



American Evaluation Association
Public Issues Form

**Evaluation in Federal Agencies: What Shapes It, and How Could the
American Evaluation Association be Part of the "What"?**

Second Annual AEA Public Issues Forum, Friday, November 9, 2007

This publication contains edited proceedings from the first Public Issues Forum sponsored by the American Evaluation Association at *Evaluation 2007: Evaluation and Learning*, held November 9, 2007 in Baltimore, Maryland, USA.

Credits:

Susan Kistler, AEA Executive Director

Michael Morris, Chairman, AEA Public Affairs Committee

Content editing: Damon Thompson, AEA Communications Director

Copyright © 2007 American Evaluation Association. All rights reserved.

Printed in the United States of America.

Table of Contents

Introduction	2
Abstract	3
Forum Participants	4
Panel Comments: Michael Morris.....	5
Panel Comments: George Grob.....	6
Panel Comments: Tom Chapel.....	9
Panel Comments: Patrick Clark	16
Panel Comments: Bernice Anderson.....	20
Panel Comments: Debra Rog.....	23
Questions and Answers.....	25
Additional Resources	29

Introduction

Welcome to the proceedings of the second American Evaluation Association Public Issues Forum. In 2006, the AEA Public Affairs Committee, led by current AEA President Bill Trochim, developed the idea of the Public Issues Forum as a new and potentially useful mechanism for engaging AEA in addressing timely and important issues for our field. A public issues forum is an AEA panel designed to encourage constructive dialogue and to engender awareness among leaders, policy-makers and the general public on issues of pressing importance regarding the role, image or implementation of evaluation in public contexts. The forums are expected to encompass a diversity of views on the selected issue, to address the issue in a civil manner, and to be an educational endeavor, striving to present nuanced and thoughtful multiple perspectives in order to achieve deeper understanding. The topic that the forum addresses is selected for its relevance and importance to the field of evaluation, for its immediacy, and because it is at the interface of evaluation and the public. The AEA Board is supporting these Public Issues Forums for a three-year trial period that began in 2006. The AEA Public Affairs Committee is responsible for identifying the issue, inviting the participants, managing and moderating the forum and disseminating the proceedings. Forums are recorded, transcribed and edited into a document posted on the AEA website www.eval.org along with relevant resources and links.

Abstract

The 2007 Public Issues Forum explores how state-of-the-art knowledge and expertise in evaluation can be more effectively linked to the formulation of evaluation policy at the federal level. Panelists from three federal agencies – the National Science Foundation, the National Institute of Justice, and the Centers for Disease Control and Prevention – address three major questions: (1) How is evaluation policy established in their agency? (2) What types of evaluation-related input would their agency welcome from a professional organization such as the American Evaluation Association? (3) What are the means through which AEA could provide such input? Against this background, panelists also discuss the following: To what extent will the 2008 Presidential election and its aftermath present opportunities for the professional evaluation community to play a greater role in the formulation of evaluation policy? What factors are likely to facilitate or hinder this influence? When a professional organization endeavors to elevate its public profile at the federal level, what cautionary tales should it be mindful of?

Forum Participants

Michael Morris (moderator), is Professor of Psychology and Coordinator of the Master's Program in Community Psychology at the University of New Haven. His areas of expertise include ethical issues in program evaluation and organizational consultation in the human services. Morris has a Ph.D. and an M.A. in Community-Social Psychology from Boston College.

George Grob is President of the Center for Public Program Evaluation and serves as a consultant to the AEA's Evaluation Policy Task Force. He has more than 35 years of professional leadership and public service in the Federal Government, mostly as an evaluator in the Department of Health and Human Services,. As HHS' Director of the Office of Evaluation and Inspections, he oversaw the production of more than 1,000 studies in a 15-year period. Grob has a Master's Degree of Arts in Mathematics from Georgetown University.

Thomas Chapel is Senior Evaluation Scientist in the Office of Workforce and Career Development at the Centers for Disease Control and Prevention. An author of a number of articles and chapters on evaluation, he has an MA in public policy and an MBA, both from the University of Minnesota.

Patrick Clark is Acting Chief of the Evaluation Research Division at the U.S. Department of Justice's National Institute of Justice. He has 30 years of experience in evaluation research in the criminal and juvenile justice systems, and has a Ph.D. in psychology from Michigan State University.

Bernice Anderson is Senior Advisor for Evaluation at the National Science Foundation. Her publications include *Breaking the Barriers: Helping Female and Minority Students Succeed in Mathematics and Science*. Dr. Anderson received her doctorate in education from Rutgers University.

Debra Rog (discussant), is an Associate Director at Westat. She has an extensive background in program evaluation research, having served as director or principal investigator for many studies addressing issues such as homelessness and housing, mental illness, child development, family violence, and living with HIV/AIDS. Rog has been elected the 2009 President of AEA. She received her Ph.D. in social psychology from Vanderbilt University.

Panel Comments: Michael Morris

My name is Michael Morris. I'm the chair of the Public Affairs Committee of the American Evaluation Association and it is a pleasure this afternoon to welcome you to the second annual Public Issues Forum. The forum is actually going to be audiotaped for the AEA website and will be in place at the website, hopefully not too long after this. And, given that we have a new Communications Director in the organization, who knows? Future Forums may be made into an HBO mini series; we have no idea where this could go. So if you would like to begin an acting career, try to get on the forum panel. We have a very exciting topic today. As AEA grows, the role it might play in the development of policy at the Federal level is increasingly being discussed by the Association, and today's session is going to focus on what currently takes place with respect to evaluation at the Federal level - both in a general sense and in some specific agencies - and also what might be the case in the future in terms of evaluation's presence and particularly the presence of the American Evaluation Association.

We have a group of individuals to discuss that who are eminently qualified for that topic, so let me introduce them very briefly to you and then we will begin.

Wendell Primus was not able to be here today, but we have a wonderful substitute for Wendell – George Grob, who was formerly Deputy Inspector General for Evaluation at the Department of Health and Human Services. And prior to that, he was the Director of Planning and Policy Coordination at HHS.

He will be followed by Thomas Chapel at the Centers for Disease Control and Prevention. Thomas is Senior Evaluation Scientist - not junior, not freshman, not sophomore - Senior Evaluation Scientist in the office of Workforce and Career Development at the Centers for Disease Control.

Patrick Clark, National Institute of Justice. He is the Acting Chief - not acting assistant chief - he is the Acting Chief of the Evaluation Research Division at the National Institute of Justice.

Bernice Anderson is at the National Science Foundation. She is a Senior Advisor for Evaluation. Once again, not freshman, sophomore or junior – *Senior* Advisor for Evaluation there.

Our discussant will be the 2009 President of the American Evaluation Association – Debra Rog. She is an Associate Director at Westat and has 25 years of experience in the field of evaluation.

Each of our presenters will have about 15 minutes to provide an overview within their domain and then that should leave us some time for questions at the end. Let me ask George to come up to the podium and we will begin. Thank you.

Panel Comments: George Grob

Wow, it's a real pleasure to see you all in the room and I'm glad to see there's such a great interest in our topic. I'm not going to delve into opinions and advice for you. My job here is to simply give you a factual orientation to set the stage for the other speakers. I want to let you know in as concrete way as I can what we mean when we talk about evaluation policy.

While the topic of Federal evaluation policy seems somewhat vague, it is possible to it when you see it. I used to think about that in the following way: that sometimes when I went to the office, I really would dress up like I am now – my best suit, dark, a tie like this, and my wife would say, “Oh, George, you going to make national policy today?”

The truth of the matter is that at the end of the day, I could tell if I was involved in national policy; I literally could tell. For example, I could say, “Well, today we sent a draft bill over to the Congress,” or, “We prepared the budget for such-and-such a thing.” My goal here today is to give you concrete examples of evaluation policy. That's how I'll try to address it for you.

Evaluation Policy: the short version is that – it's laws and rules. But it is really much broader than that. Rules is a big word here. Now I don't mean just formal regulations. I mean to refer to all kinds of rules, including administrative rules that affect the practice of evaluation. This would include funding, standards, evaluation organizations and requirements for specific studies.

To press the point further, evaluation policies can be quite detailed. They might say, “Here are the questions we want you to answer, here are the methods we want you to follow. This is the kind of person we want you to get to do it and this is how much money we're going to get you to do it and we want it done by this period of time.” Each one of these questions floats down in a natural way to all these details. But those are the kind of things we mean by policy within the government. By the way, today we're only talking about the policies of Federal Government, although the concepts are similar in state and local governments as well.

Evaluation Policy establishes the boundaries within which we practice our profession. I would really encourage you to look at that back page of our manual for the conference. You'll see the statement that Bill Trochim has written for the coming year. He's got great definitions and if any of you were treated with going to the Awards Banquet, Bill also laid it out pretty well there.

Federal policies take the form of laws and regulations, and you can really see if something has happened – a law is passed, a regulation is issued and a regulation is a real document. You can see that, one day, the Secretary or some other senior official signs a regulation. Then there are drafts; they're formal, you

make comments on them. These are very precise things, such as administrative procedure – e.g., what you have to do to approve a Medicare claim; organizational charts – a new office has been created; there wasn't an office there, now there is an office there. At one time there was no Office of Evaluation in the Inspector General's office and then one day there was an office. It is that real. This is not nebulous stuff that I'm talking about; it's very, very real.

Other examples include administrative procedures and requests for proposals, to name a couple. I will give you a good example of one that I saw and was interested in for a short time. It was a request for proposals to make grants for the Abstinence Only program and they had a requirement that between 10% and 30% of the funds had to be set aside for an evaluation and they wanted it to be the random controlled study. And, by the law, it had to be done by an evaluator who was associated with a major educational institution in the state in which the grant was being made. Furthermore, the program office went on to say and it needs to be a PhD. When we say that there are these details that affect how we practice, that's an example. There are many, many more but that's a real, live example of something you might see in an administrative document.

There are still more examples– the budget's a big one. It may prescribe how much money will be available for evaluation for an organization.

Prevailing practices are also a form of evaluation policy. Bill has referred to this. Everybody in an office does things in a certain way.. It is not written down. If you work in that office, you're going to do it that way. They'll tell you.

Evaluation plans – some offices actually issue evaluation plans. They're published. In most of the offices I worked, there was a formal plan.

Policies are established by the Congress and the Executive Branch. In the Congress, much of the work is done by committees, but the first ones, the big ones, are the general government oversight committees. There are two of them – one in the Senate and one in the House of Representatives. Then there are subject matter oversight committees. They look such things as health, education, labor, environment, and the like. And then there are the appropriations committee, which recommend how much money will be allocated for the policies and programs approved by the oversight committees.

Then there is the Government Accountability Office, GAO. That's a congressional office, it's not an executive office. First of all, it exists! They do evaluations. They don't do many audits anymore. It's almost all what they would call evaluations, and which most other people would too. They use the phrase - "performance audit". But it is truly evaluation work, or at least most of it is.

On the executive side, you have the Office of Management and Budget. This office established PART. In addition, each Federal department or agency may

have offices, money, procedures, plans for evaluation., However, some agencies have no such arrangement for evaluation.

There may also be special agencies that the Congress and Executive Branch leaders would call on to conduct evaluation studies. These include the Institute of Medicine, National Science Foundation, the Inspector Generals.

Another well-known example of evaluation policy is the Government Performance and Results Act (GPRA). That law was prepared by those oversight committees for the functioning of the entire government. It took a dozen years to get that law passed.

Another example of evaluation policy is appropriations, including set-asides. A prominent example is the setting aside of 1% of program appropriations to be used for evaluations in many agencies within the Department of Health and Human Services. Still another example one would be the so-called “high stakes” testing policy for schools. That law would have been formulated by the congressional committees that oversee education programs.

Then there are evaluation standards. A prominent example is the GAO standards for performance audits. While they are used by GAO, they are also followed by many State and Federal Agencies. In a similar vein, the Inspectors General have issued quality standards for evaluations and inspections.

Examples of Executive Branch policies include the Program Assessment Appraisal Tool. It was issued by the Office of Management and Budget OMB) and applies to almost all Federal programs and agencies. Other examples are OMB standards for and clearance of surveys, standards for the protection of human Research subjects, issued by regulation. In addition, there are evaluation units within Federal agencies. I worked in one for many years—the Office of Assistant Secretary for Planning and Evaluation at the Department of health and Human Service. Furthermore, some of the programmatic agencies within large federal departments have offices dedicated to evaluation, including the Centers for Disease Prevent and Control, the National Institutes of Health, and the Health Resources and Services Agency at HHS, to name just a few.

Another example of evaluation policy is the requirement for the used of randomized control trials for the assessment of the impact of education programs. That policy was established by regulation. It was not required by law. The Education Department came up with that rule.

Policy and practices vary considerably. Some policies apply across the whole Government – GPRA, PART, the clearance of surveys. Other policies, such as those relating to how much funding is provided for evaluation, whether there is a formal evaluation office for an agency, whether these is a requirement for evaluation plans, how studies are published vary from agency to agency.

Panel Comments: Tom Chapel

I'm Tom Chapel; I'm with the Centers for Disease Control and Prevention. What I am, as mentioned, is the very, very Senior Health Scientist for Evaluation. This is a disclaimer: Much as I'd like to be the spokesman on behalf of evaluation, as I always tell people, I'm more of the poster child. I'm the living embodiment of what we're trying to do with evaluation. As we'll point out, one of the things lacking in CDC is a strong home base for evaluation and there's ways that operates to our benefit and there's ways in which that becomes trouble – problematic.

Today I'm going to address the three questions that were set out for us by our Panel Abstract: who sets evaluation policy in the CDC environment, what can AEA do to influence the evaluation of policy, and what are some particular points of influence? The first thing to understand is that public health is very decentralized and CDC itself is very decentralized and so a short answer is, there really is no kind of home base or center which really sets evaluation policy. There's a variety of forces, that I'll go into in a second, that tend to shape how it plays out center by center, program by program.

To a degree there's anything kind of framing how we approach it that might make our approach distinctive, I think it would be our CDC evaluation framework. This was a very comprehensive and inclusive effort in the late 90's to rethink how we did evaluation. It was motivated, not so much by, or just by, our desire to do better evaluation, or more evaluation, but as much by the desire that the evaluations that we did would actually make a difference. They would actually be used for program improvement.

With that in mind, there's nothing very magical about it and it looks probably like any other applied framework that you've seen, but it was fairly influential in changing how we thought about evaluation – particularly by introducing the idea of using logic models and understanding your program, and particularly this idea of engaging stakeholders, thinking broadly about who cared about the program, who could actually make a difference. In our framework, we also adopted the evaluation standards, which we did not create but we adopted from others, and that's been very influential in asserting and forcing the point that there's no one right evaluation. There's no one right set of evidence, or way to analyze things, but things need to be case specific. Now, the good side of that is that it's made evaluation much more case specific, instance relevant, and probably led to higher use. The downside of it is that, consequently, the application of this framework is going to look very different, center by center, program by program, instance by instance, and so it becomes hard to typify CDC evaluation.

Who sets CDC policy – really normative practices, standards of practices, peer pressure to do evaluation in a certain way? In the past we probably had something closer to a home base for evaluation. We probably don't now. There are several places *not* contending for that role, I don't know that anyone's

contending at all, which is perhaps part of the problem. But there are several parts of the organization where this kind of thinking might be housed in the future. The forces shaping it [evaluation practice]- meaning when you look across our center's programs and divisions and, of course, our many front line partners doing public health work - what kind of distinguishes why evaluation looks the way it does and there's these 5 kinds of big clumps of forces or influences. I'll talk a little bit about how AEA can find teachable moments to influence those forces or to mold those forces.

The first is that CDC is a science-oriented agency. Its roots are in epidemiology. It started off as a communicable disease center, with a very strong epidemiological emphasis. That's not routinely the case anymore. A lot of what we do tends to be programmatic, but if I think about how evaluation plays out and particularly what most consultations across the agency are like, they can be very different in parts of the organization, with a very heavy science component, or where they see the production of science or translation of science as one of their major missions. And, because epidemiologists, while they may not bring randomized control trial thinking, have their own versions of very rigorous methods of cause, distribution, etc, which are often exactly what the doctor ordered for the sorts of evaluations we do. [But in] other cases [this approach runs] aground on the reality of what we're trying to do at the front line.

The second issue of science orientation is that we have so heavy an investment in and identification with epidemiology and surveillance that it's not uncommon - and this was a problem for a while, especially in infectious diseases - for people to see surveillance as evaluation. Surveillance, for those of you who are unfamiliar with public health terminology, doesn't mean the kind of thing we're worried about these days, but it means the reporting on distal outcomes, reportable diseases, how many people in a community have an STD? What are the risk factors that may be exhibited related to it? How many people have HIV? Those sorts of things, and there's a tendency to see those....first of all, God bless us that we have that distal outcome information, but because we have that, there's a tendency sometimes to see that as evaluation. And when things are going right, we're doing well on STD's, we're doing well on HIV or whatever, it works out swell, it can actually be evaluation. When things are heading south, we kind of lack all that intermediate and short-term outcome information that allows us to improve our programs.

(2) A second thing, and I don't quite know how it plays out, but in certain issues in the organization about how evaluation is practiced depends on where evaluators sit, and they sit in one of a couple of places - sometimes they sit in their own unit, so they're closely allied with the research and science side of the aisle. In other cases, they sit firmly in with the programmatic folk. I don't know that one is better than the other, but I can say that when I see evaluation practiced at front-line level with grantees and partners, that the influence of where the evaluator sits can often be evident in that.

(3) The third force, and these are not in order of importance, but the third force, (and George alludes a lot more of them than I even thought of. As I was sitting there, I was thinking, *oh, so that's where that stuff comes from*) is the many, many external forces, and, of course, we think of GPRA and PART, but also for us at CDC, legislative mandates, authorizing legislations for program XYZ, which of course, then cascades down through us as our RFP's which go out to our grantees and partners etc. I would say, in general, that [GPRA/PART] tends to be less influential on how evaluation is practiced at CDC than perhaps it might be in other places. As my friend, Michael Schooley - who may be here - said in comments when I asked him to look at my talking points, that these things tend to be general in orientation and very specific in practice, but also very different in practice by practice by practice.

So GPRA and PART, for example, as a citizen, I just celebrate those. It's very hard to dispute the logic of the orientation and motivation behind them. As they're applied in practice, they tended not to be terribly influential and also, I think - this is very disturbing - that they're not seen as evaluation. While they may influence the practice of data collection and the proof of program performance, it's not like people then go back and say, "Well, that was really an evaluative experience."

I saw a friend once say, after going through the PART experience, who said: "This has really been good, because it made me really ask hard questions and I'm going to take that back and think about it." But that's the only person I've ever heard say that in my time working on PART. Most people think of it as, "Good Lord, what do they want? How can I get this done? How can we meet the needs of this and then move on with my life?" So I'm working hard with the folk who now own the PART and GPRA stuff because they just lament this fact. They know we're sitting on all this great information. How is it we can get people to see this as evaluation and useful information?

(4) The structure of public health is a huge factor in this and I don't know we're unique in this, but I think it is definitely a factor. It's not only [that] CDC is decentralized but remember, the way public health is practiced in the United States is through a network of about 3,000 local health departments. Of course, money is funneled from us sometimes directly to them, more often through a series of states, territorial, tribal departments etc. What that means is that our programs tend to be very, very cognizant, and if they weren't, our partners would make sure that they were, of the limitations of what you can do at front-line level. And so what you end up with is a very constrained resource environment -either the inability to have the skills and capacity to do much more than sometimes basic performance measurement, even if we were to ask for something very rigorous in our proposals, and where people do have that capacity, a resentment that, gosh darnit, I've got to divert this money [to evaluation]. Every once in a while, the question of 1%, 5%, 10% going to evaluation comes up, and those

proposals often don't get legs on the programmatic side of the aisle because people understand how it's going to play out at the front-line level.

(5) Finally, and this is probably last, but not least, because it actually is the most major force I'll talk about when we talk about teachable moments: the CDC's current strategic planning efforts, which have been going on for a reasonably long period of time. They've had various iterations, [and], like all strategic planning efforts, they're taking probably longer than anticipated. This was an effort by the current director to get us as an organization to align ourselves behind some limited number of goals and objectives to get our programs to align to those, to try and define best pathways to those goals and objectives and then to measure that stuff and use it as an organized reflection.

There's a whole bunch of associated measurement with that, and here's the irony – well, the lament and irony – it does not, at this point, shape evaluation, it shapes data collection, it shapes reflection, it has everybody running around thinking about what they need to do better, but those are not seen as evaluative questions. And that's probably the biggest lament I have in my life right now, is that this performance measurement and metric stuff is growing up – it's important but it's growing up – with a whole different crowd of people than the crowd that I'm used to dealing with.

One of the things I want to talk about, when we talk about how to enter into these conversations at CDC, is how to grab these questions and make people see them as evaluation questions, as paths that have already been tried by evaluators long before the corporate consultants and others came in and had us thinking about these questions.

[On the issue of] helpful input – there's a whole bunch of stuff that would be useful to us, both as an agency and especially in trying to deal with front-line staff. First off, an endorsement that we are right, the realization that focused evaluation means, by definition, [that] there's no right evaluation, but rather that it's an application of standards case by case. There's a matching of standards and approaches to a situation. Having said that, what that means is, we have as many cases where we really need really strong rigorous models and have sometimes the lack of capacity to do it, but also the need to kind of push that models of credible evidence can be quite broad, and models of causal attribution can be quite broad and aren't always RCT's. The discussions that we've had within AEA the last couple of years about causal attribution and experimental models as a goal standard, I think those are right on target for some of the conversations we'll be having in CDC once the dust clears on goals and objectives and people start turning an eye really towards what does it mean to have performed well? What does it mean to prove that we made a difference, an impact?

The application of PART and GPRA, I think what AEA did in commenting on PART and the ways in which it can be used more effectively, was exactly on target. Likewise, the comments the AEA made – what George alluded to – on the Department of Education and the desire, the pressure for randomized control trials. I thought that was very, very helpful and I think can be very influential in a place like CDC, by offering credible experts to come and endorse - especially for the science side of the aisle – [approaches that] may seem sometimes as a too “loosey-goosey” approach to data collection and analysis.

Then finally, we have a big problem. Which is, as I said, on the one hand there are folk who see surveillance as evaluation. And of course, we don't want to preach that – as valuable as surveillance is. On the other hand, we had a panel earlier today where we talked about this challenge of how important performance measurement is, how much brand it has taken on, and yet the ability or the danger that performance measurement will trump evaluation. I think as evaluators, it's going to be very important, as credible experts to come in and say, “Look, this is a continuous cycle of improvement, there's a role for all these things to play.” But none of them trump evaluation; they're all part of the evaluative puzzle.

[Related to the question of] means for providing input, I call these looming teachable moments. The first one is the one about which I have the greatest hope – I'm sorry. I don't know if I have the greatest hope, but it holds the greatest opportunity. And that is the goals and objectives have been rolling out for a while, but I believe there's light at the end of the tunnel and that there's enough clarity on them that they will start rolling out and people at divisional level will need to start thinking about, *if that's what we're going to be as an agency, how do I fit in, how do I align?*

This idea of cascading down, aligning up, I think it's on the tip of most people's tongues, and it comes with a whole bunch of associated reporting. Right now, the associated reporting is probably not well liked, it's very high level, very complicated like many MIS systems are, but I think it's going to get better, and if evaluators can be there, I think most of what we have already talked about – logic modeling and intermediate outcomes, etc, will be very helpful.

What's happened in the absence of evaluators being present at that discussion – it's not like someone's closed the door on us, they haven't even thought about this., then I guess it's not much different than the questions that we broached with the framework—is a whole bunch of corporate and management thinking is brought to bear, and what that's done is that words are being introduced, which have their equivalents in evaluation, but are being introduced on the wrong side..

As I said this morning, I have to talk about “metrics”. No one talks about “indicators.” I can't talk about “intermediate outcomes”, because the management consultant talks about them as “success factors”. You don't have “logic models”,

you have “project charters”. Things like that, which make it seem like this is the first time we’ve done this, [that] this is a brand new question. And so I think probably teaching people that this path has been trod, there’s light at the end of the tunnel, please listen to us, we have a way forward, would be very valuable.

Second thing, the CDC framework is ten years old; I’m a huge fan of it, [and] not just because I’m paid to say that. I really think it served us well, but interestingly, its blessing is also its curse. It’s so open-ended because it needs to be, we want people to be case specific, but the downside is now that programs have been around for a while, they’re actually looking for more guidance. They’re looking for more structure. In this very open-ended landscape, can we help with focusing? How many stakeholders are too many stakeholders to engage? How long should a logic model take? How detailed should it be?

You say an evaluation focus can be about anything. It really can’t be about anything, can it? Shouldn’t we demand that it always have process and some outcome orientation, etc.?

A good friend of mine, Goldie McDonald, and I were talking about it, and she’s much less of a fan of our approach than I am, and I think what we’ll probably try to do is engage some conversations this year. Jumping ahead to my slide about opportunities for input, this is a place where experts should be coming in and saying, “This is what we like about your approach, this is what’s perilous about the approach, there are [something equivalent to] “AAA Trip Tiks” through this landscape and here’s how you match the characteristics of a situation into a strong evaluation approach.”

Peer review is a huge issue. Our current Director, Julie Gerberding, was a huge fan of peer review, has really moved us to extramural research programs, etc. It doesn’t happen that often, although a lot of evaluation stuff gets subjected to clearance. A lot of evaluators spend their lives trying to get their project classified as not evaluation, but - this is the kind of organization we are - things are either “research” or – the rest of the world – “not research.” So we try hard to justify that we’re on the “not-research” side of the aisle, mainly to avoid a lot of the processes that we talked about. Nevertheless, when things go up for clearance, that’s often where a lot of the science bent of the organization comes to bear, often for good, sometimes for ill - a mismatch with an understanding of what we’re trying to do versus the very typical, rigorous science that we apply. Having said that, it could be that this is exactly the time to start thinking about “peer review,” very loosely held. Looking at the portfolio of our evaluations, trying to look at them and see if there’s some pattern, that as external experts, we can endorse that says, “This is a good approach. That’s probably an approach that’s a little bit wanting.”

Last, but not least, publishing in “our journals.” I dare say, most of the folk doing the performance measurement at the organization do not read the American

Journal of Evaluation. Perhaps they read New Directions for Evaluation, if I've left it on their chair. But the American Journal of Public Health. [also] we have some internal in-house journals. There was a big issue on evaluation a couple of years ago. We probably ought to instigate something like that again. And that would be a really wonderful opportunity for big names in the field to influence people who read these broader vehicles and are in our leadership.

[Related to] challenges, there are two challenges: one is, there's no home base for evaluation. There's a lot to celebrate about that. I'm sure we're all in situations where, having a home base on anything, whether it's budget, legislation, science, can be as confining and constraining- establishing an orthodoxy that can be unhelpful - as it is liberating by giving you guidance. Having said that, though, I think trying to formalize much of the informal network that we have - perhaps turning our evaluators into an official working group, trying to reconstitute the evaluation working group, that established a framework way back when— I think would be useful right now because I think it would give a point of advocacy or a point of entry for outside experts to help us advocate for evaluation.

The hugest challenge I see is the most ironic – CDC loves consulting and external advice. Maybe even to a fault at times, but it's good news that our organization's very open to external advice. There's not an internal orthodoxy that people are wedded to. The issue for us is, as I mentioned, these issues facing the organization - which I think eventually will lead to evaluation “brand” down the road if we channel it right—are not seen as evaluation questions.

When I was at MBA School, we used to talk about primary versus secondary demand. Primary demand: milk is good for you. Secondary demand: buy Sealtest [brand] milk. The issue for AEA is that we have secondary demand, but we don't have primary demand. The questions being asked in the organization right now are not framed as evaluation questions. They're framed as performance measurement questions, corporate questions. How can I be more corporate? The first big challenge is to move that around so that people see this all as part of program reflection - an organized process of program reflection, and that evaluation, performance measurement, those sorts of things are all part of the same puzzle.

Once we solve that, it creates a demand for, “If this is about evaluation, how can I find the world's best evaluators to help me out?” Of course, we have them at AEA and it gives us that entrée. Thank you, those are my setup points and I'm looking forward to hearing my colleagues and some discussion later.

Panel Comments: Patrick Clark

Good evening. Thank you for having me. I'm Pat Clark and I'm on the evaluation policy staff at the National Institute of Justice. It's a pleasure to be a part of this effort. With these first slides, let me take you through how evaluation policy might be developed in the Department of Justice (points to slide with organizational chart of the Department of Justice).

Some say that this is the largest law firm in the world, and over there on the left hand side, you'll see the Civil Division under the Associate Attorney General, who's currently acting, under the Deputy Attorney General, who's currently acting...and the Attorney General who may be appointed any day now. The Office of Justice Programs is where the National Institute of Justice is situated and, as you can see, under the Office of Justice Programs, the National Institute of Justice is embedded in the programs and bureaus of OJP. We consider ourselves to be the research, development and evaluation arm of the Department of Justice, but it's a very small arm in that regard. We are to consult, and often do assist the program offices in evaluation work, but that's often something that we wait for.

Here is an organizational chart of NIJ. I'm situated in the Office of the Director, and I'm now what's referred to as an Evaluation Specialist. On the right hand side of this chart, you'll see the Office of Science and Technology (OST) in the National Institute of Justice. They are primarily in the business of forensics and technology development, and follow a scientific R&D kind of approach.

Over the last few months, since I've been moved from the Office of Research and Evaluation to the Office of the Director, we've engaged OST in a process of looking at evaluation in more in-depth terms than what they're used to. In R&D and technology development, evaluation is often a matter of usability and human factors, and what I call "Dick Clark" evaluation. That is usually the extent of their evaluation research – it's easy to dance to, it's got a good beat. Well, we're now in the process of discovering with OST what outcomes might be evaluated.

Policy in the National Institute of Justice is primarily set by the Director, who is appointed by the President and confirmed by the Senate. Dr. David Hagy, who you saw at the top of the previous chart, has been nominated, but has yet to be confirmed. But, a large portion of the policy regarding evaluation that we have in place, in use, I think is a function of outside pressure. And, our good friends at GAO are responsible for some of that pressure. In recent years, they've issued three reports critical of evaluation research at the National Institute of Justice, as you see by these titles (on the slide) here. I'm sure you can pick up these reports at your local bookstore, but in no uncertain terms, they sent a message to NIJ to improve our evaluation research.

My position was created as a result of these reports. So, I'm actually quite thankful to GAO for that nudge - To assess evaluation processes, develop approaches, and develop appropriate strategies. And, I'm all about that these days.

Following that in 2005, the National Academies, Committee on Law and Justice formed a committee on improving evaluations of anti-crime programs, primarily looking at National Institute of Justice work. We were nudged by them with 23 recommendations to improve our work in evaluation. That's having an impact on the way we do evaluation research.

Also, in the spirit of full disclosure, there is this thing called, PART. It brings a tear to my eye to tell you that the NIJ is rated by OMB as only "adequate."

What types of evaluation-related input would NIJ welcome? I've been thinking about this since receiving the invitation to be part of this forum. I think AEA is doing a fine job in that regard. Many of our staff attend your conference on an annual basis. In recent years, there have actually been some resources for staff to get in-service training and they attend the workshops at AEA. I was left with putting your charter up for this slide to say, in reaffirming your charter, you are doing a fine job. But, if you want to break into new areas, I may be able to give you some ideas.

We spend a lot of our time at the Institute for Justice trying to evaluate programs that are enacted through the process that was laid out here. And, in the area of crime and justice programs, those enactments usually come through crime legislation. You may have heard of the last large one, the 1994 Crime Act of the Clinton Administration. That actually created a number of offices, including the Office of Community Oriented Policing Services and the Office of Violence Against Women. The Bureau of Justice Assistance (BJA) was probably the first crime program office and was created under the 1968 Omnibus Crime Control Act. They're primarily responsible for what's referred to as "Byrne grants." These are block grants to local governments and police departments. They spend these funds across about 30 purpose areas.

That gets me to the challenges that we face - and I'm whining here - as it relates to evaluating criminal justice policies at the Federal level. Federal policies, at least in criminal justice, are often enacted, funded and implemented without any regard to evaluation. It's a political process that results in a crime act. Crime act formula funds are distributed through agencies like the Office of Victims for Crime, and the Office of Violence Against Women with an array of purpose areas that are very difficult to operationally define and evaluate.

Block grants are similarly awarded across many objectives without any provision for evaluation. And federal programs, especially in the area of crime and criminal justice, overlap in jurisdiction and in services. In one location there will be a little

bit of money applied toward violence against women, maybe to hire a prosecutor. You'll have a victim's compensation board. You'll have a community oriented policing service. And, it's difficult to assess the effect of any of them in terms of outcomes.

Specific program funds are marginal, minimal at best. A prosecutor is a good example. For the Violence Against Women-STOP program, a local jurisdiction can provide a prosecutor for about \$85,000. But \$85,000 a year may not make a lot of difference in terms of outcomes. Funds for the purposes of evaluation are often limited and seldom sufficient. Our base for discretionary spending on research grants will probably be about \$10 million this year. And unfunded mandates are very popular and could unexpectedly use some of that discretionary base.

Lastly, earmarks are crossed out on this slide. But, I've heard they're going to come back with names of sponsors associated with them. Earmarks exist with few conditions and no requirements for evaluation. This all makes it very difficult to do evaluation work. Do you feel my pain?

What are the means to which AEA could provide some input? A large portion of what you do is having an impact, especially with young staff who are seeking education in the areas of evaluation. But as alluded to here, we need to help folks out there discover what evaluation is and what it isn't.

Providing objective and independent information in the system is, I think, a laudable effort. And, AEA can function very well in that regard, and educate us on contemporary and cutting-edge and innovative methods. It's not uncommon for NIJ to publish a solicitation, and get a hundred applications in response. And, half of those might propose what Tom Cook calls the "workhorse" (non-equivalent group quasi-experimental) design. We just finished - reported out here yesterday - our first quasi-experiment with propensity score matching - so you are making inroads.

The Policy Task Force is really a great idea, I wanted to reiterate, your efforts are laudable.

Timing of conferences has been a topic of discussion. I've been back and forth from another conference in town this week. And, we have to work on timing of important conferences. I'm leaving town tomorrow to head to the American Society of Criminology. If we could all collaborate on schedules, I think we'd have a heck of an impact. For instance, the Association for Public Policy Analysis and Management (APPAM). One of the evaluators that I've met through my association with Stephanie Shipman and Federal Evaluators Group is at the APPAM conference this week. Many of the folks in your target group are at that conference right now, and many of them being trained in economics and in public

policy - but they really need your help. NAPHA is meeting this week as well, and their members won't have the opportunity to attend an AEA workshop.

The next slide is an e-mail that I received last Friday. I'm inundated with these constantly. These are offers to come out and be trained with in-service opportunities. I think that AEA could get into the business of doing this kind of one-day forum. Timing is everything, but across the span of a year, staff often have an opportunity to get out of the office for a few days and concentrate on a particular topic. As an example, I just had the opportunity to go to Alexandria, Virginia and attend a workshop on concept mapping. That wouldn't have happened if it hadn't been available in town.

And then there's this place on the Internet – www.expectmore.gov. If you want customers, there's a list of agencies and organizations on that website right now that have been assessed as "inadequate," or who haven't performed, with "results not demonstrated" and they need your help.

How does the presidential election affect... There's an opportunity coming here. The transition is beginning, or has begun. Political people are leaving government daily. Political appointees are leaving and in their place will be career people for some duration of time while the new Administration is moving into place. Career people appreciate what you're talking about and I think you may be able to take advantage of this opportunity.

You could also create new opportunities to interject your message. In another life I was a lobbyist for a citizen's based state-wide charter. Even though you may not have money to buy a ticket to the party, you may still meet with executive and congressional staff. Create teaching opportunities in these kinds of forums and meet with career agency staff in the interim. Go to and participate in forums, and submit some testimony, even though you may not be able to read it aloud on CSPAN, it gets on the record. And, encourage your membership to become active at the local level. Interject yourself in the process of program development, especially at the local level.

What factors may likely facilitate or hinder? How about an administration that is looking for input? How about an administration that is not? I really can't predict which way we're going to go with that. But I wish you all the luck in the world. But, please avoid being used and exploited as it relates to politics. Affirm your Charter at every turn, and I think that will serve you well. Thank you very much.

Panel Comments: Bernice Anderson

Thank you. Evaluation is about change and I will share with you how we are changing at the National Science Foundation (NSF) with regard to how we think about evaluation. As the Foundation keeps pushing the envelope about STEM education evaluation, we always need the help of AEA as NSF aims to have a leadership role in evaluation practices in government. Regarding GPRA and PART, NSF has been effective to date. However, we still need input from AEA to help us maintain our high standards with regard to evaluation.

The first question that we had to address is: how is evaluation policy established in their agency? I am from the Directorate for Education and Human Resources (EHR) and that's the directorate that has the responsibility for STEM education evaluation. Since 1992, EHR has been trying to deal with the mandate that all of its education programs must be evaluated. This Congressional mandate that started with EHR is now impacting the other research directorates at the Foundation because of OMB's PART and ACC activities that we will discuss later.

In addition to Congress asking us to evaluate our programs, the National Science Board (NSB) has taken a close examination of our programs. NSB has recommended that NSF evaluate its education investment and develop new strategies to increase their impact on the quality of STEM education. It is important to point out that only do they [Board members] want us to evaluate the investment; but they are hoping that, because of these evaluations, we'll be able to develop some new strategies. This is in line with your 2007 theme of evaluation and learning. We need to learn from these evaluations some insights for moving the frontier of science.

In addition, the Committee on Equal Opportunities in Science and Engineering (CEOSE) has examined our programs focused on equity and diversity and recommended that over the next two years the Foundation should assess the outcomes of our programs and determine their impact on broadening participation and transforming institutions. We are also asked to provide some capacity-building activities; CEOSE recognizes that, even as you require evaluation to move to a higher standard, you need to build capacity in the field. There is a need to provide additional information to our grantees regarding the reporting requirements so that the annual and final reports that they submit are evidence-based.

The second question was: what types of evaluation related input would our agency welcome from AEA? As I said, we are about change, so we need you to help us think through what we are doing correctly and what we else we need to do to improve our performance assessment framework.

While we think what we are doing is correct, there are a lot of new things that are happening in evaluation and we need your thoughts about new directions. For example, help us think through our practices of linking evaluation activities for various stakeholders (program staff, OMB, Congress, etc.).

You've already heard about GPRA, you've heard about PART and you've heard about the R&D criteria used by our Committee of Visitors (COVs). The Foundation has used expert panels for over 25 years as an effective means for assessing the quality of our merit review process, its integrity as well as the results of the award portfolio. Greater emphasis is being placed on examining the outputs and the outcomes of our projects. In the area of program evaluation, we need a lot of AEA help, in terms of being available to give advice about program evaluation designs and be willing even to conduct project -level evaluations that are appropriate and rigorous. Again, we need your input about a comprehensive evaluation activity that looks at implementation, advances the knowledge base and meets the accountability reporting requirements.

We fund independent evaluations through contracts and grants. When an evaluation study is funded as a grant, the primary purpose is to advance the field. However, just because it advances the field does not mean that NSF cannot use that same information to justify appropriate use of its resources.

Overall, EHR supports the multiple-method evaluation approach that uses mixed-methods. We fund studies using both qualitative and quantitative data. We also believe that one study won't provide you with all the answers; therefore we have a continuum of evaluation activities that is informed by the program development cycle. Most of our program evaluations are conducted for multiple purposes.

We would like for you to help us reinforce in the field that while it is important to conduct evaluations for compliance purposes, evaluations should help our project directors to understand and explain what they're doing. Also, we want to be better Federal program managers in terms of monitoring; and most of all, we want to learn from project and program evaluations. And then based on multiple evaluations, we should be able to use the evaluation findings to inform our future plans for the efficient use of the program resources. Again, AEA plays a very critical part in evaluation by emphasizing the use of evaluation to advance the knowledge base.

Let's talk about hierarchy of study designs used by the Academic Competitiveness Council (ACC). The top two tiers (experimental and quasi-experimental designs), according to ACC, define rigorous evaluation. We don't argue with that, as it relates to an impact study of an intervention. Our issue is that not all of our programs are interventions. Not all of our awards are implementation projects/studies. Because we have different types of projects supported by the Foundation, we need your help to identify a framework that goes beyond the ACC model.

For example, when a program is funding research and development (R&D) projects, should the evaluation framework differ from the ACC model? When do you use or what is the role of the randomized control trials (RCT) in the early stages of development of R&D projects? Our AEA President was most recently at a conference at the Foundation where the discussion focused on projects going through various phases of development and that it was important to not inappropriately use RCTs. We are not saying we're against using experimental/quasi-experimental designs. They have an appropriate place, but we need to know when the RCT is appropriate in terms of the maturity of the project, as well as the type of the project. We need your support in helping us to define the various types of STEM education projects that we have and what are the appropriate ways to document and assess that a particular project or program has been successful and effective.

The last question was about the way that you can provide input. Whenever we contract a program evaluation, we ask for the contractor to have an advisory group. When you get calls to serve on the advisory panel for a particular program evaluation, I would encourage you to say yes because it would be an opportunity for you share your expertise and to advance the field of evaluation. You may get a call serve as a member of a COV, the Committee of Visitors, as EHR takes a more careful examination of the evaluation effort and the results of an award portfolio. Some of the COVs will have a need for more evaluators to serve on the review team. Whenever we have evaluation planning workshops, I will be encouraging the program officers as well as the evaluation staff to look at AEA as a source for identifying evaluation experts to attend these workshops and/or serve as major presenters. When we begin our next series of evaluation seminars, we'll be reaching out to you again, hoping that you will be willing to share what you're learning - how you are advancing the field of evaluation.

This sums it up in terms of the assessment cycle for EHR. Compared to 1992, EHR is in a different place in evaluation with different programs and different needs, but we do think that AEA provides the foundation, the richness, the scholarly information that undergirds the evaluation demands for today and the future. NSF looks forward to your help as we think about innovative evaluation designs, ways to improve our monitoring systems, and increased use of evaluative research and longitudinal studies.

Finally, let's work with the community of practitioners/managers in terms of building their capabilities to help shape designs and use evaluation results. We often think of evaluation in terms of program change and success. It is important to note that at the Foundation, regardless of the results, a part of a successful evaluation is making certain that program managers know as much about evaluation as they should and know how to use the data/results for multiple purposes. This is a new area and NSF wants your help so that we can be at the forefront of the evaluation/accountability discussion.

Panel Comments: Debra Rog

I'm very pleased to be here in front of such an esteemed panel. What I tried to do was cut across rather than go through each one what I saw as some themes and what I think it says for our association. First, I want to thank George also for stepping in at the last minute and setting an overview, which I think was a very good overview, in setting the stage for having us understand more broadly what evaluation policy is all about and, in particular, what he did. And you heard it through a number of the other talks, was the number of external forces that are pushing on evaluations when the time is right for having evaluation policy and for us having a role.

We heard about evaluation policy in 3 different settings and I think really represented a range - from more of normative practice without a formal policy to one that Bernice talked about, with a much more detailed mandated policy. And the settings also have a range of approaches and, although there were different settings, there were lots of similarities. And I think what you did also hear across them, they're all affected by forces - some internal, some external - that push the direction of evaluation. Tom really spoke more clearly about some of the internal forces, in terms of the culture of the organization, in terms of surveillance or coming from some of our public health focus and how that shapes evaluation. We heard a lot about the external forces. GPRA, PART, other things that - I can't read my notes here, - GAO and other forces that helps to strengthen evaluation.

One thing I also heard, going back to thinking about setting the stage and about these forces, is the patience we have to have putting it together. Hearing that GPRA took 12 years to get started makes you know that progress can happen, but you've got to have hope and optimism.

In terms of the roles, there was a remarkable similarity despite the differences in the organizations of the roles that AEA can play, and ones that we are prepared to do. I heard a role as critical advisor - to be a sounding board, to be a support for high standards, to provide a good housekeeping seal of approval when it makes sense and when it's appropriate. To act as peer review in looking at both specific kinds of programs and initiatives and maybe more globally what the organizations are doing. I saw a role as educator. What evaluation is and is not. Again, Tom said it in terms of it's not just performance measurement, it's broader than that. Helping to clarify and educate.

I heard a lot of what we've been hearing in a number of sessions. What is really nice is that none of these organizations say that evaluation is one thing, or that there is some hierarchy of priority, but there is a real need to reinforcing and keep educating how we make method choices. Again, matching methods to questions, matching methods to contexts, matching methods to a variety of those factors.

What we also heard, is evaluators as trainers that go and actually train the people who are doing evaluation, both with the mechanisms we already have in place, but also maybe some going inside and doing some more internal training of folks, both in terms of new approaches and cutting edge approaches as well as some of the standard things that are going on.

Lastly, there was a role that I heard some of being a partner. In particular, one of the examples was maybe we need to talk about joint publications and working together, again to sort of help get the evaluation language and evaluation culture inside.

I did hear some challenges. Again, moving up the demand for evaluation, Tom talked about it being secondary demand, moving it to a primary demand, working through the complexity of some of these organizations. We saw a very complicated chart and where evaluation is in there and started figuring out – we've got three organizations here, we want to go to different policy places, how do you start to negotiate that? Places where there's lack of processes for evaluation, negotiating the different languages and how folks talk about things. The fact that we talked about whether or not, as we're getting into a new administration, the openness that's going to be there and how we're going to have to deal with that.

Then one of the people said at the end, to not be co-opted and to guide against misuse of evaluation and to keep holding to our standards as we start to get into here. So I'm very hopeful listening to this, and with greater involvement and collaboration, we might be able to avoid some of the divisiveness that's been created by some other policies that have been done by having more... either those policies that were misguided or misunderstood. It provides for evaluations that have a greater potential for being methodologically sound, appropriate to the questions and settings and are more likely to lead to good use.

I want to stop there with my comments but I want to start with the first question.

Questions and Answers

Q: I heard a lot of openness, but I still have two questions. I want to know what would be the best way for AEA to start to develop these relationships. How do we identify these teachable moments? I like that topic but how do we know that? How do we strike when the iron is hot? That's one question. The second one is what are the cautions? What do we need to be careful about?

George Grob: I'll venture an answer to the second question. I think that, in terms of the cautions, one of the cautions would be to avoid over-stepping our representative-ness of AEA on particular and detailed matters about which there may be some legitimate differences of opinion. I think we have to be very systematic in how we sort out the advice we might give on very particular questions about methods or approaches at a very concrete level, so that we don't misrepresent the organization as a whole. There's not an easy answer to this question as to how we find the common ground that everyone is comfortable with, but I think that's something we have to do and that's a caution. I think a related caution, too, is that if we begin to broker the availability of resources to help others, we have to find a way to make sure that we're confident, that we're not ending up organizing and offering bad advice, if you will, in a systematic way, so that's an answer to your second question.

I just may approach the first one, but let others answer as well.

Michael Morris: Does the panel want to comment on Debra's first question?

Tom Chapel: On the first one, in our case, there are two or three big actual moments, Bernice referred to several tangible moments, ways in which you could serve in roles that we very well may need you for. There I think it's just a matter of having a mole inside the organization that keeps you attuned to the opportunity. People need some expert advice on evaluation or this panel is being formed of experts and if I can get you in there, would you be willing to comment and speak or serve, or whatever. We're fortunate at CDC, we have one of the few regional partnerships with AEA every summer, and if we do our job and start making sure that the crowd and the audience at the institute includes people who are users of evaluation, people making performance measurement decisions, that's a great opportunity for the right people to be up there preaching and stuff. Beyond that, the idea of using our framework review as an opportunity to reflect on what's good and bad about the utilization focus we've had and it creates natural opportunities for experts from the AEA to come in.

George Grob: Let me add a little more to that. If you pop up a couple of levels and you see who's setting the policy, which might be staff members and members of Congress, or staff members at OMB, or senior staff in the department, some of whom may not be careerist, but we may see some incoming political appointees, or people of that nature, who have an interest in

this topic and I've seen that does happen, is that if we can reach those people and we can explain to them that evaluation has come a long way and that we have a lot to offer, they may believe it and they may want to, themselves, advocate for some systematic improvements within the departments where they work or within the laws that they can fashion, if we can show them that there's value in what we're offering, they may adopt it.

Bernice Anderson: I think another opportunity will be your dissemination efforts regarding this meeting. Get the word out about what has happened at this '07 meeting in terms of new learning in evaluation and lessons that are still being confirmed. Prepare the document in such a manner that it is very user- friendly for program managers, as well as policy makers. The second suggestion is give more attention to longitudinal studies and follow up studies. I know that NASA is interested in longitudinal studies and I hope that NASA and NSF would partner with or reach out to AEA to address this evaluation need.

Patrick Clark: Popping into my mind is in addition to what I already offered – the Institute for Justice used to, in the Clinton Administration, starting in 1991, we used to have quarterly professional conference series at the JW Marriott in Washington DC and we concentrated in that regard, inviting congressional staff as well as executive branch staff to these one -day, half-day series and perspective lectures in the context of criminal justice. They had an impact in the molding of the '94 Crime Act. There are many Congress people who want their staff to be knowledgeable of things like methods and staying abreast of developments in that regard, and if you look at the schedule of Congress and you schedule on the calendar that works with their calendar, it's possible to fill a room for a couple of hours with staff who have a break and can come and learn something new and different.

Michael Morris: Questions from the audience? ¹

Q: I guess the question that I have is that as AEA embarks on really trying to begin influencing federal policy, you know we always take a mixed methods kind of stance with respect to evaluation methods and I suspect that with respect to influencing policy we would want to take a mixed method type of stance and I've heard that from a number of people, and we have representatives on this panel from three Federal agencies, but I've also heard you saying, not only do you have meetings at your own level, but that there are mandates coming at you, you all mentioned OMB, PART and so on. I'm worried about the mix here and what advice, if any, do you have for us as we embark on this about how we should be thinking about...to what extent should we be trying to influence the mandate, versus to what extent we should be providing direct kind of support to people in your position and your agencies and recognizing we're going to have to do all of these things to some degree. What advice or thoughts do you have for us?

Bernice Anderson: I think you do both simultaneously. You juggle.

Patrick Clark: I'll second that. No, at both levels, seriously, I wasn't kidding. As a lobbyist, I had no money, so if you can't buy a ticket to the party, what are you going to do? Honestly, to influence those mandators, it was a matter of putting myself in the room and offering testimony and sooner or later, staff would say, "Let's get that guy here because he might have something to say..."

Tom Chapel: At the CDC, certainly there's this brush fire approach, the lower level can influence the boss, but the boss pays a lot of attention to what her boss and her boss's boss are saying. If you get into the head of the OMB examiners, and I know they change on a regular basis, it gives you a broader sense of what evaluation is and you're a credible spokesman on that. I think that does a lot of good, if you can get in the head of the PART folks and teach them how to apply PART in a case specific way so that it's less of a blunt instrument, then that would be a great benefit and pay off leverage that then seeps down to the practitioner level.

Michael Morris: We have time for one more question.

Q: The Consortium of Social Science Associations (COSSA) actively lobbies to achieve better legislation in all fields for the social sciences. One of the strategies they use, but not the only strategy, is to scan incoming legislation for provisions related to performance management, accountability, evaluation, and research. This is while the legislation is in committee, and affects budget as well as authorization committees. Is AEA considering such a strategy?²

George Grob: That's an excellent strategy and that is something we definitely need to do and we obviously need to be systematic about it, both for the mechanics of scanning, so you don't miss things and also the human factor of people...some people are more in touch with that stuff and so we need to organize ourselves so that the people who are in touch can let us know when things are rising, but I think you're absolutely right about what you're saying.

Patrick Clark: COSSA, as a lobbying organization, may have the resources that AEA, but I'm not quite sure how you're putting resources to this effort, may have more resources in that regard, but monitoring legislation is easier than it used to be, you don't have to sit on a Federal register or on a print news to see what's coming up in the hopper, so some key term searches, as it were, and you're making me think that even with part time volunteer work it may be feasible to stay on top of blue backs as they come down the line. I also am prompted to talk about a little bit of, as you brought up COSSA. The Campbell collaboration, some years ago, actually taking off of another idea, like our perspective lectures has been having every spring a little summit on the hill where they invite legislators to come and staff to come in as they talk about a variety of public policy issues that are very important to the legislature as they're working on these same policies.

The difference is that you're not in a particular subject area, so other than methods, getting an array of human services CDC type, public health, maybe even criminal justice issues to present and hold forums on.

Michael Morris: We have come to the end of our time; I'd like to take an opportunity to say a couple of things. One is to thank your panelists. They did two things that frequently do not happen in sessions. One, they actually responded to the questions that were presented to them. For someone who is as anal retentive about time management as I am, they actually kept to their allotted times for their presentations, I'd like to thank them for that and I would be remiss if I did not offer a special thanks to George Grob. When I say he appeared here on short notice, only my bladder control can tell you how short that notice was that was involved here and I will forever be in his debt for stepping in at this time, so thank you to our panel.

Endnotes

¹One comment and response were deleted due to limitations in the audio recording.

²Since this question was unable to be heard on the recording, the questioner later paraphrased via email.

Additional Resources

Website of the AEA Evaluation Policy Task Force

<http://www.eval.org/EPTF.asp>

New Directions for Evaluation

“Informing Federal Policies on Evaluation Methodology: Building the Evidence Base for Method Choice in Government Sponsored Evaluation” (edited by George Julnes and Debra Rog); April 2007 (Issue 113)

<http://www.josseybass.com/WileyCDA/WileyTitle/productCd-078799734X.html>

(Available On-line to AEA Members after Logging in to the AEA Website)

New Directions for Evaluation

“Promoting the Use of Government Evaluations in Policymaking” (edited by Rakesh Mohan and Kathleen Sullivan); March 2007 (Issue 112)

<http://www.josseybass.com/WileyCDA/WileyTitle/productCd-0787997080.html>

(Available On-line to AEA Members after Logging in to the AEA Website)

The Government Performance and Results Act of 1993

<http://govinfo.library.unt.edu/npr/library/misc/s20.html>

<http://www.whitehouse.gov/omb/part/>

The Program Assessment Rating Tool (PART) Website

<http://www.whitehouse.gov/omb/expectmore/>

The President's Executive Order on Improving Government Performance

<http://www.whitehouse.gov/news/releases/2007/11/print/20071113-9.html>