



Speakers' Bureau

Performance Measurement and Performance-Based Management: An Interview With Joseph S. Wholey (Released September 11, 1998)

By James J. Cavanagh, U.S. Department of Energy and Stuart Haggard, U.S. Department of Veterans Affairs

Joseph S. Wholey is Professor of Public Administration at the University of Southern California and Senior Advisor for Evaluation Methodology in the U.S. General Accounting Office (GAO).

The views and opinions expressed are his own and should not be construed as either the policy or position of the U.S. General Accounting Office.

Question: "Performance Measurement is a hot topic now. What does performance measurement mean to you? How does it differ from performance management?"

Answer: In writing the paper, "*Performance-Based Management: Concept, Training, and Research Implications*," for the International Symposium on Performance-Based Management and Its Training Implications in Caserta, Italy, I found myself going back and forth between performance measurement and performance management. I see performance management as including performance measurement. To me performance measurement implies a periodicity, not a one shot program evaluation study. So quite often, for management purposes, perhaps every year or every quarter, we measure how we are doing. The second thing that seems to be implicit in the notion of performance measurement is the comparison against goals or target levels of intended accomplishment. Performance-based management, on the other hand, essentially uses performance information either to manage better to improve program effectiveness or to interact with people who control resources to demonstrate what the program accomplished, so they commit the resources to keep the program running.

Question: "As government decision-makers, what are the specific requirements for a performance measurement system? (For instance, does it have to be easy to maintain, provide timely information, link to strategic plans?)"

Answer: Criteria for performance measurement systems include: (1) ability of the performance measurement system to provide valid and reliable information; (2) the utility of the information system, i.e.; are people using it for management decision-making or for other decisions of a policy nature; and (3) the cost of the performance measurement system in terms of staff time, burden on reporting entities,

and other costs. There are these three domains, and you have to keep going back and forth among the three and think about what would make a good performance measurement system under the circumstances. I do not favor performance measurement for the sake of performance measurement; only if it has a use. Is the information good enough to use for management decision-making? That is one level of quality. Is the information good enough to use in a budget or accountability dialog? That may call for a higher quality of data. There may be higher standards of validity and reliability to demonstrate to people that do not have much knowledge about the program what is its performance. When we use the performance information for management purposes, the manager compares that information with other information that the manager has. So the issue of requirements is not a simple one.

A prerequisite for having a sensible performance measurement system is that some kind of agreement exists on agency or program goals and how to achieve the goals. That is step one; the performance measurement system is step two. So first have agreement on what to accomplish at the agency level, which is often called a strategic plan, or at the program level, which can be called a strategic plan if the program is large enough or may be called an operational plan, if the program is smaller. The measurement system is relevant only in the context of some degree of agreement on what goals we are shooting for and how we are trying to get there. That tells us what kind of measures will be useful in managing for performance and communicating what performance is desired and what performance we are achieving.

Question: "What are the major problems with the performance measurement systems and metrics that you have experienced?"

Answer: Problems relate to the "flip-side" of what we spoke about earlier as performance measurement system requirements. You want a measurement system of high quality, useful, and not too costly. A lot of issues relate to feasibility of performance measurement, the cost of performance, and huge problems as to what are the right measures (which technically are called "validity" issues): Is this a valid measure of performance or not ?

A lot of practical problems occur with many Federal programs, as the services are delivered at state or local levels, and so you do not control the delivery of the service and do not control the collection of the data. So you must decide how to construct and maintain a high quality performance measurement system on a remote basis. Some times we use statistical data series to try to assess the performance of programs. More often we use information generated by those individuals operating the program itself, so there are issues relating to the quality of the data, and validity of the data. This may, at times, seem unsolvable, but sometimes "good is good enough," as the information is for a particular purpose.

We do not want to insist on data perfection and drive the cost of the measurement system out of sight, if it is not crucial to the decision-making and program management process. It is true that the Inspectors General and the Comptroller General are always concerned with data quality and accuracy, but we must always think about how the information will be used. Is the performance data accurate enough for the intended use? If so, we do not want to drive up the cost of the system to increase its accuracy without purpose. We want a few dollars left over for staff and service delivery, we cannot be spending scarce dollars only on data quality.

There is a committee that advises the Comptroller General on the "Yellow Book," which comprises the Government Auditing Standards. That Committee is looking into if and how audit standards should be changed in this new era of performance measurement. This is an emerging area, as not only do people not yet know how to do performance-based management, they also do not yet know how to audit performance measurement systems.

In the mean time, we attempt to take the same approach as your Council, sort of a best practice approach, and share reasonably good ways of doing things. We need to see whether, for example, in the Child Support Enforcement Program, the Federal government and the states have clear agreement as to the program goals. Then go back later in try to figure out what the measures of success should be. Then try to make the measurement systems measure correctly and yield comparable data. And then later on come in and audit the measurement system. The whole performance measurement and performance-based management process is evolving in stages. This is an emerging field: We should not demand absolute perfection because it is possible to kill the whole thing.

There is a lot of sophistication on the Hill. We have made much more progress than we would have expected in the last year or so, and people see the difficulty in the road ahead. We cannot predict the future, or even see what happens tomorrow, but (at least among people in leadership positions) there is an understanding that some of these things take time. One of the big issues now, for example, is crosscutting programs. Take job training, for example, where many different agencies pursue similar goals. We must find reasonable ways of measuring each agency's contribution. We need to develop some common measures and some measures specific to an agency that differentiate the unique contribution of the agency. People understand that all this cannot be done in a day -- and that maybe we need to worry about making sure that we can do better tomorrow than we did yesterday.

Question: "What types of performance measures do decision-makers need to see on a regular basis to manage and lead well? (For example, financial measures, customer satisfaction measures, process measures?)"

Answer: We need a whole ream of measures: input measures, output measures, intermediate outcome measures (customer satisfaction fits under intermediate outcome), and, in some cases, end outcome measures as well. In some of the environmental programs, for example, you can see a whole chain of measures. You can measure cost of programs; measure whether certain regulations have gotten out, not just in the federal government but in states and localities. You can see whether people pollute less than they did before. You can see whether the quality of the air or the water is better. You can see whether people's health is getting any better. There is a whole chain of measures.

Another interesting thing is that customer satisfaction measures are often good proxy measures for hard to measure outcomes, as a start. People will not settle for that forever, but it is to a clue and a start. When doing customer satisfaction surveying, it is often quite good to get information on what happened -- in the same interview. (Not just did they like the service, but what is their present status.) So you can get some additional outcome data in the very same interview, and that is a wise thing to do.

In the beginning, use the satisfaction measures as performance measures, but people at the policy level are generally not satisfied with this as a hard measure of program performance.

You must supplement the satisfaction information with additional outcome information, and maybe it will not always come from the program's performance measurement system. Sometimes, it will come from a program evaluation study. These program evaluation studies would not necessarily be done every year, maybe every few years, and you would see how the satisfaction measurement fit with some of the harder empirical outcome information. I continue to think that in addition to customer satisfaction measures, you want to develop other intermediate outcome data on early change in status -- not simply rely upon customer satisfaction. But having customer satisfaction and customer status data is an advance over simple input-output data.

Typically, organizations do not have agreement on outcome measures. We conducted one study at GAO involving 40 different programs, and tried to find out how they were doing on developing their

performance measurement systems. It turned out that organizations either can get measures from past studies or can have the help of the evaluators in constructing new measures. Both were found to be effective.

The idea of common measures is very demanding notion and presupposes that some willingness among the parties to cooperates exists. I am not fully convinced on how far we need to go in the short-term on common measures. I think that, typically, it is useful to try to have some measures in common and some unique measures. That is usually a good way to go. But to get agreed upon common measures requires that agencies work together somehow. Unless there is a convener or maybe a self-help group, and there are examples of both of these around Washington, then the common measures are going to be very difficult to come by. Sometimes the U.S. Office of Management and Budget (OMB) or someone else will invite agencies to get together to develop common measures, or sometimes the agencies will just decide that it is good for them to work together.

One thing about performance measures that has not been discussed a great deal is to what extent will the results of the performance measurements be publicized or made widely available. Often for management purposes you want to have an early warning system, where you can see how things are going. You may not want to put that preliminary data on the table. On the other hand, some agencies have taken this type of heat. The U.S. Postal Service is one that comes to mind. Every quarter they publish all sorts of data on how well they are doing in various categories. They have an external group, one of the big accounting firms, that independently gathers information on the quality of their basic services: whether they are timely in delivering the mail.

Managers need to face the question whether to have a separate internal performance measurement system, or do they want to use the same information for internal and external performance reporting, or will one be a subset of the other. One thing I advised the research agencies, some time ago, was to test their performance measurement systems first and make sure they work well before putting themselves on the line and saying these are the results that they would report to the policymakers.

Question: "If someone were to develop a step-by-step process for developing performance measures, what key characteristics need to be present? (For example, would the process have to be inexpensive to implement, easy for anybody in the organization to use, incorporate customer requirements, be benchmarked against best-in-class performance?)"

Answer: I would see the process as follows (refer to Table 1):

Table 1

Key Steps and Critical Practices in Performance-Based Management
1. Define Mission and Goals (including Outcome-Related Goals)
a. Involve key stakeholders in defining missions and goals.
b. Identify key factors that could significantly affect the achievement of the goals.
c. Align activities, core processes, and resources to help achieve the goals.
2. Measure Performance

- d. Develop a set of performance measures at each organizational level that demonstrate results, are limited to the vital few indicators for each goal at each organizational level, respond to multiple priorities, link to responsible programs, and are not too costly.
 - e. Collect sufficiently complete, accurate, and consistent data to document performance and support decisionmaking at various organizational levels.
 - f. Report performance information in a way that is useful.
3. Use Performance Information
- g. Use performance information in systems for managing the agency or program to achieve performance goals.
 - h. Communicate performance information to key stakeholders and the public.
 - i. Demonstrate effective or improved program performance.
 - j. Support resource allocation and other policy decisionmaking.
4. Reinforce Performance-Based Management
- k. Devolve decisionmaking with accountability for results.
 - l. Create incentives for improved management and performance.
 - m. Build expertise in strategic planning, performance measurement, and use of performance information in decisionmaking.
 - n. Integrate performance-based management into the culture and day-to-day activities of the organization.

Source: Adapted from U.S. General Accounting Office, Executive Guide: Effectively Implementing the Government Performance and Results Act (Washington, D.C.: Author, 1996), pp. 8-46.

Then I would focus on the "characteristics" of an appropriate performance measurement system, which I see in the following manner (refer to Table 2):

Table 2

Characteristics of Appropriate Performance Measurement Systems: Quality, Utility, and Costs Dimensions
1. <u>Demonstrate results</u> : Performance measures should tell each organizational level how well it is achieving its goals.
2. <u>Limited to the vital few</u> : The number of performance measures for each goal at a given organizational level should be limited to the vital few. Those vital few measures should cover the key performance dimensions that will enable an organization to assess accomplishments, make decisions,

realign processes, and assign accountability.

3. Respond to multiple priorities: Performance measurement systems at each organizational level should take factors such as quality, cost, customer satisfaction, and stakeholder concerns into account -- and create incentives for managers to strike the difficult balance among competing demands. Performance measurement systems should cover the performance dimensions -- the outputs and outcomes -- that are important to the primary intended users of the performance information.
4. Link to responsible programs: Performance measures should be linked to offices that have responsibility for making programs work.
5. Sufficiently complete, accurate, and consistent: As agencies implement performance measurement systems, they should balance the costs of data collection against the need to ensure that the collected data are complete, accurate, and consistent enough to document performance and support decisionmaking at various organizational levels. Performance measurement systems should meet reasonable tests of validity, reliability, and timeliness -- and should periodically be reviewed and updated.
6. Used in decisionmaking: Performance information should be used in systems for managing the agency or program to achieve performance goals, in accountability to key stakeholders and the public, in demonstrating effective or improved performance, or in supporting resource allocation and other policy decisionmaking.
7. Not too costly: Performance measurement systems should not be too costly in terms of the management and staff time required to collect, analyze, and use performance data; the costs of any contracts for data collection and analysis; the burden imposed on reporting entities; and other political and bureaucratic costs of performance measurement.

Source: Adapted from U.S. General Accounting Office, Executive Guide: Effectively Implementing the Government Performance and Results Act (Washington, D.C.: Author, 1996), pp. 24-28.

Question: "What major challenges face federal managers in successfully implementing the Government Performance and Results Act of 1993 (GPRA)?"

Answer: I pay close attention to two groups of federal executives. One is the political appointees. They face a big problem in getting clear that one of their main responsibilities is program management. They tend to be policy oriented people, who want to get legislation through or get their budget approved. One of the challenges is to get them to understand that one of their responsibilities is greater program effectiveness in a time of constrained resources. For another group of important managers, the career executives, need to ensure that they participate in the working groups, the committees that address stakeholder needs; to get key external stakeholders and people who depend on their programs involved in developing and refining program goals and performance measures. It is a challenge for managers to involve those groups in developing what different performance measures should be. The creation of the performance measure should not be limited to the chief financial officer or the chief budget officer.

Program managers and executives need to work together to demonstrate their accomplishments. They must communicate the value of what they are doing. Managers have to be players and part of the developing dialogue with customers, as well as with OMB and Congress.

The measures cannot be developed in a "windowless room." In the new GPRA environment, there needs

to be a "public-ness" and a greater interactivity in the development of goals and the strategies and the measures. Many organizations have not experienced that degree of openness; however people who have tried it have found that they can often get the stakeholders involved in enhancing their results. Some of the environmental programs, for example, involved the states and leveraged their resources and results over and above what would have been accomplished with just Federal dollars through voluntary cooperation of the stakeholders to achieve the same goals. These interactions resulted in larger or more aggressive goals and better results than the federal managers themselves would have been able to accomplish by themselves. There is a good example of this in the U.S. Coast Guard story discussed around town, where they got into an alliance with industry for better training programs to improve safety in the workplace. You may start working on a technical issue (do we have the right goals and the right measures?), and find that you can leverage common activities and actions.

One of my doctoral students looked at state efforts in performance measurement. One of the things we noticed was community partnerships in Oregon and Florida, where the creation of statewide measurable goals spurred local people to work towards immunizing the children or whatever the common goal was. The people in the public health arena have been doing this for almost 20 years with a different health objectives. You have synergies that come out of some of these technical discussions of issues, so that together we can accomplish many things that no one of us individually would be able to accomplish. This is a very powerful approach.

Question: "What role do you see for benchmarking in helping agencies implement GPRA?"

Answer: There are two roles for program evaluation under GPRA. One role helps people see how well they are doing in order to set their goals and devise their strategies for achieving those goals. The other role for program evaluation under GPRA is measuring hard to measure things. Not everything can be measured with the performance measurement system. There are all sorts of difficult to get data. You may not be able to get some data annually, and may have to initiate an evaluation study if you do not have the right sort of performance measurement schemes. You might have to bring in some expert reviewers and do some soft evaluation. Now benchmarking could help them both. Benchmarking can help people see how well other people are doing and how they are achieving certain goals. This might influence them to set similar goals and strategies and approaches to achieving those goals in their agency. Also, benchmarking can be a program evaluation scheme, as now you will have new information that you did not have before on outputs, quality of outputs, and outcomes.

Question: GPRA created the structure for agencies to move towards strategic planning and performance measurement. How is the next step of performance budgeting going to occur?

Answer: GPRA has 2 kinds of Performance Budgeting. The first one I would call Performance-Based Budgeting; that started in February. The President proposed what is intended to be a Performance-Based Budget in the sense that it stated that, with the requested dollars and FTEs, agencies would produce certain performance: outputs and outcomes. There is a government-wide performance plan with agency performance plans that connect resources to results, and that is one meaning of Performance Budgeting.

A second meaning of Performance Budgeting in GPRA is that there will be a handful of pilot projects (which have been postponed for a year now but will take eventually place) and there will be a report back. The second kind of Performance Budgeting is akin to Zero-Base Budgeting or multiple level budgeting. At two or more different resource levels, the agency would say these are the outputs and outcomes(the results) that we would produce. Then the decision maker, whether it would be OMB or an appropriations committee, could agree or disagree and fund it accordingly. The second one is what Senator Roth really loves and wants.

OMB concluded the agencies' technical ability to measure performance and technical ability to relate dollars to results was sufficiently weak that they postponed the Performance Budgeting pilot projects. I do not know how the second kind of Performance Budgeting will work out. But to the extent we do not know how to measure, it is premature to start to do this second form of Performance Budgeting, as we do not know how to associate the cost with the results. OMB was correct in putting it off. Now we will see if they will come forth and propose a certain number of pilot projects.

Question: "Can you describe how best-in-class organizations successfully link performance with resource allocation? "

Answer: The evaluation community uses the terminology the "logic model." The logic model shows the chain of causation from expenditure of resources to program activity to delivery of products and services (which are the outputs) to intermediate outcomes and end outcomes.

When the U.S. National Highway Traffic Safety Administration told me their performance goals, frankly I did not believe them because they were so outcome oriented! They set their performance goals in terms of outcomes or results for the American people in terms of crashes, injuries, and deaths on the highways. They said these are our performance goals, and then they showed chain of causation. The agency has regulatory programs, they have grant programs, they have health promotion programs, and they have other levers to try to improve safety on the highways; for example, they will get states to pass stronger laws. Maybe they have a few dollars as bait that they use to help their leverage. They will make grants to states, and then they will then get the states to enforce the laws they have on books. They will pull vehicles off the highways if they are too dangerous to drive. When there is a crash, they have this big research program to find out what caused the crash. So they have all these different ways that they can affect outcomes.

Another example is U.S. Environmental Protection Agency (EPA). They have a little program called the Chesapeake Bay program. It is an intergovernmental partnership with the District of Columbia, Maryland, Pennsylvania and Virginia. They have some measurable goals and they have the chain of causation very well worked out. Healthy people and healthy fish are the end outcomes. Better water quality is an intermediate outcome, and less pollution going into the water is another intermediate outcome. Then they get to shorter term outcomes like change in behavior of industries, agriculture and suburban population (recycling). The intergovernmental partnership, which the Governors and the federal government co-chair, makes choices as to where they are going to get the better payoff for their investments to produce better water quality and healthier fish and healthy people.

These are two fine examples that I mention as linking performance with resource allocation.

Question: "In a recent paper you presented at an International Symposium on Performance-based Management in Caserta, Italy, you stated that organizations need to, "create financial incentives for effective or improved performance; for example, by reallocating resources to higher-performing organizational units." Are you concerned that hastily shifting resources from programs that do not achieve short term objectives, without fully understanding the underlying reasons for missing the mark, might result in more harm than good?"

Answer: In assessment of performance measurement systems and performance-based management systems, what has always concerned me were the negative consequences of performance management in the Job Training Partnership Act program. The early experience with performance-based management happened at the state level, where they allocate different financial resources to suburban and urban areas based on results. They found that local service delivery areas were taking the easier clients and running them through the program, finding them jobs at better wages, and getting high scores. So they went back

and looked again at the performance measurement system and tuned it up to take better account of the difficulty of the task. They found that the more sophisticated measurement systems cut down on the negative consequences.

However, I am sufficiently concerned about the problem you raised to emphasize the use of intangible incentives primarily. There is some research, not totally conclusive at this point, that in the public sector you can go pretty far with intangible incentives. I do see a place for financial incentives, though. The State of Maryland has started interesting incentive programs. They have rewarded not only high performance among school districts, but also rewarded improved performance: where a district has been able to dramatically improve performance. When you're going to give out incentives there are two different things you want incentivized: High performance is to be incentivized, but also improved performance should be incentivized. Where you want improved performance, improved performance should also be incentivized. You don't want to just give to the rich and just let them get richer. You want to encourage people to move up. The State of Kentucky gave different schools around the state a "flag of excellence". They also had in Kentucky the "flag of progress," which was awarded if the school had started moving up over the base line. Maryland was doing it with dollars. Kentucky was doing it with flags. Both focused attention on what the ball game is.

What the Job Training Partnership Act program found was that they needed their performance-based management system to let everybody know the goal of program and what they were trying to accomplish. They use small financial incentives in Job Training Partnership Act: Up to a small percentage of the money in a state (up to approximately 5%) could be allocated based on performance, but the rest of the money was allocated based on a formula that did not include performance. My main interest is having people work toward common goals. This might be accomplished through financial incentives without doing too much harm, or might best be accomplished through intangible incentives, where they probably do less harm.

Question: "It has been observed that, "Performance measures do not say why something happened." Is Congress, and are agencies, prepared for the more in-depth program evaluation needed to determine causality?"

Answer : Interestingly enough, people in policy positions often think that performance measurement and program evaluation are the same thing, and they don't understand the difference. So, are people prepared? No.

The statute and the legislative history are clear: Program evaluation now has been enshrined in law. I was meeting with the U.S. Department of Transportation about their launching a multi-year program evaluation effort, and I think it has something to do with the statute. The statute wants to know: What is your multi-year schedule of program evaluations? That is fine, but we have to see how to use this tool to do two things: improve the quality of our programs and communicate the value of our programs. That is what performance measurement can do, but program evaluation can do it even better, because program evaluation can get at, just as you asked in your question, causation: Is the program causing the outcomes? What portion of the outcomes are being caused by the program? Or how does the program produce those outcomes? You need in-depth evaluation studies to find those things out.

So here is a case in which the statute and legislative history are wise, and they bring the federal government to a renewed interest in program evaluation. At least in the agency I just mentioned and I am sure in many others as well, program evaluation has been tried and focused on small things, but now it should be focused on bigger things. The U.S. Department of Health and Human Services spends a lot of money on program evaluation studies, for example. Let's make sure they focus on the big questions: "Are we making progress toward the major goals of programs?" "How are we doing?" "Can we do better?"

Evaluation studies can help us measure better and also can help us change our program strategies and our agency strategies so that we achieve our goals better than before.

Question: "How can agencies instill an organizational culture that focuses on results?"

Answer: That is very hard. GAO did a study on how far has performance-based management reached in changing cultural within agencies. This survey is about a year old now. It was in our June 1997 report. Since then progress has been made, but there is still a huge, long way to go. On both talking the talk, and walking the walk. I would say that the key thing to me is to what extent are people doing their business in a performance-based way. Are they managing the program activities to try to improve and demonstrate better performance. Are they allocating the resources to try to improve and demonstrate performance. Do they run their personnel system in a way that gets people's attention on good performance, on better performance? Do they run their procurement system in such a way that the contractors are helping us to achieve better performance? When we defend our budget do we talk in terms of performance? I have had some evidence back from OMB, that the conversation is changing: the conversation between OMB and agencies; the conversation between OMB and the Hill; the conversation between agencies and the Hill. Hearings are being held on performance issues. In the best cases, it is like a "full-court press" to change culture.

Within GAO, we have goals and measures, and people seem to pay attention to them. The personnel system wants to know: what have you contributed to the GAO-wide goals? I remember the U.S. Coast Guard, for example: In the fitness rating of every officer there is the question of what have you done that contributed to the Coast Guard's goals. So when you are evaluating managers and executives, performance related achievements are part of the basis of the personnel appraisal. I mentioned earlier that political appointees in management positions are still policy-oriented people. In one department, they introduced the notion of Performance Agreements between the President and the Secretary, and between the Secretary and the head of each bureau within the department. People can understand this linkage. The notion of the performance agreement, as they have in Great Britain for example, are part of what they call the Next Steps agencies. These executive agencies are pieces of the departments, in which they have CEO hired to achieve certain performance goals. They may be hired from inside the government or outside the government, and it is known what their accountabilities are. They have freedom in the personnel system, procurement, and so forth, to do what needs to be done to achieve their commitments and goals.

My solution, if there are negative consequences to be worried about, for example, corruption, civil rights violations, whatever they may be (say, with environmental programs, consequences like too many burdens on industry), is that you should bring the negative consequences into the performance framework by stating goals relating to them. You can have goals related to negative side effects to be avoided, with the goal stated in terms of minimizing or controlling the level of these negative consequences. So you can broaden your framework of what a performance culture is, and have a better chance to be successful. Otherwise there are some people who will fight you all the way because they will say that America was not set up for performance. America was set up for Liberty, and we are all afraid of losing other things that we value. So if you can show how to get those things into the equation, you will be much better off.

In a time of limited resources we have to find every way to be more performance oriented. We have to be able to find more ways of delegating more authority and having better accountability for results in order to get more things done than we could achieve ourselves. We all have to be in a performance culture, but we have to be able to get the participants to perceive that they are not going to lose too much. If you see police officers in New York City accused of violating individual rights as they work hard for Mayor Giuliani's performance goals, we have to balance those crime-reduction goals with goals

for safeguarding citizens rights.

Question: "If someone were to develop a metrics guidebook, what critical characteristics would the guidebook need to have for it to be useful? (For example, would it have to include pictures, be thin, be easy to read?) To make it as useful as possible, what should the content of a metrics guide include?"

Answer: First of all, the guidebook needs to stress that goals must be clear. In assessing performance, we use both metrics and softer assessment schemes like expert opinion, peer review, customer satisfaction, and focus groups. We have to say what we are looking for in a performance assessment system. The guidebook would have in it the goal of ensuring that performance measurement systems are not too costly, that they produce high-quality data, and they provide useful information for management purposes and for policy purposes. So the goal of the performance measurement system must be clear.

Then the guidebook would need to have examples of performance measurement systems that have proven to be practical, cost-effective, and have delivered sufficient high-quality information for decision-makers to do two things. The preeminent goal should be to improve the effectiveness of our programs. We need to have examples of how organizations have not only measured well but used those measures effectively, to improve their organizations and programs. Second, people want to see examples of performance measurement systems that have been able to communicate the value of the programs and projects to the decision-makers and to external constituencies. Performance measurement does not always have to do with getting more money, sometimes it has to do with getting more flexibility. If people can understand and trust the agency, they may be willing to give it more freedom to manage itself and achieve common goals. Most managers very much appreciate freedom and less red tape, and the ability to use their professionalism and their expertise to be innovative in how they do the things that need to be done to achieve common goals and purposes.

Flexibility is a very important matter to program managers. The guidebook could have some examples, say like where the U. S. Coast Guard goes in and shows their business plan to the Appropriations Committee, and the Committee looks at the business plan and gives them more freedom. As another example, NOAA was able to get the Committees to collapse part of their appropriation down from three budget accounts into one big account. You can manage with limited resources if you have fewer constraints, and that is very often a very good thing and a benefit to the managers. The guidebook have these types of examples that show not only technical quality in performance measurement but also systems that were able to provide cost-effective and useful data for internal management and also for representing the program in the policy process outside the agency.

Question: "If the government created a metrics guidebook, what do you think would be the best way to distribute it? (For example, should it be distributed through the U.S. Government Printing Office, hard copies to training centers, available on the Inter-Agency Benchmarking & Best Practices Council or National Performance Review Home Page?)"

Answer: All of these may be needed, they are all good ideas. Different versions of the guidebook would be helpful. If a high level manager gets interested in performance management, then a high-level executive summary would be helpful. It would help educate those at the top-level, and to communicate to those senior management to get them to be talking the talk and walking the walk.

People educate themselves if it is in their interest to do so. It will be helpful to have the guides available in many different places in different forms that are focused to different levels of need. There is training, and there is development. Professional development happens easier where people find that it is worthwhile to learn, because people care, because it facilitates their doing what they want to do.

People will seek out these types of documents, and having them generally available in a number of places will be very helpful. The Urban Institute has gotten into a consortium with ICMA and more than 40 cities around the country; they are trying to measure performance. They're trying to produce and benchmark data from all the cities, as to how is the police department doing, how is the parks department doing, and so forth. They are producing comparable data on the performance of different cities and counties, on how particular government functions are performing. You have in North Carolina a consortium of local governments across a state, and the universities are helping them.

United Way of America has developed a guidebook with the help of the Urban Institute. This has been a very valuable document. They have a short version for the Board of Directors and they have a detailed version for the people to use for practical application. There need to be a number of tools for "effective" implementation.

Question: " Some, including in GAO have taken the position that program evaluation must occur at the program office level (with the implication that corporate/agency evaluation offices are not the preferred approach.) What is your view on agency evaluation resources, and what perhaps should the agency's position be?"

Answer: When I was with the U.S. Department of Health and Human Services (HHS) years ago, I was in charge of the central evaluation unit in the Secretary's Office. At that time, there were 40 different evaluation offices around the Department for which I had functional management responsibility. I believe that both program-level and agency-level evaluation offices are necessary to be effective. In the U.S. Department of Transportation, they are making clear that the majority of their evaluations will be done at the lower levels, however, there will be evaluation activities performed in the Inspector General's office and in the budget office. You need to do the work at multiple levels.

The more important issue is doing the right evaluations: doing important evaluations instead of trivial ones. Make certain to not focus at too low a level or too small an issue to be important across the agency or across the program. Program-level offices are inclined to do smaller studies that do not give insight into overall program performance or more how to improve overall program performance. On the other hand, at the corporate level they are more inclined to do policy analysis and policy research, which again do not tell you a lot about program performance or how to improve their program results. I want the evaluations to be used mostly for program improvement and greater program effectiveness, so it is necessary that the studies be done close enough to the people who run those programs and that they be sufficiently accurate, timely, and complete to be useful to managers.

Question: " A recent NPR study of performance measurement noted that "effective performance measurement systems take time: time to design, time to implement, and time to perfect. Performance measurement systems must be approached as an iterative project in which continuous improvement is a critical and constant objective." In your opinion, will Congress afford agencies the luxury of "Time" to evolve their performance measurement programs?"

Answer: The first thing I mention is that the people at United Way of America make the same point: that effective performance measurement systems take time. They have mentioned three to five years. I do not think you can wait three to five years, however. Under the balanced budget agreement, Congress has decided roughly what resources will be allocated to roughly what things over multiyear periods. The money appropriated will be spent somewhere. It is nevertheless true that every individual agency program is somewhat at risk if in a zero sum environment, everybody in order to protect themselves has to try to accelerate progress. The people on the Hill are into a scoring approach, trying to compare how well the different agencies are implementing GPRA. So far it is focused on strategic plans and

performance plans, but later on performance reports. That, of course, is all paper. To my mind the more interesting things are whether people improve their programs and whether people communicate the programs objectives and performance better. So it is the use of performance information that is important.

Even if the Congress is willing to give the agencies more time, the agencies themselves may not be able to afford to take too much time before they need to be able to demonstrate that they know how to use performance management systems to improve their programs and demonstrate what their programs are accomplishing. I think that most players in this process are willing to accept improvement and progress over time.

Question: "In follow-up to your zero sum comment, where "more" resources serves as the determining factor in program success, do you think that will be forthcoming?"

Answer: First within the agency, then in the dialogue with OMB, and third in the dialogue with the Committees, programs will have a chance to make that case. Sometimes they will be convincing and sometimes not. Since it is zero sum and the money will be spent somewhere, to the extent they can make the case there is a chance that additional resources can be afforded them. Sometimes the agency may be well advised to run pilot efforts to demonstrate that the greater resources produce better results and better performance. Performance budgeting indicates that at a lower resource level we will produce this result and at an increased level we increase our performance. But there will have to be some evidence that a direct relationship exists before you can really be persuasive in making your case within the agencies, to OMB, or to Congress. This is a very interesting multi-level analytic challenge.

The views expressed by members of the Inter-Agency Benchmarking & Best Practices Council do not necessarily represent the views of the U.S. Department of Veterans Affairs.

