 JPART 15:633–639

Where Next? Research Directions on Performance in Public Organizations

George A. Boyne
Cardiff University

Kenneth J. Meier
Texas A&M University and Cardiff University

Laurence J. O'Toole Jr.
University of Georgia

Richard M. Walker
University of Hong Kong and Cardiff University

The dependent variable in all the articles presented in this symposium issue of *JPART* was organizational performance in public agencies. Management variables included organizational strategy, resources, leadership, goals, workforce diversity, and representation. The empirical articles also controlled for external effects on performance. Although the collection of articles presented here is a major step forward for research on public organizations, the evidence that management matters needs additional exploration, and the methods used by scholars need to be improved.

Public organizations are required to address a range of goals, some of which may be in conflict. This obligation in turn requires organizations that deliver public services to focus their attention on multiple dimensions of performance. Boyne (2002) has identified sixteen dimensions of performance in public organizations grouped into five themes—outputs, efficiency, effectiveness, responsiveness, and democratic outcomes. Four of these themes are captured in the empirical articles presented here. Output quality is considered by Andrews et al., Brewer, Chun and Rainey, and Martin and Smith, while output quantity features in the study by Andrews et al.[1] Efficiency is considered as part of a composite indicator of organizational performance by Andrews et al. All the articles consider measures of service outcomes and include managerial effectiveness, productivity (Brewer; Chun and Rainey), and formal effectiveness (Andrews et al.; Hill; Martin and Smith; Pitts). Measures of responsiveness figure in two articles—Chun and Rainey examine customer service orientation, and Andrews et al. include citizen satisfaction and consumer satisfaction.

The articles reflect the "state of the art" in the field of the determinants of performance in public organizations: the majority of the studies emphasize service delivery (Boyne 2003; Forbes and Lynn). Though knowledge on service delivery is becoming clearer, the coverage of international settings and service delivery themes remains limited.

Address correspondence to George A. Boyne at boyne@cardiff.ac.uk.

1 Individual measures are used in the articles by Andrews et al., Chun and Rainey, Martin and Smith, Hill, and Pitts. The studies by Andrews et al., Brewer, and Chun and Rainey use aggregate measures of organizational performance.

doi:10.1093/jopart/mui037
Advance Access publication on February 18, 2005

More research is required that addresses other dimensions of performance. In particular, research needs to examine the impact of management on dimensions of performance that are central to the wider effects of public service organizations, such as equity and cost-effectiveness. Some of the work has already begun (for example, in Boyne et al's. [2003] examination of the impact of public choice reforms on equity), but much more is required. The final group of measures proposed by Boyne (2002) focuses upon "democratic outcomes" and embraces probity, participation, accountability, and cost per unit of democratic outcome. To our knowledge no public management studies have taken these dimensions of performance as their dependent variables. In short, much work remains to be undertaken.

PUBLIC MANAGEMENT MATTERS FOR PERFORMANCE, BUT WHEN, WHERE, AND HOW?

The studies included in this symposium add credence to the notion that public management and public managers influence the performance of their organizations and of public programs more broadly. Management is by no means the only determinant of performance, as Forbes and Lynn remind the reader, but it is increasingly clear that managers can improve program effectiveness, sometimes in substantial ways. These findings, therefore, are generally consistent with those reported in Boyne's (2003) review of the evidence and with the accumulating recent systematic research results reported elsewhere (for instance, Meier and O'Toole 2003; Nicholson-Crotty and O'Toole 2004; O'Toole and Meier 2004).

This general proposition, needless to say, has rarely been in serious doubt, but only in the last few years have rigorous investigations begun to validate the idea. This symposium not only adds to the general evidence but also does so with supportive findings on both "subjective" (as in Brewer's investigation) and "objective" (as in the work reported by Hill and by Pitts) performance measures. Furthermore, this set of articles begins to differentiate the broad category of "public managers" and explore the impacts of management in different governmental forms, at different points of influence, and at various levels of responsibility. Local general-purpose governmental strategic approaches have impacts on service quality (Andrews et al. forthcoming), as do efforts by the managers of special-district governmental forms (the Texas school district findings). Levels within organizations have different impacts—top management matters (Hill), multiple levels have impacts (Pitts), and frontline supervisors are also consequential (Brewer)—and reflect the need for theories that capture the variety and complexity of organizational life (Walker and Enticott 2004). Managerial influence can be seen in outputs (Chun and Rainey) and outcomes (Andrews et al.).

Still, these results only catalyze the development of as-yet-unanswered questions. Managerial strategy seems to matter when it interacts with other organizational features, like representativeness (Andrews et al.), but what of numerous other possible interactive impacts of managerial strategy and behavior? Managerial aspects seem to matter for some dimensions of performance but not all (Pitts). Why? It is particularly useful to assess such patterns via multiple performance measures, in part because there may be trade-offs among certain performance criteria (Martin and Smith). Public management seems likely to display different impacts when performance is assessed in the short term than when attention is focused on the longer haul (Hill). The causal logic involved in understanding the precise forms of managerial influence over results is a major subject demanding extensive exploration.

Developing a clearer theoretical context for the possibilities and also limits of public managerial influence would be a useful starting place. Discussions in this field often implicitly suggest that managers have great potential to shape results, but one does not need to be a proponent of the organizational ecology perspective to recognize that many influences on performance are external to the administrative system and cannot realistically be shaped by the efforts of even the most capable and committed cadre of managers. This point is accepted by the researchers whose results are reported here. Management is situated within a multilevel governance system composed of many different elements (Forbes and Lynn), and the estimations regarding control variables, as introduced into the multivariate empirical studies, testify to the importance of exogenous influences. Elsewhere, scholars have shown the folly of rewarding or penalizing public authorities for performance results that extend beyond their abilities to control (Andrews et al. forthcoming), and it is important for the field to learn the range of feasible impacts that could be expected via the direct or indirect interventions of management.

Within the limits of the possible, then, what kinds of effect sizes can be expected from the efforts of public managers? Other studies have suggested that maximum managerial impacts on performance may be in the range of roughly 8 to 10 percent of overall outcomes, at least for established and relatively stable programs (for instance, O'Toole and Meier 2003), but this question has barely been explored. Some of the case study literature implies that managers may have a nonincremental (and positive) influence on performance under certain conditions, but no systematic work on public organizations has yet been conducted to test this notion.

Clearly, a major part of the research agenda to be explored in the future needs to be devoted to examining which aspects or dimensions of public management carry performance implications. Qualitative work can help in exploring the subject and identifying possible points and processes of influence, but investigating this question will require efforts to model precisely such complex performance relationships.

Public managers expend varying amounts of time and energy motivating those employees involved in program implementation, clarifying objectives, staffing and funding priority operations, resolving internal organizational conflicts, structuring and restructuring bureaus and task forces, developing strategies for improving performance in the future, and making decisions about committing inputs to production. They also deal with an array of environmental challenges and opportunities, including competition from other units, the involvement of politicians and political institutions, the needs and demands of programs' target groups and the public more generally, and—especially in the United States—the vicissitudes of litigation. Further, as mounting empirical evidence indicates (for instance, Hall and O'Toole 2004), they must aim to extract results from organizational patterns considerably more complicated than the classic Weberian setting. Public–private links of all sorts, the involvement of nonprofit organizations in the delivery of public services, and networked arrangements of some complexity often mean that the managerial challenge must also involve the development and maintenance of collaborative ties across diverse and differently situated organizational actors. Which sorts of managerial actions are likely to accrue performance impacts, in both the short and long term, are surely functions of multiple contingencies, including varying sorts of goals and types and levels of goal ambiguity; there is no one-size-fits-all set of injunctions. Two factors make matters more interestingly complicated. First, different managers perform these varied tasks with different levels of skill. And the interactions of the multiple management functions with some production-relevant variables is almost certainly nonlinear. Second, while there are multiple contingencies, the strategies managers implement may not be their own and may not be fully specified. Public managers are frequently required to adopt programs of management reform. The imposition of a set of management practices in one area may have negative consequences for organizational performance. If, as has been argued elsewhere (MacDuffie 1995), higher levels of organizational performance are achieved by bundled or simultaneous alternation to management and organizational arrangements in the full system rather than the sum of each practice individually (Ichniowski, Shaw, and Prennushi 1997), further questions are raised about how practices interact and reinforce one another or are linked to organizational functions and strategies. While some initial steps have been made toward the development of models to specify the role of "public management" in the quest for performance (Meier and O'Toole 2002, 2004; O'Toole and Meier 1999), most of the heavy lifting certainly remains ahead.

PUBLIC MANAGEMENT AND ORGANIZATIONAL PERFORMANCE: METHODOLOGICAL ISSUES

Issues of method and measurement abound in the study of public management and its linkage to organizational performance. While the articles in this symposium touch on several such issues and illustrate appropriate methods, much remains to be done. This section will discuss five methodological problems that need to be solved for progress to be made: the use of appropriate techniques to deal with multiple indicators of performance, taking issues of measurement seriously, specifying the correct relationships and functional forms, shifting the focus to optimizing rather than least squared error techniques, and adjusting research methods to reflect the Bayesian nature of public management decisions.

First, all public organizations have multiple goals, either because they are assigned multiple programs or because enabling legislation contains multiple objectives. While a single indicator of organizational performance might be the holy grail of private sector research, no single indicator can capture the complexities of public organizational performance in the twenty-first century. Multiple programs, multiple goals, and multiple indicators imply both complementarities and conflict within the organization. Although management plays a role in resolving such conflicts (to the extent they are resolved [see Simon 1947]), Martin and Smith clearly demonstrate the inadequacies of current methods in analyzing multiple indicators. The use of seemingly unrelated regression provides a reasonable solution to the problems and should become standard practice in the field. Additional concern needs to be directed to establishing the causal sequence among multiple indicators; the correlated errors problem is likely worse when one indicator serves as an input for another indicator than when there are trade-offs only because of the overall constraint imposed by limited resources or problems of underspecification. Any causal linkages between multiple indicators generates yet another source of error that will be compounded across equations.

Second, measurement issues need to be given much more consideration in the study of public management and organizational performance. Virtually every scholar concedes that a performance indicator is always a rough estimation of what we hope that programs will achieve (e.g., employment for six months versus permanent employment in jobs with growth potential), yet this issue of measurement validity is rarely considered. The implications of maximizing performance on a partially valid indicator raises the specter of suboptimization (Blau 1956) or even systematic cheating (Bohte and Meier 2000). Similarly, even the bare minimum element of measurement, reliability, gets little attention, perhaps because the focus on performance measurement has detracted from creating reliable measures of public management (for one such effort, see Meier and O'Toole forthcoming). In contrast to the work on performance measurement, the measurement effort on public management is still in its infancy. But even in the area of performance measurement, the difference between subjective and objective measures of performance creates additional difficulties. The Andrews et al. linkage between subjective performance and race (even with controls for objective performance) is troubling simply because the subjective indictor appears to have an element of racial bias in it. To the degree that the subjective indicators (and objective indicators also) contain systematic rather than random measurement error, our quantitative analyses will be biased. Such biases are likely greater to the extent that subjective measures rely on the assessments of organizational employees or the general public rather than service consumers. Organizational employees have a vested interest in positive evaluations, and the general public may lack the familiarity that service consumers have with program outputs.

Third, management theories have far outstripped empirical research, particularly in using the appropriate specification of hypothesized relationships. The research literature relies primarily on linear relationships to estimate the impact of management on performance. In areas where nonlinear relationships have been taken seriously (Meier and O'Toole 2003), data sets are rarely robust enough to deal with multiple nonlinear relationships simultaneously, since most techniques rely on estimations that induce a great deal of collinearity into the models. Similarly, theories (Perrow 1972; Woodward 1980) suggest that contingent relationships are central to management, and contingencies imply interactions among two or more variables. The estimation process for interactions again generates significant collinearity that is unlikely to be overcome in any modest-sized data set. The two obvious solutions are larger data sets with greater variation among the organizations and management practices (and thus long-term efforts to create data sets that extend over significant periods of time) and the adoption of alternative quantitative techniques such as simulation and math modeling. These alternative techniques, however, will also require more precise theories to specify relationships when the data are inadequate to do so.

Fourth, public management as a field of study is ill-served by the regression-based techniques that it adopts from the social sciences. Public management is a design practice (and a design science for those interested in the scholarly study of management) concerned as much with how things might be as how things are (Simon 1969). Management theory suggests that successful management will be able to produce greater outputs from a given level of inputs: in short, that the relationship between inputs and outputs will vary based on management. Such an emphasis cannot be picked up by techniques that specify that the same relationship holds between a resource (say, budgets) and performance across all organizations. Because least squares regression techniques do exactly this (see Meier and Gill 2000), they need to be supplemented with other techniques that allow scholars to see how individual managers and organizations optimize. Whether either of the two common approaches (Data Envelopment Analysis or Substantively Weighted Analytical Techniques) is the solution (Martin and Smith; Meier and Gill 2000; Wu, Bretschneider, and

Marc-Aurele 2002) or not is less important than that they are designed to overcome the shortcomings of techniques that cannot assess optimization.

Finally, statistical methods need to reflect the data-generation process. In public management this means recognition of one simple fact. Public management decisions are rarely unconstrained simply because bounded rationality restricts the manager from addressing all issues facing the organization. Rather, the issues that float to the top of management's agenda are determined by how well the organization is performing on its multiple dimensions. While the statement that management actions are contingent on past performance is noncontroversial in the public management literature, it implies a dramatic shift in our methods of analysis. Because managers do not operate from a blank slate and because even their need to make decisions is a function of past performance, managers operate in a Bayesian world. All managers have a set of expectations based on experience and judgment. The existence of such priors in a system where whether or not a decision is likely to even be made is contingent on past performance fits exactly the assumptions of Bayesian statistics and fits poorly with the assumptions of classical statistics (see Gill and Meier 2000).

While each of the problems in quantitative methods poses challenges to the study of public management and organizational performance, they all also provide opportunities. Quantitative methods are tools that can be used to examine the relationship between public management and organizational performance. Selecting the right tool for the right task, however, is far more important than the literature has acknowledged.

In sum, further theoretical work on the impact of public management on organizational performance needs to address multiple contingencies and multiple criteria of success and failure. In addition, our empirical methods need to reflect this theoretical complexity more closely, by using more sophisticated and nuanced techniques. A close linkage between such improvements in theory and methods holds the promise of further substantial developments in our understanding of the determinants of performance in public organizations.

REFERENCES

Andrews, Rhys, George A. Boyne, Jennifer Law, and Richard Walker. Forthcoming. External constraints on local service standards: The case of comprehensive performance assessment in English local government. *Public Administration.*

Blau, Peter M. 1956. *Bureaucracy and modern society.* New York: Random House.

Bohte, John, and Kenneth J. Meier. 2000. Goal displacement: Assessing the motivation for organizational cheating. *Public Administration Review* 60 (2): 173–82.

Boyne, George A. 2002. Concepts and indicators of local authority performance: An evaluation of the statutory framework in England and Wales. *Public Money and Management* 22 (2): 17–24.

———. 2003. Sources of public service improvement: A critical review and research agenda. *Journal of Public Administration Research and Theory* 13 (2): 767–94.

Boyne, George, Catherine Farrell, Jennifer Law, Martin Powell, and Richard Walker. 2003. *Evaluating public management reforms: Principles and practice.* Buckingham, UK: Open University Press.

Gill, Jeff, and Kenneth J. Meier. 2000. Public administration research and practice: A methodological manifesto. *Journal of Public Administration Research and Theory* 10 (1): 157–200.

Hall, Thad E., and Laurence J. O'Toole Jr. 2004. Shaping formal networks through the regulatory process. *Administration and Society* 36 (2): 1–22.

Ichniowski, Casey, Katheryn Shaw, and Giovanna Prennushi. 1997. The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review* 87 (2): 291–313.

MacDuffie, John Paul. 1995. Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial and Labor Relations Review* 48 (1): 197–221.

Meier, Kenneth J., and Jeff Gill. 2000. *What works: A new approach to program and policy analysis.* Boulder, CO: Westview Press.

Meier, Kenneth J., and Laurence J. O'Toole Jr. 2002. Public management and organizational performance: The impact of managerial quality. *Journal of Policy Analysis and Management* 21 (3): 629–43.

———. 2003. Public management and educational performance: The impact of managerial networking. *Public Administration Review* 63 (6): 689–99.

———. 2004. Conceptual issues in measuring and modeling management and its impacts on performance. In *The art of governance: Analyzing management and administration,* ed. Patricia Ingraham and Laurence J. Lynn Jr., 195–223. Washington, DC: Georgetown University Press.

———. Forthcoming. Managerial networking: Issues of measurement and research design. *Administration and Society.*

Nicholson-Crotty, Sean, and Laurence J. O'Toole Jr. 2004. Testing a model of public management and organizational performance: The case of law enforcement agencies. *Journal of Public Administration Research and Theory* 14 (1): 1–18.

O'Toole, Laurence J., Jr., and Kenneth J. Meier. 1999. Modeling the impact of public management: The implications of structural context. *Journal of Public Administration Research and Theory* 9 (3): 505–26.

———. 2003. *Plus ça change*: Public management, personnel stability, and organizational performance. *Journal of Public Administration Research and Theory* 13 (1): 43–64.

———. 2004. Desperately seeking Selznick: Cooptation and the dark side of public management in networks. *Public Administration Review* 64 (6): 681–93.

Perrow, Charles. 1972. *Complex organizations: A critical essay.* Glenview, IL: Scott Foresman.

Simon, Herbert A. 1947. *Administrative behavior*. New York: Free Press.

———. 1969. *The sciences of the artificial.* Cambridge, MA: MIT Press.

Walker, Richard M., and Gareth Enticott. 2004. Using multiple informants in public administration: Revisiting the managerial values and actions debate. *Journal of Public Administration Research and Theory* 14 (3): 417–34.

Woodward, Joan. 1980. *Industrial organization: Theory and practice.* 2d ed. New York: Oxford University Press.

Wu, Jiannan, Stuart Bretschneider, and Frederick Marc-Aurele. 2002. Estimating models of extreme behavior: A Monte Carlo comparison between SWAT and quantile regression. *Chinese Public Administration Review* 1 (2): 165–76.