Using Outcome Information to Redirect Programs

A Case Study of the Coast Guard's Pilot Project Under the Government Performance and Results Act

---

Rick Kowalewski
Commander, USCG

UNITED STATES COAST GUARD

OFFICE OF MARINE SAFETY, SECURITY
AND ENVIRONMENTAL PROTECTION

April 1996

The kinds of programs and agencies *that may find this case study relevant include:*

- *regulatory programs*
- *operational programs*
- *safety and security programs*
- *environmental programs*
- *transportation programs*
- *agencies with multiple missions and diverse activities*

Acknowledgments:

This case study was prepared with much valuable information, support, review, and advice from Coast Guard colleagues. Special thanks to Ed Ziff, Commander Chris Boegel, Lieutenant Commander Jean Butler, and Lieutenant Wyman Briggs -- who contributed many insightful comments as practitioners in the planning and evaluation effort, apart from their major contributions to the substance which the case study attempts to convey. Thanks also to George McDonald in the Office of the Secretary of Transportation, for constructively challenging the thought process and bringing together all the pilot projects in the Department for collaborative, "external" review of our work.

The American Society for Public Administration (ASPA) Government Accomplishment and Accountability Task Force provided the key trigger for this study, and the Office of Management and Budget (OMB) provided the very helpful outline to organize the thought process and presentation. Members of the National Academy of Public Administration (NAPA) provided continuing encouragement and feedback over the last two years to help hone the plans and to help better discern which parts of the Coast Guard's development process were important.

OVERVIEW

One of the central -- and most difficult -- issues presented by the Government Performance and Results Act is the shift in focus from inputs, activities, and outputs ... to *outcomes.* In retrospect, it seems only natural that this would provoke stress in the organization. Consider what outcomes are: they are not what you do as an organization, they are the real-world effects of what you do. They are, by definition, things you don't control. In our experience and from our observations of other agencies, managers are often reluctant to risk accountability for things outside their control. Outcomes also commonly cut across organizational lines. A variety of programs will influence the outcomes in areas like safety and environmental protection (and probably any of the more important outcomes of interest to the public), and this fact tends to blur traditional organizational boundaries. Maybe more significantly, the search for outcome-oriented goals is a process of examining the very basis for existence of your programs. Measurement in this environment can be particularly threatening.

In the U.S. Coast Guard GPRA pilot**,** planning and management are centered around a Business Plan with nine outcome-oriented performance goals, historical baseline measures, and broad strategies for achieving our goals. Goals are formulated by senior leadership and reviewed annually, and they are focused on affecting things the public sees, like deaths, injuries, and environmental damage. Most of the Coast Guard's goals can be demonstrated to be a "reach" in comparison with the historical trend of data, and all except one were met or exceeded in the first year of the pilot.

Tangible managerial flexibility has been the underpinning of our management approach. Since January 1994, the programs have rescinded or relaxed performance standards for all marine safety program activities not required by law, regulation, or national consistency. By substituting field discretion for old, internal specifications on how the work would be done, managerial flexibility has been increased for over one-half million work hours annually -- equal to about 8½% of program FTE. It was a fundamental principle from the outset that we couldn't hold managers accountable, in any sense, for results, if they didn't have flexibility to achieve those results.

By disaggregating high-level measures like worker fatalities and passenger-vessel casualties, we have begun using trend and risk information to redirect activities and resources toward achieving our goals. We found several surprises, including some risks we had underestimated, and some program successes we didn't realize we could claim.

We continue to grapple with the process of making this all work. Some of the challenges are persistent. But our plan gives us a better basis for dialog with our stakeholders (who share an interest in outcomes), more clear communication with our

workforce, and a better basis for making tradeoffs among programs and resources. We believe we can better support public funding for our programs. And, of course, by focusing on results, there is a greater likelihood that we'll achieve something worthwhile.

THE CONTEXT FOR OUR PLANNING

The U.S. Coast Guard is one of ten operating administrations in the Department of Transportation. Within the Coast Guard, the GPRA pilot encompasses a variety of prevention and response programs, with annual spending of about $460 million and program staffing of about 3400 -- about one-eighth of the Coast Guard totals.

- We regulate the design, construction, equipment, manning, operations and maintenance of commercial vessels, and the qualifications of their crew;
- establish and enforce standards for handling hazardous materials onboard vessels and waterfront terminals;
- negotiate international standards; and
- direct response activities to mitigate the effects of accidents or pollution.

Historically, **our program operations** have been directed through nationwide performance standards governing the scope and frequency for hundreds of activities like tankship inspections, oil transfer monitoring, and harbor patrols. National goals expressed by Congress in legislation were translated by Coast Guard Headquarters into fairly detailed program design. Funding and staffing were allocated on the basis of these standards, and activities were reported and monitored regularly.

Over the years, a number of studies were undertaken to optimize our activity levels on a nationwide basis. But we were never quite current. Our changing operating environment quickly outpaced these broad studies, and a changing mix of programs and work processes altered the equations. More importantly, local circumstances were commonly believed to be a seriously underestimated factor in risk.

Our current strategic planning effort can probably be traced to an organizational assessment in early 1992, including surveys of stakeholders, field offices, and program personnel at districts and Headquarters. The feedback we got clearly signalled the need for change in the way we do business. One broad, recurring theme was especially evident: centralized management was undercutting field focus on the purpose of their activities and limiting their ability to contribute toward the outcomes we really care about.

"Explain what the marine safety programs are trying to accomplish."
"Please set priorities for our mission, every operation can't be #1 priority."
"The responsibility for end products should be explained the first day on the job."
"Give us a goal that is challenging."

Some of the 5000 comments from the field.

Our planning process was driven from the outset by a very senior manager who had an appreciation for the value of planning, and was in a position to direct/influence the senior managers in the organization. A new strategic planning branch, with a total staff of four, provided research/analysis, planning facilitation, and development support.

We experimented with a number of planning techniques over the next year, including the use of scenarios, critical issues, value-adds, environmental assessment, and quality management. Each provided a piece of the puzzle, but none provided the comprehensive view that we sought. At our first planning retreat we developed seven critical management issues -- things like personnel management, data analysis, and stakeholder linkages -- each with specific problems we needed to resolve to assure the long-term success of the organization. Strategies were developed. But after 2½ days, there was a lingering sense that we had missed something. There were no issues that dealt directly with *safety* or *environmental protection*, our "bread-and-butter." We continued experimenting.

One thrust was a concentration on program data, centered on an effort to develop a comprehensive Data Matrix of program activities. We are a fairly data-rich organization. But we had a complex collection of systems for capturing and analyzing data, with the result that even seemingly identical questions often yielded inconsistent answers. In one instance, we found we had given three different numbers to Congress in reporting the number of oil spills which had occurred in one recent year. The Data Matrix was intended to give us a single corporate reference for management. While a significant effort, this did not address directly the question of what we *ought* to be measuring. It was more an attempt to get people to start using the data we had.

"Value-add" planning moved us in the right direction. In this process, we identified what we do, why we do it, who else does related things and how they relate, and how we might measure the special value our organization adds to the final result. Over several months, we held workshops for each of our ten program divisions. This began to focus on things that were central to our mission, but was somewhat encumbered by the organizational boundaries between divisions. We found a good deal of overlap

among the programs and their purposes. This was to become raw material for the next planning retreat.

HOW WE DEVELOPED OUR INDICATORS

Implementing GPRA can easily get bogged down by uncertainties and distractions. Managers may be ambivalent about setting goals ... more comfortable with immediate problem-solving ... reluctant to risk accountability for things outside their control ... or reluctant to commit at all when priorities are always changing. Organizations which were set up to optimize activities may be poorly aligned to address outcomes. The budget linkage is unclear. Data collection may be misdirected or incomplete, and evaluators can get mired in the pursuit of unattainable precision. In the end, performance information can be overwhelmed by other factors in decision making.

There are also a number of risks in trying to measure outcomes. You may find that you can't easily measure the outcomes. You may find that you can measure them, but that the measures aren't credible. You may find that some of your programs aren't actually achieving results, or that the results are not worth the costs. Each of these possibilities pose reasons not to start, and some incentive for "gaming."

Our aim was a Business Plan that could be used to guide priorities and drive decisions at every level of the organization. The motivation was to manage better. There was very general direction from higher levels of the Coast Guard for each Program Director to develop a Business Plan, but its form, content and schedule were undefined. There was no clear linkage to the budget process (at least at that point), and no real external drivers. The Government Performance and Results Act didn't yet exist; it was a bill (S.20) in the Senate.

The Board of Directors -- the senior program managers and commanding officers of key Headquarters commands, led by the Deputy Program Director -- assembled for the second planning retreat to outline the plan. Program managers were asked to leave their program hats at the door, and to participate with a broader perspective as a group of the most experienced officers in the general field of marine safety. In two days, the group:

- defined the mission of the organization, and drafted two strategic goals
- formulated several preliminary performance goals in each strategic area
- identified five areas of business focus as prime targets for reinvesting resources
- began associating major program initiatives with program goals
- recast the previous year's "critical issues" into "guiding management principles"

- decided to rescind or relax mission performance standards which were not mandated by law, treaty, or Federal regulation, or essential for national consistency

The planning framework was cobbled together after the first several hours of somewhat disjointed deliberations, to provide a sense of context for the pieces of the emerging plan. The framework included elements which flowed from *mission → strategic goals → performance goals → business focus → major initiatives and detailed actions*. A time horizon and organizational responsibility were associated with each element. And informing all of these were our "guiding management principles," the key strategies we needed to apply to assure the long term health and success of the organization.

A statement of our collective mission didn't exist before this effort. Working from some draft ideas and a preliminary set of criteria prepared before the retreat, the group agreed on a simple, one sentence statement of our purpose, in words that could pass the "next door neighbor" test for understandability. The mission is intended to be enduring.

> *Our mission is to protect the public, the environment, and U.S. economic interests, by preventing and mitigating marine accidents.*

Strategic goals were to follow from the mission, and aim at what we are trying to accomplish in the world, in the very long term. One strategic goal covered marine safety; another addressed environmental protection. Both cut across the existing organization, as multiple programs are needed to achieve results in either area. This is the first point at which the group began to feel discomfort associated with the cross-organizational nature of outcomes. Many expected to see strategic goals in each of our traditional program areas, like maritime licensing or port security. Very likely many wanted goals they could own and manage individually as program managers. The outcome focus upset the equilibrium.

Formulating performance goals was probably the most discomforting part of the process. These were to be specific, short-term, measurable, and usable in sorting program changes and priorities. They would collectively define success. We were looking here for things we don't control, but which we can influence substantially. We found it helpful to avoid the term "goal" (a loaded word) and ask instead "what are we trying to achieve?" We also found it useful to continually ask "why?" Why do we do this function, or why do we care about achieving this result? Completing a rulemaking project is important because it can change behavior, which can reduce accidents; accidents are important (primarily) because people can get hurt or killed. The answers to "why?" typically led us closer to the outcomes of real interest.

We excluded a number of small, minor programs in this process. Our intent was to identify a set of goals that guided priorities and essentially characterized the organization. It was important to us that our major goals not be diluted by many trivial ones.

With no historical baseline data at this point, target levels for our goals were set intuitively at first, using the professional judgment of our senior managers. Our aim was to set targets which were realistic but challenging. We wanted to force people to rethink how they did business, rather than to respond with just a little more effort. Five-year targets helped clarify that our goals would be a "reach," and also helped capture the value of many longer-term things we do, like R&D and standards-development.

Before our first Business Plan was published, we explored strategies for achieving our goals and identified five areas of business focus -- areas where our activities and resource investment were believed to be disproportionate to safety or environmental risk. We brought in our senior marine safety officers from all ten districts, asked them if they could achieve these goals, how they would go about it, and what they needed to accomplish it. They validated the goals as reasonably attainable; outlined their general strategies; and asked for *data*, especially, to help them manage risk.

Data analysis presented a surprisingly contentious issue. For several months, agreement could not be reached on whether the data analysis function ought to be distributed or centralized in the organization, or, if centralized, what part of the organization should be responsible for it. There were basic differences about what analyses needed to be done, and what needed to be done first. After lengthy discussion, one manager observed that "we seem to be like ten flies on an elephant, trying to describe the elephant."

A focused, short-term effort to *demonstrate* the development of performance measures was proposed to break the logjam. We created an ad-hoc Program Evaluation Group of "experts" -- with academic backgrounds and experience in economics, operations research, and policy analysis -- from throughout the organization. This group substantially advanced the sophistication of our baseline measures, began to disaggregate the data to identify important risk factors for targeting, outlined a range of external factors potentially affecting achievement of our goals, and identified several possible unintended effects of our programs for further study. The first major effort of the group was completed in six weeks.

Stakeholders were not directly involved in our planning process up-front. We used (implicitly) the results of stakeholder surveys from our organizational assessment, and the broad experience of our senior managers in dealing with the maritime industry,

states, other agencies, and Congress over the years. But it was not until our Business Plan was developed and published that we actively engaged our stakeholders in this process.

THE OUTCOME INDICATORS THAT WE USE

Our Strategic Goal for Marine Safety is to eliminate deaths, injuries and economic loss associated with commercial marine transportation and the shipment of military cargo. Over five years, our programs aim to achieve these major goals:

Performance Goals	**Performance Indicators**
Reduce accidental deaths and injuries from maritime casualties by 20%.	1) Worker fatalities per 100,000 workers 2) Non-worker fatalities/billion tons of commerce
Reduce the risk of passenger vessel casualty with major loss of life by 20%.	Vessel accidents per thousand passenger vessels
Reduce fatality rates aboard uninspected fishing and towing vessels halfway toward the average of the U.S. inspected fleet.	1) Fishing fatalities per 100,000 workers 2) Towing fatalities per 100,000 workers
Eliminate substandard commercial vessels from U.S. waters.	Interventions per 100,000 port calls
Reduce the vulnerability of U.S. ports and waterways to intentional damage/injury.	Compliance rate with IMO guidelines (preliminary indicator, still under development)

Our Strategic Goal for Marine Environmental Protection is to eliminate environmental damage associated with marine transportation. Over five years, our programs aim to achieve these major goals in marine environmental protection:

Performance Goals	**Performance Indicators**
Reduce the amount of oil and chemicals going into the water from maritime sources by 20%.	Gallons spilled per million gallons shipped
Reduce the discharge of plastics/garbage going into the water from maritime sources by 20%.	Pounds of marine debris recovered per mile of shoreline cleaned, in annual coastal cleanups
Reduce the total number of major and	medium spills by 50%.

Number of major/medium spills per billion tons shipped	Increase the removal of spilled oil from the water by 10%.
Gallons removed per 100 gallons spilled	

Each goal is represented with a performance baseline, including a graph of the historical trend, an itemization of what's included or excluded in the measure, tabulated data with source references, secondary (or disaggregated) measures to show significant components of the primary indicator, and notes on unresolved problems or issues with the data. Several examples are included with this case study in Attachment A.

We began with more simple measures, and refined them substantially over about six months. Our first performance goal -- to *"reduce deaths and injuries from maritime casualties by 20%"* -- provides a useful illustration of the process:

> *Our first performance plan* included a graph showing ten years of data from our Marine Safety Information System (MSIS). The graph plotted the number of deaths and injuries by year. This was raw data at this point, not scrubbed for data quality, and not normalized for changes in exposure over time. It was obviously a somewhat simplistic measure. But it reflected the state of our measurement capability at the time.
>
> *We progressively refined the scope* of our measures by clarifying what should be included and excluded. For our goal related to deaths and injuries, we began to focus on deaths as a primary indicator, in part because of problems with the reliability of injury reporting. Less serious injuries, in particular, were substantially underreported, and we had no data to discriminate the severity of injuries. Deaths also represented the most important outcome, and in some respects could be used as a proxy indicator for safety generally. We also more narrowly focused on accidental deaths, excluding homicides, suicides, and natural deaths which were sometimes included in the data (later adding the qualifier "accidental" in our goal statement, for clarity).
>
> *To normalize the data for changes in exposure*, we decided (after considerable study) to use the number of maritime workers as a denominator for our measure. This didn't address all of the fatalities under our purview, but it did address most of them, and it provided a common measure of exposure for this large group. It also provided a measure which could be compared more broadly to published measures for safety in other industries. It meant more narrowly focusing the indicator for the goal, and it also meant considerable research into non-Coast Guard data sources for measures of the maritime worker population. Eventually, we put together data from six different data sources to give us, for the first time, a

composite number of workers on U.S. vessels. Our measure was now expressed as a rate -- fatalities per 100,000 workers.

On more detailed examination of the data, we found a number of duplicate entries, miscoded data, and some simply erroneous information. We used data base tools, comparisons of data across different data bases, and even manual examination of case files, to help identify possible data problems. We found a suspicious drop in fatalities in 1990 and 1991, corresponding to the time we made a major data base conversion and change in field reporting. It was only by comparison with paper files at field offices that we were able to nail down the problem and correct the numbers.

We still had substantial fluctuation in the data from year to year -- in fact, the "noise" exceeded the "signal," or increment, that we were looking for in measuring success against our goal. The immediate problem was deciding where to set the baseline fatality rate from which we would measure. We were most interested in changes in the real, underlying risk -- apart from annual fluctuations. And we had no evidence that the fatality rate for 1993 would be anything more than coincidentally representative. We experimented with 3-, 4-, and 5-year moving averages (a simple technique for "smoothing" the data, which we used in our second performance plan), and later with logarithmic regression (a more sophisticated technique used in our third performance plan, and still used). This ultimately gave us a trendline which established a statistical reference point for our goal, and also gave us a better sense of the real trend over time.

Projecting the trendline over the next five years, we could see -- and demonstrate -- that our five-year goal of reducing fatalities by 20% was in fact a "reach," or more than what we would expect if we just continued the past trend of improvements over time. For comparison, we plotted the goal as well, assuming equal increments of change to reach our five-year goal, and these points finally provided our specific targets for each year. [The complete baseline for this goal is provided in the attachment.]

Further analysis of the data resulted in a variety of secondary indicators for our performance goals. By disaggregating the data, we began identifying high risk areas for targeting, and this provided the basis for the most important use of our performance measures -- making real changes in safety and environmental protection. This is illustrated in detail in the section on "How We Use Information on Outcomes."

Not all of our goals or measures are equally strong or outcome-oriented. Our goal for fishing vessel safety, for example, can be viewed as a strategy for achieving

our first goal of reducing accidental deaths and injuries. But it was considered by senior management to be important enough in its own right to justify calling out a separate goal. Our goal to eliminate substandard vessels can be viewed as a general strategy or an intermediate outcome toward achieving most of our other safety and environmental goals. But this goal had a special flavor of industry accountability, and the goal represented a fairly visible public commitment we had made. The measure for this goal -- essentially a violation rate -- remains somewhat short of gauging the real outcome of interest, and doesn't provide an effective way to gauge progress toward our five-year goal of zero; further work continues.

One of our measures was improved substantially by using an external data source. In the first version of our plan, our goal of reducing the amount of plastics and garbage going into the water was measured by the reported rate of compliance with international treaty, based on dockside inspections. We later found an annual report of a nationwide beach cleanup published by the Center for Marine Conservation -- a private, nonprofit group -- which was broad in scope and included a process for cataloging the debris by source. We pulled out the data for maritime sources, divided by the number of shoreline miles surveyed, and had a new measure which clearly focused on the outcome of interest -- how much junk actually gets in the water and washes up on the beach.

To a large extent, these outcome measures replace output and efficiency measures for our organization. We still have data on activities, but we don't regularly track them or hold people accountable for them. While outputs are clearly important in achieving outcomes, our overriding concern was that we change the way we do business and change what people focus on. Given the inertia to continue what we've been doing (what we know best), we felt we needed a major course correction. Besides, we didn't want to collect more data, we wanted to replace old systems and reduce our data collection overall.

HOW WE USE INFORMATION ON OUTCOMES

Our business is risk management. The main use of our performance information is to affect the outcomes themselves -- to improve safety, security, and the marine environment. We do this directly by disaggregating the data to identify risk regionally and globally, in terms of systems, vessel service, flag, classification society, owner/operator, operational circumstance, time of day or season, geography, or other factors which enable targeting our efforts.

By disaggregating our measure of worker fatality rates, we discovered several interesting details about the outcomes of our programs:

- *The National Research Council, several years earlier, had cited fishing* as the least safe industry in America. We found the actual fatality rate to be even higher than that reported by the NRC. This finding helped to validate one of our major areas of business focus.

- *More significantly, we found that commercial towing* was even less safe than fishing -- a finding that did not emerge until we put the fatality data together with employment estimates from the Bureau of Labor Statistics. The raw numbers were not very high; the rates were. We shared this finding with the major trade group for the towing industry. After an independent study to examine the employment estimates (which they thought were much too low), our estimates were essentially validated, and this led to a collaborative study to isolate the nature of the problem. It turned out that most of the casualties were deckhands, and most of them were falling overboard -- a problem which did not lend itself to an off-the-shelf inspection program for towboats. We now have a formal partnership with the towing industry to address targeted, non-regulatory solutions to the problem. The fatality rate has dropped from 77 per 100,000 workers in 1992 to 27 per 100,000 in 1995.

- *A graph of the fatality rates clearly showed clustering* of the data. Both uninspected fishing and uninspected towing vessels had relatively high fatality rates; all of the major types of inspected vessels were grouped together with relatively low rates. While clearly not conclusive, this does give at least a preliminary gauge of the effectiveness of our inspection/regulatory programs, in quantitative form.

In measuring passenger vessel safety, our first indicator was a direct measure of accidents with more than six fatalities (matching our original goal, and our biggest concern). In fact, there have been only four such incidents in the last 15 years. As a result, this primary measure was not informative in showing changes in the underlying risk for passenger vessels. Further work has focused on a more indirect measure of risk -- the rate of "precursor" accidents which have the potential for propagating into major losses of life. In this case, the trends appear to be going in the wrong direction. Both the accident rate per vessel, and the number of vessels in the fleet, have been rising -- a significant warning flag for the program. This has resulted in a subsequent change to our program goal to reduce this risk, and continuing evaluation of the reliability of our measure. We have entered into a formal partnership with the Passenger Vessel Association to address the issues.

In measuring oil and chemical pollution, we noted a subtle, but significant, change in program focus as the performance goal was formulated. The national goal, in legislation, was to eliminate oil pollution in U.S. waters. This was a simple target: zero. For two decades we focused on reducing the number of spills toward this end.

In reexamining the basic purpose, however, it was clear that the *amount* of oil or chemicals in the water was much more significant, in terms of environmental damage. Interestingly, while the number of reported spills has been rising, the volume has steadily declined. The two trends present entirely different pictures of the program, and lead to different program directions.

In disaggregating data on spill volume, we found that "small" spills (under 100 gallons) represented 93% of the number of spills, but only 3% of the total volume. But our investigations process was indiscriminate -- the same level for every spill. Our measure provided the basis for recasting the investigations program to concentrate relatively fewer resources on these small spills, for greater overall impact or return on investment.

Performance information is communicated to the field in several ways:

- Annually, we update, publish, and widely distribute our Business Plan to marine safety offices, strike teams, group offices, reserve units, training centers, and most staff offices. For each goal, the plan includes the most current data, tabulated and plotted, with selected secondary indicators to highlight key risk areas.

- Quarterly, we provide an electronic copy of the data which were used to construct the performance baselines, extended with considerably more record detail based on requirements from the field commanders. (An extract is included in Attachment A.)

- As part of the annual planning process, we hold an annual conference to bring together our District Marine Safety Officers with Headquarters program managers to evaluate progress against our plan, adjust our direction, and develop new strategies.
- Headquarters staff have been assigned "client" districts, and make field site visits throughout the year to discuss and explain the plan and to get feedback directly from operational commanders and their staffs.

Our line managers in the field use the measures in our Business Plan and quarterly data extracts to help target their activities toward national goal achievement, based on local risks. Training in risk assessment has been added to the Port Operations curriculum for staff officers at the Marine Safety School in Yorktown, VA, and our site visits by Headquarters staff typically include a demonstration of the use of spreadsheet tools to disaggregate data. This performance information is then combined with local knowledge and broad interaction with industry, other agencies, and the community to identify the best opportunities and approaches for reducing

risk.

- *Marine Safety Office Savannah, GA* had a head start with their own strategic planning several months before our first Business Plan was published. They used their new discretion to board more vessels, examine more containers and more fishing vessels, and interact in more ways with the industry to reduce the regulatory burden. In the first year, MSO Savannah cited a 33% reduction in marine casualties and oil spills compared to the previous year, and recovered almost double the percentage of spilled oil -- from 16% in 1993 to 31% in 1994.

- *Marine Safety Office Puget Sound, WA* used focused education and enforcement at the start of the commercial fishing season to significantly reduce the number and size of spills from fueling operations; they used a collaborative quality management team with the Washington State Ferry System (a very safe operation, but with high exposure as the largest ferry system in the U.S.) to resolve problems with an automation system; they put marine inspectors on the examination teams for all foreign vessels, and saw a sizable reduction in spills from these vessels. Another of their initiatives was to accept self-inspection for part of the mid-period inspection credit on vessels of several major shipping companies; this reduced the unit's time investment, but also reduced delays in the vessels' operating schedules and increased the overall level of compliance.

- *Marine Safety Office Philadelphia, PA* trained Search-and-Rescue (SAR) station personnel in environmental protection and response, to provide more efficient geographic coverage of their area. With SAR stations distributed more broadly along the coast, the Philadelphia office was able to respond faster to spills and pollution threats, and increase the available time for other prevention and enforcement activities by their regular workforce. Their training program has been adopted more broadly as a "best practice."

- *Marine Safety Office Hampton Roads, VA* inaugurated a practice of issuing "Letters of Concern" from the Captain of the Port, rather than rely solely on the deterrent effect of penalty action when problems have been uncovered. These letters provide an extra incentive for safe operation, since pilots and vessel operators do not want to find themselves in a situation where a casualty has occurred as a result of a practice that has been documented as a concern of the Captain of the Port.

Performance feedback from the field is obtained broadly through an unstructured,

narrative report from each Marine Safety Office at the end of the year. The reports describe the strategies and activities pursued over the past year toward achieving our goals, how those strategies are related to outcomes, what effects were observed, and what challenges remain. "Best practices" are distilled from the reports, and those are shared with the rest of the field to consider applying these practices in other geographic areas. This approach also serves a second purpose: it recognizes field units that manage their operations particularly well. The examples above were drawn from these reports.

Mission performance standards (aimed at activities) were the basic mechanism we had used for management control. It was becoming clear that we couldn't simply add new, *outcome*-oriented goals on top of our *activity* standards and expect our field commanders to be accountable for both. At the same time, outcome-oriented goals presented an opportunity. By clarifying what we really want to achieve, these goals could free the organization to explore new approaches. Rescinding these standards challenged conventional thinking about how to implement and manage programs, and there was an uneasy consensus to make the change. We ended up rescinding all performance standards that were not mandated by law, treaty, or regulation, or required for compelling (and specific) reasons of national consistency. The changes gave new, tangible flexibility to field commanders equivalent to ½ million work hours annually, or 8½ percent of direct program staffing (calculated from historical workload data associated with the activities that were no longer centrally-directed).

More indirectly, performance information is used in a variety of management and decision processes which ultimately are intended to affect outcomes as well. We use our plan explicitly in deciding and justifying our R&D priorities every year. We have used it in developing our multi-year budget strategy, which targeted six program activities to scale back or terminate based on their relative contribution toward achieving our goals. We have used it to demonstrate the value of our programs in the appropriations process. We have used the measures, specifically, as the basis for a research study to measure program effectiveness (trying to estimate *our* contribution to the outcomes), and similarly as the basis for an impact evaluation for our new container inspection program. Our measures are used to help adjust our goals and our plan. And, since we no longer direct and monitor activity levels, measurement and evaluation of outcomes have become the mechanism we use more generally for management control. This, in turn, reinforces the delegation of managerial flexibility to line managers throughout the organization and to the workforce generally.

Personnel evaluations in the Coast Guard have always been based, to some extent, on getting results. All officers are rated on this element. The report requires a numerical grade and narrative comments to address: "*The quality/quantity of the officer's work accomplishments ... the effectiveness or impact the results had on the officer's unit and/or the*

Coast Guard." Outcome measures provide a useful, new benchmark for this evaluation, to the extent that rating officers can describe performance in terms of real-world results. Anecdotal evidence suggests that this is beginning to happen.

By using outcome information in managing our programs, we met or exceeded seven of eight ambitious targets for safety and environmental protection in our first year, and five of seven in our second year, with no additional resources. That's clearly the most important result. But we have observed a number of other impacts as well -- some intended, some unintended:

- We have several new partnerships with our stakeholders in the maritime industry. In our experience, conflicts commonly arise over activities, such as the scope and frequency of inspections, or specific corrective actions. Outcomes, on the other hand, give us common ground in things we all care about, and it has been helpful in general to be able to agree first on the scope and magnitude of the problem before addressing solutions.

- Our workforce, in general, appears more motivated. Field reports describe a better comprehension of the programs, appreciation for greater flexibility, and a sense of excitement about clear and ambitious goals. However, these reactions have been mixed with skepticism and resistance by some at various levels of the organization. Some continue to do the same basic activities they have done for years, and may decline to reassess these activities in view of the Business Plan; they "remain to be convinced" that they need to change or that the Business Plan will endure. However, this skepticism is diminishing as new leadership at the top of the organization has succeeded the old and has embraced the Business Plan as the way we will do business.

- Data reporting is more timely and of better quality. The field is getting their data back quarterly in a form they can use, and this process is further stimulating demand and progressively improving the data.

- Our FY95 appropriation was increased above our request by $5 million and 100 FTE (about 1% and 3% of the totals, respectively), with Senate report language citing the importance of our program goals and the soundness of our Business Plan as the rationale. The Appropriations Committee further permitted the Coast Guard to discontinue the specific program activities associated with these resources, releasing most of these FTE for reinvestment in other activities based on safety or environmental risk, and generally endorsing increased managerial flexibility.

- From the outset of this planning process, it was evident that our goals

inherently cut across the organization. To achieve our fishing vessel safety goal, for example, requires efforts in developing standards, inspection, licensing, equipment approvals, search and rescue, fisheries law enforcement, radio navigation, and navigational information services. In the broad Coast Guard planning process, these program relationships have been recognized and have led to development of a more aggregated set of goals for the whole Coast Guard. This has also provided a catalyst for a current effort toward performance indicators which span all transportation programs at the Department level.

THE COSTS

The cost of our planning and evaluation has been about four full-time equivalent (FTE) staffing, or about 1/10 of one percent of direct program FTE. This estimate includes the business planning efforts of senior management, most of the projects and activities of the Strategic Planning and Analysis Branch, and the data extraction and peer review efforts of the Program Evaluation Group. Most other new things we are doing are directly replacing old things we used to do (like changing the data we collect), with no net resource costs. All of the performance measurement has been done on Pentium computers with spreadsheet software -- an additional $20,000 investment.

Within the Coast Guard organization, the Office of Marine Safety, Security, and Environmental Protection has had wide latitude to experiment with performance planning. The Coast Guard's planning process requires that each program director prepare a business plan. This is a relatively new requirement, and only very broadly prescribed. With few constraints, we shaped our business plan to suit our needs. Our first performance plan for the GPRA pilot was then distilled from the business plan (primarily); it took about ½ day to complete this translation. There have been no additional documentation or reporting requirements associated with the GPRA pilot, beyond those specified in the law.

We developed our performance plans and reports somewhat collaboratively with the other three GPRA pilots in the Department of Transportation. In a series of informal meetings, we shared our plans, challenged each other to help make our plans more readable and meaningful, and shared ideas on the process. After each round of plans or reports to OMB, we shared lessons learned more broadly in seminars for all DOT personnel. Again, these costs were quite modest. The Office of the Secretary of Transportation coordinated all these gatherings, and was pointedly nonprescriptive in the planning process.

Our interaction with Congress has been limited to staff discussions and general

reference in Congressional testimony to the goals and major strategies in our plan. This has generated no new reporting requirements.

THE LESSONS FROM OUR EXPERIENCE

While our mission, organization, or circumstances may be different from other programs or agencies, many more general principles may be transferrable. Here's what we learned about the process ...

Organizing to plan:

- *The personal involvement of senior line managers* is critical in formulating program goals. These goals will be the benchmarks which define success for the organization. They will drive all priorities, including the budget.

- *When managers approach the planning process* with the intent of driving business decisions, they can achieve a better basis for priorities, resource allocation, and dialog with stakeholders. It appears unlikely that they could achieve these gains to the same degree if the process were driven by budget concerns.

- *Outcome-oriented goals inherently cut across organizational lines*, and therefore their development is facilitated by use of a cross-organizational group. Delegation to smaller components of the organization tend to narrow the view and suboptimize results.

The planning process:

- *There is an inclination to focus on management issues* (vs. mission-oriented objectives), and these can cloud the process. These are certainly important to the long-term success of the organization, but they are at a different level than the things the public gives us money to accomplish. It may be helpful to address these management issues first, or at least separately, to avoid inappropriately mixing internal management goals with the overarching external goals.

- *GPRA requires a fundamental rethinking of programs*. An effort to document current business/activities is tempting, but implementing GPRA is more of a strategic thinking exercise. GPRA is not about what you do -- it's about why you exist.

- *The field of Program Evaluation* offers one of the more useful perspectives in strategic planning. It's a typical part of an evaluator's job to reconstruct goals as a basis for analysis, and their experience can bring great insight into this process.

Formulating goals:

- *Outcome goals and indicators* must extend beyond what the program/agency controls. By definition, things you can control are not outcomes. It helps to understand that goals/indicators aren't direct measures of a program's performance - they are a window to the external world that we're trying to influence. This fact allows managers to take the risk needed to set outcome-oriented goals.

- *It's helpful to think of goals* as independent of how the organization will achieve them. This frees the organization to explore alternative approaches to delivering products/services, and insulates the plan from frequent changes/interruptions from intervening factors and changing priorities. Good goals, in fact, can provide a useful anchor during organizational change.

Setting targets:

- *Goals are likely to be formulated,* at the beginning, with no historical baseline data to use in setting target levels. Incremental (or percent-change) goals, based on professional judgment, can offer a workable, interim solution to get started.

- *There is substantial value in setting goals that are a "reach."* Goals must be based on conceivable strategies for achieving them, but they can (probably should) also serve as a catalyst for achieving extraordinary results. Safe target levels do not provide as much intrinsic value to the public, and do not motivate employees to rethink how they do their work.

Developing measures:

- *Performance measures must be relevant to senior management*, and usable in managing activities and resources. At the same time, managers usually don't have the skills themselves to do credible program evaluation (or development of specific indicators for the goals). This is better left to staff with economics, policy analysis, or operations research backgrounds.

- *There are two basic questions* which GPRA suggests need to be answered: 1) are the intended outcomes occurring?, and 2) what is the program contribution to those outcomes? The first question is more easily answered than the second; it must be answered first in any program evaluation; and it can provide a wealth of useful management information by itself (to help focus activities). Impact evaluations are not prerequisite to goal-setting or measurement. In fact, the search for cause-effect relationships can be a frustrating distraction during the early stages of planning.

- *Imperfect measures are OK*. The process should stay flexible and reflect iterative improvements -- just getting started can provide a wealth of useful management information. And using the data improves it.

Using performance information:

- *The key to using performance information* is the ability to disaggregate the performance data -- to identify which factors or circumstances appear to contribute most to the bottom-line measure. This can give program managers a much better understanding of risk than they may have had before.

- *Managerial flexibility is inherent (and necessary) in the process*. In fact, managers can't be held accountable for achieving outcome-oriented goals without sufficient managerial flexibility to achieve those goals. This can be increased dramatically by simply reducing the organization's own internal rules and standards for activity performance. Managerial flexibility may present one of the strongest incentives to undertake good performance planning.

- *Accountability for outcomes can't be in the strictest sense*. There are always external factors and others influencing the results. The issue of accountability may be viewed more productively as a way of assuring focus and reporting on outcomes, and as a way of giving credit to people. There is strong evidence that managers will be responsive to positive incentives, and that programs can work as well or better with less control.

Putting it all together:

- *Implementing GPRA need not be a costly, onerous or complex exercise*. The Coast Guard's performance plan is 10 pages plus a series of baseline measures like those attached to this case study. Much of the process is common sense. And program managers may find that the volume of material needed to describe an outcome-based program is far less than that needed for an activity-directed one.

- *Goals and measures are part of a bigger communication process* - the idea is to communicate program value for making high level decisions on relative priorities.

- *The proof test has to be whether things are getting better*, or at least being affected positively, in the real world. Shortcomings in the process will be dwarfed by success.

SPECIAL CHALLENGES

Linkage between Coast Guard activities and outcomes remains to be more firmly established. After determining whether the intended outcomes are occurring (as demonstrated with our performance indicators), the next step is clearly to identify the Coast Guard contribution to those measured values. An R&D project completed in 1994 provides preliminary measures of effectiveness for two major program areas, and indicates a significant correlation between inspection and boarding efforts and the frequency of spills and casualties. However, researchers repeatedly point out that it is extremely difficult to model the marine transportation system, due to the influence of many factors affecting an essentially unbounded system. Studies have also identified deficiencies in our marine casualty data base which need to be resolved in order to link our activities to outcomes. These difficulties pose real challenges to quantifying organizational impact.

Our goals are generally aimed at incremental improvements, but do not provide a sense of the recurring value of our programs -- in other words, a measure of the fatalities and spills we prevent already, and the cost to maintain that level of safety/environmental protection. This will become more important as we approach smaller numbers of "bad things happening," where our ability to further reduce these outcomes diminishes. Ideally, we would have a measure of program effectiveness which contrasted "today" with the "no-program" scenario, to demonstrate overall return on investment in a credible way.

We still need to extend managerial flexibility substantially beyond the activity standards we relaxed or rescinded when our first Business Plan was published in January 1994. Organizational inertia appears to have stalled the more broad reexamination that was envisioned back then, and several new programs have been implemented using the old, centralized approach toward directing field activities. While we acknowledge that this approach may be necessary to get a new program started, we see little movement in these areas toward relinquishing "control."

Our staffing model no longer works as it was originally designed. The model for marine safety offices was based on predictable activities, whose scope and frequency were combined with estimated exposure (such as number of port calls) to derive resource requirements for each field unit. Without mission performance standards now for many of our activities, we have no basis for many parts of this calculation. To compensate for this new shortcoming, we have been experimenting with a risk-based model for staffing. But while this is relatively simple in concept, it has proven elusive in practice. The basic idea is to allocate personnel resources in a way that corresponds with differences in risk exposure. But there are problems of how to weight the difficulties presented by geography and different industry characteristics;

how to weight the values of different goals; and how to weight differences in our ability to influence the outcomes in different areas. This is especially difficult in a hybrid operating environment, where some activities are discretionary, others are standardized, and all aim at the same set of outcomes.

Accountability at the local level is limited by the somewhat coarse resolution of our performance measures. We can broadly measure success at the national level, but changes from year to year at the district or local levels are not statistically significant. A unit may have one fatality one year, and none (or four) the next. This is complicated further by a variety of factors which would affect accountability. Should a fatality be "charged" against the unit where the death occurred, where the vessel was last inspected, or where the personnel were last licensed? The problems in answering these kinds of questions limit our ability to more tightly couple organizational measures with personal performance evaluation.

WHERE WE'RE GOING FROM HERE

We continue to refine our goals and our measures, to further disaggregate and analyze the data, and to reexamine our strategies for achieving our goals. We are substantially broadening the involvement of our stakeholders in the planning and execution processes. We have already entered into several formal agreements with segments of the maritime industry, and we are pursuing opportunities for more agreements and working relationships in the context of our plan.

The Coast Guard is now extending the performance planning of the GPRA pilot to cover all of the operating programs for the agency. Originating in the program offices and driven by the senior leadership, the effort has roughly paralleled the process used with the pilot programs.

A small team of four senior planners (one from each major operating area of the Coast Guard and one from the office of the Chief of Staff) is doing background and developmental work; all of the program managers get together to develop and vet the goals and strategies; the Senior Advisory Group to the Commandant (consisting of the four senior-most flag officers below the agency head) deliberates, refines, and recommends the strategic elements of the plan; and the Commandant assesses and approves the major components as we proceed. At this point, we have completed a new statement of mission, a set of five strategic goals, preliminary performance goals and indicators. The main effort now is to refine the measures, assemble all of the

baseline data, and lay out our strategies for achieving these goals. A preliminary performance plan has been sketched out.

As we learned with the pilot, all of these outcomes cut widely across the organization, and require the actions of many programs to accomplish common objectives. In a subtle but profound way, this is recasting the focus of the organization from its traditional four "mission areas" to five crosscutting outcomes. By constructing flowcharts for each performance goal to trace how we affect the outcomes, we are gaining new perspective on the broad, multi-program contributions to outcomes.

Efforts to associate program outcomes with resources, and to portray these relationships effectively in the budget, will likely be a major focus for FY98 and FY99. Early experimentation has already shown that there is a confounding many-to-many relationship between activities and outcomes -- that is, a variety of activities are required to produce a single outcome (like safety, environmental protection, or transportation mobility), while a variety of outcomes may often result from a single activity (like inspections). The traditional method of dividing facility costs and apportioning them to programs can substantially distort the costs or savings associated with changes in the programs.

On the other hand, there is a tantalizing insight here. We have argued for years that the public gets extra value from the Coast Guard as a result of its multimission focus. We knew intuitively that this leveraged our resources, but our argument was largely qualitative. By arraying our traditional programs against outcomes, and then estimating the program costs to achieve those outcomes individually, we are beginning to be able to quantify that value. Work continues.

ATTACHMENTS:

A) Performance Indicators (illustrations) from our FY95 Performance Report
   Extract of Quarterly Data for Field Units

B) Summary of Case Study Methodology and References

Summary of Case Study Methodology and References

This case study was developed principally through the firsthand experience and observations of the author, who served as the Chief, Strategic Planning and Analysis Branch, Office of Marine Safety, Security, and Environmental Protection from 1992-95. The author wrote the Business Plans and GPRA pilot Performance Plans for FY94-96, wrote the first annual Performance Report, facilitated all senior management planning, chaired the Program Evaluation Group, chaired the team to develop Coast Guard performance goals and measures, and personally visited Coast Guard field and district offices. He has also worked with many other Federal agencies in developing their plans under GPRA.

Background references establishing the context for the planning effort include program management studies by the Coast Guard and by outside organizations (1975-1991), an organizational assessment (1992), internal policy and organization documents, the Marine Safety Staffing Standards model, the Marine Safety Data Matrix of Program Activities, "old" Program Descriptions, Program Directions and intermediate planning documents, and the various statutory authorities for our programs.

Supporting evidence for program changes and effects includes annual reports from field commanders, reports of field site visits, field requests for data, reports of annual Marine Safety Officers conferences, annual prioritization of Coast Guard R&D projects, cooperative agreements with industry trade groups, and independent studies by industry and non-governmental organizations. The most compelling evidence, of course, is in the actual measures of performance as detailed in our annual Performance Reports.

Estimates of the magnitude of managerial flexibility were based on a computation of work hours associated with the program activities that were relaxed or rescinded in 1994. Using program performance standards published in the Marine Safety Manual, work hours for each affected activity were derived from the Marine Safety Office Staffing Standards Model, then simply aggregated and compared to the total staffing available.

Extract of Quarterly Data for Field Units

Personnel casualty cases (Deaths and Injuries) - Data Fields

Vessel information:

Vessel/facility identification number
Vessel name
Vessel service
Vessel registry
Data vessel built
Registered gross tonnage
Registered length
Hull material

Victim information:

Last name
First name
Position victim held on vessel

Accident circumstances:

Year incident occurred
Year, month, and day incident occurred
River mile marker
Body of water in which incident occurred
Latitude of incident
Longitude of incident
District in which incident occurred
Type of vessel casualty or personnel casualty
Type of accident (e.g., asphyxiation, diving accident, hypothermia, scalded, etc.)
Vessel's status at time of casualty (moored, underway, etc.)
Activity in which vessel was engaged at time of casualty (fishing, offloading cargo, etc.)
Activity victim was engaged in at time of incident
Specific location (of victim) on the vessel at time of incident
Result (dead, missing, or injury)
Deaths (number)
Missing (number)
Injuries (number)
Cause of casualty (up to 3 causes listed)

Investigation details:

Extract of Quarterly Data for Field Units

MSIS marine casualty case number
Investigating unit