

**Administrative or Survey Data for Measuring Organizational Performance:
What's the Difference?**

Rhys Andrews
Cardiff University

George A. Boyne
Cardiff University

Richard M. Walker
University of Hong Kong

Prepared for the session on 'Researching Public Management and Public Service Delivery: Taking Stock of the Methodological Repertoire' at the XIV International Research Society for Public Management Conference, Berne, Switzerland.

Administrative or Survey Data for Measuring Organizational Performance: What's the Difference?

Abstract

Debate about the best way to measure performance in studies of management in public organizations is longstanding. We address this topic through a review of the evidence from 93 studies that use administrative and/or survey measures of organizational performance. We find that administrative data typically reflect the performance judgements of government (at the central, regional and local level) and regulators, while survey data is based on the perceptions of citizens, service users and public sector managers. We undertake a critical review of the twelve articles that use both administrative and survey measures of organizational performance. This reveals limited differences in the impact of management variables on the two types of performance measures. However, in those studies using survey measures, management variables are more likely to have a positive link with the performance judgements of service consumers than the judgements of managers themselves. This implies that public managers may underestimate their impact on citizens' perceptions of organizational performance.

Administrative or Survey Data for Measuring Organizational Performance: What's the Difference?

A major quest of public management researchers in recent years has been to understand why some organizations perform better than others. A wave of research studies has investigated the 'determinants' of organizational performance in various national and institutional settings, notably in the U.S. Federal government, English local government, Texas School Districts, U.S. state agencies, and Welsh local governments (Ashworth et al., 2010; Boyne et al., 2006; Walker et al., 2010). This research topic has been given further prominence by symposia in journals such as *Public Administration* (Walker and Boyne, 2009) and the *Journal of Public Administration Research and Theory* (Boyne and Walker, 2005; Boyne and Hood, forthcoming).

A fundamental issue that researchers in this field must tackle is how to define and measure organizational performance. In this paper we examine one element of this problem, namely the relative merits of measuring performance with administrative or survey data. Administrative measures are based on data collected and published by government organizations themselves, such as indicators that are mandated by higher levels of government, and scores issued by government inspectors. Major examples of these in the UK have included Best Value Performance Indicators for local authorities, and 'star ratings' awarded by inspectors of schools and hospitals. Survey measures may also be set by political superiors (such as mandatory surveys of consumer satisfaction), but are often based on primary data collected by academic researchers (usually through surveys of public officials' views of the performance of their organizations).

We review studies of public management and performance, and summarise the types of performance measures that are dominant in the field and the dimensions of performance

that have been covered. This is followed by an analysis of whether the empirical evidence on the determinants of performance is affected by the types of measures that have been used. We pay particular attention to the results of studies that have tested the same explanatory variables on both administrative and survey measures of performance. To what extent are the implications of the evidence for theory and practice shaped by the types of performance measures that are used?

Organizational Performance in the Public Sector: Conceptual Dimensions and Stakeholder Perceptions

An ideal set of measures for explaining and evaluating performance would satisfy a multitude of conceptual and technical criteria. We use the expression '*set* of measures' deliberately, because it is impossible for any single measure to meet all of the relevant criteria. To some extent, these criteria can be met by composite measures that cover many different dimensions of performance; yet such measures often introduce extra problems of weighting and combining separate indicators. In this section we set out two main conceptual criteria that can be used to map performance measures: the number of different dimensions of performance that are covered, and the variety of stakeholder perceptions that are taken into account. This is important because differences between administrative and survey measures on one or both of these criteria may lead to different evidence on the effects of public management variables. In the following section, we assess the extent to which the administrative and survey data that have been used to measure performance in empirical studies differ on these criteria in practice.

The concept of organizational performance is widely recognised as comprising many different dimensions. Therefore, a comprehensive set of administrative or survey measures would cover multiple dimensions of performance. These include economy (cost per input of given quality), quantity and quality of outputs, efficiency, effectiveness, value for money,

equity, responsiveness to service needs, and consumer and citizen satisfaction. Whether administrative or survey measures cover all of these dimensions will depend on the priorities of the stakeholders who set them.

Administrative indicators of performance are selected by governments and their agents, so their composition reflects the priorities of powerful groups within the state. This means that their coverage is likely to vary across nations and over time. For example, during the ‘boom’ years of the 1990s and early 2000s, statutory performance indicators for local authorities in the UK shifted from a focus on service inputs to outputs and outcomes (Boyne, 2002). In the next decade, as financial stringency afflicts public organizations, the focus is likely to be on efficiency and value for money. All of this implies that studies which use administrative data to test theories of the impact of management on performance are in effect providing evidence on whether management makes a difference to the achievements that are valued by political principals.

Survey measures of performance that are mandated by government are also likely to be shaped by the values of political principals (thus the emphasis on consumer satisfaction in recent years, which reflects a central tenet of NPM that services should be responsive to public preferences rather than bureaucratic preferences). Many survey measures, however, are designed by public management researchers, and therefore (implicitly or explicitly) reflect their values and beliefs about what constitutes good or bad performance. Thus when managers are asked to respond to researchers’ surveys about the performance of their organizations, the content of the questions may not reflect the aspects of performance that they themselves would prioritise. An implication is that the results from such studies reveal whether management makes a difference to performance as filtered through the values of both researchers and managers. In principle, then, it seems likely that testing the same

explanatory variables on administrative and survey measures of performance will produce different results.

This discussion leads directly to the second conceptual criterion for mapping performance measures: the range of different stakeholders whose views are incorporated in them. Organizational performance, perhaps especially in the public sector, is open to a variety of interpretations and is politically contestable. The multiple stakeholders that judge performance include political principals, funders (usually higher levels of government), professionals, managers, front-line staff and service recipients (and non-recipients). These stakeholders may disagree about the most important dimensions of performance, the indicators used to measure them, and the scores on these indicators that constitute success or failure. Thus, in addition to different results for administrative and survey measures, we should expect different patterns of empirical evidence when the same management variables are tested on survey data drawn from different stakeholder groups.

Administrative and Survey Measures in Practice

In this section of the paper we review the measures that have been used as dependent variables in studies of organizational performance in the public sector. Which dimensions of performance, and whose perceptions of performance, have been covered in these studies? To answer these questions we examined the contents of five of the leading journals that have regularly published articles that seek to identify why some organizations perform better than others. These are: *Administration and Society*, *Journal of Public Administration Research and Theory*, *Public Administration*, *Public Administration Review* and *Public Management Review*. We then ‘snowball’ from these journals to examine any public management studies that utilise both administrative and survey measures of performance. Although the ‘performance movement’ is more than two decades old as a field of management practice, it is only in the last ten years or so that public management researchers have been using the

archival data from government sources which the movement has spawned to test theoretical propositions in rigorous models. Similarly, the recognition of the need to complement and supplement the administrative data with survey data on performance is also a recent phenomenon. Therefore we restricted our search of journal articles to the period from 2000 onwards.

A thorough review of the available evidence within the selected journals requires the adaptation of additional search terms to ensure that we capture the variety of dimensions of performance. We therefore searched for effectiveness, efficiency, equity, failure, outcome, output, productivity, quality, quantity, responsiveness, results, satisfaction and trust in the title, abstract or key words of articles. Two major search engines (the SSCI and Scopus) were used to ensure that our findings were not biased by the program requirements of a single engine. Studies analysing employee outcomes, such as turnover, job satisfaction and representativeness, were excluded from our review as they measure the internal administrative achievements of public organizations, rather than their service performance *per se*.

Our search revealed 93 studies that analyse the determinants of organizational performance in the public sector. 31 of these were found in *Journal of Public Administration Research and Theory*, 28 in *Public Administration Review*, 9 in *Administration and Society*, 9 in *Public Administration* and 4 in *Public Management Review*. Other journals in which articles utilising *both* administrative and survey performance measures were found included: *International Public Management Journal*, *Policy & Politics* and *Journal of Policy Analysis and Management*. The articles are listed in table 1 which provides information on the unit of analysis and the sample size, whether the study used administrative or survey data, the dimensions of performance examined, and the stakeholders included.

[Position of table 1 about here]

The majority of the studies (57) drew upon only administrative data, and a smaller number survey data alone (24). Only a small number of studies (12) used both administrative and survey data to measure organizational performance. The studies we review were all undertaken in single countries, with most (56) being conducted in the USA (30 of which are carried out in Texas school districts), with the UK represented on fifteen occasions, Denmark in four articles, Israel in three, and Canada, Colombia, Korea, Norway and Sweden one each. The evidence covers a wide range of public services ranging from single purpose organizations, such as schools, to multipurpose organizations, such as local and state governments.

The studies that used administrative data covered the following dimensions of performance: effectiveness (e.g. measures of school test results, crime clearance rate); efficiency (e.g. cost per client served); equity (e.g. allocation of grant funding to residents in most need; student pass rates by ethnicity); output quality (e.g. pupil retention rates in schools; visual assessment of street cleanliness); output quantity (e.g. units of output); indexes (e.g. basket of performance measures and inspection reports).

These articles examined the following stakeholders' perceptions of performance: central government (including federal agencies and regional governments); national regulator (e.g. the Audit Commission in England); state and regional government (e.g. Welsh Assembly Government); local government (e.g. US county governments); independent experts (e.g. detailed inspection of service quality).

The studies using survey data typically asked respondents to assess achievements on the following dimensions of performance: effectiveness (e.g. parent's perception of their child's 'school readiness'); efficiency (e.g. how efficiently a local government is run); equity (e.g. whether the police treated people equally regardless of race); output quality (e.g. the quality of individual local services); responsiveness (e.g. ability to access services online);

satisfaction (e.g. satisfaction with treatment received); social, economic and environmental well-being (e.g. how well the organisation performs on this); and trust (e.g. confidence in an organization's ability to deliver a service). These survey-based studies covered the following stakeholders' perceptions of performance: citizens (i.e. adults residing within a given territory); clients (e.g. parents of schoolchildren); managers (e.g. senior and middle managers); employees (i.e. staff from all levels of the organization); front-line staff (i.e. staff responsible for service delivery). Four studies utilised indices combining multiple dimensions of performance, though none included all of the dimensions identified here.

We found only twelve studies that used both administrative and survey data as sources of information on organizational performance (see table 4 below). These studies included measures of effectiveness, efficiency, equity, output quality, output quantity, responsiveness, and satisfaction. The studies using administrative and survey data covered the perceptions of central government, a national regulator, state government, local government, independent experts, citizens and managers.

Overall, studies using administrative data have not included measures of trust, and survey based studies have not attempted to operationalise efficiency in a single measure. The index measures from the English local government dataset have included output quantity in the administrative index and efficiency in the survey index. Second, equity is more often operationalised in administrative datasets than in surveys. Questions about satisfaction are clearly based on the perceptions of citizens and can only be captured by surveys of the relevant constituents. Third, researchers using surveys are more likely to develop indices of performance than those using administrative data. Researchers using the latter have used index measures because they are the main performance indicator adopted by the key governmental stakeholder (Andrews et al. 2006) or to standardise measures of performance from different service areas (Andrews et al. 2008).

Some survey measures of performance are drawn from secondary sources; these have been used widely in the USA and involve the reanalysis of large scale federal government survey data (see for example Brewer 2005; Brewer and Selden 2000). Fourth, effectiveness is operationalised in a number of ways. Measures include both single item measures and indices. While a larger range of items are used to measure effectiveness in surveys, the administrative data are dominated by school pass rates, with over thirty studies (mostly of Texas School Districts) adopting this measure.

Having described the dimensions of performance and the stakeholders included in the empirical studies, we now move on to look at the relationship between them. Table 2 confirms that administrative data are drawn from governmental stakeholders, whereas survey data are drawn from a variety of groups. The only cells populated in the table are in the upper left and lower right quadrants. This indicates that the key stakeholders for administrative data are governments (and their agents in regulators) and independent experts, while the stakeholders for survey data are associated with citizens and clients and civil servants (variously defined). Studies with central government as the stakeholder include measures of effectiveness, efficiency, equity, and output quality and quantity, as do those for state/regional governments with the exclusion of output quantity. Studies where the main stakeholder is a national regulator typically use indexes to measure performance and also satisfaction. The number of studies taking their key stakeholder as local government or independent experts is very small.

[Position of table 2 about here]

Table 3 takes the analysis one stage further by examining the relationship between dimensions of performance and the aspects of organization and management that have been tested on them. Articles were coded from an inductive perspective and include only the main explanatory variable of interest (in cases where joint tests are conducted we include reference

to both variables). Fifteen areas of management and organization are identified in table 3. The table again suggests that particular research questions, or areas of organization and management, are associated with the use of administrative or survey data. Questions associated with bureaucracy, diversity, environment, HRM, leadership, market structure, networking, performance management, publicness, and structure have all been examined using administrative data only. By contrast, studies of goal ambiguity have utilised only survey data on performance. Four types of explanatory variable—citizen evaluation, management, red tape and strategic management—have been tested using both types of performance data. We now go on to investigate these studies in more detail.

[Position of table 3 about here]

Administrative or Survey Data: Does it matter?

Does the use of administrative or survey data make a difference to the evidence on the impact of management on organizational performance? A comprehensive answer to this question would be based on the findings of a large number of studies that have tested a wide range of management variables on administrative and survey measures covering all of the dimensions of performance that we discussed above. Ideally, the administrative and survey measures in each study would cover exactly the same dimensions of performance so that the relative effects of management variables on them could be compared directly. Unfortunately, such a set of empirical studies does not yet exist. Instead we must work with only 12 studies that have used both types of dependent variable; these allow us to provide a preliminary answer to the question ‘does it matter whether administrative or survey measures of performance are used?’

Before we turn to the results of these studies, it is important to note that although the evidence is limited in quantity it nevertheless has some strengths. First, the range of explanatory variables that have been tested on both administrative and survey measures of

performance includes important public management characteristics such as strategy, performance management, leadership, organizational structure, processes and red tape. Secondly, although the dimensions of performance tapped by the administrative and survey measures are not perfectly matched, they are often sufficiently close (e.g. where indexes based on large numbers of administrative measures are compared with consumer or management perceptions of overall performance) to provide a clear insight on whether the type of performance data makes any difference to the empirical evidence.

Table 4 lists the 12 studies that use both survey and administrative data, and provides information on the type of performance measure, the number of tests, and the percentage of results that support the hypothesis that management matters (or offer non-significant findings). In order to draw conclusions from this body of evidence we need to consider how the results of different studies should be combined and synthesized. The method that is used here is based on the percentage of statistical tests that support the hypothesis that the management variables under investigation are associated with organizational performance. To count as support for the hypothesis, two conditions must be satisfied. First, the result must be in the predicted direction. Second, the result must be statistically significant, that is, greater than would be likely to arise by chance alone (at the .05 significance level). If these criteria are applied to all of the tests in a single study, then a support score can be calculated. This is the number of tests that are consistent with the notion that particular management practices result in higher organizational performance as a percentage of all the tests that are reported in the study. The total number of tests conducted ranges from 2 to 40. The final step in this analytical procedure is to construct an aggregate support score across all the studies that have tested the impact of management on 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 1 or 200 tests. This unweighted mean has the advantage that

studies that conduct a large number of tests on the same data set are not given undue importance. Second, the support score for each study can be weighted (multiplied) by the number of tests in that study. In other words, equal weight is attached to each test rather than each study. This weighted mean has the advantage that studies that report only a small number of tests do not have a disproportionate influence on our analysis.

[Position of table 4 about here]

The “real” level of support for the impact of management on different types of performance measure probably lies somewhere between the unweighted and weighted figures. Although it is impossible to determine precisely where, we suspect that the weighted mean provides a more accurate picture because studies that report only a few tests tend to show higher than average percentages of significant results. Put simply, the published results may not be a representative sample of all the results that were obtained. Therefore, we emphasize the weighted support score in our evaluation of the evidence.

Table 5 presents the mean support scores for administrative and survey data. The support scores are also presented for the survey data by stakeholder group represented (citizens or customers and managers). The weighted support scores suggest that positive relationships between management variables and performance are not more likely to be uncovered by either administrative or survey data. While the number of studies we are working with is small, the weighted support score for surveys items is 44% (unweighted 54%) and 41% for administrative data (unweighted 43%). The studies that use both administrative and survey measures of performance are now reviewed, commencing with those from the stream of work on strategic management.

[Position of table 5 about here]

The studies that include strategic management, and more precisely strategy content as conceptualised by Miles and Snow (1978), examine this independent variable alongside

others in contingency models. The additional variables include representative bureaucracy, red tape, and sustainable management. Andrews et al. (2005) examine the proposition that the impact of representative bureaucracy is contingent on organizational strategy. They show that representative bureaucracy (operationalised as ‘the percentage of local authority employees from ethnic minority communities’) is negatively related to both consumer satisfaction and administrative performance, but that this effect is mitigated by an innovative and proactive strategy. The authors report similar results for administrative and survey measures of performance.

Walker and Brewer (2009) also examined the relationship between strategy content and a second variable, in this case red tape. This study also drew upon English local government as its unit of analysis. The study included index measures drawn from both administrative and survey data sets, and some comparability of the indexes was argued to be an important strength of the study. The index of survey performance is derived from a multiple informant survey of officers from the senior management team in the local authority and each service area. Officers were asked to rate the performance of their organization in comparison to others on measures of quality, efficiency, value for money, effectiveness, equity, customer satisfaction.

The study drew on the strategic management work of Miles and Snow (1978) and examined whether different strategic stances were able to offset the likely harmful effects that red tape can have on organizational performance in public agencies. Walker and Brewer (2009) concluded that prospecting can offset the detrimental impacts of red tape, but that in organizations with a reacting stance (where organizations await from instructions from the external environment) the presence of red tape worsens performance. The interaction between red tape and a strategy of focusing upon core business and efficiency neither assists nor

detracts from performance outcomes. These findings were consistent for the use of both the administrative and survey-derived measures of organizational performance.

Enticott and Walker (2008) undertake tests of the independent effects of the impact of prospecting, defending and reacting together with sustainable management (concerned with management behaviour in relation to social, economic and environmental practices). They examined the effect of these variables on the CSP score and a single item perceptual measure from managers who were asked about their organization's performance (as compared to other similar organizations) in terms of 'promoting the social, economic and environmental well-being of local people'. They find three statistically significant relationships. Using the administrative measure of performance they identify a positive relationship with prospecting, and marginally stronger results are thus recorded when the data are survey based.

Four of the articles (Andrews et al. 2009; Andrews and Boyne 2010; Walker and Boyne 2006; Walker et al. 2011) examine questions of management from a number of perspectives. Andrews et al. (2009) and Andrews and Boyne (2010) tackle questions of management capacity and its effect on performance. Both studies take the administrative measure of performance as the CSP, and the survey measure as the consumer satisfaction Best Value Performance Indicator. Andrews et al. (2009) argue and empirically demonstrate that prior experience of with dealing with immigration will provide the capacity to offset the negative impact of worker migration from the EU accession countries on local authority performance. The negative relationship is suggested because of the scale of the migration and the shock or jolt to the system it produced. Their hypothesis on the shock to the system is supported for both measures of performance, however the mitigating impact of prior experience only holds in the case of service performance (administrative data) and not for consumer satisfaction (survey data). Andrews and Boyne (2010) delve into the question of management capacity further exploring independent effects on performance and joint effects

with leadership. They report similar results for either measure of performance: management and leadership matter separately for performance and their interaction has positive effect.

Walker and Boyne (2006) and Walker et al. (2011) tackle questions about public management reforms. Walker and Boyne (2006) undertake an evaluation of the Blair government's strategy for public service improvement (rational planning, devolution and delegation, flexibility and incentives, and enhanced choice). They measure performance using the index developed by the regulator (the CSP) and survey items of managers' perceptions of performance that measure service outputs, efficiency, responsiveness and outcomes. They report more positive results for managers' assessments of performance. Walker et al. (2011) explore the capacity for the conceptual framework of 'market orientation' to explain public service performance as measured by three stakeholder groups. Market orientation consists of three aspects, customer orientation, competitor orientation and interfunctional coordination which is presented as the market or privateness extreme of the new public management reforms. They find the strongest effect of market orientation on consumer satisfaction, followed by managers' own assessment of consumer satisfaction, and lastly by the regulator of English local government (administrative data using the CSP metric).

The study by Brewer and Walker (2010) examines red tape. It again draws upon the index of performance developed by the regulator of English local government and managers' perceptions of performance. In this study Brewer and Walker (2010) disaggregate red tape and performance into their component parts and uncover evidence that is contrary to the popular notion that red tape is only harmful to organizational achievements. The distribution of the results between administrative and survey data does not suggest that one operationalization of performance is more likely to result in more positive results.

Two studies focus on citizen expectations. The Licari et al (2005) study draws upon expert assessments of park and street landscapes and contrasts this with citizen evaluations. They uncover very similar levels of rating between these two stakeholder groups. They do note some bias towards more positive assessments in more affluent areas, but conclude that citizen evaluations are not overly problematic. Folz's (2004) study of benchmarking schemes for public service improvement reaches similar conclusions. Folz (2004) argues that similarities in perspectives from different stakeholder groups can be built upon to ensure that services can be developed that meet the needs of all stakeholders.

The final two articles included in our review touch on substantive questions about management and performance and methodological concerns about the use of administrative and survey data. Meier and O'Toole (2010) tackle longstanding concerns in social science about common source bias—the question at the heart of this article. They contrast administrative scores on the TAKS and for college bound students with the perceptions of school district superintendents. Perceptions of performance were measured by asking 'compared to similar districts, my assessment of our performance is' using a five point Likert scale (excellent, above average, average, below average, inadequate) in relation to 'TAKS performance', 'college bound performance', and 'overall quality of education in the district'). Much of their analysis focuses upon relationships between performance indicators, however, there is one test where they use a measure of prospecting ('Our district is always among the first to adopt new ideas and practices'). Meier and O'Toole's findings are stark, and contrast with all the others reviewed here. They find that prospecting has a statistically insignificant relationship with administrative measures of performance but strong relationships with all three survey measures (their models controlled for a range of external constraints).

Andrews et al. (2010), like Meier and O'Toole (2010), similarly focus much of their attention on relationships between different measures of performance and different

stakeholders. Their bivariate analysis suggests that perceptions of managers are more positive on an administrative measure of service performance (the CSP as discussed above) and that managers believe that consumers are more satisfied than they themselves report. The study moves on to undertake a multivariate analysis with the dependent variable measured as the gap between the two measures of performance (e.g. difference between managers internal rating of performance and CSP) and independent measures of the organizational environment and management factors. Andrews et al. (2010) conclude that fashionable management practices are more likely to explain the performance gap for service performance than consumer satisfaction, thereby reporting stronger results for administrative than survey data.

In sum, the pattern of the evidence from studies that test the same management variables on both administrative and survey data is that the type of performance measure makes little difference to the statistical results. Management has much the same impact on different performance indicators, regardless of whether these reflect the priorities of governmental or other stakeholders. However, there is an extra and unexpected twist in the evidence. Within the survey data reported in the studies, two stakeholder groups are represented: citizens (or consumers) and managers. The weighted support scores here indicate that, if the dependent variable is managers' perceptions of organizational achievements, the studies are likely to offer only moderate support for the proposition that management variables are associated with higher levels of performance (weighted support score 38%, unweighted 46%). By contrast when the views of citizens and customers are used as dependent variables, the studies offer stronger support for the management-performance hypothesis (weighted 60%, unweighted 62%). Clearly, many more studies that compare the impact of management on performance as perceived by these two stakeholder groups is required before strong conclusions can be drawn. Nevertheless, this initial sift of the evidence suggests that public managers may underestimate the effects of their activities on

performance as perceived by citizens.

Conclusion

In this paper we have sought to identify the relative use of administrative and survey measures of performance in studies of the impact of public management on organizational achievements, and to evaluate whether the use of one type of measure rather than another makes any difference to the empirical evidence. Overall, there are twice as many studies of organizational performance drawing on administrative data as those drawing on survey data. Our review identified some interesting differences in the characteristics of the studies using one or the other type of data. Studies drawing on administrative data appear to be more likely to focus on effectiveness, efficiency and equity, while those using survey data seem to be more likely to be concerned with responsiveness, satisfaction and trust. This in turn seems to reflect the priorities of different stakeholder groups. The stakeholder perceptions on which the performance measures are based exhibit a stark divide between studies using administrative or survey data. Central, state and local governments, regulators and experts are all sources of administrative performance data but not survey data, while the reverse is true of citizens, clients, managers, employees and front-line staff.

Our review of the findings from the small set of studies that use both administrative and survey measures of performance suggests that reliance on one type of measure rather than the other does not make a great difference to the results that are obtained. Nevertheless, survey measures based on the perceptions of different stakeholders are associated with different findings: in particular, management variables appear to be more likely to influence the performance perceptions of citizens than the perceptions of managers themselves. This is a phenomenon that clearly warrants further theoretical and empirical investigation, because it

implies that public managers may not be the best judge of whether their efforts are worthwhile. Contrary, however, to assumptions that public managers are likely to make inflated claims about their accomplishments, our review of the existing evidence implies that they are not sufficiently aware of the strength of the link between their activities and citizens' perceptions of public service performance.

More broadly, our analysis reveals that much remains to be done to paint a more comprehensive picture of the determinants of performance in the public sector. Whether performance is measured with administrative or survey data, many dimensions of this basic concept have so far been operationalised in only a handful of studies, and the range of stakeholders taken into account is typically narrow and focuses either on government or on one or two other groups. In order to provide a more complete and accurate answer to the question posed in the title of this paper, future studies need to use both administrative and survey data to capture a wider range of dimensions of performance that reflect the views of a variety of stakeholders.

References

- Andersen, L. B. and Blegvad, M. (2006) Does Ownership Matter for the Delivery of Professionalized Public Services? Cost-efficiency and Effectiveness in Private and Public Dental Care for Children in Denmark. *Public Administration*, 84:1 pp147-64.
- Andersen, S. C. (2008) The Impact of Public Management Reforms on Student Performance in Danish Schools. *Public Administration*, 86:2 pp541-58.
- Andersen, S. C. and Serritzlew, S. (2007) The Unintended Effects of Private School Competition. *Journal of Public Administration Research and Theory*, 17:2 pp335-56.
- Andrews, R. and Boyne, G. A. (2010) Capacity, Leadership and Organizational Performance: Testing the Black Box Model of Public Management. *Public Administration Review*.
- Andrews, R., Boyne, G. A. and Enticott, G. (2006) Performance Failure in the Public Sector - Misfortune or Mismanagement? *Public Management Review* 8:2 pp273-96.
- Andrews, R., Boyne, G. A., Law, J. and Walker, R. M. (2005) External Constraints on Local Service Standards: The Case of Comprehensive Performance Assessment in English local government. *Public Administration*, 83:3 pp639-56.
- Andrews, R., Boyne, G. A., Law, J. and Walker, R. M. (2008) Organizational Strategy, External Regulation and Public Service Performance. *Public Administration*, 86:1 pp185-203.
- Andrews, R., Boyne, G. A., Law, J. and Walker, R. M. (2009a) Centralization, Organizational Strategy, and Public Service Performance. *Journal of Public Administration Research and Theory*, 19:1 pp57-80.
- Andrews, R., Boyne, G. A., Law, J. and Walker, R. M. (2009b) Strategy Formulation, Strategy Content and Performance. *Public Management Review*, 11:1 1-22.
- Andrews, R., Boyne, G. A., Meier, K. J., O'Toole, L. J. and Walker, R. M. (2005) Representative Bureaucracy, Organizational Strategy, and Public Service Performance: An Empirical Analysis of English Local Government. *Journal of Public Administration Research and Theory*, 15:4 pp489-504.
- Andrews, R., Boyne, G. A., Meier, K. J., O'Toole, L. J. and Walker, R. M. (2009) EU Accession and Public Service Performance. *Policy & Politics*, 37:1 pp19-37.
- Andrews, R., Boyne, G. A., Moon, J., and Walker, R. M. (2010) Assessing Organizational Performance: Exploring Differences Between Internal and External Measures. *International Public Management Journal*, 13:2.
- Andrews, R., Boyne, G. A. and Walker, R. M. (2006) Strategy Content and Organizational Performance: An Empirical Analysis. *Public Administration Review*, 66:1 pp52-63.
- Avellaneda, C. N. (2009) Municipal Performance: Does Mayoral Quality Matter? *Journal of Public Administration Research and Theory*, 19:2 pp285-312.
- Bilodeau, N., Laurin, C. and Vining, A. (2007) "Choice of Organizational Form makes a real Difference" The Impact of Corporatization on Government Agencies in Canada, *Journal of Public Administration Research and Theory*, 17:1 pp119-147.
- Bohte, J. (2001) School Bureaucracy and Student Performance at the Local Level. *Public Administration Review*, 61:1 pp92-99.
- Bohte, J. and Meier, K. J. (2000) Goal Displacement: Assessing the Motivation for Organizational Cheating. *Public Administration Review*, 60:2 pp173-81.

- Boschken, H. L. (2000) Behavior of Urban Public Authorities Operating in Competitive Markets - Policy Outcomes in Mass Transit. *Public Administration Review*, 31:6 pp726-58.
- Boyne, G.A. (2002) Concepts and Indictors of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money and Management* 22, no. 2, 17-24.
- Boyne, G. A. and Chen, A. (2007) Performance Targets and Public Service Improvement. *Journal of Public Administration Research and Theory*, 17:3 pp455-77.
- Boyne, G. A. and Hood C. (2010) Incentives: New Research on an Old Problem. *Journal of Public Administration Research and Theory* forthcoming.
- Boyne, G. A. and Meier, K. J. (2009) Environmental Turbulence, Organizational Stability, and Public Service Performance. *Administration & Society*, 40:8 pp799-824.
- Brewer, G. A. (2005) In the Eye of the Storm: Frontline Supervisors and Federal Agency Performance. *Journal of Public Administration Research and Theory*, 15:4 pp505-27.
- Brewer, G. A. and Selden, S. C. (2000) Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies. *Journal of Public Administration Research and Theory*, 10:4 pp685-711.
- Brewer, G. A. and Walker, R. M. (2010) The Impact of Red Tape on Governmental Performance: An Empirical Analysis. *Journal of Public Administration Research Theory*, 20:1 pp233-57.
- Brown, T. (2007) Coercion versus Choice: Citizen Evaluations of Public Service Quality across Methods of Consumption. *Public Administration Review*, 67:3 pp559-72.
- Chun, Y. H. and Rainey, H. G. (2005) Goal Ambiguity and Organizational Performance in US Federal Agencies. *Journal of Public Administration Research and Theory*, 15:4 pp529-57.
- Collins, B. K. and Gerber, B. J. (2008) Taken for Granted? Managing for Social Equity in Grant Programs. *Public Administration Review*, 68:6 pp1128-41.
- Fernandez, S. (2005) Developing and Testing an Integrative Framework of Public Sector Leadership: Evidence from the Public Education Arena. *Journal of Public Administration Research and Theory*, 15:2 pp197-217.
- Folz, D. H. (2004) Service Quality and Benchmarking the Performance of Municipal Services. *Public Administration Review*, 64:2 pp209-20.
- Fossett, J. and Thompson, F. J. (2006) Administrative Responsiveness to the Disadvantaged: The Case of Children's Health Insurance. *Journal of Public Administration Research and Theory*, 16:3 pp369-92.
- Garnett, J. L., Marlowe, J. and Pandey, S. K. (2008) Penetrating the Performance Predicament: Communication as a Mediator or Moderator of Organizational Culture's Impact on Public Organizational Performance. *Public Administration Review*, 68:2 pp266-81.
- Gill, J. and Meier, K. J. (2001) Ralph's Pretty-Good Grocery versus Ralph's Super Market: Separating Excellent Agencies from the Good Ones. *Public Administration Review*, 61:1 pp9-17.

- Goerdel, H. T. (2006) Taking Initiative: Proactive Management and Organizational Performance in Networked Environments. *Journal of Public Administration Research and Theory*, 16:3 pp351-67.
- Herman, R. D. and Renz, D. O. (2004) Doing Things Right: Effectiveness in Local Nonprofit Organizations, A Panel Study. *Public Administration Review*, 64:6 pp694-704.
- Hicklin, A., O'Toole, L. J. and Meier, K. J. (2008) Serpents in the Sand: Managerial Networking and Nonlinear Influences on Organizational Performance. *Journal of Public Administration Research and Theory*, 18:2 pp253-73.
- Hill, G. C. (2005) The Effects of Managerial Succession on Organizational Performance. *Journal of Public Administration Research and Theory*, 15:4 pp585-97.
- Isett, K. R., Morrissey, J. P. and Topping, S (2006) Systems Ideologies and Street-level Bureaucrats. *Public Administration Review*, 66:2 pp217-27.
- James, O. (2009) Evaluating the Expectations Disconfirmation and Expectations Anchoring Approaches to Citizen Satisfaction with Local Public Services. *Journal of Public Administration Research and Theory*, 19:1 pp107-23.
- Kelman, S. and Friedman, J. N. (2009) Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19:4 917-46.
- Kim, S. (2005) Individual-level Factors and Organizational Performance in Government Organizations. *Journal of Public Administration Research and Theory*, 15:2 pp245-61.
- Leland, S. and Smirnova, O. (2009) Reassessing Privatization Strategies 25 Years Later: Revisiting Perry and Babitsky's Comparative Performance Study of Urban Bus Transit Services. *Public Administration Review*, 69:5 pp855-67.
- Licari, M. J., McLean, W. and Rice, T. W. (2005) The Condition of Community Streets and Parks: A Comparison of Resident and Nonresident Evaluations. *Public Administration Review*, 65: 3 pp360-8.
- May, P. J. and Winter, S. C. (2007) Collaborative Service Arrangements - Patterns, Bases, and Perceived Consequences. *Public Management Review*, 9:4 pp479-502.
- Meier, K. J. and Bohte, J. (2000) Ode to Luther Gulick - Span of Control and Organizational Performance. *Administration & Society*, 32:2 pp115-37.
- Meier, K. J. and Bohte, J. (2003) Not with a Bang, but a Whimper - Explaining Organizational Failures. *Administration & Society*, 35:1 pp104-21.
- Meier, K. J. and Hicklin, A. (2008) Employee Turnover and Organizational Performance: Testing a Hypothesis from Classical Public Administration. *Journal of Public Administration Research and Theory*, 18:4 pp573-90.
- Meier, K. J., Mastracci, S. H. and Wilson, K. (2006) Gender and Emotional Labor in Public Organizations: An Empirical Examination of the Link to Performance. *Public Administration Review*, 66:6 pp899-909.
- Meier, K. J. and O'Toole, L. J. (2001) Managerial Strategies and Behavior in Networks: A Model with Evidence from US Public Education. *Journal of Public Administration Research and Theory*, 11:3 pp271-93.
- Meier, K. J. and O'Toole, L. J. (2003) Public Management and Educational Performance: The Impact of Managerial Networking. *Public Administration Review*, 63:6 pp689-99.

- Meier, K. J. and O'Toole, L. J. (2006) Political Control versus Bureaucratic Values: Reframing the Debate. *Public Administration Review*, 66:2 pp177-92.
- Meier, K. J. and O'Toole, L. J. (2008) Management Theory and Occam's Razor - How Public Organizations Buffer the Environment. *Administration & Society*, 39:8 pp931-58.
- Meier, K. J. and O'Toole, L. J. (2009) The Dog that didn't bark: How Public managers handle Environmental Shocks. *Public Administration*, 87:3 pp485-502.
- Meier, K. J. and O'Toole, L. J. (2010) I think (I am doing well), therefore I am: Assessing the Validity of Administrators' Self-Assessments of Performance. Paper presented at the Midwest Political Science Association Conference, Chicago April 22-25.
- Meier, K. J. and O'Toole, L. J. Boyne, G. A. and Walker, R. M. (2007) Strategic Management and the Performance of Public Organizations: Testing Venerable Ideas against Recent Theories *Journal of Public Administration Research and Theory*, 17:3 pp357-77.
- Meier, K. J. and O'Toole, L. J. and Goerdel, H. T. (2006) Management Activity and Program Performance: Gender as Management Capital. *Public Administration Review*, 66:1 pp24-36.
- Moynihan, D. A. and Pandey, S. K. (2005) Testing how Management Matters in an era of Government by Performance Management. *Journal of Public Administration Research and Theory*, 15:3 pp421-39.
- Nicholson-Crotty, S. and O'Toole, L. J. (2004) Public Management and Organizational Performance: The Case of Law Enforcement Agencies. *Journal of Public Administration Research and Theory*, 14:1 pp1-18.
- Nicholson-Crotty, S., Theobald, N. A. and Nicholson-Crotty, J. (2006) Disparate Measures? Public Managers and Performance Measurement Strategies. *Public Administration Review*, 66:1 pp101-13.
- Orr, M. and West, D. M. (2007) Citizen Evaluations of Local Police - Personal Experience or Symbolic Attitudes. *Administration & Society*, 38:6 pp649-68.
- O'Toole, L. J. and Meier, K. J. (2003) Plus ca Change: Public Management, Personnel Stability, and Organizational Performance. *Journal of Public Administration Research and Theory*, 13:1 43-64.
- O'Toole, L. J. and Meier, K. J. (2004a) Parkinson's Law and the New Public Management? Contracting Determinants and Service-quality Consequences in Public Education. *Public Administration Review*, 64:3 pp342-52.
- O'Toole, L. J. and Meier, K. J. (2004b) Public Management in Intergovernmental Networks: Matching Structural Networks and Managerial Networking. *Journal of Public Administration Research and Theory*, 14:4 pp469-94.
- O'Toole, L. J. and Meier, K. J. (2004c) Desperately Seeking Selznick: Cooptation and the Dark Side of Public Management in Networks. *Public Administration Review*, 64:6 pp681-93.
- O'Toole, L. J., Meier, K. J. and Nicholson-Crotty, S. (2005) Managing Upward, Downward and Outward - Networks, Hierarchical Relationships and Performance. *Public Management Review*, 7:1 pp45-68.
- Pitts, D. W. (2005) Diversity, Representation, and Performance: Evidence about Race and Ethnicity in Public Organizations. *Journal of Public Administration Research and Theory*, 15:4 pp615-31.

- Pitts, D. W. (2007) Representative Bureaucracy, Ethnicity and Public Schools - Examining the Link between Representation and Performance. *Administration & Society*, 39:4 pp497-526.
- Pitts, D. W. and Jarry, E. M. (2009) Getting to know you: Ethnic Diversity, Time and Performance in Public Organizations. *Public Administration*, 87:3 pp503-18.
- Provan, K. G., Huang, K. and Milward, H.B. (2009) The Evolution of Structural Embeddedness and Organizational Social Outcomes in a Centrally Governed Health and Human Services Network. *Journal of Public Administration Research and Theory*, 19:4 pp873-93.
- Selden, S. C. and Sowa, J. E. (2004) Testing a Multi-dimensional Model of Organizational Performance: Prospects and Problems. *Journal of Public Administration Research and Theory*, 14:3 pp395-416.
- Shingler, J., Van Loon, M. E., Alter, T. R., et al. (2008) The Importance of Subjective Data for Public Agency Performance Evaluation. *Public Administration Review*, 68:6 pp1101-11.
- Smith, K. B. and Larimer, C. W. (2004) A Mixed Relationship: Bureaucracy and School Performance. *Public Administration Review*, 64:6 pp728-36.
- Sorensen, R. J. (2007) Does Dispersed Public Ownership Impair Efficiency? The Case of Refuse Collection in Norway. *Public Administration*, 85:4 1045-58.
- Tankersley, W. B. (2000) The Impact of External Control Arrangements on Organizational Performance. *Administration & Society*, 32:3 pp282-304.
- Tolbert, C. J. and Mossberger, K. (2006) The Effects of E-government on Trust and Confidence in Government. *Public Administration Review*, 66:3 pp354-69.
- Ugboro, I. O. and Obeng, K. (2009) Board Activities, Involvement, and Public Transit Performance. *Administration & Society*, 41:2 pp235-57.
- Van Ryzin, G. G. (2006) Testing the Expectancy Disconfirmation Model of Citizen Satisfaction with Local Government. *Journal of Public Administration Research and Theory*, 16:4 pp599-611.
- Van Ryzin, G. G. and Immerwahr, S. (2007) Importance-performance Analysis of Citizen Satisfaction Surveys. *Public Administration*, 85:1 pp215-26.
- Van Ryzin, G. G., Immerwahr, S. and Altman, S. (2008) Measuring Street Cleanliness: A Comparison of New York City's Scorecard and Results from a Citizen Survey. *Public Administration Review*, 68:2 pp295-303.
- Van Ryzin, G. G., Muzzio, D., Immerwahr, S., et al. (2004) Drivers and Consequences of Citizen Satisfaction: An Application of the American Customer Satisfaction Index Model to New York City. *Public Administration Review*, 64:3 pp331-41.
- Vigoda, E. (2000) Are you being Served? The Responsiveness of Public Administration to Citizens' Demands: An Empirical Examination in Israel. *Public Administration*, 78:1 pp165-91.
- Vigoda-Gadot, E. (2007) Citizens' Perceptions of Politics and Ethics in Public Administration: A Five-year National Study of their Relationship to Satisfaction with Services, Trust in Governance, and Voice Orientations. *Journal of Public Administration Research and Theory*, 17:2 pp285-305.

- Vigoda-Gadot, E. and Mizrahi, S. (2008) Public Sector Management and the Democratic Ethos: A 5-year Study of Key Relationships in Israel. *Journal of Public Administration Research and Theory*, 18:1 pp79-107.
- Walker, R. M and Boyne, G. A. (2006) Public Management Reform and Organizational Performance: An Empirical Assessment of the UK Labour Government's Public Service Improvement Strategy. *Journal of Policy Analysis and Management*, 25:4 pp371-93.
- Walker, R. M. and Brewer, G. A. (2009) Can Management Strategy Minimize the Impact of Red Tape on Organizational Performance? *Administration & Society*, 41:4 pp423-48.
- Walker, R. M. and Enticott, G. (2008) Sustainability, Performance and Organizational Strategy: An Empirical Analysis of Public Organizations. *Business Strategy and the Environment*, 17:1 pp79-92.
- Walker, R., Boyne, G.A. and Brewer, G., eds., (2010) *Public Management and Performance: Research Directions* Cambridge University Press.
- Walker, R.M., Brewer, G., Boyne, G.A., and Avellenda, C. (2011) Market Orientation and Public Service Performance: New Public Management Gone Mad? *Public Administration Review* forthcoming.
- Zimmer, R. and Buddin, R. (2009) Is Charter School Competition in California Improving the Performance of Traditional Public Schools? *Public Administration Review*, 69:5 pp831-45.

Table 1: Studies of organizational performance in the public sector

Study	Organizations and sample size	Type of measure	Performance dimension	Stakeholder
Andersen and Biegvad, 2006	275 Danish dental care centres	Administrative	Effectiveness	Central government
Andersen, 2008	778 Danish schools	Administrative	Efficiency	Central government
Andersen and Serritzlew, 2007	1348 Danish schools	Administrative	Effectiveness	Central government
Andrews and Boyne, 2010	88 English local governments	Administrative Survey	Index	Regulator
Andrews et al, 2005a	148 English local governments	Administrative	Satisfaction	Regulator
Andrews et al, 2005b	80 English local governments	Administrative Survey	Index	Regulator
Andrews, Boyne and Enticott, 2006	120 English local governments	Administrative	Satisfaction	Regulator
Andrews, Boyne and Walker, 2006	119 English local governments	Administrative	Index	Regulator
Andrews et al, 2008	51 Welsh local government services	Administrative	Index	Regional government
Andrews et al, 2009a	47 Welsh local government services	Administrative	Index	Regional government
Andrews et al, 2009b	53 Welsh local government services	Administrative	Index	Regional government
Andrews et al, 2009	148 English local governments	Administrative Survey	Index	Regulator
Andrews et al. 2010	58 English local governments	Administrative Survey	Satisfaction	Regulator
Avellaneda, 2009	40 municipalities in one Colombian state	Administrative	Index	Regulator
Bilodeau, Laurin and Vining, 2007	11 Canadian federal agencies	Administrative	Output quality	Central government
Bohte, 2001	350 Texas school districts	Administrative	Efficiency	Central government
Bohte and Meier, 2000	476 Texas school districts	Administrative	Output quantity	State government
Boschken, 2000	42 US Public Transit Agencies	Administrative	Effectiveness	State government
Boyne and Chen, 2007	147 English local governments	Administrative	Efficiency	Central government
Boyne and Meier, 2009	1,000 Texas school districts	Administrative	Effectiveness	Central government
Brewer, 2005	US federal agencies	Survey – 2,719 federal employees	Equity	State government
			Effectiveness	Employees
			Output quality	
			Output quantity	

Brewer and Selden, 2000	US federal agencies	Survey – 9,710 federal employees	Index of: Effectiveness Output quality Output quantity	Employees
Brewer and Walker 2010 Brown, 2007	135 English local governments Columbus, Ohio	Administrative Survey Survey – 1,188 residents	Index Index and 8 PIs Output quality	Regulator Manager Citizens
Chun and Rainey, 2005	32 US Federal Agencies	Survey – 25,184 employees at all levels	Effectiveness Output quality Responsiveness	Employees
Collins and Gerber, 2008	390 Nonfrontier counties in California, Kentucky, Texas and Utah	Administrative	Equity	Local government
Fernandez, 2005	1,000 Texas school districts	Administrative	Effectiveness	State government
Folz, 2004	1,012 US cities	Administrative Survey – waste coordinators	Efficiency Output quality	State government Managers
Fossett and Thompson, 2006	17 US State governments	Administrative	Responsiveness	Experts
Garnett, Marlowe and Pandey, 2008	50 US state primary health and human service agencies	Survey – 274 information managers	Effectiveness	Managers
Gill and Meier, 2001	534 Texas school districts	Administrative	Effectiveness	State government
Goerdel, 2006	507 Texas school districts	Administrative	Effectiveness Equity Output quality	State government Managers
Herman and Renz, 2004	Health and welfare charities in one large US metropolitan area	Survey – 162 managers in year 1, 199 managers in year 2	Effectiveness Responsiveness	
Hicklin, O'Toole and Meier, 2008	1000 Texas school districts	Administrative	Effectiveness	State government
Hill, 2005	1000 Texas school districts	Administrative	Effectiveness	State government
Isett, Morrissey and Topping, 2006	13 mental healthcare providers in Virginia	Survey – 198 front-line staff	Output quality	Front-line staff
James, 2009	386 English local governments	Survey – 4067 adults	Satisfaction	Citizens
Kelman and Friedman, 2009	155 English hospital trusts	Administrative	Effectiveness	Central government

Kim, 2005	9 central government agencies, 5 provincial governments, 26 local governments in Korea	Survey – 1,739 employees	Index of: Effectiveness Efficiency Equity Output quality Satisfaction	Employees
Leland and Smirnova, 2009	545 Urban bus transit agencies	Administrative	Efficiency Output quality	Central government
Licari, McLean and Rice, 2005	99 Iowa towns	Administrative Survey – 10,798 adults	Output quality Output quality	Experts Citizens
May and Winter, 2007	159 Danish municipal employment services	Survey – 159 middle managers	Effectiveness	Managers
Meier and Bohte, 2000	678 Texas school districts	Administrative	Effectiveness	State government
Meier and Bohte, 2003	1000 Texas school districts	Administrative	Output quality	State government
Meier et al., 2007	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Meier and Hicklin, 2008	1000 Texas school districts	Administrative	Effectiveness	State government
Meier, Mastracci and Wilson, 2006	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Meier and Nicholson-Crotty, 2006	60 US metropolitan counties	Administrative	Effectiveness	Central government
Meier and O'Toole, 2001	507 Texas school districts	Administrative	Effectiveness	State government
Meier and O'Toole, 2003	507 Texas school districts	Administrative	Effectiveness	State government
Meier and O'Toole, 2006	1043 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Meier and O'Toole, 2008	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Meier and O'Toole, 2009	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Meier and O'Toole 2010	628 Texas school districts	Administrative Survey	Effectiveness Effectiveness Quality	State government
Meier, O'Toole and Goerdel, 2006	1000 Texas school districts	Administrative	Effectiveness	State government
Moynihan and Pandey, 2005	50 US state primary health and human service agencies	Survey – 274 information managers	Effectiveness	Managers

Nicholson-Crotty and O'Toole, 2004	544 US police departments	Administrative	Effectiveness	Central government
Nicholson-Crotty, Theobald and Nicholson-Crotty, 2006	711 Texas school districts	Administrative	Output quality	State government
O'Toole and Meier, 2003	507 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
O'Toole and Meier, 2004a	1000 Texas school districts	Administrative	Effectiveness	State government
O'Toole and Meier, 2004b	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
O'Toole and Meier, 2004c	507 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
O'Toole, Meier and Nicholson-Crotty, 2005	1000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Orr and West, 2007	Providence, US	Survey – 509 adults	Effectiveness Equity	Citizens
Pitts, 2005	1,000 Texas school districts	Administrative	Effectiveness Output quality	State government
Pitts, 2007	1,000 Texas school districts	Administrative	Effectiveness Equity Output quality	State government
Pitts and Jarry, 2009	1,000 Texas school districts	Administrative	Effectiveness	State government
Selden and Sowa, 2004	22 human service organizations in New York and Virginia	Survey – 319 clients	Responsiveness Satisfaction	Clients
Shingler et al. 2008	Utility Regulatory Agency, US	Survey – 892 clients	Effectiveness	Clients
Smith and Larimer, 2004	350 multi-racial Texas school districts	Administrative	Effectiveness Output quality	State government
Sorensen, 2007	211 Norwegian municipalities	Administrative	Efficiency	Local government
Tankersley, 2000	1,211 US Electricity distributors	Administrative	Efficiency	Central government
Tolbert and Mossberger, 2006	815 US residents	Survey – 815 adults	Responsiveness Trust	Citizens
Ugboro and Obeng, 2009	110 US public transit systems	Administrative	Efficiency	Central government
Van Ryzin, 2006	New York City	Survey – 615 adults	Effectiveness Output quality Satisfaction	Citizens
Van Ryzin and Immerwahr, 2007	1081 US residents	Survey – 651 adults	Effectiveness	Citizens
Van Ryzin et al., 2004	New York City	Survey – 1500 citizens	Effectiveness Output quality	Citizens

Van Ryzin, Immerwahr and Altman, 2008	New York City	Survey – 3696 citizens	Satisfaction Effectiveness	Citizens
Vigoda, 2000	One Israeli city	Survey – 281 adults	Responsiveness Satisfaction	Citizens
Vigoda-Gadot, 2007	All Israeli public services	Survey – 2281 adults	Satisfaction	Citizens
Vigoda-Gadot and Mizrahi, 2008	All Israeli public services	Survey – 2281 adults	Responsiveness Trust	Citizens
Walker and Boyne, 2006	119 English local governments	Administrative Survey	Index Output quantity and efficiency Responsiveness Outcomes	Regulator Managers
Walker and Brewer, 2009	135 English local government	Administrative Survey – 1,056 top and middle managers	Index Index	Regulator Managers
Walker and Enticott, 2008	75 English local governments	Administrative Survey	Index Well-being	Regulator Manager
Walker et al. 2011	182 English local governments	Administrative Survey	Index Satisfaction	Regulator Citizens Managers
Zimmer and Buddin, 2009	184 Californian schools	Administrative	Effectiveness	State government

Table 2: Stakeholders and dimensions of performance

		Central government	National regulator	State/regional government	Local government	Independent experts	Citizens	Clients	Managers	Employees	Front-line staff
<i>Administrative data</i>											
	Effectiveness	45	-	42	-	-	-	-	-	-	-
	Efficiency	30	-	2	50	-	-	-	-	-	-
	Equity	10	-	20	50	-	-	-	-	-	-
	Output quality	10	-	30	-	50	-	-	-	-	-
	Output quantity	5	-	-	-	-	-	-	-	-	-
	Responsiveness	-	-	-	-	50	-	-	-	-	-
	Satisfaction	-	-	-	-	-	-	-	-	-	-
	Index	-	78	5	-	-	-	-	-	-	-
<i>Survey data</i>											
	Effectiveness	-	-	-	-	-	25	100	30	17	-
	Efficiency	-	-	-	-	-	-	-	9	-	-
	Equity	-	-	-	-	-	5	-	9	-	-
	Output quality	-	-	-	-	-	20	-	4	17	100
	Output quantity	-	-	-	-	-	-	-	-	-	-
	Responsiveness	-	-	-	-	-	15	-	9	17	-
	Satisfaction	-	22	-	-	-	25	-	13	-	-
	Social, economic and environmental well-being	-	-	-	-	-	-	-	9	-	-
	Trust	-	-	-	-	-	10	-	4	-	-
	Value for money	-	-	-	-	-	-	-	-	-	-
	Index	-	-	-	-	-	-	-	13	50	-
TOTAL	%	100	100	100	100	100	100	100	100	100	100
	n	(20)	(19)	(60)	(2)	(2)	(20)	(2)	(23)	(7)	(1)

Table 3: Measure of organization and management and dimensions of performance

		Bureaucracy	Citizen expectation	Diversity	Environment	Goal ambiguity	HRM	Leadership	Management	Market structure	Networking	Performance management	Publicness	Red Tape	Strategic Management	Structure
<i>Administrative data</i>																
	Effectiveness	44	-	51	40	-	50	75	8	27	54	50	22	-	14	66
	Efficiency	-	-	-	10	-	-	-	-	9	-	-	66	-	-	-
	Equity	11	-	17	20	-	25	-	8	9	23	17	-	-	2	-
	Output quality	22	5	34	20	-	25	-	12	-	23	17	11	-	2	-
	Output quantity	-	-	-	-	-	-	-	-	-	-	17	-	-	-	-
	Responsiveness	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Satisfaction	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Index	11	-	-	10	-	-	-	8	-	-	-	-	9	50	33
<i>Survey data</i>																
	Effectiveness	-	30	-	-	33	-	-	4	-	-	-	-	9	7	-
	Efficiency	-	-	-	-	-	-	-	4	-	-	-	-	9	-	-
	Equity	-	5	-	-	33	-	-	-	-	-	-	-	9	-	-
	Output quality	-	20	-	-	-	-	-	8	-	-	-	-	9	-	-
	Output quantity	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Responsiveness	-	10	-	-	33	-	-	12	-	-	-	-	9	-	-
	Satisfaction	11	25	-	-	-	-	25	12	-	-	-	-	18	7	-
	SEE	-	-	-	-	-	-	-	-	-	-	-	-	9	-	-
	Trust	-	5	-	-	-	-	-	4	-	-	-	-	-	-	-
	Value for money	-	-	-	-	-	-	-	-	-	-	-	-	9	-	-
	Index	-	-	-	-	-	-	-	8	-	-	-	-	9	7	-
TOTAL	%	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	n	(9)	(20)	(6)	(10)	(3)	(12)	(3)	(27)	(9)	(13)	(6)	(11)	(11)	(14)	(3)

Table 4: Results for management effects in studies containing both administrative and survey measures of performance

Study	Dimension of performance	N of test	Percentage of tests		
			+	NS	-
Folz 2004	Output quality – survey	3	33	66	0
	Output quality – administrative	2	50	50	0
Andrews et al 2005	Consumer satisfaction ^a	6	50	50	0
	Index ^b	3	66	33	0
Licari, McLean and Rice, 2005	Citizen evaluation	4	100	0	0
	Expert scores	2	100	0	0
Walker and Boyne 2006	Managers perceptions ^c	27	63	37	0
	Index ^b	9	33	66	0
Enticott and Walker 2008	Managers perceptions of social economic and environmental well-being	4	25	75	0
	Index ^b	4	25	75	0
Andrews et al. 2009	Consumer satisfaction ^a	3	100	0	0
	Index ^b	3	33	66	0
Walker and Brewer 2009	Index of managers perceptions ^d	8	50	50	0
	Index ^b	8	62	38	0
Andrews and Boyne, 2010	Consumer satisfaction ^a	6	66	33	0
	Index ^b	6	66	33	0
Andrews et al. 2010	Consumer satisfaction ^a	4	25	75	0
	Index ^b	4	75	25	0
Brewer and Walker 2010	Index of managers perceptions ^d	5	40	40	20
	Single items of managers perceptions ^e	40	18	68	15
	Index ^b	5	20	60	20
Meier and O'Toole 2010	Effectiveness ^f	2	0	100	0
	Managers perception of effectiveness ^g	3	100	0	0
Walker et al. 2011	Consumer satisfaction ^a	7	72	14	14
	Managers perceptions of consumer satisfaction	7	28	72	0
	Index ^b	7	14	86	0

Notes:

a 'percentage of citizens satisfied with the overall service provided by their authority'

b Core service performance, an organizational-wide measure of performance that includes performance indicators, the results of inspection visits and assessment of formal plans.

c Measures were outputs and efficiency, responsiveness, outcomes.

d Single items of those listed in ^c.

e Index of: output quality, efficiency, value for money, effectiveness, equity, customer satisfaction, staff satisfaction, social economic and environmental well-being.

f Pass rate on the Texas Assessment of Skills and Knowledge

g Perceptual assessment of performance on the Texas Assessment of Skills and Knowledge

Table 5: Summary of weighted support scores for administrative and survey data and performance

	+	NS	-
Administrative	41%	57%	2%
Survey	44%	45%	8%
<i>Survey stakeholders</i>			
Customer	60%	28%	11%
Manager	38%	52%	7%