Ensuring the Integrity of Crucial Data

William C. Rivenbark and Carla M. Pizzarella

The manager of Carolina City scans the headlines of the daily morning newspaper while preparing for a senior staff meeting. An article about performance measures in Tarheel Town, a neighboring city, catches his attention. Tarheel Town adopted an internal performance measurement system eighteen months earlier to address a general public perception of poor service delivery. According to the article, in a city council workshop the night before, the relatively new manager of Tarheel Town presented and discussed numerous performance statistics generated by the system, among them, 3.0 minutes average response time for high-priority police calls, 142 dispatched calls per police patrol officer, and 55 percent clearance rate for UCR (uniform crime reporting) Part I crimes assigned to investigators.[1]

The manager of Carolina City opens his city's annual operating budget and turns to the section on police services. He finds the following corresponding statistics: 2.8 minutes average response time for high-priority police calls, 155 dispatched calls per police patrol officer, and 29 percent clearance rate for UCR Part I crimes assigned to investigators. He highlights the statistics in his city's budget and clips the article in the newspaper, deciding to discuss the information with senior staff. Why, he wonders, does Tarheel Town's performance in these three areas differ from the performance of Carolina City, especially in the clearance rate for UCR crimes?

PERFORMANCE MEASUREMENT WORKSHEET

Agency Name: _____

Agency Code: _____

Contact Person: _____ Phone: _____

Agency Goals:

Service Area:

Service Area Objectives:

	Input Indicators	Output Indicators	Efficiency Indicators	Service Quality Indicators	Outcome Indicators
Indicator					
Indicator Calculation					
Data Source(s)					

REMINDER: Accuracy of the information provided above determines the data's usefulness for budgetary planning, managerial decision-making, and operational improvement.

Performance measurement is now a common management tool in local government, especially in larger jurisdictions.[2] It provides a systematic approach to assessing, monitoring, and improving service delivery by creating and tracking measures of workload, efficiency, and effectiveness.[3]

In recent years, to enhance their performance measurement systems, localities also have embraced benchmarking. "Benchmarking" involves tracking performance measures over time (trend analysis), comparing performance measures with established objectives or targets, or comparing performance measures with the results of other jurisdictions.[4]

Rivenbark is an Institute of Government faculty member specializing in local government administration. Pizzarella is a research associate specializing in public budgeting and performance measurement. Contact them at rivenbark@iogmail.iog. unc.edu and pizzarella@iogmail.iog.unc.edu.

Both Carolina City and Tarheel Town use performance measurement, and informally at least, the concerned manager of Carolina City is using Tarheel Town as a benchmark. But suppose the two jurisdictions are not basing their performance statistics on comparable data. That might account for the apparent difference in performance.

Auditing or verifying the accuracy of data used to create measures of performance is an area of performance measurement and benchmarking that has received minimal attention.[5] Localities are accustomed to having their financial data audited. They are not accustomed to having their performance data audited. A critical step in performance measurement and benchmarking is the establishment of an internal or external review process to ensure that performance measures are materially accurate, reliable, and comparable.[6]

This article presents the results of a performance measurement audit conducted by staff of the North Carolina Local Government Performance Measurement Project (the North Carolina Project) for the three functional areas of police services: patrol, investigations, and emergency communications. It begins by providing a brief overview of the North Carolina Project and the purpose of the audit. It then reports the audit results, discussing five general areas that affect the accuracy and the comparability of performance data.

Background

The North Carolina Project is a benchmarking consortium of fourteen municipalities, four professional organizations, and numerous local government officials. It was established to promote performance measurement in local government, to produce comparable performance and cost data for participating units, and to encourage use of the data to improve services or processes. Institute of Government faculty and staff manage it under the guidance of a steering committee made up of representatives from each participating jurisdiction. The project annually publishes a performance and cost data report on the service areas under study and the performance measures

ABBREVIATED AUDIT REPORT ON EMERGENCY COMMUNICATIONS

The scope of the performance measurement audit for emergency communications included the service profile form for that function, related performance and cost data in the Final Report on City Services for FY 1999–2000, and on-site interviews with emergency communications personnel for each participating unit. The audit was conducted during March and April 2001, concluding with an exit conference in May 2001.

Methodology

In February 2001, project staff developed a questionnaire and distributed it to each participating municipality. They used it to obtain initial information on how the municipality was collecting and reporting performance data. They then conducted on-site visits to detail how the participating municipalities were collecting and reporting data and to determine what their capabilities were for collecting and reporting other data. Project staff used the information from the on-site visits to generate the findings and the recommendations that follow.

Findings

1. A series of events occurs between the time a call requiring dispatch is placed and the time the call is dispatched. (For the steps in handling an emergency call, see Figure 1.) Variations such as routing patterns may exist across jurisdictions.

Figure 1. **Steps in Responding to Emergency Calls**

• Caller dials E-911 (or other number).

• Call is received by telephone company switching center and routed to Emergency Communications Center.

• Call rings in Emergency Communication Center.

• Call is answered by telecommunicator (beginning of talk time).

• Call is entered into CAD (computer-aided dispatch) system and routed, if necessary, to appropriate dispatcher.

• Call is ready for dispatch (and unit begins to respond if available).

• Call is held until unit is available to respond (if response is not immediate).

• Unit is assigned and responds to call.

---

2. Interpretation of the measure "Average time from receipt of call to dispatch, for calls resulting in a dispatch" has varied in the determination of when "receipt of call" begins:

• For one jurisdiction, "receipt of call" begins when the call is registered by the telephone company.

• In another case, "receipt of call" has been interpreted as being the moment when the telephone call is answered by the telecommunicator.

• In most cases, "receipt of call" has been interpreted as being the moment when the call is first keyed into the CAD system.

3. Interpretation of the measure "Average time from receipt of call to dispatch, for calls resulting in a dispatch" has varied in the determination of when "dispatch" begins. The definition of "dispatch" and the reporting structures of CAD systems appear to differ from jurisdiction to jurisdiction:

   • For at least one jurisdiction, "dispatch" represents the time when the telecommunicator makes the call available for dispatch. In this instance, "Average time from receipt of call to dispatch . . ." does not include "stack time" (the time the call was held).

   • For other jurisdictions, "dispatch" represents the time when a responding unit has been assigned and is ready to take the call. In this instance, "Average time from receipt of call to dispatch . . ." does include stack time.

4. Interpretation of the measure "Average time from receipt of call to dispatch, for calls resulting in a dispatch" has varied slightly in the determination of which calls are included in "calls resulting in a dispatch":

   • One jurisdiction includes only E-911 calls.

   • Most jurisdictions include all calls resulting in a dispatch, which could be calls to E-911 or calls to another telephone number.

5. Variations exist in the tracking capabilities of units to arrive at "total number of incoming calls":

   • One unit is able to track all non–E-911 calls but is unable to track E-911 calls, which are transfers from another center.

   • Some units are installing a system to track the number of all incoming calls. These units have been either providing estimates or not providing the total number of incoming calls.

   • Most units have a system that tracks the number of all incoming calls.

6. Variations exist in the types of calls included in the "total number of incoming calls." These variations are due to the different functions of each emergency communications center rather than differences in interpreting the question. Some units are primary centers; others are secondary. Some transfer calls to other units of government (for example, the sheriff or Emergency Medical Services); others receive calls that were transferred from another call center.

Recommendations

1. Add the measure "Calls dispatched per telecommunicator" to correspond with the existing measure "Calls answered per telecommunicator." The dispatch function is arguably the core element of emergency communications. Inclusion of this measure would provide an additional dimension of a telecommunicator's workload.

2. Replace the measure "E-911 calls only, answered per 1,000 population" with "Calls dispatched per 1,000 population." The number of calls dispatched provides a more comprehensive base of information than the number of E-911 calls. Calls dispatched can include calls that come through both the E-911 lines and other lines.

3. Replace the measures "Cost per call answered" and "Cost per E-911 call answered" with "Cost per call dispatched." Calls dispatched provide the most meaningful basis for cost information because call dispatch is the primary function of emergency communications.

associated with each participating jurisdiction.[7]

In December 2000 the steering committee requested a performance measurement audit for the three police functions named earlier. The purpose of the audit was to ensure that all the participating municipalities were interpreting the data requests on the service profile forms (the data collection instruments used by the project) correctly and consistently and that their processes of collecting and reporting data were not leading to a false appearance of performance deficiencies.[8] In other words, the audit should determine whether a difference in performance between one jurisdiction and another was the result of actual performance or the result of different interpretations of data requests or different processes of data collection.

Project staff conducted on-site visits during March and April 2001 to obtain information on how participating units were collecting and reporting data on patrol, investigations, and emergency communications and what their capabilities were for collecting and reporting other performance statistics. Subsequently, project staff distributed draft reports of their audit to the participating units for review and feedback. In May 2001 they held an exit conference with the steering committee to finalize the audit results and recommendations. (For an abbreviated audit report on emergency communications, see the sidebar on page 29.)

Audit Results

The North Carolina Project uses a benchmarking process that was specifically designed to collect and report accurate performance data. Much time was invested in creating service definitions, constructing detailed service profile forms, and cleaning the performance data submitted by participating jurisdictions.[9] The performance measurement audit of the three police functions provided another means to the end of accurate performance data. The audit revealed that benchmarking was working, but it also identified material differences in the performance data that were attributable to both controllable and uncontrollable factors.

Five factors were found to affect the accuracy and the comparability of performance data:

- Departmental changes
- Interpretation of definitions
- Reporting models
- Reporting capabilities
- Functional boundaries

These factors apply to systems that analyze trends or set performance targets, as well as to those that look at other jurisdictions' performance, especially in localities that collect and report data annually. The natural evolution of organizations over time requires that they periodically review the processes they use for service delivery and the performance data they gather to determine how accountable they are.

Departmental Changes

Among the many challenges that the North Carolina Project has experienced since its creation in fall 1995 is turnover of local government personnel who collect the performance data. Turnover has created difficulty with service definition, data interpretation, and data accuracy. A learning curve exists in performance measurement and bench-marking, and this creates a need for regular training of new personnel in data collection and data cleaning.

The situation has been further com-plicated as departments have altered their internal structures for service deli-very, changing the types of data they produce and the processes by which they collect data. An example of such a change is an investment in new tech-nology. Since the inception of the project, several jurisdictions have changed from one kind of computer-aided dispatch (CAD) system (a system for tracking dispatched calls) to another, or upgraded their existing system, and this has affected how they gather and report data. Another example is a change in organizational structure that creates processes crossing over departmental and divisional lines. Although departmental changes may cause difficulties with data accuracy, they are natural occurrences within organiza-tions and should be accounted for in the collection of performance data.

4. Provide clear definitions of start and end times for the measure "For calls dispatched, number of seconds from receipt of call to dispatch." Reword the measure to read "For calls dispatched, number of seconds from CAD entry to dispatch." This measure would represent the interval from when the call is first keyed into the CAD system to when the call has been assigned and dispatched to a responding unit, including stack time if necessary.

5. Regarding "calls dispatched," include self-initiated calls except for administrative events, duplicates, or calls related to another call. Also include telephone responses and walk-ins.

6. Delete the measure "Sustained complaints per 100,000 calls answered" because of variations in definitions of "sustained" and "complaint" among jurisdictions.

7. If information is available only through sampling or estimation, write "N/A" (not available) on the service profile forms.

8. When possible, report raw service data on the service profile forms, allowing project staff to calculate the performance measures.

These recommendations are summarized in Table 1.

Table 1. Performance Measures for Emergency Communications

Existing	Proposed
• Total calls answered per 1,000 population	• Total calls answered per 1,000 population
• E-911 calls only, answered per 1,000 population	• Calls dispatched per 1,000 population
• Calls answered per telecommunicator	• Calls answered per telecommunicator
• Cost per call answered	• Calls dispatched per telecommunicator
• Cost per E-911 call answered	• Cost per call dispatched
• For calls dispatched, number of seconds from receipt of call to dispatch	• For calls dispatched, number of seconds from CAD entry to dispatch
• Number of seconds from initial ring to answer	• Number of seconds from initial ring to answer
• Percentage of calls answered within three rings (18 seconds)	• Percentage of calls answered within three rings (18 seconds)
• Sustained complaints per 100,000 calls answered	

Epilogue

During the exit conference in May 2001, police personnel reviewed and accepted the recommendations and the proposed performance measures contained in this report, and in July 2001 the project steering committee approved them. The service profile form for emergency communications was adjusted for data collection beginning in August 2001.

Reviewing the accuracy of data is a necessary step in performance measurement and bench-marking.

Interpretation of Definitions

The audit revealed that units were providing different types of data because of differences in their interpretation of questions on the service profile forms. One of the performance measures for police investigations is "Percentage of UCR Part I crimes cleared, of those reported." Jurisdictions are supposed to provide the number of crimes reported during the fiscal year as well as the number of cases cleared. Most were providing the total number of cases cleared within the fiscal year, regardless of when the crimes were reported; however, some were providing only the number of cases cleared that were reported during the corresponding fiscal year.

Jurisdictions also varied in their determination of when "receipt of call" began and when "dispatch" occurred, for the emergency communications measure "Number of seconds from receipt of call to dispatch." One jurisdiction began tracking calls when the telephone company registered them. Others tracked calls once they were entered into the CAD system. In terms of "dispatch," some included "stack time" (the time the call was held); others did not. Interpretation of when

receipt of call begins or when a dispatch occurs may result in a performance difference of only a few seconds. However, this is a critical measure for emergency communications and should be tracked in a comparable format.

Reporting Models

Each department had its own reporting models and systems to record data. Some variations existed in the way the data were captured, categorized, and reported. For the patrol performance measure "Average response time to high-priority calls," some units collected response times for all high-priority calls, including self-initiated calls, which have a response time of zero. Other jurisdictions omitted self-initiated calls, thereby increasing their average response time.

Many of the measures for patrol and investigations include the number of UCR Part I crimes that have been reported during the fiscal year. Approximately half of the departments tracked crime on the UCR model, whereas the other half tracked it using incident-based reporting.[10] These two reporting models classify crimes differently, resulting in comparability problems. As a result of the audit, project staff intend to address these differences in the narrative section of the project's annual report of performance and cost data for the participating jurisdictions.

Reporting Capabilities

Some units were unable to comply fully with the information requested for particular measures, instead providing estimates based on conversions of partial-year data, subsets, or samples. The limited reporting capability of CAD and other tracking systems was the primary reason for using estimates. For example, one unit did not have a

> **Performance measurement represents a management tool for analyzing the operational results of local government. It provides a framework for accountability, planning and budgeting, operational improvement, program evaluation, and allocation of limited resources.**

system installed to record when incoming calls were answered. Therefore the unit conducted sampling to provide data for the measure "Percentage of incoming calls answered within three rings." Several jurisdictions now are adding, upgrading, and replacing systems, enabling them to provide complete information in the future.

Functional Boundaries

Each service area encompassed in the North Carolina Project has a corresponding service definition that describes in detail the resources and the personnel with direct involvement in that service's activities. The service definition provides the boundaries for determining the cost of a function as well as the personnel who deliver the service. The audit revealed that police departments had a clear understanding of the personnel, functions, and resources involved in emergency communications but not of those involved in patrol and investigations. Each service definition assumes that the service is a separate and distinct function with clearly identifiable personnel. The personnel and the resources being used in patrol and investigations were not completely separate. Inaccurate cost and performance data for these two services were being created as a result.

In the case of patrol, variations existed in the methodology for determining which positions to include in the number of patrol officers. Jurisdictions varied in their inclusion or exclusion of positions such as school resource officers, supervisors, traffic officers, and other special unit positions. It became evident that patrol was a complex function that could not be defined consistently across jurisdictions. Further, the duties of patrol officers were not limited to the patrol function. The multiple roles of patrol officers

indicated a need for a broader service definition.

Although most departments had investigative units, people outside those units were handling a portion of the investigative work and case clearance. For example, many jurisdictions assigned certain types of crime investigations to patrol officers. Although the patrol officers and others were contributing to the investigative function, most units were excluding these costs from the total cost of police investigations.

As a result of the audit, the steering committee decided to combine the patrol and investigation functions into a broader service area. The need for accurate and comparable data was the overriding factor in this decision. It was understood that trend data and detailed measures for each functional area would be lost.

Conclusion

Performance measurement represents a management tool for analyzing the operational results of local government. It provides a framework for accountability, planning and budgeting, operational improvement, program evaluation, and allocation of limited resources.[11] Benchmarking expands the usefulness of performance measurement by providing local governments with a methodology to identify best practices for service delivery. Examples of North Carolina Project members using the performance and cost data to improve service provide evidence that benchmarking can yield tangible results.[12]

Two major components of performance measurement and benchmarking are collection of accurate and reliable performance data and use of those data to improve services or processes.[13] This article demonstrates the need for an additional component of performance

measurement and benchmarking: auditing of performance data.

An aspect of benchmarking highlighted by the audit was the influence of a jurisdiction's business practices on its performance, as revealed by the selected performance measures. Although business practices do not necessarily influence the accuracy and the reliability of performance data, they affect service outcome as measured by performance statistics. Each jurisdiction makes choices in service delivery, including level of service, priorities, and procedures for providing the service. An example involves the measure "Average response time to high-priority police calls." A vehicle accident without personal injury may warrant a high-priority response in one jurisdiction. Other jurisdictions may not consider this occurrence high priority because personal injury was not present. The understanding of differences in business practices is critical to identifying best practices in a bench marking initiative for service or process improvement, and to preventing comparability problems from a limited amount of information.

. . . .

The manager of Carolina City presents the newspaper article and the Carolina City police department's performance measures to senior staff, requesting that the police chief contact the Tarheel Town police department for clarification on the measures. The police chief's review reveals that the neighboring jurisdiction uses a different definition of high-priority calls, does not include telephone responses in its total dispatched calls, and uses stringent solvability factors for assigning cases to investigators.

Comparison of Performance Measures

Measure	Carolina City	Tarheel Town	Explanation of Variation
Response time to high-priority calls, in minutes	2.8	3.0	Different definitions of high-priority calls
Dispatched calls per patrol officer	155	142	Different types of calls recorded for dispatched calls
% of UCR Part I crimes cleared, of those assigned to investigators	29	55	Different methodology for assigning crimes to investigators

Involving service providers in audits of performance data is critical to success.

The police chief prepares a comparison (see page 33) and forwards it to the city manager for review.

The manager receives a number of inquiries from council members who have read the newspaper article and made their own comparisons. They are concerned about the 29 percent of UCR Part I crimes cleared, compared with Tarheel Town's rate of 55 percent. The community has experienced an increase in crime the past two years, and a low clearance rate is going to be difficult to explain to their constituents. The city manager provides an overview of the comparability problem and says that, using Tarheel Town's definitions of performance data, the police chief is calculating comparable measures for council members' review. On the basis of this experience, the city manager decides to implement an annual review of selected performance data to ensure that the three years of performance measures presented in the annual operating budget are accurate, reliable, and comparable.

Notes

1. UCR (uniform crime reporting) is a system used by law enforcement agencies to report crime at an aggregate level to the State Bureau of Investigation and the Federal Bureau of Investigation.

2. Evan Berman & XiaoHu Wang, *Performance Measurement in U.S. Counties: Capacity and Reform,* 60 Public Administration Review 409 (Sept./Oct. 2000); Theodore H. Poister & Gregory Strieb, *Performance Measurement in Municipal Government: Assessing the State of the Practice,* 59 Public Administration Review 325 (July/Aug. 1999).

3. Salomon A. Guajardo & Rosemary McDonnell, An Elected Official's Guide to Performance Measurement (Chicago: Gov't Finance Officers Ass'n, 2000).

4. David N. Ammons, Municipal Benchmarks (2d ed., Thousand Oaks, Cal.: Sage, 2001).

5. Katherine Barrett & Richard Greene, *Truth in Measurement,* 13 Governing 86 (Sept. 2000).

6. The goal of an auditor is to determine if financial data are reasonable, not absolute. "Materially accurate" is an auditing term for "reasonable," representing that financial statements are fairly presented. For more information on materiality, see Stephen J. Gauthier, Governmental Accounting, Auditing, and Financial Reporting (Chicago: Gov't Finance Officers Ass'n, 2001).

7. William C. Rivenbark & Carla M. Pizzarella, Final Report on City Services for Fiscal Year 1999–2000 (Chapel Hill: Inst. of Gov't, The Univ. of N.C. at Chapel Hill, 2001).

8. For copies of the service profile forms, *see* William C. Rivenbark, A Guide to the North Carolina Local Government Performance Measurement Project (Chapel Hill: Inst. of Gov't, The Univ. of N.C. at Chapel Hill, 2001).

9. Service definitions identify the specific activities encompassed by the service as well as the resources and the personnel involved. Data cleaning involves examining the information for reasonableness, comparing it with data from previous years, and reviewing it with service and line managers.

10. Incident-based reporting is a system used by law enforcement agencies to report crime at a detailed level on the basis of data elements and values for describing each incident or arrest.

11. Ammons, Municipal Benchmarks.

12. David N. Ammons, *Benchmarking as a Performance Measurement Tool: Experiences among Municipalities in North Carolina,* 12 Journal of Public Budgeting, Accounting & Financial Management 106 (Spring 2000).

13. Dale R. Geiger, *Practical Issues in Level of Precision and System Complexity,* 49 Government Accountants Journal 28 (Summer 2000); *Avoiding Pitfalls in Managerial Costing Implementation,* 50 Journal of Government Financial Management 27 (Spring 2001).