Evan Berman
XiaoHu Wang
University of Central Florida

Performance Measurement in U.S. Counties: Capacity for Reform

This study examines the capacity of U.S. counties to undertake performance measurement. Based on a national survey of counties with populations over 50,000, the authors address the following questions: To what extent do counties implement performance measurement? Which capacities must be present for different levels of implementation and success? What can counties do to increase their capacity for performance measurement? And, what is the effect of county structure and functions on the use of performance measurement? This study finds that the success of performance measurement is greatly affected by counties' underlying organizational capacities.

In recent years, there has been great interest in using performance measurement for increasing accountability and improving performance (Walters 1998; Holzer 1998; Ammons 1996; Greiner 1996; Keehley et al. 1997; Harris 1995). Despite efforts in many jurisdictions some observers are lowering their expectations for this management reform. Myriad challenges have been identified, such as uncertain stakeholder support and inadequate technical ability to collect and analyze performance data (Radin 1998; Theurer 1998)—problems that have plagued previous management reforms (Berman 1998; Carroll 1995; Brown, Hitchcock, and Willard 1994; Berry, Berry, and Foster 1998). This article discusses how successful implementation of performance measurement requires careful attention to the management of underlying organizational capacities for achievement.

Based on a national survey of U.S. counties with populations over 50,000, this study examines the following questions: To what extent do counties have the capacity to implement performance measurement? Which capacities must be present for different levels of implementation and success? What can counties do to increase their capacity for performance measurement? To what extent do they undertake capacity-enhancing efforts? And, what is the effect of county structure and functions on efforts to increase capacity for performance measurement?

There are several reasons to study performance measurement at the county level. First, performance measurement is a means of providing accountability to county residents, who are often more informed about municipal than county affairs. In addition, it provides accountability to higher governments: counties receive far more funding from states and the federal government than do cities (U.S. Census 1997). Second, counties have historically been referred to as the "dark continent of American politics" and, therefore, have been understudied in public administration (Svara 1993; Menzel 1996; Streib and Waugh 1991). Since the 1980s, many counties have increased their abilities, leadership roles, and functions, yet we know very little about counties in this regard. Performance measurement, including the level of technical capacity (for example, the ability to analyze data and monitor goals) reflects on the professionalism of managers. Third, counties rely to a far greater extent than cities on commission forms of government; therefore, they are useful for studying the impact of county structures and the roles of elected officials on securing necessary support for management reform (Cigler 1995; ICMA 1998; Lewis and Taylor 1994). Support from

Evan Berman *is an associate professor in the Department of Public Administration, University of Central Florida (Orlando). His research interests are productivity, ethics, and local government. He has authored numerous articles that have appeared in* Public Administration Review, Administration and Society, American Review of Public Administration, *and other leading journals. His books include* Productivity in Public and Nonprofit Organizations *(Sage, March 1998) and* Public Sector Performance *(Westview, 1999) in the ASPA Classics Series. Email: bermane@pegasus.cc.ucf.edu*

XiaoHu Wang *is an assistant professor in the Department of Public Administration, University of Central Florida (Orlando). His research and teaching interests are budgeting and management. He received his doctorate from Florida International University in 1997. He has published in* Public Budgeting and Financial Management. *Email: xwang@mail.ucf.edu*

elected officials is often important in innovation, and it is especially crucial to performance measurement because it is, in part, undertaken to provide elected officials with improved information.

Framework

In the past decade, managers have learned a great deal about implementing management reforms. Slowly, attention has shifted from the specifics of new management innovation strategies (what is performance measurement?) to strategies for implementing change (what steps do we need to take?). Today, however, more study is needed to better understand the conditions for implementing change (do we have the conditions that are required for success?) and to manage them (do we know what these conditions are?). In this context, the term *capacity* refers to organizations' ability to achieve their aims (Honadle 1980). Ensuring adequate stakeholder support and technical abilities have become key capacities for implementing management reforms.

In many ways, understanding the importance of managing stakeholder support and technical ability comes from lessons learned the hard way. Budget reforms such as the Planning-Programming-Budgeting System and Zero-Based Budgeting in the 1960s and 1970s brought such worthy concepts as goals into the budget lexicon; but were criticized as technically over-reaching (because of inadequate data collection capabilities) and were inadequately supported by political officials, who viewed these reforms as a threat to their power (Joyce 1993a; Gianakis and Stone 1997; Mikesell 1995). Information from program evaluation efforts in the 1970s was judged as being too costly, complex, and untimely.[1] Although "second generation" evaluation in the 1970s and performance measurement in the early 1980s focused on short-term analysis and generated easy-to-understand data for decision makers, these efforts were not always part of ongoing decision-making processes among stakeholders (Wholey and Hatry 1992). They needed to be institutionalized as part of budget or strategic-planning processes. Lack of integration and user involvement caused underutilization.

In the case of performance measurement, *technical (or infrastructure) capacity* refers to counties' ability to develop performance goals and measures and to overcome such conceptual barriers as distinguishing outcomes from outputs. A 1997 General Accounting Office study reported difficulty in relating long-term goals to annual performance objectives; others have also noted that the relationship between performance goals and service efforts and activities is sometimes unclear and that it changes over time (Joyce 1993b; Mascarenhas 1996; Salzer, Nixon et al. 1997; Hakes 1996). In this regard, many authors urge jurisdictions to develop measures that are practical, as well as valid, reliable, and easy to understand (Hatry et al. 1992; Ammons 1995). Often, these measures are based on operations, and data are more readily collected where information technology is widely used (Stokes and Monaco 1997; Grizzle 1982). Outcome measures often require citizen and client surveys, but many jurisdictions do not have the capacity to gather these data in scientifically valid ways: contracting for such information may be costly and untimely. Leithe (1997) also notes the frequent absence of cost-accounting systems, which are necessary for identifying unit costs. The literature discussing specific technical competencies for performance measurement is surprisingly limited, although this aspect is mentioned in older literature (Pressman and Wildsavsky 1973; Sabatier and Mazmanian 1979).

Stakeholder (or political) capacity concerns the nature of support for performance measurement (Jones and McCaffery 1997; Cope 1997). There is widespread agreement that management reforms require support from top management. Support from elected officials is critical because it forecloses back channels, legitimates reforms and new performance expectations, and helps ensure funding for new efforts. In this regard, some states, such as Minnesota, mandate the use of performance measurement by counties, often as part of county budgeting and strategic planning (Walters 1998; Center for Accountability and Performance 1998; Tigue and Strachota 1994).[2] The lack of support, especially from legislatures, is a commonly cited reason for the budget-reform failure (Mikesell 1995; Rubin 1997). Elected officials are reluctant to give support when they view it as a ploy by bureaucrats to "technicalize" their operations and thereby avoid legislative scrutiny (Kettl 1994). Some authors believe that a major difference between previous managerial reforms and recent performance measurement efforts is that the latter has support from legislatures (Broom 1995; Kettl 1994; Melkers and Willoughby 1998). Counties also require support from lower managers, because it is well known that lower managers can sabotage reform through foot-dragging and end-runs involving citizen advocates and elected officials (Gerhart 1995; Romzek 1998). Support from advisory boards increases support for management reforms, because it increases the legitimacy of the actions of elected and appointed officials.

When counties satisfy many of the above conditions, they can be said to have a high level of capacity for performance measurement. This study suggests that high capacity is a prerequisite for widespread use of performance measurement in jurisdictions. Widespread use and institutionalization cannot occur when, for example, lower managers cannot acquire data in a timely manner or when the effort does not have funding support from elected officials.

However, initial or pilot efforts do *not* require high capacity. Indeed, experiential approaches to management reform often use small-scale efforts on a trial-and-error basis to identify areas where capacity is lacking. While these efforts may have a limited initial impact, top managers often measure their success by broadened support, providing the basis for future improvement.

Which factors affect counties' ability to achieve high capacity? It is often hypothesized that management reforms are championed by public managers (who increase their contribution and visibility in the organization), and therefore they are associated with reformed, council-administrator forms of government. Paradoxically, unless elected officials are given control over important performance-measurement decisions (such as deciding which measures should be used), management reform may suffer inadequate legitimacy from these stakeholders. It may be perceived as an effort to sidestep traditional methods of accountability. By contrast, commission forms of government involve elected officials to a far greater extent in department operations; therefore, officials are more likely to be involved in decisions about performance measurement, but they may lack the ability to implement performance measurement. Thus, the net effect of the form of government on the use of performance measurement is ambiguous: a greater desire for performance measurement but inadequate political capacity in council-administrator counties, versus greater political capacity but inadequate technical capacity in commission forms of government.

Obviously, other factors affect counties' ability to use performance measurement. For example, the availability of resources for testing new ideas matters, because, like most new efforts, performance measurement requires some additional resources. Positive employee and managerial attitudes also matter, at least as perceived by top managers who are responsible for implementing performance measurement (Grifel 1994). When managers and employees resist change, top managers face additional hurdles of dealing with the sources of that resistance. Professionalism may improve the capacity for performance measurement because of heightened concern for accountability and performance. In addition, an important step toward increasing capacity is making an accurate assessment. Some diagnostic strategies might involve asking managers about their current understanding of performance measurement, or assessing the capabilities of existing information technologies to ensure the timely generation of data. Strategies for gaining stakeholder support might involve collaborating with elected officials in the design of performance measures, or asking managers about the extent to which they are willing to implement performance measurement (DeSantis and Leal 1998). These and other conditions are discussed below.

Finally, it is unknown whether having a broader range of functions increases the use of performance measurement. Lessons learned about performance measurement in one function may transfer to another, especially in counties with central reporting systems. Also, counties expend relatively more on education, welfare, public health, hospitals, and correction functions than do cities, which spend more on police, fire, parks, housing, roads, and utilities. Although it is sometimes held that designing viable outcome measures in areas such as education or social services is more challenging, we simply do not know whether developing performance measurement or capacity in these functions is, in fact, more difficult. The following section discusses the use of performance measurement in counties, followed by an analysis of the effect of capacity on its use.

Findings

How Much Performance Measurement?

This study finds that 33.6 percent of counties use some form of performance measurement (see "Methodology" box), but counties vary in the intensity of their use (Table 1). Although measures of use are somewhat arbitrary, this study defines "high intensity" as using performance measurement in at least 75 percent of functions and, within these functions, using workload as well as effectiveness or quality measures. Based on this measure, 20.1 percent of jurisdictions which use performance measurement can be said to have "high use," or 6.8 percent of all U.S. counties. If this standard of use is relaxed, so that performance measurement is used in 66 percent of all functions and only 80 percent of such functions are required to use a broad range of measures, then this percentage increases to 32.5 percent, or 10.9 percent of all counties. Thus, about 20 percent–33 percent of counties which use performance measurement can be said to have high use, or 7 percent–11 percent of *all* counties over 50,000.

Table 1 County Performance Measurement

Breadth of use	
Percent of functions that use performance measurement:[1]	
Less than 25.0 percent	13.9%
25.1–50.0 percent	8.1
50.1–75.0 percent	10.5
75.1–100 percent	67.5
Depth of use	
Percent of efforts that include outcome or quality measures, as well as workload measures:	
Less than 25.0 percent	38.3%
25.1–50.0 percent	13.9
50.1–75.0 percent	14.8
75.1–100 percent	33.0

[1]Among counties that use performance measurement. On average, counties have 12.6 of the 18 listed functions in Table 2.

Methodology

A survey was administered in 1998 to county managers regarding performance measurement use in their jurisdictions. The survey was pretested on a group of 50 county managers and, following minor changes, mailed to 856 counties with populations over 50,000 (identified through *Counties USA,* 1997). After three waves of mailing, 209 responses were received from counties using performance measurement. To determine the extent to which we canvassed counties that use performance measurement, a telephone survey was conducted among a random sample of counties that did not respond. Of the 106 nonresponding counties contacted, only 13 counties indicated that they use performance measurement in some way. Thus, (856–209)*13/106= 79 counties did not respond to the survey and use performance measurement. The survey *response rate* of counties using performance measurement is (209/209+79=) *72.5 percent.* It follows that (209+79/856=) 33.6 percent of U.S. counties use performance measurement.

Although the sampling frame is a census, the response set is a sample. Thus, it is appropriate to use statistical tests to determine whether differences in the sample can be inferred to reflect differences in the population of counties as well.

The telephone survey included some randomly selected survey items. Comparison of these responses with those of the mail survey respondents does not indicate problems of nonresponse bias. To ensure valid survey data, we also conducted follow-up telephone calls with respondents who indicated the use of a wide range of measures. Respondents were asked for specific examples and verified their survey responses. Very few changes were made as a result of the telephone interviews. This study uses a composite measure of performance measurement that is based on the breadth and depth of performance-measurement use.

Breadth of performance measurement use: Respondents were asked to identify which of the following18 typical county functions use performance measurement: police; fire; corrections; transportation; education; parks and recreation; library; economic development; code enforcement; street maintenance; animals; solid waste; health; welfare; housing; hospitals; financial administration; and personnel administration. Responses were scaled from 0 to 1, adjusted for functions that are present. A "1" indicates that all of a jurisdiction's functions use some from of performance measurement. A jurisdiction which, for example, has 15 of the above 18 functions, and which uses performance measurement in 6 of them, has a breadth-of-use score of (6/15)= 0.40.

Depth of performance measurement use: Respondents were asked to identify various types of performance measures, shown in Table 2. "Depth" is defined as the percentage of performance measurement efforts that include workload as well as effectiveness or quality measures. The latter measures are associated with outcome-based measures and suggest depth of use. Responses were scaled from 0 to 1, adjusted for functions that are used in respondent's jurisdictions. A "1" indicates that all of the respondent's functions which use performance measurement have workload, effectiveness, and quality measures. A jurisdiction in which, for example, five of its six performance measurement efforts include workload as well as effectiveness or quality measures, has a depth-of-use score of 5/6=0.83.

The Cronbach alpha measure of internal reliability of the composite measure is 0.73. This means that the measures of breadth and depth can be combined into a single, unidimensional measure of performance-measurement use.

Table 2 shows the use of performance measurement by type and by county function. Workload or output measures are most common. With the exception of economic development, housing, education, and hospitals functions—which are among the least common in counties—these measures are used in 66 percent to 75 percent of county functions. These measures are typically based on data that are routinely tracked by administrative, regulatory, or operational processes. Effectiveness or outcome measures are used in 45 percent–50 percent of functions, and quality measures are commonly used in 35 percent–45 percent of functions. Quality measures are used less frequently, as they require additional efforts such as client or citizen surveys. Among frequent county functions, county corrections are least likely to use service quality measures. This finding furthers the validity of the survey findings—presumably few jail wardens survey their prisoners' satisfaction—but it also reflects a growing trend among county jails to survey noninmate satisfaction among inmates' families, lawyers, bondsmen, and vendors.

The lowest quintile of users, based on the index of performance measurement described in the "Methodology" box use performance measurement in only 14.0 percent of their functions, of which 3.6 percent include effectiveness or quality measures. Such counties are defined as "low users" in this study. Counties with populations under 250,000 use performance measurement in 69.9 percent of their functions, as compared with 84.2 percent of functions in counties with populations greater than 250,000. There is no difference in the breadth or depth of

Table 2 Frequency of Use and Type of Performance Measurement by County Function

	Type of Performance Measures[1]			
Function	**Respondents reporting function**	**Workload or output**	**Effectiveness or outcome**	**Service quality**
Personnel	98.1%	70.7%	50.2%	43.4%
Finance	95.7	69.0	50.5	35.0
Corrections	87.6	65.0	43.7	27.3
Parks and recreation	83.7	66.9	44.0	44.0
Code enforcement	82.3	75.0	50.0	40.7
Street maintenance	81.3	74.7	52.3	45.3
Animals	80.4	65.5	46.4	38.1
Police	79.4	68.5	47.6	34.9
Solid waste	78.5	76.2	49.3	45.1
Health	70.8	68.9	52.7	44.6
Economic development	68.9	51.4	46.5	23.6
Welfare	66.0	63.0	51.4	36.2
Transportation	65.1	67.7	44.1	36.8
Library	60.1	67.7	51.2	46.5
Fire	54.1	67.3	45.1	35.4
Housing	47.4	50.0	39.0	19.0
Education	32.5	35.3	32.6	26.5
Hospitals	25.8	40.7	29.6	24.7

[1]Of counties reporting function.

performance-measurement use among functions that are "typical" county functions versus those that are more municipal in nature, nor are there differences by region or form of government.[3] Finally, various authors have called for increased professionalism in county government (see Streib and Waugh 1991). Attention to ethics includes a concern for accountability; effectiveness is an important measure of professionalism (Streib 1992). Counties that make ethics a priority have significantly greater breadth and depth of performance measurement. However, having a graduate degree in public administration, which is also a measure of professionalism, is not associated with increased use.

The above findings can be compared to those of cities. In a recent survey of cities with populations over 25,000, Poister and Streib (1999) find that 38 percent use some form of performance measurement, compared with 34 percent found in this survey. As in this study, larger jurisdictions are more likely to use performance measurement. Among cities that use performance measurement, 62.2 percent use workload or output measures, as measured across 13 typical municipal functions. Based on Table 2, such measures are used by 63.5 percent of county functions. The relative findings for the use of effectiveness measures are 45.8 percent (cities) and 45.9 percent (counties), and 39.4 percent (cities) and 36.0 percent (counties) for quality measures. Thus, cities and counties are quite similar in their breadth and depth of performance measurement.[4]

How Much Capacity?

Table 3 reports counties' capacity for performance measurement. Respondents most often agree that they can develop outcome measures (74.7 percent), relate outputs to program operations (79.2 percent), and compare results with goals (74.0 percent). These statements reflect the ability to conceptually understand and apply performance measurement. Somewhat fewer respondents have the technical ability to collect and process data: 57.5 percent report that they have adequate information technology for performance measurement, 61.5 percent of respondents can collect data in a timely way, and 63.1 percent have staff capable of analyzing performance data. Almost half of the respondents (47.4 percent) lack one or more of these technical abilities. Although only 29.1 percent of respondents can conduct valid surveys, these are often contracted out. A strong positive correlation exists between the ability to conduct valid surveys and the ability to collect performance measurement data in a timely way (chi-square= 12.6, $p < .01$).

Most counties that do performance measurement have support from the county manager (88.9 percent) as well as internal (department heads, managers, and supervisors) and external stakeholders. Elected officials frequently give their support (71.1 percent), and in some instances advisory boards (40.3 percent) and citizen advocates (40.3 percent) also support efforts. Table 3 shows that *technical capacity and stakeholder support are significantly associated with increased use of performance measurement,* based on the above index measure (see "Methodology"). Among the measures of technical capacity, the ability to relate outputs to program operations most strongly differentiates "high users" from "low users." Specifically, 97.4 percent of high users can relate outputs to program operations, compared with 53.1 percent of low users.[5] Among counties with high use, 79.8 percent report that they can collect data in a timely

Table 3 County Capacity for Performance Measurement

	Agreement %	Association with use[1]
Technical Infrastructure		
"Most departments in our jurisdiction … "		
Can relate outputs to program operations	79.2	.331**
Can develop outcome measures	74.7	.253**
Can compare actual results with program goals	74.0	.313**
Have a cost-based accounting system	66.7	–.023
Have staff capable of analyzing performance data	63.8	.221**
Can distinguish between outputs and outcomes	63.1	.170*
Can collect performance-measurement data in a timely way	61.5	.339**
Can assess the validity of performance measures	60.8	.217*
Can compare performance measures with those of other jurisdictions	59.6	.212*
Can compare performance measures across departments	58.1	.250**
Have adequate information technology for performance measurement	57.5	.172*
Can conduct scientifically valid surveys	29.1	.128
Aggregate Measure	*56.5*	*.331***
Scale statistics: alpha=.84, mean=.565, st.dev.=.314		
Stakeholder Support		
The County Manager's office supports performance measurement	88.9	.108*
Most department heads support the use of performance measurement	75.0	.351**
Elected officials support the use of performance measurement	71.1	.293**
Most managers support the use of performance measurement	62.0	.212**
Higher governments demand use of performance measurement	49.3	.112
Most supervisors support the use of performance measurement	47.6	.314**
Citizen advocates support the use of performance measurement	40.3	.252**
Citizen advisory boards support the use of performance measurement	36.8	.203**
Most employees support the use of performance measurement	36.5	.310**
Aggregate Measure	*54.5*	*.289***
Scale statistics: alpha=.73, mean=.555, st.dev.=.309		

[1] **1% significance.
*5% significance.
Note: Tau-c measures shown.

way, 67.6 percent have adequate information technology for performance measurement, and 79.5 percent have staff capable of analyzing performance measurement data.[6] These three items, along with the ability to relate outputs to operations, strongly distinguish counties in this sample: 79.3 percent of counties which have all four capabilities have a high use of performance measurement, whereas 75.0 percent of counties that lack all four capabilities are among the lowest quintile of users.

But stakeholder support matters, too. Support from county managers is important, but it does not greatly differentiate high users from low users (85.0 percent versus 79.5 percent). Even low users often have the county manager's support, which suggests that this is a prerequisite for doing performance measurement of any kind. Rather, what often differentiates high users from low users is the support of elected officials (92.9 percent versus 46.3 percent), as well as support from department heads (90.5 percent versus 48.8 percent) and supervisors (73.8 percent versus 29.3 percent). Elected-official support, for example, is consistent with the above framework and is mentioned by interviewees (below). Although less commonly sought, support from citizen advocates, advisory boards, and employees also differentiates high users from low users.[7] Mandates by higher governments do not directly increase use, but they are significantly associated with support by elected officials (chi-square=17.1, $p < .01$) and county managers (chi-square=6.8, $p < .01$). Support by elected officials and department heads, along with the above four capabilities, is shared by 88.0 percent of all counties with high performance measurement use.

In the sample, 30.6 percent of counties have *all* of the above six capabilities: (1) relating outputs to operations; (2) collecting data in a timely manner; the presence of (3) staff capable of analyzing performance data; (4) adequate information technology; and support from (5) department heads and (6) elected officials. Across all measures of capacity (Table 3), 38.7 percent of counties meet or exceed the average technical and stakeholder capacity as counties which are high users. *Thus, about one-third (31percent–39 percent) of all counties which currently use performance measurement can be said to have adequate capacity for it.* The item that most differentiates high-capacity counties from low-capability counties is the presence of staff that can analyze performance data.

These findings are supported by the interviews. One county administrator who noted the importance of both technical and political problems stated that the biggest challenge is "to get buy-in from the county commission as well as elected department heads [because] these people have very little experience in this area.… After getting their support, the greatest problem will be being able to analyze the data." Some smaller counties, which are just beginning their efforts, note the challenges of getting staff involved: "There has been no emphasis on data collection before … some staff think that it is not their job." But even counties with well-established efforts noted problems with data collection and political support. In one such county, it was said that "we like to emphasize quality, but we conduct our citizen survey only once each year, and we cannot go out every month and conduct this survey." This county is also concerned that in the event of an economic downturn, performance measurement would be abandoned by the commission to cut costs. Another large county with a large, well-known effort commented that "the main obstacle is getting everyone on board with the measures." This respondent also noted resistance from the commission to use performance measurement as part of setting budget priorities. Further findings from the interviews are reported in "Building Capacity: Interviews" box.

The level of capacity quickly drops off among the lowest users: only 9.1 percent of low users have adequate capacity. However, even low users have some capabilities that allow them to implement performance measurement at some level. Relative to those respondents who indicated that they do not use performance measurement, low users are more likely to identify program goals, distinguish outputs from outcomes, develop outcome measures, and derive performance measures from goals.[8] That is, low users have developed conceptual abilities, but they lack the broader technical and political capabilities that characterize high users. However, low users have only 1.5 years' experience with performance measurement, as compared to 4.0 years for high users. County capacity in these six areas increases over time (3.7 of six measures for counties with efforts that are less than three years old versus 4.7 of six measures for more established efforts. t=2.77, $p < .01$). This suggests that low users broaden their capabilities over time when they become more involved in performance measurement.

Finally, there are no differences in capacity based on the mix or type of county functions using performance measurement, but counties with council-appointed administrators have higher technical readiness than counties with commission forms of government, even when controlled for city size.[9] Making ethics a priority is associated with higher stakeholder support (t=2.05, $p < .05$), perhaps inspiring trust among elected and appointed officials that top managers can be trusted to fulfill their promises. Having a graduate degree is not associated with increased capacity for using performance measurement. This study also finds that technical and political capacities are higher when these efforts are led by a central-budget or finance office. When such offices are involved, respondents report, on average, 7.7 of the 12 technical capacities shown, compared with only 2.6 among counties that are not led by a budget or

Building Capacity: Interviews

Many counties using performance measurement report support from elected officials, but a frequent concern is that performance measures are not much used in allocation decisions in the budget process. Hamilton County, Ohio (population 864,000), is trying to ensure that elected officials use performance-measurement data and participate in developing performance measures by presenting them in a way that shows their relationship with key agency goals and priorities. Mecklenburg County, North Carolina (population 580,000), balances performance-measurement data with citizen responses to provide elected officials with a varied approach to program review. Other counties, such as Saline, Kansas (population 52,000), which has only recently begun its efforts, are still trying to get elected officials interested in performance measurement. In this instance, the county administrator has worked hard over several years to explain the benefits of performance measurement and has also shown a videotape series to elected officials.

In many instances, budget and senior managers have also gone to great lengths to obtain staff buy-in. Managers argue that performance measurement helps staff demonstrate accomplishment, yet interviewees report that initial performance-measurement efforts are often viewed as punitive and are not used by managers outside the budget process. To ensure a positive momentum, Fairfax County, Virginia (population 820,000), is looking at new ways to make its 2,400 indicators more relevant to departments, elected officials and others. Catawba, North Carolina (population 118,000), is also using performance measurement to stimulate innovation in service delivery, focusing on increased effectiveness and cost savings. Still, other counties such as Washington County, Minnesota (population 180,000), are seeking institutionalization by linking the performance measurement of program objectives to the individual performance appraisals of department heads and other managers.

In most counties, agencies are responsible for their own data collection. Some departments are concerned that reporting and collection requirements require considerable staff time. Hamilton County's budget office is now conducting focus groups with each department to identify and address their problems with performance measurement. A frequent complaint is that measures are not part of budgeting and other information technology systems. There also are ongoing conceptual problems about designing meaningful measures. Even counties with established efforts, such as Mecklenburg County, are still working on developing satisfactory outcome measures for such areas as general administration. Virginia Beach County, Virginia (population 393,000), and Washington County have designated a staff person in the budget office to streamline and help departments in their data collection and reporting. In Orange County, Florida (population 677,000), recent efforts have caused some departments to allocate a staff person whose primary responsibility is performance measurement. By working together, these employees solve many performance-measurement issues for their departments.

finance office. According to interviews, the budget office often lends expertise in identifying performance measures and supports the development of information technology for performance measurement.

A Strategy for Increasing Capacity

Counties greatly vary in the areas in which they lack capacity, and self-assessment is the essential first step toward increasing capacity (Van Wart 1995). Table 4 shows strategies that managers might undertake to assess the capacity of their counties. Of respondents, 61.5 percent ask managers about their ability to develop performance measures, and doing so is associated with increased technical capacity (that is, the aggregate construct shown in Table 3). Managers who assess their county's information gathering and analysis capabilities have higher levels of technical capacity. However, only half the respondents do so, and very few (9.6 percent) use any diagnostic checklist to assess their capability for performance measurement. About half of the managers assess whether they have adequate resources for performance measurement. This is associated with increased capacity. Table 4 shows that the aggregate measure of assessing technical feasibility is significantly associated with increased capacity.

One way to determine the level of stakeholder support is to ask them for it: 85.4 percent of managers in high-use counties ask elected officials for their support, in contrast to only 33.3 percent of managers in low-use counties. Managers in counties with high use of performance measurement are also more likely to ask support from advisory boards (52.6 percent versus 15.6 percent, $p < .01$). The aggregate measure of assessing stakeholder support is positively associated with stakeholder capacity. Counties with high use of performance measurement use, on average, half of the strategies shown (7.5 of 15), compared with only 1.9 strategies used by counties with low use. In short, successful counties ask around.[10]

Overall, then, the chain of causality is clear: counties with high levels of performance measurement have higher political and technical capacities, and counties that have these capacities undertake a multitude of assessment strategies. On average, cities with high technical capabilities are far more likely to ask managers about their understanding of performance measurement and ability to develop them. They are also more likely to examine budget resources for performance measurement.[11] Counties with high political capacity for performance measurement are more likely to ask lower managers whether they are willing to implement performance measurement, and they are also more likely to develop strategies to obtain the support of elected officials. Some of these differences are especially striking: for example, 57.6 percent of counties with high political capacity ask lower managers whether they are willing to implement performance measurement, whereas only 17.5 percent of counties with low capacity do so.

However, the efficacy of these managerial strategies is affected by the context in which they operate. In this regard, the form of government does matter, albeit indirectly. Elected officials in council-administrator governments are significantly less likely to participate in the design of per-

Table 4 Strategies for Assessing Capacity

	Use (%)	Association with capacity[1]
Technical Feasibility		
Asking managers about their ability to develop performance measures	61.5	.399**
Asking managers about their understanding of performance measurement	57.0	.389**
Assessing information-gathering capabilities	48.0	.363**
Examining budget resources for performance measurement	45.5	.240*
Assessing information data-analysis capabilities	44.4	.521**
Asking a consultant to assess ability to conduct performance measurement	21.1	.157
Assessing the ability of managers to conduct scientific surveys of clients	11.2	.099
Use a diagnostic checklist to assess capabilities for performance measurement	9.6	.199*
Aggregate Measure	*33.8*	*.344***
Scale statistics: alpha=.79, mean=.338, st.dev.=.280		
Stakeholder Support		
Obtaining support from elected officials for implementing performance measurement	59.2	.402**
Asking managers to identify a pilot project for performance measurement	48.7	.311**
Asking agency directors whether they are willing to implement performance measurement	46.5	.323**
Asking lower managers whether they are willing to implement performance measurement	35.9	.410**
Seeking support from advisory boards for efforts to use performance measurement	35.1	.396**
Proposing a pilot effort assess commitment	34.5	.251**
Negotiating rewards and expectations with managers regarding performance measurement	23.5	.233**
Aggregate Measure	*36.9*	*.431***
Scale statistics: alpha=.73, mean=.369, st.dev.=.295		

[1] Shown are associations with, respectively, technical and stakeholder capacity (see Table 3).
**1% significance.
*5% percent significance.
Note: Tau-c measures shown.

formance measures than those in other forms of government (28.9 percent versus 52.6 percent, $p < .01$). Further, elected officials who participate less in the design of performance measures are also less likely to support them. Indeed, participation by elected officials is negatively associated with council-administrator forms of government (tau-c = –0.25, $p < .01$), yet positively associated with commission (tau-c = 0.14, $p < .05$) and elected executive governments (tau-c = 0.12, $p < .05$). The significance of these findings is robust, even when controlling for the size of counties or the level of performance-measurement use.[12] Participation by elected officials is positively associated with increased technical and political capacity (both $p < .01$), and many respondents told us that support from elected officials for performance measurement greatly helped these efforts.

Professionalism is also of interest: making ethics a high priority in organizations is positively associated with gaining the support and participation of elected officials. By contrast, when county officials feel that they must deal with mismanagement before undertaking new initiatives (a measure of lacking professional norms and practices), they report significantly fewer efforts to assess counties' political capacity for performance measurement. Making ethics a priority is not associated with using assessment strategies regarding technical infrastructure.[13]

Does Capacity Affect Satisfaction with Performance Measurement?

Table 5 shows the outcomes of performance measurement in counties. Overall, almost half of the respondents agree or strongly agree that performance measurement increases awareness of the need for accountability (48.0 percent) and the ability to determine service efficiency, effectiveness, and timeliness (40.0 percent–43.0 percent). About one-third agree that accountability and commitment to excellence have actually increased (35.6 percent and 31.5 percent, respectively). About one-quarter believe that performance measurement has improved group decision making (26.6 percent), and some agree that it has helped them to eliminate services (16.1 percent). Further analysis shows that counties that use performance measurement experience greater benefits by using it more: for example, 63.4 percent of counties with high performance-measurement use agree or strongly agree that it has helped them to clarify program goals and objectives, as compared to only 17.7 percent of counties which have a low use (t =4.34, $p < .01$). High-use counties also agree more strongly that it has helped them increase their commitment to excellence (53.7 percent versus 14.7 percent, t =3.78, p <.01). Although, on average, respondents agree that performance measurement has positive outcomes, 49.8 percent of respondents disagree to varying degrees with one or more of the above statements.[14]

Technical infrastructure and stakeholder support increases satisfaction with outcomes (both $p < .01$, see Table 6). Counties with support from elected officials and lower managers, adequate management-information systems for performance measurement, and the ability to conduct valid surveys report significantly higher levels of satisfaction with outcomes, even when controlling for the extent that they use performance measurement ($p < .05$). Many technical capacities are also associated with examining resources for performance measurement. Additional analysis shows that counties which meet the criterion of adequate capacity for high use (see above) agree or strongly agree with 55.7 percent of the statements shown in Table 5, as

Table 5 Outcomes of Performance Measurement

Program Outcomes	Agreement
Increased awareness about the need for accountability	48.0%
Increased ability to determine service efficiency	45.0
Increased ability to determine service effectiveness	43.0
Increased ability to determine service timeliness	40.0
Established performance target levels for programs/services	40.0
Clarified agency or program goals and objectives	37.2
Improved accountability of program performance	35.6
Ability to achieve improvements despite resource constraints	32.5
Increased commitment to excellence	31.5
Improving group decision-making capabilities	26.6
Determined long-term budget needs	23.6
Eliminated services that are no longer needed	16.1
Improve timeliness of management decisions	15.5
Aggregate Measure	*33.5*
Scale statistics: alpha=.83, mean= .335 st.dev.=.324	

Note: Shown are the percent who agree or strongly agree based on the following scale:

1=strongly disagree; 2=disagree; 3= disagree somewhat; 4=don't know; 5= agree somewhat; 6=agree; 7=strongly agree.

Tau-c measures shown.

compared to 28.5 percent among counties which lack such capacity ($p < .01$). Satisfaction steadily declines among counties with both low use of performance measurement and low levels of capacity (7.7 percent).

The model further shows that negative employee attitudes are a barrier to performance measurement. Such attitudes, however, are not associated with outcomes, suggesting that this effect can be overcome once efforts are underway. Models which include making ethics a priority show that this variable is not significantly associated with either the level or outcome of performance measurement, but this variable is associated with fewer instances of negative employee attitudes ($r = -.53$, $p < .01$). Although technical capacity is not significantly associated with the level of performance measurement when, for example, controlling for stakeholder capacity, it is significantly associated with subsequent satisfaction ($p < .01$). Finally, although the form of government is insignificant in both models, the preceding analysis shows that it is associated with generating stakeholder support for performance measurement.

Table 6 Regression Analysis

Dependent variables	Level of use		Satisfaction with outcomes	
Independent Variables				
Constant	0.26	(0.10)*	–0.31	(0.10)**
Technical capacity (Table 3)	0.15	(0.10)	0.44	(0.10)**
Stakeholder capacity (Table 3)	0.34	(0.10)**	0.42	(0.09)**
Resources for generating and testing new ideas	0.00	(0.05)	0.00	(0.43)
"Employees just act busy"	–0.33	(0.07)**	0.02	(0.07)
County size (population)	0.04	(0.02)*	0.03	(0.02)***
Level of use	n.a.	n.a.	0.19	(0.08)*
Form of government[1]	0.04	(0.05)	–0.04	(0.04)
Northeast	–0.07	(0.08)	–0.04	(0.07)
South	–0.07	(0.06)	–0.07	(0.06)
West	–0.07	(0.07)	–0.09	(0.73)
R^2-Adjusted		.292		.418
N=		155		147

[1] Council-appointed administrator government =1; Other =0.

** 1% significant.

* 5% significant.

*** 10% significant. Shown are the regression coefficients, and, in parentheses, the standard errors.

Conclusion

This study finds that about one-third of counties use performance measurement and that about one-fifth of these have a high level of use. Among those that use performance measurement, about one-third (31 percent–39 percent) have an adequate level of capacity, although the level of capacity decreases sharply among low users (9 percent). Capacity requires that jurisdictions are able (1) to relate outputs to operations; (2) to collect timely data; have (3) staff capable of analyzing performance data; (4) adequate information technology; and support from (5) department heads and (6) elected officials. Support from elected officials is less forthcoming in council-administrator counties, as these officials participate less often in the development of performance measures. The results show that the absence of these conditions does not forestall initial application, but that they threaten widespread use. This research also finds that widespread use of performance measurement increases satisfaction with its impacts.

Assessing counties' capacity for performance measurement is a useful diagnostic strategy to identify areas of organizational deficiency. Many managers are accustomed to experiential learning, in which initial efforts are tried and then give way to larger ones as experience grows and obstacles are identified and overcome. Initial small-scale efforts help identify and address capacity shortfalls. Such an approach is decidedly rational in a world in which counties vary greatly in their abilities and deficiencies. The nature of many management reforms, including performance measurement, increasingly requires much groundwork before they can be successfully and fully implemented. Whether the challenge is inadequate technical abilities to collect data or the reluctance of elected officials to come on board, confronting such challenges will require time and effort. Considerable foresight is now available to help managers identify these challenges.

Notes

1. Program evaluation also required a degree of trust between elected and appointed officials, which some observers felt was lacking. According to Wildvasky (1972), evaluation could reveal shortcomings of managers, as well as the impracticality of elected officials' goals. These problems loomed large in his mind: "I started out thinking that it was bad for organizations not to evaluate, and I ended up wondering why they ever do it."
2. Likewise, Congress, through the Government Performance and Results Act of 1993, requires all federal agencies to submit annual performance data from 1999. A majority of states have legislation requiring performance-based information from state agencies.
3. However, as discussed further, the form of government indirectly affects the level of performance measurement by its effect on the capacity for performance measurement.
4. On average, counties report having used performance measurement for 4.2 years. Counties which have used performance measurement for more than 10 years use more effectiveness or quality measures: 62.1 percent of such counties report using effectiveness or quality measures in their functions, as compared with 46.4 percent of counties which use it for a shorter time. Counties have increased the number of performance measurement efforts over time. The data also suggest a growing trend of using central-budget or finance offices to spur initial uses of performance measurement: 30.7 percent more counties report that they are led by the central-budget or finance office in efforts that are less than three years old. Also, counties which have used performance measurement for less than three years use it in 82 percent of their functions when led by the budget or finance office, as compared to 57 percent of functions when led by other offices or agencies.
5. Although one might conjecture that all counties which are unable to apply performance measurement in this manner should be low users, such counties sometimes use performance measurement to compare program results with goals, which does not require this capability.
6. It might be expected that all "high use" counties would have these three capabilities, but in practice many counties make do with some manual data collection and the use of outside statistical assistance for analysis.
7. Respectively, the percentages for high and low users are: support from citizen advocates (56.1 percent versus 20.0 percent), advisory boards (45.2 percent versus 17.5 percent) and employees (57.1 percent versus 19.5 percent). All differences are significant at the 1 percent level.
8. The respective percentages compared to nonusers of these four items are 94.1 percent versus 73.7 percent, 64.3 percent versus 43.8 percent, 68.7 percent versus 46.1 percent and 73.3 percent versus 57.1 percent. All differences are significant.
9. The difference in scale means is significant at the 5 percent level: 35.9 percent versus 19.2 percent (t=2.06). No other differences by structure of government are significant.
10. The positive impact of using assessment strategies on the level of capacity is robust, even when controlling for some conditions that may affect it; for example, the effect of pervasive cynical or apathetic employee attitudes might deter the use of assessment strategies because of anticipated difficulties in getting staff to use performance measurement. Indeed, the presence of apathetic staff attitudes negatively affects the use of assessment strategies as well as the level of capacity ($p < .01$). Cynical attitudes by elected officials also decrease their support for performance measurement ($p < .05$). While such attitudes make the task of implementing management reforms more challenging for public managers, they do not affect the positive relationship between the use of assessment strategies and the level of capacity.
11. Specifically, cities with high technical capacity undertake 4.0 of the technical assessment strategies shown in Table 4, as compared to 2.1 strategies among counties low technical capacities. Counties with high political capacities undertake 5.2 of the political assessment strategies, compared to 2.9 strategies among counties with low political capacity. Both differences are statistically significant.
12. The negative effect of the council-administrator form of government on the use of performance measurement is, in large measure, offset by the propensity of larger counties to use more performance measurement.
13. Having a graduate degree is also not associated with undertaking assessment strategies.
14. Interestingly, the involvement of central-budget or finance offices is unrelated to outcome satisfaction. Satisfaction with performance measurement increases over time, too, even when the number of efforts used is controlled for.

References

Ammons, David. 1995. *Accountability for Performance: Measurement and Monitoring in Local Government.* Washington, DC: International City/County Management Association.

———. 1996. *Benchmarking for Local Government*. Thousand Oaks, CA: Sage Publications.

Berman, Evan. 1998. *Productivity in Public and Nonprofit Organizations*. Thousand Oaks, CA: Sage Publications.

Berry, Frances, William Berry, and Stephen Foster. 1998. The Determinants of Success in Implementing an Expert System in State Government. *Public Administration Review* 58(4): 293–306.

Broom, Cheryle. 1995. Performance-Based Government Models: Building a Track Record. *Public Budgeting and Finance* 15(4): 3–17.

Brown, Mark, Darcy Hitchcock, and Marsha Willard. 1994. *Why TQM Fails (And What To Do About It)*. Chicago: Irwin.

Carroll, James. 1995. The Rhetoric of Reform and Political Reality in the National Performance Review. *Public Administration Review* 55(3): 302–11.

Center for Accountability and Performance. 1998. *Performance Measurement: Concepts and Techniques*. Washington, DC: American Society for Public Administration.

Cigler, Beverly. 1995. County Governance in the 1990s. *State and Local Government Review* 27(1): 55-70

Cope, Glen H. 1997. Bureaucratic Reform and Issues of Political Responsiveness. *Journal of Public Administration Research and Theory* 7(3): 461–71.

Counties USA. 1997. A Directory of United States Counties. Detroit, MI: Omnigraphics.

Gerhart, Craig. 1995. Management Involvement. In *Accountability for Performance*, edited by D. Ammons, 187–95. Washington, DC: International City/County Management Association.

Gianakis, Jerry, and Jeffrey Stone. 1997. Measuring the Performance of the Municipal Organization: An Implementation Tale of Two Cities. *Administrative Theory and Praxis* 19(3): 73–87.

Greiner, John. 1996. Positioning Performance Measurement for the Twenty-First Century. In *Organizational Performance Measurement in the Public Sector*, edited by A. Halachmi and G. Bouckeart, 11–50. Westport, CT: Quorum.

Grifel, Stuart. 1994. Organizational Culture: Its Importance in Performance Measurement. *Public Management* 76(9): 19–20.

Grizzle, Gloria. 1982. Measuring State and Local Government Performance: Issues to Resolve before Implementing a Performance Measurement System. *State and Local Government Review* 14(September): 132–6.

Hakes, J. 1996. Comparing Outputs to Outcomes. *PA Times* 19(10): 1–2.

Hatry, Harry, et al. 1992. *How Effective Are Your Community Services? Procedures for Measuring Their Quality.* Washington, DC: The Urban Institute and International City/County Manager Association.

Harris, Jean. 1995. Service Efforts and Accomplishments Standards: Fundamental Questions of an Emerging Concept. *Public Budgeting and Finance* 15(4): 18–37.

Holzer, Marc, and Kathe Callahan. 1998. *Government at Work: Best Practices and Model Programs*. Thousand Oaks, CA: Sage Publications.

Honadle, Beth. 1981. A Capacity Building Framework: A Search for Concept and Purpose. *Public Administration Review* 41(5): 575–80.

International City/County Management Association. 1998. *Municipal Yearbook 1998*. Washington, DC: International City/County Management Association.

Jones, L.R., and Jerry McCaffery. 1997. Implementing the Chief Financial Officers Act and the Government Performance and Results Act in the Federal Government. *Public Budgeting and Finance* 17(1): 35–55.

Joyce, Philip. 1993a. The Reiterative Nature of Budget Reform: Is There Anything New in Federal Budgeting? *Public Budgeting and Finance* 14(3): 36–46.

———.1993b. Using Performance Measures for Federal Budgeting: Proposals and Prospects. *Public Budgeting and Finance* 14(4): 3–17.

Keehley, P., S. Medlin, S. MacBride, and L. Longmore. 1997. *Benchmarking for Best Practices in the Public Sector*. San Francisco, CA: Jossey-Bass.

Kettl, Donald. 1994. *Reinventing Government? Appraising the National Performance Review.* Washington, DC: The Brookings Institution.

Leithe, Joni. 1997. *Implementing Performance Measurement in Government*. Chicago, IL: Government Finance Officer Association.

Lewis, Edward, and George Taylor. 1994. Policymaking/Implementing Activities of Elected County Executives and Appointed County Administrators: Does Form of Government Make A Difference? *International Journal of Public Administration* 15(5): 935–53.

Mascarenhas, R.C. 1996. Searching for Efficiency in the Public Sector: Interim Evaluation of Performance Budgeting in New Zealand. *Public Budgeting and Finance* 16(3): 13–26.

Melkers, Julia, and Katherine Willoughby. 1998. The State of States: Performance-Based Budgeting Requirements in 47 of 50. *Public Administration Review* 58(1): 66–73.

Menzel, Donald. 1996. *The American County: Frontiers of Knowledge*. Tuscaloosa, AL: University of Alabama Press.

Mikesell, John. 1995. *Fiscal Administration: Analysis and Applications for the Public Sector.* Belmont, CA: Wadsworth.

Poister, Theodore, and Gregory Streib. 1999. Performance Measurement in Municipal Government: Assessing the State of the Practice. *Public Administration Review* 59(4): 325–35.

Pressman, Jeffrey, and Aaron Wildavsky. 1973. *Implementation*. Berkeley, CA: University of California Press.

Radin, Beryl. 1998. The Government Performance and Results Act: Hydra-Headed Monster or Flexible Management Tool. *Public Administration Review* 58(4): 307–16.

Romzek, Barbara. 1998. Where the Buck Stops. In *Transforming Government*, edited by P. Ingraham, J. Thompson, and R. Sanders, 193–219. San Francisco, CA: Jossey- Bass.

Rubin, Irene. 1997. *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing*. Chatham, NJ: Chatham House.

Sabatier, Paul, and Daniel Mazmanian. 1979. The Conditions of Effective Implementation. *Policy Analysis* 5(4): 481–504.

Salzer, Mark, Carol Nixon, L. James, A. Shut, Marc Karver, and Leonard Bickman, 1997. Validating Quality Indicators: Quality as Relationship Between Structure, Process, and Outcome. *Evaluation Review* 21(3): 293–309.

Stokes, Anthony, and Beverly Monaco. 1997. Can Technology Really Help Your Results Act Implementation? *The Public Manager* 26(3): 21–4.

Streib, Gregory. 1992. Professional Skill and Support for Democratic Principles. *Administration and Society* 24(1): 22–40.

Streib, Gregory, and William Waugh. 1991. Probing the Limits of County Reform in an Era of Scarcity: A National Survey of County Administrators and Executives. *Public Administration Quarterly* 15 (3): 378–95.

Svara, James. 1993. The Possibility of Professionalism in County Management. *International Journal of Public Administration* 16(12): 2051–80.

Theurer, Jim. 1998. Seven Pitfalls to Avoid (When Establishing Performance Measures). *Public Management* 80(7): 21–4.

Tigue, Patricia, and Dennis Strachota. 1994. *The Use of Performance Measures in City and County Budgets*. Chicago, IL: Government Finance Officers Association.

U.S. Bureau of the Census. 1997. *Census of Governments: Finances of County and Municipal Governments*. Washington, DC: U.S. Department of Commerce.

U.S. General Accounting Office. 1997. *The Result Act: Observations on the Office of Management and Budget's July 1997 Draft Strategic Plan.* Washington DC: General Accounting Office.

———.1998. *The Results Act: Assessment of the Governmentwide Performance Plan for Fiscal Year 1999.* Washington DC: General Accounting Office.

Van Wart, Montgomery. 1995. The First Step in the Reinvention Process: Assessment. *Public Administration Review* 55(5): 429–38.

Walters, Jonathan. 1998. *Measuring Up*. Washington, DC: Governing Books.

Wholey, Joseph, and Harry Hatry. 1992. The Case for Performance Monitoring. *Public Administration Review* 52(6): 604–10.

Wildavksy, Aaron. 1972. The Self-Evaluating Organization. *Public Administration Review* 41(4): 480–5.