis, some will be met while others are missed, so improvement and deterioration are likely to occur together. Moreover, if only some goals are explicit and quantified, then activities and efforts are likely to be biased towards these at the expense of others (Kerr 1976).

Another problem with the goal model is that formal organizational objectives may differ significantly from the operative objectives that are actually pursued (Perrow 1961). The replacement of official aims (e.g. 'public welfare') by unofficial aims (e.g. pursuit of bureaucratic self-interest) has been extensively criticized by public choice theorists (Boyne 1998; Self 1993), and by other opponents of services that are paternalist and unduly dominated by professionals (Cutler and Waine 1997). Nevertheless, some strands of the effectiveness literature regard the promotion of the interests of organizational members as synonymous with success. For example, Cummings (1977, p. 60) argues that 'an effective organization is one in which the greatest percentage of participants perceive themselves as free to use the organization and its subsystems as instruments for their own ends'. It seems unlikely that many policy makers or academics would accept this as a basis for defining public service improvement. Even if operative aims reflect the public interest rather than professional self-interest, it may be difficult to elicit and identify their exact content. The operative goals may exist only in the minds and oral discourses of service providers, rather than in strategic plans or policy statements (Newman 2001). This is hardly conducive to establishing what counts as improvement or whether it has occurred.

If public service objectives could be clearly stated, universal agreement on them existed, and official and operative objectives were identical, the goal

model would still confront a basic methodological problem of attribution (Pollitt 1995). For example, the population may become more healthy over time, but is this the product of better public services? Service improvement can only be validly inferred from a positive shift in outcome data once all other relevant influences on health have been taken into account. This problem can be alleviated if the focus of attention changes from outcomes to variables that are under more direct organization control (e.g. number of operations performed, cost per operation), but this effectively narrows the emphasis from *service* improvement to *organizational* improvement.

Despite these problems, the goal model has a powerful resonance with current debates about improvement. Public services are clearly expected to achieve *something* and to produce tangible results that satisfy key stakeholders. The practical issues, which are not addressed by the goal model, are how to identify these groups and their preferred outcomes and relative priorities.

The systems-resource model

This model is based on two definitions of effectiveness which emphasise that organizations are open systems that must acquire resources from their environment. (Cunningham 1978; Molnar and Rogers 1976). The first, developed primarily for private sector organizations, is *survival*. In other words, organizations that continue to exist are by definition successful, whereas those that wither are unsuccessful. This reflects the assumption that market pressures 'weed out' inefficient and ineffective firms, and that only the fit survive. Even for private firms, this is a questionable view of the determinants of longevity (Meyer and Zucker 1989). The second systems-resource definition focuses on the *preconditions* of survival. Organizations require inputs from the environment in order to survive, including goods and services from suppliers and money from customers. According to Yuchtman and Seashore (1967, p. 898) an organization is effective if it has the ability 'to exploit its environment in the acquisition of scarce and valuable resources'. Furthermore, organizations that obtain the biggest share of available resources, and thereby grow (and probably survive the longest), are deemed the most effective.

The relevance of this model to the effectiveness of private organizations has been criticized because it focuses on inputs rather than outputs or outcomes (Hall 1987; Strasser *et al.* 1981). Further problems arise if the model is applied to the public sector. In particular, the survival or growth of public organizations is no guide to whether services are improving. In the first place, organizations may be abolished (or retained) for political reasons that have little relevance to service standards. For example, it is widely acknowledged that the Greater London Council and metropolitan counties were abolished in the 1980s because they presented a political challenge to the then Conservative government, rather than because their record of service provision was poor (O'Leary 1987). The ability of a public organization to acquire resources is also likely to reflect its political support as much as its

service accomplishments. For example, funding is currently available to promote various forms of 'partnership working'. The claim that organizations which embrace this policy (and thereby receive extra resources) will deliver better services is, as yet, unproven.

Nevertheless, the systems-resource model does highlight an important policy issue. Does an injection of extra resources lead to service improvement? The current Labour government has recently stated that it intends to raise health spending in the UK to the European average in the next five years. It is striking that the objective is to boost *expenditure* on health rather than health itself. The assumption seems to be that the former will lead to the latter – but is this hypothetical connection valid? Most previous research has found little link between expenditure inputs and service standards (Dean and Peroff 1977; Hinkley and Marquette 1983; Powell 1997; Sharkansky 1967). However, such research has usually sought to draw longitudinal conclusions from cross-sectional data. The relationship between spending and performance in a single time period may be largely neutral (Figure 1). This pattern suggests that higher levels of expenditure are not associated with better services – organizations that spend more (on the right hand side of the figure) are just as likely to have below average as above average performance. However, a different picture may emerge if the longitudinal relationship is analysed. Figure 2 shows the pattern before (time t) and after (time $t + 1$) a boost in public spending on a service. In each time period the cross-sectional relationship between spending and services appears to be weak. Yet over time the performance of all organizations has moved up. The highest spenders, within a given time period, are still not performing better than the lowest, but outcomes have risen across the whole set of service providers. Thus conclusions on the spending-service link need to be based on longitudinal data. If a positive link is established empirically, then

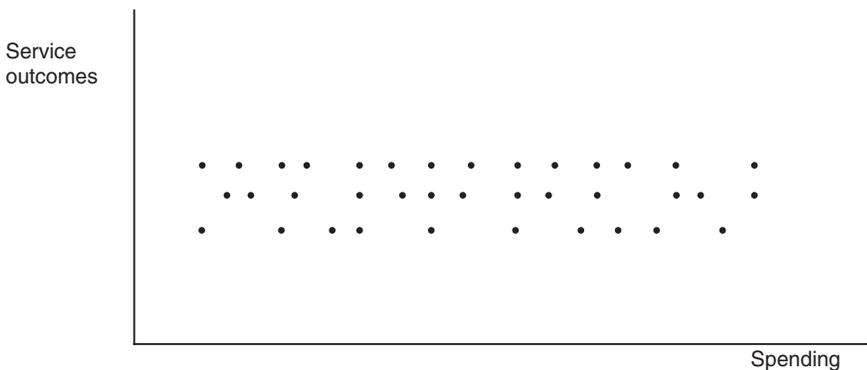


FIGURE 1 *Cross-sectional relationship between spending and service outcomes*

Note: each • plots the position of a public service provider against its spending (on the horizontal axis) and the outcomes that are achieved (on the vertical axis).

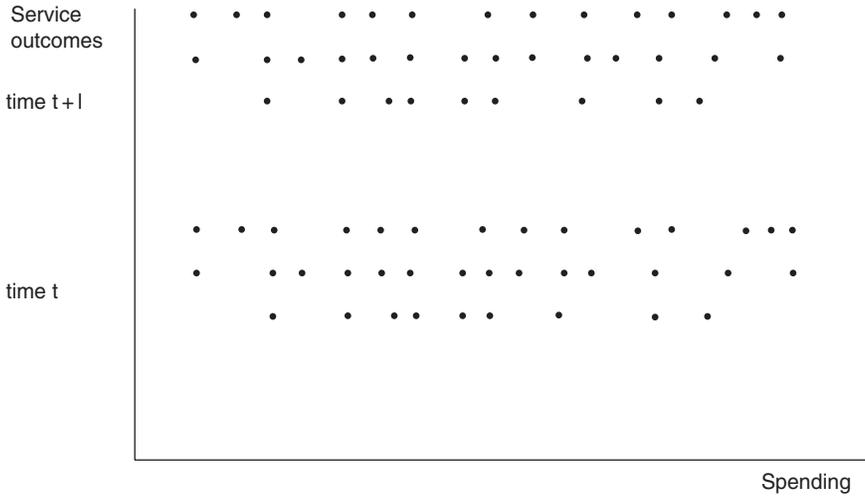


FIGURE 2 *Longitudinal and cross-sectional relationships between spending and service outcomes*

Note: time t is the period before a general boost in spending, and time t+1 is the period after.

the systems-resource model may yet have some relevance to public service improvement.

Internal processes model

If goals are nebulous and contested, and if survival or growth is weakly connected with success, then an alternative approach is to use organizational attributes as proxies for effectiveness (Price 1968; Spray 1976). This is the gist of the internal process model which focuses either on technical aspects of organizations (e.g. their information and budgeting systems) or human aspects (e.g. personnel practices that promote job satisfaction and employee welfare).

This model has a superficial relevance to public services, because outcomes are often hard to define and measure (and, as noted above, difficult to attribute exclusively to the work of service providers). Why not, then, use the achievement of higher professional or managerial standards as indicators of service improvement? There is evidence that such criteria are used by funding bodies to evaluate the performance of public agencies, and by senior managers in service delivery organizations. For example, Herman and Renz's study of health and welfare organizations in the USA found that funders and managers 'do not rely on bottomline outcomes as meaningful indicators of organisational effectiveness, preferring instead evidence of following correct procedures or doing things right' (1997, p. 200). Similarly, the Best value inspectorate in the UK is currently using 'corporate governance'

criteria to assess whether services in English and Welsh local government are likely to improve (Audit Commission 2001).

The presumption in these cases is that particular sets of internal processes are strongly related to organizational performance. The problem is, however, that evidence to substantiate this view is not available. Thus, as with the systems resource model, *potential* antecedents of service improvement are being confused with improvement itself. Yet this conclusion may need to be qualified, at least as far as the 'human resources' version of the internal process model is concerned. A motivated and contented workforce may be regarded as crucial to public service improvement for two reasons. First, good human resource management may be a precondition of organizational success. Quite simply, improvement seems unlikely if public service staff are poorly trained and inadequately rewarded. Second, even if staff development and morale are not related to service standards, the welfare of the workforce may be regarded as an end in itself. After all, if the ultimate purpose of service improvement is to enhance human happiness, then job satisfaction in the public sector contributes to this. The trick, of course, is to find an appropriate balance between the interests of various internal and external stakeholders. This issue is explored further in the context of the multiple constituency model that is discussed later in the paper.

The competing values model

This fourth perspective on organizational effectiveness is an attempt to synthesize elements of the three models that have already been discussed. The competing values model assumes that all organizations face contradictory criteria of performance (Quinn and Rohrbaugh 1981, 1983). Furthermore, these criteria can be mapped onto two dimensions that together yield four interpretations of effectiveness. The first dimension concerns the attitudes of key stakeholders towards control and flexibility: should organizational processes and performance targets be prescribed and monitored, or should managers be given discretion to develop strategies as they go along? The second dimension concerns the location of power over organizational activities: does this rest primarily inside or outside the organization?

These two dimensions are combined in the competing values framework to provide four 'ideal type' models of organizational effectiveness (see figure 3, below). The ideal types overlap closely with: (a) the goal model (focus on control, with goals in the competing values framework interpreted as externally imposed rather than internally selected); (b) the systems-resource model (emphasis on flexibility of response in order to obtain inputs from the external environment); and (c) two versions of the internal process model: technical processes (focus on control, combined with an assumption that information and budgeting systems are internally set), and human relations (internally selected to emphasize the development and welfare of staff).

The competing values model has been applied extensively in empirical research on organizational effectiveness (e.g. Kalliath *et al.* 1999; Rohrbaugh

1981). Many organizations have been mapped onto the ‘control’ and ‘power’ dimensions in order to evaluate their performance. The model’s strongest contribution to an understanding of public service improvement is that the *contradictions* between different interpretations of effectiveness become explicit (e.g. the tension between lower cost and higher quality). This echoes the ‘balanced scorecard’ for measuring effectiveness which has been extensively used in the private sector (Kaplan and Norton 1996; Butler *et al.* 1997) and has more recently been imported into the public sector (Aidemork 2001; Talbot 1999). In particular, the competing values framework highlights that it is impossible for organizations to emphasize simultaneously control and flexibility, or to be responsive primarily to both internal and external stakeholders. Rather, movement into one box in figure 3 must be traded against movement out of another. Despite this, the framework does not really go beyond the models that it synthesizes. None of the problems with the separate models are resolved by combining them in this way.

The multiple constituency model

This final model abandons the search for universal criteria of organizational effectiveness. The simplicity and order of the competing values framework is replaced by the complexity and mess of practical politics. The multiple constituency model assumes that the performance of all organizations is judged by a variety of internal and external groups (Cameron 1986; Connolly *et al.* 1980). Each of these groups uses different criteria to assess effectiveness. Even if all stakeholders use similar criteria, the relative weight that they attach to each one is likely to vary. Furthermore, the criteria and the weights shift over time as preferences change and as the balance of

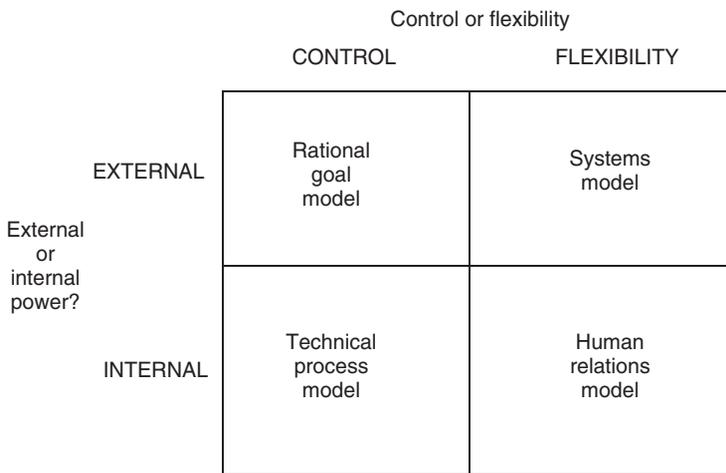


FIGURE 3 *The competing values model*

power alters between groups. An organization is effective, then, to the extent that its multiple constituencies perceive it as effective (Zammuto 1984). To some extent, this view is consistent with the systems resource model. However, success is determined not by the ability to obtain technical inputs but by the ability to mobilize and maintain political support. Gaertner and Yamaran (1983, p. 97) argue that 'an effective organization is one that is able to fashion accounts of itself and its activities in ways which constituencies find acceptable'. Thus, the multiple constituency model emphasizes organizational image and legitimacy rather than substance and service delivery.

The implications for a definition of public service improvement are profound. The idea that there is a single set of performance goals that can be universally agreed must be abandoned. Instead, improvement is exposed as an arena for political struggle between different groups that are attempting to impose their preferred criteria of success or failure on public services.

This model of improvement poses conceptual and normative problems. First, it appears to be all process and no content: the multiple constituency models tell us *how* improvement is defined (by the outcome of a political contest between various stakeholders) but not *what* it is (except whatever the dominant groups say it is). Second, both the how and the what may be questionable from an ethical viewpoint: is the existing balance of power between groups in society appropriate? Are all constituencies equally able to articulate their preferred performance criteria? If not, who is to rebalance the interests of different groups? The best available answer to this last question in a democratic society may be 'the elected government', which implies that public service improvement is whatever political leaders says it is. What, then, if the government is unwilling or unable to pronounce on the meaning of improvement?

In this case the practical application of the multiple constituency model takes the definition of public service improvement squarely into the realms of political science and market research. An operationalization of the model would require the identification of the various constituencies that have an interest in service provision and an assessment of their relative power. Furthermore, it would be necessary to elicit their definitions of improvement and the intensity of their preferences. Even if all of these empirical research tasks could be accomplished, the multiple constituency model suggests that the views of different stakeholders may be irreconcilable. Thus it would be impossible to take the final step of identifying the combination of service standards that would best satisfy net preferences across all constituencies. In other words, the content of public service improvement would remain indefinable.

TOWARDS A DEFINITION OF PUBLIC SERVICE IMPROVEMENT

The five models of organizational effectiveness provide different sets of criteria for defining public service improvement. The goal model emphasizes changes in the substantive standards of services; the system-resource model

focuses on the capacity of service providers to acquire technical inputs from their environment; the internal processes model suggests that most attention should be paid to how such inputs are used within organizations; the competing values model attempts to combine all of these perspectives into a universally applicable set of evaluative criteria; and the multiple constituency model implies that improvement exists only in the eyes of organizational stakeholders.

The preceding analysis of these models suggests that three are seriously deficient as a basis for defining public service improvement. Both the systems-resource model and the internal process model are largely concerned with potential sources of improvement rather than improvement itself. It is possible that more resources and better internal processes are correlated with improvement. Nevertheless, these variables are, at best, rough proxies for higher service standards rather than direct measures of this concept. Moreover, the components of these models are themselves vague. Which resources are most critical to improvement? What counts as 'good' internal processes? Which budgetary and information systems or styles of human resource management are most likely to lead to better services? A substantial research effort is required to answer these questions, and thereby to test the usefulness of the systems-resource and the internal process models. The competing values approach is also weak because it draws substantially on these two models. Moreover, its implicit claim that a universal set of dimensions can be used to map service improvement seems unduly grand and apolitical.

By contrast, a combination of the remaining two models is a positive basis for defining and measuring public service improvement. The strength of the goal model is that it emphasizes the *content* of improvement. Service standards have tangible elements (e.g. speed, quality, reliability) that are likely to be valued by all stakeholders, even if they disagree about the weights that should be attached to them. This 'realist' view of social phenomena (Sayer 1984; Pawson 1989) compensates for the weakness of the multiple constituency model which assumes that improvement consists only of the subjective impressions of powerful stakeholders. In effect, the goal model pulls the multiple constituency model back from the brink of the postmodernist morass of extreme relativism. The strength of the multiple constituency model is the recognition that concepts of improvement are political rather than technical. The model illuminates the *process* by which improvement is defined, rather than the content of the definition. This effectively compensates for the weakness of the goal model which neglects the origins of public service performance criteria.

This is not to argue that the two models can be smoothly and easily integrated in their conventional form. The goal model needs to be modified to reflect the existence of multiple perspectives on the content of public service improvement. In short, the traditional view that a small and consensual set of performance criteria can be used to judge the success of public organizations

must be sacrificed. Similarly, an amended version of the multiple constituency model needs to accept that some agreement is possible on whether specific dimensions of public services have improved or deteriorated, even when different stakeholders disagree about the weight to be attached to the dimensions. If a 'pure' version of either model is retained, then they cannot be blended successfully to produce a meaningful interpretation of public service improvement.

Taken together, the goal and multiple constituency models imply that a working definition of improvement must incorporate both the substance of organizational achievements, and the inherently political nature of judgements on success or failure. Thus public service improvement can be defined formally but provisionally as *a closer correspondence between perceptions of actual and desired standards of public services*. This is a working definition in the sense that it provides a first rather than a last word on this issue, and sets a direction for further inquiry. The main research questions posed by this definition include the following:

- what criteria are used by different stakeholders to judge the standards of public services, and how are the criteria selected?
- how much consensus is there on these criteria, and on their relative importance? For example, how do preferences vary by income, gender, age, ethnicity and geographical areas?
- how are actual and desired levels of performance on these criteria identified, and by whom?
- to what extent is improvement a result of a rise in actual, or a fall in desired, service standards? In other words, can perceptions of improvement be managed by lowering the expectations of stakeholders as well as raising the performance of service providers?
- how closely are perceptions of improvement related to movements in 'objective' indicators? For example, are changes in Best Value performance indicators for English local authorities positively correlated with movements in public satisfaction with local services?

An additional complication in the analysis of improvement is that judgements on success or failure are likely to vary with the circumstances of service provision. In other words, operational definitions of service improvement will be contingent rather than universal. Contingencies such as the age of organizations and the characteristics of their environments have been identified in the effectiveness literature (Quinn and Cameron 1983). Judgements on public service improvement also need to take a variety of circumstances into account. For example, organizations may find it easier to achieve a given improvement in performance if they are starting from a low base. Moreover, improvement on one criterion of performance may be contingent on deterioration on another. For example, innovation, which may be necessary for long-term success, may require a substantial level of slack resources. The latter, however, may show up as inefficiency in the short term. An

environmental contingency is that higher service standards may be attained more easily by organizations that face favourable socio-economic circumstances. Evaluations of the extent of improvement need to take such contingencies into account, especially if rewards and sanctions are allocated selectively between 'high performing' and 'failing' organizations, which is the current policy of central government in England for health authorities and local councils.

CONCLUSIONS

Public service improvement is an idea whose time has come. Indeed, it now seems remarkable that this topic has taken so long to emerge at the top of policy and academic agendas. What issue could be more central to the work of policy makers, and to research in public policy and management? Suddenly, it seems, central government is looking to the academic community in the UK and elsewhere to shed light on the determinants of improvement. However, there is little evidence that can illuminate the government's quest for better public services (Boyne 2003; Rainey and Steinbauer 1999). Academic researchers remain largely at the stage of clarifying the questions rather than providing empirical answers.

One of these questions, and perhaps the most fundamental, is 'what is public service improvement'? In this paper, five conceptual models that are relevant to this issue have been explored, and their strengths and weaknesses evaluated. Two perspectives on organizational effectiveness, the goal and multiple constituency models, can be combined to provide useful insights on the meaning of public service improvement. Taken together, they imply that improvement can be assessed on the basis of a variety of substantive criteria, but that these criteria (and the relative weights attached to them) vary across stakeholders and over time. Moreover, judgements on the extent of success or failure vary with the internal and external contexts of organizations that provide public services. Thus definitions and measures of improvement are not technical and universal but politically constructed and contingent on a variety of circumstances.

An implication for academic researchers is that the dimensions of improvement need to be more clearly and comprehensively theorized, in ways that go beyond notions of 'economy-efficiency-effectiveness' and 'inputs-outputs-outcomes' (Boyne 2002). Furthermore, the practical relevance of these conceptual dimensions needs to be checked against the views of the multiple constituencies who judge the standards of public services. How are the dimensions of improvement related to each other in practice, and how (and why) do these relationships vary across stakeholder groups? Answers to such questions may be a precondition of building realistic models of the antecedents of improvement.

A synthesis of the goal and multiple constituency models also contains important messages for policy makers. In particular, improvement is not a single phenomenon with a uniform interpretation. Accordingly, an attempt

to pursue a narrow set of goals (even if these are accompanied by a wide collection of performance indicators) may result in a huge welfare loss for many stakeholder groups. An implication is that a programme of improvement needs to be matched to the plurality and diversity of public preferences, which raises the thorny old issue of the appropriate balance between national frameworks and local discretion. This remains unresolved but needs to be reconsidered afresh in the contemporary context of promoting better public services.

Finally, public service improvement may slip down the policy agenda, as it is either partly accomplished or overtaken by other domestic and foreign policy issues. However, academics should make a sustained attempt to analyse the nature and determinants of improvement, not only because these are inherently interesting intellectual puzzles, but also to be better equipped to contribute to policy debates the next time round. For the present, it is worth emphasizing to national policy makers that you can (temporarily) take improvement out of politics, but you can't take politics out of improvement.

REFERENCES

- Aidemark, L. 2001. 'The meaning of balanced scorecards In health care organizations', *Financial Accountability and Management*, 17, 23–40.
- Audit Commission. 2001. *Changing gear*. London: Audit Commission.
- Boyne, G.A. 1998. *Public choice theory and local government*. London: Macmillan.
- Boyne, G.A. 2002. 'Concepts and indicators of local authority performance: a review of the statutory frameworks in England and Wales', *Public Money and Management*, 22, 17–24.
- Boyne, G.A. 2003. 'Sources of public service improvement: a critical review and research agenda', *Journal of Public Administration Research and Theory*, forthcoming.
- Boyne, G.A., C. Farrell, J. Law et al. 2003. *Evaluating public management reforms: principles and practice*. Buckingham: Open University Press.
- Butler, A., S. Letza and B. Neale. 1997. 'Linking the balanced scorecard to strategy', *Long Range Planning*, 30, 242–53.
- Cameron, K. 1986. 'Effectiveness as paradox: consensus and conflict in perceptions of organizational effectiveness', *Management Science*, 32, 539–53.
- Cameron, K. and D. Whetten. 1996. 'Organizational effectiveness and quality: the second generation', in J. Smart (ed.), *Higher education handbook of theory and research*. New York: Agathon Press.
- Connolly, T.E. Conlon and S. Deutsch. 1980. 'Organizational effectiveness: a multiple constituency approach', *Academy of Management Review*, 5, 211–17.
- Cummings, L. 1977. 'Emergence of the instrumental organization', in P. Goodman and J. Pennings (eds), *New perspectives on organizational effectiveness*. San Francisco: Jossey-Bass.
- Cunningham, J. 1978. 'A system-resource approach for evaluating organizational effectiveness', *Human Relations*, 31, 631–56.
- Cutler, B. and B. Waine. 1997. *Managing the welfare state*. Oxford: Berg.
- Commission for Health Improvement. 2001. *Corporate plan, 2001–2004*. London: Commission for Health Improvement.
- Dean, G. and K. Peroff. 1977. 'The spending-service cliché: an empirical evaluation', *American Politics Quarterly*, 5, 501–16.
- Gaertner, G. and S. Ramayaran. 1983. 'Organizational effectiveness: an alternative perspective', *Academy of Management Review*, 8, 97–107.

- Hall, R. 1987. *Organizations: structures, processes and outcomes*. Englewood Cliffs, NJ: Prentice Hall.
- Heffron, F. 1989. *Organization theory and public organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Herman, D. and D. Renz. 1997. 'Multiple constituencies and the social construction of nonprofit organisation effectiveness', *Nonprofit and Voluntary Sector Quarterly*, 26, 185–206.
- Hinkley, K. and J. Marquette. 1983. 'Views through a kaleidoscope: the dimensions of state welfare policy measures', *Western Political Quarterly*, 36, 466–78.
- Kalliath, T., A. Bluedorn and D. Gillespie. 1999. 'A confirmatory factor analysis of the competing values instrument', *Educational and Psychological Measurement*, 59, 143–58.
- Kaplan, R. and D. Norton. 1996. 'Linking the balanced scorecard to strategy', *California Management Review*, 39, 1, 53–74.
- Keeley, M. 1978. 'A social justice approach to organizational evaluation', *Administrative Science Quarterly*, 23, 272–92.
- Kerr, S. 1975. 'On the folly of rewarding A while hoping for B', *Academy of Management Journal*, 18, 769–83.
- Lindblom, C. 1959. 'The science of muddling through', *Public Administration Review*, 39, 517–26.
- Le Grand, J. 1991. *Equity and choice: an essay in economics and applied philosophy*. London: Harper Collins.
- Maynard-Moody, S. and C. McClintock. 1987. 'Weeding an old garden: toward a new understanding of organisational goals', *Administration and Society*, 19, 125–42.
- Meier, K., R. Wrinkle and J. Polinard. 1999. 'Equity versus excellence in organizations', *American Review of Public Administration*, 29, 5–18.
- Meyer M. and L. Zucker. 1989. *Permanently failing organizations*. London: Sage.
- Mohr, L. 1973. 'On the concept of organizational goal', *American Political Science Review*, 67, 470–81.
- Molnar, J. and D. Rogers. 1976. 'Organizational effectiveness: an empirical comparison of the goal and system resource approaches', *Sociological Quarterly*, 17, 401–13.
- National Assembly For Wales. 2001. *Wales programme for improvement: guidance for local authorities*. Cardiff: National Assembly For Wales.
- Newman, J. 2001. 'What counts is what works? Constructing evaluations of market mechanisms', *Public Administration*, 79, 89–103.
- O'Leary, B. 1987. 'Why was the GLC abolished?', *International Journal of Urban and Regional Research*, 11, 193–217.
- Pawson, R. 1989. *A measure for measures*. London: Routledge.
- Perrow, C. 1961. 'The analysis of goals in complex organisations', *American Sociological Review*, 26, 854–866.
- Pollitt, C. 1995. 'Justification by works or faith', *Evaluation*, 1, 13–35.
- Pollitt, C. and C. Bouckaert. 2000. *Public management reform*. Oxford: Oxford University Press.
- Powell, M. 1997. *Evaluating the National Health Service*. Buckingham: Open University Press.
- Price, J. 1968. *Organizational effectiveness*. Homewood, IL: Richard D. Irwin.
- Provan, K. and B. Milward. 1995. 'A preliminary theory of interorganizational network effectiveness: a comparative study of four community mental health systems', *Administrative Science Quarterly*, 40, 1–33.
- Provan, K. and B. Milward. 2001. 'Do networks really work? a framework for evaluating public sector organisational networks', *Public Administration Review*, 61, 414–23.
- Quinn, R. and K. Cameron. 1983. 'Organizational life cycles and shifting criteria of effectiveness: some preliminary evidence', *Management Science*, 29, 33–51.
- Quinn, R. and J. Rohrbaugh. 1981. 'A competing values approach to organizational effectiveness', *Public Productivity Review*, 5, 122–40.
- Quinn, R. and J. Rohrbaugh. 1983. 'A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis', *Management Science*, 29, 363–77.
- Rainey, H. 1997. *Understanding and managing public organizations*. San Francisco: Jossey-Bass.
- Rainey, H. and P. Steinbauer. 1999. 'Galloping elephants: developing elements of a theory of effective government organizations', *Journal of Public Administration Research and Theory*, 9, 1–32.
- Rhodes, R. 1996. 'The new governance: governing without government', *Political Studies*, 44, 652–67.
- Rohrbaugh, J. 1981. 'Operationalizing the competing values approach: measuring performance in the employment service', *Public Productivity Review*, 5, 141–59.
- Sayer, A. 1984. *Method in social science: a realist approach*. London: Hutchinson.
- Self, P. 1993. *Government by the market?* London: Macmillan.

- Sharkansky, I. 1967. 'Government expenditures and public services in the American states', *American Political Science Review*, 61, 1066–77.
- Simon, H. 1964. 'On the concept of organizational goal', *Administrative Science Quarterly*, 9, 1–22.
- Spray, L. 1976. *Organizational effectiveness*. Kent State, OH: Kent State University Press.
- Steers, R. 1975. 'Problems In the measurement of organizational effectiveness', *Administrative Science Quarterly*, 20, 546–58.
- Strasser, S.J. Eveland, G. Cummins, *et al.* 1981. 'Conceptualizing the goal and system models of organizational effectiveness: implications for comparative evaluation research', *Journal of Management Studies*, 18, 321–40.
- Talbot, C. 1999. 'Public performance – towards a new model?', *Public Policy and Administration*, 14, 3, 15–34.
- Tsui, A. 1990. 'A multiple-constituency model of effectiveness: an empirical examination at the human resource subunit level', *Administration Science Quarterly*, 35, 458–83.
- Walsh, K. 1991. 'Quality and public services', *Public Administration*, 69, 503–14.
- Wintour, P. 2002. 'Prescott outlines targets to end "postcode poverty"', *The Guardian* 16 January, p. 10.
- Yuchtman E. and S. Seashore. 1967. 'A system resource approach to organizational effectiveness', *American Sociological Review*, 32, 891–903.
- Zammuto, R. 1984. 'A comparison of multiple constituency models of organizational effectiveness', *Academy of Management Review*, 9, 606–616.

Date received 28 March 2002. Date accepted 2 July 2002.