Performance Management in State and Local Government

How Does *JobStat* Work?

- Weekly meetings
- 2 Job Centers
- Panel of senior agency managers, including the Commissioner

ent Report &

se Analyses:

– New Programs are Created

Wellness, Comprehensive Assessment Rehabilitation and

Compare

G. Over Time

BAYRIDGE(70) POS - Job Center

Jobstat Report, Version 4.5

December , 2004

Director: Rhonda Charles | Core 4,024 | POS Cntgt Eligibility 96.7% | Chystelle/POS Cntgt Eligibility 95.4% | % Front Line Staff on Board 100.0% | Party Sz. 22%

		Dec-04	3Mo. Avg	YTD Avg	2003 Avg
Index 4.5 Score					35.1
Employment					
PLACEMENTS					
1. Closed for Earnings		13.6	16.5	13.7	10.3
2. Budgets Completed		15.2	16.1	15.1	14.4
3. Qualified Reported Placements		28.8	32.5	29.4	25.1
RETENTION					
4. Retention Rate - 3 Mo.		87.3%	87.2%	87.9%	78.7%
5. Retention Rate - 6 Mo.		84.8%	81.8%	82.7%	76.8%

The Nelson A. Rockefeller Institute of Government

Performance Management in State and Local Government

Contents

Introduction

On February 8, 2005, the Rockefeller Institute held a public policy forum on the state and local role in performance management in New York State. The forum was co-sponsored by the Rockefeller Institute, the New York State Division of the Budget, and the Manhattan Institute. This introduction is organized bottom-up, beginning with the local level and then discussing the state and federal levels.

Speakers at the forum made me feel good. All six of the speakers presented constructive, upbeat reports on what they are doing. Their statements reflected a positive view of what can be done, and at the same time demonstrated a needed strong dose of realism on how hard it is to get good performance data that can influence state and local public management.

The speakers to a person stressed using performance management systems to monitor and ratchet up performance to achieve clear goals on a timely basis — not annually, but much more regularly (preferable on a monthly basis) — with extensive interaction between agency leaders and the managers of agency programs.

In the Dall Forsythe edited volume published by the Rockefeller Institute Press on performance management,[1] one of the major chapters on state and local performance management (of which there are several in this volume) is on the CompStat performance management system in New York City for the New York Police Department. Crime reduction is the main goal. Dennis Smith, who is a co-author with William Bratton of the chapter on CompStat, presented an update of this chapter and an appraisal of how other performance management systems, outgrowths of CompStat, are being implemented in New York City.

Swati Desai moderated the panel and presented a talk on how the JobStat system in New York City works to monitor and manage the performance of the City's 26 Job Centers for welfare and related human services. I have attended Thursday morning meetings on JobStat where the commissioner and his/her chief aides meet and interact with the heads of two of the City's job centers. I was, and continue to be, impressed by this demonstration of performance management in action — where it matters most, at the front lines.

Also at the afternoon session, Fred Wulczyn, a leader nationally on performance management for child welfare programs (foster care, adoptions, family preservation and abuse prevention), described his role in designing and helping to operate New York City's EQUIP system. This performance management system, which relies on techniques developed at the University of Chicago, has had extensive practical application. Because it has

---

1   *Quicker, Better, Cheaper?: Managing Performance in American Government*, edited by Dall W. Forsythe. See Appendix A for the Table of Contents. Go to http://www.rockinst.org/publications/ripress_books.html for more information.

been field-tested and operates with carefully scrubbed data, EQUIP is used for ranking and decision making about the sponsorship and funding of child welfare services.

All three systems — CompStat, JobStat, and EQUIP — have developed over time and operate in real time. They are success stories where success is most critical.

The morning session on state-level performance management was organized by the New York State Division of the Budget. The first speaker was Chauncey Parker, Director of New York's Division of Criminal Justice Services. He concentrated on New York City's CompStat system, praising its architect, Jack Maple, and noting that he had attended up-wards of 150 CompStat review meetings. Parker stressed what he called "the three Ds" — Defining goals clearly, having timely accurate Data, and holding people accountable in well-organized Deliberation processes. He described the CrimeStat system his office has established to partner with 15 major urban counties in New York to create similar perfor-mance management systems, focused like CompStat on crime reduction.

The second state-level speaker was Robert Fleury, First Deputy Commissioner of the Office of General Services, assisted by Rebecca Meyers. An important contribution Fleury made was to emphasize the way the mission of an agency affects its goals and management system. The Office of General Services, he said, is "a decidedly operational organization that builds, fixes, and maintains state facilities." Its performance management system is necessarily inward looking — a tool for agency management.

In the discussion of Fleury's presentation, Edward Ingoldsby, Division of Budget Chief Budget Examiner, highlighted points brought out by Fleury. Ingoldsby noted that performance management works best "on an agency-by-agency basis where you have strong commissioner level support." He added that it is difficult "to link performance management with the formal budgeting system." Performance management is not well suited as a tool for budgeting. Doing this, he said, can undermine its efficacy as a man-agement tool.

A good example of how hard it is to avoid problems in performance management if the budgetary stakes are tempting was brought out by John Reed, New York State De-partment of Civil Service. He cited a mis-specified goal for the sanitation system in New York City, the amount of refuse collected. Reed said, "they discovered they were hosing down the truck to increase the weight people were delivering."

The third speaker at the morning session was Andrew Eristoff, Commissioner of the New York State Department of Taxation and Finance. He previously served as both a City Council member and agency head at the local level in New York City, so he brought a multi-level perspective to the discussion. We use performance management "to manage our state-of-the-art taxpayer and collection call centers, to reduce waiting times, allocate resources, adjust hours, and match employee skills to caller issues." This, he said, is "em-bedded in our culture." Eristoff described the agency's "compliance continuum" and talked about the challenges involved in making such a system work well, which he said re-quires that it be "a continuing process." The latter point reflects an important generaliza-

tion — namely, that performance management has to be dynamic, with frequent adjustments of goals and measures to reflect changed conditions and policy preferences.

In the question-and-answer sessions, there was discussion about how agency executives can pull together and showcase performance management systems. The Mayor's Management Report in New York City was discussed — how it has been slimmed down over time, how it is sometimes viewed too much as a political document, and the reasons why governments have to be careful not to "over-integrate" and over simplify performance management conceptually and operationally.

Although it was not the subject of the forum, it is appropriate to add a discussion of the federal role in performance management. For both the federal and state role, my view is that their role should be primarily a leadership, catalytic, and teaching role, except for agencies where the federal government or the state has operating responsibility. (In the Forsythe volume, the chapter on performance management by the Social Security Administration is a demonstration of this point.)

Unfortunately, there is a strong tendency at the national level for the federal government to design and require the implementation of elaborate performance management systems that fail because they misunderstand the federalism terrain. Both the 1994 law passed by Congress, the Government Performance and Results Act (GPRA) and the results measurement system adopted by the Bush Administration, focused on what are called Program Assessment Rating Tools (PART), have this problem.[2]

The Bush administration frequently stresses "results" in budget documents, using PART scores to justify budget changes, which in the current fiscal environment are mostly expenditure reductions. This is unfortunate. For one thing, it can cause the kind of gaming and distortions that undermine the idea of smarter, stronger, data-driven management to improve program performance. For another, it misses a critical point. The fact that a program is underperforming doesn't mean its goals are unimportant. Maybe, to the contrary, the purposes involved are so important that more money is needed along with better managerial capability to carry them out. Performance management is best suited, as its name indicates, to managing performance. It is strongest and most useful if carried out at the level of operational responsibility.

When we decided to publish this report in hard copy, we asked all of the participants to work with us on editing their presentations and I thank them for doing so. Michael Cooper, Director of Publications, supervised the preparation of this report; Irene Pavone in my office worked with us to organize and review the material presented. I thank both of them for their help.

Richard P. Nathan

---

2   Richard P. Nathan "Presidential Address: 'Complexifying' Performance Oversight in America's Governments," *Journal of Policy Analysis and Management*, vol. 24, no. 2 (2005): 207-215. See Appendix B.

Welcome — Peter D. Salins

I am pleased to participate in this event. First of all, I just want to get through the formal part and that is welcoming you to this conference on performance management. I salute Dick Nathan and the team that put this seminar together. They did a great job in organizing it. I extend this welcome wearing numerous hats. My official welcoming capacity as the provost of the State University of New York, the parent organization of which the Rockefeller Institute is an integral and highly esteemed institutional component. I am also a fellow of the Manhattan Institute, which is one of the co-sponsors of this event. I am, unfortunately, not in any way affiliated with the Division of the Budget, so I don't wear that particular hat. In these capacities, I'd like to express my enthusiasm and support for the conference. It will be a stimulating discussion dealing with some key conceptual and substantive issues that face New York State and the budget.

But beyond my institutional hats, I wear another hat and that is as a public policy analyst. I am one of you. I have wrestled with the issues surrounding performance management and the effective deployment of resources for my entire professional life as a teacher, a scholar, an author, and an administrator. Beyond the work that I have done over the years in a large variety of housing and economic and urban issues, I try to use the tools of performance management in my current position as provost of the State University. Issues of outcomes and outputs, their measurement, deploying effective tools to influence their direction, are part of the daily fabric of my office. From my point of view, performance management at the State University has worked and it has worked effectively. Over the last eight years, the University has made enormous progress in attracting better students, giving them better education, getting them to earn their degrees in a more reasonable time frame, and helping the faculty become more effective teachers and productive researchers.

We've used a variety of performance management tools to achieve this. One very important tool that those of you who are from the Division of the Budget can appreciate is our budget allocation process, which uses a form of performance budgeting to drive academic policy across our 64 campuses and it also tackles our particular version of the state-local, in this case, system-campus division of responsibility. We have developed a variety of performance management initiatives to improve instruction, to promote greater scholarship, to increase the efficiency of our various campuses and components of the campuses, to enhance our facilities, and to create a more hospitable student environment.

I know that the country and maybe even those in our own ranks are ambivalent about policy wonkery. In the media the term "policy wonk" is generally an epitaph. But while I may not share the views of everyone here as to the appropriate extent of the government's role in some major functional areas, I definitely agree with the majority of my fellow wonks that whatever we do, the public funds must be used effectively and efficiently, whether by simplifying or "complexifying" in frameworks that we use. With that I conclude my welcome and I wish you the best for a stimulating and interesting conference.

Part I — The State and Local

Role in Performance Management

Overview — Richard P. Nathan

Today's program for me is a learning experience. I am going to listen hard and distill ideas out of your discussion on state and local government performance management. I regard my role as more important being in the audience than being a presenter, but I'm not shy after many years of being a policy wonk. I decided that what I would do is write a set of notes, which I call "Overview," about what I'm thinking for today's session, what I'm particularly listening for, and what I would like to have us expand upon. This is not a paper, but notes. But first, a commercial message. We published a book, *Quicker, Better, Cheaper: Managing Performance in American Government*, edited by Dall W. Forsythe. One of the purposes of this book is to drill down and look at performance management, as Peter Salins mentioned, and performance budgeting at the state and local levels. We specialize in much of our work in the nature and importance of American federalism when it comes to what happens to public purposes after they are decided upon.

Henry Olsen, from the Manhattan Institute, David Kidera from DOB, and I began talking about this program a long time ago. I am pleased by the way it shaped up and by the fact that we have distinguished state and local government management leaders and experts to participate in crucial ways in this program.

Now let me make some comments to set the scene. I view the genesis of the idea for this conference as studies we do at the Rockefeller Institute. Brian Stenson, our new deputy director, with a lot of help and involvement of Don Boyd and David Wright, is interested in and trying to be helpful and provide good offices on state issues involving budgeting, which is the hottest topic in state government at this time. We welcome Brian and our new Institute co-director Tom Gais, who is also in the audience today. Many people at the Rockefeller Institute conduct studies that involve looking on a broad basis at implementation and finances. That's our role: Not to advocate but to educate. To help policymakers understand how public policies are formed and then how and if they are implemented.

"Complexifying" Performance Oversight

At the 2004 annual research conference of APPAM, which is always a wonderful event for people like you in this audience. I gave a talk with the title "'Complexifying' Performance Oversight in America's Governments." I got the word "complexifying" from Senator Moynihan. I was testifying before him about his 1988 Welfare Reform bill, and I was making the point that we needed an argument to pay more attention to the fact that nobody was paying attention to it. At the Rockefeller Institute, we had issued a series of reports on the flawed implementation of that law. Pat said (it's somewhere in the record, although I think maybe it was expunged because I haven't been able to find it), "the trouble with you, Dick, is you're a complexifier." He didn't want to hear all that. I responded, "I'm very sorry but I need to say it."

A key point for today is that we need to be careful not to go overboard about performance management. People in government always have to have a crisis and have a big initiative, and we tend to talk about performance management this way as if it is going to deliver us from more evils than it could ever deliver us from, to claim more about how it can be done and what it can achieve than is reasonable to. We need to recognize how big and complex governments are, and so are their agents, nonprofit organizations. Government in the domestic sector is not privatized, but nonprofitized. To talk about and be serious about how we perform public functions and carry out public policies is complex. It involves, intrinsically, American federalism and state and local government or you don't understand the subject. That is a hard point to make to people in Washington.

When I taught at the Woodrow Wilson School, I said to the bright bushy-tailed students that I feared they would go off to some think-tank or public agency and think they could just lay it all out: "Now you go do it and be smart like we are." It doesn't work so easily. There is a cartoon from Pogo. Some of you remember the cartoon, Pogo. There was one frame where Pogo says, "Yep son, we have met the enemy and he is us."

That is a good way to make the point about how we tend to go overboard, claim too much, demand too much, and not understand enough. I was talking to a graduate student once, and I mentioned Pogo and "the enemy is us" and the student went off to the library
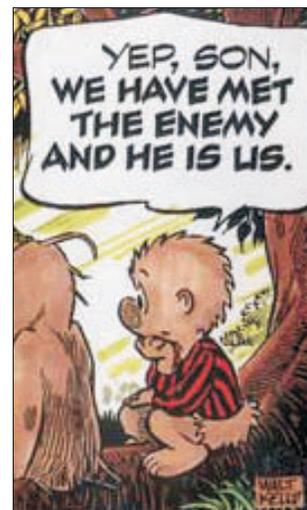
and came back and said, "I couldn't find Pogo," thinking of Pogo like the philosopher Plato. I told the student who Pogo was and he was extremely embarrassed. One day I mentioned this in his presence and he came up to me afterwards angrily and said, "Don't ever say that again and use my name." So, I won't say who it was.

Promising Too Much

"The enemy is us" refers to the over hyping and hoopla that gets associated with performance management. Some social science experts have a tendency to want to measure everything, often in ways that can't be done. I use a quotation from a very important source: the Book of Hosea in the Old Testament. "The number of people in Israel," Hosea said as they were exiting Egypt, "shall be like the sands of the sea, which can neither be measured or numbered." There are a lot of things like the "sands of the sea" that can't be numbered and measured. We have to be careful not to promise too much. Unless a program has a large purpose and a clear purpose it is going to be very hard. You can select things you want to highlight. You can be smart about it, but you mustn't over promise what performance management can do, although it can do a lot.

Next, I have an exhibit, the latest issue of *Governing* magazine, which ranks all the states on how well they manage. If you read it carefully, it's almost funny how they do it, but I still think it's a good thing. They have lots of caveats and believe me they need them. (The interesting thing is that they say New York gets a B-. Everybody either gets a B- or B+. It's like Garrison Keillor in Lake Wobegon. It made me think of that.) The authors say, and I've followed this over the years, "Credible strategic planning at the agency level, once the nonstarter in New York, has been gaining momentum. There's growing interest in performance measurement and reasonable attention to capital planning and other basic infrastructures." There you are. We're doing better and yet my understanding of state government in New York, based on knowing about it and reading about other states, is that we don't put as much emphasis on performance management per se as some states do.

Performance management wasn't "made in" New York State. One thing to think about, and this occurred to me as I was preparing for today, is that with the opening up of the New York State budget process and legislative hearings in public, a lot of what the agencies do maybe now will be on the record in a fuller way. The things that agencies do that they measure, that they use to communicate mainly about with central staff and with the DOB, may become better known and that may influence people to live up to the nice praise I just

read. Perhaps I'm not being fair and I'm not fully informed, but I think this is something at least we don't put up front in a way that some states and many cities do. So I think that it is important. The next panel will get into that with agencies. They are different in what they do. Therefore, what they can measure and how they can measure it depends a lot on the nature of their task and the politics of their function. That is further and needed "complexification."

We pay our politicians to over promise. If you want to get elected, you should say there is a crisis and second that there is a great solution for it, which only I know. That's the marketplace for politicians. I mean no disparagement. But we can compound the problem of over promising by acting as if we can measure things with exactness where we can't. When you look at the studies that have been done of the effects of public policies you know it's extraordinarily difficult. I have been fascinated by this and have written about it, too much I'm sure some people would say. The gold standard of performance measurement is "impacts." What happened because of a program that wouldn't have happened without it? The counterfactual. Well, that's very hard. You can't measure the same person doing something and not doing it. So your counterfactual has to be a randomly assigned control group and that is expensive and difficult; it takes a long time, and it often fails to take into account the fact that policy purposes differ. They differ in different times and places, and by different officials, and by different agencies.

I was for a long time involved with Manpower Demonstration Research Corporation (MDRC), an organization that I greatly respect. MDRC and many other groups do impact studies with random assignment — classical social experiments, and I believe that that is good. But you can only do social experiments when you've got an important, large enough public purpose and where political leaders are willing to spend money on that study and take the time to do it. It has to be done so that when you show something worked, it is something that can be replicated, which is a big assumption. There are indeed a lot of big assumptions here. I think it's good but people tend — this is one of my beefs — to want to do social experiments of too many things and to pretend that you can do it when you can have often only very weak interventions. It's a good thing but it's not the answer for everything that we care about.

Managing for Results

Congress thought performance management was a good thing. A decade ago (and this is unusual, because it is usually what you hear from the executive) the Congress enacted the GPRA Act — the Government Performance and Results Act of 1993, called the "Results

Act," requiring all federal agencies to do elaborate things to make good on a very reasonable proposition, which is we shouldn't judge governments by what they spend money on, but by what their results are. Very reasonable. I know from my own experience how tempting that idea is. Let's look at that for a minute. Results. Take as an example the GI Bill of Rights. Now, most people think that the education program after World War II was a very good thing and had achievements. We still believe in it. Education is the right thing to do for returning veterans, among the other things that we should do. Now what could be the impacts or outcome of the GI Bill of Rights? Are people happier? Do they live longer? Do they earn more over their lifetime? What are the impacts? Outcomes are the same as impacts, but without a control group and a counterfactual.

Consider if Congress had decided to do a study. Ten years from then they likely would have found that the GI Bill was modestly successful. Modestly significant results are all they'd find because it's so hard to measure. Politicians don't think "modest" sounds very good if we find modest results of a program. We didn't do that study. We should not have done it and we didn't. There are a lot of cases where even if we wanted to, rigorous impact assessment is something that we can't do because the program is too small or too squishy, or because we just want to march ahead as we did with the GI Bill. Even if you did a statistical comparison of veterans who got GI Bill benefits with the comparison group that didn't and then say, "Look how great they are," what good is that? The comparison group is people who maybe weren't energetic enough or smart enough or anxious enough to advance themselves by getting educational benefits. So does that prove anything? These are very complicated questions. I'm not saying you can't do useful social experiments. Just be careful.

Outcomes and Outputs

What is the difference between the outcomes and outputs of a program? The more I think about it, the more I don't know the answer. I'm supposed to know things like that. I define it in the following way: Accept for a moment that outcomes are what happens to people and outputs are what programs do. That seems okay to me. There are many things where if you just say what the outputs are I'm going to be satisfied. For example, integrating human services for needy families. You could never prove it works. Mark Ragan and I have had many discussions about this. There are some things where I just want to know who did we do this for, how much of it did we do, how much did it cost, who was served, and where were they now? We've been trying to measure the impacts and outcomes of childcare for like 35 years. I'm not sure that we're ever going to get there. There are a few studies. There are great

debates and the literature is full of complicated articles with formulaic, hard-to-read explanations of what they did, and it still isn't satisfying. I would say the same thing about after-school care. Tell me the outputs. We need to be very careful about what we measure rigorously.

The Federalism Challege

I have two more points. One is about federalism. I'm a nut about federalism and if you are, you would be happy at the Rockefeller Institute. Often any government that puts a dollar or a shoulder into something wants to measure it. The federal government is the worst offender, saying it will measure things that they have no control over — instead of asking state governments, which is what they should do, to make sure that sensible measures are made that take into account local government structures and the role of nonprofit organizations.

The second point is a related point: The question of <u>who</u> should measure performance is often as important as <u>how</u> to do it. That involves state agencies for sure as the crucial middleman of federal systems.

In New York City, when the commissioner would allow it, which now I can't do anymore, I went to the "JobStat" meetings. Swati Desai will talk about this today. I was impressed. There are 26 Job Centers in New York City. Then Commissioner Jason Turner and his assistant, Andy Bush, would sit in a room every Thursday morning with the directors of two Job Centers, and they would present and discuss data about their outputs and outcomes over time and compare them to other Job Centers. Using the same data for all centers and shifting the data used as purposes shifted, trying to control for differences in conditions in different places, the population served in different places in the City. It was a civil, important conversation. It involved the commissioner and his top aides and the directors and their staff of two job centers each week. It was a management tool, which is a very important goal of what performance measurement has to be about. If you're going to use performance measurement and decide who should get more money or should get less money, I think that's going to defeat it. It's going to be defeated. What if there is a purpose you like and it's not working? Well, maybe you should provide more money for that purpose, or maybe less money. I think if performance measurement becomes strictly a budget tool and isn't sufficiently a management tool, like school report cards, that is not good.

Part II — The State Role

Introduction — Henry Olsen

We have been conducting a number of forums across the country on performance management. I direct the Center for Civic Innovation, which was formed by then Indianapolis Mayor Steve Goldsmith ten years ago. Of course, any of you who have recognized Steve's name know that he is synonymous with performance management at the municipal level. He came into office as a privatizer and came out as an advocate of competition because what he learned was not necessarily that the private sector can do public purposes better or more efficiently, but that it was the measurement of performance and the injection of competition into provision of public services that enabled the citizens of Indianapolis to get the most bang for the buck.

I think we do have to be careful, as Dick very eloquently mentioned, of being aware of the over hype. If I could try and sum up the thrust of that aspect of Dick's speech using a quote from a distinguished graduate from the London School of Economics, Mick Jagger, "You can't always get what you want, but if you try real hard, you get what you need." I think that's often the reality of performance management in some ways. Government does things that are difficult to equate to a private sector model because you often do not produce a discrete product and do not have internalized cost structures in the way that somebody who produces soap or the distribution at Wal-Mart has.

Much of What Government Does can be Measured

As Dick also mentioned, you also don't have goals that are capable of being measured precisely because again you're not trying to move person A from point B to point C on that mounted scale, but rather to have a generalized improvement in their well-being, something that is inherently difficult to measure. But nonetheless, there is much that state government does that can be managed and can be measured. As Dick has pointed out and I've seen myself at the local level — particularly in New York City with the CompStat measurement system and management system of police services and the JobStat system of measuring the ability of welfare recipients to move into the workforce — things that were initially considered complex can, in fact, be managed using effective performance management tools. At the macro level, everyone can think of that as well. Take a look at New York State. New York

State currently spends among the most in the nation per student on operational costs for education. Sometimes New Jersey is number one; sometimes New York State is number one. Do New York State's educational outcomes measure up as an efficient use of those resources even after adjusting for different cost of living? That is one of the things that a new project of the Manhattan Institute is going to be addressing in the next year. We have recently announced the creation of the Empire Center for Public Policy, which will be looking specifically at New York State issues, not merely performance management and budgetary issues, but the wide range of public policy challenges that New York State faces to offer innovative solutions. I'd like to introduce E. J. McMahon, who is our fellow and director of that new center. If you have not already had the opportunity to introduce yourself, please take that opportunity because you will be hearing and seeing a lot more from him and the Empire Center over the next year.

Health care is another example of New York State's need for performance management. New York State spends more than California and Texas combined on per capita Medicaid services. Does it have appreciably better outcomes? Are there techniques and tools that we can use to measure those outcomes and consequently deliver the services that we need as a state at a cost that are affordable to our taxpayers? Depending on how you calculate it, New York has the first, second, or third highest per capita tax burden in the country with pressures to increase that tax burden. With deficits as far as the eye can see and with pressure for more spending on all aspects, particularly in education from the Campaign for Fiscal Equity case. Questions of performance management ought to be at front and center of state government over the next five years.

It's my distinguished pleasure to be able to introduce three people who are wrestling with these challenges today and who will be able to tell you what they have been doing and what they plan to be doing, the general challenge of making sure that New York State's taxpayers get the services they want at a price that is affordable.

In no particular order, we have with us Chauncey Parker, who is the director of Criminal Justice. As director of Criminal Justice, Mr. Parker also serves as Governor Pataki's senior advisor for criminal justice and serves as the commissioner of the New York State Division of Criminal Justice Services. Mr. Parker is a graduate of Rollins College and the Duke University Law School and serves on the board of Harlem Day Charter School. He'll be followed by Mr. Robert Fleury, the first deputy commissioner of the Office of General Services, overseeing centralized support provided to state agencies with design and

construction, public procurement of commodities, technology and services, real property management and development, and many other service functions. He served as general counsel for OGS for five years, immediately preceding his appointment. Last, we will hear from Mr. Andrew Eristoff, who is the commissioner for Taxation and Finance and previously served as the department executive deputy commissioner. He has also served as the New York City commissioner for finance during the second term of Mayor Giuliani. He was a member of the New York City Council for six years and was legislative counsel to state Senator Roy Goodman of Manhattan. Mr. Eristoff is a graduate of Princeton and of the Georgetown University Law Center. If I can turn the microphone over to Mr. Parker, I think we will go in order for about three 15-20 minute presentations and then we will open the floor up to questions of all three panelists. Thank you.

Criminal Justice and Performance Measures: What Gets Measured Gets Done

Chauncey G. Parker, NYS Division of Criminal Justice Services

Thank you. First I want to tell you that it's a pleasure and honor to be here at this event sponsored by the Manhattan Institute and the Rockefeller Institute. I think that what's terrific about the work that you do is that you're optimistic about the possibility of what government can accomplish. That's what performance measures are all about. I think that anybody who isn't an optimist would run for cover when it comes to performance measures because what we're really doing is we're holding ourselves up to count by very specific measures, holding ourselves accountable for what we do in government and what we can accomplish. I think the work you do in really insisting that we can always do better in government is really important.

I'm also honored to be here with Bob Fleury from OGS. I'd say that every good idea that we've ever had in Criminal Justice we've always borrowed from somebody else. One of the ideas that we had went through Bob and Ken Ringler. Their performance measurement software is something that we're now implementing in state criminal justice, but we really looked to them as leaders in this and using them as our radar system in criminal justice. Andrew Eristoff has done such a terrific job in public service both as a city councilman in New York City, but also as a commissioner in the New York City Department of Finance. He knows this business far better than I do. It's a pleasure to be associated with two people who have been in this business a lot longer and are accomplishing a lot more in this area.

I want to take a little while to tell you about what we're doing in criminal justice and how performance measures are everything to us in terms of where we're going. About 15 years ago, I read Reinventing Government by David Osborne and Ted Gaebler, which I'm sure many people here have read. To me this is the most important book, or among the most important books, you can read in government. It states this really basic simple principle or philosophy and that is: What gets measured gets done. That is what we think about in criminal justice. What we think about every day. It drives everything that we do. It is so clear that it is just part of human nature.

We know that as kids in class, if class participation is part of the grade then people are going to participate in class who might not otherwise participate. It's natural. The teachers say what percentage of your grade is class participation. We all listened to that because we knew it would get measured and get done. On report cards if it's pass/fail, we're going to have a different attitude when we're in school than we are if it's for a particular grade. We know in sports that whether you kick a field goal or run for a touchdown, one is worth three points and one is worth six points. You're going to have a different strategy because what get measured gets done. These examples go on and on. I think one of the best examples is in business. The best businesses, when they do well, what they measure are profit and everything that they do is geared to it. They don't manage by anecdote; they manage by profits. This is the bottom line, a clear, objective goal that everybody can look to see whether or not you are effective or not effective. But that principle is so true.

What this book talks about is that it's absolutely critical you pick the right measure. It's easy to pick all sorts of measures and collect all sorts of data, but if you pick the wrong measure then you can take your ship and sail it in a completely ineffective direction. For example, in the context of law enforcement, if you measure success by how many people you arrest, that's going to be one kind of strategy. If you measure your success by how many police officers were assigned to a particular area or how much overtime is worked, or how many kilos you seized of cocaine that leads to a particular kind of strategy. What I think is so important is the paradigm, I think this started in New York City, for what we measure in criminal justice changed really dramatically in 1994 when the police department said that we're not going to measure success by how many people we arrest, which was really the general understanding. Instead we're going to measure crime reduction.

The Three D's of Management

I think one of the pioneers, one of the people who I absolutely love, who is such a driving force behind this, was Jack Maple. If any of you ever knew or ever read about him, he unfortunately passed away a month before September 11th. He had this belief that anything was possible. He really set a process in place through Commissioner Bradley and Mayor Giuliani, and he was the person at the head of the table that set up CompStat. I became sort of nutty about it. I've been to CompStat probably 150 times and watched how they've done it. If you go to CompStat you see the magic of what can happen if you really set a very simple strategy in place. What you would look at is the three D's of management. One is that they defined their goal. Their goal was very clearly crime reduction and that's what everything depended on. It didn't matter whether you arrested more people, whether you seized more kilos, whether you worked here or worked there, the whole thing was based on a very clear objective goal. People went crazy when they first came out with this, saying, "How am I, in the 40 Precinct, possibly supposed to reduce crime?"

He said, "The second thing is that you need timely, accurate data to be able to make your strategic decisions on what crime is and where it's happening and how it's happening. That will help you anticipate crime and be more effective at reducing crime." The third principle was deliberation. There has to be a process that once you define your goal, once you get timely and accurate data, where everybody comes around a table, all the people who are involved in the process, and are held accountable on whether we are succeeding or not succeeding. That's really what I think is the essence of the CompStat process. It's what we're looking to do and we've been doing in Criminal Justice under the governor's leadership.

From CompStat to CrimeStat

If you've ever been to CompStat, I'll give you an example and there are thousands of them. But a typical example, one that I saw, was there was this spike in robberies in the area of the 40th Precinct. Their goal was very clear. Everybody at the table knew it, from the police commissioner to the police officer on the street. They knew this was the goal of crime reduction. They had data that told them there was this spike and it was up 30 percent over a 60-day period. They then used timely, accurate data. The way we used to do it in Criminal Justice, we would get our data at the end of the year and report back and say, "Isn't that interesting, robberies are up or robberies are down." CompStat used timely, accurate data every week and now we do this in Criminal Justice at the state level. You can use timely,

accurate data to see how you're doing so you can then use that information to make proactive strategic decisions. In this example, they saw that the robberies were up and they used the data and they mapped the data and they put it up on a big screen and said, "Look at the 40th Precinct. There were 100 robberies. Let's put up on the screen in red dots where those robberies occurred." They could see there were maybe 100 dots but they were all clustered in a particular area. Then they said, "Let's look at what time they occurred." They found that 80 percent of the robberies occurred between 8:00 P.M. and 2:00 A.M. on Friday and Saturday night. There was a vulnerable spot and it was by a subway station. Then they said, "Let's look at the anti-crime officers that could help." That's a tool that the police have to solve robberies, anti-crime plainclothes officers who can blend in. When they reviewed who was available, perhaps they found these staff worked Monday through Friday 9 to 5. Maple said to the police commander, "What's your plan?" The commander said, "That's when they work. They've worked that for 10 years. That's their shift. They're 9 to 5." Maple said, "Well, you tell that to somebody who gets robbed at gunpoint at 9:00 P.M. on Friday night. Our job is to protect people. Our measure is crime reduction. That's what it's all about." That kind of transparency, so focused on what the goal is, on what's working and what's not working and what the answer is. If you get this timely, accurate data, obviously, then the answer just stares you in the face. It is what I think makes all the difference. If anybody has ever been to CompStat, you see this is really the essence. There's no magic as to why crime keeps going down. Particularly in New York City, there's no magic with all their processes. Every Thursday from 8:00 A.M. to 11:00 A.M., they sit and they drive through the data and look at the crime.

That I think is a great model and I think it's such a great model that every time I've been to CompStat, you see people in the back rows. There are two or three rows of visitors. People come from places like South Africa, the Illinois State Police, and the San Diego Police Department to watch this process. And that's what we're doing in State criminal justice. We've actually set up a whole CrimeStat Bureau that is driven by the three D's. We have to define our goal, which is crime reduction. We have to get timely, accurate data to tell us how we're doing. We have to have a system for deliberating on how we're doing. When we started this process, we didn't even get crime statistics from a lot of departments until the end of the year. When you think of the money we're putting out there for criminal justice initiatives, it's too late to know that crime rates have spiked in Niagara County after the year is done. We need to know in a timely, accurate way so we can then apply the resources to be used as effectively as we possibly can. The process that we use is starting with the goal. We have our clear goal, which is to reduce crime. That is what we look at every day. It is now a condition

New York State Criminal Justice System

Reduce Crime and Improve the Effectiveness of Criminal Justice

Selected Performance Measures

Crime Rate & Volume	Local Jail and Prison Operations	Offender Re-Entry	Community Supervision
-Murder -Rape -Robbery -Aggravated Assault -Burglary -Larceny -Motor Vehicle Theft	-Prison Population -Jail Populations -Complaints Received -Assaults on Staff - Suicides/Deaths -Facility Inspections Conducted	-Inmate Documentation -Program Participation Rates -Employment Rates -Reconviction Rates	-Successful Completion Rates -Program Outcomes -Absconders at Large -Violation Processing Time -Rearrest Rates
Crime Victims Services -Claims Processing Time -Claims Assignment -Domestic Violence Training Sessions	**Criminal Aliens** -Deportation Proceedings Held -Deportation Processing Efficiency -Rearrests of Previously Deported Aliens	**DNA Databank** -Processing Time for Collections -Number of Hits	**Sex Offender Registry** -Registration Processing -Address Verification
Grant Contract Processing -Average Processing Time -Site Visits Conducted	**Traffic Safety** -Speeding Tickets Issued -DWI/DWAI Arrests -TraCs Implementation	**Local Information Sharing** -eJusticeNY Implementation -Timely Crime Statistics -Registering Investigative Targets -Intelligence Center Participation	**Law Enforcement Services** -Fingerprint Processing -Rapsheet Accuracy -Equipment Repair -Law Enforcement Training

February 7, 2005

of funding for police to provide their crime statistics within 30 days after the close of the previous month. We now have 85 percent of our departments reporting in a timely way and we're going to keep at it.

Second is that we have literally hundreds of secondary indicators, which appear on the handout (see above). If you look at this sheet of the criminal justice system, we use numerous indicators. There are secondary indicators, which we believe will either directly reduce crime as a result of doing things better than we were doing before or they will improve the services that we provide. Some of them, just an example, are that we look at something like the processing of DNA. DNA is one the most fantastic tools that we have in criminal justice. We have these unsolved crimes where you collect these DNA samples for rape, murder, and other kinds of serious violent crimes. Now in New York City they are expanding to property crime. You have this hidden key or Rosetta stone in science. You have these collected samples but you don't know whom it's connected to. Then offenders get convicted and they are required to submit a DNA sample as a result of that conviction or for a certain category of crime. It's absolutely critical that you take the offender's sample and you start to match it against the unsolved crimes and that the process be as quick as possible. Already, there are over 1,500 hits on the database and the number of hits will continue to increase as more samples are run against it. It is absolutely critical that time be as short as possible because you can solve the crime earlier or you can also prevent

someone from committing a future crime. So we want processing time to be as quick as possible. By putting a spotlight on it we are saying that the DNA secondary indicator is important to us. We've gone from 320 days to 45 days in processing time. It's just a simple management strategy.

Another thing that we looked at, for example, is how quick are we on the service side? When someone is a crime victim and they submit their claim to be paid by the state, it's a matter of service to crime victims. How quickly do they get paid? It used to take about 220 days for a crime victim to get their claim paid. Now it takes 93 days. Each month, we're obviously trying to make it even better. We measure indicator after indicator, all leading toward the ultimate goal of crime reduction.

Then what we do is in our deliberation process, we have two different tracks that we go on. One is that we go within the safe criminal justice community with strategies that deal with each one of these different areas. For example, one of the areas that we're really focusing on as a matter of public safety and as a way to further reduce crime is offender reentry, focusing on the people who leave prison and go back into the community. Our goal is to measurably reduce the number of offenders who come back to prison with a new crime as a matter of public safety. The 26,000 people who leave every year, we want to reduce the numbers that return to prison. We want to increase the number that go on to live law-abiding, productive lives. So how do we do that? One thing that we realized is that's not just a Parole challenge. We have engaged Mental Health, and the treatment of people from OASAS (Office of Alcoholism and Substance Abuse Services), Labor, the Office of Temporary Disabilities and Assistance — twelve different agencies at the commissioner level. In fact, we're meeting with them this afternoon. Twelve different commissioners sit at a table and say, "Our goal is to measurably increase the number of people on parole who successfully complete parole." That's our crime reduction goal. Now we have all these nontraditional criminal justice agencies, everybody looking at the same map at the same time, looking at the same goal and saying, "How can each of us contribute to that?" One way to contribute is, for example, to look and see whether offenders coming out of prison are Medicaid eligible. There is a period of time that they come out and they don't have Medicaid benefits. We work with other agencies to figure out a way that you can close that gap so that people coming out have their Medicaid benefits in place, and can get quick access to treatment and other needed services.

In addition, we look at how many people are leaving prison with a birth certificate. That document is absolutely critical for them so that they can get a job, get proper identification, and do everything else that they need to do. So we have all sorts of indicators, but all of us in government are looking at the same map at the same time, talking about reentry. We're doing that at the same time as a collaborative strategy, whether it's related to prisoner reentry, fugitive enforcement, or reducing auto theft. Category after category, everything we do is a collaborative effort. We have meetings where everybody sits around the table, use timely, accurate data, and then figure out how we can do better at what we do.

Operation "IMPACT"

The second thing has to do with our CrimeStat process. The governor announced in the last year a program called Operation IMPACT (Integrated Municipal Police Anti-Crime Teams) where he identified that in New York State, New York City has had dramatic crime reductions. The rest of New York State has also had dramatic crime reductions. But there are certain areas in New York State where crime has not gone down as dramatically as we think it could. Those 15 counties outside of New York City, which account for 80 percent of non-New York City crime, include Monroe County, Erie County, and Albany County. In each of those areas, we set up a partnership with federal, state, and local law enforcement led by local police to come up with strategies again with the same principle — the goal is crime reduction. They have to get timely, accurate data to see how we're doing. We have to get together and see whether or not it is effective. We have just started, over the last five months, a process where the police departments themselves come in. For example, in Dutchess County, the chief of the Poughkeepsie Police Department and the district attorney of Dutchess County come to Albany. We have literally 50 people at a table talking about crime in the city of Poughkeepsie and what we can do to measurably reduce it. How are we doing? Going back to Maple's example of where the robberies are. What's happening? What can we do together in criminal justice to reduce it? We've had Dutchess County, Albany County, and Onondaga County come in and next month Schenectady is coming in. Orange County is coming in. One after another are coming in for this collective process of looking at the same map at the same time to produce results.

The Bottom Line

Overall, this is sort of an "under the hood" look at what we're doing in criminal justice. If we step back for a moment, and I conclude with this, what is the bottom line? Looking at the

governor's record over last ten years, what is the bottom line? The bottom line is that crime is down over 50 percent. The bottom line is that if you compare the total number of crimes in New York State in 1994 to 2004, there are 400,000 fewer crimes in New York City in just that ten-year snapshot before and after. If you compare the number of violent crimes in New York State from 1994 to 2004, in 2004 there were 86,000 fewer people who were victims of violent crime. That's 86,000 fewer people robbed at gunpoint, raped, murdered, and seriously assaulted. Over 1,000 fewer people who have been murdered. These are remarkable bottom-line accomplishments. In addition, at the same time, prison population in New York State has gone down in the last five years by 8,000, which is completely inconsistent with traditional criminal justice thinking that in order to reduce crime you have to arrest more people, put more people in prison. Actually, while the national average goes up in prison population, New York is down 8,000 and falling. Third is the recidivism rate, which has been cut by a third. I think that's because of the governor's results-oriented strategy of looking at the bottom line to keep people as safe as possible. Whatever it takes, that's what we're going to do. I think that's why we've had such great success here in New York State. Thank you very much for the opportunity to talk to you for a few minutes and I wish you the best of luck.

Performance Measurement in an Operational Agency

Robert J. Fleury, NYS Office of General Services

Good morning and thank you for the opportunity to be here this morning. It is a great honor. I'm going to start off by diverting wholly from my notes by saying I learned so much just by listening to Dr. Nathan and Mr. Parker so far this morning. Let me tell you a story about learning. In mid-winter 1978, there was a young man picking up cigarette butts on the top of a mountain in southern Italy. He was an enlisted man in the Air Force. He decided that wasn't what he wanted to do with his life. He recognized that the World War II/Korean War GI Bill was about to expire. So, he said to heck with this. He took advantage of that GI Bill, and he's got a lot more gray hair today, but he happens to be here with you this morning, still learning.

I'm happy to be here representing our new Acting Commissioner, Dan Hogan, at the Office of General Services. I want to bring this down to the thousand-foot level if I can. You've heard a great deal already this morning about the analysis, the complexity, and the very complex questions that are involved when you look at performance management or

performance measures programs. At OGS, we're a very decidedly operational organization. Our performance measurement program reflects that. We started it back under Commissioner Seymour, and continued it through Commissioner Ringler and now with Acting Commissioner Hogan. The credit and the accomplishments that we've made really go to the staff and to the people who do the work every single day. I'd like to introduce to our audience today Rebecca Meyers, who is our Director of Organizational Effectiveness and who is responsible for the performance management and measurement program at OGS.

What I'd like to focus on this morning is that we are a better-managed organization than we have been. We have the data to show it. The message that I'm going to deliver to you is that it takes a great deal of discipline and a great deal of work in order to initiate, to follow through, to have a successful program. Unless you are doing this throughout the entire organization, unless your data are reliable, unless every person who is working with it does in fact believe in it, you're probably not going to have great results out of it. Like I said, OGS is a decidedly operational agency. We have a diverse mission. We build things. We fix things. We maintain things. We provide things. We're not here to initiate or advance a political or social agenda. We're here to provide the services that let others do that successfully. We have one of the largest real estate portfolios of any organization in New York State barring SUNY. We have one of the largest Heating, Ventilation, and Air Conditioning (HVAC) plants in the world operating right here in the city of Albany. We operate campuses, which include the Empire State Plaza and the Harriman Campus. We have over 60 major facilities across the state. We produce over 2,000 contracts for commodities and services that your municipalities use every single day. The road salt that goes on your roads today probably came through an OGS contract.

So you can see the measurement challenges that we face in such a diverse agency. There are many ways in which we've changed, including that we are smaller. We lost 12 percent of our workforce during a period of 90 days in early 2003. We're much more computerized, if I can use that word. We're better integrated, far more than we once were. As I said, we started working on performance management back in 1998. When we started, we did a little internal review, a survey or mini-review. It told us that we had a number of glaring holes in the management of our agency. Even though we weren't falling down on the job by any means, we didn't have a unified vision. We didn't have clear goals. We had relatively ad hoc performance information. We decided that we had to fill those gaps. I think we've gotten some dramatic performance increases. Some of you are going to be more interested in methods than about how we got there. Others are going to be more interested in the bottom

line. I'm going to try to start with the results and work backwards. I'll get to a discussion of the methods.

We've successfully improved our performance results in many areas. I'll mention a couple of them. We measure building conditions. We also do an annual survey of the people who occupy our buildings. If you work in a state office building, it has shown up on your desk. We've asked you to respond to a number of questions. What do we do with that data? We analyze it and we try to understand what are the good things that are going on in the buildings and what are the bad things that are going on in the buildings. If you work in a state building, you will have noticed that there is increased air volume moving through the men's restrooms. That's a direct result of that survey. We're getting good feedback on it. That's a minor, but important, point. Annually we summarize for our own selves, not for external consumption, what we believe we have accomplished. One of the important things we note, and one of the things that is hard to continually focus on, is the areas where we have opportunities for improvement. Nobody wants to admit that they are not doing the best job they can, but we put a lot of emphasis on this.

So our brand of measurement is derived directly from the fact that we are an operational agency. We're the ones who keep the buildings running. We're the ones who keep the power on, heat on, and hopefully nice and clean for you. The general public simply assumes that this is going to happen. They're not concerned about whether or not we're measuring. So our measurement program is inwardly focused. We understand what the expectations are and we look at them and try to analyze them and ensure that our understanding is complete. But we know that the visitors who come to the Empire State Plaza to look at the modern art collection are not measuring us. What's important though is that the public does want to know that when they go in to look at that modern art collection, they are going to go into a clean, safe, handicapped-accessible environment. Those are the things that we focus on.

A Tool in the Toolbox

The basic question for us is getting our job done. Our initial principles set up this approach. Back in 1998, there were a number of things that we learned as we started to devise and plan this system. We knew that from an operational perspective, an internal perspective, dealing with our people, there was going to be natural resistance. So we worked very hard to diminish that by first engaging our executives. Our executives were continually presenting to the agency, to its employees, the importance of the rationale for what we hoped to derive from

this system. Parenthetically, I want to point out performance measurement is not the be all and end all. It's a great thing. It's a great tool, but my perhaps limited view is it is only one leg of the three-legged stool. If you don't have a strategic plan and your goals aren't clearly defined, if you don't have business plans that tell the parts of your organization what they are going to try to accomplish, measuring it isn't going to get you anywhere. It is an important tool but it's one tool in the toolbox.

So we used our executives very effectively to push out the idea of the importance of performance measurement. Also, in constructing the plan we recognized that we weren't going to get it right the first time around. If we started shooting measurers, we weren't going to have any data measurers left by the time we got done. So the message had to be positive. Every time we would go into a performance review, we had to give encouragement. We had to be positive, regardless of whether we were ready to bang our heads against a wall or not. People react well to positive things and the negative things are going to be turnoffs. We were also counseled that success will not come quickly and will not come easily in the development and implementation of a performance measurement program. This is true even in an operational arena like ours, where there are concrete things that you can look at. Anyone who implements a program like this is going to take his or her own pathway through it. They are going to go pretty much on their own timeline. It's taken us about seven years to get where we are today. We found it very helpful to do our strategic planning on a 2 or 2½-year cycle. We found it very helpful to also do our business planning annually. Let me remark that we do that within our business units and we have five major business units within the agency. They each do their business planning on about an annual cycle. We also do our reviews of the performance of the business plans. Performance must be shown to be consistent with the business plan and with the strategic plan on a semiannual basis. This week we are going pell-mell into that cycle again. Rebecca and I and others will be doing six semiannual performance reviews during the remainder of this week and nine next week. I said before, and I will reiterate, that if you're here because you want to learn what it's like to go through the development and implementation of a performance measurement program, don't think of it as an easy thing. Rebecca and I will be working 15-, 16-, 18-hour days to get through these performance reviews because you can't do them and do the rest of your regular job at the same time. You have to put that extra effort into it. I'm also here to tell you that I'm convinced that it's worth it.

We've gotten through phases now where we've tightened up our measures. We've gotten through a phase where we realized, and Director Parker mentioned it, that the data that
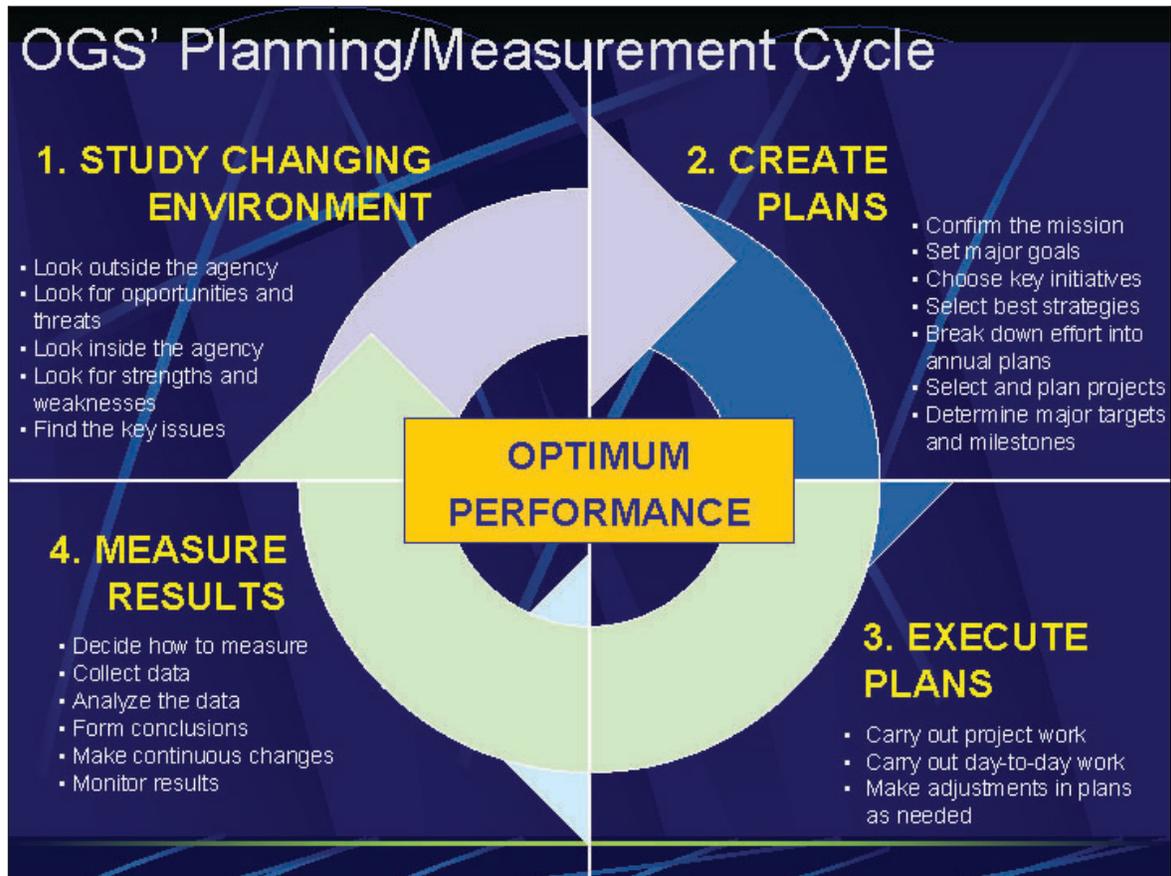
we were getting were tending to overwhelm us. There was so much of it. We have acquired and implemented a computer-based system to help us keep track of the data and to roll it up into measurements that the executives can look at (see next page). This is a nice summary slide that we use in OGS, which we developed in OGS from adaptations of other things, but it's geared to us. The key here is that there is no starting point and there is no ending point in this cycle that you see. Today in OGS, some of these things are happening all over the circle. In some cases, we're taking a look at the importance or the value, the validity, the viability of our measures and we're trying to reanalyze them. In other cases, we're working on refining and developing our business plans. In other areas, we're on execution. By the way, some of these things are not as minute or seemingly unimportant as the airflow in the men's rooms. We have a $90 million construction rehabilitation project going on right down the street here at the Alfred E. Smith Building, which is part of one of our business plans.

We're measuring results. Our schedule right now, it's mid-February and we're going to get through our semiannual reviews by early March. Next, people will be thrown into the business planning, and we will take a look at refining our strategic planning.

Project Management

Where are we going next? We've developed the ability to collect data and measure it, analyze it to drive and support our strategic vision. We've developed the ability to better link the data systems that feed this measurement cycle. We are going to improve the use of measures in managing projects as well as processes. My emphasis over the next year or two years is this: maintain and sustain performance management and performance measurement. But equally important, I want to bring into the picture another concept and that is project management. We have 1,654 employees in OGS. We have 1,654 different ways of approaching the management of a project. This is not rocket science. There are plenty of studies that have been done. There are models that have been developed. We're adapting one to OGS and we're pushing it out through the agency. So far, we've trained over 200 people in the basic principles of project management. We're creating a project management office. We will be coupling this effort with our strategic planning, business planning, and performance management. Project management is going to make us an even better organization, one that serves to an even greater degree the needs of our clients.

Lessons Learned

What have we learned? Here's a short list: It takes patience and commitment, and it can't be done overnight. The trickiest part is that three-legged stool: strategic planning, business planning, and performance management. Sound data are absolutely critical. They can feed you BS and sometimes they will try. You have to be watchful for that. Another important lesson is balancing. You can overdo it. You've got to gauge your organization's readiness, to have a sense of your organization, what's it's willing to receive from you, act upon, and where it's willing to go. You must encourage people. There are no simple answers. There is a "complexifying" nature to the whole thing. But as a leader, the way you manage your organization will definitely change. Thank you very much.

Performance Measurement in State Taxation and Finance

Andrew S. Eristoff, NYS Department of Taxation and Finance

Let me preface my remarks by saying or admitting how incredibly nervous I am today to appear before you with all these august prophets. First I want to thank Dick Nathan for inviting me to participate. Why am I nervous? First, I was one of the bright, bushy-tailed

youngsters at Princeton that Professor Nathan sought to guide many, many years ago, back in 1984. I don't know if you recall, professor, but I was in one of your seminar classes. I'm not sure just standing here how well I did, but I think it all turned out okay in the end. Second, by way of preface to my remarks, although I may find tax administration to be sexy, I will understand if all of you don't. Tax policy is interesting in that everybody is a tax policy expert I discovered. Tax administration is what we do at the Department of Taxation and Finance and that's what I intend to focus on today.

The focus of my remarks will be to address three basic questions. First, how do we at Tax use performance measures? Second, has our use of performance measures and measurement been a success? And third and finally, do I have any profound insight to offer to those of you assembled here and to other public administration experts and to people in public administration?

First, how we use performance measures. Our use of performance measures in essence supports our strategic planning process. Back in 1997, Tax embarked on a comprehensive planning process that included the development of the mission, vision, goals, and objectives and performance measures. At that time, the overall product tended to emphasize customer service. In addition to periodic updates, this process includes an annual operational plan-developing process at the divisional level. The department's divisions provide and develop in essence a business plan for the coming fiscal year. Each of those plans contains specific goals and objectives that dovetail with our strategic plan and are related to performance measures that enable us to measure progress. These measures are intended to help us access productivity, identify and improve service levels, allocate resources, and ensure accountability. Six key measurement areas include accuracy, timeliness, completed versus scheduled cost, customer perception, and employee work life quality.

The examples, I think, of our historic use of performance indicators should be fairly obvious. We use measures to manage our-state-of-the-art taxpayer and collection call centers, to reduce waiting times, allocate resources, adjust hours, and match employee skill levels to caller issues. In a similar vein, we use measures to allocate resources to reduce processing exception inventories. An exception in processing occurs when a return or other submission by a taxpayer kicks out because there is something wrong with that return. We've use measures to assess how we're doing in reducing these stacks of inventories of exceptions, and we use statistics like time-to-clear in order to help assess performance. These tools have enabled us to reduce our backlog of exceptions in the withholding tax area from more than

three years at one time to about three months today. Finally, we use measures to assess our audit programs focusing on output type statistics like the number of audits completed, revenue assessed and collected, and revenue per audit hour. Similar statistics inform our collection operations.

Now, the hard question. Has our use of performance measurement been a success? Well, I would say success is a function of expectations. In an academic, tactical sense, performance measurement is indeed working for Tax. It's embedded in our culture. Nevertheless, it remains selective and its coverage and impact is operationally incomplete. Why is that? Let me offer a hypothesis. Our experience in my judgment reflects our management expectations, which in turn reflect our traditional focus on managing risk. Each public entity has different functions and needs that necessarily inform different expectations of performance measurement. In our case, the fact that we are largely a financial widget factory — an administrative agency that processes millions of tax returns and billions in payments each year — means that we have historically tended to emphasize the management of operational risks, were they financial, political, and/or customer service. We have tended to emphasize the management of risks over program performance per se. Indeed, I would say that historically we really have not had a working definition of "program," let alone "program performance," to work with. Accordingly, to date, we've experienced more success in using measures to manage risks than to improve program performance. To use the Nathan Paradigm, as a widget factory, we've tended to measure inputs and outputs, not outcomes or impacts.

So, if success is a function of measuring expectations, how do we change expectations? That's what we're working on now. We've just completed a strategic plan three-year update and that update included normal updates and refinements to the existing plan. But in retrospect the update became a vehicle for resetting and rebooting our expectations away from a myopic focus on managing risk to one that includes managing program performance.

Supporting and Promoting Voluntary Compliance

That change is manifest in identification of an overall strategic priority for the department: For the first time, we have identified the support and promotion of voluntary compliance with the tax laws as our overall strategic objective. Our performance measurement program will now have to focus on supporting voluntary compliance. How would you do that? Yes, we can measure how much revenue we collect. To look at the statistics, you'll find about 98 percent of the revenue that the state collects comes in

The Compliance Continuum

The Compliance Continuum illustrates two important points. First, the cost of tax administration and compliance --- for both DTF and taxpayers --- generally increases as one moves from left to right. Criminal enforcement against a minority of taxpayers is critical but an inherently inefficient means of raising revenue. Second, activities on the left side of the continuum directly impact a much larger number of taxpayers acting voluntarily, and hence are more efficient.

To meet the strategic goal that all taxpayers voluntarily pay the correct tax timely, *DTF should pursue activities that move taxpayers "up" the Compliance Continuum to achieve greater voluntary compliance. This conceptual framework should help us determine our priorities on both an enterprise and a division level. DTF should always seek to move taxpayers "up" the Compliance Continuum to achieve greater voluntary compliance.*

"voluntarily." The rest comes through audit and collection and criminal enforcement activities. But voluntary compliance is extremely and notoriously hard to measure. It is often talked about in terms of a tax gap. How much are we really collecting in relation to how much we ought to collect? That's a very, very difficult concept to actually plot upon. I've come to believe that our approach must be like that of an astrophysicist who deduces the existence of an unseen planet by using the effects of their gravitational pull. We've got be more creative in how we try to assess what is really going on.

One approach that we're using involves the use of what I have called a "compliance continuum." That is the handout that I've offered today (see above). Let me just give you a brief run through, and forgive me if this is a little bit on the theoretical side but I believe it does help put things into context. Tax engages in a wide range of activities that directly or indirectly promote voluntary compliance, stretching from activities that occur before the taxpayer begins to fill out his or her return all the way to enforcement of the tax laws through criminal prosecution. We think of it as a continuum.

On the other side of the ledger, taxpayers have interactions with Tax that cover a broad range of parallel activities from reviewing forms and instructions, filling out those forms, through potential litigation and obviously criminal defense. The cost of tax administration and compliance for both Tax and taxpayers generally increases as one moves from left to

right. Criminal enforcement against a minority of taxpayers is critical, but it is inherently inefficient as a means of actually raising revenue. In a similar vein, activities on the left side of the continuum directly impact a much larger number of taxpayers acting voluntarily and hence are much more efficient.

Now, if you accept the compliance continuum as a conceptual framework that accurately describes what Tax does, its functions, then you will understand that it is in essence our "program." Thus, improving "program performance" means moving taxpayers up the continuum to achieve greater voluntary compliance. That is the desired impact or "outcome," if you will. In terms of performance measurement then, we need to identify and use measures that will help us move taxpayers up the continuum — measures that speak to the impact of audit and criminal enforcement initiatives on taxpayer behavior, such as before and after snapshots using control groups, if you would harkens back to Professor Nathan's comments. In addition, we need to use measures that help us identify and reduce processing exceptions, not just resolve them quicker, for instance, call center monitoring and surveys in order to help identify training gaps. Experimenting with changes to forms and instructions to see what the impact would be on the number of exceptions and processing problems that are created as a result. Measures that help us focus audit and collection resources, e.g., no-change audits — in other words, the number of audits that we performed that result in no additional assessment on the taxpayer. This is a key measure for us because there are different ways of reading the data. But clearly if we have a no-change audit, it represents a waste of our time and the taxpayer's time. In general, we would like to see less of those kinds of audits and more audits that are productive and ultimately reach a more highly targeted group of taxpayers for whom compliance is in fact an issue.

This is a continuing process. I wish I could stand before you today and claim that we've completed the process. But again referring to Professor Nathan's paper, it is in fact a dynamic process. It is in fact highly complex and it will never be complete. Because we have been going through this internal reevaluation of how we use performance measurement, frankly we're just at the beginning of this next — I hope higher — level of consciousness with respect to performance measurement.

Don't Get Hung Up on Numbers

Let me close with a couple of what I call "profound insights." I suspect you may not find them very profound at all, but it helps me try to work out my thoughts. First, don't get hung up

on the numbers per se. Building on Professor Nathan's point, the numbers you pick may in the end not be all that important. Of course it's great if using performance measurement leads to documented higher performance, but I think that there's real value in developing an organizational culture that supports the process of identifying, providing, using, and experimenting with measures even if you can't back it all up with pseudo-scientific data. Of course, don't underestimate the value of measures as conversation starters. This is critical and I would say CompStat at its very core is an example of how measures are used as conversation starters.

Define Performance for your Organization

My second, and last, profound insight is that your first task as a public manager seeking to use performance measures is not to identify the measures, but instead to define performance for your organization as we have been attempting to do at Tax. Each public organization obviously has a distinct definition. You need to decide what performance is for your organization and then seek to identify and use measures that support that definition. I would advise against letting performance be defined by default or by outside factors. For instance, I happen to believe that Tax for years defined performance by default as "risk." Basically, it was a default setting. We have to raise the revenue and we don't want to let anybody "get away with it." I think that it's frankly more complicated than that. An example of defining it or allowing outside actors to define performance would be tailoring all your activities so that you can qualify for federal funding. I'm not a social scientist and I have little experience in programs that are funded through the federal government. But my understanding is that there is an awful lot of dovetailing of performance and activities in order to maximize your federal funding. I don't think that is necessarily reflective of what really are the core indicators of performance for a given agency. All of this, it has been said many times over, is harder than it sounds. But I truly believe that trying is of great, great value. I wish this was original to me, but I think that it bears mentioning that you are or will become what you measure. There are many different ways of saying this but I think it's true. Start measuring something, and you will have an impact. I think that there is another side to that observation and that is that your choice in measures will have an impact upon your organizational culture. You will morph into something that reflects what you measure. That is an important thing to keep in mind as you decide your performance measurement approach.

Thank you very much for your indulgence. I'm very pleased to have this opportunity to make a few remarks. Again, I'm surprised that you would invite a tax administrator to participate, but we're honored and thank you very much.

Discussion

Question: *I'm really impressed with what I heard today. A few years ago I had conversations with Dall Forsythe, then budget director for the end of the Cuomo administration. Dall, of course, did the book that Dick Nathan was talking about. We've come a very long way baby. It really is impressive. I'm sure there is a lot more to do. I had a question about criminal justice. I was very impressed that you're working with Dutchess County and other places, but I had that New York City sort of centrism in my question. How is this working in terms of New York City? For example, those wonderful statistics about reduction in crime in the state, what percent of that reduction of crime are New York City crime reductions? And how does the state compare with that? Can the city and state, using performance management tools and the approach you described really well today, have the ability which is better than has existed in the past to work with the city? Because that is a huge challenge and good people who have worked in the state have felt that the City is just hard to deal with, and then have moved to positions in the City and find that the state is hard to deal with. How is it playing out there because it seems to me that there is a very interesting story beneath the story that you presented?*

Chauncey G. Parker: The crime reduction in New York City has outpaced the rest of the state. I think it's probably twice as much over the last 10 years within New York City as the communities outside. There's really been a very dramatic shift. In the early 1990s about 65 percent of crime was in New York City and now its now 55 percent outside of New York City. So, there's been this shift. The crime numbers are down but the percentages has shifted. I think the key of it is that the goal is crime reduction. Everybody shares that same goal whether you're in New York City or you're in Schenectady or you're in Poughkeepsie or you're in Albany. They all have to goal of crime reduction. It's a lot trickier, I think, to do it outside of New York City because you have 500 police departments. NYPD is a 37,000-person organization. Our relationship with them is excellent. We've collaborated on a lot of different issues. But they have a management tool for their organization. They really are the premiere, dominant police force in New York City.

Outside it becomes trickier because you have a lot of different people, sheriffs, chiefs in towns and villages. But you can bring them all together with this common goal, which is they all want to make their communities safer tomorrow than they are today.

Here's an example of something that we're doing with New York City but we're also doing outside. The governor announced an initiative called New York's "100 Most Wanted." We all in law enforcement are looking for fugitives. NYPD has their fugitives, marshals have their fugitives, Buffalo has their fugitives, and everybody has their fugitives, but they are all looking for their own fugitives. But the governor says why don't we put them on an integrated

list? Let's take the 100 worst, the most dangerous, put them on a list, put them up on a website, and everybody look for the same individuals that they're suppose to be looking for anyway. Let's have the eyes and ears of not just one particular department but the eyes and ears of 75,000 police officers looking at this. In the first 30 days, nine people were arrested. It worked exactly as he had envisioned it would work. Somebody from Troy who was wanted got locked up in Brooklyn, which they would have never known about but they saw him and put two and two together and saw that they were treating them with geographic asylum.

It's kind of remarkable to think that the Poughkeepsie police chief and the DA would come to Albany and sit at our CrimeStat table to talk about crime reduction. But they came and they've absolutely bought into the idea that crime reduction should be the goal, and we're all in this together and so let's figure out how we can do it. That's a little bit longer to answer your question, but it has really tied together a relationship with the City. These other departments are coming together in a way that I think they've never come together before.

Question: *Mr. Eristoff said that good measurements can be good conversation starters and that reminded me of what Commissioner Mills always says about the school report cards. He wants them to prompt school boards, parents, and teachers to get into conversations about what the numbers mean and what people can do to improve them if that's indicated. I want to ask a question about the difference between inside versus outside use of performance measures. We heard a lot today about mostly inside use and very good use of statistics by internal management in the agency and the police department and so on. My question is should the state do more in terms of outside use of statistics? Should we have a state-level version of the Mayor's Management Report of the City of New York so that numbers are reported to the public annually or whatever time period to prompt more conversations around the state about many of the issues that confront us and what the numbers tell us about those issues?*

Andrew S. Eristoff: I can take a stab at it. As a former City commissioner, I have an opinion or two about the Mayor's Management Report and the process. The short answer is yes, I think it would be great if we had a state-level equivalent to the Mayor's Management Report. However, I think that there are some additional points that need a little bit of elaboration. First, the structure of the Mayor's Management Report process isn't serving quite the original intention. It was supposed to inform the budget process. It was supposed to basically provide the City Council and advocates with statistical information and policy and program information that could then be factored into the deliberations on the Mayor's proposed budget. So there's a mid-year Management Report and then there's a final Management Report at the end of the fiscal year. That portion of it really doesn't work. However, there is a portion of it that I think really does work and that is that the MMR is a great tool for forcing reexamination of internal operations inside an agency. The very fact that we had to produce this thing, a big pain in the toches by all accounts, was an important factor not only in creating an atmosphere of accountability but in reminding everybody that performance did count and that somebody in fact was watching. So the process of assembling statistics, packaging the statistics, or defining new measures was extremely, I think, helpful. It goes back to the point that you picked up on, which was it was all a very, very useful

conversation starter. But it was a little stultified and it didn't serve the budget process the way I think the original framework might have expected.

Question: *But I think the issue of having the data available externally raises the problem of shooting the measurer. If you want to protect the people who are providing the information to a certain degree, it becomes very difficult to do that when you lose control over who's reading the information. Do you have any wisdom on that?*

Robert J. Fleury: I would also like to point out too that the data, to the degree that they are collected and used within an agency, are readily available to the public. They may have to know that they want to ask for it and perhaps that's another question. I don't know the answer to the question, should we have a statewide report, because I think it has a number of nuances. Rebecca has touched on one of them. Another one is, we want to use the data in OGS to compare ourselves to others. For instance, the Building Owners and Managers Association develops data on appropriate and efficient use of office space. So we develop information and we compare it. We're comparing it for a particular purpose, to try to get more efficient use out of state-owned and state-leased office space. When I have someone external to the agency and external to government doing a comparison for another purpose, whatever that might be, do I get the same result? I don't think so. I think you're going to get different results. You're going to get different answers. You're going to develop comments, criticisms, whatever it might be, that aren't focused on the original purpose of the development of the data and the analysis done to it. I am not advocating withholding the information. We are government. We're accountable to the people and we need to be forthcoming with our data. But I am saying that there are purposes for which we've developed these data and we need to recognize that those purposes can be driven astray by analysis that is not consistent with its original design.

Andrew S. Eristoff: There are two sides to this. One side says if you really want to use performance measures to improve your internal operations, you don't really want to go public because the data can be misinterpreted. The press will focus on the dirty laundry and the whole thing spins out of control and you won't be able to actually use it to drive performance in your agency. On the other hand, there are many of us who feel that the very act of publicizing performance statistics does two things. One it basically sends a chill through your organization to get it right so that the data get cleaned up but quick. I have on occasion used this. It's a bit of a bludgeon, but I have used it. When you threaten to say I think we should go public with this information, it's amazing how staff gets their act together and gets the right data so they can't be misinterpreted. And the other thing I really feel strongly about is that we shouldn't be ashamed to tell the truth. I mean not everything we do at Tax is working absolutely perfectly. But if we spend all this time and energy covering up or trying to stay out of the limelight and stay out of the line of fire, I think frankly that's a misdirection of our resources and energy.

I think the management report is an interesting example because the press can focus on the dirty laundry every year. They'll pick one or two things that are lousy and then they'll also go on. Every year, it's the standard story. They'll say, "And the Mayor," whoever the mayor is, "is presenting these statistics in a way that is designed to puff up his administration." So

they'll discount it for politics and political influence. Frankly, for most agencies the process becomes ho-hum. The statistics are out there and in the end, except for some very rare exceptions, the press and the public don't really zero in on the stats. Instead, you do get questions from the City Council and questions from outside analysts and folks like that. It ends up not being quite as disastrous from a PR standpoint as many agency managers would anticipate. I think that if you've got bad news, why not just go up front and be up front about it, get it over with instead of pretending it doesn't exist because it won't go away! It's like a toothache, you've got to deal with it when you have it or otherwise you're going to wind up losing a tooth.

Edward Ingoldsby: Certainly, what I would say is these are three of our leading agencies when it comes to performance measurement and that's why they were selected to participate today. If you look at the Executive Budget and the materials associated with the Executive Budget, you'll see throughout the presentation that this year in particular the governor has emphasized performance management in developing and advancing recommendations. It's from E2 municipalities where we've advanced a new program that emphasizes performance contracts, performance agreements between localities and the state director of the budget, and to the welfare area where we're handing out funding based on performance in food stamp performance and reviews. It's throughout the Executive Budget. David, is there any area you think we should mention?

David Kidera: I think we could talk a little bit about STAR (School Tax Relief Program) and post-STAR. The governor has proposed a special STAR program, which I think many of you may be familiar with, basically providing additional taxpayer relief, which I guess is a second theme throughout the budget. Additional taxpayer relief to homeowners in counties that keep their county's spending below the cap the governor is proposing on Medicaid. The same thing applies to school districts. If the school districts keep their spending below the rate of inflation, below about 3 percent on average a year, taxpayers will see an increase in their STAR benefit to the tune of like $48 million. So we're trying to link not just at the state level but also at the county level for Medicaid spending and at the school district level for school spending. If you can improve your performance in those areas in terms of controlling your costs, improving efficiencies, taxpayers would see a direct benefit in terms of an increase in their STAR benefit.

Edward Ingoldsby: Again, it's throughout the budget. I think I will just mention quickly that we are initiating a strategic planning process starting in a couple agencies. We're in exploratory phases there. So you will see some discussion of that in the coming year.

Question: *All three of you actually indicated some potential problems, whether explicitly or implicitly. For instance, implicitly you said the undercover detectives who have to change their work schedule in order to patrol the trouble spots. How do you change the culture, the climate, within the agency to make it more performance based and to deal with some of the concerns of people who work with it?*

Chauncey G. Parker: That is why I started with the NYPD. I think it is just setting a clear goal. In this specific example, an area in Brooklyn had a crack house. People used to ask

the police department, "How come your arrests are down?" This is how people talked for a long time. "Your arrests are down. Your arrests have to go up." Then you have a crack house in your particular area, you have a particular attitude toward that crack house. If you go to that precinct commander and say, "I don't care who you arrest, your car thefts are up or your assaults are up or whatever." What happened in Brooklyn is they actually figured out a way to close the place down. It was an abandoned building. All of a sudden, they're working with other City agencies and all sorts of other people who they would have never worked with otherwise to close it down because that's their strategy to reduce crime.

In terms of people, you watch any department in the NYPD, you've got Narcotics, the Detective Bureau, Patrol, and Housing, all the people who were vulcanized within the same department. They would sooner work with another department, even the FBI, than they would work in their own department because the measure was arrest and arrests are like goals in hockey. Only one person can score. So one person wants to be the arresting officer. In NYPD to change it to crime reduction is like all of a sudden saying, "Here are the new rules in hockey. The only way you can get points is by an assist." You would see people passing back and forth. I really so passionately believe that what gets measured gets done. If you set the right goal, you will change human behavior. In law enforcement, law enforcement has got to work together. The critical thing is to have it absolutely clear, use data that shows how you're doing, and then put people around the table and say, "How come it's up or how come it's down?"

Question: *I certainly agree with that, but I think that there's been almost too much emphasis on the number aspect. Your line people have to believe that your interest is in improving overall performance and they have to want to do it. If the data collection is to beat people over the head with, it's not going to work. Back in the early 1970s or 1960s, New York City tried to measure refuse collection and were delighted that it had gone up so much. But they weren't measuring how clean the streets were. They were measuring how much refuse was collected. Then they discovered that they were hosing down the truck to increase the weight that people were actually delivering. You can get a measure but your measures are never going to be that precise. It's kind of like my issues with changing the budget system. I basically don't believe that the system is so broken. Any system will fail if the people are not sincerely trying to make it work. So, I think part of getting people bought into it is making sure that they actually trust your goals. That they trust that you're not trying to make them look bad, you're not trying to figure who to lay off or how to cut, but that you're actually trying to improve the services that they joined, whatever organization they joined originally to promote. So I think the human aspect is a very important part of it.*

Robert J. Fleury: There are two things that need to be added to Director Parker's remarks. Clear articulation of a goal is critical. In order to change the culture, you have to communicate on all levels of the organization. It's going to start with the executive, but there has to be a communication plan that brings it down to every single level in the organization. Our organization is more diverse than many of yours. I actually go to staff meetings about planning and performance in the evenings where the predominant language is not English for some of our cleaning staff. I'll start out a conversation with them in Italian. You've got to communicate with them in a manner in which they understand.

The second thing that has to be added is that you have to provide positive reinforcement or reward. OGS is the only agency as far as I know that has the ability to make a monetary reward to an employee for success and we use it in a very elemental fashion. We have some people who have a hard time getting to work each day. We have a program that if they get work on time every day for a full year and also perform in their jobs well on a rating system, they get a bonus check. Do you know where our rate of employees not coming to work on time went? It dropped dramatically after we started that program.

Positive communication is also important. This is one of our employee newsletters. Right up on top of the front page, we've showcased statistics showing where performance measurement is working. We are showing employees the positive trends in tenant satisfaction with the office buildings we maintain. So there has to be communication. It has to be at every level. The executive has to be involved. The executive's management team has to be involved. You have to get down to that third level of management, people who are actually line supervisors and then you have to tell every single person in the agency, every single person in the organization. Not everyone will listen to you, but most will.

Dennis Smith: I just wanted to add to the director's comment about CompStat's role, I think we would miss a key part of that whole press to management if we don't see the idea of how it accelerated the process of learning about what works and what doesn't. If assists, which aren't after all the goal, didn't lead to winning more games then you wouldn't want to focus measurement on it. I think what we get out of this whole management performance/management revolution, if it works right, is a much faster feedback loop. Learning faster about what works and what doesn't and I think when that's communicated that changes the employees' behavior too. If they see leadership changing because of things that have happened on the ground, they take it more seriously too. But there used to be a disconnect. There were two cultures, where one culture thought they knew the answers but the people on the higher levels didn't care. CompStat brought precinct commanders and lieutenants and sergeants into One Police Plaza for regular conversations. The conversations didn't just start, they went on. When they learned something, they distributed that feedback very quickly and the places that have copied that feature of CompStat in parks management and other areas have really seen the payoff.

Chauncey G. Parker: I think that it's so important what you're saying and I think it has to work. Success is defined as much by anecdote and I firmly believe that. When Maple came in people thought he was crazy. They wanted him to leave because he believed that you could actually control crime and do all this kind of stuff. Then all of a sudden, the murders went from 2,200 down to 1,300 down to 900. This crazy guy showed them you can actually do it. It was a tangible, effective way that the message gets shown so that people see that. Then they buy into it and then that excitement becomes contagious. That's what I'm saying.

Question: *Just in simpler terms, when you have plainclothes policemen who are working 9 to 5 their whole life and then you turn around and tell them that they have to work from 8 until 2 in the morning, how do they really approach that?*

Chauncey G. Parker: I think that there's an awful lot who aren't happy. I don't think anybody likes change and there is some percent that believes we can change the world. But most people don't really like change. They like the world to go the way it is going. I think that people who are used to working 9 to 5 also don't like to have their schedules changed, but that becomes the leadership challenge for that particular precinct commander. He could not destroy the morale of the people who worked for him and rule with an iron fist and say, "From now on, you're all working this." You have to think of some creative, innovative way to get the will of the people who are actually doing the work and get the job done. Somehow they do that. Because if you break their will, you can get them working the night shifts, but they're not going to do effective jobs. Create an atmosphere to get them to respond to it, that's the leadership factor.

It is a lot easier said than done but I think that clear goals and accountability lead the best leaders to rise above the rest. If you look at a place like NYPD, the person who runs the entire operation next to the chief of the department, Gary McCarthy, he took those reins and he's only 39 years old. He went from a captain to a deputy commissioner in two or three years because he was an innovative and inspiring pro-active leader who could get the job done and not undermine the will of the people who worked there, but inspire them to do more than they were doing before him.

Question: *Another question for the speakers is what about unions? New York State is a state that is strongly organized in the public sector. To what extent and what ways do union leaders and things that involve unions come into the process of what you've been educating us about today?*

Chauncey G. Parker: The answer is they are an important part. I think unions are thrown up as an excuse sometimes as to why things can't be done. You can't do this, you can't do that because of unions, which I don't really believe at all. I think if NYPD or anywhere else makes the goals very specific then leadership and management can figure out a way to work within their relationship with the unions to get the job done. People connected with NYPD again have probably one of the biggest unions in the state. They have somehow figured out how to reduce crime 60 or 70 percent with a union. I think it's something that you work through together. It's always a factor, but I don't think it's an excuse not to get it done.

Robert J. Fleury: In OGS, we deal primarily with three unions: CSEA (Civil Service Employees Association), PEF (Public Employees Federation), and to a far lesser degree and scope NYSCOPBA (New York State Correction Officers and Police Benevolent Association, Inc.). In each case, we make a very deliberate effort to communicate with them, and involve them in the strategic planning process. Their members are represented in the business planning process and are represented in the performance measurement process. From an executive perspective, we meet with our local union leaders on a regular basis and discuss implications of the strategies that we're trying to push out into the agency. I think because of that we've developed a very good working relationship with the unions. I don't see unions as a problem. I see them as collaborators and people who can assist us. That's been true in our case.

Andrew S. Eristoff: I would generally agree with Bob. I have much more in way of issues with civil service than with public employee unions. We're blessed with an excellent, high quality workforce and we have a very good union relationship. Leadership that is responsive and helpful. I want to bring this back to this concept that I floated earlier, the compliance continuum. That's actually helped us in managing our relationship with our employees because it's allowed me to market a very simple idea, which is that all of us, no matter what function that we're performing at Tax, in some way impacts voluntary compliance. Everything that we do is on that continuum. This is a concept that I've tried to roll out in a number of different settings with different employees and different units and to a person they all get it. In many cases, they've come up to me and said, "You know, now I sort of understand how I fit in to this larger picture of what we're doing." That in turn has been very helpful in getting employee buy-in. The leadership of our unions has been extremely cooperative and supportive of that effort.

Edward Ingoldsby: Well, as Mr. Eristoff mentioned in the New York City example, it's difficult to link performance measurement with a formal budgeting system. It's been a challenge for many states and the federal government. We'll continue to emphasize it and look at ways to make it work. But clearly where performance measurement does work best is when on an agency-by-agency basis you have strong commissioner-level support for the principle of performance management. It comes from the top down and in that kind of environment as you have seen here, I think in the budgeting arena that the resources that are being allocated are being used in the best possible way.

Henry Olsen: I think this has all been very interesting and I can tell by the discussion that all of you found this very useful and for that I'm glad that we were involved. As far as the challenges are concerned, I think one thing is it can be easy to take a performance management goal system and turn it into an aspiration system. Instead of being something that is clear and discrete and consequently manages the day-to-day operations, it becomes something softer and fuzzier and aspirational, and consequently less effective That is a mission creed or focus creep that could afflict every organization, whether it's the private sector or the public sector. But the lack of competition is something that is particularly acute in the public sector, because if you have that lack of creep there's not going to be a Jet Blue that comes in and shakes up an entire industry with a new way of doing things. So, as you're going about making performance management work in New York State, I would just recommend to try and keep the focus and keep the concreteness, keep the specificity, so that two years from now you're not looking back saying, "Oh, I'm disappointed in the output or I'm disappointed in the accomplishments."

Part III — The Local Role

Introductory Comments — Swati Desai

I think it makes more sense for Dennis Smith to go first. Dennis has done a lot of work in performance measurement in New York City. In Dall Forsythe's book he has also written an article along with Commissioner Bratton on CompStat. Dennis will talk about is CompStatting the human services. It makes sense for Dennis to start on how human service agencies, three agencies in New York, have talked about and used data in performance measurements. Then I will talk specifically about our example of JobStat and how it has evolved. We have been doing JobStat for five years. Following that, Fred Wulczyn will talk about performance measurement systems and issues in the child welfare system. So I think at an operating level at least in New York City we have done a tremendous job in performance measurement and in managing performance measurement. To illustrate that, I will start with Dennis.

A Framework for Performance Management in New York City

Dennis Smith, Robert F. Wagner Graduate School of Public Service, New York University

Swati and I agreed that I would start this first because in a lot of the work that I have done on performance measurement as a scholar, I have taken advantage of the extraordinary work New York City has done creating and implementing performance management systems. Since there are two people here who have actually done that with city agencies, it seems that they should tell their own story and I would like to provide a more general framework of this.

On paper, New York City has had a performance planning and management strategy that provides for multilevel action since 1977. It provides that each agency will have an agency management plan. Those agency management plans will be broken down into the units of the agency in terms of what performance is expected from each of the units as part of a contributing factor in the performance of the agency. In addition to the agency management plan, there are monthly management reports on key indicators, reports to the mayor's office. Key people in the mayor's office review this to be able to see whether there are problems, to learn from what is being presented, and to see the broader picture. Twice a year, the results of

the whole process are presented in the Mayor's Management Report, the MMR that people were talking about this morning. I can tell you, having studied this since the late 1980s under the Koch administration, only now, in 2005, we are about getting to the point where that description is an accurate account of performance management in New York City government.

Dick Nathan has written extensively about the challenges of implementation. This is a great implementation story. When I was doing the study in 1989, there was a manual that described this whole performance planning and management system in the Mayor's Office of Operation, but nobody on the staff at the time knew about it. The system had started in 1977. The Mayor's Office of Operation had produced a report because it had been mandated. The Mayor's Management Reports (MMRs) that were mandated by charter were coming out twice a year. But the burden of my findings from looking at twelve agencies at the end of the Koch administration is that everybody reported because they had to. Mostly what they were reporting on were activities and inputs; not even outputs and certainly not outcomes. There was no focus on outcomes. More importantly, who cared? The agencies reported because basically they had to. In interviews officials usually used mild profanity in describing the Office of Operation mandated indicators reporting. They basically felt it was extra work and they didn't see it as critically relating to anything they really cared about. That was a surprise to me. I didn't believe it was as unimportant as they did. I could see how people in the City Council were using the data, although not as much as they could have used and now use it. In general, there is much more thoughtful use of this information by many actors than there used to be. But at the time, at the end of the Koch Administration, after the system had been in place for more than a decade, it really wasn't working as anything other than a low-level reporting system.

The basic insight for me was that you have to go beyond reporting on performance measurement to what I called in my first paper on the subject, performance management. This is a paper I did for an APPAM conference in the early '90s. I described first of all how I'd done this study, and I found that not much was being done to implement the management goal of performance management. I made recommendations, which by the time, as I said in the APPAM paper, I could report were being ignored by the Dinkins administration. They had little interest in the interactions that I think are key to performance management, the interactions between ideas, action, and evidence. The interplay between the elements of this triangle is a critical balancing act for getting performance management right. I watched with frustration as the potential for the collection of data and reporting of data was lost by several

MODEL C
THE COMPSTAT MODEL OF PERFORMANCE MANAGEMENT

Dennis C. Smith
NYU, Wagner School
2001
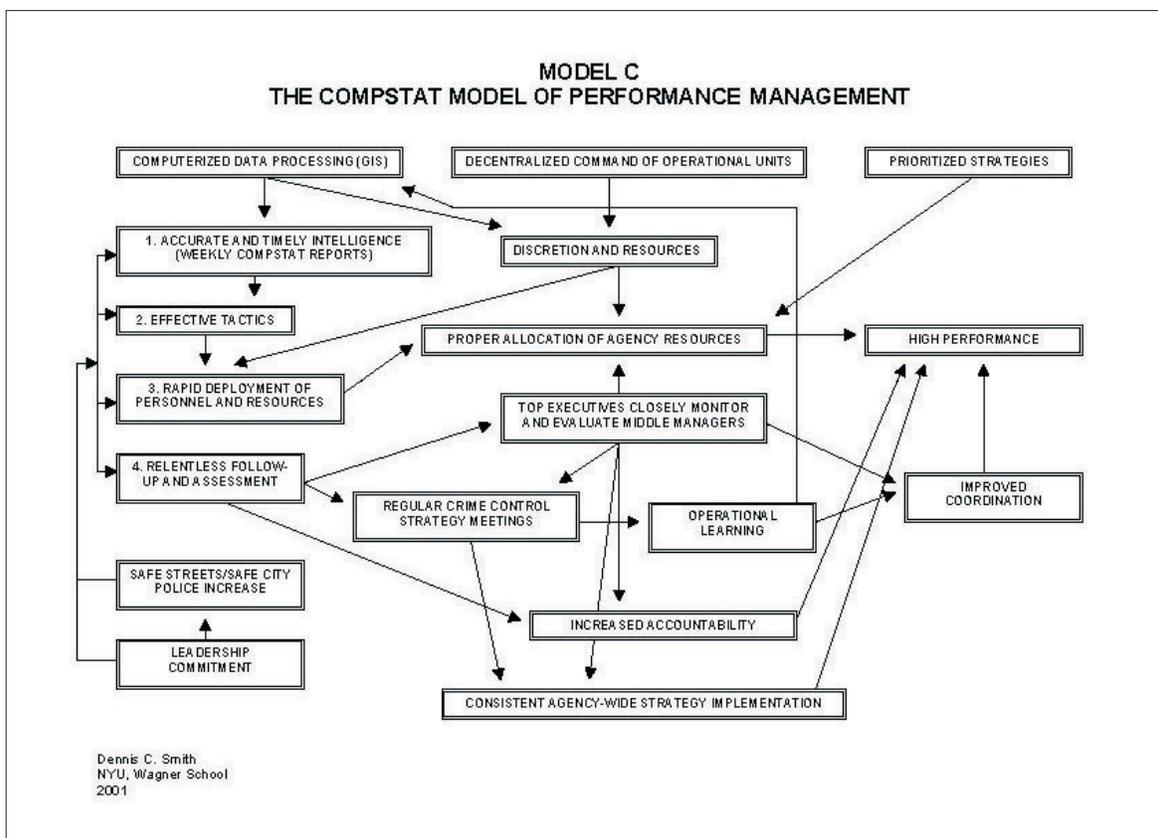
administrations. And I watched with great interest when it began to be used by Bill Bratton and Jack Maple at NYPD. Basically they got it — and they saw that the police department had a tremendous amount of information and that, if they would assert it, they would have an outcome people cared about. That outcome was reducing crime. Mayor Giuliani had run on a campaign to reduce crime. He picked a police commissioner who had demonstrated previously in a smaller scale with the New York City Transit System that he could reduce crime in New York City. Giuliani picked him to lead that and crime reduction became crystallized as the key police outcome. Tracing how that played out at NYPD was an interesting story in its own right.

Corrections and Parks

I fairly quickly got interested not only in just how it was playing at NYPD, which got a lot of attention. (I imagine that anyone who is in our field has heard about "CompStat" and has ideas about this.) Less known is the fact that other places, like Corrections under Mike Jacobson and Bernie Kerik, and Parks under Henry Stern, created their own versions relatively quickly, an approach to managing performance similar to what they were seeing happening successfully in policing. I went out and the TEAMS (Total Efficiency

Accountability Management System) at Corrections had very similar performance review meetings. They were very tough. I went to Parks and sat around the table with Commisioner Henry Stern and other Parks officials with their blazers, with their nicknames from Henry, talking over coffee about safety and cleanliness in the parks. They, too, had seen the key features of CompStat. You need to figure out what information you need that matters, how to get it in a timely fashion, how to get it to the people who need that information to convert it into action, and then to follow up with more evidence to see if it's making the difference. That's the ongoing process of learning, which I was mentioning in my question this morning, which you mustn't overlook.

The area is sometimes overlooked in connection with understanding of CompStat and the CompStat approach. If you look at the way the police department explains CompStat, they say it is four things. It is a matter of strategies. They emphasize accurate and timely intelligence, effective tactics, rapid deployment of personnel and resources, relentless follow-up, and assessment. Those are things that you heard about this morning from the New York State Criminal Justice director. But when I was looking at it, I was afraid that if people just try to copy CompStat, they would miss some key features. This includes things like having the right kind of information available in a form that you can use in a timely fashion.

When Bill Bratton became Commissioner of Police in New York City timely information was not available to decision makers. The vast mountain of information that NYPD had in 1994 when he came in was in a mainframe system. There were the "mandarins of the mainframe," as I called them, who made sure that nobody else got a piece of that action. There were no personal computers in the precincts. Data would be sent down to headquarters, processed, and months later it would be fed back to precinct managers. They weren't expected to do anything with that. Their job didn't include being responsible for reducing crime. That was an overall responsibility. Centralized 911 dispatchers dispatched their forces. Few Soviet systems were more centralized than NYPD at the time that CompStat was introduced. That had some advantage in terms of our ability to change the system; it badly needed change. And, fortunately, Bill Bratton had the New York City Police Foundation, which was a resource he could tap for money to go to a computer store and buy 76 computers and put them out in the precincts. He began to have people there inputting the data about reported crimes and arrests and so forth, locally, before it went downtown. This computerized processing of data is critical in Fred's story about child welfare. Many of the things that they're doing well today in terms of performance management would have been impossible a decade ago when their record system was completely on paper. But they

recognized that probably from the lesson of CompStat and they changed the way they maintained their information so they would use it to map their data.

The other piece was figuring out who are the actors in the organization that are in places and where, if you give them the information in time they can do something with it. The Parks Department was collecting good outcome information about safety and cleanliness for several years before they got the CompStat idea. What they were tracking with their annual collection of data on safety and cleanliness was a decline of their performance. They would do an annual survey at the end of the calendar year. They would know how their Parks system, which was in many different parks, was doing. But that was too late. By that time, things had changed, the government had changed, and it was on to the next year. Under ParkStat, they instead broke the process into manageable pieces. Each district would get reviewed and the reports on those district performances would be available in a week. Then they would have action planning meetings and they would follow up to see if the cleanliness and safety issues that had been found by the monitors had been addressed. Within a fairly short period, they went from 40 percent of the parks being measured clean and safe to 80 percent. The Corrections Department, at Rikers Island, with its 19 facilities at that time, went from being one the most dangerous jails in America to being one of the safest, using a system that said we need to have the actors in place to manage this information to use it so we can increase accountability so we can accelerate the process of learning.

Leadership was important in every place this happened. Not just top leadership. Top leadership has to have the tools and the skills, and the charisma, to be sure other leaders in the organization take notice, respond, and take some responsibility. So I would add leadership to this model.

Resources are also significant. They were often overlooked. Ray Kelly was Police Commissioner up to the time that Bratton came in. Under Ray's leadership, crime had already started to come down. That is oftentimes not noted. Crime had come down a little, which was a reversal of a long trend. But not the kind of dramatic change in crime that occurred when they introduced the new system when Bratton came in and announced that we're going to reduce crime by 10 percent. No police commissioner prior to him in recent memory had set a target for crime reduction. In fact, when the Mayor's Office of Operations tried to get the police department to participate in the process that most agencies have of setting targets, the police department resisted because they said, "Our job is to respond to crime. We will set targets for our response time, for the number of cops we have on patrol, but too many other

factors cause crime, so we're not going to set a target for that." When Bratton said, "We're going to reduce crime by 10 percent" there were audible gasps, not just in the police department, but in a lot of other sectors as well. The amazing thing is that at the end of the year, it was down 14 percent. That galvanized and sort of crystallized this concept.

Now, CompStat's success has subsequently been challenged in plenty of commentaries that you may have read. People can't believe Giuliani did anything successful or can't believe that the police could ever really have much of an impact on crime. There is criminology literature, very pervasive literature out there, which said that crime was bigger than the police, that they didn't have the technology to address crime. Experts have come back with a lot of other explanations. The economy improved, crack cocaine declined, the proportion of young people declined, so you didn't have as many people in the criminally active age group committing crime, and the corrections population went up. In the article in Dall Forsythe's book that I wrote with Bill Bratton about this, we took on those other arguments. George Kelling, who wrote Broken Windows, is also doing some work on this. Basically, we can show that the economy improved after the time CompStat started. We can show that the age trends were actually the reverse of what people thought was the alternative explanation. The ages that are normally engaged in more crime were going up when crime was going down in New York City. So that can't be the alternative explanation. We basically went one by one knocking down alternative explanations. One of the things I believe strengthens the argument that this management approach makes a difference is that when I studied other agencies, it worked there too. They weren't affected by the environmental factors that may have affected the police. There was a dramatic improvement in Corrections and Parks and Welfare management.

Citywide Accountability Project

Mayor Giuliani in his book overstates (imagine a politician overstating) the extent to which he led a transformation on this management approach throughout the city. He really didn't attend to it as a citywide thing until the very end. He created something called the Citywide Accountability Project. It was in its infancy when he left. Then Mayor Bloomberg took it on big time. It was made a central goal of his management that this approach can and should work broadly across the city. He's asking every agency to show-and-tell how it's using the CompStat model to improve the performance of the agency. Bill Grinker and I wrote about CompStating welfare and social services, picking up where we had left off in an article we wrote for the City Journal at the end of the Dinkins's administration in which we

chronicled the decline of the use of analysis and evidence in the Dinkins administration. I had already told you that did not get all that far in the Koch administration. The Dinkins administration had a variety of reasons — budgetary reasons, ideological reasons — to not invest as much in evidence and analysis. It had really declined in places like HRA when then Commissioner Bill Grinker, who was the founding director of NDRC, worked hard to build it up. They tried things like family preservation in child welfare with no analytical system in place, no ability to target, based on analysis where they should be going with this. Not surprisingly, it was not succeeding at the time we wrote "HRA Adrift: Social Spending Without Direction." Ten years later, we came in and looked at the use of CompStat for homeless services in HRA, in JobStat that Swati is going to talk about, and, yes, in child welfare with very noticeable results.

New York City as a Laboratory for Performance Management

I think that New York City is a laboratory for this whole performance management movement. We've learned a lot there. My feeling is if you can do it there, you can do it anywhere. You've heard that line before. Now you don't have to wait for the MMR to arrive in the mail and take up space on your bookshelf. The report has been dramatically improved. It has been dramatically reduced in length, with a bigger focus on outcomes. Every agency had to go through an exercise saying, "Which of the things that you do should the City care about? Should the public care about? Should the media care about? Lead your report with that." They've taken out a lot of things that used to fill Manhattan-sized phone directories, like indicators that were a blizzard and mixed in all together in a way that didn't show any priorities. They've taken that out and in many cases moved that information to the web. If you haven't gone to the New York City website you're really missed something that's an important part of this story. Go to the Mayor's Office Operations site (http://www.nyc.gov/html/ops/html/home/home.shtml) and look at the Mayor's Management Report and then followed the links to the agency reports. As a citizen of New York you can put in your ZIP code and, because of the system that has been put in place as part of the Mayor's management reporting process, you can get information about local performance, your local schools, local crime statistics. There is a lot of local neighborhood information that is available now.

One of the recommendations I made in the report during the Koch administration was that this should be the mayor's management report. It shouldn't be just the agency's management report to the mayor. Increasingly, Mayor Bloomberg has been asking, "Which of my performance objectives

for this year, because they focus on outcomes, cut across agencies?" I think that's crucial too. You heard some very good things about that kind of work, particularly in criminal justice. That is the key I think to the next level of achievement in performance management, not just managing performance within agencies, but having mayors who manage performance overall in their cities, and governors who manage performance overall, across the board.

Performance Management at the
Human Resources Administration — JobStat

Swati Desai, Office of Policy and Program Analysis,
New York City Human Resources Administration

As most of you know I am with the Human Resources Administration, which is a social service arm of New York City. It serves food stamps, Medicaid, home care, domestic violence, food pantries, and, as Dennis Smith would remind us, at one point in time Homeless Services as well as the Child Welfare were part of HRA. Our budget is $13 billion, larger than most of the states in the United States. We have about 12,000 employees. I think it's a good idea to put into perspective the performance measurement challenges.

<table>
<tr>
<td>

**New York City
Human Resources Administration**

• The Department of Social Services for New York City

• Administers Public Assistance programs (TANF and Safety Net) through 32 offices and 19 employment contracts serves over 424,000 individuals

• Provides Food Stamp benefits through 29 offices and serves over 1,000,000 individuals

• Offers Public Health Insurance through 19 Medicaid offices and serves 2.5 million people

</td>
<td>

**New York City
Human Resources Administration**

• Administers several other programs

• Manages the total budget of over $13 billion

• Employs more than 12,000 employees

</td>
</tr>
</table>

JobStat Management Challenges

Performance measurement started during the early years of welfare reform. I should say early years because I would like to remind people that the state didn't pass the welfare reform law until July of 1997. Our performance measurement system started in about 1999. When we started looking for performance data, there were tons and tons of data but you couldn't put

**Welfare Reform
The Early Years at HRA**

Primary Goals:

1. Increase Engagement

2. Establish Local Accountability

How ???

them in the proper perspective. "Numbers, numbers everywhere but not one to use." So the first thing we wanted to make sure to know was where every one of our cases was. When we started at that point in time, we had about 300,000 cases in our system, close to 800,000 recipients. We wanted to know where everybody was, or as I would say in my presentation, "It's a Sunday night, do you know where your caseload is?" That was our primary step in performance management. First was to increase engagement of individuals because we knew that we could not move them to self-sufficiency if we did not get them involved in some sort of work activity, which is required with the state welfare law as well as the TANF welfare law.

So the first thing was to increase engagement and secondly was to realize that this could be done only if we had a better local accountability. The local accountability applies in our case to the Welfare or Job Center Director of each of our offices. Our performance measurement system measures the Director, the region (which basically represents each borough), and Citywide. The performance measurement system for us consists of two major pieces. The first we call an engagement report. By the way, if somebody is worried about whether putting these components out in the open and making them public affects the process, I would strongly argue against it. We have JobStat up on our website. We have had an engagement report up on the HRA website forever. We fight with advocates through that. We fight our lawsuits with that. All our information is up there. We have an engagement meeting every Friday. We go through a variety of different indicators. It leads us to a better understanding of caseloads. Who is where? How long have they been sitting in that particular grouping or status? How can we move the cases forward from that status? What resources are needed? What kind of access to the system is needed for that?

The second piece is the JobStat report, which, as Dick mentioned, happens every Thursday. We meet with two center directors and they need to account for the performance of their centers. These performance measurements are part of managers' tasks and standards. So they have to be on board with it because at the end of the year these tasks and standards are part of their performance review. I tell them that I would love to have everybody get 100 points because it improves the overall accountability and the system achieves its goals better.

We provide resources. We provide tools. We provide best practices back to center directors to help them see how to achieve or how to improve their performance.
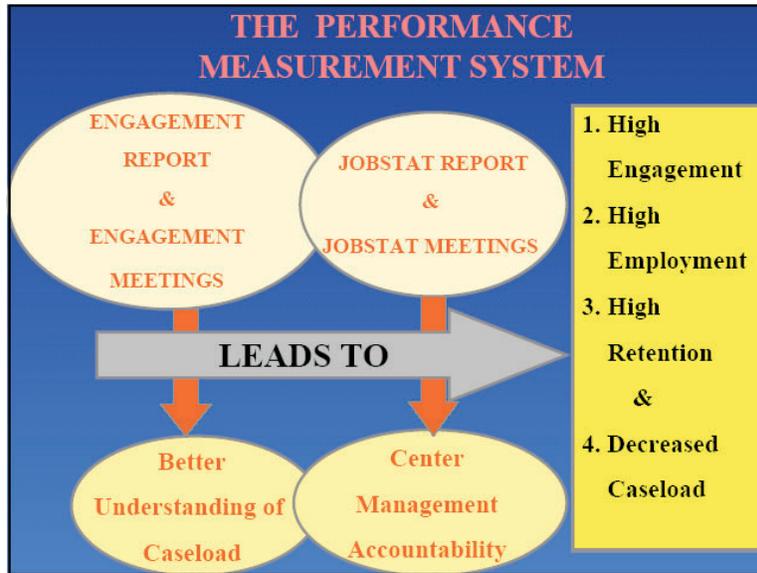
So I consider high engagement, increased employment, retention, and a decrease in caseload to be a tremendous achievement so far as it shows what we have done with our performance measurement system. We have achieved higher engagement and an engagement process moving toward self-sufficiency. The Citywide unengaged, which was once 32,000 we got down to zero. They're still zero, which means that by Sunday night we know where our caseload is and we know where we need to move it.

Defining JobStat

The next slide defines JobStat. What is JobStat? JobStat, as I said, is a performance management tool. Performance management clarifies objectives or goals, facilitates ongoing monitoring, and more importantly, it promotes learning among the best practices. If I have to say two major things about the performance measurement system that we have developed, one is accountability and second is competition, inherent competition among the directors. This was true of CompStat and really leads to better performance in the end.
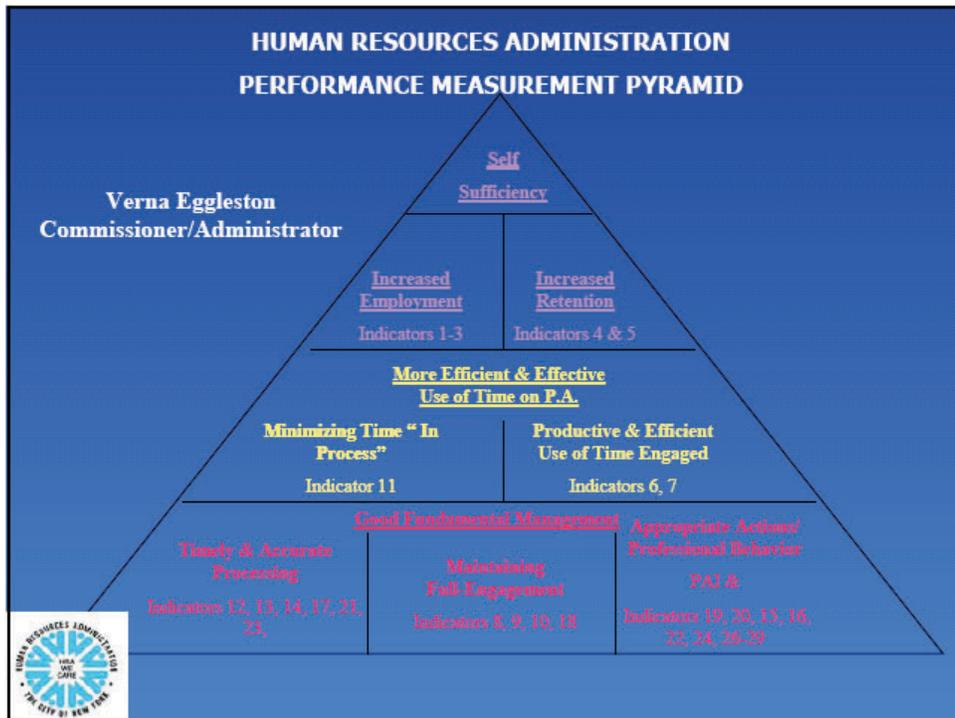
How does JobStat work? This is my JobStat program. I've made our primary outcome as ultimate self-sufficiency for our recipients. At the bottom I would call it Good Fundamental Management, which includes resources, technology, training, etc. Where do all my JobStat indicators fit in the process? We've ultimately decided that just doing the outcome or a goal was not helpful. We needed to break it down with how the goal fits into different processes. We needed to pay attention to processes in order to achieve those goals. So basically you could have an indicator to monitor retention but you can't have that if people sit in one process for very long time without intervention to move forward. We needed to measure the timeliness of processes. Then we needed to measure how efficiently those processes are done. We have what we, at HRA, call a situation room. I think in CompStat they called it the war room, didn't they? We have charts all around the room showing all kinds of maps about the public assistance density, going back to 1995 when we were at the peak of 1.1 million recipients. Now we are down to 420,000 recipients and we can show you where the decline has taken place. Obviously, I know that the high density still is in the public housing areas. We clearly know where people live.

How Does *JobStat* Work?

- Weekly meetings
- 2 Job Centers
- Panel of senior agency managers, including the Commissioner

Jobstat Report, Version 4.5

On the left side sit the administrators. On the right side sit the center managers and the regional staff. All support staff sits all around them. We have a panel of senior managers and the commissioner comes very often. Our first deputy, Ms. Pat Smith, is present at JobStat all the time. All FIA managers are basically at the meeting. We have a two-page JobStat report.

Creating Indicators

We have about 31 indicators on the report. At the top is what we call the Center Profile with center specific informational elements. We don't think it's appropriate to measure these

elements, but it's important that they know what are the different things going on, for example, census profiles. What proportion of the center's caseload is black? We get an average case size. We tell them what's their transitional childcare rate. We tell them what proportion of their caseload is receiving transitional food stamps. Some of these things are

not under their direct control, so we do not measure them for that reason. An important thing is creating indicators. It's really important to come up with indicators that have critical outcomes and processes. These are just examples and I'm not going into detail because they are really arcane details of the welfare systems. But when I do this presentation to the welfare managers across the country, they eat this up.

Setting Goals

Basically, you need to set goals. Every year we have measured goals that get set: employment, or what they call job placement, and retention (three-month retention and six-month retention). I'm proud to say last year's employment goal despite the economy was 90,000. We were able to get 87,000 placements. The three-month retention goal was 85 percent. We were able to get to 83 percent. If you did not count subsidized placement you got

86 percent. Six-month retention was 75 percent. We were able to get to 76 percent not counting subsidized placement. Each center director has his or her own goal. They need to meet their goals. If you come to my JobStat room, there is a chart that we produce on a weekly basis that shows where the center is in relation to its goal. It reminds me of the United Way goal where you have a chart and it starts filling up. It's the same idea but not quite the same chart.

Weighing Indicators

The next is weighing indicators. Our system is a very complex system. It all adds up to 100 points and each of the indicators have allocated points. The point allocation happens based on the importance of the indicators. It's not rocket science. Job placement indicators are the highest points. Job retention is the next highest. In order for them to achieve placement and retention, which are the most points, they must

complete processes effectively, which by design has a lower point allocation. The next step is to measure performance relative to your goals. These are the three employment indicators. What we do is we create what I call an acceptable range for the indicator. Say that the performance goal citywide is 90,000. Then no one would be below 80 percent of that. Points that are allocated for that particular indicator get distributed on a percentage basis. In this case, the center met its goal and received full points. From the second indicator, they only met 72 percent of their goal, so they only got 4.3 points. All of these points get added in order to come up with the score of 63.9. Over time, it also measures current month, three-month average, and the year-to-date average. You see the numbers are very current. By the beginning of February, December '04 numbers are utilized as our operating numbers on the report. By the third week of the month, the previous month's JobStat is up and running.
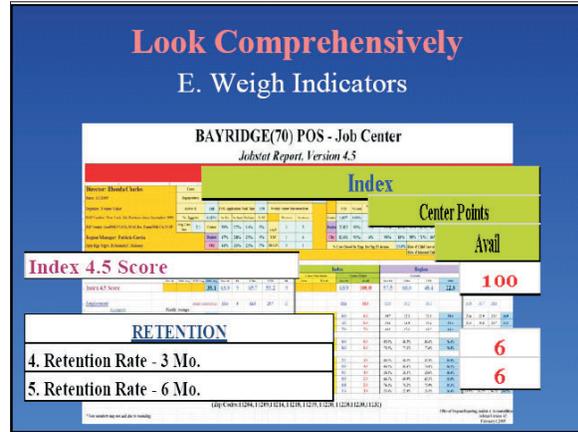
Also, I wanted to point out when we started JobStat, a lot of people said that a lot was invested into technology. Again, the argument that I make for a lot of program directors is
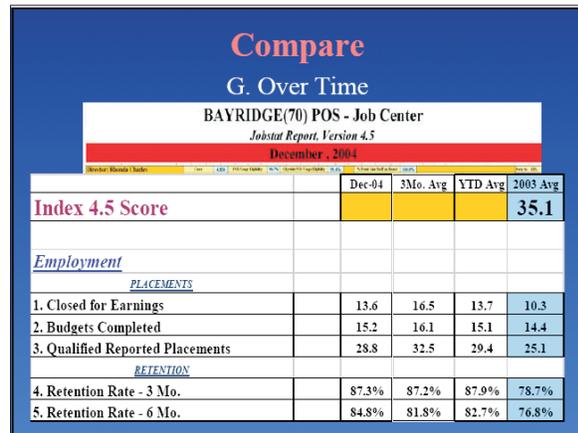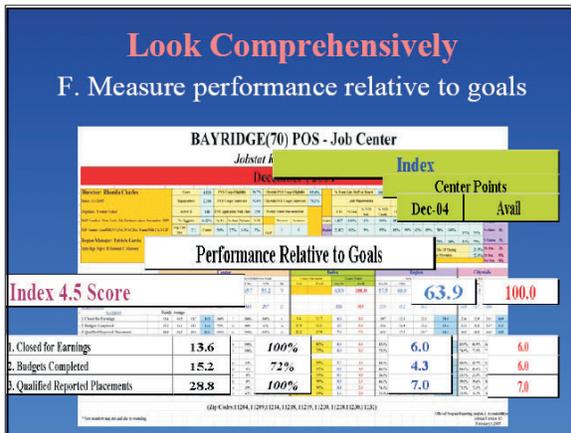
start somewhere. Even if it is a manually run report, start somewhere because what you focus on will improve and you will eventually get the necessary resources. It is really important that you start. You don't have to start with 30 indicators, start with five. Start somewhere. Start with indicators that are timely. Even if they are few, you will be able to move forward from where you stand because the power of the data will get you there.
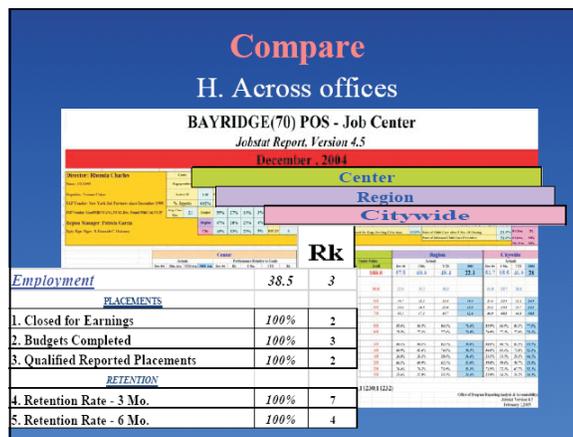
Comparing JobStat Numbers

I. Comprehensive presentation
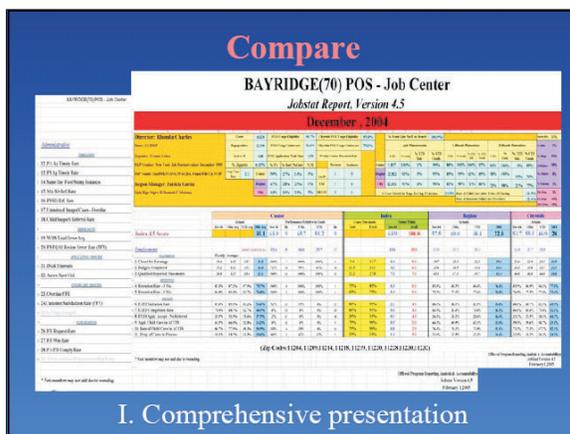
JobStat is compared across offices. We have numbers for each of our offices or centers as well as a number for the region or the borough that they come from and we have a citywide number so that they can see how they compare with the region as well as the City as a whole. This just gives you an idea of what the JobStat report looks like. We have a JobStat tracking system in place. Because I think to make performance measurement a dynamic process you have to be sure that issues that come up get resolved. That tracking sheet basically goes back to everybody who's responsible for follow-up on an issue that arises in the JobStat Meeting. With a lot of items somebody comes out and says, "Well, the computer system wasn't working." MIS people attend JobStat, they are present to hear the issue and subsequently ensure that the problem is resolved and the best practice comes out of that. We are developing an admin staff.

Creating New Programs

It's a dynamic process. We are now at JobStat Version 4.5. We have done JobStat for five years and throughout the years we've changed the indicators. We changed the goals depending on what needs to be focused on. Basically, using our own data we have developed a lot of new programs. In the last few years, we've found, in looking at our engagement report, that population is getting harder to serve and to address this need we have developed the WeCARE program. We've found some people need more intensive case management. Once again, given resources available, everybody couldn't get intensive case management. As I said this is not a mountain WITH THE SAME STONE and same consistency such that one hammer, one chisel works for all. A different consistency of stones is going to take different treatments. What do people think of JobStat? Some feel that, "this is imposed top down." Basically we feel that it's a manager's purview on how to make the process work. Also it helps them to see where they need improvement. Center managers can also ask for detailed data as we often get asked about the JobStat case list. We are then able to give them a full case list of what work needs to be done in order to improve process and performance outcome.
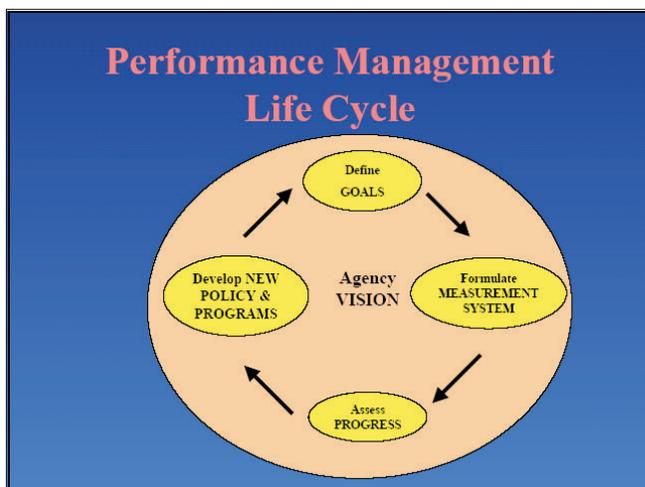
We have had 257,000 placements in three years, and 100 percent full engagement. This is because we have a tight performance measurement system. This particular performance measurement process has moved into a variety of different places. VendorStat is the performance measurement system for our employment contractors. HASAStat is for our Division of HIV and AIDS Services group. MAPStat for the Medicaid offices has just started. We also have a HomeCare VendorStat. Everybody works in a slightly different fashion. The last slide that I really want to show is the performance management life cycle. I think a number of people use this process in several different ways. Basically, it is a dynamic process. It continues to go on and as a result if you set up a performance measurement system and do not look at it and make some change, it becomes a stagnant process and it may not be helpful. The checklist, again out of performance measurement, is really important. Otherwise, you will not have people buying into the process. Reliable data in real time again is very important in order to do this. Without it, you lose your validity and credibility with the staff that are coming into the JobStat meetings and being measured against it. It's really helpful and highlights the development of management efficiencies. If I had to name my presentation I would say, "It is Management, Stupid."

Monitoring the Performance of the Child Welfare System

Fred Wulczyn, Chapin Hall Center for Children, University of Chicago

Coming last, it's a little bit hard to say anything that's new. What I would like to do is to talk a little about what we've done in New York City with regard to monitoring the performance of the child welfare system. It's very similar to what we're doing in other jurisdictions. I'll start by saying a little bit about what the child welfare system is. I might also take a moment to talk a little bit about some of the technology developments that we've been involved in at Chapin Hall Center for Children, University of Chicago. I'll talk a little bit about the system that you may have read about recently in The New York Times. The City has moved to close several contract agencies based on findings from continuous performance monitoring over the last couple years. The City has declared that, by virtue of the fact that the foster care system is a lot smaller than it used to be, the need for contract agencies is no longer what it once was. For example, in the mid-90s there were as many as 60 agencies providing care. The City no longer needs that many. In order to reduce the size of the contract agency group, they decided to use performance. The fact that the decisions are based, in part, on performance has made a bitter pill go down a little bit easier, but I would say that the pill is still fairly bitter. I'll talk about how the system developed. It is not at all dissimilar from what Swati described and what Dennis described.

Historic Roots

I will say a little bit about the child welfare system that makes it somewhat unique in this regard and then also try to draw some conclusions based on the experience that we've had as a research group that got involved in this over 20 years ago. I want to reflect on some of the challenges. What I did not hear today is in terms of historical focus. Swati actually got to some of it today, but I'd like to amplify because I think we are sort of fascinated today by our desire to manage performance. I think to be fair and to be sensible about what lies ahead in the future, we really have to see this as something that's at least 100 years old and goes back to the Progressive era. The Children's Bureau at the federal level was very interested in collecting data about the incidence of infant mortality and what the government could do to better manage that social issue. Our interest in monitoring is not necessarily new; developments in the field follow a long historical trajectory. I think if we understand it in that context, we will have a better sense for what's likely to happen in the future. I'd like to give you a little of that perspective too.

The Child Welfare System

I'll focus on child welfare in New York City since those are the data that I'm most familiar with off the top of my head. The child welfare system is responsible for the safety and permanency of children in families throughout New York City and New York State. The idea is that children should be free from abuse and neglect at the hands of their parents. That's a self-evident proposition. In the event that they are not safe in the hands of their families, then the state is obligated to provide them with protective services in one form or another. That might include placement in a substitute family or a substitute care arrangement for some period of time. But when the state takes that action, it's obviously the most intrusive form of government intervention in family life, with the possible exception of imprisonment. You're dealing with the liberty interests of children and their families. It raises constitutional issues when the state does not do a good job managing the care it provides to dependent children. The state is under an obligation to support the original family (the biological family), or to find a substitute family or some other permanent solution to the safety issues in as little time as possible. That's important, of course, because that becomes the frame for the mission or core outcomes that the system has to abide by.

Statistical Perspective

What does that look like from a statistical perspective? There are about 50,000 reports of child abuse and neglect each year in New York City. That has remained fairly constant despite all of the changes that you've seen in poverty rates and the drop in public assistance. The big changes have occurred on the foster care side. The foster care population in New York City has dropped from about 48,000 in the mid-'90s to less than 20,000 today. It is an extraordinarily important drop and it matches very closely with what we've seen in crime, public assistance, and a whole host of other social benefit programs.

What is the common element here? I think Dennis is sort of arguing that about this time there was a lot of performance management taking place and that might be the common element that accounts for it. That's an interesting hypothesis. There is though the possibility when you see something going across all of these phenomena that there is something exogenous to the system and not something that's endogenous to the system. I think that's an issue that we haven't yet really explored as social scientists. It would be well worth our time to do that, I think.

Core Mission: Child Safety

What are the core mission-critical outcomes in child welfare? We already talked a little bit about child safety. What we're talking about there is reducing the incidence of maltreatment. From a government perspective, we have a lot of trouble with that one because you can't go in and ask families, "Have you beaten your children lately?" So it relies on public reporting as a way of conveying the incidence of abuse and neglect to the public agencies that are responsible for it. We tend to look at reports of abuse and neglect as an imperfect indicator of the true incidence of maltreatment. We are then naturally interested in reducing the recurrence of maltreatment. So the government intervenes when it finds credible evidence of maltreatment. We are interested in those interventions that reduce the likelihood of recurrence. In the event that the family cannot take care of the child, the state will place the child in substitute care as I mentioned. In that case, we'd like to reduce placement. We'd like to create more effective interventions prior to placement to reduce recurrence.

Foster Care Placement

But in the event that you cannot protect children by keeping them in the home then you have to place them in foster care. If they are in foster care then the permanency objective becomes the overriding value in terms of what should happen next. We're interested in permanent homes for kids so that their caregivers will provide a safe home for them until they transition into adulthood. We're interested in stability of those placements and the formation of stable relationships. We'd rather that children not move around. We would like to increase the probability that the child will leave the system to some sort of permanent outcome, which these days includes reunification with natural family, adoption to a new family, or guardianship of one sort or another. New York does not have subsidized guardianship, but basically we're referring to kin or other adults who assume the role of guardian. Guardians need not be kin. Typically, they are kin but these are people the courts assigned to legal guardianship for the purposes of raising the children. The parental rights have not been terminated in that instance and that creates tensions that are interesting and relevant to other conversations.

The City of New York did contract at various times with up to 80 contract providers for the range of alternative living arrangements and preventative services. I'm going to talk about the providers that provide alternative living arrangements. When you think about a child placed in foster care, they are placed with another family. There might be other foster children there. They may have their own natural kids. But they are placed in that family-like

setting. That is the requirement of the law — that we seek the most family-like setting for kids. There is also group care, which is exactly what it says. That is often used in the event that the child needs some sort of therapeutic intervention in addition to safety and protection from their biological family. These 80 contract providers provide the spectrum of those services, all the way up to residential treatment.

The Administration for Children's Services

In the vein of story telling, Nick Scoppetta became head of the Administration for Children's Services (ACS) during the Giuliani administration in 1996 after a particularly horrible death of a child who was known to the child welfare system. The child welfare system did not respond appropriately or in a timely manner. Mayor Giuliani got rid of the current commissioner for Special Services for Children, changed the name of the organization (to ACS), moved it out of HRA, and created a new agency and appointed Nick Scoppetta, himself a former foster child and otherwise a distinguished resident of New York, to take over the agency. Not surprisingly, being a disciple of Mayor Giuliani, Commissioner Scoppetta hired the same consultant who the police department hired. The consultant made certain recommendations to the commissioner, not the least of which was, "Do you know what your outcomes are?"

Early in his administration, the commissioner held conversations on Monday nights with child welfare experts. One such night he asked, "What are our outcomes?" Around the table everybody said, "That's the problem, we don't know what our outcomes are." I did respectfully dissent because it was my opinion that we indeed did know what the outcomes were. We were, however, somewhat afraid to commit ourselves to the outcomes as a system because we had such a hard time managing the accountability process. If you declare what your outcomes are, well, you're just setting yourself up for your own termination.

The EQUIP System

There was a huge hurdle, getting over that issue. Then Commissioner Scoppetta had to blend that idea with the idea that, "I don't really do the work. I supervise the work but I don't do the work." He had to look at these contract agencies and say, "Okay, we have outcomes. We now need to manage you guys to do a better job." To just play out a little bit of the history in brief, after we figured out how well the agencies performed, the commissioner decided he wanted to change their performance. He realized that the way we pay for performance was a

big problem. He set out to change the way ACS provides reimbursement to contract agencies. That's another link to performance management systems that I want to talk about at the end the day or the end of my remarks. But to go back to the EQUIP (Evaluation and Quality Improvement Protocol) system: basically, it deals with three things. It deals with what the outcomes are. It deals with the process of care. That is, did the agencies do what they were being asked to do in terms of visitations, filling out paperwork, etc.? The third leg of the EQUIP system is to deal with the issue of service quality. Police going out there on the street and suppressing crime, or however they would describe that process, there is a quality component to the way the police do that. There is a quality component to the way folks interact with welfare recipients. There is a quality component to the way child welfare workers deal with families. So the third leg of the EQUIP process was to address that issue.

To make a long story short, EQUIP has been in place from the late 1990s. It's done on an annual basis. There have been over the last few years a certain group of agencies that each year occupy the bottom rungs of this performance measurement system. Increasingly, public advocacy has turned to the question, "When is the City going to stop placing children with agencies that have historically been demonstrated they don't do the same job as other agencies?"

Dennis talked about competing explanations for whether or not performance management is what has brought down the crime rates in the City. We have the same problem in terms of dealing with contract agencies and addressing the competing explanations that account for why do some kids go home 30 percent faster than the kids in other agencies. It might be because they're more difficult, but it might not be. The performance measurers are under an obligation to try to weed that out.

The Culture Change — Using Information to Justify Decisions

Again, we've talked about the development of information systems, the trajectory of all that, which look (at least in the gross detail) very similar. From the time the first piece of data was collected to the time it got used for purposes with which it was intended was a very long period of time. The explanation for that has to do with human capital formation: Getting people to know what to do with the information. You're backing out of systems that, because they've operated for so long without information, have created ways of doing things that don't require information. Now you're asking them to use information, which they've gotten very good at not needing, in order to justify decisions. That's really the culture change. For that to happen quickly you have to have people who appreciate that. And, quite frankly, in

terms of government, we're talking about the late 1980s to the late 1990s. We're talking about industry participants in general who came of age before computers. They were educated long before laptop computers, etc. Even in their educational backgrounds, the ability to link to the importance of information in any kind of systematic way was really a generation behind the evolution of information capacity. That's been part of the reason why we've seen the time lag between when the information became available and when the information was used in the manner intended.

What Makes the Child Welfare System Different

Let me talk a little bit about what makes the child welfare system different. First of all, there is a fundamental tension within the field. That is the difference between family preservation and child protection. A lot of people want to preserve families, which is a very important goal, and a lot of people want to protect children, which is ultimately the most important. But there has over time played out a very fundamental tension. The way that plays out is the difficulty around forming core mission critical outcomes. It is another reason why the system was slow to announce outcomes. Another thing that is unique, and I think probably Swati deals with this in a way that perhaps the police and other performance managers don't, is the time scale of the underlying phenomena. When kids go into foster care, they can be there for ten years. When you talk about performance change that drives from having monitored the performance, the lag time needed to observe performance changes can often outlast administrations. We have a political process that doesn't have much patience for that kind of thinking. But that is the natural phenomenon of childhood. Increasingly, the child welfare field is turning to child wellbeing as a more global outcome. Understanding what you did for a child when they were five years old in relationship to their wellbeing really requires that you wait until they made the transition into adulthood to see how they are doing in relation to the labor market. For us to engage these kinds of outcomes, we have a lot of thinking to do. In this discussion, time scale and political patience are very important.

Performance Forecasting

There is one thing that I did not hear much about today. I heard a lot about performance monitoring, but I did not hear much about performance forecasting. For systems to get better, they have to set goals and they have to hold themselves accountable to the goal. The process is both prospective as well as retrospective. Swati talked about it a little bit, but I don't think there was enough emphasis on the data for looking forward in time and establishing goals that

you can then manage in real time. Again, when you have these long time horizons, projection gets to be pretty difficult.

Linking Performance, Outcomes, and Finance

Another issue to consider is creating a direct and unambiguous link between performance, outcomes, and finance. We heard the gentleman from the Budget Division talk a little bit about making this link. I think there has to be a direct and unambiguous link to finance. One of the things that we're doing with performance-based contracting is establishing the groundwork for that. I think that is very important. I think we need to distinguish between changes that are brought about by improved (treatment) technology versus changes in the demand for services that are brought about by reduced need for services. This is a particular problem in the child welfare system. When we say we don't need as much foster care, it's not because we think we've eliminated the need that those families have. What we're saying is that the technology for serving families has changed and we're shifting it from foster care into family and community-based services (i.e., greater efficiency in the way we meet service needs). The need hasn't gone away. When performance monitoring induces greater efficiencies in service delivery, you have to worry about where the money goes, especially in a fee-for-service system, because you still have needs to meet.

Role of the Federal Government

Another issue that we deal with in the child welfare system is a very intrusive federal government. In child welfare, the federal government says how fast kids should get home from foster care. How fast they should be adopted. How likely they are to come back. How likely they are to have recurrence of maltreatment. This is not in and of itself a bad thing. It's actually been very helpful to the process. The problem is that the federal government has selected the worst conceivable measures for that. As used now, the measures could actually obscure performance improvement and lead to poorer performance. The opposite is also true. The measures used could point to improved performance when the opposite is what is really happening. One key to performance measurement systems at the local level is the way in which it relates to federal, state, and local performance. The approach adopted has to be scalable. You cannot lose your (measurement) acuity as you move from one level of government to another because for the United States to do better, New York's got to do better. For New York to do better, New York City has to do better. For New York City to do

better, the contract providers have to do better. If you're losing information as you move along the scale, you've diminished the value of your performance monitoring system.

A few closing remarks. I prefer to think about this not as performance management but as change management. In the systems of care literature, on the mental health side of things, they get those systems to work well together. Everybody is happy. All the professionals are happy with how well their systems perform. The problem is the outcomes for kids did not change. If we focus exclusively on performance, we may miss outcomes. We think that there's a link, but unless you're managing outcomes and using performance to get to better outcomes, performance becomes an end in itself and it can be translated on to the organization in a way that doesn't translate down to the clients.

In terms of skill sets, I think somewhere in your organization you're going to have to find a link between computer science, statisticians/researchers, and program and policy people. That's the critical fusion that allows the timely delivery of information. If you don't have computer scientists in the room, you don't know how to organize the data to deliver them in a timely fashion. However, if you leave that responsibility to computer experts acting in isolation, you are more likely to get information that is of little practical value to policy makers and practitioners. This has happened repeatedly in child welfare where the government has spent millions of dollars with only a marginal return. Also, this whole area of performance management is a practice area in its own right. Oftentimes, what has happened is we've observed people saying, "Performance management, that's something I do to somebody else." But there is a set of best practices for performance monitoring/performance management, and it has to be a self-reflective act. It can't simply be somebody in a high organizational position looking at somebody downstream. The practice itself has to be examined to insure the best tools are being used.

Another name to call performance management would be variation management. The thing that we're really trying to do here is manage the variation in outputs and inputs, so that we get a consistent product out to the taxpaying public. Dick always likes to recommend books, so I thought I would come along with a book recommendation, *The Full House: The Spread of Excellence From Plato to Darwin* by Stephen Jay Gould. You may know that Stephen Gould was a paleontologist. Paleontologists take, as Gould would say, the long view of change. His book is about why there are no .400 hitters in baseball anymore, which is an interesting topic in its own right. He talks about performance improvement. One of the themes of the book is the right wall of achievement. There are limits to how much

inefficiency (Gould thinks in terms of variation) we can squeeze of systems. Although he does not talk about our issues specifically, we have to think in terms of the right wall of human achievement. I think we will in time create expectations about outcomes that performance management alone will not help us achieve. Thank you.

Discussion

Question: *I was interested in your remarks on JobStat and accountability. CompStat made a difference in terms of human resources. There was inspiration. There was a turnover of a lot of people. In the case of JobStat, where the indicators are followed, does it seem to make a lot of difference in who gets promoted or who loses their job? How is this sort of carrot and stick aspect handled in terms of personnel?*

Swati Desai: Given the civil service nature of the system, you couldn't do pure hiring and firing, but demotions do take place. At times, we also have given monetary bonuses to the ten best managers. That also helps. But we also have given from time to time what we call small awards, which are small in-kind things to the centers. The centers can choose. Sometimes they choose movie tickets for the whole staff and everybody gets two movie tickets. Sometimes they choose a color printer because they see that the power of presentations is the color. They choose what they want and we give it to them. Small, but it also acknowledges them once in a while. The commissioner gets on the phone and congratulates people when they are number one and number two, especially when they move from rank 15 to rank 2. There are processes of persuasion, both monetary and non-monetary.

Question: *Does that apply to the private contractor side too? In the child welfare system, is there a tighter and tighter accountability around performance? If you don't perform, you lose the contract, etc.?*

Fred Wulczyn: Yes, they've done that in a couple of different ways. One is what they call volume management. If you are persistently underperforming, the City will reduce the contract levels for admissions or they will adjust the percentage of the volume that you could get according to your performance. They hit you on the front door. They recently implemented an enhanced rate for agencies that is a per diem rate. For high-performing agencies it's a little bit higher than it is for your mainstream agencies. The problem that they're likely to encounter there, though, is that it creates a fundamental tension in the system where the one thing that happens when you have a higher exit rate is the rate in which you lose money. If you are on a per diem system that just accelerates the rate at which money is lost. This is a process. The article in The New York Times said that there were two failed attempts. That's The Times own unique take on that. The way I look at it, having been involved in those two attempts, is that Commissioner Mattingly put forward the next iteration of a process that will one day transform the per diem system. It has to happen. This is the third time in a different context in which it will be tried. There were two very successful experiments on getting something complicated off the ground.

Dennis Smith: Bill Grinker and I talked about that in the paper on the promises and pitfalls of performance-based contracting. Part of that for the nonprofit sector was that some of the small players didn't have the management systems to be able to play. Yet in New York City at least we didn't want to lose them. For example, for the workforce development effort, SEEDCO developed an intermediary role, providing some of that management capacity to kind of cushion and support the CBOs to enable them to play in the game and do the pieces of it that they could do. There is, as Fred was just saying, a lot of learning going on. One of the slowest runners and what worries me is government and how it might make a right turn that would leave a lot of its critical partners to run off the cliff, if it isn't careful. We talk a lot about that in the City.

Swati Desai: We also have performance-based employment contracts. In the original contracts the payment scheme was set up such that they got the most for placement and slightly less for retention. At the renewal, we switched that balance. More money was paid for retention compared to placements. Clearly, the contractors that are not able to perform decided to leave the system because we reduced their referrals.

Question: *Maybe this might be for Dennis. The systems that you've been talking about are really big systems and I wonder if you see any application at a much smaller level. The city of Poughkeepsie is even big to talk about. Can you talk about application at a smaller level?*

Dennis Smith: It's always frustrating sometimes that both of us touted the book, *Reinventing Government,* and we kept using the example of one small town in California. From a New Yorker's point of view, it's the size of my apartment building's population. A close analogy is with the organization, SEEDCO, that I've been working with. We've been training managers for nonprofits on the use of performance management. A lot of them have taken to it like ducks to water, partly because people go into nonprofits because they care about the problems of people with age or with child welfare. They have their own motivation to want to focus on the outcomes and if you can give them tools that will help them demonstrably achieve their outcomes it helps them get up in the morning. But it also helps them in everything that they are trying to do now in terms of foundations asking for evidence of results, and government contractors then asking for it in performance-based contracts. They're really seeing from lots of different sides the need to do it. If you show them that it isn't high tech necessarily, it's mostly trying to be very systematic about what they're doing. It's about figuring out how different things that they are doing do or don't contribute to getting to those outcomes. Figuring out when they need that information so that they can make the differences and achieve better outcomes.

The Bowery Residents Center is a poster child for this. They've pushed it into all their programs. They are dealing with very difficult problems. It seems to me in the follow-up work that we've been doing on how the systems have been merging there, it has really been working for them. When you focus on outcomes, as you heard this morning, it seems to help and you begin to see that to succeed at what you're trying to achieve you need other contributors. So you develop a logic model of getting this reduction or this increase in some outcome. Somebody who is not part of your team has a critical piece of that action. That gets

it on the table. Then you change the group of people who are meeting to figure out how to do it. Swati and I were talking about that on the way up on the train.

Another piece that's happening now is that this is being pushed out into all the other agencies. The focus on outcomes becomes more intense. It's just impossible, even though they are not separated, for homeless agencies, HRA, and ACS not to have at some level and indeed at multiple levels people talking together to try to work together. That crosses clearly into these nonprofit organizations that they depend on to deliver the things that actually end up showing up as the product of these agencies in the Mayor's Management Report. An awful lot of that, for many agencies actually, the worker bees are nonprofit organizations delivering those outcomes. This is why I wouldn't buy into Fred's idea. I've learned a lot from what Fred said, of changed management as a substitute, because I've seen too much literature on change management where change is the end and people take a big bow because they changed something. But the thing that they changed isn't outcomes. They haven't gotten that. You have to say outcome performance management to keep that there. I think change management could be another diversion where you're talking about an activity rather than the outcome.

Question: *As a follow-up of local level, how do you overcome the problem of management being subject to re-election every two years? When your goal is getting re-elected a lot of times or you don't want your management to potentially change, how does that affect this process?*

Fred Wulczyn: I think that's the virtue, if you can capitalize on it, of the civil service system. The permanent workforce of the government provides some stability over time. What the political layer brings to that is maybe a change in view of what are the particular values that we want to espouse. The stable civil servant in local towns, villages, and even in the State of New York ought to be in a place where it can dampen some of the oscillation or swings in ideology from one elected official to the next. Getting services systems to go forward in pursuit of common outcomes is really the important part. The goal of using outcomes is to level political differences. There are no Republicans I know of who are against child safety. There are no Democrats I know of who aren't for child safety. It creates a common ground for having a conversation and the focus on the outcomes creates the modus operandi that should be relatively immune to the political vagaries of ideology. I don't want to oversimplify the problem, but I think the conversation should be about achieving outcomes with less emphasis on approaches, especially if the evidence about efficacy is blurry (as it often is in social program).

Dennis Smith: One example, a little note on Fred's point, is actually New York City's NYPD. When Bratton was thrown out in the political squabbles between Giuliani and Bratton, Safir was brought in and he didn't like CompStat. He was going to get rid of it. The police officials who had been running it threw up red flags in as many places as they could. Some people saw it and they went back to Safir and Giuliani said you've got to keep going because they didn't have anything to replace it. He didn't want to see the crime reduction improvements go away. So they held on to it. What they did with it is another story and some of it is told on this chapter on CompStat in the Dall Forsythe book. People who were in the

system, who were probably opposed to it two years earlier, then really wanted to hang on to it because they knew it was working. So they stopped a political appointee from changing it.

Question: *This session fits with what we learned about this morning. I have a couple of reactions and questions. One thing I would like is to have Fred Wulczyn speak about now that child welfare has been in the news. They dropped some providers recently. I'd like you to say more about whether it was because of their foster care performance or their performance in family care. What were the kinds of things that caused agencies to be dropped? What can you tell us about the politics of that? Is it something where people go to the mayor and use all their supposed political muscle to deal with it?*

Dennis, you talked about Commissioners Bratton and Safir. People come in to these appointed positions, which they are going to have for two or three years, and says, "Boy, I wish the people before had known what to do. I'm going to show you what to do; I'm going to change things. It's going to be different." I was listening this morning and thought that people think about government in terms of what people are arguing about — Pataki's budget or the budget that the President just sent up. I was pleased listening this morning to the three commissioners talking about serious things that they do. Things that will never get into the paper, that reflect caring leadership by political people and it made me feel good.

Fred Wulczyn: It's an interesting dynamic and it's what reminds me that we have a tendency to see and inflate the value of our own contribution to a particular moment in time. I think what has happened in the City, as I mentioned in my remarks, is they have eliminated contracts with two providers of foster care services. They just ended them. Kids in their custody are being transferred to other contract agencies. Whether they are actually moving their foster homes or not is a case-by-case issue. But the City literally stopped the agencies from providing services because the agencies in question did not provide care well enough to achieve outcomes comparable to what other agencies accomplish. We wouldn't be in this position if it were not for the fact ACS had been watching all of the agencies for a period of five years.

That was really the issues in *The Times* articles earlier in the summer or in the fall when they basically asked the question, "What are you going to do with the agencies that perform poorly?" They asked that to Commissioner Mattingly right around the time he came in. The situational context is important to understand. If the City was struggling to meet demand with capacity, I think it would have played out differently. I think part of this is a historical accident or good fortune, depending on how you look at it. I also want to stress very much that both agencies are operated by minority-controlled organizations. One was created during the foster care crisis during the late 1980s. There is a great deal of value placed on those organizations. The public's responsibility is to not pronounce an end of contracts without trying to first strengthen the organization. I like the SEEDCO model of providing backroom support, because one issue with a lot of agencies is that they are smaller and they cannot achieve economies of scale on the administrative side. Therefore I think this is the public agency's obligation to support capacity development. Then I think a public agency can be rather unsentimental when it comes to the end of the day: "Listen, the job is not getting done. We (the public agency) have an obligation to use our knowledge efficiently because that's

how you distribute public benefits most broadly." So you can be very unsympathetic about that, but you have to be in a position where you say, "We will work with you on a performance improvement plan. If you don't succeed at that point then we have to take harder corrective actions." That said, once a decision was made, all of the things that you might imagine (phone calls, appeals), all that stuff took place. I don't know if the mayor got a call, but I'm sure somebody was trying to figure out how to get to the Mayor, if somebody on the Board knew somebody who knew somebody connected to the Mayor's Office if not the mayor himself. So all that stuff plays out and it can get difficult behind the scenes.

Question: *How do we learn more about EQUIP?*

Fred Wulczyn: We have a paper. We're academics, so naturally one would have to have a paper on this sort of stuff. So we have a paper that we will be happy to send you. It came out a couple years ago. It describes mostly the outcome end of things. It explains the risk adjustment process. It explains what the outcomes were. There's another paper that we did with some folks from the Office of Children and Family Services. It describes similar work in the context of New York State and the management of county child welfare systems that we can also send you.

Question: *Here is one of the things that I'm finding as we are increasingly utilizing performance outcomes for contracting with not-for-profit organizations and human services. We're having increasing pressure on us to align that directly to the payments for those contracts, which presents a huge challenge. As we're looking at desired outcome changes in the behaviors of individuals, which tend to be more long term, and the alignment of the payment to the not-for-profit provider, we're going to be challenged to really start to shift it more to activities and outputs. Particularly to allow for some cash flow for these not-for-profits, very often these small CBOs, to be able to continue to provide child abuse prevention programs or services. While we may be looking for behavioral changes to improve parent ability to work with their children and families, etc., the expectation is that we now pay these organizations on a quarterly basis for their work so that they can pay their staff. It is going to force us to have their outcomes more defined by the activities and outputs, as opposed to the desired long-term outcomes of behavioral change. I'm doing this in the very small context of program development, unlike New York City, which in some ways has the luxury of having long-term performance that can be measured with a particular contractor. We're looking at, generally, program development initiatives that cover a three- to five-year period for a particular provider. We renew their contract on an annual basis, yet we're also being expected now to implement a performance-based system. I'm trying to figure out how we approach that without having to dumb it down to pay for activities as opposed to really trying to look at the longer-term need for behavioral changes. The Division of the Budget and the Office of Children and Family Services are increasingly expecting appropriate performances. They're increasingly encouraging agencies to align payment to performance. When we start to operationalize that, that's when it starts to rub.*

Fred Wulczyn: I think this is where the work of academic institutions can play an important part: What's the contribution of spending to quality-of-life indicators? That

draws the conversation into more of an outcome focus so that it's not just process and activities.

In the last 25 years, we've made extraordinary advances in not just having data but in the analysis of data, including spatial and temporal analysis of human data and mathematical models that help organize our thinking. That specialized knowledge today exists in universities. The diffusion process for that knowledge into government is a slow one, but a good one. That's why I think there has to be viable partnerships with universities. I wouldn't give up creating a positive measurable link to outcomes that are even spaced close together. But it's the quality of thought that goes into it that will allow you to do that.

Dennis Smith: I want to add that I think that logic models make explicit the assumptions about the relationship between activities and outputs, and outputs and outcomes. You need to draw on academic research that's been done. You can then use it to reassure yourself that you're going down a reasonable road. In the case of policing, there has been systematic research that debunked a lot of props for the old way of policing. Studies showed that a lot of things that police agencies were investing in — trying to produce rewarding response time, rewarding this and that — wasn't helping that much. It took CompStat doing something to free up the police department from what Mayor Koch focused 8 to 12 years on — trying to maintain a promised level of police officers on patrol deployment. That was his obsessive focus. Visible patrols have been shown by researchers not to make that big a difference. They stayed with that until they showed through other evidence what could really bring crime down.

Appendix A:
Table of Contents for *Quicker, Better, Cheaper?: Managing Performance in American Government*, edited by Dall W. Forsythe

Appendix B:
Presidential Address: "Complexifying" Performance Oversight in America's Governments

Presidential Address: "*Complexifying*" Performance Oversight in America's Governments

Richard P. Nathan

In keeping with my focus as President of APPAM on the "M" in our name, this talk deals with the performance management movement in American government.

Senator Daniel Patrick Moynihan, in a moment of frustration with me at a 1989 Senate hearing on welfare reform, said I am a "complexifier." In this spirit, the way I see it is that the performance management movement in American government is on the right track, but that it *oversimplifies*. I want to serve today in the role of a "complexifier" for the performance management movement. My aim is to be constructive—to suggest ways in which efforts to improve government performance can be reconciled with the pluralistic setting of U.S. public management.

Leaders in the federal government over the past 40 years have oversold simplistic systems for fulfilling public policy goals as expressed in the alphabet soup of systems like PPBS, MBO, ZBB, NPR, and GPRA.

The Alphabet Soup

PPBS stands for Lyndon Johnson's Planning-Programming-Budgeting System, adopted with much fanfare and based on private industry and Defense Department systems to assess and compare public spending options. MBO was Nixon's successor approach for Management by Objectives. ZBB was Carter's more radical initiative for zero-based budgeting to rank all spending options from the ground up in allocating government funds. NPR was Clinton's National Performance Review, which sought to focus government management and budgeting on achieving results. In 1993, a system for strategic results-based management and budgeting was enacted into law under the Government Performance and Results Act (GPRA).

For shock value, at the end of October of this presidential election year, I want to say something good about the newest alphabetically named performance management system. No matter what happens in the presidential election next week, I suggest this slogan, "*Let's not part with PART*," referring to the George W. Bush Administration's "Program Assessment Rating Tool" to compile effectiveness ratings for all federal programs.

THE PART SYSTEM

According to the FY 2005 budget, the Administration has "PART*ed*" (this verb form is used frequently in the budget) 400 programs representing 40 percent of the

budget. Two hundred programs will be added in FY 2006. These evaluations are based on 30 questions to rate each program that is assessed in four areas: 20 percent for its purposes and design being clear, 10 percent for strategic planning, 20 percent for program management, and 50 percent for results. The ultimate goal is to evaluate the performance of all federal programs (over 1,000) in this way.

PART assessments draw one of five conclusions: effective, moderately effective, adequate, ineffective, or results not demonstrated. They are available online.[1] Illustrative ratings are for Head Start "results not demonstrated," for Medicare "moderately effective," and for the Community Development Block Grant (CDBG) "ineffective." Reading PART reports (most of them are about eight to ten pages long), one is struck both by their brevity and their variation. Some, for example on food stamps, are evenhanded and draw on a range of sources from within government and outside. Others are analytically weak.

DISCUSSION

The critical challenge for the PART system and efforts like it is setting performance goals. Where do goals come from? The answers vary. In some cases, it is wishful thinking in the form of political over-promising. In other cases, and these are the ones I care about, performance goals are based on research, policy analysis, and expertise. For this big category, I believe there should be broader consultation and collaboration involving policy officials, program managers, policy analysts, academic policy researchers, and public management experts.

In situations in which the results of definitive social science experiments based on random-assignment studies can be drawn upon, they materially aid policymakers by giving them a high level of confidence about impacts. *Such studies, however, are not available, and indeed could not be conducted, for all types of public programs and for all types of affected groups and policy conditions and needs.*

Terminology is important here. An *impact* means we can show that a particular public program caused something to happen that would not have happened otherwise. An *outcome* is the word customarily used to express a program's performance goals, regardless of whether we know if its activities are additive. Frequently, two types of people are involved: policy researchers, who generally favor experimental studies of program impacts, and management experts, who focus on outcomes. The two groups often have different mindsets and skill sets. As I see it, public policy and public management researchers (especially active members of APPAM among them) should contribute to knowledge about *both* impacts and outcomes.

Leaders in the public service are called upon in many (indeed, most) situations to set and periodically adjust outcome goals for performance management based on the most pertinent public policy knowledge available, and also drawing on expertise, experience, and observation. And yet even with the best of such efforts, governments often do not deal wisely with another critical challenge discussed in this paper, the need to devise politically acceptable and workable performance goals that can't be gamed for undesirable purposes.

While I think the PART system is on the right track in focusing on individual programs as the basic building blocks for assessing managerial performance, major problems are that it is too centralized, too insular, and not sufficiently discriminating. It does not adequately take into account the great differences that exist in the

[1] See www.whitehouse.gov/omb. Accessed September 7, 2004. See also Rodriguez, 2004, pp. 56–61.

size, importance, operational character, and settings of different public programs. I have a particular *federalism* problem in this context.

THE FEDERALISM CHALLENGE OF PERFORMANCE MANAGEMENT

Most domestic programs of the federal government operate by indirection through state and local governments that in many cases contract with nonprofit and private corporations to provide public services. It is beyond the scope of this talk to survey the backward and forward bounces of decentralization in American intergovernmental relations. Suffice to say, as Martha Derthick emphasizes, there are many ways in which members of the Congress and federal Executive Branch officials attempt to influence domestic affairs by adopting narrowly focused grant-in-aid programs and imposing and enforcing conditions, regulations, and guidelines on their operation. On the other hand, there is also a long history of devolutionary efforts to broaden grants (for example, by creating block grants), loosening or not enforcing regulations, granting waivers of federal requirements that are sought by state and local governments and interest groups, and generally by virtue of the fact that policymakers often disagree on public purposes and as a result adopt vague or even contradictory goals for domestic public policies.[2]

Wherever the devolution ball bounces, the point that stands out for me is that state and local governments have to be involved in assessing and improving program performance, especially under federal grant-in-aid programs that are broad gauged and have multiple purposes, as is so often the case. Federal agency officials should work with state officials in ways that are not heavy handed. They should adopt approaches that are continuous, user-friendly, candid, and appropriately intergovernmental.

This recommendation gets me into the consideration of differences between inputs, outputs, and intermediate and end outcomes as the goals of performance management.[3] It is arguable as a general principle that under broad-gauged federal grants-in-aid, the federal government should care most about *organizational outputs*, and in turn work with states to stimulate them to assess and help them in the best ways they can to define and measure the *outcomes for individual participants* of these federally aided domestic programs in the different environments in which they operate. This is what we actually do in more cases than is acknowledged.

Governmental programs have literally thousands of iterations. Reconciling two values, the flexibility (which is inherent to the diverse governmental environment of American federalism) with accountability (which policymakers and administrators should care about and achieve), requires that the PART system be *intergovernmentally sophisticated*. But this is not the only way the performance management movement needs to be complexified.

---

[2] For useful essays on this subject, see Derthick (2001) and also her other writings.
[3] See the attached definitions by Hatry, 2001, p. 19. This part of the discussion concerns "units of analysis," as well as measured goals. For some government programs, performance management goals involve *individuals* as the units of analysis (students, job seekers, sick people). For others, and for reasons that often reflect inherent limitations of measurement, we settle for (or may even decide we prefer) *organizational* goals as the units of analysis: Did service providers do what they were supposed to (for example, social security offices, state food stamp programs)? The subject of organizational performance is treated in Nathan (2000), chapters 6, 8–10, which describe institutional evaluations of the implementation of national welfare reform policies.

THE GAMING CHALLENGE OF PERFORMANCE MANAGEMENT

The Office of Management and Budget in its on-line documentation on the PART system says it is not primarily a budget tool. I agree with this idea. The PART system should not be sold solely as a method for deciding that such and such a program doesn't work so we should cut it, or that it does work and we should add resources. That is not to say that weak programs should be retained or that they should receive more resources. There are situations in which performance findings should influence budgeting. But if PART is used mainly for this purpose, the likelihood is that it will lose its managerial efficacy.

This possibility of gaming raises difficult questions for performance management involving how to set performance goals so that they influence agency behavior in the desired ways. Doing this requires both hard and soft accountability measures. In private industry, soft decision factors are often taken into account in rewarding managers. Economists have affirmed the wisdom of blending objective and subjective performance measures in this way (Baker, Gibbons, and Murphy, 1994). Unfortunately, however, the preponderance of attention and literature on managerial oversight in government has focused on rigid numeric goal setting.

In this intense 2004 pre-election period, it is not easy to envision the kind of consultative and interactive goal setting that could wisely and in evenhanded ways bring qualitative and policy related decision factors into play in performance oversight. This is made harder right now by cuts that have taken place in the managerial staffing of domestic programs, because overseeing such goals is necessarily labor intensive. Moreover, countering the rise in rating-mania by bringing qualitative and political variables into play in performance oversight is hard to prescribe. Yet, the truth is that, like speaking prose, we do it all the time, as suggested by some of the examples discussed next.

SOME EXAMPLES

In the field of employment and training, random-assignment demonstrations have shown that focusing on jobs is effective in aiding low-income, low-skilled people, especially women. However, performance-management goals focused on these dependent variables involving, for example, job placement and tenure can have problems. They can produce cream skimming, that is, selecting the most job-ready people to be aided, people who would have found jobs anyway. Corrections for this problem have been attempted, but can backfire. In one case I know of, the result was to undermine the public employment programs of the 1970s, leading to their eventual demise (Cook et al., 1985).

In the field of K–12 education, the emphasis on numeric goal achievement has caused unintended consequences that have required political jockeying and managerial recalibration. To paraphrase Tip O'Neil, all education is local. No matter how elegantly formed, the nature of the educational process requires that national oversight goals for local schools be focused on simple and general indicators, such that they tell only part of the story about school performance. For this reason it is important that performance goals for local schools be viewed as motivational every bit as much as being viewed as tools for the close and fulsome calibration of institutional effectiveness.

In the field of health policy, experts have urged the adoption of financial incentives for effective practice and quality enhancement, but at the same time acknowledge that the processes for providing and overseeing such standards face impedi-

ments (Epstein et al., 2004). Furthermore, we are told that such goals are unlikely to achieve their intended results unless substantial amounts of new money are provided (Epstein et al., 2004). Where is the money going to come from?

Citing these examples is not meant to suggest that setting performance goals and assessing how well they are achieved is unwise. The point is that performance management is difficult, can be expensive, and worst of all can backfire. It has to be smart and it has to be flexible, adaptive, and subtle. Stimulating efforts to ratchet up program performance on the part of the armies of governmental and quasi-governmental workers at every level of government who implement public policies requires setting and treating performance goals so that they serve *both* as targets for managers and as symbols for the public, in the latter sense, symbols that are well and widely understood and accepted. Raising student performance by using performance goals to focus attention on shared values about what our schools should teach has a good effect, and yet in the process of doing this we have learned about pitfalls of over-specification and have had to make adjustments. This is as it should be; undue rigidity in setting performance goals can undermine their effectiveness.

A good illustration of the latter point is shown by the recent success of welfare reform in reducing dependency and facilitating employment. The 1996 national welfare reform law dramatically changed bureaucratic behavior. Seemingly, this was rooted in strict and specific goals about jobs, hours of work, etc. But in reality, there was a big loophole, the "caseload reduction credit." This provision of the law, plus others, enabled states to advance their work-first purposes in a manner that reflected varied state and local values and conditions. The process was incremental and typically American. The success achieved was not the result of crafty planning by calculating policymakers. It came about serendipitously in ways that were surprising to us in our implementation studies of the 1996 national welfare reform law (Nathan and Gais, 1999; Gais et al., 2001, pp. 35–69).

MODIFYING THE PART SYSTEM

To reiterate, what is appealing to me about the PART system is that it focuses on individual programs as the basic building blocks, more so than on strategic and often overly elaborate purposes as advocated under previous federal government management reforms (Radin, 2000, pp. 111–135). What is needed, however, is a more candid and flexible treatment of goals in ways that involve a range of scholarly and expert perspectives.

Ann Blalock and Burt Barnow have called for a *partnership* for connecting demonstration and evaluation studies conducted by academic experts with performance management systems:

> Our recommendation is that competent evaluation research, or applied social science research, must be coordinated with or integrated within performance management systems if precise, valid, reliable information about social programs is to be made available to decisionmakers. (Blalock & Barnow, 2001, pp. 487–519)

Blalock and Barnow point out that what they call the "evaluation movement" was developed in the crucible of academia, while the "performance management movement" has its roots in public administration and in administrative bureaucracies. They believe, and I agree, that collaboration between these two movements would yield important benefits.

We recommend that the major direction for the future is to coordinate evaluation research with performance management systems more fully, moving toward full integration of evaluations within performance management. Such integration will require that performance management systems treat evaluators not as aliens from outer space, who land only periodically to study and give advice, but as part of an interdisciplinary team. It will require that evaluators become more sensitized to managers' needs, to have ongoing information for tracking outcomes, and to express the benefits of their professional roots with greater humility. (Blalock & Barnow, 2001, pp. 487–519)

The two "movements" as described by Blalock and Barnow have different disciplinary bases. The impact/experimental movement is essentially Weberian in its assumptions about bureaucracy and implementation: Programs should be precisely controlled and replicated. By contrast, performance management seeks to establish goals as reference points for managers who are encouraged to innovate and change what they are doing in response to continuous feedback.

Just as there are many players in the bargaining processes for policymaking in American political pluralism, multiple players should be involved in deciding and adjusting the outcome goals for performance management.

Who should be the multiple players in outcome goal setting? Program managers at the appropriate levels of government (federal, state, and local) bring a needed perspective to bear about the likely effects of different performance goals in different settings. Researchers and policy analysts can also play an important role drawing on research that shows what programs are likely to work best and what their effects are when they do. But, it isn't enough to know if a program works. The responsible goal-setting officials also need to work their way through the hard questions raised earlier about how to express such findings in politically acceptable ways in performance goals that can't be gamed for undesirable purposes. Another important group of players is elected and appointed political officials; they clearly have a role to play in setting performance goals that reflect executive branch and legislative policies.

There is still another overarching aspect of this question about who should participate in performance oversight involving *auspices*. The GPRA law seeks the joint role of Congress and the Executive, whereas the Bush Administration's PART system gives the strong lead role to the Office of Management and Budget. While someone (that is, some agency) has to be in the lead, the challenge, as I see it, is not so much a challenge involving agency roles as one involving *transparency*. There needs to be a high level of transparency in sharing information about performance oversight. This is needed in the case of the PART system as it applies to other control agencies besides OMB, so that they can participate in performance oversight. The Congressional Budget Office, the General Accounting Office, and the Congressional Research Service, along with state and local budget and management offices and outside evaluators, need to have access to the underlying data that are used and understand how they are used in PART performance evaluations.

In line with the "complexifying" theme of my talk, this challenge of sharing data is made more difficult as more data are brought to bear, and more data *should be* brought to bear in performance management. An obvious and important opportunity is the availability of administrative data. We are virtually drowning in administrative data. Almost all domestic programs have reporting requirements that are extensive and detailed. Yet, despite this fact, and despite the fact that the quality of administrative data varies, not enough effort has been made to clean up these data, to scrutinize them and compile the best sources. Here, the underlying federalism

condition that states are different—that is, laboratories not just of democracy, but also of data availability—creates an environment in which multi-state studies can take advantage of opportunities to use administrative data wisely and more extensively.[4] Making fuller use of the best existing administrative data sets can have the important additional advantage of enabling evaluators to break out data for different groups of program participants in different kinds of program settings.

Another step that I believe would advance the art and practice of performance management would be to *relax* a little, loosen up the rating games. Instead of widely advertised *U.S. News and World Report*–type scorecards and report cards, people inside America's governments would do well to present performance results in more detailed and nuanced ways, explaining why and how they operate under different programs and conditions.

As a political scientist, I feel obliged at this juncture to go more deeply into my reason for putting so much emphasis in this talk on the dynamism, diversity, and fragmentation of the political milieu in which public services are provided. The reason is that so often in government *administration is policy*. Policymaking and administration are intrinsically linked. The day-to-day promulgation of rules and guidelines; the issuance of policy interpretations; the approval of grants; appointments and staffing decisions, all involve the transmission of values in ways that are more than routinely ministerial (Nathan, 1983).[5] The public administration literature is deficient in recognizing that policies and programs are changing all the time. Performance management must take this dynamism into account. It has to be seen as a *continuous* process. It has to be carried out by trial and error. It cannot be accomplished by fixed "one-size-fits-all" managerial formulas.

This is not to say that we should stop working hard on performance rating, only that these ratings should be presented in thorough and nuanced ways, and that they should be explained in statements that describe the strengths and limitations of the technology and data on which they are based.

It was not my intention in this talk to make specific recommendations for modifying the PART system. Moreover, my sense is that this should not be done in a rigid and formal way. Discussions about who is consulted and what kinds of inputs are used in setting performance are necessarily *situational*.[6]

Opening up the PART system (and state and local PART-like performance management systems) to more players and greater variations can also achieve something that is very practical and low tech. It can help deal with the information explosion in public affairs. We are virtually flooded with reports, e-mail messages, books, articles, conferences, and harangues. Performance management systems like PART, which pull together what is known about the effects of public programs, can inform and clarify debates about what America's governments do and the immense, diverse, and complex ways their actions affect social and economic conditions.

I realize that this plea for nuanced discourse, greater selectivity, flexibility, and transparency is wishful. "Complexifying" performance management in these ways

---

[4] Some large cities, notably New York City, have innovative performance management systems, which can provide valuable lessons and insights.

[5] On management tracking, I suggest Heinrich, 2002, pp. 712–725.

[6] The best book that shows how public management is *situational* is *Bureaucracy* by James Q. Wilson (2000 edition). Wilson describes organizations as needing "to acquire sufficient freedom of action." Based on his teaching and research, Wilson said he discovered that what was most needed was to permit an organization "to define its tasks as it saw best and to infuse the definition with a sense of mission." See chapter 2, "Organization Matters," pp. 14–28.

214 / *"Complexifying" Performance Oversight in America's Governments*

requires the application of judgment, and this invariably opens opportunities for political maneuver. Still, the fact of the matter is that we can never depoliticize performance management, nor should we even think of trying to do so. In the final analysis, what is needed is to bring more players into the process in ways that create a wise balance of expertise and politics. The long-term result would be to create a setting in which there is a more realistic focus on program results that gives citizens an honest, believable, and realistic portrayal of what their governments do and how well they do it.

RICHARD P. NATHAN, director, Nelson A. Rockefeller Institute of Government, State University of New York.

ATTACHMENT

Performance Management Definitions

* Inputs: Resources (that is, expenditures or employee time) used to produce outputs or outcomes.
* Outputs: Products and services delivered. Output refers to the completed products of internal activity: the amount of work done within the organization or by its contractors (such as a number of miles of road repaired or number of calls answered).
* Intermediate Outcomes: An outcome that is expected to lead to a desired end but is not an end in itself (such as service response time, which is of concern to the customer making a call but does not tell anything directly about the success of the call). A service may have multiple intermediate outcomes.
* End Outcomes: The end result that is sought (such as the community having clean streets or reduced incidence of crimes or fires). A service may have more than one end outcome.
* Efficiency, or Unit-Cost Ratio: The relationship between the amount of input (usually dollars or employee-years) and the amount of output or outcome of an activity or program. If the indicator uses outputs and not outcomes, a jurisdiction that lowers unit cost may achieve a measured increase in efficiency at the expense of the outcome of the service.
* Performance Indicator: A specific numerical measurement for each aspect of performance (for example, output or outcome) under consideration.

Source: Hatry, 1997, pp. 1–4.

REFERENCES

Baker, G., Gibbons, R., & Murphy, K.J. (1994). Subjective performance measures in optimal incentive contracts. The Quarterly Journal of Economic, 109(4) November, 1125–1156.

Blalock, A.B., & Barnow, B.S. (2001). Is the new obsession with performance management masking the truth about social programs? In D.W. Forsythe (Ed.), Quicker better cheaper: Managing performance in American government. Albany, NY: The Rockefeller Institute Press, pp. 487–519.

Cook, R.F., Adams, C.R., & Rawlins, V.L. (1985). Public service employment: The experience of a decade. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Derthick, M. (2001). Keeping the compound republic: Essays on American federalism. Washington, DC: The Brookings Institution Press.

Epstein, A.M., Lee, T.H., & Hamel, M.B. (2004). Paying physicians for high-quality care. The New England Journal of Medicine, 350(4), 406–410.

Gais, T.L., Nathan, Richard P., Lurie, Irene, & Kaplan, T. (2001). Implementation of the Personal Responsibility Act of 1996. In The World of Welfare, R., Blank and R. Haskins (Eds.), Washington, DC: Brookings Institution Press.

Hatry, H.P. (1997). Comparative performance measurement: FY 1996 data report (adapted). Washington, D.C.: International City/County Management Association.

Hatry, H.P. (2001). What types of performance information should be tracked. In D.W. Forsythe (Ed.), Quicker better cheaper: Managing performance in American government, Albany, NY: The Rockefeller Institute Press.

Heinrich, C.J. (2002). Outcome-based performance management in the public sector: implications for government accountability and effectiveness. Public Administration Review, 62(6), 712–725.

Nathan, R.P. (1983). The administrative presidency. New York, NY: Macmillan.

Nathan, R.P. (2000). Social science in government: The role of policy research. Albany, NY: The Rockefeller Institute Press (reprinted).

Nathan, R.P., & Gais, T.L. (1999). Implementing the Personal Responsibility Act of 1996: A first look. Albany, NY: The Nelson A. Rockefeller Institute of Government.

Radin, B.A., (2000). The Government Performance and Results Act and the tradition of federal management reform: Square pegs in round holes. Journal of Public Administration, Research and Theory, 10(1), 111–135.

Rodriguez, J.F. (2004). The arrival of performance budgeting. The business of government, pp. 56–61 (summer).

Wilson, J.Q. (2000). Bureaucracy. New York, NY: Basic Books.

Performance Management in State and Local Government

O n February 8, 2005, the Rockefeller Institute held a public policy forum on the state and local role in performance management in New York State. The forum was co-sponsored by the Rockefeller Institute, the New York State Division of the Budget, and the Manhattan Institute.

Welcome: Peter D. Salins, State University of New York

Part I: The State and Local Role in Performance Management

Part II: The State Role

Introduction — Henry Olsen, The Manhattan Institute

Agency Presenters:

Division of Criminal Justice Services — Chauncey G. Parker, Director

Office of General Services — Robert Fleury, First Deputy Commissioner

Department of Taxation and Finance — Commissioner Andrew S. Eristoff

Part III: The Local Role

Introduction — Swati Desai, New York City Human Resources Administration

Presentations:

Performance Management in New York City — Dennis Smith, Robert F. Wagner Graduate School of Public
　　Service, New York University

Performance Management at the Human Resources Administration: JobStat — Swati Desai, New York
　　City Human Resources Administration

Performance Management of Child Welfare Systems — Fred Wulczyn, Research Fellow at Chapin Hall
　　Center for Children, University of Chicago
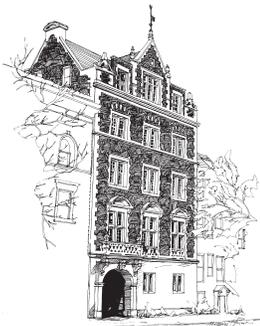
The Nelson A. Rockefeller Institute of Government
State University of New York
411 State Street, Albany, NY 12203
www.rockinst.org