

Lessons from Performance Measurement Leaders: A Sample of Larger Local Governments in North America



Credits

This paper was written by Shayne C. Kavanagh, Senior Manager of Research for the GFOA's Research and Consulting Center in Chicago, Illinois. It is a custom research project completed for the City of Calgary. To contact the GFOA about the topic of this paper or to inquire about your own custom research project, please contact Mr. Kavanagh at skavanagh@gfoa.org.

ABFM Member Reviewers

Association for Budgeting and Financial Management Reviewers

Mark D. Abrahams, CPA, President, The Abrahams Group

RADM Fred L. Ames, USCG (Ret.) P.E., Assistant U.S. Trade Representative for Administration,
Office of the U.S. Trade Representative, Executive Office of the President

Jane Beckett-Camarata, Ph.D., MPA, Senior Lecturer, Penn State University-Harrisburg

Lynda A Boswell, ICMA-CM, International Monitoring and Evaluation Consultant

Harry P. Hatry, Distinguished Fellow and Director of the Public Management Program, the Urban Institute

Jeff D. Lilly II, Director of Governmental Relations, RCR Technology Corporation, Inc

Elaine Yi Lu, Associate Professor, Department of Public Management, City University of New York

Marcus D. Mauldin, Assistant Professor, The University of Tennessee at Chattanooga

Joseph Ohren, Professor, Political Science Department, Eastern Michigan University

Irene S. Rubin, Professor Emerita, Northern Illinois University

Daniel Williams, Associate Professor, Baruch College

Case Studies

City of Baltimore, Maryland

Andrew Kleine, Chief, Bureau of the Budget and Management Research

Chad Kenney, Director, CitiStat

City of Irving, Texas

Lynda Johnson, Chief Performance Officer/HR Director

Johanna Mandel, Performance Manager

King County, Washington

Michael Jacobson, Deputy Director, Performance and Strategy

Fred Jarrett, Deputy County Executive

Christie True, Director, Department of Natural Resources and Parks

Marguerite Ro, Chief of Assessment Policy Development and Evaluation Unit, Public Health

John Klein, IT Strategic Planning Manager, Department of Information Technology

Jim Chrisinger, Director, Continuous Improvement Team; Office of Performance, Strategy and
Budget

Nancy Buonanno Grennan, Director, Human Resources Division, Department of Executive Services

Maricopa County, Arizona

Sandi Wilson, Deputy County Manager

Kirk Jaeger, MBA, Strategic Business Consultant

Janet Woolum, PhD, Strategic Initiatives Manager, Human Services Department

City of Minneapolis, Minnesota

Paul Aasen, City Coordinator

Jay Stroebel, Deputy City Coordinator

City of Windsor, Ontario

Natasha Couvillon, CGA, Manager of Performance Measurement and Financial Administration

Tony Ardovini, Deputy Treasurer, Financial Planning

Helga Reidel, Chief Administrative Officer

Harry Turnbull, Executive Director of Information Technology

Lee Anne Doyle, Executive Director/Chief Building Official

Thom Hunt, City Planner / Executive Director

Jo-Anne Gignac, City Councillor

Melinda Munro, Manager of the Office of Continuous Improvement

Ronna Warsh, Community Development & Health Commissioner

The GFOA also wishes to acknowledge all other elected officials, departmental directors and other members of the public service who participated in the study.

The survey research for this report was conducted using www.surveymethods.com.

Sponsored by the Government Finance Officers Association and the City of Calgary, Alberta.

GFOA received a grant from the City of Calgary, Alberta, to study zero-based budgeting for public employers. GFOA used the grant to conduct independent research using a survey of GFOA members, case studies, and secondary sources. The findings and resulting publication were reviewed and approved by an independent panel of GFOA members.

June 2013

This is an exceptional study. The presentation is especially well done and will be helpful to any agency that is looking for guidance as they implement performance management systems.

-Fred Ames,
Office of the U.S. Trade Representative,
Executive Office of the President

The report is excellent. It is one of the most comprehensive and insightful reports on the practical issues relating to PM that I have seen.

-Harry Hatry, the Urban Institute

The Research and Consulting Center is the management analysis and consulting arm of the Government Finance Officers Association and is nationally recognized for its comprehensive analytical and advisory services, and specialized research on state and local government finance. Since beginning operations in 1977, the GFOA Research and Consulting Center has provided advisory services to hundreds of local, county, and state governments; public utilities; elementary and secondary education systems; and transit authorities.

The GFOA's Research and Consulting Center encourages inquires about this study or about repeating the analysis for other governments — please contact Shayne Kavanagh at 312-977-9700 or skavanagh@gfoa.org.

Copyright 2013 by the
Government Finance Officers Association
203 N. LaSalle Street, Suite 2700
Chicago, Illinois 60601
www.gfoa.org

Table of Contents

1.	Introduction and Executive Summary.....	5
	Methodology	5
	A Note on the Terms Used.....	5
	What Did GFOA Find?	6
2.	Development of the Performance Measurement System.....	9
	Major Historical Milestones	9
	Generating Interest in Measurement	14
3.	Staffing and Training for a Performance Measurement System	16
	Staffing for Performance Measurement	16
	Location of Analytical Capacity	16
	Training Program	18
4.	Analytical Approaches Used in Performance Measurement.....	20
	Historical Trending and Other Methods.....	20
	Target Setting	21
	Benchmarking	22
5.	Using Performance Measurement for Decision Making.....	25
	Performance Review Sessions.....	25
	Planning and Budgeting	26
	How Departments Use Measures	29
6.	Validation of Measures	34
7.	Elected Officials’ Perspectives on Performance Measurement	36
8.	Performance Measurement Challenges and Defects.....	37
	Performance Measurement Challenges	37
	Defects of Performance Measurement Systems	40
9.	Conclusion: The Value of Performance Measurement	41
	The Lasting Impacts of Performance Measurement	41
	Overall Value of Performance Measurement	42
	Appendix 1: Research Methodology	44
	Appendix 2: General State of the Practice of Performance Measurement	47

1. Introduction and Executive Summary

The Government Finance Officers Association (GFOA) received a research grant in 2012 to study the performance measurement experiences of larger North American cities and counties with the reputation for advanced performance measurement practices. Our objective was to find out what value the participating governments perceive to experience with performance measurement and to identify what they believe are the keys to their success in using it.

Methodology

The GFOA first interviewed the budget and performance measurement staff in each organization (see accompanying sidebar for a list of the 10 case study governments). We then interviewed a variety of managers in operating departments as well as elected officials to solicit their perspectives on performance measurement. The GFOA also surveyed all management-level staff in each organization under study, both inside and outside of the budget/performance measurement office. The survey, therefore, had hundreds of individual respondents and the survey results presented in this report reflect these individuals' experiences. Appendix 1 contains further information on the research approach and the people who took part in the interviews and survey.

A Note on the Terms Used

Public-sector performance measurement uses a number of specialized terms, not all of which are commonly understood or consistently defined. For this report, we elected to simplify the terminology to avoid confusion. For example, the term performance measurement is used across this entire report, while the term performance management is omitted. This is because: 1) the scope of this report is a bit narrower than the term performance management sometimes suggests, and 2) performance management has a different meaning outside of public budgeting and finance. For example, in human resource management it refers to managing the performance of individual employees.

The GFOA also avoided referring to categories of measures that were too fine-grained or detailed. We refer primarily to two categories. The first category includes measures of day-to-day operations, such as measures of efficiency, service quality (e.g., turnaround time for citizen requests), and workload counts (e.g., number of units produced). The second category covers measures of community conditions, such as literacy rates, local perceptions of public safety, and other measures that pin-

The Case Study Organizations

- City of Baltimore, Maryland
population 619,000
- City of Irving, Texas
population 220,000
- Halifax Regional Municipality, Nova Scotia
population 414,000
- King County, Washington
population 1,970,000
- City of London, Ontario
population 366,000
- County of Maricopa, Arizona
population 3,880,000
- Miami-Dade County, Florida
population 409,000
- City of Minneapolis, Minnesota
population 387,000
- City of Windsor, Ontario
population 211,000
- City of Winnipeg, Manitoba
population 684,000

Subjects were selected on primarily based on size and comparability with Calgary, which funded the project, along key dimensions.

point the outcomes of public services that make a community livable and attractive.

What Did the GFOA Find?

Perhaps unsurprisingly, none of the research participants have conducted a formal study to quantitatively evaluate the return from the time and money spent on performance measurement. However, our survey of staff¹ in these organizations shows that they overwhelmingly believe that their experiences with performance measurement have been valuable. About 80% of respondents believe that performance measurement is either “very valuable” (about 45%) or “valuable” (about 35%), while about 10% were neutral. Only the remaining 10% of respondents thought that performance measurement was not valuable. When it came to backing up these views with the budget, the respondents were fairly consistent. About 75% said more money should be spent on performance measurement, while about 13% said current spending levels were about right.

What is the General State of the Practice of Performance Measurement?

The GFOA conducted a more general survey of GFOA members. A fuller accounting of the results is available in Appendix 2. In summary, a slight majority of respondents (60%) use performance measurement regularly, but only 10% use it extensively. Overall, the respondents seemed to find performance measurement less effective than our case study governments, but were still supportive of its use.

In terms of specific benefits, our survey respondents were most enthusiastic about the lasting impact of performance measurement for “increasing the awareness of and focus on results” and “increasing awareness of factors that affect performance” – about 75% felt these were important benefits of performance measurement. About 65% of respondents thought that performance measurement had important lasting impacts across the other areas the GFOA inquired about:

- Improving program/service quality
- Improving responsiveness to customers
- Improving efficiency of programs
- Improving effectiveness of programs
- Changing strategies to achieve results

When asked to identify the most important general benefits of performance measurement, respondents’ top choices were “better decision making” and “improved effectiveness” (in a near tie, these choices each received about 22% of the top votes for most important). “Improved transparency to the public” and “improved productivity” were virtually tied for the next spot (about 16% each).

¹ The survey was distributed to staff organization-wide that held a significant role in the performance measurement system. As a result, many different functional areas were represented in the responses, though the vast majority of respondents were management-level staff.

Interestingly, survey respondents were less enthusiastic about the effect of performance measures in budgeting and planning. For example, only 40-45% thought performance measurement was effective for “changing the level of approved budget,” “improving communication between the budget office and departments,” “reducing/eliminating ineffective services,” or “affecting cost savings.” Our interviews indicate operating department managers don’t often use measures extensively in the budget process. Instead, they tend to use more conventional decision-making criteria, such as what was spent last year, how much new money is available this year, and changes in workload. Budget/management departments, though, are continuing to work to improve this state of affairs.

Our research participants placed the most emphasis on the following elements as the keys to achieving the benefits from performance measurement:

- ✓ **Chief executive support is indispensable.** Support for performance measurement from the chief executive is essential for success.
- ✓ **Regular performance review meetings are essential.** Regular performance review meetings between departments and the chief executive are regarded as critical and even indispensable to a performance measurement system. These meetings serve to accentuate the importance of making decisions using data and provide a forum for discussion about performance.
- ✓ **Performance measurement’s greatest value is as a tool for learning.** If performance measurement is used as a punitive tool, it will face resistance. Using the system to surprise people or create “gotchas” will not endear the system to many. However, if conversations about performance are designed to be constructive, then the system will receive much better attention.
- ✓ **A centralized performance measurement coordinating authority provides a locus for performance measurement activity.** A central office should coordinate the performance measurement program (though not be “responsible” for it). This could involve establishing reporting formats, providing technical assistance to departments, and coordinating performance review meetings. It does not necessarily mean that data analysis should be centralized – in fact, only one of our research participants had centrally located analysis capabilities. Rather, it is more important that departments have the capability to use data locally.

There were also a few surprises about what was *not* a key factor for achieving benefits from performance measurement:

- ✓ **Highly sophisticated analysis.** Historical trending and comparisons to external standards were, by far, the most common and most sophisticated approaches used. Departments, however, still reported getting good results from measurement practices even without using more sophisticated measurement and analysis techniques.
- ✓ **High tech.** While many of our case study governments are using some advanced technology and have found it helpful, it was not central to their experience. In fact, in one government the CIO refused to implement a high-tech “dashboard” because the organization’s use of performance measurement had not yet caught up to the technology. This emphasized the need for the

hard, hands-on work needed to design and use a performance measurement system.

- ✓ **Robust training programs.** The predominant approach to training is one-on-one counseling between the office of management/budget and departments (as it might be needed) and on-the-job learning. Regular “courses” or seminars on performance measurement are not common. Survey respondents found the predominant approach to be adequate.

The balance of this report, divided into the following sections, further develops the ideas above and presents many other research findings:

- **Development of the Performance Measurement System.** Major milestones in the development of performance measurement systems, how to generate interest in performance measurement, and anticipated future directions for the case study governments.
- **Staffing and Training for a Performance Measurement System.** The staffing level of the central performance measurement office and the training approach for performance measurement.
- **Analytical Approaches Used in Performance Measurement.** Analytical approaches used, such as historical trending, benchmarking, and target setting.
- **Using Performance Measurement for Decision Making.** Regular performance review meetings, measurement in the budget process, and performance measures in operating departments.
- **Validation of Measures.** How concerns of reliability/validity of performance data are addressed.
- **Elected Officials’ Perspectives on Performance Measurement.** How elected officials view and use performance measures.
- **Performance Measurement Challenges and Defects.** The major challenges with performance measurement and how to address them.
- **The Value of Performance Measurement.** Reviews the value that our research subjects have received from performance measurement.

2. Development of the Performance Measurement System

This section addresses the development of a performance measurement system, including:

- **Inflection points.** These are the key points to look for (and attempt to create) when developing a system.
- **Generating interest in measurement.** These are strategies to consider for generating more interest in performance measurement among staff and elected officials.

Inflection Points

The participants in our case studies experienced a number of inflection points — points after which a significant improvement occurred after the use of performance measurement — when developing their performance measurement systems. These experiences could help guide governments aspiring to become data-driven organizations on their own journeys by demonstrating some of the characteristics of a successful approach. The milestones fall into two categories: leadership/political and measurement system design changes.

Leadership/Political

Strong executive leadership was a common element among many of the case studies. Probably the most well-known example is Baltimore’s Mayor Martin O’Malley, who spearheaded the development of CitiStat.² In Baltimore’s case, the city did not have a history of performance measurement. The mayor required that data be used to monitor performance and make decisions through the CitiStat program. Strong elected leaders were also important for other governments. Minneapolis had some experience with performance measurement since the 1980s, but only became heavily engaged in using performance data when Mayor RT Rybak came into office. Similarly, the elected county executive, Ron Simms, prioritized performance measurement in King County.

Executive leadership need not necessarily come from an elected official. Maricopa County’s widely known “managing for results” program was developed under the direction of the appointed county manager. The performance measurement systems of the City of Irving and the Regional Municipality of Halifax were also developed under the direction of an appointed chief administrative officer. In the City of Windsor, the chief information officer actually initiated performance measurement, and a new chief administrative officer expanded the use of performance measurement from there.

Regardless of whether executive leadership comes from an elected or appointed official, it is important to establish a vision for what the performance measurement system will be used to achieve and how it will get there. For example, Mayor O’Malley

² “Citistat” is a measurement system popularized by the City of Baltimore. Much has been written about Citistat by other authors. Readers unfamiliar with CitiStat are encouraged to consult these other sources. See for example: “The Citistat Model: How Data-Driven Government can Increase Efficiency and Effectiveness” by the Center for American Progress, 2007.

wanted to drastically improve front-line customer service. Hence, CitiStat – with its emphasis on measures relating to issues like pothole repair times and status of outstanding code violations – was a good way to realize this vision. A vision can be inspired by successful models of performance measurement in other governments. Many of the executive leaders in our research were inspired by CitiStat to undertake their own performance measurement initiatives. These initiatives took on their own distinctive flavor, but a real-life, successful, easy-to-understand model of performance was instrumental for generating an initial interest in performance measurement. For instance, Mayor RT Rybak and his chief administrative officer found that CitiStat matched well with their own vision for performance measurement in Minneapolis.

It is important to note, however, that models used successfully by governments cannot always be copied directly. For example, King County’s Ron Simms was intrigued by CitiStat, leading to the development of “KingStat.” However, the county found that the focus on tactical measures of day-to-day service was not fully supportive of King County’s vision to make progress on bigger-picture community conditions, such as the job retention rate of adult employment program clients, or the percent of eligible children enrolled in the Children’s Health Initiative. King County became more successful in its aim once it developed methods to supplement KingStat by tracking bigger-picture measures of community condition through a separate, complementary initiative (King County AIMs High system), which tracks community condition indicators across seven major goal areas.

Elected Versus Appointed Leadership

Miami-Dade County has had a unique experience with both elected and appointed executive leadership in performance measurement. The county’s system was initially developed by the county manager. Following a change in the county’s charter, the county switched to an elected executive system of government. The new elected executive was very enthusiastic about performance measurement and has actually furthered the use of data-driven decision making in the county. The lesson here may be that elected executives are potentially able to drive performance measurement further and faster than appointed officials because they have political power to match their administrative authority. On the other hand, a performance measurement system might have more staying power in a professionally managed organization because incoming managers are somewhat more likely than incoming elected officials to have the professional training that teaches them to value and use measures.

While much has been written about the role of executive leadership in performance measurement, much less has been said about the role of the governing board (e.g., a city council). The board can adopt policies or other legislation to support performance measurement. For example, King County’s board adopted legislation that institutionalizes particular practices (e.g., KingStat, AIMs High) and that mandates that all parts of the county government participate in the performance measurement system. Miami-Dade County adopted “Governing for Results” legislation that served a

similar purpose. The legislation balances between providing discretion to management to develop and adjust the system as may be required and requiring accountability from staff to track, use, and report performance data.

However, complications may arise if the board itself attempts to lead the development of performance measurement. For example, before Ron Simms's tenure in King County, the County Board promoted performance measurement as a way to increase the transparency of the county executive's operations. This set the system on a foundation of mistrust because one branch of government was seeking to increase the transparency of another (leading to questions of motives, how the data would be used, etc.). Years later, the board found a more constructive approach – it charged the county's independent auditor with forming a performance measurement workgroup consisting of representatives from across county government. The workgroup served as a forum to exchange ideas and to develop practical standards for the application of performance measurement across the county.

Finally, some governments found that the political stresses caused by budget pressures actually benefited performance measurement. Budget pressures require greater efficiency and difficult spending cutback decisions. In this challenging environment, performance measurement can demonstrate its value. Miami-Dade used performance measures to help identify which programs were most important and to thereby better target cutback decisions. The process began with a review of the departments' business plans, including projected performance for the upcoming year using performance measures. The measures help more precisely define the benefit that will be created by a public service. Rather than just relying on what a public service would cost the county, discussions could focus more on the "value" the department will create through its services. The business plans were then compared to the countywide strategic plan to find out where the department's proposed activities aligned with the strategic plan. Measures made it easier to understand the results of a department's activities, and therefore made it easier to assess whether those activities were contributing to the county's larger strategic goals. Activities that were not in alignment represented a more likely candidate for cutback. The measures in the business plan also had more direct use in the budget discussion. For example, measures of program output could provide a basis of discussion for how service levels would be impacted by a cutback. This enabled budget decision makers to better articulate where cuts might be most acceptable and to understand better the service reductions that would result from budget cuts.

This was a far more satisfactory approach for participants than the across-the-board-cuts. In Maricopa County, departments that made the best use of performance measures tended to fare better during cutbacks, thereby demonstrating the importance county leadership placed on measurement. The City of Windsor's use of performance measurement benefited from a somewhat different form of stress. When an extensive strike by organized labor compromised some city services, the council needed better performance data to more intelligently consider alternatives to traditional service delivery models.

Leadership/Political Checkpoints

- ✓ Successful performance measurement requires executive leadership. The leader could be either an elected or appointed official.
- ✓ The leader should articulate a vision for how performance measurement will be used and what it will achieve. Successful models from other jurisdictions are often helpful to make the vision more tangible for followers.
- ✓ It may be difficult for the governing board to play the lead role, since performance measurement is largely a managerial activity. However, the board can support measurement through appropriate policies and legislation.
- ✓ Financial challenges may provide an opportunity to show the value of performance measurement.

Measurement System Design Changes

The governments in our research cited a number of changes to the design of their performance measurement systems that were inflection points in their use of performance measurement. For many governments, the most important inflection point was the development of a strategic planning and/or business planning process. The articulation of goals for the government and its departments begs questions like: did we accomplish what we set out to do? How effective have we been?

Most governments in our research have an organization-wide strategic plan that states organization-wide goals. The strategic plan then might be supplemented by business plans that articulate objectives for the individual departments. Miami-Dade offers an illustrative example of how strategic planning and business planning were used to drive performance measurement. Miami-Dade has identified seven “strategic areas” as part of its strategic plan: public safety; transportation; recreation and culture; neighborhood and infrastructure; health and human services; economic development; and general government. The strategic plan identifies the major goals for each of these areas. Several county departments have responsibilities in each of the strategic areas. Departments prepare two-year business plans that address how they will contribute to the strategic area. This approach is intended to help make sure that the county comprehensively works on the larger issues that matter to the public. Measures are developed for the strategic areas as well as the departments to help gauge progress towards the goals.

While strategic planning and business planning is the most common inflection point for using performance measures, our case study governments also experienced a number of others, reviewed below.

Enabling Technology. A few case study governments were enthusiastic about the ability of technology to facilitate access to data and information. Miami-Dade implemented an off-the-shelf software called “Active Strategy Enterprise” that helps them better visualize, organize, and communicate the entire constellation of interlocking goals and measures that make up their system. Baltimore and Minneapolis both

cited 311/customer relationship management technology as crucial for capturing data on responsiveness to citizen requests and completion of work orders. Interestingly, though, technology has not played a pivotal role in most of the case study governments. In fact, the Windsor CIO's refusal to implement a web-based performance measurement dashboard because of the poor quality of the underlying data was seen as an inflection point because it really focused the city's attention on the need to improve data quality and use measures to make decisions before worrying about how to present data.

The Budget Office as the Locus?

Our case studies suggest that the locus may be better located outside the budget office because it deemphasizes performance measurement as just a "budget tool" and emphasizes its use as an every-day management tool.

Create institutional locus performance measurement activity. A central office of performance measurement has proven important for coordinating and supporting performance measurement activities across the organization for the case study governments (the staffing for such offices is discussed later in this report). Maricopa County, for example, experienced an inflection point when two of its performance measurement "facilitators" were moved out of the budget function and into a more dedicated performance measurement structure. This new structure had a more flexible relationship with other units in the county that have an important role in making performance measurement successful. For example, this enabled a closer relationship with human resources in order to give the performance measurement staff greater influence on the county's in-house staff training program and access to additional training capacity. This has improved the effectiveness of the performance measurement training program and has helped with beginning to integrate the county's performance measurement and performance appraisal systems. However, the central performance measurement coordinating offices found in our case studies are not large and play more of a facilitative role in performance measurement. A dedicated performance measurement office does not necessarily mean that data analysis is centralized – in fact, it is important the departments are able to analyze their own data so that they can use it for day-to-day management.

Measurement System Design Change Checkpoints

- ✓ A strategic plan is crucial for identifying overall goals for the organization. Business plans address the actions and measures for individual departments.
- ✓ Technology could be very helpful to performance measurement, but is not essential. In fact, it could even prove a harmful distraction – for example, if measurement presentation technology is prioritized over data quality.
- ✓ It is helpful to have a dedicated function to coordinate and support performance measurement, but this doesn't necessarily have to include centralized analytical capabilities.

Generating Interest in Measurement

Our research subjects took conscious action to generate interest in performance measurement so that it would be used to manage the organization. Perhaps the most important action for generating interest in measurement is regular review of the measures with departments and the chief executive, with all parties in the same room. This point was consistently made by the research subjects. Baltimore, which pioneered the CitiStat model and is well-known for its lively review meetings, noticed that the energy for performance measurement dropped noticeably when the review meetings were suspended for a six-week period while the CitiStat office took time to rework their approach. Participants missed the joint problem-solving interaction with their peers and the mayor's focus on performance measurement simply was not as visible without the meetings. This emphasizes that the meetings must be held consistently or previous gains can rapidly deteriorate. Some pointers from our research subjects on using such meetings to generate interest include:

- The chief executive must be directly involved. This makes it a lot of easier for department directors to motivate their workforce because it shows that demand for measurement comes from the very top. If the chief executive cannot attend the meeting in person, he or she should assign an authoritative proxy to be there instead. The chief executive need not attend every meeting in the organization about performance (for example, departments may wish to meet together on their own to discuss shared performance goals), but there should be at least one dedicated forum for direct chief executive involvement in performance measurement.
- The meetings should involve a little bit of “theatrical flourish” to keep them interesting and lively. For example, interesting analysis, visualizations, or challenges to status quo thinking can all serve this role.
- The meetings should provide an atmosphere where open and frank discussion is safe. To this end, one research subject recommended against making the meetings open to the public out of concern that participants would self-edit.
- Consider inviting council members to the meeting. This provides another form of visible, high-level support.

Another essential means for generating interest in performance measurement is to ensure that measurements have business value to the departments. Often, government performance measures capture summary-level information that might be of interest to a chief executive or the board, but offer little management value to the department. To be seen as legitimate, the system must provide operational benefit to departments. This quote from a survey respondent captures the issue well:

“We've created statistics that do little to help us understand our process pinch points and what we do about them. We create council reports to answer the questions posed, but think little about the processes to get that work done once council [has its question answered]. Therefore, I have to believe that performance management is simply window dressing.”

Of course, the downside to focusing on measures that departments find valuable is that these measures may focus narrowly on department operations (particularly efficiency and quality) and not provide information on how constituents' lives are improved as result of public services or on broader community conditions that might be of interest to the board and/or CEO. Consequently, the system will need to find a compromise where departments track information that they find helpful, while also satisfying the information needs of the board and CEO.

Generating Interest in Performance Measurement Checkpoints

- ✓ Foremost is to hold regular meetings where the chief executive reviews measurement results with departments. Departments should conduct their own internal meetings to prepare for the executive meetings and to drive improvement within the department.
- ✓ The measurement system should provide operational value to departments.

3. Staffing and Training for a Performance Measurement System

After the system has been developed, it must be maintained. This involves developing a staffing and structuring strategy for the central measurement office, and training staff throughout the entire government agency to use performance measures. Accordingly, this section of the report addresses the following issues with respect to maintaining performance measurement capabilities:

- **Staffing for performance measurement.** This section reviews how the case study organizations staff their performance measurement office.
- **Location of analytical capacity.** Analyzing measures is key to getting value from measurement, so this discussion focuses on where that analytical capacity is located.
- **Training program.** The approach that the case study governments have taken to training staff in the use of performance measurement is addressed.

Staffing for Performance Measurement

Staffing of the central management/budget office for performance measurement is generally fairly light, with most of the performance measurement duties falling to departments. The exception is Baltimore, which has a centralized analysis capability (though departments also have their own capacity). Exhibit 3.1, on the following page, summarizes the staffing and duties of some of the offices that participated in our research.³

Location of Analytical Capacity

Of note, Baltimore is the only case study in our research to have a dedicated, centralized analyst group. There are four analysts (though there is a lot of turnover), each of which has a portfolio of three or four agencies that they study and analyze using a “template” of standard data they obtain from departments. The City’s 311 system is the most commonly used data source for the analysts. The larger city agencies also have their own analysts. This does result in some duplicative efforts, but it also means that data-driven decision making is more widely practiced throughout the city.

The other case study governments rely much more heavily on departments to analyze their own data. The centralized performance management personnel may help analyze data in certain cases (e.g., departments with weak analytical capabilities), but not routinely for all departments. The centralized personnel also have some role in analyzing the bigger picture measures of community conditions, though even this may be shared with operating departments to some extent. Data may also be reviewed and analyzed by a department head’s supervisor and/or peers in review meetings where departments present their performance data to others.

³ Staffing is not completely comparable between the case study governments due to differences in responsibilities of the centralized offices and how duties for performance measurement are divided between central staff and staff in the operating departments.

Exhibit 3.1: Staffing of the Central Performance Measurement Office

	Staffing	Duties
Baltimore		
(pop. 619,000)	1 FTE CitiStat director, 3 FTE , analysts and 1 FTE investigator. 6 budget analysts contribute time (15-20%) on PM activities	-Coordinate CitiStat process, including gathering data submissions from department. -Coordinate and staff CitiStat review meetings -Analyze PM data -Gather & analyze outcomes for “budgeting for outcomes”
Halifax		
(pop. 414,000)	1 manager (.333 FTE), 2 planning & performance coordinators (.75 FTE each) and 1 admin assistant	-Coordinate design of PM system and submittal of measurement data -Coordinate strategic planning and goal setting
Irving		
(pop. 220,000)	Two 0.3 FTEs to coordinate the PM system	- Administering PM database & archive data - Making sure staff enter data in the database -Check data for major variations -Provide training & develop ad hoc reports
Maricopa		
(pop. 3,880,000)	1 FTE in Business Program Management role	-Work with other staff to identify, document, and prioritize business requirements for the Managing for Results Information System -Participate in the testing, evaluation of technical solutions -Coordinate the resolution of system performance issues
Miami-Dade		
(pop. 409,000)	12 total people who are about 25% dedicated to PM duties	Variety of tasks related to coordinating submittal of data, development of the strategic plan, coordination of business plan, and administering dashboard/scorecard system
Minneapolis		
(pop. 387,000)	0.25 FTE coordinator, 0.5 FTE intern	-Coordinate PM effort and Results Minneapolis meetings. -Staff results Minneapolis meetings
Windsor		
(pop. 211,000)	0.5 FTE PM coordinator, 1 FTE intern, 0.5 FTE OMBI coordinator, & 0.75 FTE admin support	-Ensure consistency across the organization in approach to benchmarking & ensure consistency in presentation -Communicate information about PM and the planning process - Facilitate internal reporting
Winnipeg		
(pop. 684,000)	1 FTE coordinates PM and 0.25 FTE reviews financial components of measures	-Coordinate with departments to submit data for budget -Coordinate benchmarking submissions - Handle special projects related to PM

Training Program

One of the more interesting findings of GFOA's research is that inadequate training on performance measurement was not regarded as a problem by survey participants (which included staff from a wide range of operating departments). In fact, it was rated the second least important potential challenge for effective performance measurement out of the ten choices provided. This result is interesting because very few of the case study governments have a particularly robust training program for performance measurement.

Maricopa County seemed to have the most robust training program. The county's "Managing for Results Strategic Consultant" (a staff position) teaches many classes for employees, on performance measurement or related to performance management, throughout the year. Classes include an overview of the county's budgeting and performance management system (called "Managing for Results"), "Performance Measurement 101," and "From Data to Decisions: Analyzing for Results." Also, when the Board of Supervisors or departments updates their strategic business plan, training on performance measures is included as part of the update process.

However, the rest of the case study governments have much less robust approaches (yet dissatisfaction with training does not seem to be apparent among survey participants). The primary modes of training include: 1) one-on-one counseling between the management/budget offices as required to help them refine their measures and 2) on-the-job training where departments simply learn by doing. Some of the following reasons were offered by interviewees as to why a less robust training program seems to be adequate:

- **Performance measurement simply is not that complicated.** To quote one interviewee, "this isn't rocket science." Extensive training is not necessary to explain the essentials of measurement. One-on-one work or hands-on experience may be more useful for learning to apply the basics than formal, classroom training.⁴ Also, as explained in more detail later in this report, departments seem largely satisfied with basic approaches to measurement and analysis, which further explains why extensive training would not be much in demand.
- **Many employees aren't interested in extensive training.** Many employees aren't interested in understanding the deeper nuances of performance measurement and are just interested in the basics of how to measure. Some employees may simply want to know to comply with the chief executive's directives to use measures!
- **Communication is more important to the adoption of performance measurement.** Communicating the fact that the organization's leaders expect data usage to inform decision making may be more important than technical training to the success of performance measurement.⁵

4 According to the National Training Laboratory (Bethel, Maine), adult learning is more effective using hands-on approaches, rather than lecture or readings.

5 This would be consistent with the emphasis many organizational change researchers/authors place on strong executive leadership as a requisite element for organizational change.

- **The case study governments generally have long-standing performance measurement systems.** Some case study governments had more robust training programs when they started performance measurement, but have reduced training as the knowledge has become institutionalized. Even those that did not start out with a robust training program have found the experience has become institutionalized.

Staffing and Training Checkpoints

- ✓ Staffing of the management/budget office for performance measurement is fairly light.
- ✓ Most of the data gathering and analysis duties fall to departments.
- ✓ A centralized analytical capacity is not a necessity. In fact, most case studies rely on departments to analyze their own data.
- ✓ Most case studies do not have particularly robust training programs, yet respondents to our survey do not see lack of training as a big problem.
- ✓ The predominant approach to training is one-on-one counseling between the office of management/budget and departments, as it might be needed, and on-the-job learning.
- ✓ The case study governments generally find the predominant approach to be adequate.

4. Analytical Approaches Used in Performance Measurement

This section addresses how our case study participants analyze performance data. It includes the following subsections:

- **Historical trending and other methods.** Monitoring the trends in measurement results over time was the most popular analytic approach. More sophisticated methods were uncommon.
- **Target setting.** Defining a standard for what constitutes a desirable measurement result is reviewed.
- **Benchmarking.** Using the performance of other organizations as point of comparison for one's own performance is discussed in the final section.

Historical Trending and Other Methods

A shared feature of all the cases is that the methods used to analyze performance data are kept fairly simple. In fact, historical trend analysis was, by far, the most common approach. Historical trending simply involves lining up the measurement results from a number of successive time periods and seeing if the results are going in a desirable or undesirable direction.

Management Tampering

This was W. Edwards Deming's term for when managers react to normal variations in performance by initiating some form of management action. In many cases, tampering makes things worse because some level of variation is normal, so the management action will likely not address root causes of sub-optimal performance and may even worsen performance. For example, it may simply add additional process steps to account for a rare special case that caused the observed variation. Hence, when using trend data managers should be aware of what constitutes acceptable variation within the process and avoid reacting too strongly to unusual cases.

More sophisticated statistical methods like correlation analysis or linear regression do not appear to be in wide use among the case studies. The reason for this is that most managers seem to be satisfied with the guidance they get from the simpler analytical methods. Past GFOA research has shown that, if there is a desire for more sophisticated analysis, it may be wise to identify a small number of critical [performance questions that would benefit by this type of analysis, then identify the precise data needed to answer the question (data collection and cleaning costs are often the majority of the expense of an analytics project, so limiting analysis only to what is precisely needed helps contain costs), and, finally, pilot analytic technology to answer this question. If and when success is achieved, the approach can be spread. The GFOA publication, *The State and Local Government Performance Management Sourcebook*, contains a case study of Oakland County, Michigan's, use of analytics and how it accomplished these steps.

Target Setting

Targets establish a threshold of acceptable performance and provide a standard against which to compare actual results. The case study governments do not necessarily use targets in a consistent manner. For example, Baltimore observes trends and looks for the trend to go the “right” way, but there are few explicit targets – though, turnaround times on work orders are a notable exception. In other case studies, the departments take the lead in setting targets. In these cases, there may or may not be negotiation between the central authorities and the department to adjust the target.

Targets are generally thought to be useful because they have motivational value, they hold people accountable, and without targets performance is too ambiguous. However, targets have important drawbacks as well. The primary problem is articulated by W. Edwards Deming, who believed that targets are largely an arbitrary number.⁶ This arbitrary number can introduce distortions into the system. For example, if the current target is set beyond current performance, the manager has three choices for how to reach the target: redesign the work; distort priorities and the process (e.g., to deemphasize aspects of operations that do not influence the target, but may still be important); or to “game the numbers,” if not outright cheat. If the manager does know how to redesign the work, the question is: why has he or she not done it before and is the target an optimal capacity? If he or she does not know how to redesign the work, the remaining two options are the only ones left. Of course, if the target is set below the program’s true capability, there is little incentive to improve.

The implications of this line of logic follow:

- Performance measurement systems should ideally be accompanied by some sort of continuous improvement approach, such as Lean or Six Sigma. This gives managers the knowledge of how to improve performance and to set performance goals more rationally.
- Performance targets must be considered very carefully. For example, managers must consider what sort of maladaptive behaviors they may be incentivizing with a target. To illustrate, one case study had set targets for the amount of time in which public complaint cases have to be “closed.” The result was that cases were marked as closed prematurely.
- Targets may be especially dangerous when set by people who are further away from the work itself (e.g., central management, the governing board). Those who are further removed have less conception of the real capability of the service, so the target is likely to be more arbitrary and, therefore, more distorting. Further, department managers are better able to spot (and respond to) the maladaptive behavior that arises from targets. For example, in Windsor targets for closure rates of code violation complaints led some staff to mark complaints as “closed” prematurely, but department managers no-

⁶ This discussion on the drawbacks of targets is taken from John Seddon, *Systems Thinking in the Public Sector* (Triarchy Press, 2008).

ticed the problem after some time and took steps to address it. Fortunately, department managers are often close enough to the front line action to notice these types of problems. However, if targets are imposed from outside, the department managers may be the ones incited to maladaptive behavior.

Benchmarking

Benchmarking is the practice of comparing one’s own performance to professional standards or to the experience of other organizations. Benchmarking is used to provide context for performance measurements (e.g., is our performance better or worse than others’) and to encourage the search for better methods of providing public service (e.g., what are the highest performing organizations doing that we can learn from?). When it comes to selecting external organizations to benchmark against, our case study governments mostly commonly benchmarked against respected local governments that are perceived to be similar along certain key dimensions (the dimensions may vary depending on what service is being benchmarked). A few governments also benchmarked against private firms for services that have a private-sector equivalent (e.g., a call center).

Benchmarking generated more mixed feelings than performance measurement generally, but was still viewed quite positively. When questioned about the frequency of benchmarking, about 19% of our survey respondents benchmark regularly and belong to a formal benchmarking support group of some kind. About 27% benchmark regularly, but do not belong to such a group. An additional almost 25% only benchmark as may be required by an external request. Finally, about 30% of survey respondents do not benchmark at all. When questioned about the value of benchmarking, about a quarter said “Benchmarking has great value and should be done as much as is practical.” Forty percent said “Benchmarking is generally helpful and should be done regularly.” The vast majority of the rest were neutral on the subject, with only a small minority overtly negative.

We did not find vast differences between people in different types of government functions when it came to opinions on benchmarking. Public works and utility-oriented services as well as community development and cultural services tended to be somewhat (20%) more likely to regularly benchmark than other types of services,

Exhibit 4.1: Survey Respondent Opinions on the Value of Benchmarking

Benchmarking has a great value and should be done as much as is practical.	24.3%
Benchmarking is usually helpful and should be done regularly.	39.1%
Benchmarking is useful in particular situations and should be applied to only those situations.	27.4%
Benchmarking has only occasional use, usually in response to special situations like elected official's request or bond issuance.	3.3%
Benchmarking is rarely or almost never useful. Time and effort are usually better spent on other things.	3.2%
No opinion	2.8%

like public safety or administrative or support services. Interestingly, however, public safety and administrative and support functions were about 15% more likely to believe benchmarking to be helpful or very helpful. However, as was described above, all groups had high opinions of benchmarking, overall.

A wide variety of benchmarking sources were used, including benchmarking collaboratives (such as the Ontario Municipal Benchmarking Initiative - OMBI), professional associations, and standards promulgated by the federal government.

While staff was generally positive about benchmarking, there was some concern about the ability to obtain “apples-to-apples” comparisons. The quote from a Minneapolis survey respondent captures the general view expressed by many:

“It's quite difficult to achieve an apples-to-apples comparative analysis. Still, to the extent possible, benchmarking keeps leadership focused on what might be possible within the industry and helps us determine to what level, at what rate of funding, and for what constituency certain services will be provided. It's one tool within the arsenal of a comprehensive approach to performance measurement/management.”

For example, Windsor used benchmarking data as the catalyst for a decision to change their staffing mix at the city's retirement home. The benchmarking data showed that the city had a high cost per customer served, so they examined how the low-cost provider did business. Analysis of the data led the city to begin using more part-time than full-time staff, among other changes.

Benchmarking can be a useful way to set goals and provide motivation. As one respondent from Halifax put it: *“If someone else can do it, so can we. [Benchmarks] can be used to encourage better performance. [Benchmarks] can identify areas for improvement.”*

Another important benefit of benchmarking is that it is a very intuitively attractive concept for many: to compare performance in your own organization to the performance of others. Hence, benchmarking could be a good way to generate interest in measurement among audiences that might otherwise be difficult to engage (e.g., hesitant managers, some elected officials) – and it also is a good way to bring measures other than cost into the performance conversation. For example, OMBI's reports are made public, which generates a good deal of interest in the community and among elected officials, in some of our case study governments. More generally, elected officials, in particular, seem to like benchmarking to support high-profile decisions like fee and ordinance updates.

Benchmarking also fosters relationships with peers in other governments. These networks can be used to find new and innovative ideas on how to improve service. For example, one survey respondent reported great value from learning about different animal control methods from across peer agencies.

However, it is important to set expectations for what benchmarking can accomplish by itself. As the quote from Minneapolis pointed out, benchmarking is just one tool for performance measurement and, though it can help place performance in a larger context, it will not help much to explain detailed reasons for performance. A respondent from King County said it well: “[Benchmarking] helps us ask “why?” as we see differences [in scores].” Put another way, benchmarking won’t substitute for management acumen to go behind the data and examine reasons for under-performance, but it can inspire a valuable questioning of the status quo.

Also, it is important to consider the drawbacks of benchmarking. The foremost drawback mentioned by our research subjects was the difficulty of generating “apples-to-apples” comparisons. Formal benchmarking collaboratives probably offer the best chance for overcoming this problem as they can help a group of governments develop consistent, compatible data definitions. However, even collaboratives are not a totally reliable solution. For example, while the Canadian participants in our research were enthusiastic about the Ontario Municipal Benchmarking Initiative, we did not find similar enthusiasm among our case study governments for any collaborative in the United States. The best solution is likely to limit benchmarking to instances where useful, valid comparisons are available, rather than attempting to benchmark all services simply for the sake of benchmarking. Not only will the comparisons be more valid, but it will reduce the costs of benchmarking, increase the value from benchmarking, and preserve the credibility of the practice. Further, respondents report that politicians tend not to be as sensitive to the comparability of problems of benchmarking as staff – they are more willing to induct benchmarking data into their decision-making despite comparability problems. Hence, it may behoove staff to avoid this problem by not providing faulty data to elected officials in the first place.

The second challenge is that benchmarks may present the illusion of easy answers to performance questions. This quote from the survey captures the problem well:

“The challenge is that good benchmarking involves work. Real work. It is a starting point, rarely an ending point; but, the final deliverable of benchmarking reports, often tables and charts used in comparisons, lend themselves to be used as determinations or answers for those seeking to support a [preconceived] position or argument.”

Analytical Approaches Checkpoint

- ✓ Analytical methods have been kept simple. Historical trending was the most popular method.
- ✓ Comparing measurements to target values can be used, but one must beware of the distorting features of target setting.
- ✓ Staff was generally positive about benchmarking, though there was some concern about the ability to obtain “apples-to-apples” comparisons.
- ✓ Benchmarking is a very intuitively attractive concept for many: to compare performance in your own organization to the performance of others. Hence, benchmarking could be a good way to generate interest in measurement among audiences that might otherwise be difficult to engage.

5. Using Performance Measurement for Decision Making

This section discusses the most important ways in which performance data is used for decision making among our case study participants. It specifically addresses:

- **Performance review sessions.** Review sessions were considered one of the most important uses of performance measurement data by a wide cross-section of our research participants.
- **Planning and budgeting.** Performance measurements have been widely promoted as a way to bring more rationality to planning and budgeting discussions.
- **How departments use performance measures.** The prior two points address organization-wide decision-making processes. This discussion focuses on how measurements are used within departments for making decisions.

Performance Review Sessions

The case study governments commonly held meetings at regular intervals (e.g., quarterly) to review performance and believed these to be of central importance to the success of their system. A review of performance might take place during the budget process, but our research participants believe the budget process is a suboptimal time for a performance review because the focus of the budget process is resource allocation, not performance improvement. Instead, a dedicated forum to review performance is important. The case study governments' experience is consistent with the findings of Harry Hatry and Elizabeth Davies in "A Guide to Data-Driven Performance Reviews."⁷ They cite the following three prerequisites for a successful review meeting approach:

Leadership sufficiently interested to support and participate in regular meetings. Primarily, this refers to direct participation by the chief executive or his or her designee in meetings to review performance data with departments.

Miami-Dade is trying an approach where department leaders within one of the county's six "strategic areas" meet without chief executive representation in order to share ideas amongst themselves. The idea is that the absence of the chief executive will allow department executives to be more frank in their discussion. This requires trust on the part of the chief executive – perhaps a different form of support than Hatry and Davies intended with this point, but important nonetheless.

A performance measurement process that provides timely and accurate data on program outputs and outcomes. Hatry and Davies' point is that performance indicators do not have to be perfect, but their data must be sufficiently valid to permit meaningful, useful discussion. As discussed elsewhere in this report, few of the case study governments have extensive data validation programs, but most are confident that the data is accurate enough to add value to decision-making processes. So, the data does not have to be perfect, but should be good enough that it represents an improvement over "gut" decision making.

7 Harry Hatry and Elizabeth Davies, "A Guide to Data-Driven Performance Reviews," a paper in the IBM Center for the Business of Government "Improving Performance Series." 2011.

Staff that, at a minimum, can assist the leader in examining the data and provide advice on issues to address at the meetings. The primary role of staff is to coordinate the agenda to make sure the meeting is run efficiently and to make sure that performance data is made available in an accessible format. The case study governments relied on both the central performance measurement office and operating departments in varying degrees to provide and format the data. Only Baltimore had a dedicated central analyst function. The accompanying side bar shows a typical agenda from Hatry and Davies' research. Interestingly, all of the case study governments sought to emphasize the need to learn about performance improvement and exchange ideas at these meetings, not use them to prosecute poor performance. This last point even includes Baltimore, which has garnered a reputation for having more confrontational reviews.

Typical Agenda Items at Performance Review Sessions

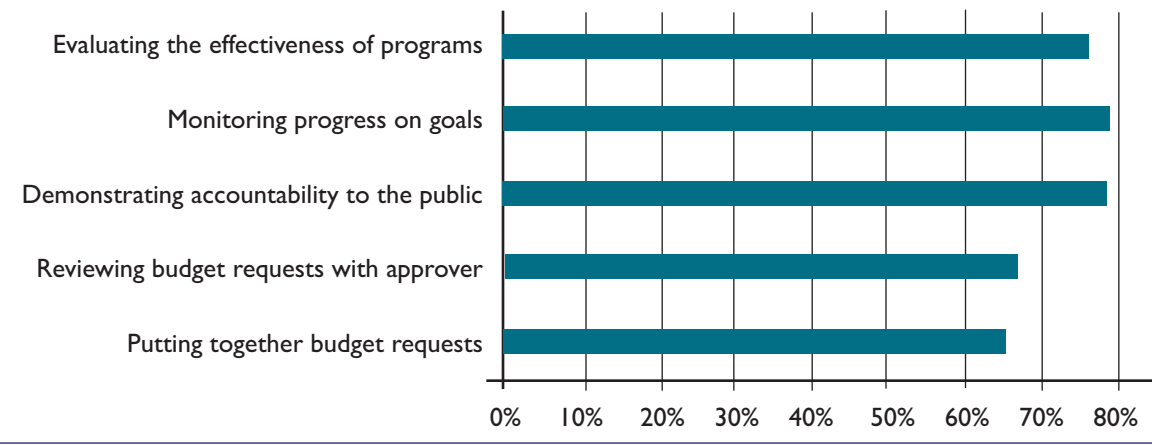
1. Introduction
2. Review action items and outstanding issues from last meeting
3. Discuss overall findings and leadership questions
4. Discuss areas or indicators displaying particularly high or low performance relative to that expected
5. Brainstorm next steps
6. Identify the action items needed

Planning and Budgeting

Performance measurement is closely associated with planning and budgeting. The theory is that performance measures supplement financial data to help the organization make decisions in light of value considerations, not just cost. However, the reality does not appear to have fully realized the promise identified by theory just yet. GFOA's survey of the case study organizations posed three sets of questions that provide insight into this issue.

The first set of questions asked how important a role performance measurement plays in various management processes. Exhibit 5.1 (on the following page) shows the percent of respondents who believe that performance measures were either important or very important for the indicated process. About 65% of respondents believe measures are important or very important for both putting together budget requests and reviewing budget requests with the budget approval authority. Monitoring progress on goals and demonstrating accountability to the public score higher – almost 80%. This is probably not too surprising given the enthusiasm the GFOA found for regular performance review meetings, as discussed earlier.

Exhibit 5.1: Percent of respondents who think performance measures are important or very important to various management processes



The second set of questions asked how effective performance measurement has been in producing results for better communication. Exhibit 5.2 shows the percent of respondents that thought performance measurement is “effective” or “very effective” for improving the indicated facets of communication. Only 40% think that performance measurement has improved communication with the budget office, but much higher percentages believe it has been helpful in other areas. In our interviews, many operating department staff advised us that the most important improvement budget offices could make in their performance measurement approach was to put forth a more concerted effort to learn the business of operating departments. Their view was that measures can only advance communications so far – it is important that budget staff get out of the office and visit the departments to see how the work really occurs.

Exhibit 5.2: Percent of respondents who think performance measures are effective or very effective for improving communication processes

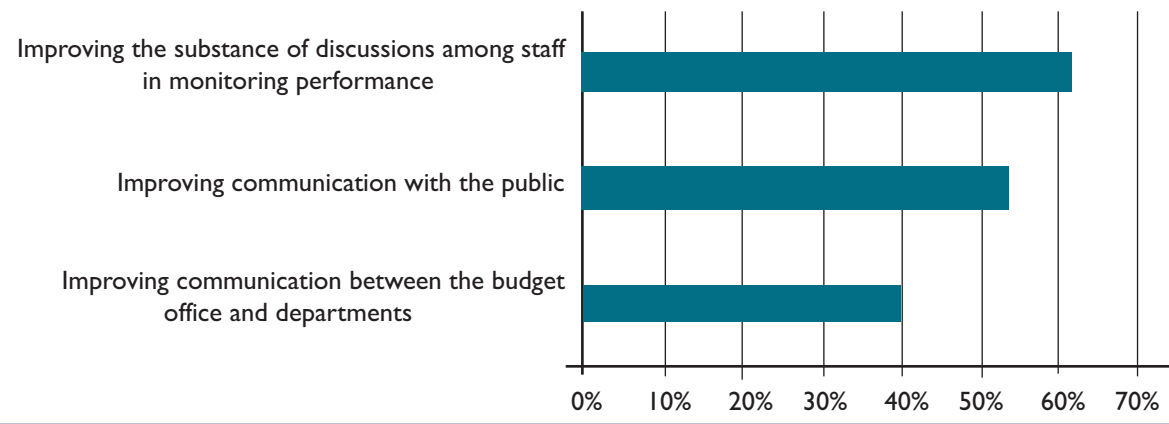
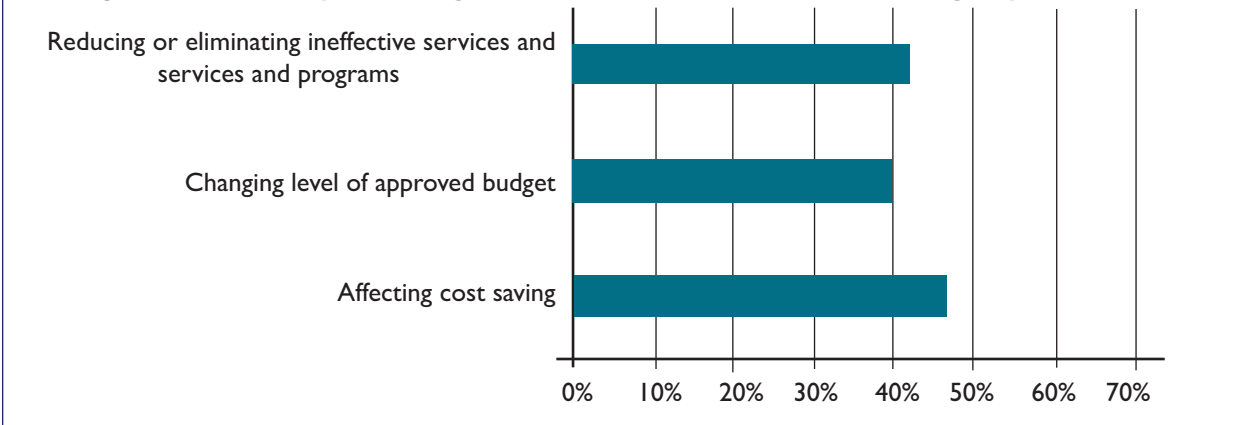


Exhibit 5.3: Percent of respondents who think performance measures are effective or very effective for producing the indicated results in the budget process



The third set of questions asked about performance measurement’s effectiveness for producing results for budgeting, specifically. Exhibit 5.3 shows the percent of respondents that thought performance measurement is “effective” or “very effective” for producing the indicated results in budgeting. Performance measurement was rated as “effective” by less than half for all three areas.

Our interviews suggest that performance measures may not have rated as highly as might have been expected for two primary reasons. First, with respect to cost savings in particular, respondents may perceive that performance measures do not lead directly to cost savings – rather they only suggest where cost savings might be found and then it is up to management acumen to actually affect the savings – or the measures may simply confirm management’s intuition on where cost savings are available. Further, affecting these savings may not necessarily be perceived as a “budgetary” activity.

Second, and more importantly, is that budgets are still largely formulated by departments using conventional decision criteria, such as how much was spent last year, how much revenue is available to fund increases, and the workload the department or program is experiencing. Of course, the offices of management/budget are not satisfied with this state of affairs and many of the case study governments are trying new ways of budgeting to improve the integration of measurement with financial decision making.

For example, Windsor and Winnipeg are emphasizing lines of service (programs), rather than departments in their budgeting. The idea is that it is easier to connect performance to programs than departments, so it will be easier to take into account performance information when allocating resources. Baltimore and Minneapolis are using “Budgeting for Outcomes,” a method of budgeting that uses outcome measures to make allocation decisions based on the effectiveness of programs for meeting organizational goals. The idea behind this budgeting approach is that measures that focus on day-to-day operations are not as useful in budgeting as measures of bigger-

picture community conditions. Measures of community condition spark discussion about which programs and services the government should provide in order to make the biggest difference in the lives of constituents, rather than just how to best perform the programs and services that are already being provided.

Examples of Performance Measures Making a Financial Difference

- In Winnipeg, performance data highlighted the fact that city paramedics spent excessive time waiting at hospitals for patient transfer of care. As a result, management introduced a policy to charge the hospitals for wait times.
- In Baltimore, managers suspected that the cost-effectiveness of their approach to sweeping sidewalks could be improved by increasing the ratio of capital to labor used in the program. They used performance measurement techniques to investigate the idea further and found that cost savings were, in fact, possible.
- One of the ways Winnipeg used benchmarking data was to provide information to decision makers on tree pruning cycles. Recently, funding has been increased to reduce the once-every-14-years cycle to once every seven years.
- In Winnipeg, performance data highlighted the efficiencies gained on the Handi-Transit service by making regular buses more accessible, coupled with a policy of providing free regular transit to Handi-Transit registrants with unlimited eligibility.

How Departments Use Measures

Perhaps unsurprisingly, our interviews and surveys revealed a range of experiences on how department managers use performance measures. We encountered departments that did not use measures at all, except to comply with reporting requirements that come from the central budgeting/management office, and we found departments whose use of measures went far beyond the requirements of the organization-wide system and used measures regularly for monitoring goals and managing and deploying operational resources.

Accordingly, one of the most consistent messages we received from departments was that a performance measurement system must maintain flexibility – no single solution will work for all departments. Departments believe it is important for the budget office to keep an open mind and work with them to understand the key outcomes the department produces (though getting deep into the details of the department operations was not seen as necessary). The intent behind this advice was to avoid a one-size-fits-all approach to performance measurement, where all departments must conform to very specific standardized formats, methods, measures, etc., to the point where it reduces the management value of the system. For example, it might be helpful for a department to have a checklist of criteria that a measure should meet. On the other hand, a system that specified the precise measures to use, or that placed too many or narrow parameters on the types and number of measures allowed would be burdensome.

A related point made by department managers was that the performance measurement system should provide flexibility to change measures over time. Departments' management needs and strategies change, so they should be able to change their measurement system accordingly.

Perhaps the deeper message from the department managers was that the budget/management office must accept that some departments will take performance measurement more seriously than others. Attempts to "force" the less enthusiastic departments into good performance measurement by way of overly detailed processes and formatting requirements will likely result only in rote compliance, while stifling innovation among the more enthusiastic departments.

Another consistent finding was that staff believes measures of efficiency and service quality are more useful than community outcome measures – about 70% found the former to be very important or important to their management approach, while only around 60% said the same about the latter. Even workload/output measures were deemed more important than outcome measures, with 64% support. Of course, recent performance measurement professional and academic literature has advanced outcome measures as the most important type of measure, though it would seem that department managers do not share this view. Our interview and survey offered a number of explanations, such as:

- Departments are immersed in operational work, so prefer measures that give them feedback on their day-to-day responsibilities.
- Data for outcome measures tend not to be available as regularly or as timely as data for other types of measures. In fact, our survey showed that outcome measures are used less frequently by departments, which supports this assertion. Also, some of our interviewees believe that they can only determine outcomes from long-term studies of clientele, which may not be practical to conduct.
- Some departments may not see the connection between their work and larger outcomes, so see outcomes measures as having limited relevance. They see the efficiency and quality of operations as being within their control, but do not see outcomes in the same way. This may be especially true when the outcome measure gauges a larger community condition that the department's work only partially affects.

However, departments can take an interest in outcome measures. For instance, Minneapolis is one case study where outcome measures were rated as more important than other types of measures by staff. Minneapolis operates a CitiStat-style measurement system called "Results Minneapolis;" however, in a departure from "text book" CitiStat, Minneapolis increased emphasis on outcome measures. CitiStat-style⁷ systems are well-known for their ability to focus participants on performance measures, and Minneapolis appears to have used that capability to help focus on outcome measures in particular.

Use of outcome measures could improve organically, over time, as managers gain experience with performance measurement and begin to naturally realize the need for more meaningful measures. For example, the emergency management department in one case study used to measure, via a community survey, the percent of residents who report being prepared for a disaster. However, it found this data to be an unreliable indicator of actual preparedness, so it changed its measure to percent of people prepared to be home alone, on their own for 72 hours without water or electricity (also asked via survey). This more specific measure better informed the department on whether residents were really becoming prepared for a serious disaster.

Departments, overall, reported getting good results from measurement practices – Exhibit 5.1 showed that over 75% believed measurement to be important for monitoring goals and evaluating the effectiveness of programs. Here are three key practices for how departments use measures:

Get staff involved in measurement. Departments that report high levels of success with performance measures have made a substantial effort to get staff below the department director level involved in analyzing and monitoring data. For example, the Maricopa County Transportation Department developed a scoreboard of key measures, based on how their services impacted customers. Examples of some of the measures on the scoreboard include:

- Estimated time and money saved by drivers (e.g., commuters, etc.) as result of improvements to roadways.
- Percent of traffic signal incidents responded to within two hours
- Percent of potholes filled within 48 hours

The scoreboard was presented to all 200 employees in a department-wide meeting and updates to the scoreboard were regularly posted in prominent locations in the department each month. The staff's reception to the scoreboard was lukewarm at first, but more employees have been convinced as to the merit of the system as time has gone on and benefits have materialized. The department also holds a Monday morning performance meeting with the division managers where the scoreboard is reviewed. The department director had tried monthly meetings, but found that they were too infrequent to reinforce the importance of using performance data to manage the organization. Division managers and their subordinates also have clearly delineated responsibilities and deadlines for collecting data.

Perform data analysis locally. None of the case study governments, with the exception of Baltimore (which has four positions for analyzing CitiStat data), have a large performance analysis capability in the management/budget office. Rather, departments are expected to analyze their own data. Their superior knowledge of the context behind the measures leads to more insightful analysis, potentially. Department managers reported to the GFOA that the central management/budget office is more helpful in providing a general framework in which to report measures to the chief executive and policy board than in conducting analysis.

Look behind the measure using “quality” techniques. Measures are also often referred to as “indicators.” This is instructive because performance measures indicate where additional analytical and managerial attention might be required. Hence, the real value of performance measurement that departments have experienced comes not from performing increasingly sophisticated analysis of measurement data, but rather from finding out more about what the data they have is telling them by going behind the numbers. A number of tools from what has been referred to as the “quality” movement⁸ were used by departments for this purpose. There are a number of such tools available, but some of the more common ones mentioned in the case study governments include:

- **Root cause analysis.** This type of technique looks below the immediate cause of under-performance to the root cause. For example, the “5 whys” technique asks the analyst to question why a given condition exists, then question why that underlying condition exists, and to continue to question why each following underlying condition exists until the root cause is reached. Five is the rule-of-thumb number of iterations it is generally thought to take to reach the root cause. More sophisticated methods such as “fault trees” and “fish bone” diagrams are available as well. These methods help to visualize causes of underperformance and provide more structure to the analysis. However, more sophisticated tools are not always needed. For example, Maricopa’s traffic department investigated why their responsiveness to traffic signal incidents had deteriorated for two months straight and found that two repair trucks had been undergoing maintenance during this time period (the repair equipment on the trucks is difficult and time-consuming to service). As a result, the department formed a rental agreement with an outside firm to provide access to equipment for periods when the department’s trucks would be down for an extended period.
- **Process mapping.** Process mapping tools are used to visualize the steps in a process and draw analysts’ attention to potential points of failure such as bottlenecks, hand-offs of work between people, and non-value adding process steps. Process mapping is helpful because it may be the first time the participants in the process have actually seen the entire process and it helps generate discussion about why the process is the way that it is. Process mapping is much more effective when it is part of broader approach to quality improvement, such as Lean or Six Sigma, because these broader approaches provide tools to analyze causes of underperformance and to generate solutions.
- **Set variance tolerances.** Establish upper and lower parameters for what constitutes acceptable, normal variation in a measure. Management’s attention should be drawn to the measure when normal variation is exceeded. This prevents management from taking action in response to normal variation in the measure. Some level of variation is unavoidable, and it can be counterproductive for management to tamper with a process in response to normal variation.

⁸ The quality movement is closely associated with W. Edwards Deming and has evolved into the modern practices of Six Sigma and Lean.

Measurements in Decision-Making Checkpoints

- ✓ Regular meetings between departments and the chief executive to review performance are essential to keeping performance data at the forefront and to integrating measurement data into decision making.
- ✓ GFOA's research finds less effective use of performance measurement in budgeting than might be expected given the emphasis of performance measurement in budgeting best practices.
- ✓ Departments don't appear to be using measures very extensively in the budget process. They tend to use more conventional decision-making criteria, such as what was spent last year, how much new money is available this year, and changes in workload. Budget/management departments are continuing to work to improve this state of affairs.
- ✓ Case study governments are using various budgeting reform – such as budgeting for outcomes and program budgeting – to improve the connection between budgeting and performance.
- ✓ Departments report getting good results from performance measurement. Staff involvement, locally performed analysis, and looking behind measures with quality analysis techniques appear to be among the most important factors for getting the best results.
- ✓ A performance measurement system should provide flexibility to accommodate the varying needs of different departments. A one-size-fits-all approach likely will not work.
- ✓ Departments generally prefer efficiency and service quality measures to outcome measures. This contradicts the direction of professional and academic literature, which promotes outcome measures over other measurement types. Hence, a performance measurement system will need to respect this preference and evolve towards greater use of outcome measures over time

6. Validation of Measures

Are the measurement data valid and reliable? This is a persistent question about measurement data, one that our case study participants encountered as a challenge. None of them completely solved the issue, though the most common approaches to checking the validity and reliability of measures included:

Face validity check. This asks if the measurement results just feels or looks right. The idea behind this approach is that, through regular immersion in the data, performance measurement staff will have built up an intuitive feel for the data that will help them to spot anomalies.

Do not create incentives for departments to falsify results. Targets that are centrally set or measures used for punishment create incentives to distort measures to meet targets. Conversely, de-centrally set targets (or no targets at all) and emphasizing measurement for learning has less potential for distorted incentives.

Create incentives for accurate results. It is better to develop a system where data-driven decision making is in the best interest of the departments so they want accurate data for themselves, not to satisfy a mandate from a central authority. This does not mean that incentives have to be limited to superior performance – rather, incentives can be offered for using data, rather than intuition, to make management decisions. The leading example of this is to treat budget requests that are supported by data more favorably than those that are not.

Maricopa County Audits

Maricopa is the only case study that has an established a history of regular audits of performance data. The internal audit department conducts a measurement certification to examine measures after they are produced. This includes an examination of source data and calculations. The auditor examines a variety of departments each year.

Use data entry controls. Many problems in the data are likely due to human error, rather than malicious intent. Controls on data entry can reduce human error. This might include data validation checks at the source (e.g., setting a warning flag if an entry exceeds the normal range of variation) or automating data capture entirely.

For the most part, however, the case study governments trust that departments are producing accurate data. The rationale is that while the measures may not be perfect, they are likely accurate enough that decisions based on them are an improvement over relying on intuition. Most case study governments report that they would, ideally, like to undertake more rigorous, regular auditing, but lack the resources.

Two less resource-intensive approaches that case study governments are considering to help with data validation include:

- **Create data definitions.** In some cases, the meaning of a measure may not be understood in the same way by all parties to the measurement system.

This could lead to different results when different parties use the same measure. Data definitions provide more precise meaning to the terms behind the measures. Definitions help promote comparability and consistency in performance measures because the components of the measures are similarly defined.

- **Enlist customers.** Customers may be able to help validate performance information. For example, Baltimore has launched a mobile 311 application that allows citizens to send pictures of public service problems. This allows the city to compare what the citizens are showing them to what the city's 311 data says, without having to send staff to inspect the site personally.

For many case study governments, a larger concern than checking data for validity is that the right things are being measured in the first place. These governments make the point that the measurement system needs to be flexible to drop measures that aren't helpful and to add new ones as needed. Sticking to less useful measures just because they already exist will reduce the perceived and substantive value of the system. For example, in one case, a city was measuring whether bus service was on time by simply recording the driver's opinion of whether the bus was on time or not. While the bus drivers may have been honest in their assessment, people are usually overoptimistic when assessing their own performance.⁹ Unsurprisingly, the measure, though perhaps an accurate reflection of driver opinion, was not an accurate reflection of reality.

Also, maintaining alignment between budgets, business plans, strategic plans, and measures requires regular attention. These components of the planning and budgeting system can easily become disjointed. As an illustration, the public works department for one of the case study governments was using employee headcount as a measure of input. However, layoffs did not negatively impact output, quality, or effectiveness – thereby calling into question, in the case study organization's opinion, the usefulness of the headcount measure as an indicator of performance capacity or effectiveness. In other words, although the headcount measure may have been maintained with perfect accuracy, it was not providing value to decision making.

Validation of Measures Checkpoints

- ✓ Most case study governments do not have regular, rigorous validation of performance data. Rather, they use some less rigorous methods and otherwise trust that the data is accurate enough to improve decision making.
- ✓ While the case study governments might like to undertake more rigorous, regular review of validation of measurement data, they lack the resources and believe that is better to spend the resources they do have on making sure the right things are being measured in the right way.

⁹ A vast array of cognitive science research shows that “overconfidence bias” is endemic in human thinking.

7. Elected Officials' Perspectives on Performance Measurement

Elected officials have a complex relationship with performance measures. As one elected official we interviewed put it, “the narrative is the currency of elected officials.” Performance measures, however, may contradict anecdotal evidence. Elected officials can still benefit from measures, and, in fact, our case study interviews indicate that people believe that performance measurement has improved the quality of decision making on the elected board. Here are some of the techniques that our interviewees believe have helped make performance measures more useful to elected officials:

- **Supplement measures with narratives.** A narrative or story provides context to the measures and helps them resonate more effectively with elected officials.
- **Address community conditions with performance measures.** Community condition measures (e.g., perceptions of public safety, unemployment rates) are especially relevant to elected officials' policy-making role. Measuring community conditions through a citizen survey may be especially compelling for elected officials.
- **Make sure measures reflect elected officials' strategic priorities.** This means not only aligning measures with the elected officials' goals through a strategic plan, but updating the plan frequently enough that all elected officials have a chance to provide input into the plan.
- **Emphasize both cost and effectiveness in financial decision making.** Most boards are already conscious of cost. Performance measurement can help them become more cognizant of performance issues. Raise both cost impacts and the impact on citizens when considering de-funding (or funding) a program.
- **Use operational measures to reassure elected officials that things are on track.** Ideally, elected officials will leave matters of program administration to professional managers; however, elected officials also have an obligation to ensure their constituents that standards of basic stewardship are being followed. Performance measures can help reassure them that basic standards of program efficiency and quality are being met.

Elected Official Checkpoints

- ✓ Elected officials often prefer narrative, anecdotal evidence to performance measures, but performance measures can still be helpful.
- ✓ Measures should address community conditions in order to support elected officials in their policy-making role.
- ✓ Strategic plans should be formulated and updated regularly to make sure performance measures reflect elected officials' priorities.

8. Performance Measurement Challenges and Defects

This section of the report reviews common challenges that our research participants experienced with performance measurement, as well as the outright defects they have encountered.

Performance Measurement Challenges

Our survey asked respondents to rank the most important potential challenges of performance measurement. The top three, in order of importance, are below:

- Excessive time/cost to collect data.
- Being held accountable for results that staff can't adequately control.
- Lack of available data.

Excessive time/cost to collect data. This is a top concern for almost all of the case study governments. The basic concern is that, especially in a time of fiscal austerity, staff does not have the time to collect data or administrative positions (that would normally have data collection responsibilities) may have been cut. Our research did reveal a few deeper potential causes and solutions.

- **Data capture technology.** Our interviews indicated that our survey results may, in large part, be due to perceived inadequacy in the technology used to capture data. For example, one case study government just finished an ERP implementation for human capital management, payroll time and labor, procurement, and financial and budget systems. Until recently, it would have taken a manager a week to find out how many people worked in his or her department. While that problem has been fixed, it does illustrate that getting access to data can be frustrating, especially when it comes to more advanced types of data or non-financial data (since these types of systems usually lag behind). Another example is double-entry or re-keying of data. Frustrations could develop if performance data must be entered into multiple systems or re-entered from a system-of-record to a performance measurement database. Determining what data are relevant and then investing in systems to capture high quality data at the source of performance may be a wise initial investment in performance measurement technology and the survey results seem to support this conclusion. 311/CRM systems were seen as a critically important technology for many of the case study governments because they capture many aspects of customer satisfaction – an important measure of effectiveness for many services.
- **Stop collecting data that is irrelevant.** If departments are asked to report on measures that they don't use for management, they will naturally see that time/cost as a waste. This concern can be reduced if the management/budget office works with departments to agree on measures that are useful. This doesn't mean that only operational measures should be collected, however. It does mean that departments should understand how measures of organization-wide importance are used in the decision-making process and departments should be engaged in that decision-making process, such as via strategic planning (the approach used by Miami-Dade to involve departments in strategic planning was described earlier in this report).

- **Sampling techniques.** Sampling methods can be used to generalize about a population from a limited number of observations. Staff training on sampling techniques may help cut down on the number of observations required in order to estimate performance levels.

Being held accountable for results that staff cannot control. This is a classic challenge to performance measurement, especially for measures of bigger picture community conditions and outcomes – the influencers of measurement results are thought to be too multi-faceted and outside the immediate control of operating departments. However, if the entire organization is held responsible for the measure, the dilemma becomes that if “everyone” is responsible, then no one is really responsible.

The first potential approach is for a performance measurement system to attempt to create a link between higher-level organizational goals/measures and the goals/measures of subunits, with the implication that the efforts of subunits’ actions would also impact the organization-wide goals/measures.¹⁰ The case study governments using this approach offered a number of examples of how they have addressed the accountability challenge, though they acknowledge that the challenge has not yet been completely solved.

- **Develop good communication links to departments about the strategic plan.** When everyone is clear on what the goals being measured are, they are more likely to be willing and able to pull in the same direction. Our survey of staff in the case study governments showed that when organization-wide goals are clearer to employees, they believe the goals to be more helpful to their job. When organization-wide goals are clearer they also believe that their own actions are aligned with the governing board’s vision.
- **Identify specific projects/actions that will further the organization-wide goals.** Accountability for specific projects and tasks can be assigned to specific individuals. So, while these individuals may not be responsible for the entire goal, they are at least responsible for actions that are thought to have a positive influence on the goal.
- **Use logic modeling.** Logic modeling asks departments to logically chart out the inputs they receive, the actions taken to transform those inputs into outputs, and how those outputs translate into outcomes for the customers. Logic models can be used to show departments how their actions contribute to an organization-wide goal.
- **Use cross-functional teams.** Teams of individuals from across the departments that have a hand in achieving the goal can be formed. For example, Minneapolis recruits department directors who are enthusiastic about solving a community issue to be on a team responsible for achieving the related citywide-goal. The team becomes jointly accountable for the goal.

The second approach, which is a bit different from the first, would be to separate measures of community-wide condition from departmental performance under the

¹⁰ In the work of the National Performance Management Advisory Commission, this system is called a “cascading system.” The commission’s report describes this system in more detail. See www.pmcommission.org.

premise that departments can't control these measures adequately to be held solely responsible for them.¹¹ Under this approach, the job of department directors is less about influencing the measure through direct government action and more about providing the leadership to convene and coordinate the various actors (e.g., other government agencies, non-profits, private enterprise, etc.) whose help will be needed to achieve the desired community condition. Under this approach, "moving the needle" on the measure becomes a shared responsibility of the various actors in the community who each have a part in reaching the desired condition. The government's own services should still be measured and their contribution to achieving the community condition acknowledged, but government should stop short of trying to draw a direct line between its own programs and the larger community condition that is ultimately desired. This is because, in many cases, government programs alone will not be sufficient to reach the condition – recruiting partners to help will be essential and these partners will have to work with government as a team to make the impact.

Lack of Available Data. This challenge is that performance data may simply not be available to answer the performance questions of interest. Foremost is to accept this as a reality of performance measurement – there will almost always be additional data that would be of interest. Author Mark Friedman suggests creating a data development agenda, which is used to record and prioritize unavailable data of interest. Beyond that, managers can avail themselves of various techniques to generate measurement data.¹²

- **Decomposition.** Decompose a complex problem into simpler component pieces. It may be easier to measure the components, which could help you understand the larger whole.
- **Secondary research.** See how similar problems have been measured elsewhere, such as other governments that measure performance.
- **Experiment.** Use small control and experimental groups to test a hypothesis about performance.

Challenges to Performance Measurement Checkpoints

- ✓ Excessive time/cost to collect data is a top concern for almost all of the case study governments. The basic concern is that, especially in a time of fiscal austerity, staff does not have the time to collect data or administrative positions may have been cut.
- ✓ Being held accountable for results that staff cannot control is a classic challenge to performance measurement: the influencers of measurement results are too multi-faceted and are outside the immediate control of the department.
- ✓ Performance data may simply not be available for questions of interest. This is a reality of performance measurement.

11 This approach is advocated by Mark Friedman in *Trying Hard is Not Good Enough*. Friedman believes that measures of community condition are essential, but that government action is only part of the way to change community conditions. Accordingly, departments should not be held "responsible" for such measures. Rather, public managers should focus on identifying and recruiting partner organizations that can help achieve the outcomes that the community desires.

12 Douglas Hubbard, *How to Measure Anything* (Hoboken, NJ: Wiley, 2010).

Defects of Performance Measurement Systems

In addition to challenges to performance measurement, the survey also asked about outright defects. Two defects stood out from the rest in our survey: “Just going through the motions to comply with requirements” and “Resources necessary to achieve the desired level of performance are not available.”

Just going through the motions to comply with requirements. Respondents perceive that many staff simply offer rote compliance with reporting requirements promulgated by the management/budget office, rather than using measures for more substantive purposes. Interviews with departments would appear to indicate that this defect is unavoidable, to some extent. Some managers simply do not have the temperament to use measures to drive decisions and prefer to use their intuition. Short of a change in managers, which does not happen often or easily in many government organizations, this is a situation that will not change. So, while organizational leaders can and should emphasize the importance of measurements, this research finding suggests that they should also temper their expectations for just how much real enthusiasm they can expect from staff.

Resources necessary to achieve the desired level of performance are not available. This defect is also probably unavoidable to some extent in an environment of financial scarcity. However, it is probably also necessary to work against the perception that improvement can only be made by spending more money on a program. For instance, Mark Friedman’s “Results Accountability” approach makes a point of asking managers to generate “low or no cost ideas” for improving results – and if no such ideas are generated, Friedman believes that managers simply aren’t trying hard enough.¹³ Lean process improvement is another tool designed to achieve drastic improvements in performance without necessarily making large additional investments in a service (and may actually reduce costs).

Defects of Performance Measurement Checkpoints

- ✓ Respondents perceive that many staff simply just goes through the motions to comply with reporting requirements promulgated by the management/budget office, rather than using measures for more substantive purposes. While organizational leaders can and should emphasize the importance of measurements, this research finding suggests that they should also temper their expectations for just how much real enthusiasm they can expect from staff.
- ✓ A lack of resources necessary to achieve the desired level of performance is also probably unavoidable to some extent in an environment of financial scarcity. However, it is probably also necessary to work against the perception that improvement can only be made by spending more money on a program. Lean process improvement is an example of a tool that can achieve drastic improvements, while potentially lowering costs.

¹³ See Mark Friedman, *Trying Hard is Not Good Enough* (Santa Fe, New Mexico: FPSI Publishing, 2005).

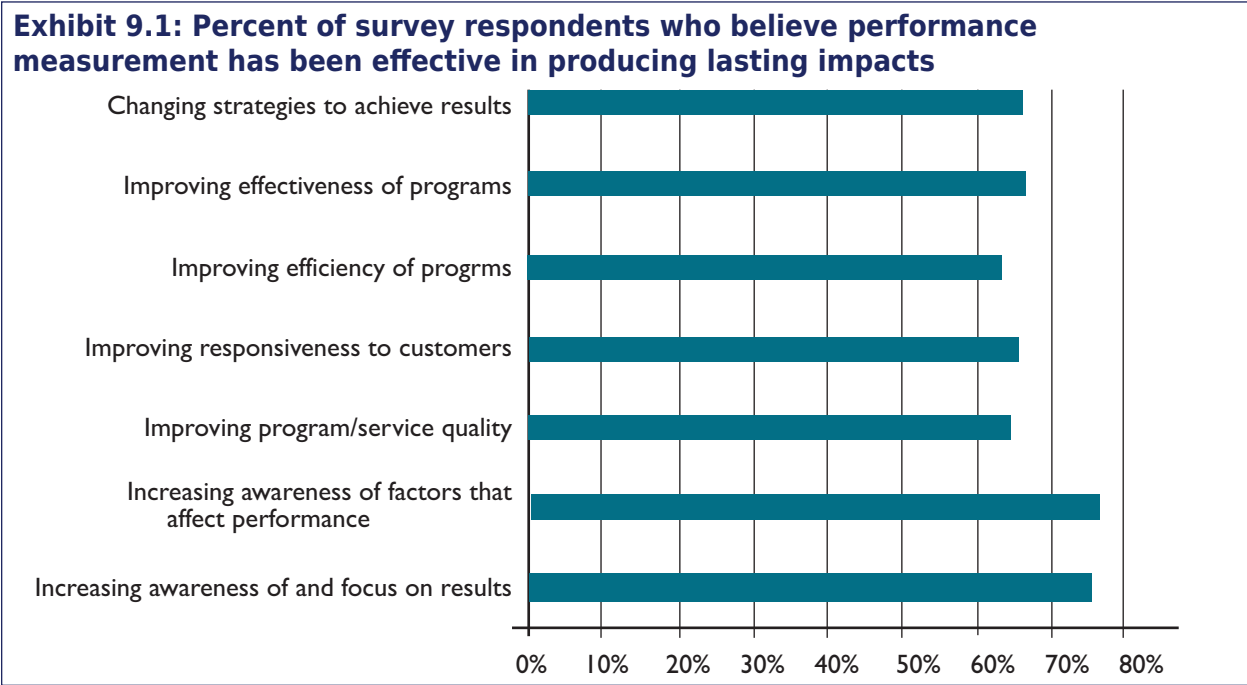
9. Conclusion: The Value of Performance Measurement

Perhaps unsurprisingly, none of our case study governments have conducted a formal return on investment analysis on performance measurement activities, but our surveys and interviews do provide a picture of our research subjects' views on the value of performance measurement. Our surveys and interviews covered two perspectives on value, including:

- The lasting impacts of performance measurement
- Overall value of performance measurement

The Lasting Impacts of Performance Measurement

Exhibit 9.1 shows the percent of survey respondents from across all case studies who believe that performance measures are effective or very effective for achieving the indicated lasting impacts. Our survey demonstrated that the respondents were most enthusiastic about the ability of performance measurement to increase awareness of and focus on results, and to increase awareness of factors that affect performance results.

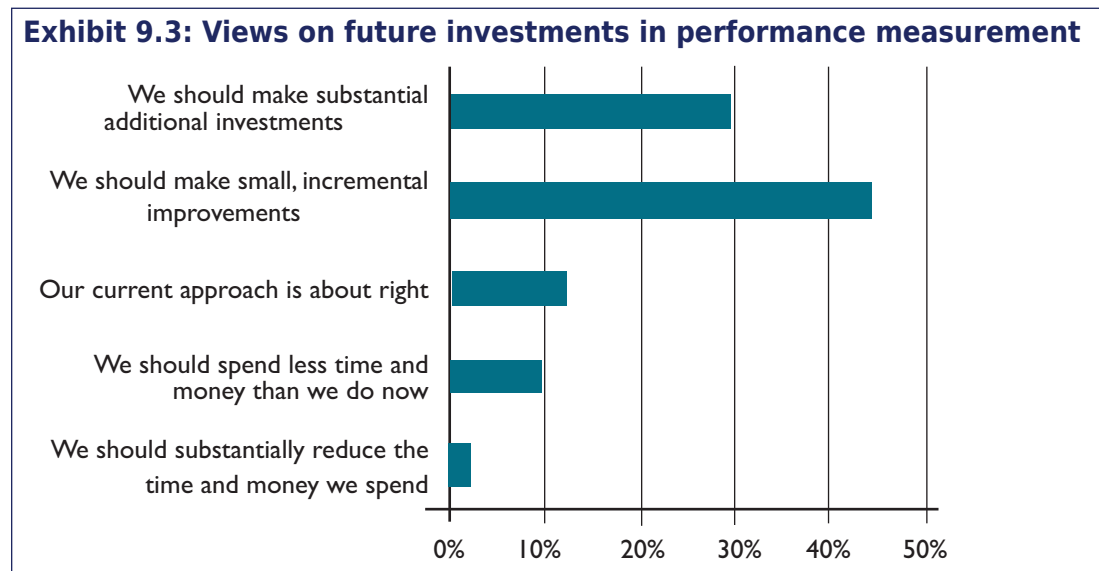
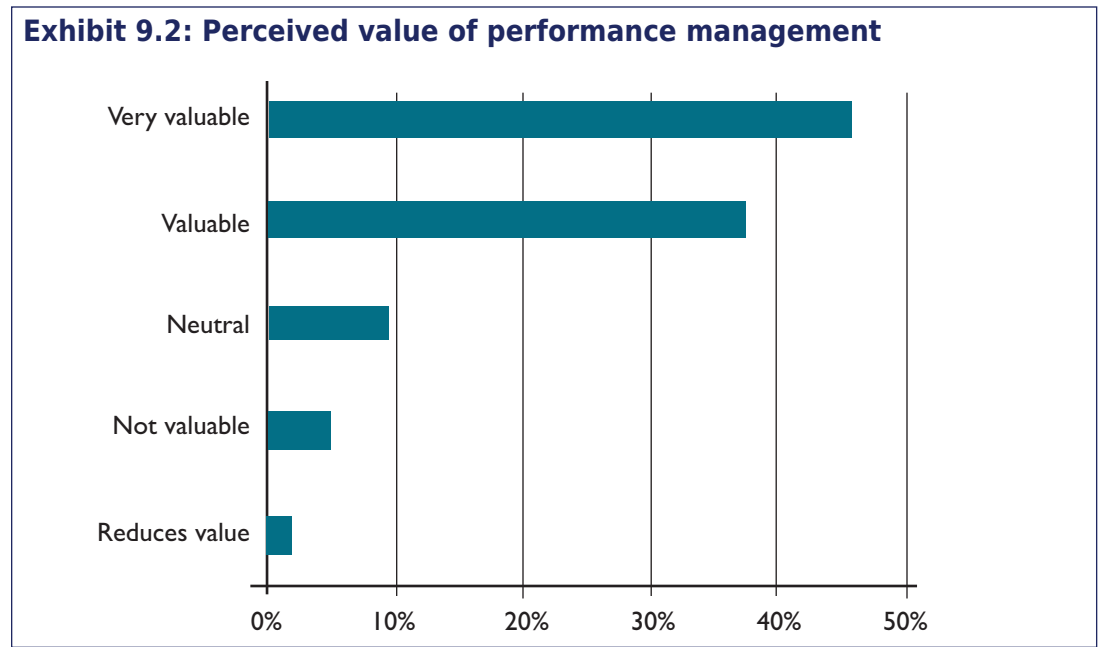


These results are probably not surprising given the enthusiasm shown in our case study governments for regular performance review meetings. Of course, a substantial portion of respondents also believe that performance measurements can create lasting impacts in the other areas our survey asked about! For example, Minneapolis improved program quality by using performance data to observe a natural experiment for what it terms the “chronic livability offender” (i.e., people that are chronic offend-

ers in misdemeanors like public drunkenness or trespassing). Active supervision by probation officers was implemented in the downtown neighborhood and was shown to reduce recidivism by 50% or more (perhaps even up to 70%). Based on these results, Minneapolis would like to expand the program to other neighborhoods.

Overall Value of Performance Measurement

Finally, the survey asked directly about the perceived value of performance measurement in two ways. Exhibit 9.2 addresses how valuable performance measurement is perceived to be, while Exhibit 9.3 asks what level of future investment the organization should make in performance measurement.



Both exhibits show that respondents were quite enthusiastic about performance measurement. The free form comments on the survey and interviews showed that respondents most appreciated the clarity brought to organizational priorities by measurement, the objectivity of the standards of evaluation, the support for operational improvement, and the mitigation of “political” influences over decision making. Here are a few quotes from the surveys and interviews that represent these opinions:

- *“Performance measurement numbers are invaluable as a conversation starter about performance issues” - Baltimore*
- *“Performance measurement provides the county with factual evidence of what is actually happening with a program or process so decisions can be made based on the facts not the politics” – Miami-Dade*
- *“It provides managers a very good barometer about how effective their operation is running and how well their services link to the city's mission, vision and values, and strategic goals.” – Irving*
- *“It's valuable to compare our results year over year to see if improvements are being made.” – Windsor*
- *“Historically, success or failure has been measured in subjective terms. Development of a systematic program of [performance measures] endorsed by council and clearly communicated to taxpayers will allow more objective measurement of success and provide for more effective decision making.” – Halifax*
- *“We are using measurement data to help support tough decisions. Measures are included in any business case now and some of the councilors really look for it.” - Winnipeg*

Value of Performance Measurement Checkpoints

- ✓ Our research subjects overwhelmingly believe that their experiences with performance measurement have been valuable. They also believe further investments should be made in measurement.
- ✓ The greatest value of measurement appears to be in raising awareness of and promoting constructive conversation about performance issues.
- ✓ The research subjects were least enthusiastic about performance measurement's impact on budgeting. However, the management/budget offices are all seeking to improve on this state of affairs.

Appendix 1 - Research Methodology

GFOA used the following sources of information for its research:

Interviews. Personal interviews were conducted over the phone with the following stakeholders for each case study:

- **Central management/budget office.** A series of three interviews was held with the management/budget office for each case study. The first interview covered basic orientation information about the organization. The second interviewed covered more detailed descriptive information about the research subject's practices. The third asked more evaluative questions about the effectiveness of the case study government's practices, including reactions to the results of GFOA's survey of the organization's staff (see below).
- **Elected officials.** The GFOA interviewed a sample of six elected officials from across the case study governments in order to get their views on the value that performance measurement and benchmarking provides to elected officials. The management/budget office helped the GFOA get in touch with the elected officials, but was not otherwise involved in the interview.
- **Department heads.** The GFOA interviewed a sample of 16 heads of operating departments from across the case study governments in order to get their views on the value that performance measurement and benchmarking provides to department heads. The management/budget office helped the GFOA get in touch with the department heads, but was not otherwise involved in the interview.
- **HR officials.** The GFOA interviewed HR officials from each case study government about individual performance appraisal practices and incentive pay practices.
- **IT officials.** The GFOA interviewed IT officials from each case study government about current and anticipated use of technology to support performance measurement.

Document review. The GFOA reviewed relevant documents for each case study, such as budget documents, strategic plans, and performance measurement websites.

Survey. The GFOA conducted a survey of the staff of the case study organizations in order to get their views on performance measurement. The survey recipients were selected by the management/budget office and, in most cases, represented management staff from throughout the organization who have a material role in the performance measurement system. The case study organizations felt that line staff were not sufficiently involved in performance measurement to be able to answer the questions. The table on the following pages summarizes the participation.

TOTAL - 491 respondents

By Department	Percentage
Public Safety/Justice	22%
Community Development/Cultural	6%
Central Services	30%
Public Works / Utility Services	35%
Legal/Policy Services	7%
Total	100%

Total

By Role	Percentage
Exec. Management	23%
Deputy Dept. Head	27%
Mid-level mgmt.	37%
Front-line Supervisor	4%
Non-supervisory staff	9%
Total	100%

Windsor - 63 respondents

By Department	Percentage
Public Safety/Justice	10%
Community Development/Cultural	5%
Central Services	33%
Public Works / Utility Services	44%
Legal/Policy Services	8%
Total	100%

Windsor

By Role	Percentage
Exec. Management	13%
Deputy Dept. Head	5%
Mid-level mgmt.	37%
Front-line Supervisor	14%
Non-supervisory staff	32%
Total	100%

Halifax - 50 respondents

By Department	Percentage
Public Safety/Justice	26%
Community Development/Cultural	4%
Central Services	34%
Public Works / Utility Services	26%
Legal/Policy Services	10%
Total	100%

Halifax

By Role	Percentage
Exec. Management	20%
Deputy Dept. Head	32%
Mid-level mgmt.	36%
Front-line Supervisor	4%
Non-supervisory staff	8%
Total	100%

Irving - 28 respondents

By Department	Percentage
Public Safety/Justice	21%
Community Development/Cultural	14%
Central Services	21%
Public Works / Utility Services	43%
Legal/Policy Services	0%
Total	100%

Irving

By Role	Percentage
Exec. Management	29%
Deputy Dept. Head	32%
Mid-level mgmt.	25%
Front-line Supervisor	11%
Non-supervisory staff	4%
Total	100%

Miami-Dade - 67 respondents

By Department	Percentage
Public Safety/Justice	22%
Community Development/Cultural	4%
Central Services	45%
Public Works / Utility Services	28%
Legal/Policy Services	0%
Total	100%

Miami-Dade

By Role	Percentage
Exec. Management	21%
Deputy Dept. Head	25%
Mid-level mgmt.	36%
Front-line Supervisor	3%
Non-supervisory staff	15%
Total	100%

Winnipeg - 22 respondents

By Department	Percentage
Public Safety/Justice	14%
Community Development/Cultural	5%
Central Services	27%
Public Works / Utility Services	50%
Legal/Policy Services	5%
Total	100%

Winnipeg

By Role	Percentage
Exec. Management	23%
Deputy Dept. Head	18%
Mid-level mgmt.	50%
Front-line Supervisor	0%
Non-supervisory staff	9%
Total	100%

Minneapolis - 41 respondents

By Department	Percentage
Public Safety/Justice	15%
Community Development/Cultural	15%
Central Services	22%
Public Works / Utility Services	37%
Legal/Policy Services	12%
Total	100%

Minneapolis

By Role	Percentage
Exec. Management	37%
Deputy Dept. Head	46%
Mid-level mgmt.	15%
Front-line Supervisor	0%
Non-supervisory staff	2%
Total	100%

King County - 19 respondents

By Department	Percentage
Public Safety/Justice	21%
Community Development/Cultural	11%
Central Services	53%
Public Works / Utility Services	16%
Legal/Policy Services	0%
Total	100%

King County

By Role	Percentage
Exec. Management	53%
Deputy Dept. Head	32%
Mid-level mgmt.	16%
Front-line Supervisor	0%
Non-supervisory staff	0%
Total	100%

Maricopa County - 201 respondents

By Department	Percentage
Public Safety/Justice	28%
Community Development/Cultural	3%
Central Services	24%
Public Works / Utility Services	35%
Legal/Policy Services	9%
Total	100%

Maricopa

By Role	Percentage
Exec. Management	22%
Deputy Dept. Head	29%
Mid-level mgmt.	44%
Front-line Supervisor	1%
Non-supervisory staff	3%
Total	100%

Appendix 2. General State of the Practice of Performance Measurement

The GFOA conducted a short survey of a random portion of its membership on the general state of the practice of performance measurement. Results indicated that the majority of our survey respondents use performance measures, but only 10% describe themselves as using performance measurement extensively (see table below). The survey also showed that about half have been using performance measurement for more than five years and 80% have been using it for more than two years.

We don't use performance measurement	38.37%
Performance measures play some role	51.16%
We use performance measurement extensively	10.47%

The table below describes the percentage of respondents who believe that performance measurement is effective or very effective for the indicated purpose. Interestingly, the general respondents tend to rate performance measurement as less effective than our case study governments. For example, almost 80% of the case study governments rated performance measurement as effective for monitoring progress on goals and demonstrating accountability to the public, compared to 66% of respondents to our more general survey.

Departments making budget requests	55.0%
Chief executive reviewing budget requests	51.7%
Chief executive developing recommended budget	53.3%
Legislative body considering proposed budget	41.7%
Demonstrating accountability to the public	60.0%
Deciding what level of service to offer	56.7%
Determining if a service is efficient	58.3%
Monitoring progress on goals / objectives	66.7%
Discussion of performance between departments & central management	51.7%
Evaluating the effectiveness of programs to meet their purpose	63.3%
Management and deployment of operational resources	51.7%
Evaluating individual employee performance	35.6%
Increasing awareness of and focus on results	57.9%
Increasing awareness of factors that affect performance results	64.9%
Improving program/service quality	49.1%
Improving responsiveness to customers	52.6%
Improving effectiveness of department programs	52.7%

We also asked about benchmarking. A majority said it was of great value or helpful, but still a slightly lower proportion than the case study governments.

Value of Benchmarking

Great value	16%
Helpful	43%
Useful	35%
Occasional use	4%
Rarely or never useful	2%

Use of Benchmarking

We don't benchmark	4%
We benchmark as needed	53%
We benchmark regularly, but don't belong to a formal cooperative or support group	27%
We benchmark regularly & belong to a formal benchmarking cooperative or support group	16%

Finally, we asked about the value of performance measurement. Some 75% of respondents still believe that they should increase funding for performance measurement, a result that is similar to that of the case studies. However, 30% of case study respondents believe the increased funding should be substantial, compared to only 22% of general respondents.