AGA
CORPORATE PARTNER ADVISORY GROUP RESEARCH

State and Local Governments' Use of Performance Measures to Improve Service Delivery

AGA CPAG Research Series:
Report No. 23
November 2009

Acknowledgements

About the Author

Harold I. Steinberg, CGFM, CPA, the writer of this report and the primary researcher for this project, is a retired partner from KPMG LLP, where he first headed the firm's state and local government practice, and then its federal government practice. He was also the first Acting Controller/Deputy Controller for the Office of Federal Financial Management, U.S. Office of Management and Budget.

Steinberg's initial experience in performance measurement was in Sunnyvale, CA, where he developed an approach for using performance measures to report the performance of local government services. He subsequently led the Financial Accounting Standards Board (FASB) research study, aimed at determining whether governments and other not-for-profit organizations could report their organizational performance using service efforts and accomplishments information, which culminated in FASB's publication of Reporting of Service Efforts and Accomplishments[1]; served as a member of the Governmental Accounting Standards Board (GASB) Service Efforts and Accomplishments Task Force that provided advice to the GASB staff for the development of criteria for reporting service efforts and accomplishments; and organized and directed the Association of Government Accountants' programs to evaluate federal agencies' Performance and Accountability Reports and state and local governments' performance reports against defined sets of criteria, provide recommendations for improving the reports, and publicly recognize the outstanding reports.

AGA wishes to thank the members of the Advisory Committee established for this project for their help and contribution to the research: **Katherine Barrett,** Barrett and Greene, Inc.; **Barbara J. Cohn Berman,** National Center for Civic Innovation; **Geoff DePriest, PMP,** Crowe Horwath LLP; **Sharon W. Erickson, CGFM, CIA,** City of San Jose; **James R. (Jay) Fountain Jr., CGFM, CPA,** Governmental Accounting Standards Board; **Richard Greene,** Barrett and Greene, Inc.; **Harry Hatry,** The Urban Institute; **Michael (Mike) Lawson,** International City/County Management Association Center for Performance Measurement; **Anthony (Tony) Longo,** City of New York, NY; AGA Past National President **Sam M. McCall, Ph.D., CGFM, CPA, CIA, CGAP,** City of Tallahassee; **Bert Nuehring, CPA,** Crowe Horwath LLP; **John J. Radford, CGFM, CIA, CFE,** State of Oregon; **Jonathan Walters,** journalist, and the following: **Eveanna Barry,** Director of Performance Reporting, AGA; **Marc Holzer,** Rutgers University; and **Anna D. Gowans Miller, MBA, CPA,** Director of Research, AGA.

Corporate Partner Advisory Group Leadership

Chairman

Hank Steininger, CGFM, CPA

Managing Partner, Global Public Sector, Grant Thornton LLP

Vice Chairman

John Cherbini, CGFM, CPA

Partner, KPMG LLP

AGA Professional Staff

Relmond Van Daniker, DBA, CPA

Executive Director

Anna D. Gowans Miller, MBA, CPA

Director of Research

Susan Fritzlen

Director of Corporate Partner Program

Marie Force

Director of Communications

Christina Camara

Publications Manager

AGA is Proud to Recognize the Firm Supporting this Effort:

Crowe Horwath™

AGA's Corporate Partner Advisory Group Research Program: Building the Bridge Between Government and Industry

AGA's Corporate Partner Advisory Group (CPAG), executive director and director of research are creating research projects of value to governments, industry and the entire AGA membership. These studies are expected to result in reports assessing current and/or best practices and make recommendations for future improvements in federal, state and local governmental accounting, auditing and financial management. CPAG members support AGA research through either cooperative or sponsored research projects. "By undertaking research, AGA is fulfilling its mission as a thought leader in advancing government accountability," said AGA Executive Director Relmond Van Daniker, DBA, CPA. "This is one of numerous research initiatives that will benefit government and bridge the gap between the public and private sectors."

The CPAG was organized in 2001 as a business element within AGA. The mission of the CPAG is to bring industry and government executives together to exchange information, support professional development, improve communications and understanding, solve issues and build partnership and trust, thereby enhancing AGA's focus on advancing government accountability. Corporate member involvement in the CPAG is limited to organizations that sign up for the AGA Corporate Partner membership program. For more information on the research program, please visit *www.agacgfm.org/research/default.aspx* or contact Anna Miller at *amiller@agacgfm.org*.

Table Of Contents

Executive Summary

This project explores state and local governments' use of performance measures to monitor the delivery of services and make adjustments, if necessary, to improve the delivery and assure the achievement of desired results. The project is also intended to determine how performance measures are used for improving service delivery and describe their efforts in such a way that other governments can adopt the practice.

The project had three parts: identification of the elements of performance management used to improve service delivery; development of case studies describing five governments' successful use of performance measures to improve service delivery; and execution of an online survey to determine the extent to which the performance management elements are used.

Key Findings

Government units of all sizes and types are using performance measures as a tool for improving service delivery. Large governments and their component agencies typically operate systematic, structured processes. Smaller government entities are often able to achieve similar results by adhering closely to basic principles. In both cases, success appears to depend on the commitment and involvement of a chief executive who sees the process as a tool for improved performance, and not just a compliance activity. Throughout the text of the report, some acronyms, words and terms are in bold face.

The process is typically started with the selection and use of **output measures** (that is, measures of the *quantity* of services provided) to monitor the delivery of services. The ideal would be to select and use **outcome measures** (that is, measures of the *results* associated with the provision of services), but there appears to be a general reluctance to do so. Government personnel often believe they do not completely control outcomes, and outcome data can be difficult to obtain. The key, therefore, is to use consistent measures from period to period in order to sustain attention to the process, while recognizing that measures can and should be modified as necessary to reflect what all stakeholders, especially citizens, want from the programs.

Other key success factors include:

- regular and frequent analysis of the performance results data in comparison to prior periods, targets or benchmarks.
- regular reviews of the analysis and results by the chief executive and/or his designee with the responsible agency heads.
- agreements with the agency heads on steps to be taken when the data reveal the need and opportunities for improvement.

Examples of changes in service delivery and concomitant service improvements resulting from analysis of the data received and discussions with personnel in five governments interviewed, are presented in detail in this report.

Recommendations

The research identified several elements that would need to be present in a system to use performance measures to improve service delivery. Some elements are essential to the process of using performance measures to improve service delivery; some support the process, but it can still be effective without them; and some are ancillary to the process. Ancillary is used to mean that although the elements are considered to be integral to performance management and could enhance a process to use performance measures to improve service delivery, they were not always used by the governments interviewed and surveyed. The elements in all three categories are identified and explained further in the report's recommendations section.

Note: Bold-faced acronyms, words and terms are defined in the Glossary of Acronyms and Terms on page 17.

Introduction

Concern with governments' performance has existed for a long time. As early as the beginning of the 20th century, groups such as the New York Bureau of Municipal Research suggested that governments should report to citizens not only in financial terms, but also on the performance of governments programs, activities and functions. In the 1930s, the **International City/County Management Association (ICMA)** championed "municipal reporting" as a way to directly or indirectly report to citizens on a government's accomplishments.

Although interest in applying measures to a government's programs has been uneven over the years, several recent developments are again helping to focus attention on performance measurement. Examples include the **Governmental Accounting Standards Board's (GASB)** recent issuance of suggested guidelines for voluntary reporting of performance information; the enactment of the American Recovery and Reinvestment Act (ARRA) and its requirement for performance information, particularly about jobs created or retained; the establishment of a **National Performance Management Advisory Commission** and its release of a performance management framework for state and local governments cited in Appendix C; the work of ICMA's Center for Performance Measurement (ICMA-CPM) and other regional performance measurement consortia; the visibility given New York City and other governments' **Stat** performance measurement systems; countless articles in *Governing* and Pew Center[2] reports; and work in other countries such as Canada, United Kingdom, New Zealand and Australia.

There are three major uses for performance measures in government. They can:

- Demonstrate accountability to citizens, elected officials and other interested parties by reporting the service efforts and accomplishments (also identified as performance and results) associated with a government's programs and activities.
- Improve the allocation of resources by using performance information to inform the development and enactment of a government's budget.
- Assure the achievement of desired results by using performance information to monitor the delivery of services and make adjustments, if necessary, in the service delivery.

AGA conducts research in areas that will be helpful to its members and others who are involved in governmental financial and performance management. Since the increasing focus on performance measures seems to concentrate on the public reporting piece, AGA and the sponsor of this research, AGA Corporate Partner Crowe Horwath LLP, decided that a research project focused on the third major use—to ascertain how performance measures can and are being used by state and local governments to improve the delivery of services and achievement of results—would spur other governments to consider adopting the process.

This research project identifies and explains the various elements in using performance measures to improve delivery of programs and services. It then presents the results of interviews with five jurisdictions that already operate such systems and an online survey of persons associated with state and local government management. This information sheds light on the extent to which and how the identified elements are used. (The research project's methodology is presented in Appendix A.)

Some call this process performance management. Although the terms performance management and performance measurement often are used interchangeably, performance management is much broader. It encompasses such elements as strategic planning, performance budgeting and performance reporting. As indicated, this project is limited to the use of performance measures to improve service delivery.

It should also be noted that the federal government has been implementing performance management for several years, pursuant to the requirements of the **Government Performance and Results Act** and the **Program Assessment Rating Tool**, a management initiative established by the executive branch. Federal agencies have made tremendous strides in making performance results information available to the public and others. There is concern, however, that the information, while widely available, is not used by management or other decision makers. Although this research was limited to state and local governments' use of performance measures, the recommendations for using performance measures to improve service delivery in this research report could be equally useful to federal agencies.

Key Findings

The following findings about the use of performance measures to improve the delivery of services were developed during the three parts of the project: building the list of elements, conducting interviews in five jurisdictions, and reviewing the results of the online survey. The first set of findings relate to the process of using performance measures to improve service delivery; the second are examples of actual improvements made by each of the five interviewed jurisdictions as a result of using the process.

- Performance measures are used to improve service delivery at both the state and local levels.
- The commitment of the government's executive leadership is essential to the process. Unless senior levels of management use the measures to monitor and drive performance changes, the effort will not produce the desired result of improving service delivery to the citizens and other constituents.
- Larger jurisdictions, with more resources available, tend to operate more formal programs to use performance measures to improve service delivery. However, this does not mean governments should throw money at the program. To the contrary, many small but agile governments are leaders in using measures to review performance and make decisions to improve.

- The process should not be used for assigning blame or finding fault with others' performance. This is especially important since one person's desire for accountability can be perceived by others, and particularly those on the front line of service delivery, as merely pointing fingers.
- Although a government might conclude it obtained better service delivery or a process improvement from collecting and using performance data to trigger a change in the way the program operates, it needs to ascertain the explicit change in inputs, outputs, and/or outcomes to confirm that the improvement actually occurred.
- Governments typically start the performance measurement process with output measures recognizing they cannot completely control the outcomes (that is, there are factors beyond the service itself that influence the desired results), and eventually move toward using outcome measures. The acceptance of outputs rather than outcomes may be bemoaned by some performance measurement experts, but is acceptable, particularly when the practicalities for those responsible for initiating the process prevent starting with outcome measures.
- Consistency, that is, the continuing use of the same measures, is important to sustain attention to the process.
- The selection of performance measures is typically iterative, which causes the need for balance between the need to update to a more appropriate measure on one hand, and consistency on the other hand.
- Comparisons with prior periods' performance tend to get more attention than comparisons to targets.
- Reporting performance results to the public and oversight bodies, as well as using the measures internally, appears to stimulate a desire for performance improvement.
- **311 systems** and other complaint systems are useful to governments for gathering information to act upon.
- The chief executive's use of performance measures for the evaluation and recognition of department heads can be an effective motivator, particularly if incentives, such as staffing flexibilities, additional equipment and bonuses, can be partially based on meeting performance targets and goals.
- To assure continuation of the use of performance measures for improving service delivery beyond the tenure of the chief executive who introduces it, the systems must be driven down into the organization such that the processes are in place, and the people want to continue using the processes.

The following are examples of the results of using performance measures to improve service delivery gathered from the five interviewed jurisdictions. Descriptions of additional results in the five jurisdictions are included in Appendix B, which presents case studies drawn from the five interviews.

- ***New York City***—The streets are regularly inspected and rated for cleanliness. Measures that showed a low level of cleanliness stimulated department management to staff the cleaning activity with a minimum number of people, authorizing overtime if necessary, rather than attempt to clean the streets with people left over from the collection activity. The results, as verified by the ratings, were cleaner streets.
- ***Baltimore***—Trash was identified as building up in front of residences. Two alternatives were adopted: increasing the emphasis on the issuance of citations to the landlords, rather than the tenants, in order to get the landlords to assume responsibility to work with their tenants to reduce trash; and establishing protocols for alerting public housing managers to be alert to the use of their Dumpsters by nearby private residence occupants. The results were less trash build-up.
- ***Westminster***—For Colorado's municipalities, the primary revenue source for general governmental operations is the sales and use tax. For the City of Westminster, it provides 64 percent of general fund revenues. The Sales Tax Division's analysis of the revenue per audit hour resulted in a shift of resources from auditing non-construction entities, which returned $167 per audit hour, to more audits of construction entities, which returned $588 per audit hour.
- ***Fishers***—The town decided snow removal performance, as revealed by comparisons to the National Citizen Survey (NCS), was not satisfactory. It added contractual services to clear the secondary streets, redesigned the routes to avoid clearing unnecessary roads, and provided contractors with maps to assure they had the correct routes. Secondary street residents expressed satisfaction with the results.
- ***Washington State***—Performance data revealed that, while overall traffic accidents on state highways were declining, motorcycle accidents were increasing. This led several agencies (Department of Licensing, Department of Transportation, Traffic Safety Commission, Washington State Patrol) to work together to provide more education for manufacturers and operators and stricter licensing, resulting in fewer accidents.

Recommendations

The research showed that the process of using performance measures to improve service delivery includes several components. Some of these components appear to be essential to the success of improving service delivery through performance measures. Other components appear to support the process, without being critical. A third set of components relating more to performance management (a broader topic than performance measurement) appear to significantly enhance the success of improving service delivery but were generally not referenced by participants in the interviews or survey. Within this report, these components are categorized as essential, supporting and ancillary elements. Individuals with extensive experience working with performance measures may use different terms or include different components within these categories. The categorization in this paper is based directly on the input from interview and survey participants, and is believed to accurately reflect how these components are currently being used by state and local government.

Essential Elements

The following elements have been identified as essential.

- **Chief executive commitment and involvement in the overall process.** The chief executive must be committed to and involved in the process. Recognizing the demands on a chief executive's time and attention, he (or she) can exhibit this commitment through a single, senior official that the department heads and others know speaks for the chief executive. The chief executive must back that individual's actions completely and clearly and not accept attempts to circumvent that individual's decisions and actions.
- **Relevant measures of at least outputs and eventually outcomes.** Relevant measures must be selected with which service delivery can be measured, analyzed and improvements sought. The initial measures can be of outputs, but measures of outcomes should be selected and used as soon thereafter as possible. Also, the measures can be selected either by the chief executive, or most likely his designee, or the agency heads with approval of the chief executive (or his designee). Eventually, the service recipients should also be consulted to assure the measures reflect what they seek from the services.
- **Periodic review and revision of the performance measures.** Notwithstanding that the same measures should be used year-to-year in order to sustain attention to the process and provide a basis for ascertaining trends, the measures should be periodically reviewed and revised as necessary to assure they reflect changing expectations for the programs.
- **Frequent, regular collection of performance data.** Results data for each performance measure should be regularly collected during the year in order that the data can be analyzed and adjustments made to enable the programs to continue to perform as expected.
- **Comparison of performance data to prior periods.** The regularly collected performance results data should be compared to at least the prior period to ascertain whether performance is improving or at least remaining stable; declining a minimal, but acceptable amount; or declining more than an acceptable amount.[3]
- **Regular review and analysis of performance results to ascertain the reasons for less-than-desired performance and identify the opportunities for improvement.** Comparisons of data that reveal performance is less than desired or declining at more than an acceptable rate should trigger an analysis of the results. The analysis should include determining the reasons for the less-than-desired performance. At the same time, the instances of better-than-expected and/or improving performance should be noted and the reasons ascertained in order that effective processes can be replicated.
- **Chief executive and other senior management participation in the reviews.** The chief executive (or his designee) and other senior management should be personally and directly involved in the reviews of the performance results. The reviews should culminate in agreements to adopt courses of action to reverse the declining performance and/or achieve the desired results.
- **Agreement between chief executive and department managers on improvement plans.** Explicit agreements need to be secured from the department heads that they will undertake the plans embodied in the arrived-at decisions. Time frames should be included in the agreements.
- **Follow up on progress (or the lack thereof) of improvement plans.** There must be an organized and formal follow-up subsequent to the agreement to assure the plans are implemented and progress toward the expected results are achieved. The follow-up should be conducted by the chief executive or his/her representative.

Supporting Elements

The following are the elements that many of the interviewed and surveyed governments included in the process but, due to less-than-universal use, appeared less necessary.

- **Explicit targets for the performance measures.** As long as the primary concern is with the direction of performance, as indicated by comparisons of results with prior periods' results, there appears to be less of a need to establish targets for performance. Nonetheless, for some measures, targets might be set in order to define desired degrees of improvement.
- **Frequent comparison of performance data to targets.** The absence of targets for performance precludes the ability to compare performance to targets. However, if there are targets, comparisons should be made.
- **Comparison of performance data to the corresponding data from similar jurisdictions.** States frequently compare their performance to like states. Also, some local governments belong to consortia that collect and provide data with which the local governments can compare themselves to one another. These inter-entity comparisons, however, are only the third most prevalent type of comparison.
- **Intragovernmental comparisons (in states and other larger jurisdictions).** Larger governments, which would include states, frequently disaggregate performance results data geographically, demographically or otherwise to identify weaker performing segments on which to focus the analysis. They also use the information to spur improvement in the poorer performing segments.
- **Support staff involvement in the reviews of performance results.** The initial review and analysis of results is generally performed by a person or organization that reports directly to the chief executive. That person, as well as other support staff—for example, budget, personnel, IT, supply—are then available to participate in meetings the chief executive holds with department heads to review performance.

- **Budget reviews and deliberations considering the performance targets and results.** The establishment of targets and the achievement of results can be used to influence budget decisions. However, the successful use of performance measures to improve service delivery does not require that budgets be established based on planned performance.[4]
- **Budget resources allocated to programs rather than solely to object classes.** Budget resources are typically appropriated first to organizations, and then to **object classes** (for example, salaries, supplies, travel) within the organizations. Appropriating to programs within the organizations helps to keep the focus on program performance.
- **Programs for obtaining and considering constituents' views.** Periodic surveys, seeking feedback on specific issues, focus groups, complaint/compliment tracking systems, town meetings, etc. can effectively complement the more objective data the government obtains from measuring performance results. They also enable the government to know how the customers view the quality and quantity of the services.
- **Some process for assuring the data's reliability.** Although most governments use internal management reviews to assure the reliability of their performance results data, this is likely to be little more than a "sniff test." Using a more structured process, such as reviews of the controls underlying the performance data collection systems or external reviews of selected data, would increase confidence in the reported performance results.
- **Regular external dissemination of performance results.** Regular reporting of performance results, based on the performance measures, to the legislature or other governing body and/or the public prompts the elected and appointed officials and citizens to push for better performance. It also stimulates the government's employees to improve their performance.

Ancillary Elements

As stated, the project's first step was to prepare a list of all the elements that could possibly be appropriate for a process intended to use performance measures to improve service delivery. The list was exhaustive. However, after the Advisory Committee's review, plus the interviews and online survey, it was realized that some of the elements were generally not used in the process, or that the process being used had not yet matured to a point where the element would be given higher priority. In this report, they are termed 'ancillary.' They are:

- **Strategic plan for the jurisdiction or departments.** The governments and/or departments might have a strategic plan, but it apparently does not drive the establishment of the performance measures used to determine and improve service delivery.
- **Legislative branch involvement in the process.** Since performance improvement focuses on the day-to-day functioning of government, there may be little need for legislative involvement. In fact, it appears that although legislators are inherently interested in programs' outcomes, they generally appear to have little direct participation in the process other than instances where performance data are cited during budget hearings.
- **Cost accounting system and cost of services.** Maximizing efficiency and effectiveness requires knowledge of the costs of services and outputs/outcomes. This, in turn, requires cost accounting systems, or at least end-of-the-year cost finding techniques. However, since service improvement is frequently not cost-dependent and cost data are absent in so many governments, this element is currently not essential. On the other hand, governments faced with a tight economic environment and limited resources might deem this element to be highly desirable.
- **Personnel evaluations, promotions, and compensation adjustments consider program performance.** The existence of externalities, the possibility of counter-productive results, the lack of experience with credible evaluation processes, plus the influence of labor unions, precludes attempts to establish a relationship between promotions and compensation adjustments and program performance/improvement. Basing personnel evaluations on measurement-based program performance is currently generally limited to evaluations of the department heads and other senior personnel.
- **External stakeholder involvement in the process.** External stakeholders (taxpayer groups, leagues of women voters, municipal research organizations, for example) are concerned with results, but have less concern over how the results are obtained. Their involvement, which might not occur until after the process is under way, would be to confirm that the government has identified and is measuring the results they want.

Summary of Case Studies

Five Jurisdictions' Approaches to Using Performance Measures to Improve Service Delivery

As stated, the findings are partly based on interviews conducted with five governments, differing in type and size, that already use performance measures to improve service delivery. The purpose of the interviews was to ascertain what these governments do and how they do it in order to pass this information on to other governments interested in using performance measures to improve service delivery. The responses were noted, but not independently verified.

This section summarizes each of the interviews and provides a name, phone number and e-mail address for obtaining additional information. It then provides overall observations resulting from the interviews that should be considered by governments wishing to initiate the process. Detailed descriptions of the history, organization, process and elements of the programs to use performance measures to improve service delivery in each of the five governments are presented in Appendix B.

New York, NY

The process in New York grew out of a 1977 City Charter revision requiring the mayor to publish a bi-annual report on performance of the city's services. The initial reports focused on services for which the agency managers already had data. A 2002 restructuring of the process limited the amount of information each agency could present in the report, replaced their narratives with measurable performance results data, and required, for declining performance, explanations of why and what would be done to reverse the trend.

In 2006, the city adopted the Citywide Performance Reporting (CPR) program. This program is directed by the Office of Operations, which reports directly to the mayor. It uses 530 customer-oriented outcome measures (as well as additional efficiency measures) that were selected collaboratively by the Office of Operations and each agency, but with the final say residing with the former. The connection between the performance measures and union contracts necessitated some involvement from the Office of Labor Relations and the unions.

The heart of CPR is an online interactive dashboard that groups the outcome measures into government functions and presents information about performance, trends, meaning of the measures, etc. Each agency head is required to print his or her agency's report bi-weekly and submit it to the deputy mayor, and then to meet with the deputy mayor to review remedial action plans for measures that declined more than 10 percent from the prior period.

The dashboard is available to the public at *www.nyc.gov/html/ops/cpr/html/home/home.shtml*. This accessibility, the detail in the dashboard, and its ease of use enable the public to monitor the city's performance. This is a major driver for improving performance.

Contact: Tony Longo, Office of Operations

212.788.1677

tlongo@cityhall.nyc.gov

Baltimore, MD

The program to use performance measures to improve service delivery in Baltimore started in 2000 following the election of a new mayor, who was advised by a veteran of the New York City program. It is called CityStat and is directed by the deputy mayor for administration.

There are two key elements in CityStat:

- A template for each department that lists performance and other data for the past two weeks, several prior two-week periods, several prior years and percentage changes for both. The templates are prepared by a CityStat analyst working under the direction of the deputy mayor for administration.
- Bi-weekly meetings held by the mayor and/or deputy mayors to review the performance of individual departments. The mayor/deputy mayor is prepared for each meeting with a 10-page memo written by a CityStat analyst. The memo provides, for selected issues, the background of the issue, performance data and questions to ask. Following the meeting, the CityStat analyst sends a one-page memo to the department head listing follow-up actions and responsibilities agreed to at the meeting.

The program was recently expanded to include meetings of several department and other agency heads involved in a policy area—for example CleanStat, which focuses on cleaning up the city; or GunStat, which focuses on reducing gun violence—and sometimes includes persons from outside the Baltimore government, such as representatives of the state attorney's office and federal law enforcement officials.

Contact: Christopher Thomaskutty, deputy mayor for administration

410.396.5176

chris.thomaskutty@baltimorecity.gov

Westminster, CO

Westminster initiated a program to use performance measures in 1999. It used mostly workload measures, which were not used to drive management decisions. In 2001, a new city manager arrived who, among other things, wanted to use performance measures to improve service delivery. He had the city join ICMA's Center for Performance Measurement.

The process is driven by a senior management analyst, who is assisted by a performance measurement team, chaired by the senior management analyst. The team is composed of at least one individual from each agency's management structure (for example, senior analyst, project coordinator, deputy agency head). Its role is to monitor the performance measure reporting process; share experiences in driving the process down within the agencies and linking the agencies' performance measures to the city's strategic plan; identify training needs; and work with the regional performance measurement consortium.

Each agency is expected to establish performance measures consistent with the city's strategic plan. A "revamp" process reviews two agencies' measures a year and, if necessary, updates the measures. Instead of regularly reviewing the performance results for each of the agencies with the agency heads, the manager expects the agency heads to use the performance measures to manage and improve municipal services. Also, each agency head must prepare a report to the manager each year describing how his or her agency uses performance measures to better understand, justify or improve operations.

A course titled "Performance Measurement 101" is given to all new employees within one year of joining the city. It is a half-day course, the purpose of which is to improve employees' understanding of performance measurement. It does so by describing what performance measurement is, why the city is doing performance measurement, and how city agencies are doing performance measurement.

Contact: Aric Otzelberger, senior management analyst

303.658.2004

aotzelbe@cityofwestminster.us

Fishers, IN

A program to use performance measures to improve service delivery started in Fishers in 2005 upon the arrival of a new town manager. He had participated in ICMA's Center for Performance Measurement in his prior position. The deputy town manager was assigned the primary coordinator position and oversees the process.

Initially, the city used all the measures included in ICMA's lists of suitable performance measures that the town believed were applicable. It subsequently recognized that many of the measures were not appropriate for Fishers. The primary coordinator met with each agency head and asked him or her to select measures based on what they believed is important for the agency to compare to ICMA-CPM's national median. The police and fire departments also considered measures they used for their national accreditation programs.

Like Westminster, the city does not perform centralized reviews and analysis of performance results. Rather, the use of performance measures to improve management is left to the agencies. The primary coordinator meets individually with representatives of each service area and reviews the data, but more to increase his understanding of the measures than to ascertain the levels of performance.

The agencies are instructed to educate new employees in performance measurement.

Contact: Nathan George, deputy town manager

317.595.3117

georgen@fishers.in.us

State of Washington

The program to use performance measures to improve service delivery started in 2004, upon the election of a new governor and issuance of an executive order requiring each agency to develop clear, relevant and easy-to-understand measures that show whether programs are successful; hold regular problem-solving sessions within the agency to improve performance; and regularly report to the governor on performance. These requirements have since been embodied in legislation. Overseeing the program is the Office of Government Management Accountability and Performance (GMAP), which is located within the Governor's Executive Policy Office.

Performance measures were initially developed in six policy areas (economic vitality, government efficiency, health care, safety, transportation, and vulnerable children and adults) by measurement teams led by a GMAP staff person and composed of people from the agencies who work in the policy area and representatives from the governor's policy and budget offices. These measures were then proposed to the agencies as "straw men." To provide assurance that the measures reflect what is important to the citizens and that the agencies are focusing on the right goals, the state held citizen meetings, workshops and town hall sessions throughout the state in 2006 and 2007. The assurance was obtained by asking the participants, "What is important to you?" "How would you know if you got it?" and "What information do you need to judge whether your state government is working?"

At the heart of GMAP are meetings, called GMAP Forums, of the governor and her leadership team (chief of staff, deputy chief of staff, policy director, budget director, GMAP chief, legislative director, directors of the Departments of Personnel, Information Services and Risk Management) with the agency directors involved in each of the policy areas. The forums, which are held regularly and are open to the public, review the past quarter's progress toward achieving specific results.

To prepare for each forum, a measurement team, comprised of a lead GMAP analyst, analysts from each involved agency, and analysts from the Office of Financial Management, meet two or three times and draft a GMAP performance report. These reports present three types of information: data for each performance measure; an analysis of performance that is prepared primarily by the agency and includes comparisons with targets, prior periods and other states; and action plans that show how performance will be improved. The action plans list what will be done, who is responsible to get it done and when it will be completed.

After the forum, the GMAP analyst sends the agency directors a follow-up memo, listing the items the agencies will have to report back on before or during the next forum.

The agencies also operate internal GMAPs, pursuant to a requirement in the Governor's Executive Order.

Contact: Jill Satran, Office of Government Management Accountability and Performance

360.902.0849

jill.satran@gov.wa.gov

Overall Observations

The following are overall observations of the way a successful program uses performance measures to improve service delivery. They are drawn from the five interviews and the subsequent development of the case studies appearing in Appendix B.

- The successful use of performance measures to improve service delivery is generally due to the desire of a forward-looking chief executive.
- The involvement of the chief executive and his or her deputies in the process increases their understanding of how the departments work and how the services are delivered and performance is monitored. The involvement also greatly increases the attention of the departments' leadership.
- The persons driving the process tend to be long-term public administration, private sector, or not-for-profit professionals (but not necessarily part of the jurisdiction's civil service). Experience in the budget office can be useful.

- As much as governments exist to produce outcomes, the systems tend to start with output measures, then evolve to outcome measures.
- The measures change periodically to reflect changing realizations of what is important to the public.
- The jurisdictions provide their employees with training on performance measures and their use.
- The budgets appear not to be linked to performance. Monies are generally appropriated to agencies and objects rather than to programs for which measures would be established.
- There is not much use of cost measures.
- Program performance is not tied to employee compensation, although it is frequently tied to advancement.
- The governments do not appear to be overly concerned about the reliability of the data. They rely mainly on internal management reviews of the data; constant working with the data to provide a sense of its reliability; public availability of the results, which enables others to question numbers they believe are problematic; and a belief (even if not justifiable) that the numbers cannot be falsified significantly.
- Legislatures do not exhibit much interest in using performance measures to improve service delivery.
- It appears that, despite the availability of the performance data, most of the jurisdictions do not issue service efforts and accomplishments reports to their citizens.

Summary of Survey Results

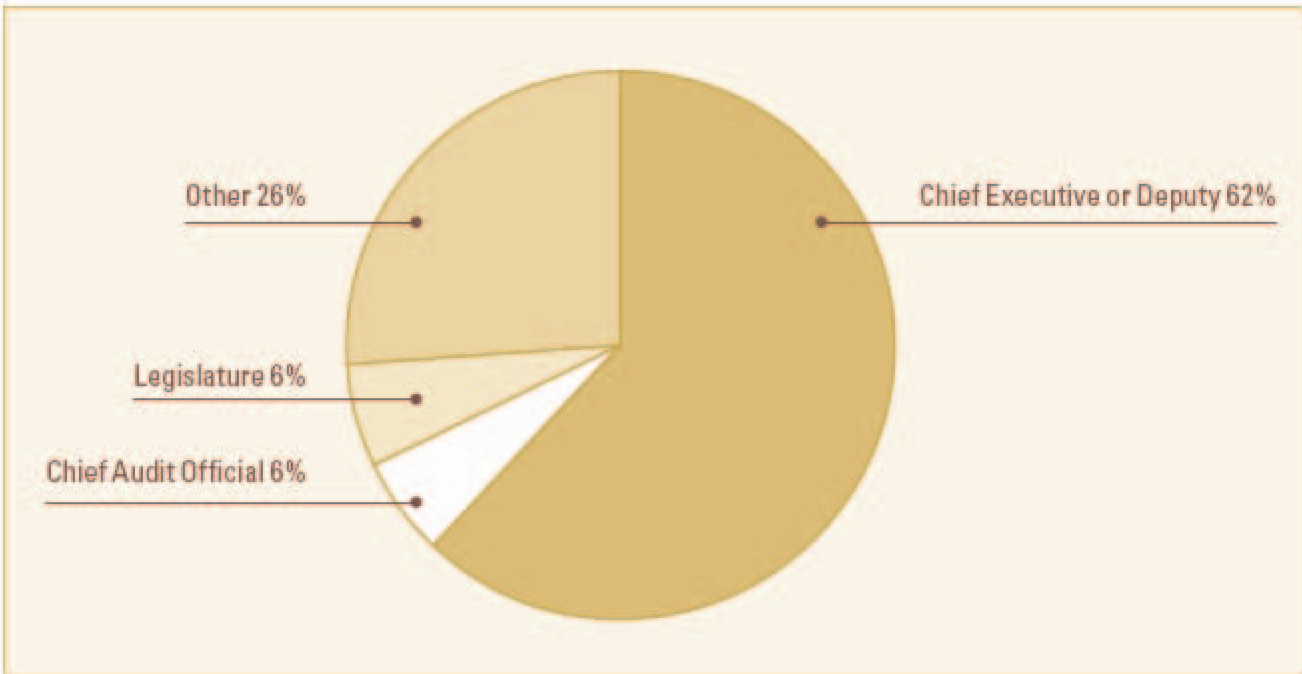
The findings are also based on an online survey constructed to determine the extent to which state and local governments use the elements of the process for using performance measures to improve service delivery. The survey generated responses from 253 persons, 175 of whom answered that the government for which they were responding uses performance measures to evaluate the effectiveness of government programs and services and improve their delivery. While the responses do not represent a statistical sample, they provide an overall sense of how performance measures are used to improve the delivery of services.

Basic Information

As stated, 175 respondents answered that they use performance measures to evaluate the effectiveness of government programs and services and improve service delivery. It appears that slightly more of the respondents were from states or state agencies than were from counties, cities, and other general and special purpose local governments. The forms of governance for the general-purpose local governments were also evenly divided among strong mayor, council-manager and commission.

Of the 175, 149 also reported they use performance measures to report accountability to the public, 144 reported they use performance measures to inform budget preparation and enactment, and 70 reported they use performance measures for other purposes. Examples of the other purposes are establishing performance contracts, providing awards

Figure 1: To Whom Does the Unit Report?



and other recognition, and operating gain-sharing programs.

Eighty stated they started using performance measures to improve the delivery of services since 2001; 49 at an earlier date; 44 do not know when the government started the process; and two said they do not use performance measures. Almost half stated the reason the process started was the chief executive's desire and almost half identified a legal requirement-statute or ordinance or legislative body mandate.

Organizational Matters

Two-thirds of the respondents answered that a unit within the government oversees the use of performance measures to improve the delivery of services. Many of the units' names suggest a connection to the government's budget office. However, some carry such names as Office of Strategic Planning and Performance Management, Performance Management Unit, and Management Performance and Accountability, which suggest they could be operating separate from the budget office. Eighty-six respondents answered the question on whom the unit reports to: 62 percent said the chief executive or his/her deputy; 6 percent said the chief audit official; 6 percent the legislature; and 26 percent others such as a chief operating officer, budget officer or chief financial officer. See *Figure 1*.

Fifty-three percent of the 88 persons responding to a question about the status of the person heading the unit said that the person is in the career service, with 38 percent being political appointees and the rest elected officials. This differs from the interview results, which showed the persons heading the unit were appointees. See *Figure 2*.

Slightly more than one-third of the individuals responding to a question about the length of time they headed the unit stated two to four years; slightly less than a quarter responded less than two years, 10 percent, four to six years; and 28 percent, more than six years. Eighty percent of these individuals' prior experience was in government, mostly the budget function, but also general management, departmental management, performance management and audit. The other 20 percent's experience was in the private sector.

The units vary in size: Of the 77 governments that answered the question, 25 have one to three persons, 18 have four to six, 11 have six to 10, 19 have 10 to 25, and four have more than 25 persons in the unit. The prevailing experience of the persons working in the unit is departmental management or the government's budget office.

Most respondents do not use task forces, teams or committees to drive the use of performance measures to improve the delivery of services. The purpose of the task forces, teams and committees where they were used is as follows.

- Advise the chief executive or legislature on such matters as the implementation of a managing for results program.
- Provide a long-term strategic vision.
- Set performance measures; ensure they are written to attain specific results.
- Gather and analyze performance data and results.
- Make recommendations for improvement.
- Monitor implementation of improvement plans.

Establishing the Performance Measures

Sixty-eight percent of the respondents state that the initial selection of the performance measures involved the agencies providing the services. This was closely followed by the unit that oversees the use of performance measures to improve the delivery of services (45 percent) and employees of the agencies that deliver the services (44 percent). Other groups reported as participating in the initial selection of the performance measures were the chief executive (31 percent); legislative body (22 percent); service recipients (7 percent); and others, such as academics, consultants, etc. (13 percent). (The higher than 100 percent total indicates that more than one group was involved in the initial selection of the performance measures.)

The respondents indicated there was a fairly even distribution among the types of measures in the initial selection. They identified 0-20 percent of the measures as inputs, 0-20 percent as workload, 0-20 percent as outcomes, and 21-40 percent as outputs. It should be realized, however, that even though the survey provided definitions for the different types of measures, one person will identify a measure as an outcome while another person will identify the same measure as an output. Accordingly, the respondents were asked for examples of outcome indicators. Many would be considered outcomes: for example, library circulation and visitation rates; air pollutant levels, permit approval within 12 months, overall condition index for streets, prison recidivism rates. However, others were workloads or outputs, for example, number of retirements processed, number of calls taken, books shelved per minute.

Figure 2: Who Established the Initial Performance Measures?

Who Established the Initial Performance Measures?	Responses
Agencies providing the service	64
Unit that oversees the use of performance measures to improve service delivery	42
Employees of the agencies that provide the services	41
Chief Executive	29
Legislative body	21
Service recipients	7
Other, e. g., academics, consultants	12

In almost 60 percent of the instances, agencies were not allowed to revise the measures during the initial operation of the process, although there were instances of revisions. Examples of these are: if a closer tie to the performance goals can be demonstrated, if there are significant changes in the service delivery methodology, when better data are available, if indicated by a performance audit's findings, and at the start of a new period after discussion with the performance management unit and/or approval by the agency head or the budget office.

The previous data notwithstanding, over 80 percent of the 85 persons responding to the questions as to whether the initial performance measures were revised reported they were. These revisions appear to have been made after the process was in place for at least a few years. The reasons were to be able to monitor additional services (60 percent); the measures are no longer applicable (59 percent); to reflect a realization of what is important to the public (53 percent); to move to more outcome measures (51 percent); to narrow the number of measures to a more workable amount (37 percent); and for other reasons such as suggestions from citizens, city employees and elected officials; or to conform to a new governor's desire to narrow the focus to key functions (17 percent). (The higher than 100 percent total indicates there frequently was more than one reason for revising the initial measures.) See *Figure 3*.

A slightly smaller number of persons answered the question about the types of measures used after the revision. As expected, they responded that the percent of input measures decreased and the percent of outcome measures increased. It is hard to tell from the responses whether the proportion of measures that were workloads and outputs increased or decreased.

Performance Results Data

Eighty-three persons responded to the question asking whether data for the defined performance measures are regularly collected, tabulated and provided to managers in a usable format. Ninety percent said yes. *Figure 4* presents the sources that eighty-three respondents identified. Obviously in some governments, there are several sources.

Consistent with what was determined from the interviews, the overwhelming method used to check the reliability of the data is internal management review of the data (84 percent). Other methods reported were: central review of the data for anomalies or abnormalities (35 percent); auditor verification of the data (35 percent); department head certification of the data's reliability (34 percent); reviews of controls underlying the data collection systems (33 percent); comparisons with sources of similar data (30 percent); and following up complaints of unreliable data (28 percent). See *Figure 5*.

Review of Performance Results

The most common method for reviewing the performance results, identified by 45 percent of the 76 persons that responded to this question, is frequent, periodic meetings between the chief executive and/or the head of the unit that oversees the use of performance measures and each agency head. Most of these governments were states and other large governments. Forty-three percent of the respondents said they rely on the chief executive's review of reports of agencies' performance results followed by establishment of performance targets for the ensuing period, usually a year. Thirty-four percent use other methods, such as internal reviews of performance measures, periodic performance audits and annual legislative reviews during the budget process.

The meetings generally occurred monthly (24 percent) or quarterly (29 percent). In eight percent of the cases, the meetings were held weekly; and in six percent, bi-weekly. Thirty-three percent of the respondents answered bi-annually, annually or that they did not know.

Forty-eight persons answered the question about the chief executive's attendance at the meetings. Fifty-four percent responded always; 23 percent, frequently; 10 percent, about half; four percent, hardly ever; and eight percent, never.

Forty-four percent of the 52 persons who responded to a question about support staff (that is, budget, personnel, IT and supply) attendance at the meetings stated a significant extent; 27 percent, to some extent; and 29 percent, never.

The meetings are evidently not open to the public, according to 37 of 52 respondents to a question about that matter. This could suggest that the meetings were more like "ad hoc" meetings—that is, "while you're in the office, there is something I want to discuss"—as opposed to formal sessions.

Seventy-five persons identified what the performance results data are compared to. Seventy-nine percent said to targets, which is only slightly more than the 76 percent who said comparisons are made to prior periods' performance. Forty-one percent said to similar jurisdictions and 25 percent to other benchmarks, such as ICMA-CPM

Figure 3: Why Were the Initial Measures Revised?

Why Were the Initial Measures Revised?	Responses
To be able to monitor additional services	42
Measures no longer applicable	41
To reflect realization of what is important to the public	37
To move to more outcome measures	36
To narrow the number of measures to a more workable number	26
Other	12

standards, trade association best practices and federal standards.

Disaggregated results are sometimes considered by 49 percent of 67 persons who responded to a question about disaggregation. They are always considered, where appropriate, by 30 percent of the respondents to this question; hardly ever by nine percent; and never by 12 percent. (It is highly likely that the majority of the persons who did not respond to this question also don't aggregate.) Sixty-nine percent of the disaggregation was reported for being by time of year, month, week or day; 49 percent geographic; 40 percent demographically, 20 percent by income level.

There were 71 responses to the question, "What type of preparation is performed prior to the reviews of performance results?" Ninety-three percent stated there is a review and analysis of the performance data by the unit that oversees the use of performance measures and 21 percent said a memorandum is prepared for the chief executive. Ninety-one percent of 77 respondents stated the reviews seek the reasons for less than desired performance

Sixty-seven percent of 97 respondents said agency heads generally conduct periodic meetings to review performance results.

Follow-up on the Performance Reviews

Conscientious follow-up after the reviews and/or meetings between the chief executive and agency heads is essential to assuring the service improvements are achieved. Of the 74 persons who answered a question about the type of follow-up that is performed, 45 percent stated there are additional meetings between the chief executive and agency heads. Thirty-six percent stated a memorandum is prepared summarizing the decisions and action steps agreed to at the initial meeting; and 35 percent stated improvement plans are prepared for the instances of less-than-desired performance.

As a result, 54 percent of 76 respondents stated that agency heads sometimes cite performance measures and results as reasons for revising work processes, materials used, service delivery, etc.; 29 percent, frequently; 15 percent, hardly ever; and 3 percent, never.

Figure 4: Sources of Performance Data

Source	Responses
Agencies' operating records	75
Periodic customer surveys	43
Complaint/compliment tracking programs	29
Feedback on specific issues	22
Comment cards	17
Independent or central management inspections	16
Citizen or customer focus groups	15
Trained observers	14

Training

A question about training generated 78 responses, of which 69 percent stated that training on the nature and use of performance measures for managing and delivering government services is provided in their government. Thirty-eight percent of a slightly larger number stated the training is provided to agency heads, 36 percent to all employees; 11 percent to all new employees; and 42 percent to other, such as new supervisors, finance and budget personnel, or people involved in the performance measurement process. For 80 percent of the respondents, the training is less than 10 hours a year for both department heads and all new employees. Approximately 11 percent stated they provide 11 to 25 hours a year, and 8 percent said they provide more than 26 hours a year. The training was reported as mandatory by 40 percent of the respondents.

The training is provided by in-house personnel for 78 percent of the respondents to this question; professional associations for 25 percent; for-profit contractors for 13 percent; and colleges and universities or other not-for-profit organizations for 10 percent.

According to the titles of the training programs, the subject matter appears to encompass such topics as managing for results and performance management, strategic planning, introduction to performance measurement systems and using data and statistical tools to analyze and improve performance.

Miscellaneous

Seventy-two persons answered the question about the existence of a government-wide strategic plan; 58 respondents or 81 percent said yes, the government had one. Sixty-four percent of those having a government-wide strategic plan said it was updated within the past year; 17 percent , one to two years ago; 6 percent, two to three years ago; 5 percent, three to four years ago; and 8 percent, further back. Ninety-one percent of the 58 said the performance measures are intended to be consistent with the strategic plan.

Seventy-one persons answered the question about the existence of agency strategic plans. Fifty-one percent said they existed for all the agencies, 38 percent said for some departments, and 11 percent said for no agency. Ninety-five percent of the 63 governments with at least some agency strategic plans said the performance measures are intended to be consistent with the strategic plan.

Fifty-nine percent of 68 respondents said the government's budget is organized according to the government's programs, which again is different from the interviewees' responses. Ninety percent responded that the central budget reviews and deliberations consider the performance measures and results, with the response split evenly between all instances and a few instances. Examples of how the performance measures are factored into the budget process are as follows.

- Agencies use performance metrics to demonstrate the need for new positions.
- Budget requests include reports of performance measures, which are used as part of the deliberation process.
- Key performance measures, targets and current results are presented along side the financial budget requests.

Seventy-six percent of 63 respondents responded that the budget resources are allocated to programs rather than solely to object classes.

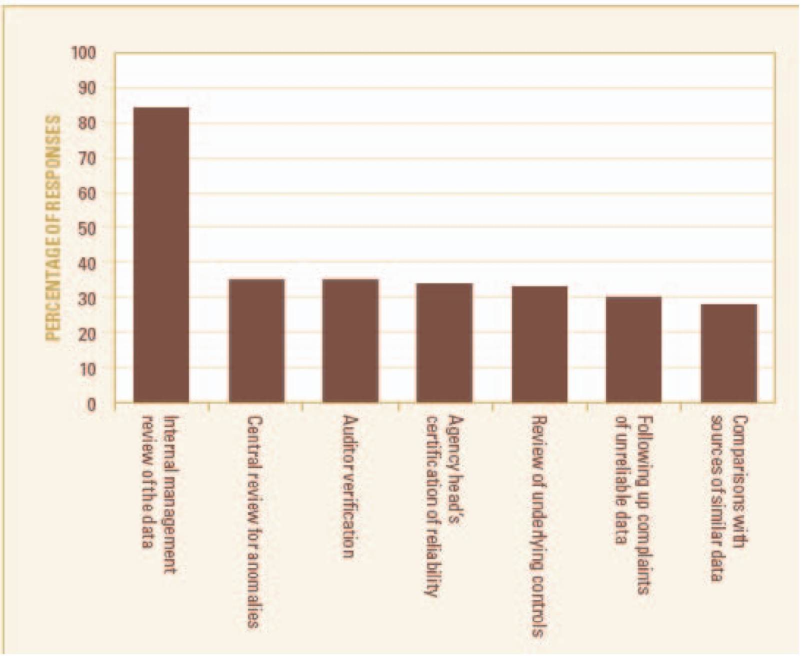
Eighty-seven percent of 63 respondents indicated their governments are fully or at least substantially able to ascertain the various types of direct costs of the services they provide; 63 percent made that claim for indirect costs; and 38 percent for imputed costs.

Forty-nine percent responded this is the result of a cost accounting system integrated with the financial accounting system, 19 percent the result of year-end cost finding; and 32 percent the result of other, such as investigation and analysis using the accounting and payroll systems or cost allocation plans approved by the federal government.

Eleven percent stated the government is fully (to a very great extent or completely) able to ascertain the costs of individual outputs or outcomes; 34 percent substantially (to a great extent or a lot); 36 percent partially (to some extent or somewhat); and 19 percent never. (This is another response in which the survey respondents differed from the interviewees.)

The extent to which personnel evaluations are reported as addressing program performance is split fairly evenly among the 66 persons who responded to this question. Twenty-one percent said usually; 23 percent said very often; 27 percent said sometimes; and 29 percent said rarely. As expected, the extent to which promotions and compensation levels reflect program performance is somewhat less: usually (11 percent), very often (17 percent), sometimes (27 percent), and rarely (45 percent).

Figure 5: How is the Data Reliability Assured?



The respondents did say they regularly disseminate performance results and other performance information to the public by various means. Seventy percent of the 60 persons who responded to this question do so through the budget documents; 55 percent use regularly published performance reports; 42 percent use the Internet; and 32 percent use press releases for notable accomplishments. Thus, public interest in reported performance results has been a catalyst in improving service delivery

Legislative bodies' interest in performance measurement was reported as ranging from "None at all" to "They love it." In between were such comments as, "There are members on the council that have been providing some general direction related to performance measurement," and "The legislative body monitors the performance measurements and deliberates on the measures during departmental presentations and the budget process." On the other side, comments such as, "Some members are interested as time allows," "Few members are extremely interested and read the reports," and "I'm not sure that our council has any more than 'limited' knowledge about the organization's performance management system," demonstrate opposite experiences.

Seventeen respondents stated they submitted their performance results to data bases maintained by such groups as ICMA, federal agencies, the American Water Works Association, the Public Library Association, state and regional agencies, and "like-minded performance management experts."

Continuation of Performance Measurement

A critical issue when operating a system to use performance measures to improve service delivery is whether the system will continue after the departure of the "champion" who initiated the program. This will depend to a large extent on the agency heads' acceptance of the process and the extent to which the system's procedures have been integrated into the government's routine operations.

Eighty-three percent of 71 respondents to a question pertaining to institutionalizing the process answered that their agency heads have very much or somewhat bought into the use of performance measurement to improve the delivery of services. Only 17 percent responded that their agency heads have hardly or not at all bought in.

Steps that have been taken to assure the continuation of the use of performance measures to improve the delivery of services following the current administration's departure include enacting legislation requiring the process; introducing technology where feasible; making the results accessible to the public, which creates demand for continuation; using training to demonstrate to mid-level supervisors the effectiveness of performance management and what's in it for them; and establishing positions and units that will outlast the current administration.

Concluding Questions

Two final questions were asked: "In your opinion, what effect has the use of performance measures had on improving the delivery of services by your government?" and "What is an example of how your government used performance measures to evaluate the effectiveness of a program or service and improve its delivery?"

Thirty-two percent of the respondents stated that the use of performance management has had a significant effect on improving the delivery of services; 54 percent said a modest effect; and only 14 percent said no effect. See *Figure 6*.

The following are a representative sample of how the respondents reported they used performance measures to evaluate the effectiveness of a program or service and improve its delivery.

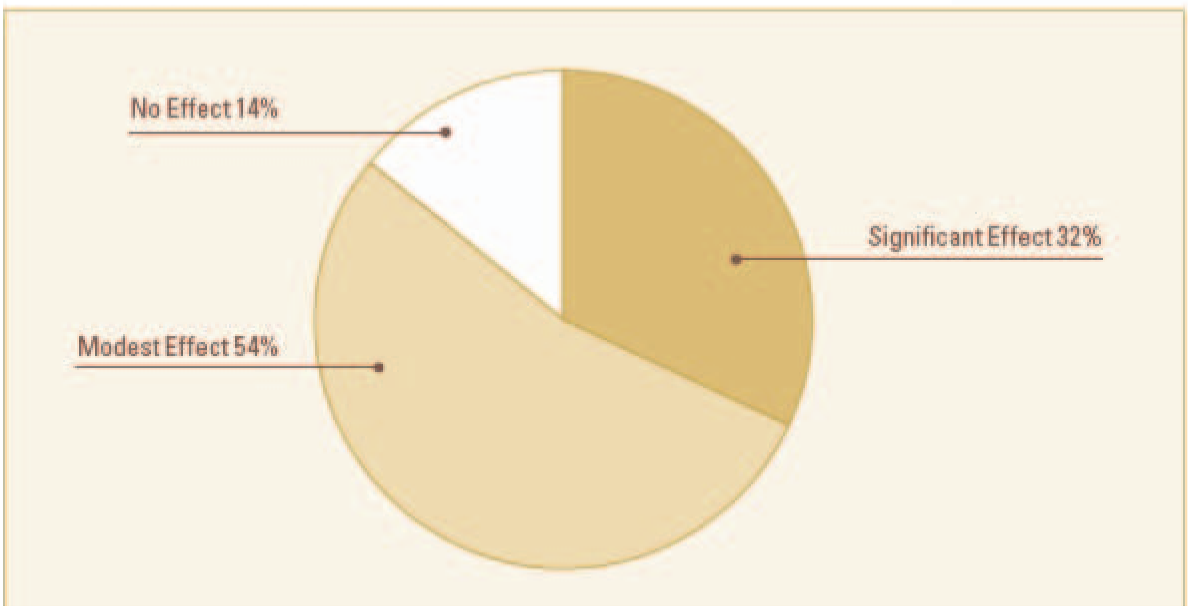
- Reduction in crime.
- Requiring taxpayers to electronically file payments of a significant amount reduced paperwork, simplified the process and enabled quicker responses to issues.
- Monitoring completion of facility safety inspections resulted in an increase in the completion rate from 70 percent to 100 percent.
- Reduced book processing turnaround time from 21 days to eight days.
- Accounting for the percentage of water drove the water system rehabilitation efforts.
- Reduced the time for issuance of air quality new source construction permits from 62 days to six days and eliminated a 600-application backlog.
- Child re-abuse rates reduced 30 percent by improving social workers' response times, which was made possible by reducing the time the social workers spent waiting for court reviews.
- Accelerated invoice payments by changing the invoice approval process.
- Closed a boot camp that showed no change in recidivism rates.

Conclusion

The notion of performance measures for government services has been discussed for many years. The current attention being given to the subject by such organizations as the GASB, ICMA and legislation like the ARRA is likely to greatly expand the use by governments of performance measures to report performance to their citizens and others. The same measures can also be used by governments to improve service delivery and achieve desired results.

This report demonstrates that such use of performance measures has already been adopted by many governments. It also provides guidance with which other governments can start to follow suit. The most significant piece missing is the desire and discipline to start and maintain the process—and that is ultimately up to each government body.

Figure 6: What Effect Did Performance Measures Have on Improving Service Delivery?



Glossary of Acronyms and Terms

311 System: A system entailing a non-emergency telephone number to which callers can report or obtain information about the government's services

GASB: Governmental Accounting Standards Board, the independent organization that establishes and issues standards of accounting and financial reporting for U.S. state and local governments

GPRA: Government Performance and Results Act of 1993 (P.L. 103-62).

ICMA: International City/County Management Association, the professional and educational organization for appointed managers, administrators and assistants in cities, towns, counties and regional entities

NPMAC: The National Performance Management Advisory Commission is a consortium of the following organizations: Association of School Business Officials International; National Association of State Budget Officers; Council of State Governments; Government Finance Officers Association; International City/County Management Association; National Association of Counties; National Association of State Auditors, Comptrollers, and Treasurers; National Center for State Courts; National Conference of State Legislatures; National League of Cities; and the U.S. Conference of Mayors.

Object Classes: Object classes are the budget classifications that represent the types of services or goods for which monies are spent, for example, personal services, supplies, travel, etc.

Outcome Measure: A measure of the results associated with the provision of services

Output Measure: A measure of the quantity of services provided

PART: Performance Assessment and Rating Tool. A presidential initiative led by the Office of Management and Budget, which rated executive branch agency performance in the early 2000s.

Stat: A performance management system based on the New York City CompStat model that has since been adapted by other jurisdictions

Appendix A: Methodology

The research approach consisted of three parts: identification of the elements of using performance measures to improve service delivery; development of case studies describing five governments' successful use of performance measures to improve service delivery; and administration of an online survey to determine the extent to which the elements are used.

The identification of the elements of using performance measures to improve service delivery started with the drafting of an extensive straw man list of possible elements. The list was reviewed and commented on by the project's advisory committee. It was also used as the foundation for the interviews with the five governments. The reactions of the advisory committee members and the use of the components by the interviewed governments resulted in dividing the list into three categories: essential, supporting and ancillary. The definitions for each of the three categories is presented in the body of the report.

The case studies were developed by conducting interviews in five jurisdictions that have been using performance measures for several years to improve service delivery. The jurisdictions were New York City, where the Stat movement started; Baltimore, which adopted a similar approach; two smaller council-manager governments-Westminster, CO (110,000 population) and Fishers, IN (67,000 population); and the state of Washington, which also adapted the Stat approach. In each instance, the interview was with the person who oversees the process.

The first step in the interview part was to develop a questionnaire based on the straw-man list of possible components. The questionnaire was sent to each interviewee prior to the interview. Three of the governments responded with written materials describing their use of performance measures to improve service delivery. The materials were reviewed and the questionnaire for that government tailored to reflect the content of the written materials.

Two of the interviews were face-to-face: New York City and Baltimore. New York City arranged interviews at the citywide and agency levels. Baltimore preceded the interview with the opportunity to attend a CityStat meeting. The others were phone interviews lasting approximately two hours each.

The case studies were prepared immediately following the interviews and sent to the interviewees for review. Comments were received and reflected in the final case studies.

The online survey was conducted to determine the extent to which the components of the process for using performance measures to improve service delivery were used by governments. It was developed with Survey Monkey and sent to the 4,600 members of AGA listed as affiliated with a state or local government. In addition, ICMA announced the availability of and provided a link to the survey in the electronic ICMA newsletter, which is regularly sent to its 8,000 members; and Rutgers University provided a link in the Network Newsletter, which is regularly sent to the federal, state and local practitioners in performance measurement reporting and performance management as well as public officials, researchers, students and citizens interested in public performance measurement and reporting that belong to Rutgers' Public Performance Measurement and Reporting Network.

There were 175 respondents who answered yes to the question, "Does the government use performance measures to evaluate the effectiveness of government programs and services and improve their delivery?" These were the basis for the analysis in this report.

The online survey was opened by 253 individuals, with 230 responding in some fashion. For this research, the focus was on the 175 responding affirmatively to the question "Does the government use performance measures to evaluate the effectiveness of government programs and services and improve their delivery?" Information provided by these 175 responses provides an overall sense of how performance measures are currently being used within state and local government.

Appendix B: Case Studies

The following are descriptions of the history, organization, process and elements of the programs to use performance measures to improve service delivery in five governments. This information was obtained from interviews conducted with the person directing the program in each of the five governments. The information was not independently verified.

New York, NY

In New York City, the genesis for the use of performance measures to improve the delivery of services was a 1977 revision to the City Charter that requires the mayor to publish a bi-annual report on the performance of city services. This report, titled the Mayor's Management Report (MMR), presented performance information for each of the approximately 45 agencies and organizations that report directly to the mayor. It was long and heavy on narratives of what had been done. Data was presented for about 600 indicators/measures that basically reflected what the agency managers considered easiest to obtain data for. The measures were typically not outcome-oriented.

The MMR was therefore restructured in 2002. After consultation with the Citizens Budget Commission and other groups, the MMR was simplified to a document to which the city council and taxpayers could relate. The size was dramatically reduced by requiring each agency to report its performance in no more than four pages; the verbiage was dramatically reduced and had to relate to the measurable results of performance reported for approximately 1,000 indicators (which has since increased to just over 1,200); and if performance was declining, explanations of why and what will be done to improve the performance are provided. Four hundred copies of the report are printed for city officials, city watchdog groups, official visitors seeking to learn about the city's operations, colleges, etc.

The most significant step to using performance measures to improve the delivery of services was in 2006 with the development of Citywide Performance Reporting (CPR) as an online interactive dashboard. CPR entailed selecting approximately 530 critical outcome measures from the 1,200 measures in the MMR. The selection was the result of a collaborative effort between the mayor's Office of Operations and each agency, with the final decision residing with the Office of Operations (or the appropriate deputy mayor). The goal was to choose customer-oriented measures of outcomes, while including some measures related to efficiency, plus remaining sensitive to budget allocations wherever possible. Because of the connection between performance measures and union contracts, there was some involvement from the Office of Labor Relations and the unions. The city council, although attentive to the reports that are issued, was not involved in the selection of the measures.

It should be pointed out that measures have been changed occasionally to reflect changing realizations of what is important to the public. For example, absolute numbers have been changed to percentage improvements for such services as responding to abuse reports; percent completed within a specified time frame has been changed to average days to complete for such services as filling potholes. Changes are held to a minimum, however, in order to maximize historical comparisons.

The heart of CPR is the "online interactive dashboard" the Office of Operations maintains using the critical outcome measures and the results for each measure. The dashboard, which is available to the public at *www.nyc.gov/html/ops/cpr/html/home/home.shtml*, is a graphical representation of agency performance, including pie charts and color-coding to make positive or negative performance trends obvious. The critical outcome measures are grouped into government functions that cut across agencies and disciplines to reveal the overall picture about city government performance. Detailed information is provided for each measure, including an explanation of what the measure means, its reporting frequency and other useful details. A drill-down capability is provided that allows users to review comparative trends over a five-year period. CPR's accessibility, detail and ease of use enable the public to monitor the city's performance. This is a major driver for improving performance. It also increases agency heads' accountability, and has them pay more attention to results.

Specifically, a city resident-or anyone else-can go to CPR; click on the theme with which he is concerned and ascertain the agencies that provide services for that theme; select the desired agency; and then see a color-coded pie chart showing the performance for that agency's measures. Green indicates the percentage of the agency's performance measures that are improving or stable; yellow indicates the percentage declining 10 percent or less; red indicates the percentage declining more than 10 percent; grey indicates the percentage for which there are no data yet; and white indicates the percentage for which there is no desired direction for the measure.

Each measure is also listed individually with numbers showing performance for the latest month and year, the changes in performance from the previous month and year, and the timeliness of the information. Each measure is color coded with its performance trend to enable a quick identification of measures with which there should be concern.

The other way CPR drives improved performance is that each agency head prints his (or her) agency's report and attaches it to a bi-weekly memo that he (or she) submits to the deputy mayor to whom he reports. He then meets with the deputy mayor and, among other matters, discusses the measures marked in red. The agency head (or other senior manager) generally will already have a remedial action or plan for improvement underway, or at least under development. The meeting with the deputy mayor and/or staff will serve to get their feedback on the plans from a policy standpoint, which might lead to a revision or updating of the plan, and possibly set targets and/or deadlines for the action plan. The mayor's senior staff frequently participates in these meetings. In addition, the mayor's Office of Operations provides exception reports, that is, reports of the indicators marked in red, to the mayor and the deputy mayors that oversee the agencies for follow-up. The Office of Operations might also contact agencies directly and ask what is happening to cause the declining performance and then pass the information to the appropriate deputy mayors.

Other items worth noting:

- In New York, persons in management positions are long-term management professionals in government, the private sector and/or not-for-profit organizations, rather than belonging to the civil service. For example, the prior experience of the director of the Office of Operations was with the city's Office of Management and Budget (OMB). Other employees in the Office of Operations are also not civil service, but would likely be retained through different administrations, assuming the interest in CPR and performance measurement remains.
- There does not seem to be much worry about the reliability of the reported performance results. It is felt that the numbers cannot be falsified significantly. Constant working with the data provides a sense of its reliability. Public availability of the results enables others to question numbers they believe are problematic. Some of the numbers are audited subsequently. Finally, there are occasional discussions among the interested parties on the "correctness" of the indicators, for example, if the measures are always green, the indicator might not be the right indicator.
- Although the city uses performance measures extensively to drive performance, it does not use performance-based budgeting. The council appropriates monies primarily to agencies, and in only a few instances to the program level. The mayor has selected agency heads who he believes are the best in their field, and therefore will work within their budgets to achieve their performance goals. The city would like to link the budget allocations to performance levels, but believes this is too difficult at this time.
- Some agencies have cost measures, particularly the agencies concerned with infrastructure (for example, Departments of Environmental Protection, Buildings, Transportation, Housing Preservation and Development). OMB is trying to get more cost measures. It would particularly like to have costs per output.
- Program performance is not tied to employee compensation; it is contrary to the unions' policies. However, program performance in some agencies is tied to advancement. If the city pushes agency heads on performance, they will push performance down within their agencies.
- Citizen input on performance is obtained in several ways. The first is from the extensive complaint tracking conducted with the 311 system.

  Second, a Citizens Budget Commission Roundtable stimulated the city to undertake a customer satisfaction survey to gather information about priorities, impressions, etc. Twenty-three thousand responses were obtained, with a minimum of 300 from each of the city's 59 community districts. This is believed to be a representative sample. The results were transmitted to the agencies to be sent to their districts and used to alter programs.

  A third way began during the summer of 2009 when 10 city agencies started using feedback comment cards to measure customer satisfaction. Participating agencies include, among others, the Departments of Buildings, Environmental Protection and Transportation; the Taxi and Limousine Commission, the New York City Housing Authority, and the mayor's Office of Film, Theater and Broadcasting. Agencies place the comment cards at their walk-in facilities in order that customers can rate the following service elements at each location: convenience and accessibility; cleanliness; staff professionalism and courtesy; and clarity of rules and important information. Customers also rate the overall customer service received. The cards are two-sided; the reverse side lists the questions in Spanish. Agencies can apply the results of this survey to satisfy the new mayoral requirement to measure customer satisfaction at least once a year and to modify service delivery as appropriate.
- Some external groups have asked for certain measures and formats. Although the city also held a roundtable on performance-based budgeting, the attendees were not involved in selecting performance measures

The Agency Perspective

The Department of Sanitation New York City (DSNY) has four service areas: street cleaning, waste and recyclables collection, waste disposal and snow removal. DSNY is comprised of seven borough areas, broken down into 59 districts, which is further broken down into 233 sections. The following provides insight into how performance measures have driven improvements in the delivery of three of the four DSNY services.

Street Cleaning. DSNY started measuring performance in 1975 when the Fund for the City of New York conceived of and initiated a system for rating street cleanliness. The Fund took pictures of different levels of street cleanliness and asked members of the public what they thought was the level of cleanliness represented by each picture. It then developed a seven-level scale for cleanliness. A clean street with no litter would receive a rating of 1.0; a street with highly concentrated litter in which there are no gaps in the piles, that is, the litter is in a straight line along and over the curb, would receive a 3.0. Levels of cleanliness between these two extremes would receive points between 1.0 and 3.0. The Fund inspected and rated a sample of streets throughout the city and issued reports on the cleanliness of the streets.

The inspection and rating of street cleanliness has been assumed by the mayor's Office of Operations. The office and DSNY have established a goal that 85 percent of the streets citywide are rated acceptably clean, that is, a rating below 1.5; and that in each section, 80 percent of the streets are rated acceptably clean. The inspections are performed by seven inspectors from the mayor's Office of Operations who twice a month examine a random sample of 7,000 city blocks out of the 120,000 blocks in the city. They compare each street to the photos of varying levels of street cleanliness and assign the street a numerical score of 1.0 to 3.0. The rating is input into a handheld device. The street cleanliness ratings will be moved shortly to the Street Conditions Observation Unit (SCOUT) program, which is a team of inspectors based in the mayor's Office of Operations. Its mission will be to drive every city street once per month and report to the 311 center, conditions that negatively impact quality of life (dirty streets, pot holes, missing signs, graffiti). Conditions will be reported to the relevant agencies for appropriate corrective action, in the same way 311 handles complaints from the public.)

When this program started in 1975, 72.4 percent of the streets were rated acceptably clean. The percent dropped to 52 percent during the fiscal crisis of the late '70s, but then slowly started to rise. It is now 95 percent, with all districts reporting street cleanliness above 90 percent.

The availability to the commissioner of sanitation, and eventually the Office of Operations, deputy mayor and mayor, of a measure that indicates the cleanliness of the streets stimulated various actions that over time improved the cleanliness of the streets. For example:

- The district superintendents are held accountable for the cleanliness of the streets within their districts. They have to present maps showing how they clean the streets: that is, mechanical sweepers, manual sweepers, the extent of alternative side of the street parking, etc. The maps are reviewed with the borough chiefs and eventually with the commissioner and deputy commissioner for operations.
- If poor performance is noted, e-mails are sent to the district superintendent and borough chief, asking "what's up?" Analyses are performed to ascertain the reasons for the poor performance, for example, insufficient litter baskets. People may be brought to headquarters to discuss the shortfall. At the end of the year, general discussions are held of the problems encountered. In short, lots of information is shared.
- The low levels of street cleanliness stimulated DSNY to switch from staffing street cleaning with people left over from collection to requiring staffing of cleaning with a minimum number of people and authorizing overtime if necessary.
- When streets receive poor ratings, enforcement personnel make visits to enforce the requirement that property owners clean 18 inches into the street. Since the city's Office of Management and Budget projects revenues from summonses, DSNY establishes productivity standards for summonses per day in areas it knows are dirty. Also, summonses are no longer written by hand. Illegible handwriting was often used as an excuse to ignore a summons. Summonses are now produced by hand-held computers, which addresses the illegibility problem as well as automates transmission of the data.
- Although the union contracts that cover all employees up through the borough chiefs preclude tying compensation to performance, street cleanliness ratings have influenced promotions.

Collection. The performance measures for collections focus on productivity, that is, tons collected per truck per shift. The citywide target is 10.7 tons; the year-to-date result is 9.9 tons, but the decline is due not to a loss of productivity, or even more recycling, but to less available trash (for example, lighter packaging, smaller newspapers). In 2005, 54,000 tons were collected each week; so far in 2009, only 48,000 tons have been collected each week.

A target is set for each district after considerable analysis that takes into consideration such factors as the season, day of week, length of routes, numbers of stops and numbers of turns while driving. The target is then negotiated with the union citywide. More tons collected means more pay. To further increase productivity and lower overall costs, the persons on the trucks are paid an additional $5 a route if they transport the garbage to the disposal/transfer site rather than leave that trip to others.

Collection reports showing refuse tons collected per truck route and shift are prepared weekly; other reports are prepared weekly and daily. Not only are the collection results monitored constantly by management. The department also uses the historical data during labor-management meetings and/or formal labor negotiations to inform the union of the need to change the routes in order to improve productivity.

Similar measures are in place for collecting recyclables. The collections target for recyclables is 6.7 tons per truck shift; however the actual is only 5.6 tons. Lower quantities are part of the reason. The other is that for the performance measure "percent of total curbside and containerized waste diverted to recyclable," the target is a diversion rate of 25 percent, but currently only 16 percent is being diverted. In other words, the city is capturing as recyclables only 45 percent of the potential recyclables. Hence, it is also measuring and reporting the numbers of districts with curbside and containerized waste diversion rates between 0.0 percent and 4.9 percent, between 5.0 and 9.9 percent, and greater than 25 percent. It then analyzes the recyclable collections results by district, season, type of neighborhood (industrial, commercial or residential) and income strata. Outreach is performed citywide when funding is available. Problem districts receive special attention in additional outreach staff and/or increased enforcement, for example.

Snow Removal. The performance measure for snow removal is the percentages of primary, secondary and tertiary streets cleared at specific points in time. The key to improving performance is to prepare different plans (clearing, sanding, salting), down to the district level, for the different types of storms that can occur and the day during the week of the occurrence. Unfortunately, the accuracy of the forecast is typically the major contributor to performance. Hence, the commissioner meets with senior staff after every storm to determine whether and how the plans could be changed to improve snow removal.

Other items worth noting:

- The commissioner is a political appointee, but has risen through the ranks of DSNY since 1960. The deputy commissioner for finance, management and administration, who is responsible for the gathering, analysis and reporting of performance results, serves at the pleasure of the commissioner. He oversees 15 uniformed and 15 civilian employees who analyze the operations and performance results. The blending of knowledge and experiences of these two types of personnel conducting the analysis greatly increases the department's insight into the operations and results. Indeed, DSNY's entire management structure mixes uniformed persons and civilians to provide a good balance and interrelationships.
- Data are collected daily and maintained on Sanitation Control Analysis Network (SCAN). This is a 30-year-old system that is being replaced by the multi-million dollar Sanitation Management Analysis Research Tool (SMART). Data are input into the system from hand-held devices.
- DSNY uses activity-based costing to obtain the fully loaded costs of waste collection, recyclables collection and disposal. The real benefit of this analysis is that it encompasses a comprehensive review of the department's collection practices that provides information the department might not otherwise have. For instance, the analysis revealed that it costs more to collect recyclables, even after factoring in the revenues for paper and the lower tipping charge for metal, glass and plastic. (Because of the environmental factors, the city would still collect and process recyclables separately.) Another advantage to the activity-based costing is that it enables DSNY to report in each year's MMR the collection cost per ton and the disposal cost per ton.
- The 311 system has been a major supplier of complaints, particularly complaints for missed collections, but also for street cleanliness and employee behavior. DSNY has been able to respond to these complaints. As a result, the number of complaints has been decreasing. Also, if a preponderance of complaints are from a particular district, inspectors are dispatched to that district to ascertain what is occurring.

One result of the above use of performance measures to improve service delivery, as well as increase productivity, is that the number of DSNY FTEs has decreased from 11,000 in the 1970s to 7,000 in 2009.

Baltimore, MD

Baltimore's use of performance measures to improve the delivery of services is called CityStat. It was initiated in 2000 following the election of a new mayor, based upon advice provided to him by a veteran of New York City's CompStat program.

The director of the CityStat program is the city's deputy mayor for administration. The CityStat Office was originally staffed by seven analysts, but now has four. The director has a master's degree in public policy and urban planning. His experience in performance measurement has been as a CitiStat Analyst for four years. The analysts' previous work has been in fields other than public administration, for example, military law, management consulting and teaching. The director believes these experiences enable the analysts to think about more than cost containment. Each CityStat analyst is assigned two to three agencies.

There are two key elements in CityStat.

- The first element is a template. It lists, for each department, performance and other data for the past two weeks, several prior two-week periods, several prior years and percentage changes for both. Some of the data are standard for all departments, for example, expenditures, numbers of personnel and related matters. Most of the data are for measures that represent the many outputs the department produces. These measures were initially selected by the mayor and a panel that asked the departments what measures they think are appropriate, how would they know their service delivery is improving, what do they use to manage their operations, etc.

  Baltimore's measures are primarily output measures. An effort is under way to present more outcome measures, using citizen surveys to establish what the public expects from the programs.
- The second element is bi-weekly meetings held by the mayor and/or deputy mayors to review performance of individual departments. Initially the meetings were for just the departments, for example, Police, Fire, Housing and Community Development. The current mayor expanded the scope to also hold meetings with the several agency heads involved in a policy area, for example, CleanStat, which focuses on cleaning up the city; or GunStat, which focuses on reducing gun violence and involves persons from other than the city government (state Attorney's Office, federal law enforcement officials).

  The meetings are open to the public and there can be many attendees. They have not received much attention from the media. Council members will occasionally sit in, but their interest is on the periphery of the process.

  The process begins prior to each meeting with the CityStat analyst updating his or her understanding of the performance of a department scheduled for review by reviewing its latest template. Drawing on his or her knowledge of the mayor's and deputy mayor's interests, he or she prepares a 10-page memorandum addressing selected issues. For each issue, the memo presents the background of the issue, performance data and pictures if appropriate, questions the mayor and the senior staff might want to ask, and a suggested time to spend on each issue. The mayor and her team are provided the memo in advance of the meeting. This enables them to acquire, in minutes, knowledge the analyst spent hours researching and developing. The department heads are not supposed to see the memo in advance of the meeting. This compels them to rely on their knowledge of the issues to respond to questions. Nonetheless, some analysts probably alert the department heads to what will be on the agenda.

The meetings are attended sometimes, but not always, by the mayor; chief of staff; the deputy mayor for administration/director of CityStat or other deputy mayors; and representatives from the city solicitor's office, finance, information technology, human resources and the labor department. The latter representative is present in case work rules or union-related issues arise. The reviewed agency head is accompanied by his or her management team.

The mayor and staff use the meetings with the department heads and their staffs to discuss some or all of the issues in the memorandum. The discussion appears to have three major benefits.

- The discussion of a department's activities and problems by many persons within and outside the department stimulates everyone's thinking, helps them to arrive at solutions and identifies better ways to do things.
- The involvement of the mayor and deputy mayors increases their understanding of how the departments work and how the services are delivered and performance is monitored, plus greatly increases the attention of the departments' leadership.
- The meetings provide the department heads with an avenue to the mayor in which they can identify their problems and work out solutions.

Following the meeting, the CityStat analyst sends a one-page memo to the department head listing follow-up actions and responsibilities. That memo, plus the information provided by the city's 311 program, are referred to when the analyst prepares for the next CityStat meeting.

Examples of some changes that resulted from a meeting involving housing and community development are:

- Trash was identified as building up in front of residences. The discussion explored such alternatives as increasing emphasis on issuance of citations to the landlords rather than the tenants in order to get the landlords to assume responsibility to work with their tenants to reduce trash; and establishing protocols for alerting public housing managers to be alert to use of their Dumpsters by nearby private residence occupants.
- During the spring and summer months, high grass and weeds on city-owned and vacant lots becomes a problem. One solution discussed was dispatching city crews to mow the grass and weeds and place a lien for the costs on privately owned property.
- The performance data shows the city owns an increasing number of properties it would like to dispose of. A possible solution to stimulating more dispositions was identified, namely adopting a highly visible, publicly displayed graphic measure.

Other items worth noting:

- Reliability of the data is primarily assured by the analysts "putting eyes on the data." If anomalies are seen, a spot check of the data would be conducted. Complaints about data reliability could result in a subsequent review.
- The city does not have a strategic plan document. Some departments have strategic plans; others do not.
- Each week, the city contacts 500 persons who called the 311 system and asks about the timeliness of the response to the subject of the call, the quality of the response and the overall 311 experience.
- The city has traditionally used a line item budget approach in which budget requests were based primarily on the prior year's spending. It is moving toward an "outcomes budgeting" approach. This has been motivated, in part, by the need to live within shrinking resources. Outcomes budgeting will help by identifying opportunities for partnering among agencies and providing better rationales for the spending requests.
- CityStat focuses on improving services; hence there are no cost measures in the templates. It is felt that the adoption of outcomes budgeting will enable the analysts to obtain a better understanding of costs and include those in the analysis.

It is clear that for a program like CityStat to work effectively, there must be executive-level commitment; the program has to be kept simple; and focus must be maintained. To avoid failure, the government should not try to do too much; it should avoid introducing too much technology into the process; and it should keep the process separate from the finance function to avoid sending the wrong message.

Westminster, CO

Westminster, CO, is a suburb of Denver with an estimated population of over 110,000. It has a council-manager form of government.

Performance measures were introduced into Westminster in 1999 after the then-city manager realized the extent of literature on the subject and directed each division manager to develop at least two operational measures. The measures were not linked to the city's strategic plan or priorities. Most of them were workload measures that really did not measure the effectiveness or efficiency of the city's operations, (for example, lane miles to be plowed). The measures were not being used to drive management decisions or telling the story of the city's operations. Hence, the effort was not successful.

A new city manager arrived in 2001 and wanted to use performance measures as a management tool, specifically to improve service delivery. The new manager also wanted to use performance measures as a communication tool to inform the city council and citizens on service delivery successes and challenges. The council, at that time, did not know what performance measurement was or that it was going on. Hence, in 2002 the city started publishing an annual performance measurement report titled: *Take A Closer Look: How Performance Measures Build A Better City*. At the same time, the city joined ICMA's Center for Performance Measurement. Training in and examples of performance measures were provided to the city's departments and resulted in much sharper measures. The annual performance measurement report enabled the staff to start sharing performance outcomes with the city council.

The city manager is a strong proponent in performance measures, not only for public accountability, but also for improving service delivery. He believes performance measures should be used positively, not punitively. He also believes that performance measures encourage department heads and other senior managers to know what is happening in their departments, what the problems are, and what can be done about them; rather than to put their heads in the sand. He is striving to drive the use of performance measures down into the agencies such that when he leaves, the processes are in place to keep the use of performance measures going, and the agency heads and staffs want to continue using performance measures.

The departments establish the measures consistent with Westminster's strategic plan. The council invested considerable effort in development of the strategic plan, but has no direct involvement in establishment of the performance measures. Nor does the public. However, the strategic planning process provides the policy direction for staff to develop measures that support and measure how well the city is doing in achievement of the strategic plan's goals and objectives.

The city manager's approach is to have the agency heads use performance measures to manage and improve municipal services, rather than personally regularly review the performance results for each of the agencies. Each year, each agency head prepares a report to the manager describing how his or her agency uses performance measures to better understand, justify or improve operations. The manager reviews the report with the agency head and sets targets for the performance measures for the ensuing year. The manager reflects reporting and use of performance measures in the agency heads' performance appraisals.

Assisting the city manager and reporting to him is the senior management analyst for performance measurement. That person has been with the city for three years. He has a master's of public administration (MPA) and previously worked in another city's performance measures program. Although he has other responsibilities, for example, budget and legislative issues, he is a big believer in performance measures and is on the executive board of the Colorado Performance Measurement Consortium.

The city also established a performance measurement team, chaired by the senior management analyst and composed of at least one individual from each department, with some departments having more members. The department representatives range from senior analysts, to project coordinators, to deputy agency heads, but all are somewhere in their agencies' management structure. The team's role is to monitor the performance measures reporting process; share experiences in driving the process down within their agencies and linking the agencies' performance measures to the city's strategic plan; identify training needs; and work with the Colorado Performance Measurement Consortium.

Most department heads have by now adopted and are using performance measures as a management tool to improve efficiency and effectiveness, as well as to "tell their story." Some, however, are still in the process of buying into the notion.

One big contributor to the success of performance measures is the "Performance Measurement 101" training given to all new employees within one year of joining the city (although for some it might not be for two to three years). The purpose of the training is to improve employees' understanding of performance measurement. The course, which takes a half day, describes what performance measurement is, why the city is doing performance measurement (for example, it is better for the departments, it enables the agencies to tell their story, it helps the employees do their jobs); and how the departments are doing performance measurement. They then review their own departments' measures and are asked what, if anything, is missing. The city has obtained suggestions for good measures from the training.

The departments can revise their performance measures whenever they want. However, there is a "revamp" process in which the senior management analyst and one other person work with the department's performance measure person to review and, if necessary, update the department's measures. They then meet with the department head to present and obtain approval for the changes. The revamp is applied to two departments per year.

Examples of how performance measures have been used to improve the delivery of services and/or use of resources include the following.

- The Fire Department was concerned with its overall response time. It examined alternatives for two of the three elements of response time: dispatch time and turnout time. For dispatch, it assigned one of the dispatchers to handle only the fire/EMS requests and provided her with special training. The average dispatch time dropped from 1 minute 24 seconds in 2002 to 39 seconds in 2007. (An increase to 58 seconds in 2008 was due to new persons assigned to the fire/EMS dispatch. This signaled the need for more training.) To reduce the turnout time from 1 minute 25 seconds in 2002, the department installed rugged computers on the trucks so the firefighters can enter the service information as they ride to the scene of the request, rather than have to take the time before leaving the station. This resulted in a reduction to 53 seconds for 2007.
- Police Department overtime expense was running high compared to other Colorado communities and trending even higher. This detracted from the department's ability to provide other services. Sworn officers were moved to four 10-hour shifts, which enabled work previously done during overtime be completed during shift overlaps and on common days. This enabled overtime costs to be reduced by 24 percent or $217,000. The new shift schedule has also provided other tangible benefits, such as

improved morale among sworn officers who now enjoy a more regular and predictable work schedule.

- Colorado's municipalities' primary revenue source for general governmental operations is sales and use tax; it provides 64 percent of general fund revenues in the City of Westminster. The Sales Tax Division's analysis of the revenue per audit hour resulted in a shift of resources from auditing non-construction entities, which return $167 per audit hour, to more audits of construction entities, which return $588 per audit hour.
- Also in the city's Sales Tax Division, an analysis of how auditors were spending time stimulated a shift of certain activities from audit staff to other staff, which enabled the auditors to spend 65 percent rather than 53 percent of their time auditing. For the full year following this shift of staff resources, the City collected an additional $800,000 in audit revenue compared to the previous year.
- An analysis of measures of the age of the infrastructure, particularly water and sewer lines, was used to convince the council of the need to increase water sewer usage rates to provide the monies for the anticipated increased maintenance costs and rehabilitation needs. As a result of the increased maintenance and rehabilitation, the number of annual water main breaks dropped from 129 (2002) to 70 (2008) and the number of feet of sewer line rehabilitated has increased from 10,000 feet per year to 36,000 feet per year. This more aggressive rehabilitation of sewer lines led to a 6.6 percent reduction in sewer defects per mile in 2008 compared to 2006.
- Noting high and increasing vehicle operating costs per mile, the city locked in fuel prices to control this element of operating costs.

Performance measures are considered during the hiring process to assure that potential employees' values are consistent with what the measures show the positions require. For example, for an IT position, technical skills are typically very important. However, the department's performance measures show customer service is an important concern, so customer service skills are also deemed necessary and tested for during the assessment processes for candidates for IT positions.

Performance measures are also used for coaching employees when the results of customer satisfaction surveys indicate their performance has strayed from what the performance measures call for.

Citizen surveys are conducted regularly through the National Research Center, Inc. High-level questions- quality of life, value for taxes levied, for example-are asked, as well as specific questions aimed at revealing major problems and to confirm results revealed by the city's performance measures. For instance, the city recently noted an increase in graffiti. The citizen survey confirmed the citizens' concern. The city established a graffiti abatement program; drafted an ordinance making the act of graffiti a crime; and appointing a Police Department employee to coordinate graffiti education, the schedules and equipment for graffiti removal, court appointments for violations of the graffiti law, etc. The city will use performance measures in the future to help gauge the level of success with its graffiti abatement program.

As for reliability of the reported performance data, the primary coordinator schedules meetings or makes follow-up inquiries with service area specialists on any data that appear to be incorrect or abnormal. He uses city council staff reports, agenda memorandum, and budget documents to cross check information. The city's performance measurement revamp process for two departments each year looks at the data collection process and seeks to improve the accuracy of data reporting methods. Recently, as part of a performance measurement revamp process, the city's Police Department refined several reports to provide more accurate data on response times and crime data. Finally, members of the Performance Measurement Team and the senior management analyst review the performance data included in the service area templates before they are sent to ICMA. The current year's submittal is compared to the past year's templates to check for accurate data, significant changes or figures that look unusual.

Performance measures are used to inform the budget process and help in decisions to shift resources from one area to another. Westminster does not use performance-based budgeting; rather it looks for the general "bang for the buck" received from each function. Nor does it have cost accounting to any great degree. It has some efficiency measures, but they are based on direct costs.

The city publishes a 21-page annual performance report for citizens and the council, titled "Take a Closer Look."

Fishers, IN

Fishers, IN, is a suburb of Indianapolis, with a population of just under 67,000.

The town had limited experience working with and using performance measures prior to the arrival, in 2005, of a new town manager. The manager had participated in the ICMA-CPM program in his prior manager position and decided it would be beneficial for Fishers. A deputy town manager was hired in January 2006 and assigned the primary coordinator position to oversee the performance measurement process and the preparation of the reports for ICMA-CPM. This individual has an MPA and has worked in various manager, acting manager and assistant manager positions.

Initially, the town used all the measures included in ICMA-CPM's lists of suitable performance measures that the town believed were applicable. Neither the council nor citizens participated at that time. The process enabled Fishers to compare its performance with similar jurisdictions. Another early benefit was that performance measures helped departments, such as Code Enforcement, to get their personnel to be "on the same page." Individuals who previously could not identify what was meant by performance, began to understand what performance measures are.

Subsequently, the town recognized that many of the numerous measures in the ICMA-CPM lists did not make sense for Fishers to report to ICMA-CPM. The primary coordinator met with each department head to discuss the measures and ask what was important for that department for an internal cross-comparison of top key measures. This is a cross-year comparison of the various measures using only Fishers' data and the ICMA-CPM national median. They took the top 10, more or less, measures from each department. The Police and Fire Departments also considered the measures used for their national accreditation programs. Fishers then signed a contract with ICMA-CPM to obtain a cross-year comparison, which is in addition to the annual reporting process

Although Fishers has several master and other plans, the town's strategic plan was last prepared in 1998. (There is some effort under way to update the strategic plan.) Hence, the performance measures are not related to the strategic plan.

The use of performance measures to improve management is left to the departments. There is no centralized regular or rigorous review and analysis of the performance results. The primary coordinator meets individually with representatives of each service area and reviews the data, but more to obtain a better understanding of the measures than to ascertain the levels of performance. Fishers relies, for the most part, on the town's overall performance, which is felt to be quite high.

Comparisons of performance are mainly with other jurisdictions, using the data from the ICMA CPM. Fishers is looking at trends, but only for the measures in the ICMA CPM set. There are no official targets, although department heads think in terms of "we can improve."

The departments have been instructed to educate new employees on performance measurement. Also, the training for new hires touches on performance measurement. However, there is no specific content defined for the education.

Some departments, for example, the Development Department, are reported to discuss performance results, such as time to handle complaints, during their bi-weekly staff meetings.

The use of technology to accumulate and present performance data is not as comprehensive as Fishers would like, primarily because of a lack of resources. Several software programs are used for tracking results, but there are no town-wide reports, and thus no regular meetings to review the results.

The manager's evaluations of the department heads focus on the individuals' performance, but have been moving toward performance of the department. Consideration is being given to including the department heads' recognition of standards for their function, and eventually use of performance measures to improve service delivery.

The following are examples of where performance measures have been used to improve the delivery of services and/or use of resources.

- The data in facilities management work orders were analyzed and found to provide a good understanding of the time and materials necessary to complete a job, particularly when individuals are being pulled away to work on other jobs. This enabled engineering to improve its budgeting for the use of time and materials.
- Information Technology started surveying its customers and asking them, "How did the job go?" The surveys are now performed continuously. Using the information, Information Technology has been able to increase customer service and satisfaction.
- The Fire Department uses data on the numbers, types and locations of fire and EMS (Emergency Medical Service) calls to relocate vehicles, thus assuring use of the proper vehicles and reducing response times.
- The town decided that snow removal performance, as revealed by the National Citizen Survey (NCS), was not sufficient. It added contractual services to clear the secondary streets, redesigned the routes to avoid clearing unnecessary roads, and provided contractors with maps to assure they had the correct routes.
- Noting that police response times had increased, the Police Department increased the number of patrol districts and the officers on patrol. To increase case closure rates, also a performance measure, it consolidated the investigation divisions and realigned staff leadership.
- Human Resources data revealed a high employee turnover rate in comparison to other communities. To reduce turnover, the town designed new pay packages, expanded benefits, increased training and initiated "stay" interviews to ascertain why employees remained at Fishers.

Fishers participated in the NCS to gain feedback from its residents. It found the feedback so useful that it decided to increase its participation from every fourth year to every second year. Also, the town manager conducted an internal survey to ascertain how the departments, supervisors and staffs are working together.

There is no central, overall or defined process to assure the reliability of the data, that is left to the departments. The primary coordinator does not consider that reliability could be a problem as data is reported as it stands.

Fishers' budget document is a line-item budget organized by department and divisions within departments. Performance measures and results are included, and even emphasized, in the document. There is no attempt to prepare a performance-based budget.

The costs of services are determined using spreadsheets and cost-finding techniques. The size of Fishers and the resulting level of resources precludes Fishers from expending monies for comprehensive systems and processes.

Fishers does not issue a performance report to its citizens, although it issued the NCS results and posts considerable information on the Internet. The resources required are one reason. Another reason is that the primary benefit for using performance measures is considered to be as a tool to be used by the town manager and his staff to improve services.

State of Washington

The governor is the driving force behind Washington's use of performance measures to improve the delivery of services. In her previous position as the state's attorney general, she read *Leadership*[5] by New York Mayor Rudy Giuliani. She was impressed with the description of Compstat (New York's process for using performance measures to monitor and improve government services) and implemented a similar system, called AGMAP, at the attorney general's office. Upon election as governor in 2004, she decided to apply statewide this model of improving management using measures and data.

Accordingly, on Feb. 21, 2005, the governor issued Executive Order 05-02 establishing the Government Management Accountability and Performance (GMAP) program. The Executive Order required, among other things, that each agency develop clear, relevant and easy-to-understand measures that show whether programs are successful; hold regular problem-solving sessions within the agency to improve performance; and regularly report to the governor on performance. The requirements have also been embodied in legislation enacted subsequent to issuance of the Executive Order.

The Executive Order established an Office of Government Management Accountability and Performance to lead the effort. GMAP is located within the governor's Executive Policy Office. It is staffed with an appointed office chief, whose background is as a budget adviser, four (initially six) analysts, two management consultants, a program administrator and administrative staff. Analysts have a dual role. They help agencies develop their performance reports (a consultant role) and then independently analyze the information provided by the agencies, including the data, analysis and action plans, and advise the governor and her senior staff accordingly (an evaluator role). The program administrator's role is to maintain DataView, a web-based tool used to manage agency performance.

Performance measures were developed by measurement teams established in six policy areas (economic vitality, government efficiency, health care, safety, transportation, and vulnerable children and adults). Teams were led by a GMAP staff person and composed of people from the agencies that work in the policy area and representatives from the governor's Policy and Budget offices. The GMAP staff person did considerable research, for example reviewing national literature; the state's Priorities of Government materials; the agencies' strategic plans, budgets and activity reports. Logic models were used to identify desired outcome measures and then connect the outcome measures to possible output, process and input measures. These measures were proposed to the agencies as "straw men."

The rationale underlying the use of the logic models is that while outcomes are the most desirable measures for showing results, outcomes cannot always be measured frequently enough; the data are often stale; and while an agency can influence the results of an outcome measure, it typically cannot control the results. On the other hand, outputs, which can be controlled, frequently have a major influence on an outcome. Therefore output, process and input measures are accepted, particularly if there is likely to be a high correlation between the results of the output/process/input measures and the outcomes. Selecting the measures for each policy area took six to eight weeks initially. Measures continue to evolve and are replaced as needed.

An example of this logic link can be seen with the child-abuse program. The outcome measure for preventing child abuse is instances of re-abuse, which is an indicator of keeping kids safe. The best output measure is one that measures whether responses to child abuse reports are provided soon enough to assure that stories have not changed and bruises have not faded. Reducing the target for interviewing a child from within 10 days of a reported abuse to within 24 or 72 hours of the reported abuse, depending on the risk of immediate harm to the child, greatly reduced instances of re-abuse.

The heart of GMAP is the meetings the governor and her leadership team (chief of staff, deputy chief of staff, policy director, budget director, GMAP chief, legislative director, directors of Departments of Personnel, Information Services and Risk Management) hold with the agency directors involved in each of the policy areas to review the past quarter's progress toward achieving specific results. The meetings, called GMAP forums, are held regularly and are open to the public.

Preparation for each forum is considerable. Each policy area has a measurement team, comprised of a lead GMAP analyst, analysts from each involved agency, and analysts from the Office of Financial Management. The team meets two or three times prior to the forum to prepare the GMAP performance report and guide preparation for the forum. This means the agency persons need to know what is in the report so there are no surprises, what is likely to be discussed, what the governor is currently interested in, etc.

The report presents three primary types of information: data, analysis and action plans. It may also present background or contextual information. Specifically, there is typically a chart or table presenting data for each performance measure; analysis of the data written by the agency with the advice of a GMAP analyst; and, if needed, an action plan that shows how performance will be improved. These action plans list what will be done, who is responsible to get it done, and when it will be completed. The report is distributed to the governor and the leadership team members prior to the forum.

The analysis presents several kinds of comparisons:

- Targets, and particularly target ranges, for those instances where targets are appropriate. The GMAP office works with the agencies to establish reasonable, not impossible targets, consistent with the governor's expectations.
- Prior period results, with an emphasis on trends, not a single period.
- Other states, which is a comparison the governor frequently focuses on.

- Pre-existing standards set by the federal government or the state legislature.
- Other industry benchmarks.

With regard to disaggregation, agencies tend to present one average number for the entire state. The GMAP analysts know the disaggregated data are available within the departments and encourage the agencies to present them to the governor.

The reports present the same core sets of measures each time. However, the governor has only about an hour for each forum and uses her time to focus on what is currently important-which is frequently the programs not going well. She asks questions based on the performance report, suggests new measures, asks for measures broken down by categories (regions, for example), and/or directs the agency to do something that is not already on an action plan. She will also ask the department heads, "What can the leadership team do to help you?" Following the forum, the GMAP analyst sends the agency directors a follow-up memo listing the items the agencies will have to report back on before or during the next forum.

Originally, a forum was held for a different policy area approximately every two weeks. They were complex, organized meetings open to the public. Because the forums consumed a significant amount of the governor's and agencies' time, they were not held on a regular schedule, and agencies could not easily integrate the process of performance reporting into their regular business cycle. As a result, another change was made. GMAP reports are now prepared quarterly for each policy area and public forums are held as the governor has time.

There are many advantages to forums:

- The presence of the governor and her staff in the room sends a powerful message to the departments that the administration is committed to management improvement.
- The tendency to look at the same measures again and again, accompanied by the relentless follow-up on the implementation of action plans, is a powerful driver for improvement, particularly in poor performing areas.
- Convening the forums by issue area rather than department results in agencies reaching out from their silos and addressing matters they might not otherwise address.
- The governor's willingness to address problems in an open setting encourages agency directors to embrace transparency
- Agency directors report improved cooperation among agencies, a more disciplined and data-driven approach to day-to-day decision-making, increased clarity of mission, and a renewed sense of confidence in telling the story of government to customers, clients and the public.

Agencies are required by the governor's Executive Order to operate "internal GMAPs." The governor does not dictate the format for the internal GMAPs, so there is wide variety. Some are mini-versions of the governor's GMAP, complete with reports and forums. Agencies that operate with boards frequently have their management make presentations to the boards. Others incorporate performance measure reviews into their existing management processes. All, however, rely on data to demonstrate movement toward the agencies' strategic goals.

In 2006 and 2007, the state undertook the Citizen Engagement project, a series of citizen meetings, workshops and town halls held throughout the state. The purpose of this exercise is to confirm that the state is using outcome measures that reflect what is important to the citizens and that the agencies are focusing on the right goals. This insight is obtained by asking the participants, "What is important to you?" "How would you know if you got it?" and "What information do you need to judge whether your state government is working?," and then using anonymous interactive polling, small group work sessions and debate to prioritize which issues matter the most to the participants and the importance of the performance indicators in each policy area.

It should be realized, however, that GMAP uses additional measures. Citizens focus on high-level outcomes. They do not care how the results are achieved, as long as they are achieved efficiently. Hence, management needs more measures to manage and achieve these efficiencies.

The legislature has not paid much attention to GMAP, although some members and their staffs have shown an interest, followed its progress and attended some forums. Whether GMAP performance reports are used in legislation is debated.

In addition to the aforementioned example involving responding to instances of child abuse, there are several other examples of performance measures being used to improve the delivery of services and/or use of resources. Some are:

- Performance data revealed that, while overall traffic accidents on state highways were declining, motorcycle accidents were increasing. This motivated several agencies (Department of Licensing, Department of Transportation, Traffic Safety Commission, Washington State Patrol) to work together to provide more education for manufacturers and operators and stricter licensing which, in turn, resulted in fewer accidents.
- Court wait times prevented social workers from visiting children and their families. Representatives of the King County Court system, Children's Services, and the attorney general came together to modify the court process. The result was a 25 percent reduction in social workers' court wait time.

Other items worth noting are:

- Washington does not have a state-level strategic plan, although agency-level strategic plans are prepared.
- There is an indication that with budget resources scarce, the legislature is looking for programs that are not effective and there is a slow cultural change toward accepting evidence-based input. For example, the director of Chil-

dren's Services brought a GMAP report to a hearing conducted by a legislative committee looking into children's issues and used GMAP measures to successfully support a budget request for more employees. Other departments are beginning to understand how their GMAP reports could be used in the same manner. In a stringent budget environment, GMAP can build the case for at least maintaining existing budget levels.

- Washington's accounting systems do not readily support activity-based costing.
- It is up to the agencies to assure the reliability of their data, and GMAP relies on the agencies. Hence, data reliability is occasionally of some concern. The agencies often complain their data are not perfect, but the GMAP office directs them to go with what they have while they work to improve the data quality. GMAP's position is "start with what we have now, because waiting for the elusive perfect data or perfect measure can be an infinite barrier."
- The governor expects personnel evaluations to consider program performance as the merit system is supposed to be the foundation for compensation levels. Whether promotion and compensation levels actually reflect program performance may vary.
- GMAP forums have provided a motivational focus for the state's employees. Agency heads have shown videos of the forums to their agencies' staffs, who are impressed that the governor is paying attention to and understands what the agencies are doing. This increases motivation and morale. The governor's practice of writing "kudo" letters based on performance also helps.
- There is considerable training in performance measurement for state employees, particularly for supervisors who need to know about GMAP. The state's personnel department provides extensive training in addition to online training and guides on GMAP's website including: *Using Logic Models to Bring Together Planning, Evaluation, and Action-A Downloadable "How To" Guide and Performance Measure Definition and Identification-A Downloadable "How To" Guide*.
- GMAP has an extensive website. Website visits are tracked and GMAP believes that agencies currently use the site to find GMAP resources, training and performance reports more than citizens use the website. GMAP is moving toward more citizen use of the website. This will be a challenge as performance reports are written for the governor's leadership team and may not be very clear to a non-state government audience.

In conclusion, GMAP is highly replicable, primarily because it is an adaptation of another system-Compstat and its offshoot, CitiStat. For a similar system to work, there must be commitment from the government's chief executive; a willingness to put the issues on the table and have conversations with the right people in the room (and it doesn't have to be a public forum); and management teams that accept the "geekiness" of performance measures, recognizing that they are not something extra, but represent basic day-to-day management.

Appendix C, End Notes

Appendix C: Helpful References

Dr. Robert Behn, *What All Mayors Would Like To Know About Baltimore's CitiStat Performance Strategy*, The IBM Center for the Business of Government, Managing for Performance and Results Series, 2007

Sam Michael McCall, *An Analysis of Local Government Performance Measurement Reports*, Doctoral Dissertation, May 2009.

A Performance Management Framework for State and Local Government: From Measurement and Reporting to Management and Improving, Public Review Draft, National Performance Management Advisory Commission, July 2009.

End Notes

1. Paul K. Brace, Robert Elkin, Daniel D. Robinson and Harold I. Steinberg, *Reporting of Service Efforts and Accomplishments*, FASB Research Report, 1980.

2. See the Pew Center on the States website at *www.pewcenteronthestates.org/initiatives_detail.aspx?initiativeID=36072.*

3. Governments establishing a system to use performance measures to improve service delivery are likely to have performance data for prior periods even in the first year of operating the system. Performance data comprise only one element in the system.

4. Although the research and the resulting defined system is concerned only with the use of performance measures to improve service delivery and not to inform budget development and enactment, the incorporation of this and the subsequent element into the system is likely to enhance the system's effectiveness.

5. Rudolph W. Giuliani with Ken Kurson, *Leadership*, Miramax books, Hyperion, New York, 2002, pp. 71-82.

AGA CPAG Research Reports Previously Published

No. 1, March 2005: *Audit Federal Financial Controls: Sooner Rather than* Later?

No. 2, July 2005: *Financial Management Shared Services: A Guide for Federal Users*

No. 3, November 2005: *Trends in Technology*

No. 4, April 2006: *The Federal Purchase Card: Use, Policy and Practice*

No. 5, June 2006: *Challenges in Performance Auditing: How a State Auditor with Intriguing New Performance Authority is Meeting Them*

No. 6, June 2006: *PAR—The Report We Hate to Love*

No. 7, February 2007: *The State Purchase Card: Uses, Policies and Best Practices*

No. 8, March 2007: *Federal Real Property Asset Management*

No. 9, May 2007: *Should State and Local Governments Strengthen Financial Controls by Applying SOX-Like Requirements?*

No. 10, April 2007: *Process-Based Financial Reporting*

No. 11, May 2007: *The State Travel Card—Uses, Policies and Best Practices*

No. 12, June 2007: *Trends in Technology—2007 Review*

No. 13, June 2007: *The Federal Travel Card—Uses, Policies and Best Practices*

No. 14, January 2008: *21st Century Financial Managers—A New Mix of Skills and Educational Levels?*

No. 15, July 2008: *SAS 70 Reports: Are they Useful and Can They Be Improved?*

No. 16, Sept. 2008: *XBRL and Public Sector Financial Reporting: Standardized Business Reporting: The Oregon CAFR Project*

No. 17, Nov. 2008: *Characteristics of Effective Audit Committees in Federal, State and Local Governments*

No. 18, Jan. 2009: *Grants Management: How XBRL Can Help*

No. 19, Feb. 2009: *Procuring Audit Services in Government: A Practical Guide to Making the Right Decision*

No. 20, March 2009: *Performance-Based Management*

No. 21, June 2009: *Trends in Technology—2009 Review*

No. 22, Sept. 2009: *Managerial Cost Accounting in the Federal Government: Providing Useful Information for Decision Making*

AGA®

Advancing Government Accountability

Association of Government Accountants

2208 Mount Vernon Avenue
Alexandria, VA 22301

PH: 703.684.6931
TF: 800.AGA.7211
FX: 703.548.9367

www.agacgfm.org
agamembers@agacgfm.org