MAY 2003

How Federal Programs Use Outcome Information: Opportunities for Federal Managers

Managing for Results Series

Harry P. Hatry
Elaine Morley
Shelli B. Rossman
Urban Institute
Washington, D.C.

and

Joseph S. Wholey
Professor of Public Administration
University of Southern California

IBM Endowment for
The Business of Government

MANAGING FOR RESULTS SERIES

How Federal Programs Use Outcome Information: Opportunities for Federal Managers

Harry P. Hatry
Elaine Morley
Shelli B. Rossman
Urban Institute
Washington, D.C.

and

Joseph S. Wholey
Professor of Public Administration
University of Southern California

May 2003

Research supported by:

IBM Endowment for
The Business of Government

National Academy of Public Administration

TABLE OF CONTENTS

FOREWORD

May 2003

On behalf of the IBM Endowment for The Business of Government and the National Academy of Public Administration's Performance Consortium, we are pleased to present this report, "How Federal Programs Use Outcome Information: Opportunities for Federal Managers," by Harry Hatry, Elaine Morley, Shelli B. Rossman, and Joseph Wholey. This important report was jointly funded by both organizations.

This year marks the 10th anniversary of the passage of the Government Performance and Results Act. The authors of this law had hoped that by requiring agencies to develop strategic and annual plans and report on performance outcomes, government would change the way it did business. The intent was to encourage program managers, budgeters, and policy makers to increase the use of performance and outcome information in their decision-making processes.

This report demonstrates that program managers across government are indeed collecting and using program outcome data to make management decisions on how to best get results in delivering their programs to the American people. The authors of this report found that managers are using outcome information to trigger corrective actions, identify and encourage "best practices," motivate employees, and make planning and budget decisions.

The report is targeted to federal program managers. It argues that program managers have the greatest stake in obtaining and using the types of information described in this report. The report demonstrates to program managers that many of their peers are successfully using performance outcome information to manage their programs. Based on the successes reflected in the 16 case studies described, the authors present recommendations on how the use of such information can be expanded. We trust that this report will be informative and useful to managers across government as they increase their use of outcome information.

Paul Lawrence
Co-Chair
IBM Endowment for
the Business of Government
paul.lawrence@us.ibm.com

Philip Burgess
President
National Academy of Public Administration
pburgess@napawash.org

Purpose, Scope, and Method

Purpose

The primary audience for this report is federal program managers. The principal aim is to provide information and recommendations to them that will encourage and better enable them to effectively use regularly collected outcome information to improve their programs. (See "Definitions for Terms Used in This Report.")

The Government Performance and Results Act of 1993 (GPRA) requires federal agencies to provide annual data on the outcomes of each major federal program. Each year federal programs provide annual performance plans and annual performance and accountability reports.

In 2002, the Office of Management and Budget (OMB) introduced the Program Assessment Rating Tool (PART) as an additional requirement in which OMB develops, with input from federal agencies, ratings of the results produced by individual federal programs.

Outcome information such as described in this report can be used to help programs in responding to both the PART and GPRA processes. (Use of outcome information to improve service delivery is one of the stated purposes of GPRA).

The primary focus, thus far, of both these efforts has been to strengthen accountability for program results. Regularly collecting such outcome information, however, inevitability requires substantial time and effort, and often additional cost to federal programs to obtain and report the outcome information. The emphasis to date has *not* been on encouraging program managers to use the regularly collected information to improve their services and thus the results of those services.

Many federal program managers lament that they have little authority. However, they can and do regularly make many decisions and take action much more quickly than policy officials. The following quote in the *Washington Post* from such a policy official, Tommy G. Thompson, secretary of the Department of Health and Human Services (HHS), illustrates that point.

> … If I have an idea, I have to vet it through all the various divisions and agencies in this department alone. Then, if I can get any degree of unanimity of support for my idea, then it goes over to the supergod called the Office of Management and Budget. And they vote you down nine times out of 10, just to show you who the boss is. Then if you do get it through OMB, then it goes to the White House and the intelligentsia over there. They want to show you that they're in charge, so you usually have a very difficult time getting [it] through them. If the president likes the idea, it goes on to Congress, and if Congress ever does approve it, then it's time to retire.[1]

It is our belief that federal policy makers have used outcome information for many years, even before GPRA was passed in 1993. Outcome data are cer-

tainly not new to the federal government. It has traditionally collected such data as the following:

- Numbers and rates of traffic accidents and traffic mortality rates
- Incidence of various diseases
- Rates of infant mortality and low-weight births
- Employment and unemployment data
- Numbers of reported child abuse cases
- National Assessment of Educational Progress (NAEP) and other scores relating to educational achievement
- School dropout rates
- Number and rates of reported crimes (grouped by various categories of crime)

In all these cases, it appears obvious that federal programs have on occasion (and sometimes frequently) taken action based on outcome data. These are a few of the better known outcome data items. Although the federal government has collected some outcome data for many years, such data have tended to be less used by individual program managers than by higher level government officials for major policy decisions.

Even before GPRA, some federal program managers made use of outcome information to help them improve their services. Some federal programs have been collecting data on outcomes for many years. It is just common sense to use feedback on the results of what you are doing to help decide what the program should do in the future.

Despite limited control over what they do, federal program managers can have considerable influence over the program's procedures. They have a role in developing the proposed budgets for their programs, and they are key to motivating their own personnel and contractors to perform. Although such actions are not nearly as "global" or influential as those of policy officials, they nevertheless can have important and significant effects on the public.

It is a premise of this report that a major use, if not the major use, of regularly collected outcome

Definitions for Terms Used in This Report

Input: Resources (expenditures or employee time) used to produce outputs and outcomes.

Output: Products and services delivered. Outputs are completed products of internal activity: the amount of work done within the organization or by its contractors (such as miles of road repaired or number of calls answered).

Outcome: An event, occurrence, or condition that is outside the activity or program itself and is of direct importance to program customers or the public. We also include indicators of service quality, those of importance to customers, under this category.

Intermediate outcome: An outcome that is expected to lead to a desired end but is not an end in itself (such as service response time, which is of concern to the customer making a call or requesting a service but does not indicate anything directly about the success of the call or request). A program may have multiple intermediate outcomes.

End outcome: The end result that is sought (such as the community having cleaner air or reduced incidence of disease). A program may have more than one end outcome.

Outcome indicator: A numerical measure of the amount or frequency of a particular outcome.

Performance indicator: A specific numerical measurement for one aspect of performance (for example, output or outcome) under consideration.

information should be by program managers themselves to improve the effectiveness of their programs. This report describes federal agency internal uses of regularly collected outcome information. The intent of this report is to encourage greater use of regularly collected outcome information by federal program managers to improve the outcomes of their programs.

This report identifies a number of examples of such use and provides recommendations for more effective use of outcome information by program managers in the future.

Scope

We limited our effort to examples that:

- Used outcome data, where outcomes include "intermediate" or "end" outcomes.
- Used outcome data that were regularly collected, that is, the outcome information was obtained at least annually. Thus, we excluded examples in which the outcome data used by the program were obtained from ad hoc, one-time program evaluations. One-time program evaluations can be of considerable help to program managers, but are beyond the scope of this report.
- Used the outcome data internally. We have not included examples in which the sole use was to provide reports to higher levels for their use (usually an accountability use) or examples of outcome data being posted on websites or otherwise reported to the public.

We have included examples that began before GPRA was passed, that are very recent, and that may only now be emerging. We included examples whether or not the outcome data used were included in the program's GPRA reports.

We have included examples of three categories of federal programs:

- Programs that provide services directly to citizens or to benefit citizens
- Programs that provide services indirectly, that is, programs that work through other levels of government or other organizations that in turn provide the service to the ultimate customers
- Regulatory programs

Table 1 identifies cases included in the report and categorizes each one.

We did not seek examples from foreign service or defense programs. However, outcome information should be of considerable use to these programs.

Nor have we included uses made of regularly collected outcome information by higher level public officials—such as OMB, the White House, and Congress—to help with resource allocation, policy decisions, and so on. Such high-level uses are very important but are not the subject of this report. A few of the examples included may have involved approval at the agency level, however, and one or two involved legislative action.

Method

We sought to identify a small number of programs that had or are currently making some use of regularly collected outcome information.

We obtained leads for examples from many sources, such as suggestions from current and former members of OMB and the General Accounting Office (GAO), and our own knowledge of federal programs. In some instances, we began with written reports or other publications. For each example, we sought telephone or personal interviews with persons in the agency to obtain details on the uses and results. We also examined relevant program documents relating to each example. We did, however, have to rely heavily on respondents' self-reporting as to the extent of their program's use of outcome information.

The persons we interviewed included program managers, members of their staffs, and "central" staff personnel who were familiar with at least some of the details.

In some cases, to find persons able to provide information on the program's uses of the outcome data, we had to obtain referrals from numerous individuals until we found the appropriate respondent. Often, some form of clearance was needed before we could obtain the information. One federal agency declined to participate, not wanting any publicity even if it would be favorable.

Table 1: Programs Included in This Report

Department and Program	Category
1. Agriculture—Fruit Fly Exclusion and Detection	Regulatory/direct service
2. Education—Adult Education and Literacy	Indirect service
3. Education—Migrant Education	Indirect service
4. Education—Special Education and Rehabilitative Services	Indirect service
5. EPA—Air and Radiation	Regulatory
6. EPA—Enforcement and Compliance Assurance	Regulatory
7. HHS—FDA's Generic Drugs	Regulatory
8. HUD—Public and Indian Housing	Indirect service
9. Interior—BLM's Wild Horse and Burro	Direct service/regulatory
10. Interior—National Park Service	Direct service
11. Labor—Occupational Safety & Health	Regulatory
12. USPS	Direct service
13. SSA—Supplemental Security Income	Direct service
14. Transportation—Coast Guard's Marine Safety, Security, and Environmental Protection	Regulatory
15. Treasury—IRS Toll-Free Customer Service	Direct service
16. Veterans Affairs—Veterans Health Administration	Direct service

Using this process generated varying numbers of examples and varying levels of detail about them. Some agencies were able to provide more contacts, such as regional or state office personnel who provided examples of use of outcome information at those levels. Our resources permitted us to obtain examples from only a few such contacts.

Due to differences in the number of respondents and their willingness or ability to spend time providing detailed information, the degree of information provided in the examples presented in this report varies considerably.

Due to these limitations, the examples do not reflect the full use of outcome information by these agencies or programs.

Remainder of This Report

The next section, "Summary of Findings," provides a summary of our findings and identifies obstacles to the use of outcome data by programs.

Following that is "Recommendations to Federal Program Managers," which provides recommendations to federal programs and their managers for improving the use and usefulness of regularly collected outcome information.

The final section, "Federal Program Case Studies," is the heart of this report. It provides descriptions of 16 programs. For each federal program for which we found an example of use, we first describe the program and its objectives. We then provide a summary of the type of uses program managers made of regularly collected outcome data that were identified through this process. We continue with basic

information on the outcome measurement process the program used, including identifying particular outcome indicators used by the program. Then we describe examples of specific uses the program made of the outcome data. When available, information on effects of the uses of the outcome data is provided.

We hope that this report will encourage program managers and personnel throughout the federal government to consider outcome information as giving them considerably enhanced opportunities to improve services.

Acknowledgments

This effort was supported with funds from the IBM Endowment for The Business of Government and the National Academy of Public Administration's Performance Consortium. Mark Abramson and John Kamensky of the Endowment and Chris Wye of the Academy provided highly useful suggestions as our team began its work.

We received excellent cooperation from a number of federal personnel, who provided the basic information for this report. These are listed in the appendix in alphabetical order by agency. Other personnel at these agencies also provided assistance in the form of referrals to colleagues who might be able to provide information.

Finally, we give special thanks to Joe Wholey, one of our team members, whose assistance was provided *pro bono*. He generously gave of his time to help document examples for the report and contribute to its writing.

We greatly appreciate the contributions of those mentioned here and any others inadvertently omitted, and acknowledge our responsibility for any errors herein.

Acronym List

APHIS	Animal and Plant Health Inspection Service
BLM	Bureau of Land Management
CAMP	College Assistance Migrant Program
CAP	corrective action plan
CNO	chief network officer
CSCS	Consumer Service Card System
CSM	Customer Satisfaction Measurement
DAEL	Division of Adult Education and Literacy
EPA	Environmental Protection Agency
EXFC	External First-Class Measurement System
FDA	Food and Drug Administration
GAO	General Accounting Office
GED	general equivalency diploma
GPRA	Government Performance and Results Act of 1993
H&S	health and safety
HEP	High School Equivalency Program
HHS	Department of Health and Human Services
HMA	herd management area
HUD	Department of Housing and Urban Development
IRS	Internal Revenue Service
LWDII	lost workday injury and illness
MOA	memorandum of agreement
NAEP	National Assessment of Educational Progress
NPS	National Park Service
NRS	National Reporting System
OAR	Office of Air and Radiation
ODIS	Origin–Destination Information Sampling System
OECA	Office of Enforcement and Compliance Assurance
OGD	Office of Generic Drugs
OMB	Office of Management and Budget
OSHA	Occupational Safety & Health Administration
PART	Program Assessment Rating Tool
PHAS	Public Housing Assessment System
PIH	Office of Public and Indian Housing
PIP	program improvement plan
PWI	Projects With Industry
REAC	Real Estate Assessment Center
RSA	Rehabilitation Services Administration
SSA	Social Security Administration
SSI	Supplemental Security Income
SST	site-specific targeting
TARC	Troubled Agency Recovery Center
USPS	United States Postal Service
VA	Veterans Administration
VHA	Veterans Health Administration
VISN	Veterans Integrated Service Network
VR	vocational rehabilitation
VSC	Visitor Survey Card
VSP	Visitor Services Project
WEB-EIS	WEB–Ease of Use
WHBP	Wild Horse and Burro Program

Summary of Findings

Here we identify three major overarching findings of this work. We continue by discussing problems in identifying examples of use, and we conclude with observations.

Many Federal Programs Have Already Made Use of Regularly Collected Outcome Data to Help Them Improve Their Programs

Do federal agencies collect outcome data solely to respond to reporting requirements? This report indicates that by no means is this universally the case. The last section of this report, "Federal Program Examples," presents the 16 individual federal programs we examined and their uses of outcome data to help them improve their programs.

We found examples of use of outcome information in all three *categories of federal programs* we examined: programs that provide services directly to ultimate customers (whether citizens or businesses); programs that provide indirect services by working through other levels of government or other organizations that, in turn, provide service to the ultimate customers; and regulatory programs. (See Table 2 for more detail. Table 1 lists our categorizations for each example.) Our small group of examples indicates that considerable potential exists for the use of regularly collected outcome information in all three categories.

The examples of use include a range of *program sizes*—small, medium, and large—and describe applications of outcome information in terms of the likely effects of the use. Whether small or large, all the uses described appear to be ones likely to be ultimately beneficial to the public. Even managers of small federal programs can make a difference, and they should be encouraged to make use of the outcome data on their programs.

Some of the examples involve *field offices/facilities*. Outcome information has been passed down to local offices for their use in their decision making or to motivate local offices to improve outcomes. Usually the data provided to the local field offices include data on other similar offices, thus also providing a motivational effect for field offices to keep up with, or exceed, the outcomes of their peers. See examples 15 (Internal Revenue Service) and 13 (Supplemental Security Income).

Federal Managers Have Used Outcome Data in a Variety of Ways

Federal managers reported using outcome data in numerous ways. Federal managers we interviewed often identified three or more uses of outcome data for a single program. Table 2 provides a summary of the types of uses we found, along with an identification of the particular examples in Table 1.

Table 2: Uses of Outcome Information by Federal Agencies

The numbers after each use correspond to the program numbers summarized in Table 1. The uses identified in the examples do not always map cleanly into these use categories. Some overlap exists among these categories.

Trigger Corrective Action	• Identify problem areas and modify service provision/operational practices (present in numerous examples) • Identify root causes of problems and develop action plans to address them (1, 15) • Trigger enforcement activities (5, 8, 11) • Identify grantee technical assistance and compliance assistance needs (2, 3, 4, 8, 11) • Develop training or guidelines for regulated entities (14) • Identify staff training needs and provide training (10) • Reduce or de-fund poor performers (grantees or contractors) (4) • Require grantees to provide corrective action plans (5, 8) • Evaluate the extent to which changes in practices and policies have led to improvements in outcomes (9, 13) • Identify the need for policy or legislative changes (4, 13) • Identify underserved "client" groups (4)
Identify and Encourage "Best Practices"	• Identify successful grantee practices (4, 8) • Disseminate good practices information (4, 8)
Motivate	• Motivate staff (present in numerous examples) • Develop employee performance agreements (16) • Use as basis for "How are we doing?" meetings (4, 6, 7, 10, 16) • Recognize and reward high-performing federal offices or employees (12, 13, 16) • Recognize and reward high-performing grantees or regulated entities (2, 8) • Motivate grantees or regulated entities (2, 3, 4, 8, 11)
Plan and Budget	• Allocate resources and set priorities (1, 2, 5, 8, 11, 16) • Develop plans and set targets (9, 15) • Justify requests for funds (1, 3, 10) • Determine grantee funding (3) • Inform budget projections (2)

We believe the uses discussed in this report are just the tip of the iceberg. Many of these examples may appear to be routine and obvious—hardly earth-shaking. But that is the point. As outcome information increasingly becomes routinely available to federal program managers, they are likely to use the information routinely in their decisions, such as to help them allocate their inevitably limited resources.

A few of our examples include use of outcome data not only *in aggregate form* but also for *individual cases*. For example, in the Occupational Safety & Health Administration (OSHA) and Supplemental Security Income (SSI) examples, the outcome data are reported by individual employer or claimant. Department of Housing and Urban Development (HUD) housing assessment data are provided for individual public housing authorities. When impor-

tant problems are found in these particular cases, federal employees can subsequently take action to correct those problems, as indicated in those examples.

Many Obstacles to Outcome Data Use Do Indeed Exist

Although our findings on use of outcome data are encouraging, obstacles exist to the use of outcome data. Here we identify four major types of obstacles. "Recommendations to Federal Program Managers," the next section, provides suggestions for alleviating many of these obstacles and problems in outcome data use.

Lack of Authority or Interest to Make Changes

We have frequently heard the view that federal program managers and their staffs have little authority to make changes or interest to do so. Obviously, upper level officials in the executive branch and Congress have considerably more and broader authority. In effect, program managers have only delegated authority. They, of course, must act within their (and the department lawyers') interpretations of their program's legislative wording and their perceptions of what executive branch and congressional leaders want. Seldom are the limitations of their authority fully fleshed out. The incentives for program managers tend to be to walk cautiously and not upset applecarts.

Nevertheless, we believe that the examples presented here make a real case that program managers have sufficient responsibility, if not the actual authority, to make many changes. They may not have sole authority to make changes; they may need the approval of upper level government officials to make certain types of change. Nevertheless, the program manager has a major role in identifying and implementing such improvements.

Limited Understanding of Use of Outcome Data

It was difficult for many program managers to identify uses of outcome data for this report because they did not appear to recognize that they were using these data. Similarly, it appears that at least some managers may not be aware of the variety of ways in which they can use outcome data, or which particular data can be used for specific purposes. We hope the examples provided in this report help address this.

Outcome Data Problems

Outcome data use is sometimes limited by data problems:

- Some data are old by the time they reach program managers; data that are two or three years old do not have the currency program managers usually need to react.
- Even if the outcome data are not "old," the timing at which the data become available may not appear to permit certain uses. For example, decisions on continuation grants due at the end of a federal program year may need to be made before outcome data are available for that program year.
- Some of the data may not be broken out in sufficient detail to be useful to particular program managers.
- Some data may not be sufficiently specific to particular federal programs; the data may be about issues that primarily provide information for higher level use. The data collected may be on a much broader scale than that needed by the program. For example, national samples may be obtained that do not provide the detailed data needed by specific programs.
- Many federal programs, such as those that work through other levels of government, require long periods of time before they can be expected to yield the major outcomes sought. Other data on intermediate outcomes are also needed to provide program managers with actionable information in the short term.

Fear of "Rocking the Boat"

Reluctance to make changes works against use of outcome information. This can be caused by the program's, or upper management's, fear of rocking the boat or being perceived as running an ineffective program.

Problems in Identifying Examples of Use

We found that it was not always easy to get people to identify their internal use of outcome data. In our discussions with agency personnel about their programs (including personnel from programs that we have not included in this report), it was difficult for them to identify ways they used outcome data, beyond responding to upper level requirements to submit the data. We believe this was due to one of these factors:

- They actually have not been using the data.
- They have used the data but do not recognize that they have used it.

There are probably several reasons for the latter. It is common sense to use available outcome information as feedback on programs to help make decisions about those programs. Managers may do this unconsciously, and such uses may not come to mind when they are asked how they use outcome information. And because outcome data are only part of the information managers consider before taking action, they may not recall the role that the outcome information played.

Another reason for not recognizing the use of outcome information is that, in some cases, programs use too narrow a view as to what can be labeled an outcome. This tends to occur in situations in which the program considered an indicator as being an "output" when it could legitimately be considered as an "intermediate outcome." An example is that of the Food and Drug Administration (FDA). It labeled a key performance indicator—"length of time to approve (or reject) prescription drugs"—for which data have been frequently used by the administration to make program adjustments, as an output. However, because the length of time is of considerable importance to citizens who might be helped by using the drug, it can legitimately be considered an intermediate outcome indicator.

Observation

We believe that the examples in "Federal Program Examples" touch only the surface of the numerous uses commonly being made by federal program managers of regularly collected outcome information.

But even this modest effort captured a rich variety of uses. The types of uses shown in "Uses of Outcome Information by Federal Agencies" provide examples that cover a great variety of ways that federal program managers can use outcome information to better manage and ultimately improve services to the public.

Recommendations to Federal Program Managers

In this section we identify three major areas for action. Each includes a number of action steps that federal program managers should consider to increase the use, and usefulness, of the outcome information they receive. Recommendations 1 through 6 focus on making data more useful to managers. Some of these actions need to be implemented by those who provide outcome data to program managers. Program managers may want to request that these actions be adopted to make data more useful to them. Recommendations 7 and 8 focus on providing guidance or training to encourage greater use of outcome information. These actions generally need to be undertaken by higher level managers in an agency. Recommendations 9 through 17 are examples of widely applicable uses of outcome information that can be made by federal program managers.

Make Outcome Data More Useful

1. Consider the uses presented in Table 2.
Assess the extent to which each is applicable, in some version, to the outcome information collected for your program. Implement those uses if you are not already using outcome information this way.

As noted previously, even managers of small federal programs can make a difference. Be encouraged to make use of the outcome data relating to your programs.

2. Break out (disaggregate) the outcome information you receive by important customer and service characteristics, if that is not already being done.
Consider breakout categories of customers such as geographical location, age, gender, income group, race/ethnicity, educational level, and household size. Consider service characteristics breakouts such as by specific office, facility, or other managerial units (so each manager and supervisor has outcome data specific to his or her area of responsibility), and by the type and/or amount of service provided (so you can identify which are leading to successful outcomes and which are not).

The resulting information can be used in many ways, particularly in identifying problem areas and allocating resources to improve the outcomes for customer groups whose outcomes lag behind other groups. See examples 14 (Coast Guard) and 4 (Rehabilitation Services Administration).

Regulatory agencies may similarly use breakout information to identify types of regulated entities or geographic areas with poorer outcomes in order to focus their efforts on these poorer performers. See examples 5 (EPA's Office of Air and Radiation) and 11 (Occupational Safety and Health).

A key element in effective use of outcome information is to continually ask why differences occur and then assess whether actions can be taken to improve results.

3. Request reasonably frequent outcome reports (at least quarterly, for example).
This permits program personnel to spot problems earlier, pursue special studies to identify the reasons for the shortfalls, and suggest improvements.

4. Seek data that are timely.
Try not to settle for reports on outcomes that are delayed unduly. Develop procedures to obtain outcome information more quickly after the period for which the data are being reported. Timely feedback is likely to be more useful to programs and gives the program personnel a greater sense of being able to affect outcomes. Advances in information technology (IT) should increasingly help programs speed up collection and analysis of outcome data, and improve the data's quality and timeliness.

5. Present outcome data in user-friendly forms.
For example, include bar charts or other graphic presentations in addition to, or in place of, data tables or spreadsheets. User-friendly presentations help managers identify good or poor performance, or spot variations in performance among different client groups or geographic regions. Such presentations make it more likely that managers or other staff will actually use the data.

6. Include data from previous time periods (such as the last year or two) in outcome information reports.
This enables managers to see trends in outcomes over time, which helps them identify where outcomes are declining and what action may be needed.

Provide Help to Staff to Encourage Greater Use

7. Develop guidance materials on the use of outcome information.
Guidance materials might provide agency- or program-specific examples of ways managers and staff can use outcome information routinely collected by that agency or program. Guidance materials could be developed in the form of manuals (see example 6—EPA's Office of Enforcement and Compliance Assurance), computer-based material, or memos from higher level administrators. Consider identifying categories of common management decisions and linking them to specific outcome information that can be used to guide managers in making those decisions.

8. Provide training in uses of outcome information for managers and other staff who can use outcome information.
This could be incorporated in training routinely provided for new managers and staff. Training for long-standing staff and managers could be offered in the form of workshops or seminars focused on outcome data use.

Adapt Widely Applicable Uses of Outcome Information

9. Hold regular "How are we doing?" sessions with staff soon after each outcome report becomes available.
Discuss what appears to be working well and not so well, and seek suggestions for improvements. In later sessions, identify whether the improvement actions led to the desired improvements in outcomes. See examples 4 (Rehabilitation Services Administration) and 10 (National Park Service).

10. Identify and reward offices, grantees, and facilities with good outcomes.
These might include offices or grantees that continue, over multiple reporting periods, to meet or exceed their outcome targets. Recognition of such achievements can serve as a reward, especially if other forms of reward are not possible.

11. Develop grant allocation guidelines that reward improved performance.
See example 8 (Public and Indian Housing).

12. Use the outcome data to help identify successful ("best") practices within the agency or on the part of grantees or regulated facilities.
Promulgate those successful practices to other offices, grantees, and facilities. See examples 8 (Public and Indian Housing) and 4 (Rehabilitation Services Administration).

13. Use outcome data to help identify common problems and, if possible, solutions.
See examples 15 (Internal Revenue Service) and 8 (Public and Indian Housing).

14. Include a question at the end of regularly used survey questionnaires that asks respondents to provide suggestions for improving the service.
This applies to programs that survey clients (whether these "clients" are other organizations through which services are delivered to the end clients, or are the end clients themselves). Use that information as another input into your program decisions.

15. Use outcome information to assess whether a new program or procedure is working well.
See example 9 (Bureau of Land Management). Similarly, use outcome information to determine the extent to which changes in practices and policies have led to improved outcomes. See example 13 (Supplemental Security Income).

16. Use outcome information to identify needs for training for staff or technical assistance for grantees.
See examples 8 (Public and Indian Housing), 10 (National Park Service), and 4 (Rehabilitation Services Administration).

17. Use outcome information to help prioritize use of resources.
This use can be appropriate to improve programs or offices whose outcomes are lower than others or to address "underserved" client groups. See examples 5 (EPA's Office of Air and Radiation), 8 (Public and Indian Housing), and 11 (Occupational Safety and Health).

Summary

Many, if not most, of the uses that program managers make of outcome information are not as "glamorous" as large-scale changes made by upper level government officials (whether departmental, White House, or congressional). The uses made by programs tend to be "lower level" and of smaller scale. *Nevertheless, program managers can take actions that are substantive and important to the clients and the outcomes of their programs.* Program managers can make a difference, as illustrated in the examples that follow.

Federal Program Case Studies

This section contains case studies of uses of regularly collected outcome information by federal agencies and programs. We include examples from 16 federal programs, drawn from 11 departments or agencies.

The case studies are presented alphabetically by federal department/agency. In some cases we have case studies from more than one program in a department or agency. Separate descriptions are provided for each program.

The programs included are listed in Table 1, which categorizes each program as to whether it:

- Directly provides services
- Indirectly provides services by working though other levels of government or private organizations that provide services
- Is regulatory

In a few cases, a program appears to fall into two of these categories.

Categorizing examples by these types can help managers get ideas from programs in the same category as their own. However, managers might often obtain ideas for their own programs from cases drawn from different categories.

For many of these programs, the outcome indicators are far from ideal. Many of the indicators used are "intermediate," not "end," outcomes. However, focusing on intermediate outcomes often presents good opportunities for programs to make improvements that are expected to lead to improved ends. With limited program manager experience to date in outcome measurement and in the use of outcome data, even small-scale use represents progress.

People disagree on what is and is not an outcome. In general, *outcomes* represent the effects of a program's activities that directly affect citizens or nonfederal "customers" in some way. *Intermediate outcomes,* although not the final results, are nevertheless effects that are expected to lead to end outcomes. As suggested in the preceding paragraph, for this report we have been fairly liberal in including program products as outcomes.

1. Department of Agriculture Animal and Plant Health Inspection Service, Fruit Fly Exclusion and Detection Program[2]

Program and Objectives

A key mission of the Animal and Plant Health Inspection Service (APHIS) is to protect America's animal and plant resources by safeguarding resources from exotic invasive pests and diseases. An APHIS strategic goal is to safeguard U.S. plant and animal resources against the introduction of foreign pests and diseases, while meeting international trade obligations. APHIS received a FY 2002 appropriation of $36.8 million for Fruit Fly Exclusion and Detection, for both domestic and international activities.

An APHIS objective is to control and eradicate fruit flies, primarily the Mediterranean fruit fly (Medfly) and Mexican fruit fly, in foreign countries where they may pose a serious threat to U.S. agriculture as well as in the United States, by conducting detection and prevention activities. APHIS operates a cooperative program with Mexico and Guatemala to meet this objective.

A key performance goal is to minimize fruit fly outbreaks in Mexico and in the eradication area in Guatemala. Program strategies are to spray with environmentally sound insecticide, to release sterile fruit flies to lower population growth, and to destroy fly-infested fruit. Mexican aircraft and APHIS contractor aircraft dispense sterile insects in the problem areas. The main strategies to maintain the Medfly barrier in Guatemala are quarantine posts at key transportation points and the release of sterile Medflies.

Types of Uses of Outcome Data

The APHIS fruit fly program has used regularly collected outcome information in:

- Allocating and reallocating resources
- Redirecting program activities
- Justifying requests for additional resources

Outcome Measurement Process

A primary outcome indicator for the program is the number of outbreaks of Medflies, used as an indicator of prevention and eradication effectiveness. The specific indicators include the number of Medfly outbreaks in Mexico and the number of Medfly outbreaks in the eradication area in Guatemala.

Program employees in Mexico and Guatemala maintain and collect data from over 27,000 Medfly traps to assess the status of current infestations and identify new Medfly detection sites. The traps are checked at least weekly, and advanced geographic information systems are used to map out weekly surveillance results, which are reported to program managers at Medfly offices in Mexico and Guatemala for analysis and response.

Use of Outcome Data to Allocate and Reallocate Resources

When outcome data indicate a problem, AHPIS mobilizes quickly, allocating field personnel, vehicles, supplies, and other resources to the problem area. The APHIS regional office reallocates resources, increases the number of field personnel, and increases trap density to delimit the outbreak—saturating the area with traps and personnel to determine how big the problem is. APHIS shifts products (sterile insects) from one country to another to combat the problem.

Use of Outcome Data to Redirect Program Activities

In FY 1999, when weekly detection reports showed a sudden outbreak of Medflies in Mexico on the Guatemala border (and Medfly outbreaks occurred in California and Florida), program personnel initiated emergency eradication activities. A group of technical experts from the United States, Mexico, and Guatemala was dispatched to study current programs, suggest long-range goals and strategies, and recommend specific technical interventions in Mexico and Guatemala. The review team identified causes for the outbreak and made recommendations that changed the trapping and spraying programs. Main eradication strategies included strict enforcement of quarantines, aerial bait applications to reduce the number of Medflies, and the release of millions of sterile Medflies.

Use of Outcome Data to Justify Requests for Additional Resources

When problems are sufficiently large, APHIS requests contingency funds from the Commodity Credit Corporation. These funds are controlled at the level of the secretary and may require OMB approval. In FY 2002, APHIS received $23.1 million in Commodity Credit Corporation emergency funds for Medfly operations in Mexico and Guatemala. Mexico and Guatemala put in resources equivalent to the emergency funds supplied by the United States. These additional funds have allowed APHIS to increase trap density per square mile 25-fold (allowing APHIS to go from detection trapping to delimiting trapping) and to increase production of sterile insects in Mexico and Guatemala.

Results

Medfly outbreaks in Mexico went from 25 in 1997 to 254 in 1998, but decreased to 180 in 1999 and to less than 100 in 2000.

2. Department of Education Adult Education and Literacy[3]

Program and Objectives

The Division of Adult Education and Literacy (DAEL) in the Office of Vocational and Adult Education of the Department of Education provides grants to state governments for adult basic and secondary education and English literacy instruction for persons ages 16 through 60 who do not have a secondary school diploma or its equivalent. The purpose is to increase literacy and completion of a secondary school diploma or its equivalent, such as a general equivalency diploma (GED).

Types of Uses of Outcome Data

In 2000 DAEL introduced a National Reporting System (NRS) that requires each state receiving program funding to report annually on a number of outcome indicators. So far, DAEL has used that outcome information to:

- Identify which states to visit for fact finding and to provide limited technical assistance.
- Determine which states will receive annual monetary incentive funds. The data are being used along with outcome indicator data provided by states for other federal programs covered by the Workforce Investment Act. (The other agencies include the Department of Labor's employment program and the Department of Education's vocational and technical education program funded by Perkins.) State performance on adult education outcomes, along with the outcomes of these other programs, helps determine the amount of reward, if any, that goes to each state each year.
- Make the Department of Education's budget projections for adult education and literacy (as well as to meet its performance reporting requirements under GPRA).

Outcome Measurement Process

Each state receiving funds is required to provide annual data on the following adult education outcome indicators:

- Number and percentage of clients showing improvement in literacy level, based on tests administered locally
- Number and percentage of clients going on to further education and training
- Number and percentage of clients who completed high school or an equivalent secondary credential (such as a GED)
- Number and percentage of clients who obtained employment
- Number and percentage of clients who retained a job or advanced on the job

DAEL has provided manuals to the states describing the outcome indicators and data collection procedures in detail. Training has also been provided to state and local adult education practitioners over the past three years. Numerous national conferences and meetings have been held with state-level personnel on implementing the NRS, and multiple train-the-trainer workshops have been held to provide the information and tools needed by states to train local program personnel. An online training program targeting local administrators and teachers on the NRS has been available for the past three years.

Last year four regional trainings on data quality were held, and a data quality guidebook was developed and made available to the states. To further support high-quality data reported through the NRS, the Division of Adult Education and Literacy recently published a set of data quality standards that identify the policies, processes, and materials that states and local programs should have in place to collect valid and reliable data for the NRS.

Specific Uses of the Outcome Data

FY 2001 was the first year of reporting by the states. During FY 2002, DAEL staff made a small number of site visits for "fact finding," to examine

implementation, and to provide a limited amount of technical assistance. DAEL used a number of criteria to select the states to visit. The state's performance on the outcome indicators was one of the principal criteria. It was key to the selection of about half the states.

In the states visited, DAEL staff attempted to assess the extent to which lower than expected performance was due to performance problems or to reporting artifacts.

DAEL plans in FY 2003 to provide technical assistance and training to states on a more targeted basis. Performance on the outcome indicators is currently expected to be one of the major determinants.

Note: Stimulated at least in part by the federal incentives to states, at least one state, Kentucky, has implemented its own in-state incentive system for its local county adult education programs based on the federal adult education outcomes. Its Department for Adult Education and Literacy has been providing annual incentive funding of 5 percent to 10 percent to any local county program that meets a specified percentage of its outcome targets for the year. The "bonus" that a local program receives depends on the percentage of targets met for the year. Thus, the federal effort has led to a parallel process being implemented by at least one state for its local programs.

3. Department of Education Migrant Education[4]

Program and Objectives

The Department of Education's Office of Migrant Education has used outcome information for both its High School Equivalency Program (HEP) and its College Assistance Migrant Program (CAMP). Both programs are discretionary grant programs. They each provide five-year, noncompeting continuation grants to institutions of higher education or private nonprofit agencies working in cooperation with those institutions.

HEP seeks to help low-income migrant and seasonal farm workers and their children gain high school diplomas or equivalence certificates (GEDs) and enter postsecondary education or other training programs, or be placed in career positions or in the military. Grantees provide the appropriate educational, academic, and personal counseling, placement, and support services.

CAMP seeks to help eligible migrant and seasonal farm workers successfully complete higher education. It provides financial and support services to migrant and seasonal farm workers and their children to help them successfully complete the first academic year of study in an institution of higher education. Services include academic counseling, tutoring, financial aid, and housing assistance.

Types of Uses of Outcome Data

The Office of Migrant Education uses outcome information to:

- Help determine funding for the coming year for each grantee
- Help determine whether it needs to provide technical assistance
- Stimulate grantees to achieve outcome targets for the program, including those that the grantees had set for themselves

Outcome Measurement Process

For each program, the Office of Migrant Education has established a set of outcome indicators on which each grantee is required to submit annual performance reports. It tracks each grantee's performance through the data collected in those reports. Grantees are required to provide these data in their annual performance reports before they can receive their noncompeting continuation grant awards for subsequent years.

The key outcome indicators for HEP are:

- Number of students served by the program
- Number of students who received their GED
- Number of GED recipients who were placed in institutions of higher education or other training programs
- Number of GED recipients who were placed in career positions or the military

The outcome indicators for CAMP are:

- Number of students served by the program
- Number who completed the first year of college in good standing
- Number who continued in postsecondary education
- Number who completed a college education (results for this outcome indicator cannot be obtained until the last year or two of the grant)

However, the local programs have had too few resources to undertake much follow-up of students after they have left the program, so the data on the third and fourth outcome indicators for both programs is incomplete. (The most complete data are obtained on the third CAMP outcome indicator, the number of students who continued for at least one year in postsecondary education.)

Use of Outcome Data to Help Determine Funding

Each applicant for grants is required to provide its estimates (that is, targets) for the first two outcome indicators listed here for each program. Applicants provide estimates for each of the five years of the grant for these two indicators. Targets are not required for the other, longer term outcome indicators. However, the grantees are required to report on these outcome indicators in their annual performance reports.

When the office reviews the data from the performance reports, it compares the results with the grantee's targets, in addition to reviewing the data on the other outcome indicators. Progress in these outcomes is a major consideration when determining funding for the coming year and allocating any funds remaining at the end of the year.

In some instances, the office has increased funding for projects that were successful in meeting all their objectives and requested additional funds. In FY 2002, HEP had approximately 60 grantees; CAMP had approximately 40. In program year 2001, the office received performance reports from 81 grantees. It increased the funding for about 25 percent of these after reviewing the annual reports. For example, projects that had exceeded their objectives and needed additional funding to serve more students were given increases in funding. In some instances, the office has reduced funding if the project was having difficulty meeting its objectives and did not show a need for additional funding. In program year 2001, the Office of Migrant Education decreased the funding for about 10 percent of these after reviewing the reports. For example, funding was reduced for grantees who had large amounts of carryover funds because they were not able to serve all the students that they projected serving.

Use of Outcome Data to Determine the Need for Technical Assistance

The office also uses the performance results to help identify the need of grantees for technical assistance. (It avoids reducing funding unless poor performance is flagrant and technical assistance is not expected to help sufficiently.) For example, HEP and CAMP do not have sufficient resources to do on-site monitoring visits to all projects. The outcome data are a major factor in choosing site visits.

Use of Outcome Data to Stimulate Grantees to Achieve Targets

The outcome data are one of the subjects for discussions undertaken by program officers in their on-site visits to grantees, such as the extent to which grantees have met their outcome targets in the past reporting period. This is intended to stimulate grantees to achieve their targets. Use of grantee progress in meeting targets as a factor in funding decisions (as discussed previously) is also expected to stimulate grantees to achieve their targets.

4. Department of Education Office of Special Education and Rehabilitative Services, Rehabilitation Services Administration[5]

Program and Objectives

Rehabilitation Services Administration (RSA) in the Office of Special Education and Rehabilitative Services of the Department of Education administers grant programs for vocational rehabilitation services for individuals with disabilities. This example includes two programs: State Grants and Projects With Industry (PWI). The State Grants program allocates funds to state vocational rehabilitation (VR) agencies by formula. State VR agencies provide a wide range of specialized training and employment services aimed at helping people with disabilities to obtain employment.

The PWI program allocates discretionary grant funds competitively to a variety of organizations, such as nonprofit organizations and labor organizations that have partnerships with business and industry. PWI grantees provide similar employment services as state VR agencies.

RSA monitors performance, including outcome information, for recipients of its funds. Monitoring and oversight functions are primarily conducted by RSA personnel in 10 regional offices.

Types of Uses of Outcome Data

Uses of outcome information by RSA provided here include:

- Triggering corrective action
- Identifying technical assistance needs
- De-funding poor performers
- Holding "How are we doing?" meetings to update progress on goals
- Motivating grantees
- Identifying high performers and successful practices to help poorer performers
- Revising policy

Outcome Measurement Process

RSA developed new requirements for annual reporting of performance indicators by state VR agencies effective FY 2000.[6] The new requirements specified particular indicators, standards, and actions for failure to meet standards. State VR agencies had previously reported a variety of data, including some outcomes that were incorporated into the new set of indicators, such as the number of clients who obtained employment after receiving their services.

State agencies are required to report their performance within 60 days after the end of each fiscal year. Six of the indicators are related to employment and can be considered intermediate or end outcome indicators.[7] To achieve successful performance, state agencies must meet or exceed the performance levels set for four of the six indicators, including two of the following three "primary" indicators, which address the quality of employment outcomes:

- The percentage of individuals who exit the program with employment earnings at or above the minimum wage (considered competitive employment)
- The proportion of individuals who achieve competitive employment who have significant disabilities
- The average hourly earnings of all individuals who exit the VR program in competitive employment with earning levels equivalent to at least the minimum wage as a ratio of the state's average hourly earnings for all employed individuals in the state

Other indicators include an "employment outcome" indicator, measured as the percentage of individuals who become employed after receiving services and who retain that job for 90 days. Another indicator assesses improvement in state agency performance over time. This is reported as the ratio of individuals who leave the program with employment outcomes during the current performance period and the previous performance period. Another outcome indicator is the change in the percentage of individuals who are primarily self-supporting from entering to exiting the program—that is, whose primary source of support is personal income rather than, for example, public support.

RSA established minimum performance levels for each indicator in addition to establishing criteria for passing the overall standard for successful performance (described previously). For example, in order to pass the indicator on competitive employment, a state VR agency must meet or exceed the minimum performance level of 72.6 percent. In other words, 72.6 percent of individuals placed in employment by the agency must be placed in competitive employment.

A similar system of performance indicators and standards was implemented for RSA's PWI program (described later in this example).

Use of Outcome Data to Trigger Corrective Action and to Identify Technical Assistance Needs

State VR agencies that do not pass the overall standard are required to develop a program improvement plan (PIP). In the PIP the agencies identify steps they will take to help them pass the indicators for which they had failing scores. If a state VR agency does not develop a PIP or does not substantially comply with the terms of its PIP, RSA has the ability to reduce payments to that agency, or to make no further payments to that agency, until it submits a PIP or raises its performance to meet the minimal satisfactory level.

Regional RSA staff who monitor state agencies also use the outcome information to identify potential problems or questionable practices even where PIPs are not required. For example, a state agency might have a failing or near-failing score on one or two indicators. In such cases, RSA representatives encourage the agency to focus on improving those outcomes and provide technical assistance (suggestions) to help it do so.

Regional RSA offices have program specialists who monitor state VR agencies in the state(s) to which they are assigned. These RSA specialists work closely with state VR agencies to develop PIPs using a variety of information to identify factors that contribute to the deficiencies of these agencies. The RSA specialists provide additional oversight and technical assistance to state VR agencies that are required to submit and carry out PIPs. Because this system has not been in effect for long, RSA has had few cases in which PIPs were required. The following example describes the initial steps taken in the Dallas regional office related to one state agency recently required to enter the PIP process. This process was still in the early stages at the time this report was written. The following describes the initial steps:

- The RSA representative conducted an initial site visit to collect and review additional data (such as caseload size and client characteristics) to identify factors contributing to the agency's low scores. The RSA representative held discussions with state agency managers to learn their perspective on factors affecting their outcomes.
- The RSA representative will prepare a report to the state agency outlining the results of the review and factors felt to be causing the problems. This report will provide recommendations for corrective action.
- The state agency will develop a PIP to correct the performance problems. It will provide data related to steps called for in the PIP to the regional office at least every six months. The RSA representative expects to monitor data more frequently, most likely quarterly.
- The RSA representative will provide technical assistance to help the state agency carry out steps identified in the PIP as needed.

The following examples illustrate uses of outcome information by RSA representatives in regional offices in cases where there were low or failing scores but PIPs were not required.

An RSA representative in Dallas reviewed outcome information for a state agency that had focused on improving its successful closures—where clients received employment. However, the agency apparently did not realize that its "unsuccessful" closures were increasing even more than its successful closures until the RSA representative pointed this out and suggested the agency address it. The state VR agency had its staff make more regular contact with clients, particularly those who had finished training but had not yet obtained employment. The prior low levels of contact with such clients apparently caused the agency to lose track of some clients, which resulted in their being counted among the

cases without employment outcomes. The increased contact led to improvement in this agency's outcomes by increasing the number of successful closures and decreasing unsuccessful ones.

A similar example was provided by an RSA representative in the Boston regional office. One state agency had increased the number of cases successfully closed over the years, but the number of unsuccessful closures also was increasing. The RSA representative advised them in 2001 that this could lead to a failing score on the indicator for percentage of clients who become employed. She suggested the agency assess why clients were not achieving employment to determine whether changes in practice were needed. The agency was unable to do this because it did not have staff it could free up. The agency did not achieve the standard for that indicator in 2001. At that point the RSA representative suggested the agency reassign one staff member on a half-time basis to work on identifying reasons for this problem so the agency could make changes to address it. The agency agreed with this suggestion and is currently investigating options for implementing it.

An RSA program specialist in the San Francisco regional office provided examples of two cases in which he used breakout data to identify potential problems in state VR agency practices and suggest changes to those agencies. "Breaking out" data refers to disaggregating outcome data to report outcomes for specific types of clients, geographic areas, service providers, and so forth. Developing breakouts in this case required using the original data reported by the agency, rather than the indicator data the national RSA office provides to regions (which is not broken out).

This program specialist reviewed outcome data for cases closed due to employment, broken out by ethnicity of clients, for a VR agency in the Pacific Islands. The breakout data indicated very few cases closed were for Caucasians. Similarly, breakout data on clients served indicated that the agency served almost no Caucasians, although approximately 25 percent of the population in the area it served was Caucasian.

The VR specialist raised this issue with officials of that agency, who apparently were unaware of this pattern. The agency subsequently hired a Caucasian staff member and provided staff training to ensure that its outreach and recruiting efforts addressed all ethnic groups. Based on discussions with personnel of that VR agency, the program specialist estimated that approximately 10 percent of clients served are now Caucasian. (Outcome data are not yet available for periods after these changes went into effect.)

This VR specialist also reviewed breakout data on the ages of clients with successful case closures. Data for one state indicated that a considerable number were over age 65 and were concentrated in the "homemaker" category of employment.[8] This suggested that older clients with disabilities who were using agency services were not seeking employment outside the home. Since employment is a key focus of RSA, the RSA specialist raised this issue with officials of that state agency and recommended they reconsider their priorities. That agency has redirected activities of its staff to focus on services for clients who are interested in obtaining employment outside the home.

Breaking outcome data out by relevant categories makes it possible to identify differences among outcomes achieved by, for example, different types of clients (or whatever breakout category is used). This makes outcome data considerably more useful to managers.

Use of Outcome Data to De-Fund Poor Performers

RSA's discretionary grant program, PWI, uses an indicator system similar to that used for state VR agencies. PWI generally provides grants for up to five years. Grantees are required to provide annual reports on their performance in five areas. The two primary indicators are (1) the percentage of clients served who achieve a competitive employment outcome and (2) the change in client earnings from program entry to exiting the program. Three secondary indicators include (1) the percentage of clients with employment outcomes who have significant disabilities, (2) the percentage placed who previously were unemployed, and (3) the average cost per placement. (The latter is not an outcome indicator.)

PWI grantees must have passing scores on both primary indicators and two of the three secondary indicators to remain eligible for continued funding. Grantees who do not achieve such scores have a six-month period in which to improve their performance. PWI project officers located in RSA's Washington, D.C., office oversee grantees and provide technical assistance through teleconferences during the six-month period to help them improve their outcomes. Regional office staff may similarly provide technical assistance to such PWI agencies in their regions. If the grantee does not improve at the end of those six months, the grant is terminated at the end of that fiscal year. Two or three PWI grantees are reportedly de-funded each year.

Use of Outcome Data in "How Are We Doing?" Meetings with Grantees and to Provide Technical Assistance

PWI project officers in the Washington, D.C., office use outcome information to conduct quarterly conference calls with individual PWI grantees to monitor performance. Regional RSA personnel who oversee PWI projects in their states participate in these calls when available. These are, in effect, "How are we doing?" meetings that focus on the grantees' progress toward meeting performance targets.

PWI grantees provide quarterly updates on various outcome information during these calls. PWI project officers review actual performance for the quarter against targets and discuss what the grantee has been doing to address any indicators with poor performance. They discuss any problems or barriers the grantee has been experiencing that may have affected performance and provide suggestions (technical assistance) to help grantees improve their performance.

Regional RSA representatives similarly review outcome indicators and hold "How are we doing?" meetings with VR agencies in states they oversee. These meetings may be held in person or over the telephone.

One RSA representative in the Seattle regional office uses outcome indicator information for annual meetings she conducts with officials of the two state VR agencies she supervises. She looks at trends in indicators by reviewing data from prior years as well as current year data, to enable her to discuss areas in which performance is deteriorating. She emphasizes any indicators for which the agency's outcomes are close to the "failing" level. The RSA representative seeks explanations for these problems and brainstorms with agency managers to identify steps they might take to improve their outcomes. Following are examples of changes made by two agencies as a result of her approach.

One state VR agency was not performing well in the indicator for the ratio of average wage of its clients who attained employment to the state average wage. Its performance on this ratio is affected by the high average wage in that state. The RSA representative suggested the agency place greater emphasis on providing career counseling and placing clients in "career" jobs rather than minimum wage jobs. Similarly, she indicated it would be appropriate to help place clients in training programs that would take longer but result in better paying jobs. That state agency provided training on the indicators to its staff to help them understand the outcomes they should be trying to achieve. It also changed wording in documents its staff use to develop client plans to incorporate language about career change. This encourages staff to talk about career development and career jobs with clients.

Another state agency monitored by this RSA representative had failing scores on the indicator for the percentage of cases successfully closed due to employment of the client. Because of this, the RSA representative pulled a sample of files of closed cases on a visit to this agency to seek explanations for unsuccessful closures.[9] She also sought the opinions of the state agency's managers. In this case it appeared that many of the unsuccessful closures were due to agency staff losing contact with clients. The RSA representative provided technical assistance for steps the agency could take to alleviate this problem. These include (1) obtaining contact information for relatives or other secondary contacts who would know how to locate the client if the agency has difficulty reaching him or her in the future; (2) periodically updating contact information on the client and secondary contacts; and (3) making greater effort to reach clients who have not responded to letters attempting to contact them (such as phone calls to the client or the secondary contacts).

Use of Outcome Data to Motivate Grantees

RSA's performance-monitoring system is intended to motivate state VR agencies and PWI grantees to achieve target outcomes. Doing so enables them to avoid the consequences associated with failing scores. As described earlier, for PWI grantees, consequences may include loss of funding. Consequences for state VR agencies include developing and following a PIP and additional oversight by the regional office. There also is the potential for reduced payments for failure to comply with the plan.

To avoid the temptation to focus on clients with less serious disabilities in order to achieve higher success rates, one of the outcome indicators focuses on outcomes for those clients with significant disabilities. This, in effect, motivates grantees to address this population.

Thus far, RSA is not providing rewards or incentives for high performers. However, the department is examining the possibility of incentive grants that might use performance on selected outcome indicators among the eligibility criteria.

Use of Outcome Data to Identify High Performers, Identify Successful Practices, and Help Poor Performers

In August 2002, RSA initiated an effort to identify factors related to success, or lack of it, in achieving desired program outcomes. The purpose of the effort is to improve the performance of state agencies that receive federal funding for VR.

The beginning point for this VR effort is the outcome data that RSA has been collecting from state VR agencies—particularly data on client success in obtaining employment and the level of earnings. RSA is starting to use these data to identify state agencies that have performed well and those whose outcomes have been poor. The highest and lowest performing agencies are being examined to identify the major factors contributing to the high or low performance level. It is expected that this will provide useful information as to what steps the federal, state, and local service agencies can take to improve the effectiveness of VR services. As of this writing, this effort is in progress, with completion expected within about one year.

Smaller scale efforts to identify high-performing agencies and factors that affect their performance are occurring at the regional level, independent of the national effort just described. For example, a VR specialist in the Seattle regional office reported using high scores on outcome indicators to identify agencies that are doing well, and holding discussions with managers in their "How are we doing?" meetings to develop a better understanding of why their scores are high.

For example, one agency that focused on serving blind clients had higher average wages for employed clients than most others. In discussing reasons for this, the specialist learned this agency made assistive technology available to its clients early in the program, rather than waiting until they got a job (which is the more common practice). This reportedly increased the clients' job options and enabled them to get better paying jobs.

RSA personnel in this regional office informally discuss indicator data with each other and share examples of practices used by high performers. Thus, when providing technical assistance to agencies each oversees, they can provide specific examples of what an agency in another state is doing in particular areas, such as client outreach or training.

In addition, managers of agencies with high performance scores, such as the agency serving blind clients who had higher than average wages, are asked to discuss their practices at biannual meetings of chiefs of state agencies held by this regional office. This, in effect, disseminates "good practices" information and enables managers of other state agencies to learn from their peers. It also provides recognition for high performers.

Use of Outcome Data to Revise Policy

RSA revised its policy related to its program for grants to states effective January 2001. The policy revision was based, in part, on outcome data routinely provided to RSA by grantees.

The policy revision was that state VR agencies would no longer be able to count clients placed in "sheltered" employment among their employment outcomes.[10] This is consistent with an earlier amendment to the Rehabilitation Act of 1973

emphasizing "integrated" (in effect, mainstream) employment rather than sheltered or segregated employment (in facilities that only employ persons with disabilities, usually at below minimum wage).

Data from state VR agencies indicated that fewer clients were being placed in sheltered employment settings in recent years, which was one factor leading to this policy revision. Less than 3 percent of clients included in the outcome indicator for competitive employment (that is, clients employed at minimum wage or above) came from sheltered employment.

Although this policy change is too recent for there to be data demonstrating improved competitive employment outcomes, State Grants program staff anticipate it will lead to increased numbers of clients on the competitive employment indicator and no clients in sheltered workshops.

5. Environmental Protection Agency Office of Air and Radiation[11]

Program and Objectives

Under the Clean Air Act, the Environmental Protection Agency (EPA) sets air quality standards and limits to protect public health and welfare. The latter includes damage to animals, crops, vegetation, and buildings. EPA's Office of Air and Radiation (OAR) develops national programs, policies, and regulations for controlling air pollution and radiation exposure. It routinely collects outcome data on ambient (outside) air quality.

Under authorization of the Clean Air Act Amendments of 1990, EPA set national air quality standards for six principal air pollutants (called criteria pollutants). EPA also established deadlines for communities to achieve those standards. Standards are established based on existing levels of those pollutants and on scientific evidence about pollutant levels that have adverse effects on health. EPA works in cooperation with state, tribal, and local governments to ensure these standards are met.

Types of Uses of Outcome Data

Examples of OAR use of outcome information described here are to:

- Develop recommendations for pollution control legislation
- Develop or revise standards, rules, or regulations
- Trigger corrective action
- Monitor improvement

Outcome Measurement Process

OAR measures concentration of six criteria pollutants (carbon monoxide, lead, nitrogen dioxide, ozone, particulate matter, and sulfur dioxide) at monitoring stations across the United States.[12] It tracks trends in air quality based on these measurements. These data are reported by EPA region, by monitoring location, and by county. Geographic reporting of data enables EPA managers to identify areas with poorer outcomes so attention can be focused there, as illustrated in some of the following examples.

Use of Outcome Data to Develop Recommendations for Legislation and to Develop or Revise Standards, Rules, and Regulations

OAR develops or modifies regulations and rules to address pollutant levels that exceed standards. One example is the passage of EPA-initiated legislation to phase out lead in gasoline, a primary source of lead emissions that contributed to areas across the United States exceeding lead pollution standards. Initial legislation went into effect in 1973, calling for gradual reduction in lead content to one-tenth of a gram per gallon by 1986. Starting in 1975, newly manufactured passenger cars and light trucks were required to have emission control systems that necessitated operation with unleaded fuel. The average air quality concentration for lead was 94 percent lower in 2001 than in 1982, largely related to the phase-out of leaded gasoline.

EPA's Regional Ozone Transport rule is a more recent example of use of air quality monitoring data broken out by geographic area. Air quality outcome data for the early 1990s showed that cities in the northeast had not met the standards for ozone. OAR determined that this was due, in part, to pollutants from other states that enter their boundaries. OAR used computer modeling to identify specific "upwind" areas that contributed to pollution in particular areas.

OAR developed the Regional Ozone Transport rule in 1995 to address ozone pollution that was transported from upwind locations. It established amounts by which 22 states and the District of Columbia must reduce their nitrogen oxide emissions to reduce their contribution to air quality problems. The rule specifies a "budget" or maximum amount of emission and target dates for achieving these levels. The rule also recommended that the most cost-effective way to achieve those standards would be to install "scrubbers" for the major source of these pollutants, coal-fired power plants.

No data are yet available showing improved outcomes resulting from this rule. Litigation initiated when the rule was introduced delayed its implementation. It has recently been upheld and is going into effect.

Use of Outcome Data to Trigger Corrective Action

Under the Clean Air Act Amendments of 1990, OAR established a system to trigger corrective action for areas that had not attained the pollution standards for one or more of the six principal air pollutants. This example is another illustration of use of outcome information reported by specific geographic areas to focus attention and corrective action on the areas whose outcomes need improvement. Nonattainment areas are primarily urban areas.[13] Areas are designated for nonattainment separately for each pollutant for which they exceed the standard. The most common pollutant for nonattainment designation is ozone, which is used for this example.

EPA rates the degree of each area's nonattainment on a scale ranging from "marginal" to "extreme," based on levels of the respective pollutant. Nonattainment areas may be moved to a higher rating category (for example, from "moderate" to "serious") if their pollution outcome data worsen sufficiently to place them in the next category on the scale. EPA stipulates a variety of requirements to help reduce levels of that pollutant to attain the standard. One example of such requirements is that the state require automobile emission inspections. The requirements increase at each worsened category of nonattainment (that is, higher level of pollution).

EPA requires that states prepare and implement a state implementation plan to achieve and maintain air quality standards in nonattainment areas. These plans are submitted to EPA regional offices for approval. States also are required to submit regularly collected outcome data on pollutant levels for which they have not attained standards. (Frequency of reporting varies by type of pollutant.) State and local authorities establish requirements for controlling and monitoring air pollution within the nonattainment areas, following the requirements set by EPA for each level of nonattainment.

Areas can be removed from nonattainment status when they reduce pollution to the standard level for three years. At that time, the state can apply for the area to be classified as "maintenance" status. The state must submit a maintenance plan for the area, identifying steps that will be taken to maintain pollution levels at or below the standard.

Approximately 400 counties were designated as nonattainment areas in 1990 for exceeding the one-hour ozone standard (which measures peak concentrations of ozone). They were given until 1999 to improve their outcome for that pollutant to the standard. If not, they would be moved to the next higher level of nonattainment.[14] As of 2002, 157 counties originally designated as nonattainment areas for ozone had been moved to maintenance status. This represents 38 percent of the approximately 400 counties designated as nonattainment areas.

6. Environmental Protection Agency Office of Enforcement and Compliance Assurance[15]

Program and Objectives

EPA's Office of Enforcement and Compliance Assurance (OECA) works with EPA regional offices, state environmental agencies, and tribal agencies to monitor and enforce compliance with environmental laws on the part of regulated entities. OECA uses a variety of mechanisms, including conducting inspections, referring cases for civil or criminal enforcement, issuing penalty orders, and assessing penalties (fines). OECA also provides various forms of compliance assistance (technical assistance) to help regulated entities and others reduce their pollutants and improve their compliance with regulations. The long-term objective is to improve the environment and, thus, human health.

Types of Uses of Outcome Data

This example illustrates use of regularly collected outcome information by OECA to:

- Focus staff and resources
- Set priorities
- Perform program-specific performance analysis
- Hold "How are we doing?" meetings with agency personnel

The program uses both aggregated outcome data (for regions and industries) and pollutant data related to individual facilities.

Outcome Measurement Process

In 1997, OECA initiated the National Performance Measures Strategy to improve the information it obtained to guide its work. It specifically sought to include outcome indicators and performance indicators among its regularly collected data. Prior to this, output indicators, such as inspections conducted, civil or criminal actions initiated, and penalties assessed, had been OECA's primary form of regularly collected data.

OECA used stakeholder input to develop three categories of performance measures. The outcome category of measures focuses on changes in the behavior of regulated populations. These include the following:

- Number of pounds of pollutants reduced by regulated entities. EPA provided training on how to calculate reductions in pollutants to its staff.[16]
- Rates of noncompliance (reported separately for particular types of regulated sectors).
- Responses of significant violators, including (1) the average number of days for significant violators to return to compliance or enter enforceable plans or agreements and (2) the percentage of significant violators with new or recurrent significant violations within two years of previous enforcement action.

Use of Outcome Data to Focus Staff and Resources

A key reason for including outcome information is to focus attention and resources on outcomes rather than activities. For example, in the past, personnel might have focused efforts on cases they believed could be solved quickly, which would increase the number of cases resolved. Including outcome indicators was intended to encourage staff to focus on cases that will have the greatest impact in terms of pollutants reduced rather than on number of cases resolved.

Use of Outcome Data to Set Priorities

OECA biannually identifies a number of national enforcement and compliance program priorities. This includes identifying priority industry sectors on which to focus its efforts. Data used to help select these industries[17] include such outcome information as the quantity of hazardous waste generated each year in an industry and patterns of noncompliance (such as the proportion of inspections in which actionable violations were found for a particular industry). Potential environmental and human health risks associated with that industry's emissions are also considered in setting priorities.

For FY 1999 and FY 2000, OECA identified 11 priority industry sectors, including concentrated animal-feeding operations, automotive service and repair shops, petroleum-refining facilities, coal-fired power plants, and dry-cleaning establishments. Ten

of the 11 industries were retained for the subsequent two-year cycle, 2001–2002.

Results

One example of improved outcomes in a priority industry is the settlements EPA reached in 2001 with four major petroleum refiners covering 27 refineries. As part of the settlement, the refiners agreed to add air pollution controls and operational changes that will result in estimated annual reductions of 87,000 tons of sulfur dioxide and 49,500 tons of nitrogen dioxide, among reductions in other pollutants.[18]

OECA also uses outcome information as one factor in setting priorities for its staff in EPA regional offices. Due to resource limitations, regional offices may not have sufficient staff or funds to conduct inspections in all facilities or initiate enforcement actions in all cases where that is applicable. OECA enters into a memorandum of agreement (MOA) with regional offices specifying the types of activities regions will perform. Regional offices agree to put additional effort (resources) into industries on the priority list. For example, regions may increase the number of inspections in these industries, and prioritize taking enforcement actions (when applicable) in these industries, over industries that are not on the priority list.

Regional offices negotiate with OECA on the balance of resources to be used toward nationally identified priorities and regional priorities in the MOA. For example, in some cases there may be few national priority industry facilities located in a particular region. In such cases, pursuing regional priorities is expected to lead to greater improvements in outcomes.

For example, the Region 7 office in Kansas City reported it plans to use outcome indicators in setting regional priorities for FY 2004. It expects to identify industries for which inspections and enforcement would lead to the largest reductions in the amount of pollutants emitted. In addition, it plans to target enforcement action in geographic areas with higher levels of health problems affected by pollution, such as asthma. It will overlay maps depicting such health data with maps of facilities in priority industries to prioritize particular facilities for inspection and enforcement.

Use of Outcome Data to Perform Program-Specific Performance Analysis

OECA has developed a guidance manual for using performance measurement data as a management tool.[19] The manual identifies key performance-based questions and identifies relevant measurement data needed to support each question.

The guidance manual has been used as a template for a pilot Performance Analysis for a segment of EPA's Water Program. In this pilot performance analysis, staff used outcome information on data for facilities that were not in compliance with limits on discharges of pollutants. These outcome data were broken out by major type of facility: municipal, industrial, or federal. Breakout data showed that federal facilities had higher noncompliance with standards than other types of facilities. Because EPA does not have authority to impose penalties on federal facilities, staff conducting the pilot test came up with two recommendations for addressing this. First, they recommended staff seek additional explanatory information on the root causes of the noncompliance. Second, they recommended staff identify alternative tools, such as compliance assistance (technical assistance), to improve compliance at federal facilities. This is still in the recommendation stage, so no data on the results of this assistance are available.

Use of Outcome Data in "How Are We Doing?" Meetings

Since implementing the National Performance Measures Strategy, OECA administrators have conducted several "How are we doing?" meetings with administrators of its regional offices to review their performance.

OECA analytic staff compile data broken out by region twice a year and provide the breakouts to OECA senior managers. Senior managers use these data when conducting site visits to respective regions to enable them to point out areas where improvement is needed and discuss ways to address this with regional staff.

7. Department of Health and Human Services Food and Drug Administration Human Drugs Program, Generic Drugs[20]

Program and Objectives

The Food and Drug Administration's (FDA) Human Drugs Program is responsible for ensuring that drug products used for prevention, diagnosis, and treatment of disease are safe and effective. A key aspect of its work is to review and approve (or reject) new prescription and generic drugs. This example addresses one program of the FDA, the Office of Generic Drugs (OGD), and focuses only on timeliness of reviews. Timely review and approval enable quicker citizen access to new drug products and an increased number of therapeutic options for medical professionals. Timely review and approval of generic drugs make available less expensive treatment options than prescription drugs.

Types of Uses of Outcome Data

This example illustrates use of regularly collected outcome information by OGD to:

- Restructure staff assignments
- Revise procedures
- Motivate and focus employees

Outcome Measurement Process

FDA collects outcome information on the time it takes to review new drug applications (called review time) and sets performance targets for this time. FDA considers this an output measure, although we feel it is appropriate to consider it an intermediate outcome measure because of its importance to citizens whose health care may benefit from use of new drug products.

FDA has set standards for review time for prescription drugs and generic drugs, which fall under separate offices and have different standards. For FY 1999, for example, OGD's performance goal was to review 55 percent of generic drug applications within six months (180 days) of receipt of application. This goal increased to 75 percent in 2002.

The director of OGD receives monthly reports on data related to GPRA goals, including the percentage of generic drug applications reviewed within 180 days of receipt.

Use of Outcome Data to Restructure Staff Assignments

OGD restructured staff assignments for conducting drug reviews to enable it to complete reviews in a more timely fashion. The restructuring involved creating an additional team and team leader in each of its two review divisions, resulting in a total of 10 teams in two divisions. OGD was able to create the additional teams in part because approximately 25 staff were added to OGD during 2000–2001. (Congress provided FDA with additional funding for staff because of concerns about review and approval time for generic drugs. In 1999, OGD was completing review on only 35 percent of applications within 180 days.)

The restructuring created positions for additional project managers. Project managers monitor work progress and bring problems to the attention of team leaders and division directors. Team leaders coordinate the work flow of their teams, make assignments for review of drugs within teams, and serve as reviewers. Thus, the new positions for project managers and team leaders contribute to timely review of applications.

Use of Outcome Data to Revise Procedures

During the fall of 2002, the monthly reports indicated that the percentage of applications reviewed within 180 days was decreasing slightly, from over 80 percent to a percentage in the high 70s. To address this decrease, the OGD director instituted a procedural change affecting the assignment of generic drug applications for review.

Generic drug applications are assigned to one of 10 teams based on the type of drug involved. The application stays with that team until the review is completed. At times, a particular team may develop a backlog, which may result in the office failing to reach its target. The procedure was revised to enable an application to be reassigned to another team if it appears the original team will not complete the review within 180 days.

Under the revised procedure, project managers are responsible for monitoring the status of applications and identifying applications whose reviews had not yet begun 100 days after assignment to a team. Project managers bring these to the attention of the team leader and division director. The division director reassigns these applications to another team.

Reassigning the application to another team at the 100-day point leaves the new team 80 days in which to conduct the review, which is generally sufficient time for review. The older application is reassigned to a team that has a smaller queue than the team to which it was originally assigned. The new team is expected to place higher priority on the reassigned application than on previously received applications in its own queue.

The director also emphasizes to project managers and team leaders that they need to keep focused on the 180-day standard.

Use of Outcome Data to Motivate and Focus Employees

The OGD director identifies any slippage in the review-time data in his or her weekly or biweekly meetings with division directors. This serves to focus them on the importance of the outcome data. The group also discusses ways to address any slippage. (This is a form of the "How are we doing?" meeting.) Division directors are responsible for looking at the status of applications for each team under them, to identify any that are overdue, and to seek reasons for that.

When the OGD director saw the percentage of applications reviewed within 180 days decreased slightly during 2002, he held "retreats" with staff of each division. The focus of these meetings was to ensure that all staff were aware of the goal, to emphasize the importance of meeting the goal, and to further motivate staff to achieve it. The slippage in number of reviews conducted within 180 days did not continue, and OGD returned to having over 80 percent of its reviews completed within that time period.

Results

Due in part to the changes identified in this example, OGD's percentage of applications reviewed within 180 days has increased from 35 percent in early 2000 to 87 percent in mid-2002.

8. Department of Housing and Urban Development Public and Indian Housing[21]

Purpose and Objectives

The Department of Housing and Urban Development (HUD) seeks to promote adequate and affordable housing, economic opportunity, and suitable living environments free from discrimination. Office of Public and Indian Housing (PIH) programs serve low-income families and individuals who live in public housing, Native American housing, and Section 8-assisted housing. The examples here focus on public housing. PIH provides operating subsidies and capital improvement funds to public housing authorities so they can provide housing for low-income families. The examples provided here are related to PIH objectives to monitor and improve physical and related conditions in public housing.

Types of Uses of Outcome Data

Examples of PIH use of outcome information provided here are to:

- Trigger corrective action
- Identify technical assistance needs
- Prioritize use of resources
- Motivate grantees
- Recognize and reward high performers
- Identify high performers to help poor performers
- Identify common problems and their solutions

Outcome Measurement Process

HUD established the Real Estate Assessment Center (REAC) in approximately 1998. REAC's primary function is to annually collect, validate, and assess information on public housing through the Public Housing Assessment System (PHAS). This information is provided to program offices and field offices to enable them to identify risks and direct resources to improve the quality of public housing.

PIH uses this assessment system to collect and score data on four aspects of public housing: (1) physical condition, (2) financial condition, (3) management operations, and (4) resident satisfaction. The scores for physical condition of property and resident satisfaction provide data for outcome indicators. The other two components are important indicators of an agency's ability to operate public housing but are not themselves outcome indicators. However, vacancy rates and the length of time units remain vacant, which also are considered outcome indicators in this report, are included under management operations.

Each component of a particular Public Housing Authority (PHA) property is scored separately. Then the scores for each property are combined into an overall score for that PHA. Resident satisfaction counts for a maximum of 10 points; the other three components count for a maximum of 30 points each. Total maximum score is 100 points. The initial PHAS scores were derived for the quarter ending March 31, 1999.[22] Scores were considered advisory through June 30, 2001, while a variety of issues raised about the scoring system were addressed. Data collection for physical condition and resident satisfaction are described below.

Physical Condition: A "trained observer" rating system is used annually to obtain scores for this component.[23] REAC trains inspectors (who are contractors) to rate physical conditions in five areas: site, building exterior, building systems (such as heating, electricity, and water), common areas, and dwelling units (such as apartments). The Uniform Physical Characteristics System specifies eight to 17 items to be inspected in each of the five areas.

Inspectors use hand-held computers both to read definitions of deficiencies and to record deficiencies they find. Inspectors rate the severity of each deficiency, using a three-level scale.

The physical condition score for a particular PHA property is based on a point system intended to reflect the relative importance of each area and item inspected. Points for deficiencies—adjusted for severity and the importance of the deficiency for the item inspected—are subtracted from the maximum score achievable. The overall numerical score is calculated by subtracting the sum of deductions for health and safety (H&S) deficiencies from the sum of the individual area points. A code letter is added to the score to identify the kinds of H&S deficiencies observed.

Resident Satisfaction: PHAS annually surveys residents to obtain data for this subsystem. This survey addresses overall satisfaction, maintenance and repair, communication, safety, services, and appearance of the development.

Questions address such specific issues as the number of problems residents had with particular items, such as plumbing, and how long it took for repairs to be made. Residents are asked to rate how safe they feel in specific areas of the project and to identify items, such as broken locks, that contribute to crime in their development. Residents are asked to rate their satisfaction with the upkeep of specific areas, such as common areas and building exteriors, and to identify the frequency of problems, such as graffiti and abandoned cars.

HUD designates an organization as the survey administrator. This organization uses a computerized process to select a random sample of tenants of each public housing authority to receive the survey. The administrator mails the questionnaires, using a multi-stage process. This includes an initial letter advising the resident of the survey; the questionnaire itself; a reminder postcard; and a second questionnaire sent to nonrespondents. Residents return completed surveys to the survey administrator for tabulation. HUD keeps individual responses confidential. Only tabulated survey data are provided to the PHA.

Five of the 10 points for the resident survey subsystem are based on the PHA's help in implementation of the survey, including publicizing it to residents, and development of a follow-up plan to address low scores, when applicable (discussed later in this example).

After scores are computed for each subsystem, PHAs are classified into one of three categories:

- High performers have overall scores of 90 percent or greater and at least 60 percent of the points available under each subsystem.
- Standard performers have overall scores of not less than 60 percent overall and not less than 60 percent in the physical, financial, and management subsystems.
- Troubled performers have a score of less than 60 percent overall, or less than 60 percent in the physical, financial, or management subsystems.

Use of Outcome Data to Trigger Corrective Action and to Identify Technical Assistance Needs

PIH addresses corrective action for scores on the resident satisfaction separately from scores on physical condition or overall scores. Corrective action related to life-threatening deficiencies receives different treatment. This section describes each of these corrective action situations separately, with particular emphasis on the approach HUD developed to assist PHAs in the troubled performer category.

Improvement Plans for "Standard" Performers

Local PHAs in the "standard" performer category (that have overall scores of at least 60 percent but less than 70 percent) are required to take corrective action. These PHAs are required to submit an improvement plan to their HUD field office, which oversees their corrective actions. The plans identify specific steps the PHA will take. The plans include quarterly targets for those steps. These PHAs are required to submit quarterly progress reports to their field office.

Field office staff provide technical assistance—suggestions and guidance—to help PHAs achieve the intended improvements. For example, field staff may provide suggestions regarding appropriate timing to replace items such as boilers, kitchen appliances, or bath fixtures. Field office staff may provide technical assistance through telephone calls or letters or during site visits, depending on the proximity of the PHA to the field office.

Field offices have the option of employing a procedure similar to the improvement plan for PHAs whose scores are not low enough to require that step. Field offices may require PHAs with scores between 18 and 21 (out of 30) in the physical, financial, or management components to develop a corrective action plan (CAP). These are similar to improvement plans in that the PHA identifies specific steps it will take to improve performance. The PHA then submits quarterly progress reports related

to its CAP to the field office. The field office provides technical assistance as needed to help PHA management improve its score for the next year.

The New Jersey field office provided an example of using the CAP process to improve PHA outcomes. The Newark Housing Authority participated in the CAP process in New Jersey because it had a score of 17 (out of 30) in the physical assessment component in its 2001 inspection. Its CAP identified specific deficiencies in buildings and housing units that would be corrected. The PHA corrected those deficiencies and achieved a score of 24 in that component in 2002, after which it was released from CAP reporting.

Help for Troubled Performers

HUD established two Troubled Agency Recovery Centers (TARCs) to help PHAs whose scores place them in the troubled performer category. (This status is based on overall scores or scores in one or more of the physical condition, financial, or management subsections.) TARCs in Cleveland and Memphis provide oversight and technical assistance to help troubled performers attain standard scores within a two-year period. If a troubled agency fails to reach the goals established with TARC within the two-year period, it is referred to the HUD Enforcement Center. The latter may impose sanctions to ensure compliance or place the PHA in receivership.

TARCs assign a team of at least three specialists to each troubled agency. The team includes financial analysts, engineering specialists, and housing revitalization specialists. The team conducts an initial site visit and assessment that generally takes three to four working days. After this, the TARC works with the PHA to develop a MOA and CAP, which identify specific PHA actions and milestones.

The MOA also identifies technical assistance and training needs. Training and technical assistance may be provided by TARC specialists or HUD contractors, or obtained locally using HUD grant funding. TARC team members provide a considerable amount of technical assistance informally over the telephone, particularly during the initial months of supervision for each PHA. One TARC supervisor estimated that a team typically provides eight to 10 hours per week of telephone technical assistance to a PHA. One or more TARC team members may be on-site once per month, or once every other month, after the initial site visit. Following are some examples of technical assistance provided by the Memphis and Cleveland TARCs to housing authorities under their supervision.

The *Alexandria, Louisiana,* public housing authority was referred to the Memphis TARC in 1999 with a very low score on the physical condition component and vacancy rate problems. It had an overall PHAS score of 47 (out of 100) and a score of 7 (out of 30) in the physical condition component. The Memphis TARC provided technical assistance and training over approximately two years, including:

- Working with PHA management to develop a maintenance plan and identify and address "bugs" in their existing work-order system.
- Sending the PHA maintenance supervisor for training in operating a maintenance order system. (A maintenance order system includes establishing and maintaining a work-order log, developing and following systems for inspection, assigning work, and tracking work through completion.)
- Providing informal training to the maintenance workers on working within a maintenance order system.
- Assisting in developing a job description for maintenance staff and for other positions in the PHA.
- Helping the executive director evaluate applicants for the maintenance supervisor position when the original supervisor left.
- Pointing out relatively minor improvements, such as improving the cleanliness of the facility and repairing sidewalks, that could be readily addressed and that would make the facility more appealing and increase their score.

The TARC team recommended the PHA place approximately 150 units that were in substandard condition in a HUD-sponsored modernization program.

TARC had the PHA develop a long-term vacancy reduction plan and enter into an agreement with a contractor to manage getting these units modernized and ready for occupancy or demolished. TARC helped this PHA address an excessive number of units with short-term vacancies. These are vacancies in units that need only such work as minor repairs, painting, and cleaning to make them available. TARC required PHA management to prepare a plan specifying the number of units that would be readied and leased each week, and to submit a weekly electronic report to TARC reporting their progress against that plan. The number of short-term vacant units was reduced from approximately 150 in 1999 to 30 units in 2000.

Results: After TARC assistance, this PHA obtained a physical condition score of 25 out of 30 and an overall score of 80, placing it in the standard performer category in 2000.

The *Merryville, Louisiana,* PHA was referred to the Memphis TARC in 1999 with an overall score of 58 (out of 100) and a physical condition score of 14 (out of 30). As in the prior example, this PHA had a large number of vacancies and lacked a properly functioning maintenance work-order system. The Memphis TARC provided technical assistance to this PHA over a two-year period.

TARC team members provided technical assistance to enable the PHA to understand and use the computerized work-order system it had. TARC also provided technical assistance related to maintaining work-order documents to provide work crews with the information needed to perform their tasks. TARC also provided "common sense" recommendations, for example, to hire temporary contractors to prepare vacant units for rental when the number of units needing such work exceeds what PHA staff can handle.

Results: After being under TARC supervision, this PHA reduced the number of vacant units from 38 (out of 90 units) to zero, making more units available for families in need of housing. The PHA's overall score increased to 75 in 2000, placing it in the standard performer category, and its physical condition score increased to 23 out of 30.

The *Waynoka, Oklahoma,* PHA entered the TARC system in 1999 with an overall score of only 46, largely related to its high vacancy rate. The TARC team provided technical assistance focused on developing and using a work-order system similar to the examples previously described. TARC also contracted with a private organization to provide training for the PHA's new executive director (hired when the PHA was assigned to the TARC). TARC team members helped the new executive director develop a contract for maintenance services with an outside provider when she decided to contract out for this service.

Results: This PHA's overall score increased to 93 in 2000, placing it in the high performer category. Its vacancy rate fell from 41 percent to zero.

The *Wellston, Missouri,* PHA came under the Cleveland TARC in June 1998. Its overall score was 37, and its physical condition score was 11. Approximately one-third of its 210 units were vacant, and some had been vacant for long periods and were boarded up. The Cleveland TARC provided technical assistance and training to help the PHA address maintenance and vacancy issues. Technical assistance included:

- Inspecting units and writing work orders to identify the type of work needed (such as preparing a unit for rental, or emergency or non-emergency repairs for current tenants). This enables maintenance staff and management to determine whether particular jobs can be performed by maintenance staff or whether contractors are needed for larger or more extensive jobs. The executive director and maintenance director were sent to a two-day training program on how to conduct inspections.
- Installing a computer system to handle work orders, maintain waiting lists for units, track rent payments, maintain tenant information, and so forth. The PHA did not have a computer system in place prior to TARC assistance.
- Developing a tracking system so management would know the status of various work orders and establishing maximum times for completion of different categories of work.

- Developing a resident organization, which had not existed at this PHA. Elections for officers were held, and this group has been meeting regularly. The resident organization worked with local police to initiate a neighborhood crime watch group for the housing complex.

Results: After receiving technical assistance and training through TARC and adopting the maintenance work-order system, the overall PHAS score for this PHA as of December 31, 2001, (the most recent available) increased to 89—one point below the high performer category. Its physical condition score was 25 out of 30, and its resident satisfaction score was 9 out of 10. It had no vacancies and now has a waiting list of potential tenants.

Life-Threatening Problems

When the physical inspection identifies deficiencies that pose serious health or safety hazards (referred to as exigent health and fire safety hazards), the numerical score is reduced by the sum of deductions for health and safety deficiencies. A code letter is added to the score to identify the kinds of health and safety deficiencies observed. The REAC inspector leaves a written form with the PHA identifying the problem. The PHA is required to correct the problem within 24 hours. It must provide written certification of the correction to its field office within 72 hours. The next time HUD field office staff or quality assurance staff are on-site at that PHA, they check that work was completed.

Resident Satisfaction Survey Scores

HUD requires PHAs to develop follow-up plans for corrective action for all sections of the resident satisfaction questionnaire whose scores fall below 75 percent of the possible points for that section. The survey sections are maintenance and repair, communication, safety, services, and neighborhood appearance.

PHAs submit their follow-up plans electronically to HUD. The follow-up plan identifies the actions they will take in the low-scoring components, the target dates for completion, and the funding source (if required) they will use for the action. If a PHA does not develop and certify a follow-up plan, it loses some of the points for this component.

HUD has developed written guidelines for addressing issues under the various survey components. These guidelines are available to PHAs on HUD's REAC website. The guidelines are a form of technical assistance for addressing low satisfaction scores. For example, recommendations to address low scores in maintenance and repair include developing a system of accounts for work orders, training maintenance staff, partnering with a high-scoring PHA, and visiting the local HUD office to discuss methods for improvement. Communicating with residents is a recommended action under most sections.

Use of Outcome Data to Prioritize Use of Resources

Staff in PIH's New Jersey field office reported using the outcome information to prioritize resource use in two ways. The first example involves prioritizing use of field office resources. The second involves field office staff having PHAs prioritize use of their HUD-provided capital improvement resources.

New Jersey field office staff report using PHAS scores to prioritize PHAs for site visits, thus using limited travel and staff resources to oversee the PHAs in greatest need of supervision and assistance. HUD expects field offices to use PHAS scores, in combination with other information, such as the level of funding and compliance and audit findings, to conduct "risk assessments" to select the PHAs that will receive site visits each year.

Similarly, HUD's Kansas City "hub" office, which coordinates field offices in four states, reports using PHAS scores as one factor in assessing risk of the PHAs in the four states. This information is used to allocate resources among the four states for staff travel to conduct compliance-monitoring visits, technical assistance, and training.

New Jersey field office staff uses the scores for different components of the PHAS to identify the aspects of the PHA on which the site visit should focus, such as particular buildings or aspects of the physical condition component. This enables them to concentrate their efforts on components in need of improvement.

The New Jersey field office asks PHAs to use their scores on the physical condition component to prioritize spending of their capital improvement funds. Each PHA receives formula grant funds from HUD for capital improvements. PHAs are required to submit one-year and five-year plans for use of these funds to their field office. The New Jersey field office requires housing authorities whose physical condition scores place them in the failing category (scores lower than 18) or substandard categories (between 18 and 21) to identify on their plans how their use of capital improvement funds will increase their scores.

Results

The field office director believes that having PHAs focus capital improvement funding on areas with low scores has contributed to the overall improvement in physical condition scores in New Jersey since the inception of PHAS. During the first four quarters when PHAS scores were calculated (quarters ending September 1998 through June 1999[24]) 45 housing authorities in New Jersey failed the physical component. During the same quarters ending in 2000 and 2001, only four PHAs in New Jersey failed the physical component. (Scores were not yet available for those quarters through 2002.) During the first four quarters, two New Jersey PHAs were designated as troubled, and only five were designated as high performers. In the most recent four quarters, no PHAs received scores placing them in the troubled performer category, and 46 were designated as high performers.

Use of Outcome Data to Motivate Grantees and to Recognize and Reward High Performers

The PHAS scoring system is intended to motivate PHA managers to address problems *before* they appear as deficiencies on their assessments to avoid requirements for corrective action. The prospect of being assigned to a TARC may be a particular incentive, since this status involves considerably more supervision by HUD and reporting to HUD.

Because PHAS scores are available online, PHAs know that their field office can readily monitor their performance by computer. New Jersey field office officials report this serves as an incentive to improve scores.

HUD also provides incentives and recognition to high performers to encourage PHAs to improve their scores. High performers are defined as having overall scores of 90 percent or greater and at least 60 percent of the points available under each of the four PHAS components.

The primary incentive for high performers is receiving a 3 percent capital funds "bonus" over the amount they otherwise would have received from HUD for capital funds. PHAs that receive a score of 24 or more (out of 30) on the physical condition component also are subject to less frequent physical inspections. They are inspected every other year rather than annually. This enables HUD to focus its inspection resources where there is greater need.

High performers may receive public recognition of various kinds from their field offices, such as presenting plaques or other recognition at meetings of PHAs in their state. Some field offices may recognize high performers on their websites or in newsletters.

The Kentucky PIH office recognizes high performers with certificates at a recognition ceremony during its annual statewide meeting of public housing authorities. This office provides recognition to a "public housing authority of the year" and a "small public housing authority of the year." Authorities must be high performers on the PHAS to be eligible for this recognition, which includes presentation of a plaque during one of the conferences and press releases for local media.

The New Jersey field office has publicized both high- and low-performing PHAs in letters sent to all PHAs. This is intended to recognize high performers and motivate low performers to improve their scores. Field office management feels this has contributed to improvement in scores in their state.

Results

The number of both PHA properties and units falling below the defined substandard threshold of 60 decreased nationally when comparing scores for FYEs 9/30/99 through 6/30/00 with scores for FYEs 9/30/00 through 6/30/01. The percentage of properties with scores below 60 has decreased from 16.8 percent to 9.3 percent, and the percentage of units

in properties with scores below 60 has decreased from 26.7 percent to 16.4 percent. The average physical score has increased from 75 to 81. There also has been a 30 percent reduction in frequency of the most prevalent deficiencies (including problems with paint, stoves, hardware/locks, and sinks) during that period.

Use of Outcome Data to Identify High Performers and to Help Poor Performers

This use of outcome information by the New Jersey PIH field office is similar to identifying "best practices" so others can adopt them. Approximately two years ago, this field office began arranging "partnerships" between housing authorities that had low scores and were believed to have problems with management capability and a nearby PHA with standard or high scores perceived to have good management. Field office staff helps develop a formal agreement similar to a consulting agreement between the paired authorities. The high-performing agency agrees to help perform management functions, while the low-performing PHA maintains responsibility for PHA operations. The intent is to improve the management capacity of the low-performing PHA.

This practice was first implemented with the Cape May PHA and the nearby Vineland PHA approximately two years ago. The field office worked with the two authorities to forge a formal agreement in which Vineland PHA agreed to manage the Cape May PHA for a one-year period (which was extended for a second year). In 1998, the Cape May PHA had an overall PHAS score of 83 and a physical condition score of 17 (passing score is 18). In 2002, that PHA had an overall score of 96 and a physical component score of 28, placing it in the high performer category.

Since this initial partnering effort, the New Jersey field office has brought about similar arrangements with a few other pairs of housing authorities. These arrangements are too recent to yield data on improved scores.

Use of Outcome Data to Identify Common Problems and Their Solutions

The Kentucky PIH office recognized that a frequently identified fire safety hazard during PHAS inspections in its state was window air conditioners, which prevent the use of windows as emergency exits. The Covington, Kentucky, PHA was one agency that received such a violation. That PHA devised an inexpensive way to address this problem by installing a bracket that residents could unlatch to push the air conditioner out of the way in an emergency.

Because the same hazard has been cited in other PHAs and because window air conditioners are commonly used in PHAs, Kentucky PIH personnel made efforts to generate awareness of this solution. The Kentucky PIH construction and maintenance analyst has referred other PHAs with similar violations to the Covington PHA's maintenance staff to learn how to install and use the bracket. In addition, the Kentucky PHI office has publicized this solution at the Annual Maintenance Conference. When new instances of this deficiency are reported to the field office, information about this solution is provided to the individual PHA that had this problem.

9. Department of the Interior Bureau of Land Management Wild Horse and Burro Program[25]

Program and Objectives

The Bureau of Land Management (BLM) in the Department of the Interior is responsible for managing the nation's public lands. BLM addresses multiple uses on public lands, such as grazing, recreation, and commercial use, while seeking to sustain the health, diversity, and productivity of the public lands. The primary responsibilities of BLM's Wild Horse and Burro Program (WHBP) are to preserve and protect wild horses and burros and to manage healthy rangelands. WHBP activities include placing wild horses and burros for adoption.

Types of Uses of Outcome Data

This section provides examples of use of outcome data drawn from WHBP to:

- Develop plans
- Trigger corrective action
- Help evaluate new procedures before deciding whether to permanently adopt them (this use is rare)

Outcome Measurement Process

The examples in this section are based on use of two forms of outcome information regularly collected by BLM. BLM field staff monitor vegetation growth on an annual basis in arid areas.[26] They also conduct a census of horses and burros in herd management areas every three to four years.

WHBP also follows up on the condition of animals after they are placed for adoption—this process is described further in the use example. Collection of such "postprogram" outcome data is uncommon.

Use of Outcome Data to Develop Plans

BLM uses outcome information on land conditions and populations of various types of animals to help establish levels for different kinds of use that can be supported on specific BLM lands. BLM periodically sets appropriate management levels for the number of wild horses or burros that will sustain a "thriving natural ecological balance" on specific properties. BLM similarly establishes levels for the number of livestock for which ranchers are given permits to graze on specific rangelands.

Use of Outcome Data to Trigger Corrective Action

BLM uses regularly collected outcome information to identify types, levels, and locations of program activities that need to be implemented to restore outcomes to desired levels. This section focuses on corrective actions taken by WHBP. Outcome information also triggers corrective action in other BLM programs, such as revising the number of livestock permitted to graze in specific areas.

When outcome data on vegetation levels worsen or the number of wild horses or burros in particular areas exceeds the planned herd size, this triggers removal of "excess" wild horses or burros. Animals up to five years of age are offered to the public for adoption at periodic adoption events (auctions). Older animals, which are not considered desirable for adoption, are maintained in specific BLM properties.

A recent example from WHBP is the improvement in vegetation conditions in the Cibola-Trigo herd management area (HMA) in the Sonoran Desert near Yuma, Arizona, after examination of data led to removal of excess wild burros. BLM staff annually monitor loss of new growth of vegetation (primarily due to animal foraging) in that HMA. The objective for that HMA is to keep loss of new growth to no more than 20 percent. The average percentage of new growth loss for five key plant species in 2000 in two major portions of that HMA was 47 percent and 39 percent, respectively. Excess wild burros were removed from these two areas in 2000. Following that, the average percentage of new growth loss for those five species in the two major portions fell to 18 percent and 8 percent in 2001, exceeding the target.

WHBP also follows up on the condition of animals after they are placed for adoption. Title for animals placed with adopters does not pass until one year after the adopter takes possession. During that period, WHBP staff conduct at least one compliance check, usually an in-person inspection. If

animals are not appropriately cared for and the adopter does not correct the situation, BLM repossesses the animal. Adopters are also required to submit certification from a veterinarian, humane society, or similar entity at the end of that year. BLM grants title to the adopter after receiving this certification.

This example illustrates two forms of corrective action triggered by outcome information: removal of excess wild horses and burros and repossessing animals that are not appropriately cared for after being placed with adopters. The latter is unique in that it involves collection of "postprogram" outcome data and using that data to trigger corrective action.

Use of Outcome Data to Help Evaluate a New Procedure

WHBP staff in Arizona introduced a temporary change to their adoption program in FY 2002 in an effort to improve adoption outcomes. A problem facing WHBP was a "backlog" of available animals. Adoption events do not always result in adoption of sufficient numbers of animals. In FY 2001, for example, 342 animals were adopted in Arizona events, below the state program's target of 365 animals.

In the latter months of FY 2002, Arizona's WHBP implemented a "buy-a-buddy" program on a trial basis. Under this program, persons who adopt a horse or burro (at the minimum bid price of $125 or more) may participate in a lottery drawing to select a second horse or burro from the remaining animals. This second animal, which serves as a "buddy" to the first, may be adopted for $25. Implementing this pilot program required a temporary rule change by the Arizona WHBP to permit reduction in the adoption fee below the $125 minimum.

Only two of the state's FY 2002 adoption events included the buy-a-buddy program, because it was implemented late in the year. BLM staff estimate that 23 percent more animals were adopted at the two events when the buy-a-buddy system was in place than would have occurred without it. This estimate is based on the number of buddy animals adopted. Total number of animals adopted in FY 2002 increased over the prior year and exceeded by a substantial amount the year's target. Arizona's adoption target for FY 2002 was 360, but 445 animals were actually adopted. This includes adoptions at all 10 events held that year.

Based on this initial experience, Arizona's WHBP is continuing the buddy program in FY 2003. Arizona WHBP staff plan to review the post-adoption compliance data for time periods after the implementation of the buddy program to assess whether any undesirable effects occurred in terms of outcomes for adopted animals. Staff will review post-adoption data to see whether a larger percentage of noncompliant cases or repossessed animals occur than in prior years. WHBP staff feel that such increases might indicate that people adopt the second animal because of its low cost although they are not able to adequately care for it. Post-adoption data are not yet available because this program was implemented very recently.

The buy-a-buddy program has the potential for another benefit. If more animals are adopted at each event, WHBP may be able to save money by holding fewer adoption events.

These examples have two uncommon features. First, few agencies (federal government or others) have used regularly collected outcome information to help evaluate new procedures or programs. Second, some of the data that WHBP staff will use to assess the new procedure are based on follow-up data on the animal's condition after the initial adoption occurs. Few agencies regularly collect postprogram outcome data to monitor client conditions after leaving their program. WHBP routinely monitors animals' condition over a period of 12 months after placement for adoption.

10. Department of the Interior National Park Service[27]

Program and Objectives

The National Park Service (NPS) maintains the national park system and its facilities both to provide recreational and educational experiences to citizens and to help preserve the natural features and wildlife of the park lands.

Types of Uses of Outcome Data

This section provides examples of use of outcome information drawn from two national parks: Point Reyes National Seashore and Shenandoah National Park. Uses described here are:

- Motivating employees
- Justifying requests for funds
- Upgrading and maintaining physical plant
- Modifying service provision (times and locations when particular recreation opportunities are available)
- Educating and informing customers
- Training employees
- Modifying employee health and safety practices

Outcome Measurement Process

One source of regular outcome information for the NPS is an annual visitor satisfaction survey, known as the visitor survey card (VSC). This survey is conducted in all sites that have visitors.

The visitor survey is brief and is printed on a mail-in survey card. The VSC asks visitors to rate several types of park facilities (such as visitor centers, exhibits, and rest rooms); visitor services (such as maps or brochures, ranger programs, and commercial services); and recreational opportunities (such as sightseeing and outdoor recreation). Ratings on such features can be used by park supervisors and division chiefs to identify areas that may need improvement.

Two questions on the VSC address two NPS GPRA goals: visitor satisfaction and visitor understanding and appreciation.

- The question used for reporting visitor satisfaction asks visitors to rate the overall quality of facilities, services, and recreational opportunities using a scale of very good, good, average, poor, and very poor.
- The question addressing understanding appreciation is an open-ended question: "This park was established because of its significance to the nation. In your opinion, what is the national significance of this park?"

The VSC was first implemented in FY 1998 and is managed by the Visitor Services Project (VSP), part of the NPS Social Science Program. VSP staff developed the survey and detailed instructions (in the form of a VSC workbook) for park staff, who are responsible for conducting the surveys at their parks. Each park distributes 400 surveys to a random sample of visitors during one month of the year (the survey month for a given park is identified in the instructions). The VSC workbook provides instructions for selecting the random sample and procedures for distributing the survey to visitors. For example, cards must be handed to randomly selected visitors, not left in visitor centers for anyone to pick up.

VSP staff analyze data for most survey questions and report to the respective parks, also supplying the survey cards for that park. Staff at each park are responsible for analyzing the responses to the open-ended questions related to park significance. The VSC workbook provides detailed instructions for compiling responses to that question and matching them to items included in that park's previously identified set of significant features.

The NPS also regularly collects and reports data on injuries and accidents involving visitors or park employees. Such data are reported under the NPS goal that visitors safely enjoy park facilities.

Use of Outcome Data to Hold "How Are We Doing?" Meetings and to Motivate Employees

At Shenandoah National Park, management staff go over the annual VSC data at one of their regular staff meetings. This serves as a "How are we doing?" meeting. When VSC data for 2001 became available, the chief of interpretation and education at this park compiled a spreadsheet summarizing the park's satisfactory ratings on each question for

the four years in which the survey had been conducted. She did this to provide staff with a sense of how they were doing over time (trend analysis).

The chief of interpretation and education inserted notes on this table to highlight major changes—positive or negative. For example, one comment pointed out that the "poor" and "very poor" ratings for rest rooms had decreased (the responses for those rating categories were not included on the spreadsheet itself). Another note indicated that "poor" and "very poor" ratings had increased to 1998 levels in another category. Seeing increased satisfaction ratings reportedly make staff feel that the work they do is making an impact. In effect, such outcome information motivates staff.

Comments made on visitor comment cards, available at numerous locations throughout the park, also are tabulated and summarized at this park. Shenandoah park managers have made changes to the information provided to visitors as a result of open-ended comments on these cards. Here is one such example.

For several years, visitors frequently mentioned on the comment cards, or directly to park rangers or other staff, that the roadside areas looked "shaggy" or unkempt. However, the park was purposely not mowing these lands in an attempt to restore growth of wildflowers that originally bordered the road and to attract wildlife to those areas. Park staff began including explanations for the untrimmed roadsides in their printed materials for visitors and in other educational forums, such as ranger education presentations and exhibits. This led to a considerable reduction in the number of negative comments about roadside conditions.

The data from comment cards cannot be considered to be outcome data or to be representative of the visitor population. However, park staff use these comments to help them better understand the source of negative or positive data on the VSC. For example, comment cards have helped them identify the location of problems, such as specific rest-room facilities that have cleanliness problems. This enables them to target steps to correct the problems.

The superintendent of Point Reyes National Seashore also reported holding "How are we doing?" meetings when the annual VSC data are received. The superintendent reviews the data and discusses them at one of the weekly management team meetings. He focuses on identifying any changes (trends) from prior data and the scores that are relatively lower than other service characteristics in that facility. For example, scores on facility conditions—or conditions of a particular type of facility—may not be low in absolute terms, but may be lower than satisfaction with services. The management team identifies ways to bring about improvements.

Use of Outcome Data to Justify Requests for Funds and for Upgrading and Maintaining Physical Plant

Point Reyes has emphasized steps to address visitor satisfaction with rest rooms in recent years. Park management took a variety of steps to improve rest-room facilities, including upgrading fixtures (such as toilets and sinks). Some of this work was done with funds already allocated to the park budget. Other work was done with competitive funds the park sought from NPS. When applying for competitive funds, the park superintendent cited the satisfaction scores from the VSC as one of his arguments in support of funding. The park also increased the frequency of rest-room cleaning to twice per day on weekends in areas with larger amounts of visitors. They also put paper towels in some rest rooms—in addition to the energy saving hand dryers already in place—because many visitors complained about the lack of paper towels in the comments section of the survey card.

Results

Ratings of satisfaction with rest-room conditions on visitor survey cards increased from 66 percent rating these facilities as very good or good in 2000 to 74 percent rating them as very good or good in 2002. Ratings of poor and very poor decreased from 8 percent of respondents in 2000 to 4 percent in 2002.

Use of Outcome Data to Modify Service Provision and to Educate and Inform Customers

Point Reyes management also monitors data on park species, particularly endangered species. This is a form of trained observer rating. It has used that data to identify the need to take steps to manage

human recreation activities to protect particular species. Following are two examples of recent changes triggered by this type of outcome data. Note that the outcome data had been collected for many years, preceding GPRA.

The number of harbor seal pups at one of the "pupping colonies" in the park was decreasing through the early and mid-1990s. For example, the number of pups fell from approximately 250 in 1992 to approximately 150 in 1994. The number of pups in two other colonies remained relatively stable during the same time period. The primary difference between the three areas was visitor use. The area with decreased number of pups was used for kayaking.

In 1996, park management restricted kayaking during pupping season (mid-March through early July). Park management held meetings with kayak outfitters (which rent kayaks and also lead organized kayak tours in the park) and used their data on pupping to educate them about the problem and gain their support for the restriction on kayaking. The outfitters have been supportive of the change because they and the kayakers want to be able to see wildlife when using the park. Outfitters and park staff have referred visitors to other areas for kayaking.

Results

After kayaks were restricted, the number of pups in that colony increased to approximately 275 in 1996 and 1997. The number decreased in all three colonies in 1998 due to effects of El Niño but increased in 1999 (the latest data available) in all three colonies.

Data on the park's western snowy plover population (an endangered species) indicated that the survival of chicks in the nesting area in the park decreased beginning in 2000. The number of chicks fledged (the stage where they are able to leave the nest, approximately two months old) fell from approximately 24 in 1999, to approximately 15 in 2000, and to approximately 10 in 2001.

Park management implemented a visitor education program in 2001 to reduce human disturbance in the nesting area. On weekends, a ranger/education specialist was posted at the trailhead leading to the nesting area, or directly at the edge of the area, to educate visitors about the plover nesting area, the need to avoid walking there, and not bringing dogs in the area or taking them off their leash. Staff felt direct contact with visitors was a more useful approach than posting signs, which could easily be ignored.

Results

After taking these steps, data show an increased number of chicks fledged—approximately 17 for the 2002 nesting period.

Use of Outcome Data to Train Employees and to Modify Employee Health and Safety Practices

Data on injuries and accidents involving visitors or park employees are routinely collected by the NPS. Employee safety also is an important consideration for managers and may be considered an intermediate outcome.

The Point Reyes management team meets monthly to review accident data, which are compiled by park staff and entered into the NPS management information system on a monthly basis. Following are examples of changes triggered by high levels of employee "lost time" due to particular types of accident or injury.

One of the major causes of employee lost time in Point Reyes in recent years is poison oak infections. Management targeted that as an area to address. It did so in a number of ways. It provided extra training for staff who performed trail work in areas affected by poison oak, including Youth Conservation Corps employees who work on trails during the summer months. The training emphasized how to recognize and avoid contact with poison oak. The park also supplied lotions and products for washing after potential contact with poison oak.

Park management also initiated the practice of having staff change clothing immediately after returning from working in the field and "bag" their uniforms to take them home to launder. This prevents spreading poison oak while performing other tasks. In addition, park management kept employees known to be sensitive to poison oak from working in those areas.

Results

Recent outcome data indicate that management's efforts appear to have contributed to improvement in this area. The superintendent reported that in 1997, 48 percent of lost time days were related to poison oak. This decreased to 12 percent in 2001, and there were no lost time days associated with poison oak as of this writing in 2002.

The amount of time lost due to carpal tunnel or ergonomic-related injuries has been increasing in recent years. To address this, management arranged for several training sessions focused on preventing such injuries. Some training sessions were provided in 2001 and 2002; others are scheduled for later in 2002 or 2003. Reduction in lost time from these types of injuries has not yet been apparent in the data.

11. Department of Labor Occupational Safety and Health[28]

Program and Objectives

The mission of the Occupational Safety & Health Administration (OSHA) is to save lives, prevent injuries, and protect the health of America's workers. To meet these objectives, OSHA establishes and enforces protective standards and provides technical assistance and confidential consultation services. OSHA and its state partners have approximately 2,100 inspectors, as well as other technical and support personnel, throughout the country.

Types of Uses of Outcome Data

OSHA uses regularly collected outcome information for the following interrelated purposes:

- Develop its annual site-specific targeting (SST) inspection plans
- Direct compliance activities, such as identifying and notifying employers who need, and are eligible to receive, compliance assistance
- Trigger enforcement actions
- Encourage employers to improve health and safety

OSHA uses a combination of both aggregate injury-rate data and the rates for individual employers as bases for action.

Outcome Measurement Process

OSHA conducts an annual nationwide survey of establishment-specific injury and illness incidents from approximately 80,000 establishments. Employers with at least 40 employees are required to maintain and report (electronically or by mail) counts of injuries and illnesses, hours worked, and number of cases with days lost to illness or injury.

OSHA then calculates the lost workday injury and illness (LWDII) rate, the combined number of lost workday injuries and illness days per 200,000 hours worked (equivalent to 100 full-time employees). The national average rate for private industry in 2000 was 3.0.

Use of Outcome Data to Develop Site-Specific Targeting Inspection Plans

LWDII rates are used to select work sites for inspection. For example, the 2001 Data Initiative collected injury and illness data from 2000 for use in generating the 2002 Site-Specific Targeting Plan (SST-02) Primary Inspection List. The national office provides each area office with software and databases containing the establishments of the primary inspection list as well as secondary inspection lists.

Use of Outcome Data to Direct Compliance Activities

The plan selects all work sites with a rate of 14 or higher to undergo inspection (approximately 3,000 sites). In addition, to check reporting quality on employers reporting low injury rates, OSHA randomly samples approximately 200 establishments with low rates (rates between 0.0 and 8.0) for inclusion on the primary inspection list. In addition, if an area office is able to complete its primary inspection list before the expiration of the annual SST plan, it is expected to inspect additional establishments selected from the secondary inspection list, which includes employers with LWDII rates less than 14.0 and at least 8.0.

Inspections cover both safety and health compliance, including verification of employers' reported data. Record-keeping violations are cited. Walk-throughs may be conducted to interview employees in order to confirm and verify the injury and illness experience.

Use of Outcome Data to Trigger Enforcement Actions

Any serious violations of health or safety requirements must be investigated and can result in the issuance of citations. OSHA has abatement criteria standards that identify what actions employers must take to gain compliance status and what documentation must be provided. Employers are not required to certify abatement for violations they correct during an on-site inspection, but they are required to provide abatement documentation for any serious violation. In the event that an employer does not respond accordingly, the case can be reviewed for further enforcement action.

Use of Outcome Data to Encourage Employers to Improve Health and Safety

The injury rate data are also used to identify and encourage employers with moderately high rates (8.0 through 14.0) to voluntarily improve their health and safety compliance. Letters are sent to employers in this category suggesting that they hire outside health and safety experts or make use of a state consultation program, which OSHA funds. The program, which is confidential and free, is designed for employers with fewer than 250 workers. A list of employers receiving such letters is posted on the OSHA website.

12. United States Postal Service[29]

Program and Objectives

The United States Postal Service (USPS) seeks to bind the nation together by facilitating both access to mail services in all communities and the prompt, reliable delivery of business, personal, and educational correspondence. Since the Postal Reorganization Act of 1970, the Postal Service has been undergoing transformations designed to ensure operations are financed through postage paid by users rather than by tax revenue subsidies. As the business environment of the Postal Service has evolved, the agency has recognized the fundamental requirement to provide services that meet the needs of its various business and consumer customers.

Types of Uses of Outcome Data

Uses described here are to:

- Identify opportunities to improve customer satisfaction
- Improve service performance
- Motivate personnel, including use of monetary incentives
- Report performance to key leaders and the public

Outcome Measurement Process

The Postal Service uses a variety of outcome measurement approaches to measure both customer satisfaction and the effectiveness of core business activities. For example, three systems measure customer satisfaction: Consumer Service Card System (CSCS), Customer Satisfaction Measurement (CSM), and WEB–Ease of Use (WEB-EIS). CSCS is used to monitor customer service issues. It tracks and reports problems, suggestions, complaints, and information requests submitted by postal customers. CSCS maintains year-to-date and summary history files for internal use. The CSM system represents the perceptions of all U.S. consumer and commercial customers, whether or not they use postal service or products. Surveys are conducted by the Gallup Organization on various postal issues from the consumer's perspective (such as accuracy and consistency of delivery; retail clerk courtesy, knowledge, and responsiveness to customers; and telephone courtesy and accuracy of information provided). Such information is used for both strategic and tactical decision making, and to identify opportunities to improve customer satisfaction. WEB-EIS is derived from the CSM process. This system is for internal use only and focuses on information to improve USPS interaction with consumer and commercial customers.

USPS regards first-class mail as one of its flagship products. Consequently, some emphasis has been placed on achieving efficient and effective delivery of first-class mail. For example, USPS uses an Origin–Destination Information Sampling System (ODIS) to sample mail at Postal Service exit points and record information on postmark dates and other characteristics of internal interest. ODIS is considered a diagnostic tool that field managers can access online to see volume, how long mail is taking, and whether mail is on time.

Additionally, the External First-Class Measurement System (EXFC) is used to measure the timeliness and consistency of first-class mail delivery. EXFC is performed by an external contractor and provides an independent assessment of the time elapsed between when a piece of first-class mail enters the mail stream via a collection box and its delivery at one of the 138 million homes, businesses, or post office boxes served by the Postal Service six days per week.

EXFC was developed in 1989 in response to suggestions from customers that an independent measurement system was needed to assess service as experienced by customers. The process tracks the time from when mail is picked up to mailbox delivery performance for first-class mail. EXFC uses sampling procedures to select a panel of 463 Zip Code areas (based on geographic and volume density) from which 90 percent of first-class mail volume originates and to which 80 percent of the first-class mail is delivered. Testing is done continuously throughout the year, with equal volumes of test mail in each postal quarter. Test mail is generated with a variety of characteristics (different sizes, shapes, print types, colors, and window types). Some test mail is pre-barcoded, as well. The test mail is indistinguishable from real mail by USPS employees. Each piece is assigned an identifying number to enable tracking through every step of

the process. A panel of "droppers" located in each area of the country mail the test materials at a predetermined, randomly assigned time and location. Test mail is received by a panel of "trained reporters" representing households, small businesses, and post office boxes. They report the exact receipt date of test mail.

The system is designed to produce statistically representative measures for each performance cluster each quarter. The primary performance measures are the percentage of mail delivered within the service standard and the average number of delivery days. Results are reported by service standard for each of 85 performance clusters in eight USPS area offices and one major metropolitan area in the United States.

Use of Outcome Data to Improve Performance

The Postal Service has for many years been very sensitive to its key outcome data (particularly origin-to-destination times), probably as much as or more than any federal agency, because of the high level of visibility of its service to citizens and public officials. Its quarterly reports of delivery times have been issued publicly for many years. We had difficulty pinning down specific examples of use. Nevertheless, we believe the Postal Service should be included in any identification of programs that have made significant use of regularly collected outcome information.

The Postal Service has defined 85 management units by geographic areas called performance clusters. In addition, it sets annual performance standards for timely delivery of first-class mail. For example, as part of its Annual Performance Plan for 2000, USPS stipulated that it would achieve overnight first-class delivery on-time performance of at least 93 percent and two- and three-day first-class delivery on-time performance of at least 87 percent.

The analysis of EXFC data provides measures of service performance nationally, by area office, and for performance clusters. These data are compared with service delivery standards on a quarterly basis.

Use of Outcome Data to Motivate Employees

USPS has a history of using information such as financial performance and overnight mail delivery scores to judge and reward the performance of the agency's senior executives. Bonuses have been awarded for performance improvements; for example, in 1997, 1,000 postal executives received bonuses of up to 12 percent of their annual salaries as a result of performance in the preceding year.

In addition, the Postal Service has implemented a group incentive plan, participated in by most of the nonbargaining Postal Service employees. The plan rewards results for sustained financial performance and continuous improvement in customer service and the workplace. Targets are established annually. Management reviews the targets selected for inclusion and determines the weighting of each indicator. Performance levels compared with the targets are reported throughout the year, with final calculation completed at the end of the fiscal year. Once the overall performance values are determined, USPS funds the incentive pool: the higher the values, the larger the potential incentive payment. Incentive payments may be increased by exceeding targets. Actual performance, by unit, is compared with targets, and payment amounts are determined. Provisions are in place to exclude "nonperforming" employees from benefiting from group performance rewards.

Use of Outcome Data to Report to Key Leaders and the Public

The USPS Corporate Financial Planning group issues ODIS Quarterly Statistics Reports. The Office of the Consumer Advocate reports the Gallup poll findings and EXFC results to the Postal Service Board of Governors each postal quarter. In addition, the findings are issued as quarterly press releases and are archived on the USPS website (www.usps.com) to inform the public. For example, in January 2003, it was reported that five performance clusters achieved on-time delivery performance scores of 96 percent, and an additional 33 performance clusters delivered first-class mail on time an average of 95 percent of the time during the preceding quarter. Overall, first-class mail had an on-time delivery performance score of 94 percent—the third consecutive quarter in which this benchmark was reached.

13. Social Security Administration Supplemental Security Income Program[30]

Program and Objectives

The Social Security Administration's (SSA's) Supplemental Security Income (SSI) Program provides supplemental income for eligible aged, blind, and disabled citizens. SSI has for many years been tracking such outcomes as payment accuracy, amount of overpayments, and amount of overpayment collections—as well as indicators of service quality to customers, such as application processing times, accuracy of responses to telephone inquiries, and caller satisfaction with SSI services. At the end of FY 2002, there were 6.8 million beneficiaries receiving SSI payments, totaling over $30 billion for FY 2002.

SSI is administered through 1,300 field offices in 10 regions. Some of these data, broken out by office, are fed back to individual offices.

Types of Uses of Outcome Data

SSI uses regularly collected outcome data to:

- Make a variety of operational enhancements
- Identify the need for policy changes
- Evaluate the extent to which changes in practices and policies have led to improvements in outcomes
- Motivate field office managers and their staffs
- Identify needed legislative changes

This effort has been stimulated in part by the program's listing by the GAO in 1997 on its "high risk" list. GAO criticized SSI for an overemphasis on production and service over program integrity.

Some of the outcome information is used both in aggregated data (such as data for a whole field office and data for all a particular category of claimants) and to enable claims representatives to provide assistance to individual claimants in cases where processing times or other problems appear excessive.

Outcome Measurement Process

For many years SSA been tracking such outcomes as payment accuracy, amount of overpayments, and amount of overpayment collections. It has been tracking indicators of service quality to customers, such as application processing times, accuracy of responses to telephone inquiries, and caller satisfaction with SSA services provided to SSI applicants and beneficiaries.

At least annually, but for many indicators more frequently (such as monthly), SSI reports outcome data on:

- SSI outlays
- Dollar amount of detections of new overpayments to beneficiaries and as a percentage of outlays
- Overpayment collections, both the dollar amounts and as a percentage of outlays
- Overall payment accuracy (overpayments and underpayments) based on a random sample of cases)
- Projected dollar amount of overpayments and underpayments to beneficiaries based on the sample results
- Application processing time (from the time the applicant submits a completed application until payment goes out or is denied), broken out by field office and region
- Accuracy and courteousness of responses to queries (based on monitoring a random sample of clients calling the 800 or local office number)
- Client satisfaction with the helpfulness of the service (based on a national random sample of cases)

SSA also provides weekly workload data on SSI claims, post-eligibility, and other work processed by each SSA field office.

In November 2002, SSI completed processing approximately 174,000 applications for blind and disabled applicants. The national average processing time for these decisions (including approvals and denials) was approximately 101 days. (About

82 percent of these applications required a medical decision to determine whether the applicants met the disability requirements of the Social Security Act. The remaining applications did not require a medical decision and were denied because some nonmedical factor of eligibility was not met.)

For some of these indicators, data, such as processing times, are tabulated and reported by individual field office. The processing-time reports are provided to each of the 10 regions (showing the average times for each region as well as the national average). Processing-time reports are also sent to each field office (showing its own averages, those of field offices in its area and region, and the nationwide average). The processing-time data are also broken down by such factors as (1) whether the claims were for aged or for blind and disabled persons; (2) for aged claims, the percentage of claims processed within seven, 14, 30, and 60 days; (3) whether the claims were awarded or denied; and (4) whether the time was in the field office or also included the number of days the claim was being processed outside the field office.

Use of Outcome Data to Make Operational Improvements

SSI has made a number of operational improvements based on the data. To prevent overpayments to beneficiaries, SSI, with the help of the data, identified the two top reasons for overpayments: unreported wages and unreported bank accounts. SSI began running quarterly matches against wage data maintained by the Office of Child Support Enforcement. It also began receiving electronic reports from prison facilities to identify recipients no longer eligible for SSI payments.

SSI has increased the number of its periodic redeterminations of financial eligibility (periodic reviews of the nondisability factors of SSI eligibility), especially of the more complicated error-prone cases. For example, cases with an expectation of medical improvement are electronically flagged for review at frequencies related to the amount of improvement that is expected. For some of this activity, SSI is seeking added funding, using the outcome data to help it make its case.

To improve collection of debt, SSI is taking steps such as distinguishing those debts that are in an active repayment agreement with the debtor and those debts for which SSI is seeking to obtain a repayment agreement.

Field office staff in at least one office (a California field office) have used the processing times, as well as workload information, to identify needs to shift the office's resources between processing "aged" and "blind and disabled" claims.

Use of Outcome Data to Identify Need for Policy Changes

SSI has begun focusing on policy changes needed to simplify the program as a way to prevent payment errors. Some such changes require legislation changes. For example, SSI hopes to reduce cumbersome monthly wage reporting requirements by using an annualized average. Under existing law, wages earned by SSI recipients are used to help determine the recipient's benefit. Therefore, recipients must report any changes in their monthly wages within a very short time frame in order for the field office to input the wage data to affect the payment before it is issued. SSA is analyzing the option of assigning wages to each month using the average wage over the number of months worked at a specific job, or over the calendar year, whichever is less. Wage averaging would require legislation.

Use of Outcome Data to Motivate Staff

The extensive data on processing times and other outcomes provided to regions and field offices are intended to stimulate federal staff focus on outcomes such as processing times. Each field office and region is provided comparison data enabling an office to compare its own outcomes with the national average, other regions, and other field offices.

The outcome data have also been used in some field offices to help allocate the field office's pool of funds for awards to individuals within the office. Offices have used both monetary and time-off awards.

Results

SSI reports substantial improvements on its outcomes since 1998. New detections as a percentage of SSI outlays went from 4.4 percent in FY 1998 to

5.9 percent in FY 2001, a 34 percent increase. This amounts to an improvement of $477 million more in FY 2001 than if the FY 1998 4.4 percent detection rate had occurred.

The amount of overpayment collections as a percentage of outlays has improved each year since FY 1998, going from 1.8 percent in FY 1998 to 2.4 percent in FY 2001, a 33 percent increase. This amounts to an improvement of $205 million more in FY 2001 than if the FY 1998 1.8 percent collection rate had occurred.

14. Department of Transportation Coast Guard Marine Safety, Security, and Environmental Protection

Note: The following is one of the first examples, if not the first example, of the use of outcome information derived from GPRA. This material is adapted from a 1996 GAO report.[31] We felt that any report of use by federal programs to improve their services with the help of outcome data would be remiss not to include this classic example.

Program and Objectives

The mission of the Coast Guard's Office of Marine Safety, Security, and Environmental Protection is to protect the public, the environment, and U.S. economic interests through the prevention and mitigation of marine incidents.

Encouraged by the outcome focus of GPRA, the Coast Guard changed its focus from outputs to outcomes in its first business plan, dated January 1994. It recognized that the mission of its marine safety program was to save lives, not to conduct more and better inspections.

Types of Uses of Outcome Data

The Coast Guard used the outcome data in this example to:

- Identify problem areas
- Make program changes based on that information

Outcome Measurement Process

Traditionally, the Coast Guard based its marine safety efforts on inspections and certifications of vessels. It measured its performance by counting outputs, such as the number of prior inspections and outstanding inspection results. Under GPRA, it shifted its focus to accidents and casualties.

In particular, it began to break out (disaggregate) the data both by industry (such as the towing industry) and by estimates of the causes of the accidents and casualties. For example, categories of causes included equipment and material failures, human factors, and environmental factors.

Use of Outcome Data to Identify Problem Areas

When the data on marine accidents were broken out by cause, the data indicated that accidents and casualties were more often caused by human error, not by deficiencies in the vessels or by other factors.

In addition, the program found that the towing industry had a significant role in marine safety. For example, towing industry data for 1982 through 1991 showed that 18 percent of reported casualties were caused by equipment and material failures, 20 percent by environmental and other factors, and 62 percent by human factors.

Use of Outcome Data to Make Program Changes

The program began to work with the towing industry to build the knowledge and skills of entry-level crew members in the industry. The Coast Guard and the towing industry jointly developed training and guidelines to reduce the causes of fatalities.

The Coast Guard also indicated that it had given greater authority to field commanders, and it invested in activities that went directly to the goal of reducing risks on the water.

Results

This joint effort led to a significant decline in the reported towing industry fatality rate: from 91 per 100,000 industry employees in 1990 to 27 per 100,000 in 1995.

15. Department of the Treasury Internal Revenue Service Toll-Free Customer Service Program[32]

Program and Objectives

One of the Internal Revenue Service's (IRS) programs is the provision of a toll-free telephone number for citizens to use to obtain help determining their taxes. Major concerns are the accuracy and helpfulness of the information provided by IRS personnel to taxpayer inquiries.

IRS has established a process for sampling those calls and monitoring the responses of the IRS customer service representatives. It interviews random samples of taxpayer callers to ask them about their experiences with those calls.

Types of Uses of Outcome Data

IRS is using the toll-free program to test a process in which the outcome information (in this case, call-response accuracy data) is used to identify the root causes of problems and then develop action plans for correcting those problems.

Uses of call accuracy outcome information, thus far, have been to:

- Help guide program changes
- Monitor success of these efforts
- Identify major problem areas, particularly by breaking out the data into considerably more specific problem areas
- Set accuracy targets for local sites
- Establish a basis for annual action plans developed by each local site to address problems and serve as a basis for continuous improvement efforts

Outcome Measurement Process

Here we cover two procedures used by this IRS program:

- Central staff monitor samples of calls received by the toll-free numbers to identify the accuracy and quality of the responses given by local call sites.
- Samples of callers, whose calls have been monitored, are interviewed to obtain information on their call experience.

Procedure 1

A central staff of about 80 employees listens in on a randomly selected sample of telephone calls to the IRS toll-free lines. The samples monitored cover each of the IRS 15 call sites throughout the country. (These call sites focus on individual taxpayers rather than those who are self-employed or run small businesses.) The call monitor rates the accuracy of the response given by the IRS customer service representative. The rater makes an overall rating as to whether the informant gave a correct or incorrect response. The rater identifies any errors the customer service representative made and classifies errors by a pre-designated category. In filing year 2001, over 6 million tax law calls were received; in 2002, over 6.5 million calls were received. Of these calls, 15,202 were monitored in 2001; 26,940 were monitored in 2002. The calls monitored in 2001 had a "quality" rating of 75 percent; the rating was 81 percent for 2002. A call is counted as being of "quality" if the information provided was error-free and handled in accordance with all the proper IRS procedures.

This call-monitoring procedure has been used for a few years. Previously, the IRS used "testers," IRS employees, to role-play as taxpayers asking for information. However, that procedure was believed to have validity problems and was replaced by the call-monitoring process.

Beginning with the FY 2001 filing season, IRS has begun testing a considerably more systematic process for tabulating errors and using that information to develop action plans for correcting significant problems. It has used a contractor to conduct the basic analysis. Errors are tabulated in many different ways, such as by type of error, by the topic the taxpayer asked about, by reason for the error, by site, and by month.

The findings of the call-monitoring effort are reported on the IRS intranet. Individual call sites can now access the data pertaining to their calls and to other call sites.

IRS is planning to widen the application of this process to other IRS programs. It expects to use its own employees for the analysis.

Procedure 2
A random sample of callers is selected and routed to an automated survey line. Callers are asked a number of questions about 12 characteristics of their call experience, including professionalism, courteousness of employees, ease of getting through by phone, completeness of the information, and ease of understanding the information.

Uses of the Outcome Data

Uses of Procedure 1 Information
The root-cause analysis used by IRS illustrates the value of federal programs being able to break out their outcome data by many characteristics of the program and customer. The process enables the program to mine the data in many useful ways to pinpoint the nature of problems, suggesting later corrective action.

The root-cause analysis procedure, undertaken for the FY 2001 filing season (the first year in which this process was used), found that one error type, lack of complete probing, occurred more than twice as often as any other type of error. It found that this was the case for all 15 field offices. It also found that this was the case regardless of the topic (whether the question was about filing/dependents, estimated tax, interest/dividends, small business, pensions or Social Security benefits, and so on). This indicated that the problem was not a local problem but an overall system process problem.

The analysis found that within individual application types, errors focused on certain reasons for the error. For example, for errors related to "filing/dependents," two call types accounted for over 61 percent of the errors, with most of those errors due to IRS personnel not using specific pages of the IRS customer service guide.

Based on this information, IRS added actions such as the following to the IRS 2002 Action Plan:

- Improve the content of specific pages of the IRS guide for IRS personnel responding to queries
- Provide training to customer service representatives on these pages at several points before and during tax-filing season
- Conduct local, site-level reviews to ensure the use of these key pages by IRS personnel who respond to queries

The root-cause analysis went further. It also examined error rates by month of the filing season (January through June). It found that 75 percent of the call volume occurred during the first four months of the year, the time that error rates were the highest. From other sources, the analysis found that the majority of training occurs in late December and January, indicating that most calls come in when the IRS personnel are least able to answer the questions correctly. This timing problem had been identified earlier from data for the 2000 filing year.

The root-cause analysis for the next year, the 2002 filing year, found that 66 percent of all defects were due to lack of use of the IRS guide by IRS personnel responsible for responding to queries. The program found that the guide was not distributed early enough for training and that IRS personnel did not have enough training on the guide.

For FY 2002, actions taken included the following:

- The guides were improved.
- Training was started earlier, with added training in the top errors.
- Error-rate targets were established for each of the 15 field offices.

Results: One site, Jacksonville, had already taken proactive measures based on the data on error rates, and it held training a month earlier, in November and January. This resulted in a substantial improvement (8 percentage points) in the error rate during January 2001 from January 2000, and 2 percentage points of improvement for each of the next three months. Overall its error rate was about 10 percentage points better than the national average for 2001.

The overall quality rate for tax law queries in filing year 2002 was about 7 percentage points higher than that for 2001 (from about 74 percent accuracy

rate to about 81 percent). For January 2002, the rate was about 9 percentage points higher than that for January 2001 (from about 70 percent to 79 percent). This amounts to about 750,000 fewer incorrect responses given to customers.

The FY 2002 Action Plan targeted three tax subjects as improvement opportunities, based on the analysis we described. The error rates for these applications for FY 2002 were 30 percent to 42 percent lower than for FY 2001.

Procedure 1 information can be used for target setting and preparing annual action plans. IRS, in addition to the root-cause analysis, also uses the data on accuracy and quality to set and monitor local office targets and, thereby, encourage local offices to seek continuous improvement. Each of the 15 call sites has its own targets on call quality. These targets are based on the prior year's quality levels. The toll-free program has its own national target. Local sites far from the national target are given goals for larger increases than local sites at or near the national target.

Each local site is required to prepare a quality action plan at the beginning of each fiscal year. These plans are supposed to address the sites plans for training and addressing errors found in the previous years.

Uses of Procedure 2 Information

IRS has continued to get less than desired favorable responses, particularly to its questions about ease of getting answers. To improve the results, the program made adjustments beginning in December 2002 to ease a caller's ability to get through to the right party. Actions include adding new phone lines and changing the routing for some types of calls. For example, callers responding to an IRS notice are given in the notice a special telephone number so they do not have to go through unnecessary telephone options. Calls from businesses are provided a separate toll-free number than that for taxpayers with individual income tax questions.

16. Department of Veterans Affairs Veterans Health[33]

Program and Objectives

The Veterans Health Administration (VHA) is the largest health care system in the United States, caring for more than 4 million veterans each year with approximately 180,000 staff members and a budget of more than $21 billion. Since 1994, VHA has been embarked on an ambitious effort to improve service quality. VHA is now organized into 21 Veterans Integrated Service Networks (VISNs), and each has substantial operational autonomy including contracting for health-care services and generating revenue by selling excess services. Each VISN includes several medical centers, associated outpatient clinics, and a number of nursing homes—a total of more than 600 facilities.

Types of Uses of Outcome Data

VHA has used regularly collected quality and outcome information in:

- Personnel management and motivating employees, including use of performance agreements
- Allocating and reallocating resources
- Redirecting program activities

Outcome Measurement Process

Key performance indicators for VHA cover:

- Patient satisfaction (both inpatient satisfaction and outpatient satisfaction)
- Waiting times (for example, percentage of primary care appointments scheduled within 30 days of the desired date, percentage of specialist appointments scheduled within 30 days of the desired date, and percentage of patients who report being seen by a provider within 20 minutes of their scheduled appointment)
- Health-problem prevention index
- Chronic disease care index

Over the past several years, as part of efforts to improve the quality of patient care and as the practice of medicine has changed, new performance indicators have been developed. The current prevention index (Prevention Index II) comprises nine indicators that focus on immunizations, cancer screening, colorectal cancer screening, and so on. The current chronic disease care index (Chronic Disease Care Index II) comprises 23 indicators that focus on how well VHA follows nationally recognized guidelines for heart disease, hypertension, diabetes, depression, congestive heart failure, chronic lung disease, smoking cessation, and so on. Both indices track the extent to which each veteran is receiving care that follows clinical practice guidelines that are linked to better health-care outcomes.

Use of Outcome Data for Personnel Management and Motivating Employees

VHA links rewards and recognition to performance data, including regularly collected quality and outcome indicators. Directors of VHA's integrated service networks and other executives in the networks function have negotiated performance agreements that give considerable weight to quality and outcome objectives. For FY 2000, for example, the performance agreements for network directors (based on the network performance plans) consisted of three parts: (1) core competencies such as interpersonal effectiveness have a weight of 20 percent; (2) areas of special emphasis for VHA such as patient safety have a weight of 30 percent; (3) performance goals and performance targets for health-care quality, satisfaction, waiting times, costs, and so on, have a weight of 50 percent.

To encourage performance focus at lower levels, VISNs implement performance agreements between the VISN directors and individuals reporting to them such as medical center directors, clinical managers, and quality managers. Performance targets for different facilities and their executives may vary, depending on the degree of difficulty of the tasks they face. The performance agreements for nurse executives include indicators and targets on which nurses have impact; those for chief financial officers focus on cost-efficiency issues.

VHA has a standardized approach to monitoring VISN directors' progress during the year. It consists of formal reporting of progress toward the health-care–related goals and areas of special interest.

Network directors are given quarterly feedback on performance in terms of the indicators in their network performance plan, as well as feedback on additional performance indicators and formal network performance reports. These reports are followed by meetings with VHA's chief network officer (CNO) to discuss performance so that there are no surprises at the end of the year. To evaluate the performance of VISN directors each year, the CNO and the undersecretary for health use both quantitative and qualitative information, including progress toward VHA's health-care–related goals, the director's self-assessment describing actions and accomplishments that reflect each core competency, and the CNO's assessment of the director's management competencies included in the agreement.

The executives' evaluations and bonuses depend to some degree on program quality and outcomes. When facilities are not doing well and there appear to be leadership problems, executives may move to new assignments within VHA or may retire.

Use of Outcome Data for Allocating and Reallocating Resources

Resource distribution to VHA's health-care networks is determined primarily by the number of veterans being served. Performance information is used in resource allocation decisions at the health-care network level during budget execution, that is, after VHA receives its appropriations and the funds are allocated to the network. Network directors allocate and reallocate resources to meet performance goals including quality and outcome goals.

To reduce waiting times for service at specific facilities (for example, in rural areas that are having difficulty recruiting sufficient staff), networks have reallocated resources including rotating clinicians through the facility.

To improve coverage of the receipt of retinal eye examinations by diabetic veterans, one network invested in machines that record test results and, if necessary, transmit them to an ophthalmologist at another location, thereby increasing the network's capacity for meeting this performance target.

Some networks give bonuses to facilities that have done well on the performance indicators and have contributed to the network's performance (for example, by sharing information on best practice approaches).

Use of Outcome Data to Redirect Program Activities

In monthly performance review meetings with the deputy secretary and the network directors, the undersecretary for health reports progress to date on key indicators of financial performance and program performance, the status of major projects, and key workload indicators. In these meetings, the head of VHA reports on overall national performance and on the performance of each of the 21 VISNs. The meetings are used to share best practices, focusing on areas that are not going well and discussing steps that VHA and individual networks are taking to rectify the problems.

Earlier, to improve the outcomes of cardiac procedures, VHA looked at factors associated with morbidity before surgery and outcomes 30 days after surgery, and fed that information back to surgery departments to stimulate improvements in outcomes.

Results of Use of Outcome Data

VHA has made significant progress in implementing nationally recognized clinical interventions applicable to heart disease, hypertension, pulmonary disease, and other chronic conditions, and it has made significant progress in implementing interventions such as immunizations, cancer screening, tobacco consumption screening, and alcohol screening that are nationally recognized for prevention and early detection of disease. Scores on Chronic Disease Care Index I rose from 44 percent in FY 1996 to 81 percent in FY 1999. Scores on Prevention Index I rose from 34 percent in FY 1996 to 79 percent in FY 1998. In FY 2001, scores on the Chronic Disease Care Index II and the Prevention Index II were 77 percent and 80 percent, respectively.

VHA lowered mortality rates for cardiac procedures by 13 percent over an eight-year period. Mortality rates are now at least as good as in the private sector.

In FY 2001, the percentage of inpatients rating VA health services as good or excellent was 64 percent; for outpatients, 65 percent. In FY 2001, the percentage of primary care appointments scheduled within 30 days of desired date was 87 percent; the percentage of specialty appointments scheduled within 30 days of desired date was 84 percent; and the percentage of patients who reported being seen within 20 minutes of scheduled appointments was 63 percent.

Appendix: Acknowledgment of Federal Personnel

The following list identifies federal personnel and others who provided the basic information for this report. They are listed in alphabetical order by agency.

Department of Agriculture
(Phone Locator: 202/720-8732)

- Glenn A. Germaine, Animal and Plant Health Inspection Service
- Dr. Ed Gersabeck, Animal and Plant Health Inspection Service
- Dr. Stephen A. Knight, Animal and Plant Health Inspection Service
- William O. Macheel, Animal and Plant Health Inspection Service

Department of Education
(Phone Locator: 202/401-2000)

- RoseAnn Ashby, Rehabilitation Services Administration, Basic State Grants Branch, Program Administration Division
- Michael Dean, Office of Vocational and Adult Education
- Martha Garber, Rehabilitation Services Administration, Dallas Regional Office
- Fred Isbister, Rehabilitation Services Administration, Projects With Industry Program
- Mary Jane Kane, Rehabilitation Services Administration, Projects With Industry Program
- Diana Koreski, Rehabilitation Services Administration, Seattle Regional Office
- Mary Anne Langton, Rehabilitation Services Administration, Boston Regional Office
- Ann Manheimer, Office of Management, Performance and Process Improvement Services
- Bill Schubauer, Rehabilitation Services Administration, San Francisco Regional Office
- Mary Suazo, Office of Migrant Education

Department of Health and Human Services
(Phone Locator: 202/619-0257)

- Jane Axelrod, Food and Drug Administration, Center for Drug Evaluation and Research
- Butch Bosin, Food and Drug Administration, Office of Policy and Planning
- Gary Buehler, Food and Drug Administration, Office of Generic Drugs

Department of Housing and Urban Development
(Phone Locator: 202/708-5004

- George Davanel, New Jersey Field Office
- Edward DiPaula, New Jersey Field Office
- Elizabeth A. Hanson, Real Estate Assessment Center

- William Hall, Memphis Troubled Agency Recovery Center
- Ann Jefferson, Wellston, Missouri, Public Housing Authority
- Patricia Knight, Cleveland Troubled Agency Recovery Center
- Carmen F. Valenti, New Jersey Field Office
- Art Wasson, Kentucky Hub/Field Office

Department of Labor
(Phone Locator: 202/693-5000)
- Joseph J. DuBois, Occupational Safety and Health Administration, Office of Statistical Analysis
- Thomas Galassi, Occupational Safety and Health Administration, Office of Enforcement

Department of the Interior
(Phone Locator: 202/208-3100)
- Jennifer Hoger, University of Idaho, NPS Visitor Survey Card Project
- Kelly Grissom, Bureau of Land Management, Arizona Office
- Karen Michaud, Shenandoah National Park
- Don Neubacher, Point Reyes National Seashore
- Roger Oyler, Bureau of Land Management, Yuma (Arizona) Field Office
- Mary Pyles, Bureau of Land Management, Arizona Office
- Lili Thomas, Bureau of Land Management, Nevada Office

Department of the Treasury
(Phone Locator: 202/622-2000)
- Patricia LaPosta, Internal Revenue Service, Customer Accounts Service
- Joseph H. Myers, Internal Revenue Service, Strategic Planning and Policy Development, Small Business and Self-Employment
- Dorene A. Viglione, Internal Revenue Service, Wage and Investment Customer Accounts Service

Department of Veterans Affairs
(Phone Locator: 202/273-5400)
- Stanlie Daniels, Veterans Health Administration, Office of Quality and Performance
- Mark Russell, Office of the Secretary

Environmental Protection Agency
(Phone Locator: 202/260-2090)
- Dan Fiorino, Office of Policy, Economics and Innovation, National Environmental Performance Track Program
- Arthur Horowitz, Office of Enforcement and Compliance Assurance
- Jenny Noonan, Office of Air and Radiation, Office of Air Quality Planning and Standards
- Dan Palmer, Office of Enforcement and Compliance Assurance, Office of Compliance
- Amy Porter, Office of Enforcement and Compliance Assurance, Office of Compliance
- Cecilia Tapia, EPA Region 7 (Kansas City)
- Robert Tolpa, Office of Enforcement and Compliance Assurance, Office of Compliance, Planning Branch
- Lynn Vendinello, Office of Enforcement and Compliance Assurance, Office of Compliance, Information Utilization and Targeting Branch

Social Security Administration
(Phone Locator: 410/965-2982)
- Bob Marks, Office of Quality Assurance and Performance Assessment
- Paul Hallinger, Operations Division

Endnotes

1. *Washington Post,* "In the Loop," January 24, 2003, p. A3.

2. Sources include APHIS staff; Animal and Plant Health Inspection Service, "Moscamed Scientific Review Team Report," December 1998; Animal and Plant Health Inspection Service, "Fiscal 1999 Annual Program Performance Report"; Animal and Plant Health Inspection Service, "FY 2001 and FY 2002 Annual Performance Plans," March 9, 2001; and U.S. General Accounting Office, *Program Evaluation: Studies Helped Agencies Measure or Explain Program Performance* (GAO/GGD-00-204, September 2000).

3. Sources: U.S. Department of Education, Office of Vocational and Adult Education; and "National Reporting System for Adult Education: Implementation Guidelines," Division of Adult Education and Literacy, Office of Vocational and Adult Education, U. S. Department of Education, June 2000.

4. Sources: Office of Migrant Education, U.S. Department of Education, and excerpts from "Revised Department 2002 Budget HEP CAMP Nov 2002."

5. Sources include Rehabilitation Services Administration staff in Washington, D.C., and in regional offices in Boston, Dallas, San Francisco, and Seattle; U.S. Department of Education, Office of Special Education and Rehabilitative Services, Rehabilitation Services Administration, S*tate Vocational Rehabilitation Services Program: FY 2002 Monitoring and Technical Assistance Guide,* February 2002; Federal Register, Department of Education: 34 CFR Part 361, *The State Vocational Rehabilitation Services Program: Final Rule*. June 5, 2000; Ann Manheimer, *Using Performance Information to Manage Department of Education Grant Programs* (August 10, 2002 draft).

6. FY 1999 data were reported in this indicator format on a test basis. The new system went into effect with FY 2000 data.

7. The system also includes an indicator to assess the percentage of minorities served by the state agency.

8. Rehabilitation services provided to youth or younger adults are often classified in this category, since they enable the client to care for themselves at home, which enable a family member to obtain employment rather than care for the client.

9. The VR specialist usually reviews a random sample of files when conducting site visits to state agencies. In this case, a larger sample was drawn of cases closed without an employment outcome because the outcome indicator discussed indicated a problem might exist.

10. This policy revision does not preclude clients who need or prefer sheltered employment from that option. State VR agencies can refer such clients to state or local programs funded through other sources, such as HHS.

11. Sources include Office of Air and Radiation, Office of Air Quality Planning and Standards; www.epa.gov/cgi-bin: *Six Principal Pollutants; Nonattainment Areas; Latest Findings on National Air Quality: 2001 Status and Trends;* www.epa.gov/oar/oaqps/peg_caa05.html, *The Plain English Guide to the Clean Air Act;* and www.epa.gov/airlinks/airlinks2.html, *Regional Ozone Transport*.

12. Monitoring stations are operated by state, tribal, and local government agencies and some federal agencies, including EPA.

13. The type of entity designated as a nonattainment area may vary. It may be a city, an entire urban area, or one or more counties.

14. Additional locations have not been designated as nonattainment since the initial group was identified. The list is felt to contain the major areas that exceed standards.

15. Sources include OECA staff in the Office of Compliance and Office of Compliance Information

Utilization and Targeting Branch and Planning Branch, and in Region 7, and EPA *FY2001 EPA Annual Performance Report; EPA Annual Report on Enforcement and Compliance Assurance Accomplishments in 1999;* www.epa.gov/compliance/planning/results: *National Performance Measures Strategy—Final Report for Public Distribution,* December 1997; *Compliance and Enforcement Progress in FY 2001: Detailed Summary.*

16. Methods for calculating estimated pounds of pollutants reduced have been developed by OECA and vary by type of medium (such as air or water).

17. The priority-setting process also includes input from EPA regional offices, states and tribes, and stakeholders.

18. Source: Environmental Protection Agency, *Compliance and Enforcement Progress in FY 2001: Detailed Summary.*

19. EPA, Office of Compliance, OECA, *Using Performance Measurement Data as a Management Tool,* June 2002.

20. Sources include Office of Generic Drugs and U.S. Food and Drug Administration, *Performance Plan and Summary: FY 2002 Annual Performance Plan and Summary; FY 2001 Revised Final Performance Plan; FY 2000 Annual Performance Report,* April 2001.

21. Sources include Real Estate Assessment Center; TARC offices in Memphis and Cleveland; and PIH offices in New Jersey and Kentucky; www.hud.gov/offices/reac: "REAC at a Glance"; www.hud.gov/offices/pih/offices/otar: Frequently Asked Questions: Office of Troubled Agency Recovery.

22. All PHAs within a state are scored over the course of a year. Each PHA is scored annually and during the same quarter of the year.

23. This section draws heavily on "Assessing Physical Condition: An Overview of the Scoring Process," HUD Real Estate Assessment Center www.hud.gov/reac (May 28, 2001). See this document for a detailed discussion of the scoring system.

24. Each PHA is scored during the same quarter in successive years, so during the course of a year all PHAs are scored. The data presented here are for the first four quarters after PHAS scores were issued for New Jersey, a period that bridges two different fiscal years and the most recent available data covering the same PHAs (also bridging two fiscal years).

25. Sources include Wild Horse and Burro Program staff; www.wildhorseandburro.blm.gov, and Bureau of Land Management 2000-2005 Strategic Plan.

26. The frequency of vegetation monitoring varies in different areas, based on growth seasons or sensitivity of the area. Monitoring of vegetation loss is done annually at the end of the growing season in arid areas, as in the example provided here.

27. Sources include Jennifer L. Hoger, 2001 *Visitor Survey Card Workbook,* University of Idaho, Cooperative Park Studies Unit, January 2001; Visitor Survey Card Project; and officials from Point Reyes National Seashore and Shenandoah National Park.

28. Sources: Staff of OSHA's Office of Enforcement and Office of Statistical Analysis; OSHA Directives on Site-Specific Targeting 2002 and on Abatement Verification Regulation, 29CFR 1903.19; and Occupational Safety and Health Administration website: osha.gov.

29. Sources: Office of the Consumer Advocate USPS; Revenue and Volume Reporting, Corporate Financial Planning, USPS; "Origin-Destination Information System Quarterly Statistics Report," various Postal Quarters, USPS; and "United States Postal Service Five-Year Strategic Plan (FY 2001-2005)," USPS.

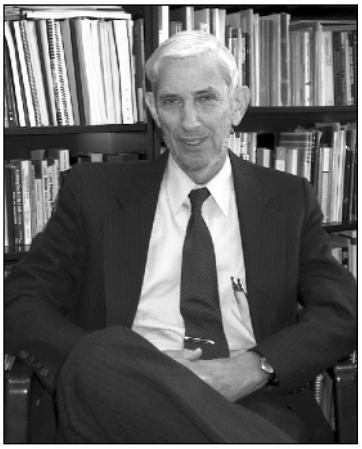
30. Sources include The Social Security Administration's June 2002 report "SSI Corrective Action Plan" SSA; and personnel of SSA's Office of Quality Assurance and Performance Assessment, and of the Operations Division.

31. "Executive Guide: Effectively Implementing the Government Performance and Results Act," GAO/GGD-96-118, June 1996, pp. 36-37.

32. Sources include "Root Cause Analysis of Quality Problems in Toll Free," Booz Allen & Hamilton, SSA Report 01-W&I-4Q-38, 9/28/2001; "EQ Impact Analysis of FY 02 Quality Improvement Actions," Booz Allen & Hamilton, SSA Report 02-W&I-3Q-22, 5/17/2002; and IRS staff in the Wage and Investment Customer Accounts Service, Customer Accounts Service, and Small Business and Self Employment.

33. Sources: Staff in the Office of Quality and Performance and Performance Analysis Services of VHA; U.S. Department of Veterans Affairs, *Departmental Performance Plan, FY 2003,* February 2002; U.S. Department of Veterans Affairs, *Annual Performance Report, FY 2001,* March 2002; U.S. General Accounting Office, *The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven* (GAO/GGD-97-109, June 1997); U.S. General Accounting Office, *Managing for Results: Emerging Benefits From Selected Agencies' Use of Performance Agreements* (GAO-01-115, October 2000); U.S. General Accounting Office, *Managing for Results: Efforts to Strengthen the Link Between Resources and Results at the Veterans Health Administration* (GAO-03-10, December 2002).

ABOUT THE AUTHORS



Harry P. Hatry is a Principal Research Associate and Director of the Public Management Program for the Urban Institute in Washington, D.C. He has been a leader in developing performance management/ measurement and evaluation procedures for public agencies since the early 1970s. He has provided assistance on Government Performance and Results Act–related activities to the U.S. Departments of Education, Justice, and Health and Human Services, and the Environmental Protection Agency's National Estuary Program.

He is a fellow of the National Academy of Public Administration. He has been a member of the U.S. Office of Management and Budget's Performance Measurement Advisory Council, the U.S. Department of Education's Evaluation Review (Advisory) Panel, the United Way of America's Task Force on Outcome Measurement, and the Independent Sector's Measures National Advisory Board.

Mr. Hatry received the 1985 Elmer B. Staats Award for Excellence in Program Evaluation and the 1984 American Society for Public Administration Award as the "outstanding contributor to the literature of management science and policy science." In 1993, he was a recipient of a National Public Service Award sponsored by the American Society for Public Administration (ASPA) and the National Academy of Public Administration (NAPA). In 1996, he received the "Evaluator of the Year" award from the Washington Evaluators Association. In 1999 the Center for Accountability and Performance of ASPA presented him with a lifetime achievement award for his work in performance measurement and established the "Harry Hatry Award for Distinguished Practice in Performance Measurement." In 2000, he was a recipient of the "50th Anniversary Einhorn-Gary" award from the Association of Government Accountants for "outstanding service to government financial professionals and sustained commitment to advancing government accountability."

Elaine Morley is a Senior Research Associate in the Public Management Program of the Metropolitan Housing and Communities Policy Center of the Urban Institute. She has had considerable experience with outcome and performance measurement of government agencies and nonprofit organizations. Recent publications in this area include *Making Use of Outcome Information for Improving Services: Recommendations for Nonprofit Organizations* (Urban Institute, 2002); *A Look at Outcome Measurement in Nonprofit Organizations* (The Independent Sector, 2001); and *Comparative Performance Measurement* (Urban Institute Press, 2001).

Dr. Morley has participated in numerous Urban Institute projects addressing public management and performance measurement. One recent effort involved developing outcome measurement procedures for State Boards of Nursing. She also participated in the Urban Institute's project with the International City/County Management Association's consortium to develop a comparative performance measurement process for various basic municipal services. Another project addressed identification of excellent management practices in community development agencies.

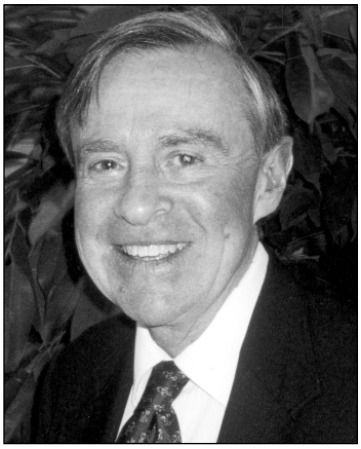
Prior to joining the Urban Institute in 1988, Dr. Morley was an associate professor of public administration and policy analysis at Southern Illinois University at Edwardsville, where her work included public-sector productivity and public-service delivery. She holds a Ph.D. in social science from Syracuse University.

Shelli B. Rossman is a Senior Research Associate in the Justice Policy Center at the Urban Institute. Ms. Rossman has more than 20 years of research and management experience on projects for federal, state, and local governments, as well as private-sector clients in the areas of criminal justice, public health, and safety. She has expertise in the design and conduct of public policy research, program evaluation, and performance measurement. Ms. Rossman has conducted training on performance management for the Thailand Office of Civil Service Commission; the Asian Development Bank Pacific Public Management program; the Foundation for Advanced Studies on International Development in Tokyo, Japan; and the World Bank Institute in its efforts to introduce performance measurement to several ministries in the Republic of Kyrgyzstan. She is co-author, with Mr. Hatry, of *Timely, Cost-Effective, and Useful Outcome Data* (NAPA, June 2000).

Ms. Rossman has conducted several performance-monitoring demonstration projects in the United States and in international settings. Currently, she is working on a U.S. Agency for International Development project in Russia that is implementing performance measurement in two municipal government settings: one focused on juvenile justice and the other on neighborhood services. Ms. Rossman also served as project director for the second phase of technical assistance to the Civil Service Commission of Thailand; this effort focused on piloting three performance-monitoring projects for the Ministry of Industry. Earlier, she served as senior staff for the Phase 1, which provided assistance to four projects within the Departments of Local Administration and Agriculture Extension.

Ms. Rossman has provided technical assistance to several U.S. federal agencies, including the Civil Rights Division of the Department of Justice; the National Science Foundation's Program for Persons with Disabilities; and the Office of Elementary and Secondary Education and the Office of Planning and Evaluation Services in the Department of Education—in their efforts to respond to GPRA requirements and to develop internal capacity to improve management decision making through reliance on performance

indicators. She recently acted as senior staff on a Centers for Disease Control project to develop performance indicators and guidance in using performance measures to track the outcomes of community-based organizations' HIV prevention services. For the Lila Wallace Foundation Urban Parks Initiative, she consulted on the development of a manual to assist park administrators in using outcome evaluation to inform marketing, maintenance, and other management decision making. Earlier, she was a task leader for the Urban Institute's project with the International City/County Management Association consortium to develop cross-jurisdictional comparative performance measures in multiple municipal service areas, including police, fire, parks and recreation, libraries, solid waste collection, and internal support services. She also served as a senior member of the Urban Institute's team designing and testing performance measurement procedures for the Office of Watersheds, Oceans, and Wetlands, Division of Oceans and Coastal Protection, Environmental Protection Agency.



Joseph S. Wholey is Professor of Public Administration at the University of Southern California, where his work focuses on the use of strategic planning, performance measurement, and program evaluation to improve government performance and accountability. Previously, he served as senior advisor for performance and accountability at the U.S. General Accounting Office, and as senior advisor to the deputy director for Management at the U.S. Office of Management and Budget. He is senior author or editor of eight books, including *Evaluation and Effective Public Management, Performance and Credibility, Improving Government Performance,* and the *Handbook of Practical Program Evaluation*.

Dr. Wholey served as deputy assistant secretary for planning and evaluation at the U.S. Department of Health and Human Services, and as director of program evaluation studies at the Urban Institute. He held elective office for eight years as a member of the Arlington (Virginia) County Board, serving as chairman for three years. He was chairman of the Washington Metropolitan Area Transit Authority, chairman of the Virginia Board of Social Services, president of Hospice of Northern Virginia, and chairman of the Arlington Partnership for Affordable Housing.

Dr. Wholey is a fellow of the National Academy of Public Administration. He holds a B.A. in mathematics from Catholic University, and an M.A. in mathematics and a Ph.D. in philosophy from Harvard University.

ABOUT THE ACADEMY

The National Academy of Public Administration is an independent, nonpartisan organization chartered by Congress to assist federal, state, and local governments in improving their effectiveness, efficiency, and accountability. For more than 30 years, the Academy has met the challenge of cultivating excellence in public management and administration.

Federal agencies, Congress, state and local governments, and education and philanthropic institutions frequently seek the Academy's assistance in addressing both short-term and long-term challenges—including budgeting and finance, alternative corporate structures, performance measurement, human resources management, information technology, devolution, strategic planning, and managing for results.

The National Academy's Performance Consortium is a membership-based organization made up of 26 federal offices and functions that have joined together under the auspices of The National Academy of Public Administration to support an annual program of peer-to-peer exchange of practices relating to performance-based government through practice papers, workshops, discussion forums, seminars, and conferences.

For additional information, contact:

Philip Burgess
President
National Academy of Public Administration
1100 New York Avenue NW, Suite 1090 East
Washington, DC 20005
(202) 347-3190
fax: (202) 393-0993

e-mail: academy@napawash.org
website: www.napawash.org

KEY CONTACT INFORMATION

To contact the authors:

Harry Hatry
Principal Research Associate
The Urban Institute
2100 M Street, NW
Washington, DC 20037
(202) 261-5521

e-mail: hhatry@ui.urban.org

Elaine Morley
Senior Research Associate
The Urban Institute
2100 M Street, NW
Washington, DC 20037
(202) 261-5629

e-mail: emorley@ui.urban.org

Shelli Rossman
Senior Research Associate
The Urban Institute
2100 M Street, NW
Washington, DC 20037
(202) 261-5525

e-mail: srossman@ui.urban.org

Joseph Wholey
Professor, Public Administration
University of Southern California
2103 21st Rd North
Arlington, VA 22201
(703) 524-0022

e-mail: wholey@usc.edu

GRANT REPORTS

E-Government

Supercharging the Employment Agency: An Investigation of the Use of Information and Communication Technology to Improve the Service of State Employment Agencies (December 2000)

Anthony M. Townsend

Assessing a State's Readiness for Global Electronic Commerce: Lessons from the Ohio Experience (January 2001)

J. Pari Sabety
Steven I. Gordon

Privacy Strategies for Electronic Government (January 2001)

Janine S. Hiller
France Bélanger

Commerce Comes to Government on the Desktop: E-Commerce Applications in the Public Sector (February 2001)

Genie N. L. Stowers

The Use of the Internet in Government Service Delivery (February 2001)

Steven Cohen
William Eimicke

State Web Portals: Delivering and Financing E-Service (January 2002)

Diana Burley Gant
Jon P. Gant
Craig L. Johnson

Internet Voting: Bringing Elections to the Desktop (February 2002)

Robert S. Done

Leveraging Technology in the Service of Diplomacy: Innovation in the Department of State (March 2002)

Barry Fulton

Federal Intranet Work Sites: An Interim Assessment (June 2002)

Julianne G. Mahler
Priscilla M. Regan

The State of Federal Websites: The Pursuit of Excellence (August 2002)

Genie N. L. Stowers

State Government E-Procurement in the Information Age: Issues, Practices, and Trends (September 2002)

M. Jae Moon

Preparing for Wireless and Mobile Technologies in Government (October 2002)

Ai-Mei Chang
P. K. Kannan

Public-Sector Information Security: A Call to Action for Public-Sector CIOs (October 2002, 2nd ed.)

Don Heiman

The Auction Model: How the Public Sector Can Leverage the Power of E-Commerce Through Dynamic Pricing (November 2002, 2nd ed.)

David C. Wyld

The Promise of E-Learning in Africa: The Potential for Public-Private Partnerships (January 2003)

Norman LaRocque
Michael Latham

Digitally Integrating the Government Supply Chain: E-Procurement, E-Finance, and E-Logistics (February 2003)

Jacques S. Gansler
William Lucyshyn
Kimberly M. Ross

Using Technology to Increase Citizen Participation in Government: The Use of Models and Simulation (April 2003)

John O'Looney

Financial Management

Credit Scoring and Loan Scoring: Tools for Improved Management of Federal Credit Programs (July 1999)

Thomas H. Stanton

Using Activity-Based Costing to Manage More Effectively (January 2000)

Michael H. Granof
David E. Platt
Igor Vaysman

Audited Financial Statements: Getting and Sustaining "Clean" Opinions (July 2001)

Douglas A. Brook

An Introduction to Financial Risk Management in Government (August 2001)

Richard J. Buttimer, Jr.

Human Capital

Profiles in Excellence: Conversations with the Best of America's Career Executive Service (November 1999)

Mark W. Huddleston

Reflections on Mobility: Case Studies of Six Federal Executives (May 2000)

Michael D. Serlin

Managing Telecommuting in the Federal Government: An Interim Report (June 2000)

Gina Vega
Louis Brennan

Using Virtual Teams to Manage Complex Projects: A Case Study of the Radioactive Waste Management Project (August 2000)

Samuel M. DeMarie

A Learning-Based Approach to Leading Change (December 2000)

Barry Sugarman

Labor-Management Partnerships: A New Approach to Collaborative Management (July 2001)

Barry Rubin
Richard Rubin

Winning the Best and Brightest: Increasing the Attraction of Public Service (July 2001)

Carol Chetkovich

Organizations Growing Leaders: Best Practices and Principles in the Public Service (December 2001)

Ray Blunt

A Weapon in the War for Talent: Using Special Authorities to Recruit Crucial Personnel (December 2001)

Hal G. Rainey

A Changing Workforce: Understanding Diversity Programs in the Federal Government (December 2001)

Katherine C. Naff
J. Edward Kellough

Life after Civil Service Reform: The Texas, Georgia, and Florida Experiences (October 2002)

Jonathan Walters

Leaders Growing Leaders: Preparing the Next Generation of Public Service Executives (November 2002, 3rd ed.)

Ray Blunt

The Defense Leadership and Management Program: Taking Career Development Seriously (December 2002)

Joseph A. Ferrara
Mark C. Rom

The Influence of Organizational Commitment on Officer Retention: A 12-Year Study of U.S. Army Officers (December 2002)

Stephanie C. Payne
Ann H. Huffman
Trueman R. Tremble, Jr.

Human Capital Reform: 21st Century Requirements for the United States Agency for International Development (March 2003)

Anthony C. E. Quainton
Amanda M. Fulmer

Modernizing Human Resource Management in the Federal Government: The IRS Model (April 2003)

James R. Thompson
Hal G. Rainey

Managing for Results

Corporate Strategic Planning in Government: Lessons from the United States Air Force (November 2000)

Colin Campbell

Using Evaluation to Support Performance Management: A Guide for Federal Executives (January 2001)

Kathryn Newcomer
Mary Ann Scheirer

Managing for Outcomes: Milestone Contracting in Oklahoma (January 2001)

Peter Frumkin

The Challenge of Developing Cross-Agency Measures: A Case Study of the Office of National Drug Control Policy (August 2001)

Patrick J. Murphy
John Carnevale

The Potential of the Government Performance and Results Act as a Tool to Manage Third-Party Government (August 2001)

David G. Frederickson

Using Performance Data for Accountability: The New York City Police Department's CompStat Model of Police Management (August 2001)

Paul E. O'Connell

Moving Toward More Capable Government: A Guide to Organizational Design (June 2002)

Thomas H. Stanton

Performance Management: A "Start Where You Are, Use What You Have" Guide (October 2002)

Chris Wye

The Baltimore CitiStat Program: Performance and Accountability (May 2003)

Lenneal J. Henderson

How Federal Programs Use Outcome Information: Opportunities for Federal Managers (May 2003)

Harry P. Hatry
Elaine Morley
Shelli B. Rossman
Joseph S. Wholey

New Ways to Manage Innovation

Managing Workfare: The Case of the Work Experience Program in the New York City Parks Department (June 1999)

Steven Cohen

New Tools for Improving Government Regulation: An Assessment of Emissions Trading and Other Market-Based Regulatory Tools (October 1999)

Gary C. Bryner

Religious Organizations, Anti-Poverty Relief, and Charitable Choice: A Feasibility Study of Faith-Based Welfare Reform in Mississippi (November 1999)

John P. Bartkowski
Helen A. Regis

Business Improvement Districts and Innovative Service Delivery (November 1999)

Jerry Mitchell

An Assessment of Brownfield Redevelopment Policies: The Michigan Experience (November 1999)

Richard C. Hula

San Diego County's Innovation Program: Using Competition and a Whole Lot More to Improve Public Services (January 2000)

William B. Eimicke

Innovation in the Administration of Public Airports (March 2000)

Scott E. Tarry

Entrepreneurial Government: Bureaucrats as Businesspeople (May 2000)

Anne Laurent

Rethinking U.S. Environmental Protection Policy: Management Challenges for a New Administration (November 2000)

Dennis A. Rondinelli

The Challenge of Innovating in Government (February 2001)

Sandford Borins

Understanding Innovation: What Inspires It? What Makes It Successful? (December 2001)

Jonathan Walters

Government Management of Information Mega-Technology: Lessons from the Internal Revenue Service's Tax Systems Modernization (March 2002)

Barry Bozeman

Procurement

Determining a Level Playing Field for Public-Private Competition (November 1999)

Lawrence L. Martin

Implementing State Contracts for Social Services: An Assessment of the Kansas Experience (May 2000)

Jocelyn M. Johnston
Barbara S. Romzek

A Vision of the Government as a World-Class Buyer: Major Procurement Issues for the Coming Decade (January 2002)

Jacques S. Gansler

Contracting for the 21st Century: A Partnership Model (January 2002)

Wendell C. Lawther

Franchise Funds in the Federal Government: Ending the Monopoly in Service Provision (February 2002)

John J. Callahan

Making Performance-Based Contracting Perform: What the Federal Government Can Learn from State and Local Governments (November 2002, 2nd ed.)

Lawrence L. Martin

Moving to Public-Private Partnerships: Learning from Experience around the World (February 2003)

Trefor P. Williams

IT Outsourcing: A Primer for Public Managers (February 2003)

Yu-Che Chen
James Perry

The Procurement Partnership Model: Moving to a Team-Based Approach (February 2003)

Kathryn G. Denhardt

Networks, Collaboration, and Partnerships

Leveraging Networks to Meet National Goals: FEMA and the Safe Construction Networks (March 2002)

William L. Waugh, Jr.

21st-Century Government and the Challenge of Homeland Defense (June 2002)

Elaine C. Kamarck

Assessing Partnerships: New Forms of Collaboration (March 2003)

Robert Klitgaard
Gregory F. Treverton

Leveraging Networks: A Guide for Public Managers Working across Organizations (March 2003)

Robert Agranoff

Extraordinary Results on National Goals: Networks and Partnerships in the Bureau of Primary Health Care's 100%/0 Campaign (March 2003)

John Scanlon

Public-Private Strategic Partnerships: The U.S. Postal Service-Federal Express Alliance (May 2003)

Oded Shenkar

Transforming Organizations

The Importance of Leadership: The Role of School Principals (September 1999)

Paul Teske
Mark Schneider

Leadership for Change: Case Studies in American Local Government (September 1999)

Robert B. Denhardt
Janet Vinzant Denhardt

Managing Decentralized Departments: The Case of the U.S. Department of Health and Human Services (October 1999)

Beryl A. Radin

Transforming Government: The Renewal and Revitalization of the Federal Emergency Management Agency (April 2000)

R. Steven Daniels
Carolyn L. Clark-Daniels

Transforming Government: Creating the New Defense Procurement System (April 2000)

Kimberly A. Harokopus

Trans-Atlantic Experiences in Health Reform: The United Kingdom's National Health Service and the United States Veterans Health Administration (May 2000)

Marilyn A. DeLuca

Transforming Government: The Revitalization of the Veterans Health Administration (June 2000)

Gary J. Young

The Challenge of Managing Across Boundaries: The Case of the Office of the Secretary in the U.S. Department of Health and Human Services (November 2000)

Beryl A. Radin

Creating a Culture of Innovation: 10 Lessons from America's Best Run City (January 2001)

Janet Vinzant Denhardt
Robert B. Denhardt

Transforming Government: Dan Goldin and the Remaking of NASA (March 2001)

W. Henry Lambright

Managing Across Boundaries: A Case Study of Dr. Helene Gayle and the AIDS Epidemic (January 2002)

Norma M. Riccucci

Managing "Big Science": A Case Study of the Human Genome Project (March 2002)

W. Henry Lambright

The Power of Frontline Workers in Transforming Government: The Upstate New York Veterans Healthcare Network (April 2003)

Timothy J. Hoff

SPECIAL REPORTS

Government in the 21st Century

David M. Walker

Results of the Government Leadership Survey: A 1999 Survey of Federal Executives (June 1999)

Mark A. Abramson
Steven A. Clyburn
Elizabeth Mercier

Creating a Government for the 21st Century (March 2000)

Stephen Goldsmith

The President's Management Council: An Important Management Innovation (December 2000)

Margaret L. Yao

Toward a 21st Century Public Service: Reports from Four Forums (January 2001)

Mark A. Abramson, Editor

Becoming an Effective Political Executive: 7 Lessons from Experienced Appointees (January 2001)

Judith E. Michaels

The Changing Role of Government: Implications for Managing in a New World (December 2001)

David Halberstam

BOOKS*

E-Government 2001
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and Grady E. Means, editors

E-Government 2003
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and Therese L. Morin, editors

Human Capital 2002
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and Nicole Willenz Gardner, editors

Innovation
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and Ian Littman, editors

Leaders
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and Kevin M. Bacon, editors

Managing for Results 2002
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and John Kamensky, editors

Memos to the President: Management Advice from the Nation's Top Public Administrators (Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson, editor

The Procurement Revolution
(Rowman & Littlefield Publishers, Inc., 2003)

Mark A. Abramson and Roland S. Harris III, editors

Transforming Organizations
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and Paul R. Lawrence, editors

* Available at bookstores, online booksellers, and from the publisher (www.rowmanlittlefield.com or 800-462-6420).

About IBM Business Consulting Services

With more than 60,000 consultants and professional staff in more than 160 countries globally, IBM Business Consulting Services is the world's largest consulting services organization. IBM Business Consulting Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build and run those solutions in a way that delivers bottom-line business value.

About the Endowment

Through grants for research, the IBM Endowment for The Business of Government stimulates research and facilitates discussion on new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

Founded in 1998, the Endowment is one of the ways that IBM seeks to advance knowledge on how to improve public sector effectiveness. The IBM Endowment focuses on the future of the operation and management of the public sector.

For additional information, contact:
Mark A. Abramson
Executive Director
IBM Endowment for The Business of Government
1616 North Fort Myer Drive
Arlington, VA 22209
(703) 741-1077, fax: (703) 741-1076

e-mail: endowment@businessofgovernment.org
website: www.businessofgovernment.org

IBM Endowment for
The Business of Government

1616 North Fort Myer Drive
Arlington, VA 22209-3195

PRST STD
US Postage
P A I D
Permit 1112
Merrifield, VA