III. External Accountability

1. External Accountability as a Goal of Performance Measurement Systems.

Among the purposes that performance measurement might have is the goal of meeting citizens' demands that governmental organizations be accountable to them for achieving the citizens' goals, and providing high quality services to them. Let's call this the goal of "external accountability."

A. As an End in Itself

Presumably, achieving the goal of becoming accountable to citizens is important to government as an end in itself. This basic assumption was contained in an exchange between Rhoda Karpatkin and David Walker. David was talking about how he was developing performance measures to guide the GAO towards improved performance. Rhoda asked a question that went to the core of her understanding of what the point of performance measurement systems should be:

> [You stated] that your ultimate client is the American people....[H]ow are you going to use the satisfaction of the American people, or some service to the American people, as an element in performance [measurement]?

In a liberal society in which government is understood to be "by" and "for" the people, it is an intrinsic virtue for government to be accountable to those who invest their money and liberty in government operations. There may even be some individual satisfaction created among citizens if they can be assured that government is performing well. The only way these self-evidently valuable goals can be accomplished is by producing information for citizens about government performance in terms that are meaningful to them.

B. As a Way to Build Trust in Government

Achieving external accountability may also be important as a way of re-building trust in government. This was the concern of Dean Joseph Nye:

> There is a feeling among citizens that government is not performing well. That has undermined their confidence in government. ...Perhaps with information technology and performance measurement, we can create a new basis for regaining trust in government.

Of course, not all people think building trust in government is a good thing. Indeed, much of our recent politics has been about the question of to what degree a liberal society can and should trust government to achieve public purposes. But one can easily agree that whatever decision makes about the size of government's role in creating a just and effective society, it ought to earn the confidence of its citizens in meeting those more or less limited conceptions of its role.

C. As a Way of Making Government More Responsive to Citizen Aspirations

Achieving external accountability might also be a device that elected representatives of the people and the managers they appoint to achieve their aims could use to help make their organizations more responsive to citizens' concerns and aspirations. Pam Syfert noted:

> It seems to me that the thing we're struggling with is performance measurement is maybe part of a bigger issue about how government runs. Getting back to what Rhoda was saying, what do citizens want from government, how is that reflected in what the elected officials say? And then how do the people in government respond to that?
>
> I sometimes see an alignment between the citizens and the elected officials. But their goals do not always align with the goals of the people that work for government. Sometimes we fail to align ourselves with what they want.
>
> ... We were having a problem with pedestrian fatalities in our city, (We actually have more pedestrian fatalities than murders in Charlotte! It's because we're growing so fast and we have transportation issues.) Our elected officials and the community came to us and said, "we think we have a problem." ...

We put together something that was pretty innovative, and told the citizens that we're going to do some things we've never done before. Then, about two weeks later our traffic engineers told us what *their* priorities were. "What do the citizens really know about the priorities," they asked. "This is what is in *our* work plan."

2. External Accountability to Whom?

It may be obvious that "external accountability" is an important goal of any performance measurement system. What is much less obvious, however, is precisely who the external actors are to whom accountability is owed, and what particular kind of accountability is owed to them.

A. Accountability to "Citizens" and "Customers"

It is natural these days to say that government should take a lesson from private industry and become more accountable to its "customers." And it is natural to assume that the customers of government are those who receive services from government: those who get their social security checks, or wait in line at registries of motor vehicles, or call the police. It is natural to make this assumption because these individuals are most like those who are customers in the private sector: they are the ones that the organization meets "downstream" in individual service transactions, and they typically "benefit" from the government operation. (Even the person waiting in the registry line eventually ends up with the right to own and operate a motor vehicle!)

But there are at least two important differences between these "public sector customers" and "private sector customers:" the public sector customers described above do not choose to *pay* for the service at the point they receive service. If they pay for it at all, they pay for it as a *taxpayer*. A taxpayer differs importantly from a customer in that a taxpayer is levied a particular charge based on his ability to pay for an aggregate bundle of services provided by government. He is not asked as an individual whether he would like to buy a particular service. He is told he must. Because taxpayers do not voluntarily decide to pay for the service, we do not know on the basis of their choice to use the service how much they value it.

Moreover, there are many "customers" of government who meet government operations as individuals at the "service end" of the organization who do not get services that benefit them, but instead receive obligations that they must meet at some cost to themselves. As Dutch Leonard observed:

> ... it makes a big difference whether the performance you're trying to assess is a performance in *delivering a service*, or a performance in *delivering an obligation*. That's a distinction we sometimes make.... [T]he way in which you would treat information ... about ... the quality of the interaction would depend on whether the transactee was an inmate in a prison or someone you're trying to enforce the law on, versus somebody to whom you were trying to directly deliver a service.

Presumably, it is not the ultimate aim of tax collecting agencies, or environmental protection agencies, or police departments to maximize the satisfaction of taxpayers, polluters, or criminals. These people not only do not pay for government operations; they also do not necessarily benefit. Or, if they benefit, they benefit not as individual clients of the organization, but instead as citizens who benefit from the organization's efforts to create attractive aggregate conditions in the society.

It might be useful, then, in thinking about the actors to whom public sector organizations might owe accountability to distinguish between: 1) those who meet the organization "downstream" at the operational end of the organization and receive either services or obligations on the one hand; and 2) those who monitor the organization's operations "upstream" at what might be thought of as the "aggregate reporting end of the organization. We might call the individuals who are downstream in the organization "clients" of the organization (to help make it clear that they are not necessarily paying for the services, nor benefiting from them, nor that their satisfaction is the point of the effort). We might call those "upstream" in the organization the "authorizers" of the operation: in particular, citizens and taxpayers.

These observations about "upstream authorizers" and "downstream clients" might seem remote to our discussion, but let me try to link this idea to our discussion through some comments made in the course of the discussion. Take Rhoda's early question to David Walker (later cited by Pam Syfert).

Rhoda assumed that the ultimate "client" for performance measurement -- the person whose interests we were

6

trying to satisfy -- was "the American people." She then went on to say that the aim could be to "satisfy" the American people, or "provide some service" to the American people. Arguably, the distinction between "satisfying the American people" on one hand, or "providing service" to them on the other could refer to the distinction between: 1) reassuring an aggregate group of citizens that their aspirations for government were being achieved (by, for example, reporting on the aggregate performance of the organization); and 2) providing individual citizens with discrete services that they could evaluate as individual beneficiaries of the service. If this were the distinction that made here use these two phrases, then her distinction would parallel the distinction I am suggesting between the "upstream authorizers" on one hand, and the "downstream clients" on the other.

Pam Syfert, operating in the relatively intimate context of local government also reported on a common experience that helps make this distinction between "upstream authorizers" on one hand and "downstream clients" on the other.

> In local government, ... your performance... is measured on a daily basis. I can't go to the grocery store or Starbucks .... without being measured by the people I work for --- the citizens of Charlotte. So, I hear how well we're doing or how poorly we're doing. I introduce all my speeches by saying I will take service requests after the speech is over.. I have found that many times I'll try to talk about the good things we've done in the city by focusing on what we call our focus areas: economic development, transportation and so on. These are bigger, longer range, strategic issues. Some of the citizens will sit there and say, "Yeah, yeah, yeah, good, good, good. But you know, there is a street light I've wanted for years and I want to know how to get it." That's the reality that we work in in local government and I think that has driven us to try and be more performance oriented.

In essence, Pam finds she has to deal with citizens evaluating their individual experiences with government as "downstream clients" rather than citizens evaluating government in terms of its overall performance as "upstream authorizers."

The distinction between "upstream authorizers" on one hand and "downstream clients" on the other has a potentially useful analogy in business. Business, of course, is pre-occupied with "downstream customers." Customers are important to business both as those individuals whose decisions to spend money on their products and services keep them in business. They are also important as those people whose satisfaction serves as the overall social justification for the enterprise. But business (publicly owned and traded companies) also have "upstream authorizers" who are important to them. Their "upstream authorizers" are the shareholders who own the firm, and the Wall Street analysts who issue commentaries on the past performance and future potential of the company. These people are important because they provide some of the resources needed by the firm, and because their valuations of the firm's potential in the future is reflected in the stock price of the firm.

What is interesting about these "upstream authorizers" in business, however, is that they do not value the firm on the basis of their experience as an individual customer. They value it in terms of its aggregate performance: its ability to meet targets in the past, and its current position and plans for the future. They do not consume individual products or services of the firms as customers; they "consume" *reports* of the organization and its potential. Based on those reports (with the help and advice of analysts), these "upstream authorizers" give it their evaluation by buying and selling stock.

So, if we were interested in using business analogies in government, we might think that government ought to be interested in reporting "upstream" to its "owners," "investors," and "authorizers" as well as "downstream" to its customers. We might also understand that the "upstream authorizers" of governmental organizations might, as part owners of the enterprise, judge the value of the enterprise not simply as individual customers, but instead as owners who are counting on the aggregate capabilities of the organization to help them realize their larger aspirations.

One final note: the distinction between upstream authorizers and downstream clients is not one based on discrete groups of people. Often, a single individual will be both an upstream authorizer and a downstream client. For example, in social security, I am both a citizen/taxpayer and an (eventual, I hope) downstream client. The distinction, instead, is one rooted in one's perspective and the values that are attached to that perspective. When I am thinking like a customer (which I spend most of my life doing), I tend to focus on what is my narrow interests. When I am thinking like a citizen or taxpayer, I am (at least in principle) to some degree thinking about larger issues: the long run, the kind of society I would like to be a part of, etc.

Note that some government operations have the characteristics of being relatively universal services: e.g. social security. For these, the distinction between upstream authorizers and downstream clients is small. We might expect the

upstream authorizers to will that the downstream clients be well treated. For other government operations, the service is not general, and may not even be a service. For these, the upstream authorizers might find it hard to imagine that they would ever be a client, and thus decide that such clients need not be treated particularly well.

This may also be true for different levels of government. It may be that in local government, the link between upstream authorizers and downstream clients is much easier to make than at state levels, and at state levels easier than at federal levels. Viewed from this perspective, one might say that the federal government will be much more focused on satisfying upstream authorizers than downstream customers. And the opposite might be true at the local level. Which of these sets of interests ought to be privileged is a very interesting question. One might naturally assume that it is the downstream client. But the difficulty with that is that the downstream client does not pay. If the upstream authorizers have difficulty imagining that they might be unemployed, or a welfare client, or a criminal offender, they might decide that the satisfaction of these individuals should be given less importance than other objectives such as saving tax money, or putting pressure on clients to change their behavior so it becomes less problematic.]

B. The Uncertain Relationship Between Government Performance and Public/Citizen Confidence in Government

There was another frustration that dogged our conversation that might be related to the important difference in perspective between "upstream authorizers" (who have impressions about and receive reports about the aggregate performance of government), and "downstream clients" (who have concrete, individual experience with governments. Given the difference in these perspectives, there was the awkward possibility that these two were not closely connected at all. Joe Nye put the point sharply:

> Suppose we improve performance in all these federal bureaucracies and no one noticed? ... There is a real problem in translating performance gains into increased public confidence in government. I mean there are some significant improvements of *outcomes* in the U.S. political system in which the federal government has played a major role. Yet, public opinion polls show that the public doesn't give the government any credit for them. If you look at the better standard of living for the elderly, or at the better quality of water and air, people answering questions on polls deny that the government has had anything to do with these results. So there are significant *outcomes* which are really important performance improvements where the credit isn't there so there is no change in the confidence levels.

Note that one of the difficulties here might be the relationship between performance of government understood as the achievement of *publicly desired outcomes* on one hand, and the performance of government understood as *satisfying individual clients* on the other. In principle, citizens as clients could feel the impact of social security on their lives: they could get the check each month with little fuss. But it is a little bit harder for citizens as "upstream authorizers" to recognize when the air and water have gotten cleaner nationally. It may even be hard for them to notice when overall levels of poverty among the elderly have gone down. These latter are *aggregate conditions* in the society: outcomes desired by citizens as overseers and measured in aggregate terms.

[Some important questions that emerge from these observations are:
1) to what extent do citizens form their views of government performance based on concrete experiences as clients, or on the basis of more general reports about how government is performing;
2) to what extent do real improvements in service to clients result in improved satisfaction with government;
3) to what extent can citizens notice when government is achieving important aggregate outcomes such as cleaning the air and water or making an urban environment safer, and how are those perceptions factored into perceptions of government;
4) what role do performance measurement systems play in causing agencies to become more effective in pursuing goals linked to client satisfaction on one hand and the achievement of aggregate results on the other, and which of these kinds of results are they now geared to;
5) to what extent can performance measurement systems that capture the aggregate performance of government agencies in achieving outcomes or providing high quality experiences to clients also serve the function of satisfying citizen overseers that the organizations are operating in the way they intended and thus improve trust in government?]

Paul Light also weighed in on the tough issue of the relationship between government performance on one hand and trust in government on the other. He reported:

In the survey [the Pew Charitable Trust] did last year on public opinion towards trust in government, the

single most powerful predictor of trust was the extent to which respondents thought government was doing a good job. Performance counts.

Now Marvin here, when the Pew survey came out, took out ads all over America saying congratulations to the postal service because they were at 89 percent approval, 89 percent, up 15 percent from ten years ago. And then the postal service workers union took out ads that said since we're at 89 percent you shouldn't contract out any more of our jobs.

But one of the things I'm struggling with ...[is] educating elites and opinion leaders on the value of changing the conversation about performance [from global issues, or topics of the day] to specific measures [of organizational performance].... If they were to come into this room and hear us talking, would say, "Good God almighty, no! [I don't want to talk about numbers that record the real performance of government.] I'd rather talk about the "national mood."

I struggle with how to make performance measurement actionable. .... But how do you translate that up so you can go out to the public and affect confidence in government?

Mark Moore suggested that two possible pictures were emerging:

Moore: What's the relationship between performance measurement on the one hand and our ability either to improve the performance of government or improve the standing of government with the citizens?

One picture of this relationship is that performance measurement is an important management tool that could actually be used to improve the performance of an agency. The theory is that if an agency's performance could be improved, those who transact with the organization will have a good experience, and will accumulate a better opinion of government. Performance measurement focused on customer service will improve customer service and that builds support in the minds of the citizens for government.

There is a potentially important problem in this theory which Dutch has alluded to: namely, that some of the people that government serves get obligations rather than services. Now, a successful "obligation encounter" might well include at least a tolerance of what had happened to people, or at least a sense of fair treatment. So you could imagine that even an obligation encounter could build faith and confidence in government if it were done intelligently and well.

But even if obligation encounters can be made attractive, one can produce aggregate effects on public perceptions of government only one at a time through individual concrete experiences with government agencies.

There is a second picture [of the relationship between performance measurement and citizen satisfaction]. In this second picture, individual citizens don't just transact as clients with the government. Individual citizens also are electors, lobbyists, authorizers, and taxpayers of government. It seems to me that if we're looking for business analogies that at that stage the citizen has become more like an investor than like a customer. Then you could ask, if we're thinking about citizens as electors, authorizers, lobbyists, investor, owners of government, what is it that they would be looking for? Perhaps performance measurement systems are important not only because they drive organizations to improve performance. They might also be important because they could produce a report that is offered to the investors, owners, citizens, authorizers.

One question we should be addressing is whether we have a picture of what that performance measurement system is, or what that report would look like that is given to our authorizers. A related question is whether a report issued to authorizers would be as valuable in building trust in government and helping government to perform better as using and internally oriented performance measurement system to drive performance so that people have a good concrete experience?

Joe Nye suggested the distinction is very important:

... because polls have been done on people's confidence in government. They are asked the question,

9

"What are your encounters with the bureaucracy?" Two-thirds of them say, "not bad." But when they are asked, "what do you think of the bureaucracy?" two-thirds say, "It's terrible." There is a gap between what people report from their personal experience and what they report [in response to more general questions]. Lipsett and Snider pointed this out a decade ago. [What citizens think of government] is a socially mediated experience. Paul is right that performance matters. The post office and the military enormously increased their performance, so when they advertised there was something to advertise. But it's also right that it wasn't just the individual contact of the citizen on Mark's first model, it was the second model which actually made the difference. I think that's, as we think of performance measures and how we bring it out, if we're interested in how it relates to confidence we'd have to think of Mark's second model, not just his first model.

Pam Syfert reported on an effort by the City of Charlotte to align perceptions of performance with actual performance levels. Charlotee has distributed a "State of the City Report, " now called a "Performance Report," for the past three years. Syfert reported several challenges in preparing the report. One challenge is making sure her staff does NOT turn the performance report into a PR report:

To have credibility, it has to show not only where Charlotte is being successful but where we still have challenges and have not met some of the goals we have for those strategic themes. I think the other thing that ....is it's a little too much [about] how much activity we did.

A second challenge is to report to the community as the community perceives of itself:

We're doing a neighborhood assessment with the university, which is going to start showing the changes in neighborhood on about 12 different dimensions: public safety to housing. .... I think one of the challenges in these kind of reports is having information that number one the citizens are interested in and number two, is not blowing your horn and risking your credibility. ....

A third challenge is distribution:

We distribute it as an insert into the newspaper, the *Charlotte Observer.* We send it to quite a list of neighborhood leaders and people involved in our committees. It's put on display in the libraries and the government center, et cetera.

A fourth challenge is getting feedback on citizen perceptions, as well as assessing the value of the report:

... we do a lot of surveying... through the University of North Carolina at Charlotte... annual surveys. And we did a specific communication survey last year that followed up how many people had seen some of the information the city provides and was it useful, et cetera. And then we also do the usual clip and send or e-mail us, send to our web page or call us, was this information helpful to you? And of course we have some citizen groups, we knew their name right away, that mailed in and said, how much did this cost and why are you wasting our time with it?

Session members suggested it would be important to call such things accountability documents, not public education. It is like a report to investors.

The discussion with Pam Syfert reveals the difficulty of trying to use performance measures to increase trust in government by reporting to "upstream authorizers" about aggregate performance as well as trying to ensure that "downstream clients" get good services, and thereby develop a favorable view of government. The difficulty lies in:

1) the great variety of "upstream authorizers" who are the targets of these reports;
2) the varied and variable interests that they have; and
3) the huge amount of noise (competing sources of information) that exists in the system that communicates to citizens about the performance of government.

Thus, one of the challenges of performance measurement to ensure external accountability is to be clear about who is in this environment; what they want to and are entitled to know; and how messages might be tailored to be effective in communicating with them.

D. "Upstream Authorizers:" Representatives of Citizens?/ Democratic Accountability

Over the course of the meeting, we developed an implicit map of the important "upstream authorizers" of government agencies. Rhoda Kapatkin got the key actors on the table:

> I think we can particularly talk about the *politician's*, the *press'* and the *citizens'* reaction.

Later, she added "interest groups" as an important actor in the external environment.

1. Politicians

Among the most important "authorizers," are the elected representatives of the people -- both legislatures and elected chief executives (Presidents, Governors, and Mayors). Politicians showed up early in our discussion, and not in a particularly favorable light. In describing his efforts to improve the performance of the Post Office through the use of internal performance measurement systems, Marv Runyon described what he experienced as political interference in his managerial work:

> Now, downside [of using performance in government]. When you start making a report like this [on the speed of mail delivery within a city] the public gets involved. You wouldn't like to have the public involved. I didn't like it. I'm from private business. It didn't set well with me. But [the reports] got in the newspapers. And so the *Des Moines Register* (or whatever) publishes this report, "Des Moines is 83, and the average is 85." [That leads them to ask me] "What are you going to do about Des Moines?" So you've got to answer that. You have oversight committees and the oversight committee members live somewhere and they want their score higher. "What are you going to do in my district?" [they ask].

> So you have 535 bosses, that's all the Congress. Then you have the mayors and the governors. They're all your bosses. So they get all this information and they start wanting the answers. And they keep wanting the answers until they get to where they think [the performance] is acceptable. That's a downside to this type of thing, it does happen, it would be better if it didn't happen that way. But the members of Congress start to micro-manage the post office, or try to. And if you don't stand up to it then they're going to do it. So that's sort of a downside to measuring. I know we'll talk about it later because I saw in the readings a lot of things about that.

Bob Kaplan also noted that an important difference between the public and the private sector was that the private sector could collect a lot of information about the organization's performance without the threat of having that information become public, or be used by overseers to "micro-manage" the operation, but it was not clear that this was true in the public sector:

> Since I've been here the last 24 hours, I've learned a lot about the inability of shielding what you do internally from outside scrutiny. This, I think, is Marvin's point. I mean the private sector we can run these performance management systems and keep them away from the [external] financial analysts and just operate the way we think we can as long as we periodically show them a financial performance and deliver on that.

> So this idea of openness and measurement providing the opportunity for micro managing by people outside the organization but who have an interest in the organization strikes me as a very interesting and profound problem..... And I think that is a real, this issue of confidentiality, the distinction between the data you need to manage your organization versus the public's right to know is something different in the public sector that does not occur, near as I can tell, in the private sector.

Ed DeSeve reminded the group of why it was that public sector executives were exposed to frequent, but inconsistent, fickle, and incoherent oversight, and how that differed from the authorizing environment that surrounded private sector executives.

> The [U.S] Constitution created a system of checks and balances. We are supposed to have an executive branch and a legislative branch. They aren't necessarily supposed to be working cooperatively. That wasn't what the Constitution created them to do. So there is a conflict built in.

At the same time we have political parties, so the conflict is magnified. [The conflict is] not just between the executive and the legislature; it is also between Republican and Democratic parties.

... As a result, we [in government] have to operate on a strategic plan that is founded on a fundamental organization that is [not designed to produce a coherent mission]. It is based on lack of definition of mission -- pluralism for sure, and conflict. [This creates] a very different [managerial] environment that we must recognize.

[Even if the mission is described coherently for a while, it changes over time] A recent outbreak in tuberculosis certainly focuses the Congress and focused NIH on a disease we thought was quiescent. As result, we [suddenly] have new priorities. And the public is very concerned about [our response to that epidemic]. How do we measure that concern and public confidence as we try to execute those programs?

We don't have a model that allows for that conversation to occur and begins to define good oversight versus bad oversight.

Shelley Metzenbaum went on to remind us of another weakness of our pluralist structure in terms of its ability to capture citizens' purposes in a reliable way:

It's hard to create stable and coherent performance measures] where you've got pluralism of mission. [It is also hard] where your [proposed] performance measures may just be getting a voice for a certain constituency that isn't at the table or needs to be heard in the debate.

And Dutch Leonard made the point that it wasn't obvious that politicians were really interested in defining the purposes of public sector organizations with sufficient precision that their performance could be reliably captured in some performance measurement system:

It will not always be in the interests of politicians to clarify exactly what they want done, because it's much easier to fuzz it up, to not have to decide, to not have to figure out who won. We [politicians with conflicting values and purposes] put both [values and purposes] into the legislation or into the discussion about what we meant for you to do. That way each of us can point to the fact that we won. Our core constituents can be happy and we can continue to do battle against the forces of evil.

And often it's more valuable in politics to be on the side of your constituent's angels, whichever side that happens to be, than to actually win. Because after all you come back and show your scars, and they need you to keep going to fight this great fight because few are willing to fight it and the forces of evil are still in control....

...Showing people that it's actually in their interest to offer the degree of clarity needed for the agencies to then be able to develop measures, translate that into action propositions, seems to me to be not something we can assume, but something we have to take as one of the challenges within this domain, to build up our argument why to advocate for it might be useful politically to add to the degree of clarity offered as guidance to the agencies.

Elaine Kamarck offered some observations about how politicians sometimes think about performance measurement issues:

Let me give you some examples from what we've heard of performance measures which are not useful to politicians. Marvin brought up this morning, correctly, the notion of cutting 13,000 post offices to save $2 billion. Each of those 13,000 is in somebody's district and would be felt by someone. The $2 billion is just part of that trillion dollar mess that is the federal government. Nobody gets credit for [a $2 billion savings that comes at the expense of local ambitions]. So as a political proposition that one is just dead end, ...

.... [A]re there any performance measures that in fact have political appeal? Once in a while in fact, there are. Early on, the vice president decided that he would say clearly that he would cut 252,000 people from the federal government, he said that in September of 1993. Most of his political advisors said don't say it,

don't say it because what if you fail? If you fail, it will be a problem. The calculation was made that if we didn't say it we would never have believers to force the government to actually do it... . And using the levers in this respect, ... it was worth it because we knew that in 1996, people, actual average voters would like the fact that it was the smallest government since John Kennedy was elected president. And the president's speeches in the fall, the best economy in a generation, smallest government since John F. Kennedy was president. That's all he had to say and he said it everywhere.

A couple of other things, crime drops, clearly Giuliani, a Republican mayor of the most Democratic city in the United States has in fact benefited enormously from the fact that there are real noticeable, measurable drops in crime. Giuliani is using performance measures, actually getting results, publicizing those results, results that people actually care about, as opposed to ones they don't, has probably had terrifically bad political consequences for the Democratic party of New York, for many years to come. They were fairly pathetic in the last election, with one exception.

So there are instances where in fact the use of these measures to actually change organizational behavior, to actually move an agenda and get some results have political benefits.

My point is that the screen through which a politician looks at performance measurement is really totally unrelated to the operations, to any "good government" criterion at all. It is purely related to an electoral calculation and a public calculation about what it is that is going to help or is going to hurt, in an electoral sense.

... when we look back at previous similar attempts in government at this, the first thing we noticed was they never were paid any attention to by any elected official anywhere, Congress, assistant secretaries, secretaries, PBO, et cetera. They all became essentially bureaucratic exercises and therefore, they failed. They didn't have any longevity. They failed because they never made that leap from the bureaucratic level to the political level.

If this exercise is to succeed and if we believe, and I think there is plenty of evidence that there are actually some good government potentials in performance measurement. If we would like it to succeed because we believe it has some good government potential then we have to be able to figure out ways for performance measurement to make the leap from the bureaucratic to the political and enter into a political calculus in dealing with this. This is a very different way of thinking.

Harry Hatry, too, came to the defense of the openness of the political process to at least receiving and responding to if not actively demanding performance measurement:

... The question for all of us is can you make the performance information part of the political discussion? ... I've seen ... many cases where ... the data do enter into elected official's and legislator's decisions. North Carolina brought in a great deal more money from the state legislature after they pointed out that the immunization figures for kids in the state were at the lowest level of any state in the country, and it immediately caused attention from the legislature. A lot of state legislatures have been into education, partly because of the education test score issue and the drop out rate issue. The Job Corps, at the federal level, housing allowance experiments, there are a lot of cases where data has indeed played a role in the legislator's decisions.

So I think it's partly our problem of how we can get data to be entered into the political debate, at least for those issues that aren't so hot. And right now we're in a congressional situation where God knows what data would do. But historically there have been occasions where legislation has been passed, authorization legislation and appropriation legislation, where indeed some of the data has indeed played a role.

In sum, if politicians cannot be counted on to be intrinsically interested in performance measurement, but are interested in electoral outcomes; and if citizens are interested in being able to measure the performance of government in terms that are meaningful to them (either as "upstream authorizers" or "downstream clients"), then we ought to be able to get politicians to pay attention to performance measures.

In order to do so, however, we might need some help. That help could come from aggregations of citizens who get interested in performance measurement in general, or in particular areas. It could also come from the media if they were interested in reporting performance measures. The reason is that interest groups and media coverage are linked to votes. And if politicians are interested in votes, then they ought to be interested in what interest groups and the press think are important. We will look at what the group had to say about these "authorizers below.

2. Interest Groups

Session members' accounts of the difficulty that pluralist politics made for developing and setting out coherent missions that could be measured hinted that interest groups might be as much of a problem as politicians. In this view, their interests were too parochial to be easily blended into a coherent mission. We talked repeatedly about the challenge faced by the Forest Service surrounded by particular interest groups who want to use the assets of the National Forest for varied, mutually inconsistent purposes:

Rhoda Karpatkin made a strong argument against that view. She argued for thinking about the interests of interest groups as things that might be important to have reflected in performance measurement systems.

From where I sit there is another perspective. I come back to Ed [DeSeve's] question to me about the Congress and members of Congress and of state legislatures or city councils or other elected officials. [These individuals] are not only in an academic sense the elected spokespeople of the people; they are, as well, the point of entry by which *interest groups* can be heard. Interest groups generally don't lobby deputy secretaries and so on and I've never heard of an interest group lobbying the GAO.

But interest groups are heard regularly and in a publicly accepted way [by elected legislators.] It's part of the process before elected officials. I had the feeling this morning that that is really experienced as a negative by the people who are in the trenches, both of the elected officials and civil society, if I can use that expression, are both interests from which you need protection. And if they were to evaporate overnight, nobody would be any the worse.

But the way I see it, as difficult and occasionally dysfunctional as the system is, the essence of it is that citizen groups can be heard by their elected officials and both the hearing process and the lobbying process and the propaganda process, all of that plays a role in both improving performance and improving our democracy.

... When Harris Wolford ran for office and appeared to have been elected in Pennsylvania because the people cared about health care reform, that message was heard by the president. It didn't work ... [but] there are other reasons why it didn't work. But it was a way of expressing something.

And thinking about the GAO, its report in I think it was `91, on waste in the health care system, was an immensely powerful tool for interest groups to try and make our health care system better. This takes place over big issues and small issues. It seems to me we should explore how these interests can connect better, not how people who are working on performance measures can protect themselves from these interests.

... the people who are creating and implementing these measures care about the outcomes as much as interest groups do. ... If they are not doing that, if they are doing something that is quite counter to what somebody wants, then my answer is that the appointed officials need to recognize that this is the give and take of citizen and government and it's very important to be respectful of it.

Mort Downey seconded Rhoda's comment about the potentially powerful role that interest groups could play in strengthening his agency's efforts to develop and them execute a strategic plan:

I wanted to come back to Rhoda's comment about the interest groups because I think they play a very useful role in doing this right. And just one example, when we were at the station doing our strategic plan, which is usually what everything is run from, we went to great lengths to get real feedback, not just paper feedback but real feedback from everybody we do business with. We made them comment on the plan, we then made all of our top team sit down and go through those comments and react to them.

Then we asked and invited all of the DOT interest groups, the auto manufacturers, the safety people, the environmental groups and the tanker operators, to one big session where they had to look at it the same way we do, not just their narrow piece but how does it all come together.

We've had, since then, a lot better conversation about what is the context of the decisions about their issue in the broader context, but without that effort on our part we just would have been stuck in the usual fashion of just looking, letting them define only the narrow agenda, but it really was up to us to reach out to define the broader agenda.

Jane Garvey sounded a complementary theme about the value of engaging private sector interest groups when they are the regulated parties:

Get industry to the table! They are your customers. They are the people who are going to have to work with you. The key part was bringing in the partners that were going to work and shape this with us. We've been, I think, able to do that. But I will tell you (and as Mort knows better than anyone) it is a consensus that is somewhat fragile. It takes a lot of work and a lot of staying to it and staying with it, both with the unions but also with the industry. It is such a competitive industry and there is often a tendency within the airline industry for them to go it alone. And keeping this consensus together has been extraordinarily important...

So we've got a steering committee that is working very, very conscientiously on developing a set of metrics, we've got a core set of principles that has to do with system efficiencies, delays, safety, increased productivity, and it's that core set of principles that that steering committee is working with us on. So that as we move forward and as we get each one of these automation tools in place we will then have, we think, a good set of measurements.

This is also very close to the strategy that Jeff Tryens has been following on a larger scale in trying to mobilize thousands of Oregonians to become and remain committed to improving performance on a limited number of key indicators that they take as the important conditions they would like to see realized in Oregon.

3. The Media

Developing interests in performance measurement through interest groups is one way of interesting politicians in developing and using performance measurement systems. Trying to interest the media in using performance measurements is another approach. Here, the group seemed perhaps the most skeptical. They thought the press was simply not interested in presenting a balanced view of government performance. Two stories made the point.

Elaine Kamarck noted the press' intense interest in failure and lack of interest in success:

[One national news show] wanted to add me to their list of sources for bad stories about government, because we had proved... very adept at finding silly government stories and publicizing them. ...

I said, I was interested in this proposition if every time they used one of our stories, they would then talk about what the government was doing ... to try and fix the problem that was causing the story." They had absolutely no interest in this and the conversation ended.

Peter Harkness noted that his organization gets lots of calls from the national news after covering a negative story but seldom after a positive one.

Randy Johnson summed up the sentiment: "Bad news is news, good news isn't."

Tony Williams focused on a different problem with the press -- it's unreasonable expectations for performance:

So we started making progress [on cleaning up Washington D.C. financial systems]. By the end of the year, in the public's mind, had made progress in all these different areas. But as we expected, we had "qualified opinion" on our audit. So the newspaper runs a story, "Financial Chaos reigns in the District". Well, this was a low point for my organization, because we actually felt like we were making a lot of

traction and this was an obscenely high standard. Then we also had the control board saying, why isn't the financial system in yet? I'm like, "Hello? There was nothing there!"

But rather than whining, I went to our people and I said, we've got to maintain, somewhere in this government, the confidence of people that something is happening. So I went before the press and I said, on the one hand I think that this is unfair, it's really unfair to say, okay, we're going to take the situation we inherited, have a financial system in place and have a clean opinion on our audit all in one year.

But having said that, I've set a goal for my organization that if we haven't cleaned up the audit in another year, I'm going to leave. And if I leave it's not going to be an opera where I jump off the building, everyone looks and says, oh, how tragic. If I leave, all the managers are going to leave with me. So it radiated clearly throughout the organization, accountability in the organization.

So to some degree we actually established a performance measurement in the process, in the sense that we developed a strategy for ourselves and we had twelve goals, we organized teams around each goal. And it was basically elementary, Bob, you're in charge of this goal, cleaning up the audit. If you clean up the audit I'm going to reward you, if you don't get the audit clean you've got to go. And I'll be there with you. You've got to have a goal of at least having in place the beginnings of a new financial system, if you get it in place so you know the routine at a certain level.

Williams wasn't content with just this press strategy, however, he also pursued the Tryens' strategy of wide consultation -- in effect, going behind the press directly to the people.

So I think there were some successes. I think another key success was in hundreds and hundreds of meetings, getting the public to understand what was at stake and that there were these milestones along the road to recovery and that there was a connection in the public's mind to cleaning up the finances and the underlying fiscal situation. The public definitely got, after a while, I think, that we spent more than we took in, which meant that, we were spending more than we took in, we were spending too much but we were also taking in too little. And a big accomplishment was, and those of you who know the Washington situation, the big accomplishment was getting the people to understand that there was a connection between not taking in enough and how we actually performed as a government, it wasn't all our treatment by the federal government.

3. Strategies for Engaging "Upstream Authorizers:" Citizens, Interest Groups, the Media, and Elected Representatives

Given that one of the important goals of performance measurement is to facilitate external accountability to citizens (both as upstream authorizers and downstream clients), to those who report on governmental activity to citizens (the media), and to citizen associations (interest groups), and to the elected representatives of the people (legislators, chief political executives and their appointees), an important question is whether anyone had any ideas about how best to do this. Some ideas have already emerged in the stories told above, but here we try to catalogue some of the ideas that emerged. Joe Dear made, perhaps, the most basic and important point:

I think there are different audiences for performance measures. Different performance measures are meaningful to the different audiences. I mean, [one has to decide whether a report, or an underlying set of measures or both] is legislative, is it media, is it internal, is it oversight, maybe they're the same, maybe not.

A. The Role of Political Management and Leadership in Shaping a Mandate and Terms in Which an Agency will Be Held Accountable

An interesting discussion also broke out about the role of either political leadership being exercised either by elected chief executives or agency managers in trying to shape both the mandate under which they operated, and the terms in which they would be evaluated by their formal and informal overseers. Not all accepted the importance of this task. When Steve Kelman asked Marv Runyon a question about how the managers got enough time to focus internal attention on performance measures, the following exchange occurred:

Runyon I had to spend about 20 percent of my time on the Hill, which I thought was--

Moore: Time well spent.

(Laughter)

Runyon: I didn't say that! I did spend probably 70 or 80 percent of my time running the operation because that's what we did! I considered we ran a company. I thought I should be spending a *hundred percent* of my time on running that company.

As noted above, however, many thought that spending time building a constituency and setting up an agreement with overseers about what they would do and how they would be held accountable was an important task. Mort Downey, for example, observed the following:

> We also used [a strategic plan linked to performance measures] to basically drive our legislative effort, as we were working on our highway and transportation bill last year. Congress agreed with us on some very targeted programs, state grants that said if you bring down your cost of medical care, which you'd like to do anyway by virtue of seat belt use that reduces injuries, you will get a bonus from the federal government. That helps us get those laws done.

David Walker, thinking about how his organization had to exercise more constructive oversight over executive branch agencies:

> I think it's incumbent on the oversight agencies as well to rethink what their role is. For example, I met with Rossotti (IRS) about two weeks ago and they are pretty hunkered down right now and he has got a pretty tough job. So I agreed with Rossotti that what we would do, and I'm trying to change the paradigm of how GAO looks, we're going to look at them more strategically.

Mike Farrell, reflecting on the experience of the NYPD in developing its COMPSTAT system, also thought that political leadership in negotiating the terms of accountability with overseers was critical:

> It's really incumbent upon political leadership to run interference on those issues that in a contentious environment, there are going to be loads of different, either formal oversight bodies or others with a particular perspective and issues that are of particular concern to them. And you can't let them reframe the question as it has been defined in trying to reach the outcome measures that you've decided on. And giving operational agencies and units the freedom in a sense to be able to work towards those goals without having to be reactive on a constant basis to issues that while perhaps important are really apart from the outcome that's desired.
>
> One example, again, the Mayor of New York being particularly focused on crime reduction, when the debates arise in the city council or with other good government groups and the like who have maybe had a pet measure or a pet issue in terms of oversight over the years, whether it be patrol strength or response time or any one of a number of other issues. In the past, the agency would be reactive whenever one of those issues came up and we felt a little bit at risk, there would be a lot of scrambling and all the operational units would have to work to try and improve on that measure, and really lose sight of the principle goal. And what I see is a change is that now we have the political weight and will to, to some degree, put those issues aside and say, well, they may be important, but in a secondary or tertiary sort of way. And for the folks who are doing the work, keep focused on the principle mission, the principle goal. . And it certainly helps them from being diverted.

But the need for leadership may be as much a constraint as an opportunity, for it may not be forthcoming in the public sector. It may not be only the politicians who are either uninterested or leery of performance measurement. Walter Broadnax focused our attention on the potential limitation of leadership in government organizations:

> I was very envious [of the point Dave made this morning about the importance of CEO leadership for performance measurement] because I didn't know who the analog is to your CEO who was knocking

17

down those walls as you ran behind him to do this work. I don't know who the analog is in the federal government to the chief executive and chairman of the board, I don't know who it is, is it the secretary of the department, is it the president? We had the vice president out in front on a lot of these issues and he was banging on a lot of doors.

But at least as I listen to you [in a private company], I heard a lot of those doors and walls really falling down, and I think of those doors and walls really falling down, and I think there were some dents in moving some walls around. ... but [in the federal government] I don't see a huge cloud of dust [even as the Vice-President] knocked them all down.

B. The Role of "Advertising" and "Accountability"
There was another interesting discussion about the extent to which leaders and government agencies were capable of "advertising" their accomplishments. This was an important part of the discussion of Charlotte's efforts to represent government performance through the newspaper insert. But the conversation went on a bit in some interesting directions.

Joe Nye, focusing on the issue of the ways in which performance might or might not, could or could not be made to enhance trust in government, made the following observations:

It's interesting. If you ask what agencies at the federal level get the highest ratings for performance from citizens, it turns out that the two that get the highest ratings are the biggest bureaucracies of all, the military and the post office. They get ratings in the 56 to 75 percent high degree of confidence. So, you ask what is it [that allows them to have such high ratings]? It's partly that both those large bureaucracies improved their performance.

But it's also that they went into the markets and started advertising. FedEx gave Marvin the chance to go out and say how good the post office was on television. And getting rid of the draft and having to go into the labor markets gave the military the chance to go out and put on ads saying "Be All You Can Be". And having to compete in those markets and use advertising and marketing was a way to dramatize the performance improvements. This is not pure Madison Avenue. If the improvements hadn't been there, people would have seen through it. But the fact that you could combine those two things makes a real difference.

I think one of the interesting questions, if you try to link up performance improvement with confidence is, what are the linkages? Well, it's not likely Congress is going to give money to the IRS to advertise itself, or many other agencies. So you have to rely more on the free media. If you take cases like New York and Guiliani, essentially Guiliani finding it in his interest to pick up a performance measurement and run on it was a way of feathering his own nest but also a way of restoring people's sense of confidence in the quality of life in New York and in the feeling that here is something they really cared about that has improved. And that has a beneficial effect on confidence in government.

So Elaine's comment about the fact that there is a direct incentive for politicians to pick some measures and not others but the ones they are going to pick are the ones they work on, is simply something we should take as a fact of life and not see as a bad thing but see as a good thing and ask which types of performance measurements can we hitch to which kinds of political incentives so the bandwagon actually rolls forward.

And you do get a number of examples, Indianapolis is cited earlier, the case where Steve Goldsmith was able to do this very successfully. But Tony Williams this morning had mentioned the importance of small milestones, not of waiting for the whole thing to come. The fact that these interim milestones showing success are extremely important to develop the confidence that allows the politician to go on to the next one.

So I think that the key question for a strategy performance measure is not just a technocratic strategy, it's a strategy, as Elaine said, to make the linkage between sensible measures of outcome rather than output, that can be picked up by politicians which can then dramatize to the public that things actually are getting better. And that won't be all of them, civil service reform may be one you'll never be able to do, given the nature of the American system. But there is some that you can do and we ought to be looking for those

because that linkage process is really critical.

The idea that the government should "advertise" its accomplishments as a way of rebuilding trust in government naturally raised some concerns -- even some hackles. But the same point sounds much different if we change the concept from "advertising," to the idea of "embracing accountability by widely reporting on the performance of government." Rhoda Karpatkin once again helped us find our way through the thickets:

> I want to return to the discussion about advertising and the problems about advertising. I think Mark put his finger on it when he said, maybe you didn't hear him, accountability. And I think Mort as well, when he described how he was engaging the interest groups in a very direct and proactive way. The word you hear a lot of now with respect to government among the people I work with is the word "transparency."

> If you put on one sheet of paper the words "accountability," "transparency" and "advertising," you can come out with a concept that says that what is important for the government to do, whether it's state or local, where it's easier, or the federal government, is to demonstrate tangibly that it is accountable and that its operations are transparent.

> And you do that by providing (and that's the fourth word) *information* to people. You can call it advertising, you can call it PR or you can call it information. At Consumers Union the department at our organization that tells the public about us is called the Office of Public Information. And to me, this need to be transparent is a compelling need on the part of government and is now being demanded.

> So, how do you tell about your performance measures? You do it in the interest of accountability and transparency. And you have to use the commonly accepted and workable public communication techniques. And that means you have to reach out to the press, to the media, you have to use op-eds, you have to campaign and fight with the TV people who only want the gotcha articles.

> And you have to do it in a very consistent and organized way and you have to bring in, as Mort did, the interest groups, not once a year, but on a regular basis. Whether you use advisory committees or informal groups or personal contacts, or something more formal, that kind of feedback begins to get you to a different place in your relationship with the public.

> And then I think, I guess I hope, that it has an effect therefore on elected officials, because the public can affect elected officials. This, to me, is very important. Here is an anecdote, it's a day of anecdotes.