Sources of Public Service Improvement: A Critical Review and Research Agenda

George A. Boyne
Cardiff Business School

ABSTRACT

Evidence from sixty-five empirical studies of the determinants of public service performance is critically reviewed. The statistical results are grouped on the basis of five theoretical perspectives: resources, regulation, markets, organization, and management. The analysis suggests that the most likely sources of service improvement are extra resources and better management. A research agenda for further work is identified, and recommendations are made to enhance the theoretical and methodological quality of studies of public service improvement.

Governments across the globe are searching for ways to improve public services. During the last twenty years there has been a pandemic of public-sector reforms, many of which are associated with the new public management (Pollitt and Bouckaert 2000). Which, if any, of the existing approaches to public service improvement actually work? How much empirical evidence is there on the impact of the various reform strategies, and how valid is the evidence? What are the main issues that need to be resolved in future work on the sources of public service improvement? The intent of this article is to answer these questions.

In the first section of the article, the meaning of public service improvement is analyzed and criteria for evaluating the effects of different routes toward this "Holy Grail" of public administration theory and practice are identified. In the second section, five major theoretical perspectives on the sources of improvement are outlined. These cover resources, regulation, markets, organization, and management. Taken together, these five perspectives offer the basis for a more comprehensive theoretical model of how to provide better public services. In the third section of the article, the methods and results of empirical studies that have tested the five theories are critically reviewed. This leads to conclusions about the extent of our knowledge of how to achieve service improvement and about an agenda for the substance and methods of the vast research program that is required on this topic.

An earlier version of this article was presented at a conference on Networks, Management, and New Patterns of Governance in Barcelona in October 2002. I am grateful to Carolyn Hill, Ken Meier, and Larry O'Toole for their comments on the article and to the *Journal of Public Administration Research and Theory*'s three anonymous reviewers for their helpful suggestions.

DOI: 10.1093/jopart/mug027

Journal of Public Administration Research and Theory, Vol. 13, no. 3, pp. 367–394

WHAT IS PUBLIC SERVICE IMPROVEMENT?

The concept of service improvement is inherently political and contestable (Boyne 2003). The performance of public service providers is judged by multiple constituencies (e.g., consumers, taxpayers, staff, and politicians). Furthermore, each of these constituencies may use different criteria to judge the standard of public services and may apply different weights to the same criterion (Boschken 1994; Heffron 1989; Rainey 1997). It follows that there is no fixed and universally applicable set of criteria for evaluating whether improvement has occurred. Nevertheless, public services have tangible elements (e.g., quantity, speed of delivery, effectiveness) that are likely to be valued by all constituencies, even if the valuations differ between groups or over time. This realist view of social phenomena (Pawson 1989; Sayer 1984) assumes that improvement does not consist simply of the subjective and irreconcilable impressions of different stakeholders.

Preliminary criteria of service improvement can be derived from the larger body of literature on the conceptualization and measurement of organizational performance in the public sector (Ammons 2001; Carter, Day, and Klein 1992). A review of this literature (Boyne 2002) has identified the following "headline" dimensions of service performance:

- quantity of outputs (e.g., number of operations performed in hospitals, hours of teaching delivered in schools, number of houses built)
- quality of outputs (e.g., speed and reliability of service, courtesy of staff)
- efficiency (ratio of outputs to financial inputs)
- equity (fairness of the distribution of service costs and benefits between different groups)
- outcomes (e.g., percentage of pupils passing exams, percent of hospital patients treated successfully)
- value for money (cost per unit of outcome)
- consumer satisfaction (which may be a proxy for some or all of the above, depending on the questions posed to service users)

If all other variables are equal, an upward shift in any of these aspects of service performance can be taken as evidence of improvement. This set of criteria will therefore be used to help analyze the results of the empirical studies that are reviewed in the third part of the article.

THEORETICAL PERSPECTIVES ON SERVICE IMPROVEMENT

There is no comprehensive and established theory of service improvement. The five perspectives outlined here have been derived in two ways: first, on an inductive basis from the hypotheses that have been tested in empirical studies of service performance and second, from recent work that has sought to develop models of organizational performance in the public sector (Boyne and Dahya 2002; Heinrich and Lynn 2000; Meier and O'Toole 1999; Rainey and Steinbauer 1999). It is important to note that the five perspectives discussed in the following section are theories in the loose sense that they identify sets of variables that

are believed to influence performance. Rigorous causal reasoning and integrated sets of precise propositions do not characterize the literature on organizational success in the public sector.

Resources

The idea that more resources will lead to better results is perhaps the simplest theory of public service improvement. A strong version of this theory suggests that higher public expenditure is a sufficient condition for improvement because this must result in a higher quantity and/or quality of public services. A weak version suggests that more spending is a necessary but not sufficient condition: the resources must be effectively managed in order to deliver the maximum potential benefits. In either case, the proposition is that the relationship between resources and service improvement is positive.

The plausibility of this view was, however, undermined by developments in practice and theory in the 1970s and 1980s. First, the "fiscal crisis of the state" (Gough 1979) led to criticisms of the post–World War II growth in public spending, which seemed to be producing an ever-increasing tax burden with no tangible benefits. Second, this apparent phenomenon was explained by public choice theorists as a product of self-interested behavior by bureaucrats that led to a bloated and inefficient public sector (Tullock 1965; Downs 1967; Niskanen 1971). Thus the flow of extra resources evaporated in the bureaucracy. The proposition that follows is that the relationship between higher spending and service improvement is, at best, insignificant. Indeed, service efficiency can be expected to deteriorate as the size of the public budget expands.

Regulation

It has been widely argued that a distinctive feature of management in the public sector is the extent of regulation by external bodies (see Boyne 2002). Public service providers are not free to choose their own processes and strategies but instead must work within policy constraints set by higher political authorities (Hood et al. 1998). The most basic form of regulation is simply laws that impose obligations or prohibit specific activities. Beyond this, political bodies wield a range of regulatory instruments that include audit, inspection, financial controls, performance indicators, plans, and annual reports (Ashworth, Boyne, and Walker 2002).

Such regulation is usually justified on grounds of accountability: public organizations are funded partly by tax revenues and have a range of responsibilities (e.g., for probity and due process) that go beyond service provision. It seems likely, then, that there is a positive effect of regulation on accountability. However, the relationship between regulation and service performance is less obvious. In particular, this relationship may be mediated by the expertise of the regulators (Boyne, Day, and Walker 2002). If regulators know better than local agencies how to improve services, then the impact of regulation is likely to be positive. In contrast, if local expertise is superior, then regulation is likely to be counterproductive. It is therefore difficult to specify an a priori proposition on the relationship between regulation and public service improvement.

Even if the general theoretical impact of regulation were positive, it is possible that a proliferation of regulators would be harmful to service standards because the delivery agents would be subject to conflicting pressures from multiple principals (Barrow 1996; Douma

and Schreuder 1992). Different regulatory bodies might impose different expectations and restrictions on local service agencies and thereby confuse and demotivate them. Similar effects might follow from the imposition of more and more methods of regulation. In other words, there may be diminishing returns to the number of regulators and the number of regulatory instruments deployed by each of them.

Market Structure

The need to inject greater competition into service markets was the dominant theme of public management reform during the 1980s and 1990s (Boyne 1998; Hilke 1993). The basic argument is longstanding and fairly simple: just as competition is believed to promote efficiency, innovation, and consumer responsiveness in the private sector, so too will these benefits follow from market forces in the public sector (McKean 1965). Competition implies rivalry between alternative suppliers for a share of a market. If competitive behavior rather than ownership of organizations is the key variable, then better services should ensue regardless of whether the rivalry is between public organizations or between public and private organizations.

Whereas traditional economic theory (and its public choice offspring) suggests a positive relationship between competition and service performance, this view has been widely challenged. The major theoretical critique is based on Williamson's (1975) work on the relative merits of markets and hierarchies as modes of economic organization. In effect, Williamson provides a contingency theory of competition. Transactions through markets are likely to work when the number of suppliers is high and information costs and asset specificity are low. When the reverse conditions apply, service provision through a hierarchy (i.e., a traditional public-sector monopolist) is likely to be a better alternative (Blank 2000). Competition has also been widely criticized for its effects on equity (Le Grand and Bartlett 1993; Udehn 1996). Even if more efficiency and responsiveness are achieved, there may be adverse effects on disadvantaged groups in society (e.g., poor, ill, or elderly people—the very groups who are likely to be most in need of public services in the first place but who are also difficult and expensive to treat).

In sum, existing theory suggests that competition has some beneficial effects in the public sector but that these effects will vary with transaction costs and across dimensions of service performance. It seems unlikely, therefore, that the empirical impact of competition on service improvement will be uniformly positive.

Organization

The concepts of reorganization and reform are often used interchangeably (Caiden 1991). This may be because changes in the organization of public services are a favorite strategy of governments, especially after a turnover of political elites (Pollitt 1984). New organizational structures can be regarded as the traditional method of redesigning public service provision and may have a symbolic importance beyond any tangible effects on performance (e.g., they may be used to signal new policy priorities).

Two aspects of the organization of public services have conventionally been targeted by reformers. The first is size—at various times the benefits of large or small organizations have been emphasized. For example, traditional arguments on local government structure suggest that consolidation of small units will produce benefits of service coordination and

economies of scale. In contrast, public choice theorists point to the benefits of responsiveness and efficiency that supposedly follow from fragmentation (Boyne 1998). No matter which side is right, it is possible that the relationship between size and performance is nonlinear—excessively small or large organizations may be less successful than their medium-sized counterparts. A second conventional target of reform is the internal structure of public service providers as indicated by their extent of formalization (e.g., reliance on rules) and centralization of power. Organization theory suggests that formalized and centralized structures work best in simple and stable environments (Bozeman 1982; Dawson 1996). If this is so, the relationship between internal structure and service performance is likely to vary with the circumstances faced by public organizations.

A further organizational variable that has gained prominence in recent years is the *external* structure through which services are provided. This reflects the emergence of a mixed economy of service delivery that comprises public, private, and voluntary organizations. Public agencies may be part of a network or partnership or may contract out parts of their services to other providers. Whether such external forms of organizational structure lead to better performance may depend on whether they are able to mobilize more skills and resources than a single organization acting in isolation.

Management

Despite the proliferation of academic work on the meaning and extent of new public management, the relationship between management and public service performance has not been extensively theorized. In contrast, there is a voluminous and sophisticated literature on management in private organizations. This suggests that management has a significant impact on the financial success of private companies (Hansen and Wernerfelt 1989). The main management variables can be classified as follows:

- leadership styles and expertise—Much has been written on the potential effect of charismatic and transformational leadership on organizational achievements (Finkelstein and Hambrick 1996). The acid test of this perspective is whether the turnover of managerial elites (executive succession) is related to changes in service performance (Boyne and Dahya 2002).
- organizational culture—The growth of the literature on culture was sparked by Peters and Waterman (1982) and given a public-sector spin by Osborne and Gaebler (1992). The core idea is that organizations that focus on results rather than procedures and have an external rather than an internal orientation are likely to perform better.
- human resource management (HRM)—The fundamental contrast here is between hard and soft styles of HRM (Legge 1995). The former approach treats staff as instruments that can be manipulated to obtain organizational ends; the latter approach can be regarded as enlightened and humanistic, paying more attention to the needs and aspirations of individuals. It has been argued that a soft style of HRM is more likely to lead to a satisfied and motivated workforce and thereby to better organizational performance (Delaney and Godard 2001).
- strategy processes—There is a long-running debate in public administration literature on the relative merits of rational and incremental approaches to strategy

formulation and implementation (Lindblom 1959; Dror 1968). Rational planning has frequently been dismissed as impossible to implement for technical and political reasons (Boyne et al. 2003; Wildavsky 1973). However, there is extensive evidence of a positive relationship between planning and performance in private firms (Boyne 2001). It remains to be seen whether this also applies to public organizations that implement planning systems.

- strategy content—This concept refers to what organizations actually do in pursuit of their objectives. The literature on private organizations suggests that strategy content can be conceptualized at two levels. First, in terms of a "strategic stance," are organizations prospectors, defenders, or rectors (Miles and Snow 1978)? Second, in terms of "strategic actions," what steps do organizations take to operationalize their stance? These can include changing markets, products, or procedures (Porter 1980). Little theoretical attention has been devoted to the potential impact of strategy content on the performance of public organizations, so it is difficult to predict how this variable may affect service improvement.

Determinants of Service Performance: The Empirical Evidence

The purpose of this section of the article is to summarize and evaluate the results of empirical studies of public service performance. The emphasis is on statistical evidence rather than qualitative research for two reasons. First, the theories outlined in the previous section posit general relationships between a number of explanatory variables and service performance. Thus it is important to evaluate their validity across a wide range of public organizations rather than a small number of case studies. Second, although qualitative studies provide useful insights on organizational performance in the public sector, such work has recently been summarized by Rainey and Steinbauer (1999), so it is unnecessary to repeat that exercise here.

Empirical studies on service performance were identified in four main stages. First, a comprehensive manual search of four leading public administration journals was conducted for the period 1970 to 2002. These journals were *Administration and Society*, *Journal of Public Administration Research and Theory, Public Administration*, and *Public Administration Review*. Second, a "keywords in title or abstract" search was undertaken through the Web of Science database (http://wos.mimas.ac.uk), which contains the contents of the world's major social science journals (from fields such as economics, management, political science, and sociology). The keywords were *performance* (and its variants) in combination with terms reflecting the theoretical perspectives discussed in the previous section (e.g., *finance*, *resources*, *staff*, *regulation*, *competition*, *markets*, *scale*, *structure*, *size*, *contracts*, *networks*, and *management*). The search was limited to journal articles in order to provide a rough quality control on the statistical evidence. The assumption here is that work that has been subjected to peer review is more likely to meet the basic requirements of theoretical and methodological rigor. This criterion excludes studies published by government agencies and international organizations such as the Organisation for Economic Cooperation and Development (OECD) and the World Bank. Third, journal articles identified in stages one and two that were cited in the sources were obtained. These procedures mean that unpublished papers on public service performance are omitted from the analysis. The results summarized in the following discussion may therefore overstate the validity of the theoretical

models on the assumption that papers are more likely to be published if they present statistically significant results. The magnitude of this bias is unknown, but estimates in other fields suggest that it is small (Rosenthal 1991).

The final stage of the search strategy was to apply a set of conceptual and methodological criteria to more than three hundred journal articles that were identified through stages one to three. Studies were selected for inclusion in the critical review if they met the following conditions:

(a) They contained a statistical test of the determinants of service performance.

(b) They reported the statistical significance of their findings. The debate on the role of significance tests in decisions on the validity of hypotheses has recently been reopened (Gill 1999). However, all tests produce an estimated coefficient that is different from absolute zero. Therefore, some criterion is necessary for judging whether the coefficient is larger than might have occurred at random. As Winch and Campbell (1970, 206) state, "[I]t is very important to have a formal and non-subjective way of deciding whether a given set of data shows haphazard or systematic variation." The alternative to formal significance tests, they argue, "is to trust the intuitive judgement of the investigator" (Winch and Campbell 1970, 206). Moreover, unless a standard criterion is applied to the results of a variety of studies, it is impossible to assess their cumulative contribution to knowledge. Tests of statistical significance are widely accepted as an appropriate criterion for this purpose.

(c) They contained dependent variables that corresponded with at least one of the seven dimensions of service performance outlined in the second section of this article.

(d) They tested explanatory variables that reflected at least one of the theoretical perspectives discussed in the second section of this article. This criterion excludes studies that focus on the "microtechnology" of production in public services (e.g., specific techniques for teaching math in schools or procedures for treating particular medical ailments). Rather the emphasis here is on categories of explanatory variables that are potentially relevant to all public services.

The search strategy and the application of these criteria yielded a set of sixty-five empirical studies on public service performance. The remaining journal articles were conceptual discussions of organizational performance, case studies of particular organizations, or specialist analyses of the microtechnology of service production in subfields such as health, transportation, and education. Before turning to the results of the studies, two preliminary issues must be discussed: What are their general characteristics, and how should their results be combined and synthesized?

Most of the empirical studies focus on the performance of public organizations in the United States. The breakdown of studies between nations is as follows United States, 54; United Kingdom, 6; Sweden, 2; and 1 each for Australia, Finland, and Holland. Thus the geographical generalizability of the results depends on whether the determinants of public service performance are similar in the United States and in other countries. Furthermore, almost half of the studies (thirty) are on education, perhaps because outcome data (in the form of exam results) are readily available for this service. Other services that have been analyzed

include health care, housing, police, and firefighting. The methodology used in almost all of the studies is multivariate statistical modeling, sometimes combined with stochastic frontier analysis (e.g., Bjurek, Kjulin, and Gustafsson 1992; Chakraborty, Biswas, and Lewis 2001; Kirjavainen and Loikannen 1998). All but four of the studies (Midwinter and McVicar 1993; Molnar and Rogers 1976; Rushing 1974; Skogan 1976) take account of either the external or the internal characteristics of public service organizations, and many studies use statistical controls for both types of potentially confounding effects on their results. External characteristics are variables that influence the degree of difficulty in providing public services (e.g., the socioeconomic context of schools). Internal characteristics are aspects of organizations (e.g., their size and structure) and their staff (e.g., age, gender, and values).

How can general conclusions be drawn from the results of these diverse empirical studies? A formal meta-analysis is inappropriate because of the wide variation in the definition and operational status of independent and dependent variables. Meta-analytical techniques are best applied to the results of "normal science," which is a stage of development that research on public service improvement scarcely aspires to as yet. The method that is used here is based on the percentage of statistical tests that support the theoretical perspectives outlined in the second section of this article. In order to count as support for one of the theories, evidence must fulfill two conditions. First, the relationship must be in the predicted direction (e.g., more resources lead to better services). Second, the difference must be statistically significant; that is, it must be greater than would be likely to arise by chance alone (the 0.05 significance level is used in all of the empirical studies). If these criteria are applied to all of the tests in a single study, then a support score can be calculated. This is the percentage of all the tests that are reported in a study that are consistent with a service improvement hypothesis.

The final step in this analytical procedure is to construct an aggregate support score across all the studies that have tested the impact of an explanatory variable on service performance. This can be done in at least two ways (Rosenthal 1991). First, the support score for each study can be treated equally, regardless of whether it contains one test or 100 tests. This unweighted mean has the advantage in that undue importance is not attached to studies that conduct a large number of tests on the same data set. Second, the support score for each study can be weighted by the number of tests in that study. This weighted mean has the advantage in that studies that report only one (possibly idiosyncratic) test are not given undue emphasis.

In the following analysis, both the weighted and the unweighted mean support scores are reported. The real level of support for the theoretical perspectives on improvement probably lies somewhere between these figures. Although it is impossible to determine precisely where, there are grounds for suspecting that the weighted mean provides a more accurate picture. In general, studies that report the results of only one or two tests find stronger support for the hypothesized influences on performance. These studies do not appear to be methodologically stronger than studies that report multiple tests (e.g., in their measures of performance or number of control variables). It therefore seems likely that their results are unrepresentative of the general relationship between the explanatory variables and service performance.

RESOURCES

The relationship between resources and service performance has been analyzed extensively: the impact of financial resources has been examined in eighteen studies (see table 1), and the

Table 1
Impact of Financial Resources on Service Performance

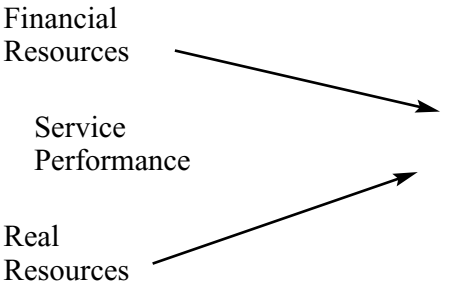
Study	Organizations and Sample Size	Measure of Financial Resources	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Sharhansky, 1967	48 U.S. states	Spending per capita	Output quantity	13	46	54	0	Yes	Yes
	48 U.S. states	Spending per capita	Output quality	1	0	100	0	Yes	Yes
	48 U.S. states	Spending per capita	Outcomes	18	5	73	22	Yes	Yes
	48 U.S. states	Change in spending	Change in quantity	12	8	84	8	Yes	Yes
	48 U.S. states	Change in spending	Change in outcomes	15	7	86	7	Yes	Yes
Bidwell and Kasarda, 1975	104 school districts in Colorado	Spending per pupil	Outcomes	2	100	0	0	Yes	Yes
Skogan, 1976	386 U.S. city police depts.	Spending per capita	Outcomes	3	100	0	0	No	No
	386 U.S. city police depts.	Spending per capita	Efficiency	3	100	0	0	No	No
Dean and Peroff, 1977	48 U.S. states	Spending per capita	Output quantity	2	100	0	0	Yes	Yes
Christensen and Sachs, 1980	100 counties in North Carolina	Spending per capita	Output quality	1	0	100	0	No	Yes
Christensen and Taylor, 1982	100 counties in North Carolina	Spending per capita	Output quality	3	0	100	0	No	Yes
	100 counties in North Carolina	Spending per capita	Outcomes	3	0	100	0	No	Yes
Choi, Allison, and Munson, 1985	44 state-owned hospitals in U.S	Spending per member of staff	Efficiency	1	0	100	0	Yes	Yes
Ostroff, 1992	298 schools in U.S. and Canada	Spending per pupil	Outcomes	7	14	86	0	Yes	Yes
	298 schools in U.S. and Canada		Consumer satisfaction	2	0	100	0	Yes	Yes
Couch, Shughart, and Williams, 1993	100 counties in North Carolina	Spending per pupil	Outcomes	1	0	100	0	No	Yes
Smith and Meier, 1994	School systems in37–49 U.S. states	Spending per pupil	Outcomes	3	33	67	0	Yes	Yes
Newmark, 1995	100 counties in North Carolina	Spending per pupil	Outcomes	7	57	43	0	No	Yes
Meier and Keiser, 1996	50 U.S. states	Spending per capita	Output quantity	1	100	0	0	Yes	Yes
Dee, 1998	4,488 school districts in U.S	Spending per pupil	Outcomes	1	100	0	0	No	Yes
Meier, Wrinkle, and Polinard, 1999	527 school districts in Texas	Spending per pupil	Outcomes	1	0	100	0	No	Yes
Meier, Polinard, and Wrinkle, 2000	1,000 school districts in Texas	Spending per pupil	Outcomes	1	0	100	1	Yes	Yes
Bohte, 2001	350 school districts in Texas	Spending per pupil	Outcomes	5	80	20	0	No	Yes
Machado, 2001	38 alcohol abuse agencies in Maine	Spending per patient	Outcomes	2	0	100	0	No	Yes
Bradley, Jones, and Millington, 2001	2,657 secondary schools in England	Spending per pupil	Outcomes	6	100	0	0	No	Yes
	2,657 secondary schools in England	Change in spending per pupil	Change in outcomes	1	100	0	0	No	Yes

[a]Mean scores: +, weighted 31%, unweighted 41%; NS, weighted 65%, unweighted 58%; –, weighted 4%, unweighted 1%.

impact of real resources has been tested in twenty-six studies (see table 2).[1] If an explanatory variable had no effect on performance, then in a large set of similar studies, around ninety percent of the results would be insignificant, with perhaps five percent showing a positive effect and five percent showing a negative effect (see Mock and Weisberg 1992). The actual distribution of the test results for resources (between positive, insignificant, and negative coefficients) is substantially different from a random distribution. The mean support score for a positive relationship between financial resources and service performance is thirty-one percent (weighted) to forty-one percent (unweighted). In contrast, the mean support score for a negative relationship is no higher than would be likely to occur by chance in a series of tests (see table 1). Similarly, the level of support for a positive impact of real resources is thirty-four to forty-six percent, whereas that for a negative effect is only seven percent.

In sum, there is virtually no indication that extra resources lead to poorer services, and there is only moderate support for the proposition that this is one route to better services. Nevertheless, it must be emphasized that a majority of the empirical evidence reveals no significant relationship between either financial resources or real resources and service performance. Taken at face value, this is consistent with public choice theorists' claim that extra money is frittered away by extravagant and inefficient bureaucrats. However, an important flaw in many of the empirical tests undermines this conclusion. The typical study includes measures of both financial resources and real resources in a single statistical equation. In other words, the evidence shows the effect of financial resources when controlling for real resources, and vice versa. The underlying direct effects model of financial resources that is being used in these studies is shown in figure 1. In contrast, a more plausible model is that financial resources allow real resources to be purchased and that the latter in turn influence

Figure 1
Direct Effects Model of Financial Resources



1 The tables exclude studies of resource effects in education summarized by Hanushek (1996) as follows:

	Percent Positive	Percent Insignificant	Percent Negative
Financial resources	18	74	8
Real resources	25	68	7

The broad pattern is similar to that in tables 1 and 2: most significant results are positive, but a majority of the evidence implies an insignificant impact of resources on performance.

Table 2
Impact of Real Resources on Service Performance

Study	Organizations and Sample Size	Measure of Real Resources	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Sharhansky, 1967	48 U.S. states per capita	Quantity of staff	Output quantity	6	17	83	0	Yes	Yes
	48 U.S. states per capita	Quantity of staff	Outcomes	5	0	60	40	Yes	Yes
Rushing, 1974	26 public hospitals in Tennessee	Staff quality	Efficiency	1	0	100	0	No	No
Bidwell and Kasarda, 1975	104 school districts in Colorado	Teacher to pupil ratio	Outcomes	2	100	0	0	Yes	Yes
	104 school districts in Colorado	Staff quality	Outcomes	2	50	50	0	Yes	Yes
Skogan, 1976	386 U.S. city police depts.	Quantity of staff per capita	Outcomes	3	100	0	0	No	No
		Quantity of staff per capita	Efficiency	3	100	0	0	No	No
Dean and Peroff, 1977	48 U.S. states	Quantity of staff per capita	Quantity of outputs	2	50	50	0	Yes	Yes
Whetten, 1978	67 manpower agencies in New York	Staff quality	Quantity of outputs	1	0	0	100	Yes	Yes
Mehay, 1979	Police depts. in 71 cities in L.A. County	Police officers per capita	Outcomes	2	50	50	0	Yes	Yes
Coulter, 1979	Fire depts. in 324 U.S. cities	Staff per capita	Outcomes	2	50	0	50	Yes	Yes
	Fire depts. in 324 U.S. cities	Staff quality	Outcomes	2	0	100	0	Yes	Yes
Christensen and Sachs, 1980	100 counties in North Carolina	Staff per capita	Quality	1	100	0	0	No	Yes
Weisman and Nathanson, 1985	78 county family planning clinics in Maryland	Staff quality	Consumer satisfaction	1	100	0	0	Yes	No
	78 county family planning clinics in Maryland	Staff quality	Outcomes	1	100	0	0	Yes	No
Walker and Williams, 1986	70 local authority housing departments in England	Quantity of staff	Quality	1	0	100	0	Yes	Yes
	70 local authority housing departments in England	Staff quality	Quality	1	0	100	0	Yes	Yes
Bjurek, Kjulin, and Gustafsson, 1992	194 day care centers for children in Sweden	Staff quality	Efficiency	1	0	0	100	Yes	Yes
Ostroff, 1992	298 schools in U.S. and Canada	Teacher to pupil ratio	Outcomes	7	28	72	0	Yes	Yes
	298 schools in U.S. and Canada	Teacher to pupil ratio	Consumer satisfaction	2	0	100	0	Yes	Yes
Smith and Meier, 1994	School systems in 37–49 U.S. states	Quantity of staff	Outcomes	3	0	100	0	Yes	Yes
Heck and Mayor, 1993	235 schools in one U.S. state	Staff quality	Outcomes	4	75	25	0	Yes	Yes

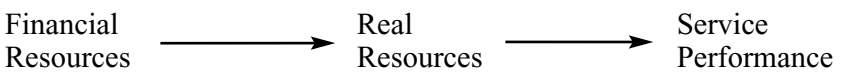
Continued

Table 2 (*continued*)
Impact of Real Resources on Service Performance

Study	Organizations and Sample Size	Measure of Real Resources	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Staley and Blair, 1995	266 school districts in Ohio	Staff quality	Outcomes	1	100	0	0	No	Yes
		Quantity of staff	Outcomes	1	0	100	0	No	Yes
Arum, 1996	51 U.S. states	Teacher to pupil ratio	Outcomes	1	100	0	0	No	Yes
Meier and Keiser, 1996	50 U.S. states	Quantity of staff	Output quantity	2	50	50	0	Yes	Yes
Zigarelli, 1996	1,100 schools in USA	Staff quality	Outcomes	1	100	0	0	Yes	Yes
Zanzig, 1997	337 school districts in California	Staff quality	Outcomes	1	100	0	0	No	Yes
Bradley and Taylor, 1998	1,307–1,580 schools in England	Teacher to pupil ratio	Outcomes	4	25	75	0	Yes	Yes
		Staff quality	Outcomes	4	25	75	0	Yes	Yes
	2,881 schools in England	Teacher to pupil ratio	Change in Outcomes	2	100	0	0	Yes	Yes
		Staff quality	Outcomes	2	0	100	0	Yes	Yes
Meier, Wrinkle, and Polinard, 1999	527 schools districts in Texas	Staff quality	Outcomes	2	0	100	0	No	Yes
Meier, Polinard, and Wrinkle, 2000	1,000 school districts in Texas	Staff quality	Outcomes	1	0	100	0	Yes	Yes
Chakraborty, Biswas, and Lewis, 2001	40 school districts in Utah	Teacher to pupil ratio	Outcomes	1	0	100	0	No	Yes
		Staff quality	Outcomes	2	0	100	0	No	Yes
Bradley, Jones, and Millington, 2001	2,657 schools in England	Teacher to pupil ratio	Outcomes	6	33	50	17	No	Yes
	2,657 schools in England	Change in teacher to pupil ratio	Change in outcomes	1	100	0	0	No	Yes
Meier and O'Toole, 2001	507 school districts in Texas	Staff quality	Outcomes	2	100	0	0	No	Yes

[a]Mean scores: +, weighted 34%, unweighted 46%; NS, weighted, 59%, unweighted, 41%; –, weighted 7%, unweighted, 7%.

Figure 2
Indirect Effects Model of Financial Resources



service performance. This indirect effects model is illustrated in figure 2. The logic of this second model suggests that the correlation between the two links in the resource chain is positive (and probably strong) and that including measures of both in the same equation will reduce their statistical significance. Furthermore, the indirect effect of financial resources through real resources is missed completely. These issues need to be explored in further empirical work, but the safest interim judgment is that the existing evidence understates the positive effect of resources on service performance.

Regulation

Only three empirical studies have examined the impact of regulation on public service performance (see table 3). Molnar and Rogers (1976) find that the number of regulatory bodies is associated with higher performance by development agencies. However, only the simple correlation between these variables is reported. Furthermore, the explanatory variable is a dichotomy (one or many higher bodies), so it is impossible to discern whether there are diminishing or eventually negative returns to regulation. The other two studies provide a mixture of positive, insignificant, and negative results for the tightness of regulatory constraints on university hospitals (D'Aunno, Hooijberg, and Munson 1991) and federal agencies (Wolf 1993). This pattern may be partly attributable to the organizations analyzed or to the measures of performance, but there are too few studies to disentangle these effects on their results. The evidence on regulatory arrangements is clearly sparse, weak, and incomplete. The biggest gap is that the impact of the nature and number of regulatory instruments is yet to be investigated. Existing empirical research provides very little basis for conclusions on whether regulatory reforms would lead to service improvement or to deterioration.

Market Structure

The validity of this theoretical perspective has been investigated in eighteen studies (see table 4). Measures of competition include the number of suppliers in a local market (e.g., Bradley, Jones, and Millington 2001; Christensen and Sachs 1980; Staley and Blair 1995), the distribution of market shares (e.g., Borland and Howsen 1992; Grosskopf et al. 2001), rivalry between public and private suppliers (e.g., Arum 1996; Domberger, Hall, and Li 1995), and service managers' perceptions of the intensity of competition (D'Aunno, Hooijberg, and Munson 1991).

The evidence on the impact of competition on service performance is very mixed. The mean support score for a positive impact is 27 percent to 38 percent, and the score for a negative impact is 24 percent (weighted and unweighted). The effects of market structure do not vary systematically with whether competition is between public agencies or between public and private providers. Nor is competition consistently linked to any specific dimension of service performance (e.g., efficiency, quality, outcomes). Even studies that use similar measures of competition and performance for the same service report very different

Table 3
Impact of Regulation on Service Performance

Study	Organizations and Sample Size	Measure of Regulation	Dimension of Performance	Number of Tests	Percentage of Tests			Internal Controls?	External Controls?
					+	NS	–		
Molnar and Rogers, 1976	110 development agencies in Iowa	Number of regulatory bodies	Outcomes	2	100	0	0	No	No
D'Aunno, Hooijberg, and Munson, 1991	35 state-owned university hospitals in U.S.	Influence of higher bodies	Consumer satisfaction	2	50	50	0	Yes	No
	35 state-owned university hospitals in U.S.	Influence of higher bodies	Output quality	2	0	100	0	Yes	No
Wolf, 1993	44 U.S. federal agencies	Influence of higher bodies	Outcomes	1	0	0	100	Yes	No

Table 4
Impact of Market Structure on Service Performance

Study	Organizations and Sample Size	Measure of Market Structure	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Christensen and Sachs, 1980	100 counties in North Carolina	Local government units per county	Output quality	1	0	0	100	No	Yes
Choi, Allison, and Munson, 1985	44 state-owned hospitals in U.S.	Number of hospitals in local market	Efficiency	1	100	0	0	Yes	Yes
D'Aunno, Hooijberg, and Munson, 1991	35 state-owned university hospitals in U.S.	Managers' perceptions of competition	Consumer satisfaction	2	0	0	100	Yes	No
		Managers' perceptions of competition	Output quality	2	0	100	0	Yes	No
Fowler and Walberg, 1991	293 schools in New Jersey	Schools in school district	Outcomes	16	6	31	63	No	Yes
Borland and Howsen, 1992	170 school districts in Kentucky	Dispersion of market shares	Outcomes	1	0	100	0	Yes	No
Couch, Shughart, and Williams, 1993	100 counties in North Carolina	Percent children in private schools	Outcomes	1	100	0	0	No	Yes
Wolf, 1993	44 U.S. federal agencies	Competition with other federal agencies	Outcomes	1	100	0	0	Yes	No

Continued

Table 4 (*continued*)
Impact of Market Structure on Service Performance

Study	Organizations and Sample Size	Measure of Market Structure	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Domberger, Hall, and Li, 1995	61 public service contracts in Australia	Competition for service contracts	Output quality	4	25	75	0	Yes	No
		Competition for service contracts	Efficiency	4	50	50	0	Yes	No
Newmark, 1995	100 counties in North Carolina	Percent children in private schools	Outcomes	7	0	100	0	No	Yes
Ruggiero, Duncombe, and Miner, 1995	636 school districts New York	Percent children in private schools	Value for money	6	0	67	33	Yes	Yes
Staley and Blair, 1995	266 school districts in Ohio	Competition							
		Number of adjacent districts	Outcomes	1	0	100	0	No	Yes
		Outcomes in adjacent districts	Outcomes	1	100	0	0	No	Yes
Arum, 1996	51 U.S. states	Percent children in private schools	Outcomes	1	100	0	0	No	Yes
Duncombe, Miner, and Ruggiero, 1997	585 school districts in New York	Percent pupils in private schools	Value for money	1	0	0	100	Yes	Yes
		Number schools in district		1	0	100	0	Yes	Yes
Zanzig, 1997	337 school districts in California	Number school districts in county	Outcomes	1	100	0	0	No	Yes
		Dispersion of market shares		1	0	100	0	No	Yes
		Percent children in private schools		1	0	0	100	No	Yes
Dee, 1998	4,488 school districts in U.S.	Percent children in private schools	Outcomes	1	100	0	0	No	Yes
Grosskopf et al., 2001	302 school districts in Texas	Dispersion of market shares	Outcomes	4	50	0	50	No	Yes
Bradley, Jones, and Millington, 2001	2,657 schools in England	Schools in local market	Outcomes	54	30	54	16	No	Yes
		Schools in local market	Change in outcomes	9	44	45	11	No	Yes

[a]Mean scores: +, weighted 27%, unweighted 38%; NS, weighted 49%, unweighted 38%; –, weighted 24%, unweighted 24%.

results. For example, Couch, Shughart, and Williams (1993) find that the percentage of children in private schools is positively related to exam performance in U.S. public schools, whereas Zanzig (1997) finds that this relationship is negative. The distribution of the test results between positive and negative coefficients flatly contradicts arguments that competition is uniformly good or bad for public services. Rather the evidence is closer to a contingency view of the benefits and costs of competition. However, existing studies provide little insight on the external or internal characteristics of public service organizations that moderate the impact of competition.

Organization

The impact of organizational structure on service performance has been tested in ten studies (see table 5). The measures of internal structure are the formalization of procedures, centralization of power and span of control. The evidence is sparse, but there is consistent support for a positive relationship between centralization and service performance (whether the latter variable is measured as output quantity, efficiency or outcomes). This is enough to cast a small doubt on the fashion for decentralization within public service organizations in recent years, although the number of tests is far from sufficient to call for a reversal of this trend. In contrast, only one of the four studies of formalization finds a positive effect on performance (the other results are all statistically insignificant). This pattern in the evidence may be attributable to differences across the environmental context of the sample organizations, but this issue has not been explored in the empirical studies.

Support for the argument that external aspects of organizational structure make a positive difference to service performance is generally weak (see table 5). Service contracting[2] is found to have an insignificant effect on the efficiency of hospitals (Alexander and Rundell 1985) and the efficiency and service quality of public transit agencies (Perry and Babitsky 1986), a negative effect on the outcomes achieved by police departments (Mehay 1979), and an insignificant or negative effect on school outcomes (O'Toole and Meier 2002b). There is slightly more evidence that characteristics of networks significantly influence public services. Whetten (1978) finds that the outputs produced by manpower agencies are positively related to the size of the network in which they operate (but not to the level of cooperation within the network); Meier and O'Toole (2001) find that educational outcomes are better if school districts have more frequent contact with other bodies. To the extent that these fragments of evidence allow any conclusion, it is that networks rather than contracts are associated with higher service performance.

The results of eighteen studies that have tested for a linear relationship between organizational size and performance are summarized in table 6. The measures of size include population, number of staff, capacity (number of hospital beds), and number of service users (school pupils). The results offer little comfort to the advocates of large or small organizations: around two-thirds of the size coefficients are insignificant, and the percentages of positive and negative results do not differ greatly (weighted mean scores of 17 percent and 14 percent, respectively). Furthermore the impact of size does not appear to be linked systematically to type of service (e.g., schools) or specific dimensions of performance (e.g., output quantity or outcomes).

Economic theory suggests that the benefits of organizational growth will eventually

2 It should be noted here that although there are many empirical studies of service contracting, almost all of them focus on expenditure rather than performance (Boyne 1998).

Table 5
Impact of Organizational Structure on Service Performance

Study	Organizations and Sample Size	Measure of Strategy Content	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Whetten, 1978	67 manpower agencies in New York	Internal organization:							
		(a) formalization	Quantity of outputs	1	100	0	0	Yes	Yes
		(b) centralization	Quantity of outputs	1	100	0	0	Yes	Yes
		External organization:							
		(a) size of network	Quantity of outputs	1	100	0	0	Yes	Yes
		(b) cooperation within network	Quantity of outputs	1	0	100	0	Yes	Yes
Mehay, 1979	Police depts. in 71 cities in L.A. County	External organization: service contracting	Outcomes	2	0	0	100	Yes	Yes
Glisson and Martin, 1980	30 organizations in one U.S. city	Internal organization:							
		(a) formalization	Efficiency	1	0	100	0	Yes	No
		(b) centralization	Efficiency	1	100	0	0	Yes	No
Alexander and Rundell, 1985	276 public hospitals in U.S.	External organization: contract management	Efficiency	1	0	100	0	Yes	Yes
Perry and Babitsky, 1986	249 public transit agencies in U.S.	External Organization: contract management	Efficiency	6	0	100	0	Yes	No
			Quality	2	0	100	0	Yes	No
Lan and Rainey, 1992	17 public organizations in New York	Internal organization formalization	Outcomes	1	0	100	0	Yes	No
Wolf, 1993	44 U.S. federal agencies	Internal organization:							
		(a) formalization	Outcomes	1	0	100	0	Yes	No
		(b) centralization	Outcomes	1	100	0	0	Yes	No
Meier and Bohte, 2000	678 school districts in Texas	Internal organization: span of control	Outcomes	2	100	0	0	Yes	Yes
Meier and O'Toole, 2001	507 school districts in Texas	External organization: frequency of contact with other bodies	Outcomes	1	100	0	0	No	Yes
O'Toole and Meier, 2002b	507 school districts in Texas	External organization: service contracting	Outcomes	2	0	50	50	Yes	Yes

[a]Mean scores (internal structure): +, weighted 57%, unweighted 62%; NS, weighted 43%, unweighted 38%; –, weighted 0, unweighted 0. Mean scores (external structure): +, weighted 11%, unweighted 25%; NS, weighted 81%, unweighted 56%; –, weighted 8%, unweighted 19%.

Table 6
Tests for Linear Effect of Size on Service Performance

Study	Organizations and Sample Size	Measure of Size	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Sharhansky, 1967	48 U.S. states	Population	Output quantity	6	0	67	33	Yes	Yes
			Outcomes	5	0	60	40	Yes	Yes
Bidwell and Kasarda, 1975	104 school districts in Colorado	Pupils	Outcomes	2	0	100	0	Yes	Yes
Dean and Peroff, 1977	48 U.S. states	Population	Output quantity	2	0	100	0	Yes	Yes
Whetten, 1978	67 manpower agencies in New York	Number of staff	Output quantity	1	100	0	0	Yes	Yes
Christensen and Sachs, 1980	100 counties in North Carolina	Number of staff	Output quality	1	100	0	0	No	Yes
Glisson and Martin, 1980	30 organizations in one U.S. city	Number of staff	Efficiency	1	100	0	0	Yes	No
Lovrich, 1985	Unspecified number of U.S. cities	Number of staff	Output quality	2	0	100	0	No	Yes
Walker and Williams, 1986	70 local authority housing departments in England	Population	Output quality	1	0	100	0	Yes	Yes
D'Aunno, Hooijberg, and Munson, 1991	35 state-owned university hospitals in U.S.	Number of hospital beds	Consumer satisfaction	2	0	100	0	Yes	No
		Number of hospital beds	Output quality	2	0	100	0	Yes	No
Fowler and Walberg, 1991	293 schools in New Jersey	Pupils	Outcomes	16	0	64	36	No	Yes
Lan and Rainey, 1992	17 public organizations in New York	Number of staff	Outcomes	1	0	100	0	Yes	No
Heck and Mayor, 1993	235 schools in one U.S. state	School size	Outcomes	4	0	25	75	Yes	Yes
Midwinter and McVicar, 1993	155 local authority library depts. in Great Britain	Population	Output quantity	6	0	100	0	No	No
			Output quality	14	21	65	14	No	No
Luyten,1994	58 U.S. schools, 95 Swedish schools, 228 Dutch schools	School size	Outcomes	5	0	100	0	No	Yes
Smith and Meier, 1994	School systems in 37–49 U.S. states	School size	Outcomes	3	33	33	33	Yes	Yes
Ruggiero, Duncombe, and Miner, 1995	636 school districts in New York	School size	Value for money	6	17	83	0	Yes	Yes
Duncombe, Miner, and Ruggiero, 1997	585 school districts in New York	School size	Value for money	1	0	100	0	Yes	Yes
Bradley, Jones, and Millington, 2001	2,657 schools in England	School size	Outcomes	6	100	0	0	No	Yes
		Change in school size	Change in outcomes	1	100	0	0	No	Yes

[a]Mean scores: +, weighted 17%, unweighted 26%; NS, weighted 69%, unweighted 63%; –, weighted 14%, unweighted 11%.

be offset by costs such as managerial overheads and bureaucratic rigidity (Boyne 1996b). The results of five studies that have explored the nonlinear relationship between size and performance are summarized in table 7. These have produced a greater percentage of significant results than studies that have examined only linear size effects, which suggests that further tests should include quadratic terms (or other equivalent procedures). Although a majority of the tests reported in table 7 are consistent with a nonlinear impact of organizational size, the pattern in the evidence is complex: almost as many tests indicate that performance at first falls with size and then eventually rises as indicate the reverse. The critical thresholds at which service performance begins to improve (or deteriorate) are also unclear. Thus whether reformers are better advised to break up large public agencies or amalgamate small ones remains obscure.

Management

Measures of various aspects of management have been included in eleven empirical studies (see table 8). Although the body of evidence is small and incomplete, it suggests that managerial variables make a difference to service performance. Indeed, this theoretical perspective on service performance receives stronger support than the other four that have already been discussed. The support score for a positive effect of management is 55 percent to 60 percent (whereas that for a negative effect is only 8 percent to 9 percent).

The impact of HRM has been analyzed in eight studies. The measures include staff satisfaction, performance-related pay, personnel stability, job security, staff morale, and job security. It is difficult to generalize from the evidence because the measures are so diverse, but the tests suggest that soft HRM aspects (satisfaction, morale) are more important than hard aspects (job security, performance-related pay). Evidence on other management variables is even sparser. Nevertheless, the four studies of leadership find positive results (Brewer and Selden 2000; Meier and O'Toole 2002a; Wolf 1993; Zigarelli 1996) as do the two studies of organizational culture (Brewer and Selden 2000; Zigarelli 1996). The only test of strategy content suggests that prospectors achieve better outcomes (Wolf 1993), whereas the extent of strategic planning makes only a marginal difference to various dimensions of service performance (Boyne and Gould-Williams 2003).

These strands of evidence are enough to justify far more sustained and comprehensive research on the impact of management on public services. The statistical results are consistent with the view that management matters, HRM, organizational culture, leadership, and strategy content may be important solutions to the problems of public service performance. Furthermore the impact of these variables on performance in the private sector has been researched widely, so there is a substantial theoretical and methodological base for further work on the public sector (see, e.g., articles in *Academy of Management Journal, Administrative Science Quarterly*, and *Strategic Management Journal*).

Summary and Research Agenda

The quantity of current knowledge on the determinants of public service performance is summarized in table 9. This table shows the number of studies that have tested the relationship between one or more of the explanatory variables and at least one dimension of service performance. The major point that emerges from this summary is simply the paucity of existing evidence. Twenty of the fifty-six boxes are completely empty, and an additional

Table 7
Tests for Nonlinear Effect of Size on Service Performance

Study	Organizations and Sample Size	Measure of Size	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Boyne, 1996a	105–400 local authorities in England	Absolute quantity of service outputs:						No	Yes
		size	Quality	10	10	50	40		
		size^2		10	60	30	10		
		size	Efficiency	3	33	0	67		
		size^2		3	67	0	33		
Ferrier and Valdmanis, 1996	219 public hospitals in U.S.	Number of hospital patients:	Efficiency					Yes	Yes
		size		4	50	25	25		
		size^2		4	25	25	50		
Bradley and Taylor, 1998	1,307–1,580 schools in England	School size:						Yes	Yes
		pupils	Outcomes	4	100	0	0		
		pupils^2		4	0	0	100		
	2,881 schools in England	pupils	Change in outcomes	2	100	0	0		
		pupils^2		2	0	100	0		
Kirjavainen and Loikannen, 1998	291 schools in Finland	School size:	Outcomes						
		pupils		1	0	100	0	No	Yes
		pupils^2		1	0	100	0	No	Yes
Meier and Bohte, 2000	678 school districts in Texas	School size:	Outcomes						
		pupils		1	0	0	100	Yes	Yes
		pupils^2		1	100	0	0	Yes	Yes

[a]Mean scores (size): +, weighted 40%, unweighted 42%; NS, weighted 30%, unweighted 25%; –, weighted 30%, unweighted 33%. Mean scores (size^2): +, weighted 35%, unweighted 36%; NS, weighted 32%, unweighted 36%; –, weighted 32%, unweighted 28%.

Table 8
Tests of Management on Service Performance

Study	Organizations and Sample Size	Measure of Management	Dimension of Performance	Number of Tests	Percentage of Tests[a]			Internal Controls?	External Controls?
					+	NS	–		
Weisman and Nathanson, 1985	78 county family planning clinics in Maryland	HRM	Consumer satisfaction	1	100	0	0	Yes	No
		(staff satisfaction)	Outcomes	1	100	0	0	Yes	No
Anderson, Shughart, and Tollinson, 1991	48 U.S. states	HRM (PRP)	Outcomes	2	0	0	100	Yes	Yes
Lan and Rainey, 1992	17 public organizations in New York	HRM (staff satisfaction)	Outcomes	1	0	100	0	Yes	No
Ostroff, 1992	298 schools in U.S. and Canada	HRM (staff satisfaction)	Outcomes	7	84	14	0	Yes	Yes
			Consumer satisfaction	2	50	50	0	Yes	Yes
Wolf, 1993	44 U.S. federal agencies	Leadership skills	Outcomes	1	100	0	0	Yes	No
		Strategic stance (prospector)	Outcomes	1	100	0	0	Yes	No
Ruggiero, Duncombe, and Miner, 1995	636 school districts in New York	HRM (job security)	Value for money	6	0	67	33	Yes	Yes
Zigarelli, 1996	1,000 schools in U.S.	Culture (pro-achievement)	Outcomes	1	100	0	0	Yes	Yes
		Leadership influence		1	100	0	0	Yes	Yes
		HRM (staff morale)		1	100	0	0	Yes	Yes
Brewer and Selden, 2000	23 U.S. federal agencies	Culture	Organizational performance	4	100	0	0	Yes	No
		HRM		4	50	50	0	Yes	No
		Leadership		1	100	0	0	Yes	No
Meier and O'Toole, 2002b	1,000 school districts in Texas	Leadership quality	Outcomes	11	90	10	0	Yes	Yes
O'Toole and Meier, 2002a	507 school districts in Texas	Personnel stability	Outcomes	22	50	50	0	Yes	Yes
Boyne and Gould-Williams, 2003	72 local authority departments in Wales	Strategy processes (extent of planning)	Service quality	5	20	60	20	Yes	Yes
			Efficiency	5	20	80	0	Yes	Yes
			Value for money	5	20	60	20	Yes	Yes
			Consumer satisfaction	5	20	60	20	Yes	Yes

[a]Mean scores: +, weighted 55%, unweighted 60%; NS, weighted 37%, unweighted 31%; –, weighted 8%, unweighted 9%.

Table 9
Extent of Existing Knowledge on Determinants of Service Performance

	Quantity	Quality	Efficiency	Equity	Outcomes	VFM	Consumer Satisfaction
Financial resources	3	3	2	0	14	0	1
Real resources	4	3	3	0	22	0	2
Regulation	0	1	0	0	2	0	1
Market structure	0	3	2	0	10	2	1
Size	4	6	3	0	11	2	1
Internal organization	1	1	1	0	3	0	0
External organization	0	3	2	0	3	0	0
Management	0	1	1	0	8	2	2

Note: Numerical values denote the number of studies that have examined the relationship between an explanatory variable and a dimension of service performance.

twenty-one have only one or two entries. The zeros or low numbers in most of the boxes constitute a research agenda in themselves. The explanatory variables that have been tested most frequently are resources, market structure, and organizational size. In contrast, few tests of regulation and internal or external structure have been undertaken. The measures of performance focus largely on the quantity and quality of outputs and the level of outcomes (especially pass rates for examinations in schools). Much less is known about influences on efficiency, value for money, and consumer satisfaction. Most strikingly, equity, the criterion of performance that may be regarded as uniquely relevant to public services, has not been examined separately (although it is subsumed within Brewer and Selden's [2000] multidimensional measure of organizational achievements).

The quality of the evidence is problematic in several ways. First, it is difficult to judge the validity of the operational measures in many studies because the underlying theoretical models are often vague. A fundamental preliminary task is to develop clearer and more sophisticated theoretical frameworks for variables such as resources, regulation, and management. At the same time, it is important to unpack further the concept of public service improvement. It may then be possible to generate and test more precise propositions regarding links between specific explanatory variables and particular aspects of service performance. For example, it may be hypothesized that regulation improves service quality and outcomes but is so expensive that it leads to lower efficiency and less value for money.

Second, a central weakness of most of the existing statistical results is that they are derived from cross-sectional models. Strictly speaking, they do not address the issue of improvement because changes over time in service standards are not examined. Only a handful of studies use measures of changes in performance as their dependent variables (including, ironically, Sharhansky's [1967] pathbreaking and ambitious study of service outputs and outcomes in the United States). The dearth of dynamic models raises questions about the direction of causality in some of the positive relationships between the explanatory variables and service performance. For example, the link between resources and performance may partly reflect the ability of successful organizations to acquire extra funds and staff. Similarly, the positive impact of size of network may arise because high-performing organizations find it easier to attract potential partners. The pattern of cause and effect is also questionable in negative statistical relationships. For example, high levels of regulation may be a consequence rather than an antecedent of poor performance. Such problems could

be tackled through systems of simultaneous equations, but it may be difficult to isolate instrumental variables that are truly exogenous. In this case, it would be better to examine changes in performance over time, which would also allow the investigation of the lagged effects of explanatory variables.

Third, the relationships between the five theoretical perspectives need to be explored. The standard statistical model in current studies is additive: all the variables are included together in a single equation, which estimates their separate effects when the others are held constant. At least two types of more complex relationship need to be analyzed.

1. Variables derived from some of the five theoretical perspectives may be causally related. For example, more resources may lead to increases in the size of service providers, more competition may lead to a change in the culture of organizations, and more regulation may constrain strategy processes and content (e.g., by suppressing innovation). The existence and consequences of such causal relationships needs to be explored through structural equations models and path analysis.
2. Some variables may moderate the impact of others. For example, the effect of competition on consumer satisfaction may depend on the level of regulation (service providers that are tightly regulated may lack the freedom to respond to consumer preferences, even if market structures are formally more open); furthermore, the impact of external organizational structure on effectiveness may vary with managerial behavior (e.g., leadership skills may be necessary to take advantage of the potential benefits of membership of a large network). Mediative relationships such as these can be examined through interaction terms in a single equation (Wright 1976) or by partitioning samples of organizations and testing whether the coefficients for explanatory variables differ significantly across subgroups.

Finally, there is an obvious and pressing need to extend the evidence to cover a wider range of political systems and services. Existing studies focus disproportionately on the United States and on education. Clearly it would be rash to assume that findings on the determinants of improvement in one nation or in one service can be generalized to all other contexts.

CONCLUSION

Few issues are as central to the discipline and profession of public administration as service improvement. The achievement of better services is often used to justify the introduction of public-sector reforms and to evaluate their impact. This article has provided a critical inventory of empirical knowledge on the determinants of public service improvement and identified an agenda for further research. It is clear that the quantity of the existing evidence is meager and that the methodologies used suffer from a variety of weaknesses. Nevertheless, it is possible to identify some important messages for academic theory and management practice.

The major theoretical implication of the analysis in this article is that public service performance is subject to systematic influences. Interorganizational differences in service outputs and outcomes are not simply random or beyond scientific explanation. The extreme contingency view that every organisationorganization is unique, and so the search for gen-

eral theory is inappropriate, is not supported by the evidence. In the terminology of Almond and Genco (1977), service improvement resembles a clock (with partly predictable movements and mechanisms) rather than a cloud (diffuse and drifting, with no consistent form). The challenge for researchers is to conceptualize and measure more clearly the relationship between the numbers on the face of the clock and the underlying causal processes.

The major practical lesson of the review of the evidence is that two of the five sets of variables emerge as the most consistent influences on performance: resources and management. The statistical results for the other theoretical perspectives are thin and/or contradictory. Thus the best advice to reformers may be to leave regulatory arrangements, organizational structure, size, and market structure as they are. Changes to these variables are largely a shot in the dark and could equally lead to poorer rather than better performance. In contrast, existing evidence provides some basis for believing that more money and better management are likely to lead to service improvement. How much money and what forms of management are two of the crucial issues that need to be explored in further research.

REFERENCES

Alexander, J., and T. Rundell. 1985. Public hospitals under contract management. *Medical Care* 23:209–19.

Almond, G., and S. Genco. 1977. Clouds, clocks and the study of politics. *World Politics* 29:489–522.

Ammons, D. 2001. *Municipal benchmarks*. London: Sage.

Anderson, G., W. Shughart, and R. Tollinson. 1991. Educational achievement and the cost of bureaucracy. *Journal of Economic Behaviour and Organization* 15:29–45.

Arum, R. 1996. Do private schools force public schools to compete? *American Sociological Review* 61:29–46.

Ashworth, R., G. Boyne, and R. Walker. 2002. Regulatory problems in the public sector: Theories and cases. *Policy and Politics* 30:195–212.

Barrow, M. 1996. Public services and the theory of regulation. *Policy and Politics* 24:263–76.

Bidwell, C., and J. Kasarda. 1975. School district organisation and student achievement. *American Sociological Review* 40:55–70.

Bjurek, H., U. Kjulin, and B. Gustafsson. 1992. Efficiency, productivity and determinants of inefficiency at public day care centres in Sweden. *Scandinavian Journal of Economics* 94:173–87.

Blank, R. 2000 When can public policy makers rely on private markets? The effective provision of social services. *Economic Journal* 110:C34–C39.

Bohte, J. 2001. School bureaucracy and school performance at the local level. *Public Administration Review* 61:92–99.

Borland, M., and R. Howsen. 1992. Student academic achievement and the degree of market concentration in education. *Economics of Education Review* 11:31–39.

Boschken, H. 1994. Organizational performance and multiple constituencies. *Public Administration Review* 54:308–14.

Boyne, G. A. 1996a. Scale, performance and the new public management: An empirical analysis of local authority services. *Journal of Management Studies* 33:809–26.

———. 1996b. *Constraints, choices and public policies*. London: JAI Press.

Boyne, G.A. 1998. *Public choice theory and local government: A comparative analysis of the UK and USA*. London: MacMillan.

———. 2001. Planning, performance and public services. *Public Administration* 79:73–88.

———. 2002. Concepts and indicators of local authority performance: An evaluation of the statutory frameworks in England and Wales. *Public Money and Management* 22, no. 2:17–24.

———. 2003. What is public service improvement? *Public Administration*, forthcoming.

Boyne, G. A., and J. Dahya. 2002. Executive succession and organizational performance in the public sector. *Public Administration* 80:179–200.

Boyne, G. A., P. Day, and R. Walker. 2002. The evaluation of public service inspection: A theoretical framework. *Urban Studies* 39:1197–1212.

Boyne, G. A., and J. Gould-Williams. 2003. Planning and performance in public organizations: An empirical analysis. *Public Management Review*, forthcoming.

Boyne, G. A., J. Gould-Williams, J. Law, and R. Walker. 2003. Problems of rational planning in public organizations: An empirical assessment of the conventional wisdom. *Administration and Society*, forthcoming.

Bozeman, B. 1982. Organization structure and the effectiveness of public agencies. *International Journal of Public Administration* 4:235–96.

Bradley, S., G. Jones, and J. Millington. 2001. The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research* 13, no. 5:545–68.

Bradley, S., and J. Taylor. 1998. The effect of school size on exam performance in secondary schools. *Oxford Bulletin of Economics and Statistics* 60:291–324.

Brewer, G., and S. Selden. 2000. Why elephants gallop: Assessing and predicting organizational performance in federal agencies. *Journal of Public Administration Research and Theory* 10:685–711.

Caiden, G. 1991. *Administrative reform comes of age*. Berlin: De Gruyter.

Carter, N., P. Day, and R. Klein. 1992. *How organizations measure success*. London: Routledge.

Chakraborty, K., B. Biswas, and C. Lewis. 2001. Measurement of technical efficiency in public education: A stochastic and non-stochastic production function approach. *Southern Economic Journal* 67:889–905.

Choi, T., R. Allison, and F. Munson. 1985. Impact of environment on state university hospital performance. *Medical Care* 23:855–71.

Christensen, J., and C. Sachs. 1980. The impact of government size and number of administrative units on the quality of public services. *Administrative Science Quarterly* 25:89–101.

Christensen, J., and G. Taylor. 1982. Determinants, expenditures and performance of common public services. *Rural Sociology* 47:147–63.

Couch, J., W. Shughart, and A. Williams. 1993. Private school enrolment and public school performance. *Public Choice* 76:301–12.

Coulter, P. 1979. Organizational effectiveness in the public sector: The example of municipal fire protection. *Administrative Science Quarterly* 24:65–81.

D'Aunno, T., R. Hooijberg, and F. Munson. 1991. Decision making, goal consensus, and effectiveness in university hospitals. *Hospital and Health Service Administration* 36:505–23.

Dawson, S. 1996. *Analysing organizations*. London: MacMillan.

Dean, G., and K. Peroff. 1977. The spending-service cliché. *American Politics Quarterly* 5:501–16.

Dee, T. 1998. Competition and the quality of public schools. *Economics of Education Review* 17:419–27.

Delaney, J., and J. Godard. 2001. An industrial relations perspective on the high-performance paradigm. *Human Resource Management Review* 11:395–429.

Domberger, S., C. Hall, and E. Li. 1995. The determinants of price and quality in competitively tendered contracts. *The Economic Journal* 105:1454–70.

Douma, S., and H. Schreuder. 1992. *Economic approaches to organizations*. New York: Prentice-Hall.

Downs, A. 1967. *Inside bureaucracy*. Boston: Little, Brown.

Dror, Y. 1968. *Public policymaking re-examined*. New York: Intext.

Duncombe, W., J. Miner, and J. Ruggiero. 1997. Empirical evaluation of bureaucratic models of inefficiency. *Public Choice* 93:1–18.

Ferrier, G., and V. Valdmanis. 1996. Rural hospital performance and its correlates. *Journal of Productivity Analysis* 7:63–80.

Finkelstein, S., and D. Hambrick. 1996. *Strategic leadership*. St Paul, Minn.: West Publishing Company.

Fowler, W., and H. Walberg. 1991. School size, characteristics and outcomes. *Educational Evaluation and Policy Analysis* 13:189–202.

Gill, J. 1999. The insignificance of null hypothesis significance testing. *Political Research Quarterly* 52:647–74.

Glisson, C., and P. Martin. 1980. Productivity and efficiency in human service organizations as related to structure, size and age. *Academy of Management Journal* 23:21–37.
Gough, I. 1979. *The political economy of the welfare state*. Oxford, England: Martin Robertson.
Grosskopf, S., K. Hayes, L. Taylor, and W. Weber. 2001. On the determinants of school district efficiency: Competition and monitoring. *Journal of Urban Economics* 49:453–78.
Hansen, G., and B. Wernerfelt. 1989. Determinants of firm performance: The relative importance of economic and organizational factors. *Strategic Management Journal* 10:399–411.
Hanushek, E. 1996. Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis* 19:141–64.
Heck, R., and R. Mayor. 1993. School characteristics, school academic indicators and student outcomes: Implication for policies to improve schools. *Journal of Education Policy* 8:143–54.
Heffron, F. 1989. *Organization theory and public organizations*. Englewood Cliffs, N.J.: Prentice Hall.
Heinrich, C., and L. Lynn, eds. 2000. *Governance and performance*. Washington D.C.: Georgetown University Press.
Hilke, J. 1993. *Competition in government—Financed services*. New York: Quorum Books.
Hood, C., O. James, G. Jones, C. Scott, and T. Travers. 1998. *Regulation inside government*. Oxford, England: Oxford University Press.
Kirjavainen, T., and H. Loikannen. 1998. Efficiency differences of finish senior secondary schools: An application of DEA and TOBIT analysis. *Economics of Education Review* 17:377–94.
Lan, Z., and H. Rainey. 1992. Goals, rules and effectiveness in public, private and hybrid organisations: More evidence on frequent assertions about differences. *Journal of Public Administration Research and Theory* 2:5–28.
Legge, K. 1995. *Human resource management: Rhetorics and realities*. London: MacMillan.
Le Grand, J., and W. Bartlett. 1993. *Quasi-markets and social policy*. London: MacMillan.
Lindblom, C. 1959. The science of muddling through. *Public Administration Review* 39:517–26.
Lovrich, N. 1985. Scale and performance in government operations: An empirical assessment of public choice prescriptions. *Public Administration Quarterly* 9:163–85.
Luyten, H. 1994. School size effects on achievement in secondary education: Evidence from the Netherlands, Sweden and the USA. *School Effectiveness and School Improvement* 5:75–99.
Machado, M. 2001. Dollars and performance: Treating alcohol misuse in Maine. *Journal of Health Economics* 20:639–68.
McKean, R. 1965. The unseen hand in government. *American Economic Review* 55:496–506.
Mehay, S. 1979. Intergovernmental contracting for municipal police services: An empirical analysis. *Land Economics* 55:59–72.
Meier, K., and J. Bohte. 2000. Ode to Luther Gulick: Span of control and organizational performance. *Administration and Society* 32:115–37.
Meier, K., and L. Keiser. 1996. Public administration as a science of the artificial: A methodology for prescription. *Public Administration Review* 56:459–66.
Meier, K., and L. O'Toole. 1999. Modelling the impact of public management: The implications of structural context. *Journal of Public Administration Research and Theory* 9:505–26.
———. 2001. Managerial strategies and behaviour in networks: A model with evidence from U.S. public education. *Journal of Public Administration Research and Theory* 11:271–93.
———. 2002a. Public management and organizational performance: The impact of managerial quality. *Journal of Policy Analysis and Management* 21:629–43.
———. 2002b. Public management and educational performance: The impact of managerial networking. *Public Administration Review*, forthcoming.
Meier, K., J. Polinard, and R. Wrinkle. 2000. Bureaucracy and organizational performance: Causality arguments about public schools. *American Journal of Political Science* 44:590–602.
Meier, K., R. Wrinkle, and J. Polinard. 1999. Equity versus excellence in organizations: A substantively weighted least squares analysis. *American Review of Public Administration* 29:5–18.
Midwinter, A., and M. McVicar. 1993. Population size and functional efficiency in public library authorities: The statistical evidence. *Journal of Librarianship and Information Science* 25:187–96.
Miles, R., and C. Snow. 1978. *Organizational strategy, structure and process*. New York: McGraw-Hill.

Mock, C., and H. Weisberg. 1992. Political innumeracy: Encounters with coincidence, improbability and chance. *American Journal of Political Science* 36:1023–46.
Molnar, J., and D. Rogers. 1976. Organizational effectiveness: An empirical comparison of the goal and system resource approaches. *The Sociological Quarterly* 17:401–13.
Newmark, C. 1995. Another look at whether private schools influence public school quality. *Public Choice* 82:365–73.
Niskanen, W. 1971. *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
Osborne, D., and T. Gaebler. 1992. *Reinventing government*. Reading Mass.: Addison-Wesley.
Ostroff, C. 1992. The relationship between satisfaction, attitudes and performance: An organizational level analysis. *Journal of Applied Psychology* 77:963–74.
O'Toole, L., and K. Meier. 2002a. *Plus ça change:* Public management, personnel stability, and organizational performance. *Journal of Public Administration Research and Theory*, 13, no. 1:43–64.
———. 2002b. Parkinson's law and the new public management? Contracting determinants and service quality consequences in public education. *Public Administration Review*, forthcoming.
Pawson, R. 1989. *A measure for measures*. London: Routledge.
Perry, J., and T. Babitsky. 1986. Comparative performance in urban bus transit: Assessing privatization strategies. *Public Administration Review* 46:57–66.
Peters, T., and R. Waterman. 1982. *In search of excellence*. New York: Harper and Row.
Pollitt, C. 1984. *Manipulating the machine*. London: Allen & Unwin.
Pollitt, C., and G. Bouckaert. 2000. *Public management reform*. Oxford, England: Oxford University Press.
Porter, M. 1980. *Competitive Strategy*. New York: Free Press.
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.
Rosenthal, R. 1991. *Meta-analytical procedures for social research*. London: Sage.
Ruggiero, J., W. Duncombe, and J. Miner. 1995. On the measurement and causes of technical inefficiency in local public services: With an application to public education. *Journal of Public Administration Research and Theory* 5:403–28.
Rushing, W. 1974. Differences in profit and non-profit organizations: A study of effectiveness and efficiency in general short-stay hospitals. *Administrative Science Quarterly* 19:474–84.
Sayer, A. 1984. *Method in social science: A realist approach*. London: Hutchinson.
Sharhansky, I. 1967. Government Expenditures and public services in the American states. *American Political Science Review* 61:1066–77.
Skogan, W. 1976. Efficiency and effectiveness in big-city police departments. *Public Administration Review* 54:551–58.
Smith, K., and K. Meier. 1994. Politics, bureaucrats and schools. *Public Administration Review* 54:551–58.
Staley, S., and J. Blair. 1995. Institutions, quality competition and public service provision: The case of public education. *Constitutional Political Economy* 6:21–33.
Tullock, G. 1965. *The politics of bureaucracy.* Washington D.C.: Public Affairs Press.
Udehn, L. 1996. *The limits of public choice.* London: Routledge.
Walker, R., and J. Williams. 1986. Housing benefits: Some determinants of administrative performance. *Policy and Politics* 14:309–34.
Weisman, C., and Nathanson, C. 1985. Professional satisfaction and client outcomes. *Medical Care* 23:1179–92.
Whetten, D. 1978. Coping with incompatible expectations: An integrated view of role conflict. *Administrative Science Quarterly* 23:254–71.
Wildavsky, A. 1973. If planning is everything, maybe it's nothing. *Policy Sciences* 4:127–53.
Williamson, O. 1975. *Markets and hierarchies: Analysis and antitrust implications.* New York: The Free Press.
Winch, R., and D. Campbell. 1970. Proof? No. Evidence? Yes. The significance of tests of significance. In *The significance test controversy*, edited by D. Morrison and M. Henkel. London: Butterworths.

Wolf, P. 1993. A case survey of bureaucratic effectiveness in U.S. cabinet agencies: Preliminary results. *Journal of Public Administration Research and Theory* 3:161–81.

Wright, G. 1976. Linear models for evaluating conditional relationships. *American Journal of Political Science* 20:349–73.

Zanzig, B. 1997. Measuring the impact of competition in local government education markets on the cognitive achievements of students. *Economics of Education Review* 16:431–41.

Zigarelli, M. 1996. An empirical test of conclusions from effective schools research. *The Journal of Educational Research* 90:103–111.