

2011

American Evaluation Association



EPTF EVALUATION BRIEFING BOOK

This briefing book contains documents solicited and compiled for the American Evaluation Association's Board of Director's Internal EPTF Evaluation Task Force in preparation for the Expert Panel Evaluative Review of the work of the AEA Evaluation Policy Task Force (EPTF).

TABLE OF CONTENTS

Preface.....	3
AEA Evaluation Policy Task Force Charge	4
EPTF Task Force Member Profiles	10
Performance Summary - AEA Evaluation Policy Task Force July 2007 to November 2011	13
I. Introduction	13
II. Consultative Campaign (Advocacy)	16
III. Public Presence Initiative (Branding)	32
Appendix A: Assessment and Recommendations to the AEA Board of Directors – July 2009	36
Appendix B: 2011 PowerPoint Summary of EPTF Activities and Accomplishments	45
Appendix C: An Evaluation Roadmap for a More Effective Government, September 2010	55
Appendix D: Evaluation Roadmap Citations and Use Record	68
Appendix E: Shea Cover Letter, March 2007	72
Appendix F: Comments on What Constitutes Strong Evidence of a Program’s Effectiveness? March 7, 2008.....	74
Appendix G: Orszag Cover Letter, February 3, 2009.....	87
Appendix H: Program Evaluation for Management and Results.....	88
Appendix I: Comments to Senate and House, June 2009.....	89
Appendix J: Comments on Paperwork Reduction Act, December 2009	103
Appendix K: Comments on Home Visitation Program, August 2010.....	110
Appendix L: Hydrofracking Letter, March 2011	121
Appendix M: Consultants Log of Outside Meetings and Contacts.....	122
Appendix N: AEA Newsletter Policy Watch Columns	156
Appendix O: Recommended Processes for AEA Board and Member Involvement in Public Documents on Evaluation Policy.....	187
Appendix P: Draft Talking Points	189
Appendix Q: Program Evaluation for Management and Results.....	194
Appendix R: Statements Approved by AEA Board, July 2010	195
EPTF Budget Expenditures	196
Evaluation Policy Advisor’s Quarterly Activity Reports.....	197

PREFACE

This section purposefully left blank so that the AEA Board of Director's Internal EPTF Evaluation Task Force may place this document within a broader context and provide further guidance as it feels appropriate.

AEA EVALUATION POLICY TASK FORCE CHARGE

June 2007 (Scroll down for Updates as of July 2009)

Available online at: <http://www.eval.org/EPTF.charge.asp>

BACKGROUND

At its Winter 2007 meeting, the Board of Directors of the American Evaluation Association (AEA) discussed its interest in the Association enhancing its ability to identify and influence evaluation policies that have a broad effect on evaluation practice. To that end the Board established an Evaluation Policy Task Force that can advise AEA on how best to proceed in this arena.

In advance of the Task Force, an Advisory Group was convened over the Spring of 2007 to address the following tasks:

- Discuss the boundaries and scope of the Task Force
- Develop a draft charge for the Task Force
- Identify and recommend potential Task Force members

The results of the Advisory Group's discussions informed the current Charge to the Evaluation Policy Task Force.

SCOPE OF WORK

The Evaluation Policy Task Force is appointed to assist AEA in developing an ongoing capability to influence evaluation policies that are critically important to the practice of evaluation. The following outlines the intended scope of this work.

- **Focus on Evaluation Policies.** The Evaluation Policy Task Force should focus on *evaluation* policy, rather than policy in general. While evaluation can help inform substantive policies in a wide range of areas, and this is a recognized central purpose for much evaluation, influencing substantive policy is not the emphasis of this Task Force effort. Instead, the Task Force should concentrate on *evaluation* policies. Examples of general areas of evaluation policy might include (although may not be limited to) policies regarding:
 - **Evaluation definition.** How, if at all, is evaluation defined in an agency or in legislation? In such contexts, how is evaluation formally distinguished from or related to other functions such as program planning, monitoring, performance measurement or implementation?
 - **Requirements of evaluation.** When are evaluations required? What programs or entities are required to have evaluations? How often are evaluations scheduled? What procedures are used to determine when or whether evaluation takes place?
 - **Evaluation methods.** What approaches or methods of evaluation are recommended or required by legislation or regulation, for what types of programs or initiatives?
 - **Human resources regarding evaluation.** What requirements exist for people who conduct evaluations? What types of training, experience or background are required?
 - **Evaluation budgets.** What are the standards for budgeting for evaluation work?

- **Evaluation implementation.** What types of evaluation implementation issues are guided by policies? For instance, when are internal versus external evaluations required and how are these defined?
- **Evaluation ethics.** What are the policies for addressing ethical issues in evaluation?
- **Focus on U.S. Federal Evaluation Policy.** At least initially, it would be desirable to focus this Task Force on the Federal level because evaluation policy decisions at that level have broad implications for many AEA members and for the field generally. It may be possible in the future to expand this effort to other arenas (nonprofits, state-governments, businesses, international, etc.), but we expect that initially at least the effort will be focused on Federal policy.
- **Target Only a Few Specific Legislative and Executive Policies.** It will be necessary to limit the effort to one or two potentially important evaluation policy formulation opportunities on both the congressional and administrative side of the Federal government. For instance, we might identify one or two major pieces of legislation and one or two areas where administrative regulations are likely over the next few years.
- **Focus on Selected Substantive Areas.** The workgroup anticipates that it will be necessary to focus the Task Force’s work on evaluation policy within a limited set of substantive areas. For instance, the fields of education and health evaluation are areas of primary interest to a broad range of AEA members where AEA could build on its current strengths. The Task Force should focus its evaluation policy shaping activities in substantive areas such as these that are related to AEA’s history and current emphases.
- **Be Appropriately Opportunistic.** Given the preceding thoughts about limiting the Task Force’s work, we recognize the value of remaining flexible and of responding (thoughtfully and selectively) as unexpected opportunities arise that might fall outside the stated initial scope. The Task Force should look to take advantage of such opportunities and should consult with the AEA Board as such arise.

TASKS AND RESPONSIBILITIES

The Task Force will advise and assist the AEA Board, Executive Committee and AEA staff in addressing two primary parallel tasks. The first task will consist of a focused, targeted *consultative campaign* designed to identify and provide useful consultation to U.S. federal legislators and executive branch staff on legislation and regulation in a select number of areas. The second and parallel task will be to develop the *public presence* that AEA projects in direct support of the consultative campaign, in particular through the development of collateral materials that can be utilized in that campaign, and perhaps through other forms of outreach as well (e.g., website presence).

The primary role of the Task Force will be to provide overall guidance on these two tasks. *The Task Force is not responsible for implementation of the tasks*, although some Task Force members may be directly involved in the implementation. To accomplish the work envisioned here, AEA intends to engage the services of a consultant who will play a major role in implementation of the consultative campaign. The Task Force is intended to advise and support this consultant, and to make policy recommendations to the AEA Executive Committee and Board on both of the tasks as appropriate.

Consultative Campaign Task. The primary purpose of this task is to influence federal evaluation policy in a manner consistent with the Mission of AEA in select areas identified by this Task Force. This task is intended to be a relatively “quiet” effort designed to develop experience and expertise on how best to influence federal evaluation policy directly.

Specific sub-tasks on which the Task Force will provide input and guidance include:

- Identifying and summarizing attempts to influence federal policy by other professional organizations that are similar in nature and scope to this effort.
- Identify and coordinate with evaluators within the Federal Government who are connected to the specific evaluation policy context.
- Advising the AEA Board on legal and ethical issues and concerns associated with efforts to influence evaluation policy.
- Developing processes for engaging and communicating with the AEA membership on this evaluation policy effort, including communicating with membership and inviting their input at sessions at the annual conference of the American Evaluation Association.
- Formulating and making recommendations to the AEA Board as appropriate.
- Consulting on evaluation policy planning processes, including methods for: scanning and identifying prospective policy opportunities; managing networks of contacts; entering into and managing specific policy influence efforts; responding to requests for input on policy formulation; monitoring policy efforts; and managing relationships with any consultants engaged by AEA to accomplish this task.
- Providing guidance and advice on efforts to scan federal legislative and regulatory opportunities for relevant policies that could be targets for influence.
- Advising on the selection of specific policy opportunities.
- Advising on how best to manage the emerging network of contacts in both the legislative and executive branch that can assist AEA in influencing evaluation policy.
- Monitoring and providing input on the evaluation policy efforts and progress made.
- Reporting to the Executive Committee and Board on progress of the consultative campaign.
- Preparing and making recommendations to AEA Board on how the evaluation policy influence effort should be implemented subsequent to the two-year appointment period of the Task Force.

Because of the potentially broad nature of this work, the consultative campaign will initially be limited to only a few policy targets. The consultant, AEA Staff, and any Task Force members directly involved in implementation will be primarily responsible for identifying an initial set of potential policy targets and will bring these to the Task Force for consideration. The Task Force will be expected to play a major role in discussing and advising on what policies should be targeted in this effort. On the congressional/legislative side, this might involve: identifying one or two major areas of upcoming legislation (e.g., one each in education and health); identifying congressional staff members that are critical to the legislation; developing contacts and building a network of personal relationships with staff members; conducting background research on evaluation policies that currently exist in the area or other relevant aligned areas; identifying potential policies that would enhance evaluation; making contact with identified staff; and, responding as needed to any requests

for input on policy formulation (e.g., testifying at congressional hearings; drafting potential language for legislation). On the executive side, this effort might involve: identifying one or two major areas where evaluation policy is likely to be formulated within the next few years; determining who the major agencies and people are who will formulate the policies; determining how direct personal contact will be made with them; advising on the development of a network of personal relationships; conducting background research on existing relevant evaluation policies; and, consulting on potential policy changes or regulatory language. The Task Force will provide general guidance and advice on these activities; implementation will be handled by the consultant, AEA staff, and select Task Force members appointed to engage in these activities.

Public Presence Task. The primary purpose of this task is to support the consultative campaign through efforts that help position AEA as the leading U.S. association in evaluation and through the development of collateral materials that represent AEA and can be utilized in discussions with policymakers. Again, while the Task Force will be expected to advise and provide guidance on this effort, and may elect to draft materials that support it, the primary implementation will be the responsibility of the consultant, AEA staff, and Task Force members who are directly involved in implementation. The Task Force will be expected to provide advice and guidance on the following:

- Development of general “talking points” on evaluation and its role that would be likely to have broad consensus from AEA members and could be used in consultative work on evaluation policies (possible examples of talking points might be, for example, that: evaluation is a critically important endeavor; legislation and regulation should have explicit written requirements for evaluation of federally-funded programs; a broad range of evaluation methods are likely to be applicable and legislation and regulation should require that multiple methods and approaches appropriate to the program being assessed should be considered; etc.)
- Development and maintenance of a roster of leading evaluators who can be called upon as needed as spokespersons or contact persons for the media, for meetings with legislators or executives in government, for testifying in formal committees or meetings, and/or as consultants on evaluation policy issues.
- Development and enhancement of AEA collateral materials (e.g., brochures, Guiding Principles, public statements) that will support the consultative campaign
- Enhancement of the AEA website so that it is more useful for influencing evaluation policy formulation and positioning AEA as a leader in advising on evaluation policy. This should include working with the consultant to develop a “Policy Watch” function for notifying and engaging AEA members regarding important impending evaluation policy changes and initiatives.

The Task Force is not expected to address the AEA’s public presence generally. Their efforts should be concentrated on guidance that will specifically support the consultative campaign. In this work, the Task Force will communicate and collaborate with other AEA Committees such as the Public Affairs or Professional Development committees on areas of common relevance.

APPOINTMENT, TIME COMMITMENT, STRUCTURE AND SUPPORT

The Evaluation Policy Task Force is appointed for a two-year period beginning in July, 2007. The Evaluation Policy Task Force is a Task Force of the AEA Executive Committee and, as such, will report through the EC and may use the EC as a sounding Board as desired.

The Task Force will consist of 5 members, the Executive Director, and a liaison from the Public Affairs Committee (either the Chair or a PAC member). The Chair of the Task Force will be a member of the Executive Committee. The Evaluation Policy Consultant will be integrally involved in the Task Force and will participate in meetings and discussions as appropriate. The Task Force Chair and AEA Executive Director will identify potential members, subject to approval of the Executive Committee, with the following criteria to guide selection of members:

- Commitment to supporting the mission and goals of the American Evaluation Association.
- Knowledge of and a history of prior involvement with the American Evaluation Association.
- Familiarity with the field of evaluation and capacity to understand and represent the field to others.
- A broad perspective on evaluation and willingness to advocate for the many and diverse views of evaluation.
- Knowledge of the federal policymaking process.
- Experience with policy development initiatives in the Federal Government.
- Experience with public presentations of evaluation to a variety of audiences.
- Diversity of the Task Force and representativeness of the breadth of members and interests of AEA.

The Task Force will provide a written annual report to the AEA Board for review at its Winter Board meeting and updates as needed through the Task Force Chair at other Board meetings. At the completion of the first year of their work, the Task Force will be expected to report to the Board with their recommendations about the feasibility and desirability of continuation of this effort and the best mechanisms for doing so (e.g., continuation of the Task Force; establishment of a standing committee or assignment of this effort as a subcommittee to an existing committee). The AEA Board will provide ongoing feedback and guidance based on its reviews of these reports.

The Task Force will be expected to convene at least once every two months over the two year appointment period. Most of these meetings will be by teleconference, although the Task Force should attempt to meet face-to-face at least annually at the AEA annual conference. The Task Force will set its own agenda and specific activities in pursuit of the tasks outlined above. The Task Force is encouraged to develop work groups that invite people who are not on the Task Force to assist in work to address specific tasks as needed.

The Task Force will most likely need to develop a mechanism that enables rapid response to situations that arise under the consultative campaign. For instance, they might wish to develop a roster of senior evaluators who are willing to be called to help respond in the event that a meeting is needed to be held immediately, legislative or statutory language needs to be drafted quickly or public testimony is requested.

The AEA office will provide administrative support for the Task Force, including assistance in setting and managing meetings and Task Force materials.

AEA will provide contracted consultative support for technical assistance in identifying, contacting and managing specific policy shaping opportunities.

Over the course of the two year timeframe, the success of this initiative will be judged on: the level and appropriateness of the activities; the extent of its influence on specific policies and policy language; and, the quality and potential value of the network developed that connects AEA, the consultant, and policymakers.

MODIFICATIONS AS OF JULY 2009

The Evaluation Policy Task Force is appointed for an **additional** two-year period beginning in July, **2009**.

The Task Force will consist of **up to 10 members, including** the Executive Director, **a member of the Presidential rotation**, and a liaison from the Public Affairs Committee (either the Chair or a PAC member). **The Chair of the Task Force will be a member of the Executive Committee.**

EPTF TASK FORCE MEMBER PROFILES

Eleanor Chelimsky (2007 – Present): Eleanor has been an evaluator for many years, first at NATO, where she worked at the international level to perform defense studies and improve the data systems available to support evaluation; then at the MITRE Corporation, where she directed evaluations for executive branch agencies, and finally, at a legislative agency, the GAO, where she developed and ran the Program Evaluation and Methodology Division (PEMD), an evaluation unit of about one hundred social scientists, doing evaluations for the Congress. Under her leadership, PEMD pioneered the use of new methodologies and performed studies in almost every subject area. She also worked to improve GAO's general capabilities in evaluation and helped a number of different countries establish their own evaluation policies and institutions. Eleanor is the lead author on numerous reports and is the lead editor on *Evaluation for the 21st Century: A Handbook* (SAGE, 1997, with Shadish). She is a former President of the Evaluation Research Society and of the American Evaluation Association, and received GAO's top honor, the Comptroller General's Award.

Leslie Cooksy (2007 – 2010): 2010 AEA President Leslie Cooksy currently serves as Associate Professor in the School of Education at the University of Delaware. She is jointly appointed in the Delaware Education Research and Development Center. From 2005-2007 she was under contract to the California Department of Health Services, where she managed the Tobacco Control Evaluation Center that provided evaluation technical assistance and training activities for 100 county and local public health projects. From 2002-2005, Leslie was the Director of the Center for Community Research and Service at the University of Delaware School of Urban Affairs and Public Policy, responsible for leading, planning, and managing their portfolio of work. Earlier in her career, as a social science analyst at the US GAO, she conducted studies of evaluation policy, including the evaluation set-aside of the Public Health Service. A founding member of the EPTF, Leslie brought with her experience with the AEA Public Affairs Committee, which had started some policy-related work. She stepped down in 2010 to focus her efforts on her duties as AEA's 2010 President.

Katherine Dawes (2007 – Present): Katherine has served as the Director of the Evaluation Support Division in the United States Environmental Protection Agency (EPA) since 2000. This division conducts – and enables partners throughout EPA to more effectively conduct – program evaluations and analyses that inform management decisions, enhance organizational learning, promote innovation and foster environmental results. As the Division Director, Katherine is particularly engaged in networking with all levels of government, academia and non-governmental organizations to promote new developments and new information regarding innovation analysis and environmental evaluation. Katherine is an active member of the American Evaluation Association and its Environmental Program Evaluation Topical Interest Group. She is also a founding member and former Trustee of the Environmental Leadership Program, a non-profit, non-partisan organization designed to train and support the next generation of environmental leaders.

Patrick Grasso (2007 – Present): Patrick G. Grasso is the current Chair of the American Evaluation Association's Evaluation Policy Task Force. He is an evaluation and management consultant working primarily with international development organizations, and served as Advisor to the Director of the

Independent Evaluation Group (IEG) of the World Bank from 2004 until 2008. Prior to joining the World Bank, he was Director of Evaluation and Learning Resources at the Pew Charitable Trusts. Earlier, he spent eleven years at the U.S. General Accounting Office as an Assistant Director in the Program Evaluation and Methodology Division, where he led a team responsible for evaluating community development programs and public management initiatives. He holds a Ph.D. in Political Science from the University of Wisconsin, and is a member of the editorial board of the *American Journal of Evaluation*.

Jennifer Greene (2010 – Present): With the rotation off of the EPTF in 2010 of Leslie Cooksy, 2011 AEA President Jennifer Greene joined the EPTF in fulfillment of the requirement that a member of AEA’s Presidential rotation (either the President-elect, sitting President, or Past-president) serve on the task force. Jennifer has been an evaluation scholar-practitioner for 35 years. Within AEA, she served with Gary Henry as co-editor-in-chief of *New Directions for Evaluation*, and she has been on the AEA board and active in multiple task forces related to AEA’s diversity and multicultural initiatives. As a professor of educational psychology at the University of Illinois at Urbana-Champaign, Jennifer’s research and practice interests emphasize the advancement of responsiveness and democratic values through evaluation and the thoughtful use of qualitative and mixed methods approaches to evaluation. Jennifer was the 2003 recipient of AEA’s Lazarsfeld Award for Contributions to Evaluation Theory, the co-editor of the 2006 *Sage Handbook on Evaluation* (with Ian Shaw and Mel Mark), and the author of a 2007 book on *Mixed Methods in Social Inquiry*.

George Grob (2007 – Present): George is the paid consultant to AEA working with the EPTF. He is the President of the Center for Public Program Evaluation, an evaluation and policy development firm focused on helping public programs work efficiently and effectively. Prior to establishing this consultancy, he was the Executive Director of the Citizens’ Health Care Working Group. This commission was charged with listening to citizens across the country to hear what they think about our nation’s health care system. Based on their concerns, the Working Group sent recommendations to the President and the Congress. In 1988 he joined the Office of Inspector General where he held the top positions in the Office of Evaluation and Inspections, becoming its Director in 1994. He oversaw the production of more than 1,000 studies in a 15-year period. George has testified more than two dozen times before Congress and assisted other public officials in doing so as well. Over the period of 2007-2011, George served as the primary author and public outreach liaison for the EPTF.

Susan Kistler (2007 – Present): Susan is the Executive Director of the American Evaluation Association. Having worked with AEA for the past 13 years, she is positioned to provide the task force with assistance regarding planning, logistics, member input, and information dissemination within the field.

Melvin Mark (2007 – Present): Mel is Professor and Head of Psychology at the Pennsylvania State University. A former President of the American Evaluation Association and former editor of the *American Journal of Evaluation*, he holds a Doctorate from Northwestern. Mel is the co-author or co-editor of key volumes in the field including *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs* (Jossey Bass, 2000, with Henry and Julnes), *What counts as credible evidence in applied research and evaluation practice?* (SAGE, 2008, with Donaldson and

Christie), and *Social psychology and evaluation* (Guilford, 2011, with Donaldson and Campbell). Mel's interest in evaluation policy stems from his broader interest in the both the precursors and consequences of evaluation. He is a co-editor and contributor to the New Direction for Evaluation issue on *Evaluation policy and evaluation practice* (Jossey-Bass, 2009, with Trochim and Cooksy). A suggestion by Mel at an AEA panel helped stimulate the creation of the AEA Topical Interest Group on Evaluation Policy.

Stephanie Shipman (2007 – Present): Stephanie is an Assistant Director in the Center for Evaluation Methods and Issues, Applied Research and Methods Team, in the U.S. Government Accountability Office (GAO). She directs studies of federal agencies' performance measurement and program evaluation activities, and methods for solving analytic challenges in program assessment, as part of the Center's mission to further program evaluation in the federal government. She also consults with federal agencies and foreign governments on program evaluation methods and policies. She previously conducted evaluations of federal programs supporting children and families in GAO's Program Evaluation and Methodology Division. In 2008, she received AEA's Alva and Gunnar Myrdal Government Award. She is also a founding member of and coordinator for the Federal Evaluators group, an informal network of several hundred evaluation officials from agencies across the federal government.

William Trochim (2007 – 2011): Bill was President-elect of AEA in 2007 and President in 2008 during which time he shepherded the concept of a policy-focused working group within AEA to its fruition as the EPTF. He served as the EPTF's first chair from 2007-2009, and remained on the task force through early 2011 as an active member. Bill is the Director of Evaluation for the Weill Cornell Clinical and Translational Science Center and participates actively in the NIH Clinical and Translational Science Awards national evaluation. He is developing evaluation systems for the HIV/AIDS Clinical Trials Networks of the Division of AIDS (DAIDS) of the National Institute of Allergies and Infectious Diseases (NIAID). He is actively engaged in research with the National Science Foundation incorporating systems approaches in the evaluation of Science, Technology, Engineering and Mathematics (STEM) education programs. And, as Director of Evaluation for Extension and Outreach at Cornell, he works with extension programs throughout New York State on planning and implementing evaluation systems for their outreach and education programs.

PERFORMANCE SUMMARY - AEA EVALUATION POLICY TASK FORCE JULY 2007 TO NOVEMBER 2011

I. INTRODUCTION

At its Winter 2007 meeting, the Board of Directors of the American Evaluation Association (AEA) discussed its interest in the Association enhancing its ability to identify and influence evaluation policies that have a broad effect on evaluation practice. To that end the Board established an Evaluation Policy Task Force to advise AEA on how best to proceed in this arena.

In advance of the Task Force, an Advisory Group was convened over the Spring of 2007 to address the following tasks:

- Discuss the boundaries and scope of the Task Force
- Develop a draft charge for the Task Force
- Identify and recommend potential Task Force members

THE CHARGE

The results of the Advisory Group's discussions informed the Board's [Charge to the Evaluation Policy Task Force](#). The Board's Charge established the Evaluation Policy Task Force (EPTF) and outlined its scope of work to:

- Focus on evaluation policies, including
 - Evaluation Definition
 - Evaluation Methods
 - Human Resources Regarding Evaluation
 - Evaluation Budgets
 - Evaluation Implementation, and
 - Evaluation Ethics
- Focus on U.S. Federal Evaluation Policy
- Target Only a Few Specific Legislative and Executive Policies
- Focus on Selected Substantive Areas, and
- Be Appropriately Opportunistic

THE TWOFOLD MISSION

It further tasked the EPTF to advise and assist the AEA Board, Executive Committee, and AEA staff in addressing two primary parallel tasks:

- **A consultative campaign** designed to identify and provide useful consultation to U.S. federal legislators and executive branch staff on legislation and regulation in a select number of areas, and
- **A public presence initiative** that AEA projects in direct support of the consultative campaign, in particular through the development of collateral materials that can be utilized in that campaign, and perhaps through other forms of outreach as well (e.g., website presence).

APPOINTMENT, TIME COMMITMENT, STRUCTURE AND SUPPORT

The Evaluation Policy Task Force was appointed for a two-year period beginning in July, 2007, as a task force of the AEA Executive Committee, reporting through the Executive Committee and using it as a sounding Board as desired.

The Task Force originally consisted of 5 members, the Executive Director, and a liaison from the Public Affairs Committee. This was later amended to provide for up to 10 members, including the Executive Director, a member of the Presidential Rotation, and a liaison to the Public Affairs Committee. All members of the Task Force have provided their services pro bono, with the exception of George Grob, a paid consultant working with the EPTF, and Susan Kistler, AEA's Executive Director.

The Charge also authorized hiring an Evaluation Policy Consultant to be integrally involved in the Task Force and participate in meetings and discussions as appropriate.

INITIAL ASSESSMENT

Over the course of the two year timeframe, the success of this initiative was to be judged on:

- the level and appropriateness of the activities
- the extent of its influence on specific policies and policy language, and
- the quality and potential value of the network developed that connects AEA, the consultant, and policymakers.

In accordance with this requirement, The EPTF provided its [Assessment and Recommendations](#) to the Board for its July 2009 meeting.

In response, the Board decided to

- continue the EPTF for two additional years, to be followed by an evaluation to be performed under the Board's direction

- increase to 10 the limit on the number of members of the EPTF
- change the requirement that the Chair be a member of the Executive Committee to a requirement that at least one member of the EPTF be a member of the Presidential rotation
- requested the EPTF to develop processes for vetting documents and statements of the EPTF, and
- authorize the Executive Director and the EPTF to routinely advocate for:
 - Broad use of evaluation in public programs, especially those of the Federal Government
 - Use of multi-method approaches
 - Tailoring evaluation across the life-span of a program
 - Adequate funding for evaluation, and
 - Use of qualified, experienced evaluators as appropriate.

The Board also [amended the EPTF's charge](#) to reflect the structural aspects of these decisions.

CURRENT ASSESSMENT

In accordance with the above decisions, the Board is conducting its evaluation of the EPTF starting in the fall of 2011. This paper has been assembled by the Evaluation Policy Consultant in consultation with the EPTF as an initial resource for the Board in conducting its review.

ORGANIZATION OF THIS DOCUMENT

In addition to the material provided above about the origins and objectives of the EPTF, this document provides information about the EPTF's activities and impacts since its inception in June 2007. It is organized around the two prongs of the EPTF's mission:

- 1) the consultative campaign (advocacy), and
- 2) the public presence initiative (branding).

This paper contains links to numerous documents relevant to a deeper understanding of the text. It also contains links to a set of key references and resources materials that contain a detailed exposition of all the activities of the EPTF since its beginning. These key resources include, for example, all the quarterly reports of the Evaluation Policy Consultant, all the "Policy Watch" columns published in the AEA monthly newsletter, and a log of all the Evaluation Policy Consultant's external meetings.

SUGGESTION TO THE READER

A key resource is a [2011 PowerPoint summary of EPTF activities and accomplishments](#) over its lifespan. It may be useful to begin with this PowerPoint summary in order to put all the activities discussed in the next two sections into perspective. It can be found at [Appendix B](#).

Another key resource is AEA's [Evaluation Roadmap for a More Effective Government](#). It is mentioned frequently throughout this document. It would probably be beneficial to read it before reading the rest of the material gathered here. It can be found at [Appendix C](#).

II. CONSULTATIVE CAMPAIGN (ADVOCACY)

Summary

We believe the EPTF consultative campaign has had significant influence in shaping the landscape for federal evaluation, especially considering how recently it was launched. The EPTF has worked with a broad spectrum of agencies and legislative initiatives on both the executive branch including

- Office of Management and Budget
- USAID
- Department of Health and Human Services
- House Commerce and Ways and Means Committees, and
- Senate Finance and Health, Education, Labor, and Pensions.

Through its direct consultation and documents like its highly regarded Roadmap, the EPTF has directly or indirectly influenced agency evaluation policies, transformed attitudes and messages about evaluation, shaped the discussion around and language of major federal executive directives and legislation. This includes:

- The Obama Administration's guidance on impact evaluation
- FYs 2011 and 2012 budgets for evaluation
- Appropriations language supporting evaluation of international HIV/AIDS programs
- Oversight of health care insurance reforms
- GAO's evaluation of international feeding programs
- USAID's evaluation policy, and
- State Department's evaluation policy

In addition the EPTF prepared and provided formal comments on:

- Simplification of the Paperwork Reduction Act procedures
- HHS home visiting programs
- GAO Auditing Standards
- Evaluator independence

After initial organizational meetings of the EPTF in the fall of 2007, the consultative campaign began in earnest in February 2008 with contacting senior policy officials of the Bush Administration and then picked up speed and influence through contacts with senior officials of Obama Administration appointees shortly after their arrival in February 2009. These activities and other events contributed in various ways and with varying levels of influence to a stream of evaluation policies being issued by the Obama Administration and the Congress over the next three years.

In this section we will chronicle the EPTF's evaluation policy advocacy efforts and the issuance of new evaluation policies, especially within the Obama Administration, certain congressional committees, and the Government Accountability Office. In addition we will try to explain the connection between these new policies and the EPTF's consultative campaign. In order to facilitate that discussion, it is useful to first discuss the variety of ways that policy advocacy can have an impact on policy. Section A provides an introduction to the various modes of policy advocacy along with a list of associated EPTF advocacy activities. Section B provides a fuller description of the EPTF activities themselves.

A. RELATIONSHIP OF POLICY ADVOCACY TO POLICY MAKING IN GENERAL AND WITH RESPECT TO SPECIFIC EPTF AND AEA POLICY ADVOCACY INITIATIVES

The use of policy advocacy by policymakers follows the same patterns that are generally recognized as characterizing the use of evaluation findings themselves, which are, in fact, often policy recommendations. There is, for example, "instrumental use", in which influence is direct, clear and sometimes acknowledged. Then there is "enlightenment use" in which the path to policymaking is usually obscure, diffused among other types of advocacies (including lobbying and various political pressures), and rarely, if ever, acknowledged. Yet even so, the use is there, and can sometimes be traced, via records and/or interviews.

Policy advocates try to affect government policies in a number of different ways, under various circumstances, with mixed results, and with varying degrees of certainty about the connection between the advocacy and the enacted policies. The following framework may be useful for thinking about policy advocacy efforts and their results.

General Influencing Mechanisms. Policy advocates often take advantage of opportunities to provide official comments on emerging policies. Sometimes, these official comments are solicited by the policy makers (such as requests for comments on proposed regulations) but sometimes may be offered to policy makers by way of unsolicited letters, public statements, and the like. This is the case, for example with AEA's:

- Response to the Office of Management and Budget's (OMB) formal requests for comments on the Paperwork Reduction Act
- Comments to the Department of Health and Human Services' (HHS) on its Home Visiting Program
- Comments to the Government Accountability Office's (GAO) on its Auditing Standards (Yellow Book)
- Letter to the editor of the New York Times of evaluator independence,
- Tables and background information on evaluation requirements in the PEPFAR Reauthorization Act provided to senior representatives of the Office of the USAID Commissioner and USAID,

- The letter, signed by AEA's Presidential Rotation to Peter Orszag, the Obama Administration's Director of OMB on evaluation policy in general, and containing AEA's Evaluation Roadmap for a More Effective Government, and
- Meetings on two separate occasions with congressional staff working on the GPRA Reauthorization Act, and
- Meetings with Senate staff working on two evaluation bills now under development.

It is not generally possible to claim a direct link between such comments and policies ultimately made. Instead, the comments are part of a broader effort, complementing the input of many other organizations. Sometimes policy success comes from the consonance with that of other commenters, or if the policy makers, later in the policy process, provide a detailed accounting of all the comments received and that reveals a connection to the comments to the final policy adopted. While the effect of such general influencing mechanisms cannot always be directly ascertained, it is still quite important to establish the advocate's positions for the consideration of policy makers. These positions tend to accumulate and may gradually become the conventional wisdom in a policy arena. It is also certain that if such advice is not provided there will not be an attributable policy effect.

Furthermore, in some cases, like AEA's *Evaluation Roadmap for a More Effective Government*, a document becomes widely circulated and highly influential, as evidenced by the [Roadmap Citation and Use Record](#). In fact, the Roadmap was probably the most influential document prepared by the EPTF as it played an important role in both the consultative campaign and the public presence initiative. It was developed early in the life of the EPTF as an explanation of evaluation policy and as an advocacy document that was presented to Peter Orszag, the incoming Director of the Office of Management and Budget. (See discussion on page 20.) In the consultative campaign it was provided to policy makers both to explain and promote the development of evaluation policies across the Federal Government or in specific areas such as health care reform and international development. At the same time, it served as a standing document which could be easily accessed and used by anyone interested in evaluation policy either in general or for specific purposes, even if AEA or the EPTF was not actively promoting a particular policy. The Roadmap is referred to throughout this document,

The Roadmap is particularly noteworthy because the way in which it was produced. As noted above, it arose directly from the need for a compelling case for evaluation as an inherent feature of good government at the time that officials of the Obama Administration first arrived in Washington. EPTF members were keenly aware of the critical importance of making a good first impression during these early days. Typically, the major management themes of a new administration are developed and announced at this time. The Roadmap reflects the intense participation by all EPTF members through numerous conference calls, emails, and drafts, all carried out initially against an extremely tight deadline. Then, after the initial draft was sent to the incoming Director of OMB, it was, at the request of the Board and in keeping with the desire of the EPTF, re-written after consultation first with the AEA Board, then with the AEA membership at large, and then again with the Board. It contains both broad and detailed evaluation policy about the purpose, need for, nature of, and administrative structures and resources for evaluation within the Federal Government. Remarkably, despite its scope and detail, this was accomplished within the restriction that it could only contain consensus statements that the EPTF and Board agreed to and were broadly accepted by the AEA membership.

Policy Involvement. Sometimes policy advocates are invited to play an insider's role in policy making. This was the case, for example, in

- OMB's Associate Director for Administration and Government Performance staff inviting the AEA to submit comments on OMB's PART Guidance in time to meet a publication deadline
- OMB staff inviting the EPTF to provide a short paper on evaluation policy to share with incoming staff of the Obama Administration at a meeting kicking off the Administration's internal examination of evaluation policy
- The invitation by the House Appropriations Committee for legislative language to be included in the House Senate Conference Committee's appropriations bill, and
- The invitation from the House Foreign Affairs Committee to provide detailed principles and legislative language for an evaluation policy for USAID

For these EPTF activities, a good case can be made that the subsequent publication of policies was linked to its advocacy efforts.

The overall tone reflected what might be called mixed success. On the one hand, attention was paid to the Roadmap's emphases on taking evaluation seriously, linking it to decision-making, opening it up to the public, and using it in context. Also, to measure program performance and ascertain which programs work, the Administration planned to lean on evaluations that were professionally done and wide open to public view. On the other hand, there was no reflection of the Roadmap's advocacy for mixed methods, the emphasis being instead on the experimental design as the preferred method. However, OMB's Memorandum to the Heads of Executive Departments and Agencies did remedy this problem (see below).

Of course, such policies are also likely to have been affected by the input of other policy advocates as well, and seldom appear word for word like the letters submitted to the policy makers.

Policy Instrumentality. In some cases the interaction of the advocate and the policy makers e.g., congressional staff, is highly interactive, involving the exchange and mutual refinement of a series of draft papers and legislative language. In these cases, draft and final legislation reflecting the policies exchanged cannot be doubted, even if the language is not exactly that of advocates but is clearly in keeping with policy exchanges. This is the case with

- the PEPFAR Appropriations bill
- the House Foreign Assistance Act Reauthorization bill
- Office of Inspector General (OIG) oversight of the health insurance reforms in Title I of the Affordable Care Act (health care reform act), and
- NIH's proposed evaluation policy for clinical and translational science awards, now under development with EPTF guidance being fed in on an ongoing basis

In the case of the of the OIG oversight of the health insurance reforms, that provision simply would not have happened had it not been for EPTF advocacy, including evening telephone exchanges with key congressional staff in the days running up to the consolidation of the House and Senate versions of the

bill. This is a particularly noteworthy example of the AEA having influence on evaluation policy through the work of the EPTF.

Policy Attribution. In some case, the influence of policy advocacy is explicitly acknowledged by the policy makers, and one can therefore be relatively confident in inferring its impact. This is the case for:

- USAID's evaluation policy, for which the key policy official responsible for drafting the policy acknowledged in a public presentation the importance of the *Evaluation Roadmap* in developing the agency's policy, and
- the statement of a senior State department official in a private meeting of the importance and efficacy of the Lundy/EPTF ad hoc's groups efforts in the pressing the Congress for evaluation in incentivizing the State department to issue its own policies.

Policy Effects. In some cases, an advocate's policies are implicitly or explicitly used by policy makers or advisors in carrying out government functions. This is the case for:

- the subsequent authorization of 25 HHS Office of Inspector General (OIG) evaluators to carry out the authorized evaluation work, and
- GAO's explicit references to the use of the Evaluation Roadmap as a standard for its evaluations of international feeding programs.

B. CHRONOLOGY OF THE EPTF'S POLICY ADVOCACY ACTIVITIES WITH REFLECTIONS ON THEIR RESULTS

The following describes in greater detail the advocacy activities of the EPTF alluded to in Section A.

Contacts with Bush Administration Officials

In February 2008, then EPTF Chair and AEA President William Trochim along with Evaluation Policy Consultant George Grob met with Robert Shea, Associate Director for Administration and Government Performance, to share insights about evaluation Policy. Shea's office was responsible for the Bush Administration's *Program Assessment Rating Tool* (PART) which required periodic assessments of performance of all Federal programs. These assessments, performed by OMB staff, centered around performance measures and the body of evaluative information available for each program. The PART assessments represented a significant advance in attention to the formal evaluation of federal programs but they were controversial in several respects. The criticisms centered on what many believed was an over reliance on performance measures without attention to richer and more nuanced evaluation reports; an emphasis on randomized control trials as the only acceptable form of impact evaluation; a suspicion (warranted or not) that the assessments were being developed primarily to use as evidence to support budget reductions for domestic programs; a significant paperwork burden to produce the recurring reports underlying the assessments; and the lack of interest by Congress, which largely was left out of the process.

Shea was receptive to AEA's overtures and immediately requested written advice from AEA on the appropriateness of randomized control trials and other forms of impact analysis. He asked for such advice within a week of the meeting to be used in developing a new PART Guidance that his office was planning to issue within the next few weeks. The result was the EPTF's first formal position paper on evaluation policy sent to senior Federal officials. In both the [Cover Letter](#), and more importantly in the document itself, [Comments on What Constitutes Strong Evidence of a Program's Effectiveness?](#), the EPTF advocated a more balanced use of methods in evaluating program impact. In particular, the document:

- provided a more balanced assessment of RCTs
- emphasized the need for mixed methods to assess program value
- called for realistic assessment of feasibility and resources for impact studies, and
- expressed the need to address equity and human research subject concerns.

The EPTF letter gradually gained attention and was the first formal document that established AEA as an authoritative source on matters of evaluation. For example, it was cited by the Government Accountability Office's analysis of the relative merits and appropriateness of various types of impact evaluation, summarized in its November 2009 report, [A Variety of Rigorous Methods Can Help Identify Effective Interventions](#) (external link).

Soon afterward, Shea invited Trochim to make one of the first presentations to the newly established Evaluation Committee of the Council of Performance Improvement Officers later established by the Bush Administration.

These initial meetings also established ongoing connections between AEA and the Office of Management and Budget senior career staff members that laid the groundwork for more significant policy interactions with the Obama Administration.

Contacts with Obama Administration Officials and Emergence of New Evaluation Policies

The EPTF lost no time making contact with the incoming Obama Administration. With advance preparation during the time period after the President Obama's election and his taking office in January 2009, the EPTF prepared to reach out to incoming senior policy officials of the new administration. On February 3, 2009, the AEA Presidential Rotation [sent a letter](#) with an early version of AEA's [Evaluation Roadmap for a More Effective Government](#) (external link to archival version of *Roadmap*) prepared by the EPTF to Peter Orszag, the incoming Director of Management and Budget, within days of his taking office.

Then, on February 9, at the invitation of a senior careerist at OMB, the EPTF sent a one page summary of evaluation policies for consideration at a meeting of senior careerists and an incoming Obama appointee to shape evaluation policy for the new Administration. That paper, [Program Evaluation for Management and Results](#), was made available to the attendees of that meeting.

From that point on a series of important evaluation policies were issued by the Obama Administration that reflected a much broader appreciation of the importance of evaluation as a key feature of good

government. The evaluation policies issued by the Bush Administrations had focused on performance indicators and randomized control trials for impact analysis. However, the Obama policies reflected Roadmap principles. While they initially focused on impact evaluation, they gradually expanded outward to include periodic evaluations to improve programs and to analyze trend data collected through performance indicators. Furthermore they incorporated many concepts embedded in AEA's *Evaluation Roadmap*, such as linking evaluation to government decision making, establishing formal evaluation organizational structures to institutionalize and gradually improve evaluation, developing the capacities of evaluators, using a variety of methods, safeguarding evaluator independence, and sharing evaluation results with the public.

2010 Budget. First out of the gate was President Obama's 2010 budget. It proposed to make better use of evaluation in education programs, increasing funding for "rigorous evaluation" as a first step to doubling the Department of Education's support for education research. More broadly, the Administration wanted to "fundamentally reconfigure the Program Assessment Rating Tool." They aimed to "open up the insular performance measurement process to the public, the Congress and outside experts" and "eliminate ideological performance goals and replace them with goals Americans care about and that are based on congressional intent and feedback from the people served by Government programs. Programs will not be measured in isolation, but assessed in the context of other programs that are serving the same population or meeting the same goals."

The overall tone reflected the Roadmap's emphases on taking evaluation seriously, linking it to decision making, opening it up to the public, and using it in context. To measure program performance and ascertain which programs work, the Administration planned to lean on evaluations that were professionally done and wide open to public view.

Science and Technology Priorities. [An August 4, 2009, memorandum](#) (external link) for the heads of Executive Departments and Agencies issued jointly by Peter Orszag, Director of the Office of Management and Budget (OMB) and John Holdren, Director, Office of Science and Technology singled out evaluation as an essential ingredient in meeting the President's science and technology priorities. It put evaluation center stage in decision making. "In their budget submissions, agencies should describe the expected outcomes from their research in relation to these four practical challenges and cross-cutting areas, providing quantitative metrics where possible, and describe how they plan to evaluate the success of various techniques to increase support for high-risk research.

"Budget submissions should also describe how agencies are strengthening their capacity to rigorously evaluate their programs to determine what has been demonstrated to work and what has not."

The guidance saw evaluation as a tool to enable "... agencies to eliminate or reduce funding for less-effective, lower-quality, or lower-priority programs ..." as well as "... target investments toward high-performing programs."

Increased Emphasis on Program Evaluation. Perhaps the most sweeping of the initial policy statements was OMB Director Peter Orszag's October 7, 2009, memorandum for the heads of executive departments and agencies on [Increased Emphasis on Program Evaluation](#) (external link). It focused

initially on "social, educational, economic, and similar programs whose expenditures are aimed at improving life outcomes (such as improving health or increasing productivity) for individuals."

Unlike earlier pronouncements about evaluation in previous Administrations which focused on randomized control trials for impact evaluations, this memorandum reflected broader Roadmap themes such as the emphasis that "evaluations can help policymakers and agency managers strengthen the design and operations of programs," and its intention to improve the institutional capacity of Federal agencies to perform evaluations, ask for an assessment of agency capacity to conduct rigorous evaluation, promote independent evaluations, and attract and retain talented researchers in an office with standing within the agency. To assist in building this capacity, it reconstituted the Inter-agency Evaluation Working Group under the Performance Improvement Council.

Also consistent with EPTF comments to OMB on PART, it avoided reference to a "gold standard" method. Instead it established that a key goal of the Working Group would be to "help agencies determine the most rigorous study designs appropriate for different programs given their size, stage of development, and other factors."

2011 Budget. President Obama's 2011 budget turned evaluation policy into a very practical matter, clearly authorizing and promoting evaluation as an integral part of decision making and putting dollars into it.

Far and away, the most impressive evaluation message is a set of overarching policies found in an annex to the budget called *Analytic Perspectives*. The parts that are relevant to evaluation are the first two sections (labeled 7 and 8) in a chapter called "Performance and Management."

Section 7, Delivering High-Performance Government, provides a new look at performance measurement and substitutes a new performance management system that builds on but also replaces the PART system used by the previous Administration.

[Section 8, Program Evaluation](#) (external link), provides detailed steps and explanations to implement the policies enunciated by OMB Director Peter Orszag in his October 7 memorandum and discussed above. The introductory paragraph puts evaluation in the context of performance management. "Performance measurement is a critical tool managers use to improve performance, but often cannot conclusively answer questions about how outcomes would differ in the absence of a program or if a program had been administered in a different way. That is where program evaluations play a critical role."

The second paragraph is equally telling. "A central pillar of good government is a culture where answering . . . questions [about how well programs work] is a fundamental part of program design and where agencies have the capacity to use evidence to invest more in what works and less in what does not."

The budget allocates approximately \$100 million to 17 agencies to conduct new evaluations with strong study designs that address important, actionable questions or to strengthen agency capacity to support such strong evaluations.

2012 Budget. The President's 2012 budget reinforces his commitment to evaluation, including proposals for a selection of strong evaluations in the Departments of Education, Health and Human Services, and Treasury; and at the Millennium Challenge Corporation, the United States Agency for International Development, the National Aeronautics and Space Administration, and the Office of Personnel Management.

While the Obama Administration certainly would have strengthened evaluation policy in any case, it appears that the content and direction of that policy was influenced by the EPTF's work. These policies seem to reflect principles embedded in the Roadmap, especially a more flexible approach to methods. In some sense, presence of the EPTF gave AEA a place at the metaphorical policy-making table.

Evaluation Policies for International Programs

PEPFAR (International HIV/AIDS Relief). In March 2008, Victor Dukay of the Lundy Foundation contacted AEA member Jody Fitzpatrick requesting help in convincing Congress to include evaluation funding in the reauthorization of PEPFAR. Jody relayed the request to the Evaluation Policy Task Force (EPTF) Chair, and the EPTF went to work through its Evaluation Policy Consultant, George Grob. George, along with Patrick Grasso, Victor Dukay of the of the Lundy Foundation, Matt Oresman of Patton Boggs, and Michelle Orza, project leader of the Institute of Medicine's evaluation of the first five years of the PEPFAR program, formed an ad hoc work group to promote evaluation of the PEPFAR program.

The work group prepared [proposed appropriations language](#) (external link) to promote funding for evaluation of the PEPFAR program. The Appropriations Conference Committee endorsed the idea, including similar language in its [2009 appropriations bill \(as summarized on page 2 of its appropriations report\)](#) (external link). The exact language is: “*Evaluation of global HIV programs.—* USAID and OGAC are expected to increase funding over time for operations research, impact evaluation research, and program monitoring to ensure that interventions and approaches to service delivery are evidence-based and continuously improved over time.”

Once this official guidance became law, the work group prepared [materials about statutory requirements for PEPFAR evaluation](#) (external link), and briefed numerous staff of the U.S. HIV/AIDS Coordinator and USAID and other State department officials about the requirements. Today, evaluation is built into the management of the PEPFAR program. Much of this would have happened without the work group's efforts, but their work did intensify interest and fortify allocation of evaluation funds for evaluation.

House Foreign Assistance Reauthorization Bill. Shortly after gaining appropriations language for PEPFAR and promoting implementation of evaluation requirements in the newly enacted PEPFAR reauthorization, staff representing Representative Howard Berman, Chair of the House Committee on

Foreign Affairs, approached the Lundy/EPTF ad hoc work group for advice in developing comprehensive evaluation requirements for U.S. international development programs. This language was to be inserted into a foreign assistance reauthorization bill that the committee was drafting.

In response, the Lundy/EPTF ad hoc working group sent the committee a [proposed foreign assistance evaluation policy](#) and [an executive summary](#) (external links) of it. These documents provided a set of legislative specification tailored to foreign assistance.

As a direct result of the input provided by the Lundy/EPTF ad hoc group, the version of the "[Foreign Relations Authorization Act, Fiscal Years 2010 and 2011](#)" (HR 2139) (external link) passed by the committee included evaluation policies very similar to those that the Lundy/EPTF working group had sent them (See page 359 of the bill.) Some of the key evaluation provisions, which are found in section 3 of the bill, are:

- A requirement for the President to develop and implement a rigorous system to monitor and evaluate the effectiveness and efficiency of United States foreign assistance
- The establishment of measurable performance goals
- Criteria for selection of programs to be subject to various evaluation methodologies
- Establishment of an evaluation organization unit in each Federal agency involved in foreign assistance activities
- Requirements to apply the lessons learned and results from evaluation activities in the planning and implementation of foreign assistance programs
- Requirements to publish all evaluation plans and reports
- Requirements for annual evaluation plans
- Consultations among Federal agencies, governments of host countries, international and indigenous nongovernmental organizations, and other relevant stakeholders
- Capacity building for evaluation in Federal agencies and for recipient countries
- Annual budgeting for evaluation
- Establishment of a Foreign Assistance Advisory Council with biennial reports of its activities to the President and the Congress
- Annual reports from the President to the Congress on the use of evaluation
- Definitions of key evaluation terms
- A 5% set aside of foreign assistance funds to pay for evaluations

Senate Foreign Assistance Reauthorization Bill. The Senate soon followed suit with its own reauthorization bill, "[Foreign Assistance Revitalization and Accountability Act of 2009](#)" (S. 1524) (external link). It filled a gap in the House bill by establishing two new entities to coordinate evaluation of foreign assistance programs. In section 5, it would establish an Office for Learning, Evaluation, and Analysis in Development that will develop, design, coordinate, guide, and conduct the complete range of evaluation activities relating primarily to the United States Agency for International Development (USAID). In section 6, it established a Council on Research and Evaluation of Foreign Assistance to conduct evaluations of the impact of foreign assistance programs carried out by any Federal agency or international and multilateral assistance programs receiving financial assistance from the United States. However, it contained none of the overarching evaluation provisions included in the House bill.

Neither the House nor the Senate Reauthorization bills were enacted into law, but they did set the stage for evaluation policies to be issued by the Executive Branch--USAID and the Department of State-- as discussed next.

USAID. In January, 2011, USAID issued [its evaluation policy](#) (external link). It calls for:

- Integration of evaluation and program planning in international development programs
- Requirements for evaluations of major programs and untested interventions
- Acknowledgement of need for both quantitative and qualitative methods
- 3% set aside of major program funds for evaluation.

During a presentation unveiling the policy at Georgetown University, at which both George Grob and Patrick Grasso were present, the USAID official who led the development of the USAID policy acknowledged the influence of [AEA's Evaluation Roadmap](#).

The State Department. Equally impressive is the [State Department's evaluation policy](#) (external link) of both development and diplomacy programs published in May. Among the highlights are:

- An evaluation framework covering all programs, projects, and activities in bureaus and missions
- Evaluation requirements for major program areas
- Requirements for evaluation plans
- Emphasis on evaluator independence

The USAID policy directly and the State Department policy indirectly followed intensive efforts by the EPTF over a three year period involving funding authorities for international HIV/AIDS programs, reauthorization language in House and Senate bills, and meetings with State Department officials. In a meeting with Vic Dukay a senior State Department official told him that it was the earlier work of Lundy/EPTF ad hoc work group described above in pressing for evaluation policy with the Congress that led to the Department's own evaluation policy.

Health Care Reform

The EPTF had identified health care reform legislation as a target for its initiative to ensure that appropriate evaluation policies are embedded in major Federal program legislation. Therefore, as soon as the President and the Congress began to work on this legislation, the EPTF began to follow it closely. Five separate congressional committees were drafting bills which, over a year long period, were ultimately combined into two, the *Patient Protection and Affordable Care Act* and the *Health Care and Education Reconciliation Act of 2010*. They were signed by President Obama on March 23, 2010.

The EPTF's Evaluation Policy Consultant, George Grob, followed each bill and prepared numerous emails and formal correspondence to Members of these committees and their staffs. For example, EPTF Chair William Trochim sent comments on emerging legislation to members of the [Senate Finance Committee and House Committees](#) with comments on the leading bills. Grob spoke with leading staff members of both political parties and sent them emails at various times on particular topics,

eventually focusing on the most crucial issue--the lack of any oversight mechanisms for the health insurance reforms (Title I of the *Patient Protection and Affordable Care Act*). The emails are too numerous to include here but are available upon request. Here is a sampling of them.

- [To Chuck Clapton \(Senate\) 6/22/2009](#) (external link)
- [To Chuck Clapton \(Senate\) 10/6/2009](#) (external link)
- [To Andrew Dawson, Marci Harris, and Cybele Bjorkland \(House\)10/6/2009](#) (external link)
- [To Andy Schneider \(House\) 9/14/2009](#) (external link)
- [To EPTF with my proposed markup of HR 3200 10/2/2009](#) (external link)
- [To Andy Schneider with markup of HR3200 IG Provisions 10/6/2009](#) (external link)

The last two emails are of special importance. They reflect a concentration on the oversight of the health insurance provisions (the key health care reform proposals). At the time these emails were prepared the House and Senate were within days of going into conference to come up with a single bill that both chambers could agree on. At that time the oversight of health insurance provisions was lacking in the Senate bill and weak in the House bill.

The EPTF authorized its Evaluation Policy Consultant to offer a markup of the House bill to strengthen the oversight provisions. Based on this proposal, the conference committee took up the issue of oversight and included a single line in the follow up bill which reflected the main amendment that we were proposing:

“SEC. 1559. OVERSIGHT. The Inspector General of the Department of Health and Human Services shall have oversight authority with respect to the administration and implementation of this title as it relates to such Department.”

As a result of this single line, all of the health insurance reforms will be subject to scrutiny by the Inspector General (IG) of the Department of Health and Human Services. The persons who put this section into the bill knew full well that they were authorizing evaluations as well as audits and investigations.

GPRA Modernization Act

On January 4, the President signed HR 2142, the [GPRA Modernization Act of 2010](#) (external link). It amends the Government Performance and Results Act (GPRA) and other statutory provisions related to performance reports, and incorporates some broad principles underlying the Program Assessment Rating Tool (PART) of President Bush's administration and many of President Obama's policies related to a highly performing government. The Act provides a three tiered approach to performance management that includes four-year strategic plans, annual performance plans, and high priority goals.

Of special interest to evaluators are the following provisions:

- Evaluation policies carried over from the original GPRA legislation, including requirements to describe program evaluations used in establishing or revising general goals and objectives in

agencies' strategic plans, to provide a schedule for future program evaluations, to evaluate agency performance plans against performance goals, and to include in annual performance reports a summary of relevant program evaluation findings.

- New evaluation requirements: A requirement for the "Director of the Office of Personnel Management, in consultation with the Performance Improvement Council, . . . [to] identify the key skills and competencies needed by Federal Government personnel for developing goals, evaluating programs, and analyzing and using performance information . . ."
- New roles for the Office of Management and Budget (OMB): The responsibility to assess program performance and to inform the agency, the Congress, and the Government Accountability Office of unmet goals. The head of the agency may need to prepare plans to correct performance deficiencies.
- Transparency: The establishment of a Federal website to publish performance goals and assessments.

The EPTF cannot claim credit for this reauthorization. However, the EPTF's Evaluation Policy Consultant, had visited on two separate occasions congressional staff members working on early versions of this bill, providing copies of AEA's Evaluation Roadmap, answering questions about evaluation policy, suggesting broader coverage of evaluation, and promoting the development of the evaluation community. The bill reflects the general tenor of advice provided at those meetings.

GAO Review of International Feeding Programs

Another way in which the EPTF's work has affected evaluation policy is through GAO's adoption of AEA's Roadmap as a standard, or criterion, for evaluating the oversight of two international feeding programs. This can be seen in their reports on:

- [The McGovern-Dole Food for Education Program](#), (external link) and
- [International Food Assistance: USAID Is Taking Actions to Improve Monitoring and Evaluation of Nonemergency Food Aid, but Weaknesses in Planning Could Impede Efforts](#) (external link)

Clinical and Translational Research

The Clinical and Translational Science Awards (CTSA) of the National Institutes of Health (NIH), which encompasses 60 institutional sites and the national consortium made up of representatives from the sites and from NIH, is designed to enable innovative research teams to speed discovery and advance science aimed at improving our nation's health, tackling complex medical and research challenges and turning discoveries into practical solutions for patients. In 2011, the CTSA Consortium reached its expected size of 60 medical research institutions located throughout the nation with a total budget of \$500 million, the largest program at NIH.

The Consortium's Evaluation Key Function Committee has undertaken the task of preparing a set of principles to promote making evaluation an integral part of the initiative. It is developing evaluation guidelines that are roughly equivalent to a more detailed version of the AEA Evaluation Roadmap,

adapted for clinical and translational research. The committee has invited EPTF input to the paper. EPTF members already have provided input to its early drafts, George Grob is serving as an advisor to the committee, and he and EPTF Chair Patrick Grass have made a presentation to consortium members assembled in the D.C. area. [A draft of the paper](#) (external link) which acknowledges the EPTF's input, is now undergoing a final round of changes. When completed, it will advance evaluation policies and principles for this important component of national biomedical research. As in many other cases, AEA's Evaluation Roadmap was used as a starting off point of departure for adaptation to the CTSA program.

Potential or Emerging Legislation

The EPTF's Evaluation Policy Consultant or EPTF members meet with agency or congressional staff to discuss potential or emerging evaluation policies. Two recent examples are discussions with congressional staff developing evaluation legislation on behalf of Senators Mark Udall, Thomas Carper, and James M. Inhofe. All three were considering introducing legislation to strengthen the practice of evaluation across the entire Executive Branch. They have not yet introduced draft bills, but have asked that the EPTF stay in touch should their ideas move forward. It is possible that nothing will come of these initial efforts. However, exactly the same thing had happened with earlier versions of legislation that initially focused narrowly on PART-like requirements but that gradually merged into the broader GPRA Modernization Act discussed above.

AEA Comments on Proposed Regulations and Other Policies

Commenting on proposed legislation is not the only way that AEA can influence evaluation policy. The EPTF's comments on OMB's PART Guidance, discussed above in the section on *Contacts with Bush Administration Officials*, is one example of affecting policy through comments on Executive Branch policies and procedures. Other arenas for influence include proposed regulations and administrative procedures. The EPTF has been active in these fields too, by supporting the AEA Board and Presidential Rotation in making public comments on policies emerging through these mechanisms.

Paperwork Reduction Act. On October 27, 2009, OMB published a [request for comments](#) (external link) on the requirements of the [Federal Paperwork Reduction Act](#) (external link). Under the Act, all surveys and other data collection instruments issued by or on behalf of any Federal agency to more than 9 responders must be approved in advance by OMB. OMB sought comments on reducing current paperwork burdens, especially on small entities; increasing the practical utility of information collected by the Federal Government; ensuring accurate burden estimates; and preventing unintended adverse consequences.

The EPTF invited AEA members to offer advice in preparing comments in response to OMB's open invitation for input. AEA's comments were approved by AEA's Board of Directors and signed by AEA's Presidential Rotation. In [its comments](#), AEA recognized the importance of minimizing paperwork burdens on the American public. However, AEA believes that the current paperwork review process has unintended negative impacts, particularly in denying Federal managers and policy makers timely

access to information that is critically needed to address emerging problems and take advantage of promising approaches. AEA offered several options to speed up the reviews while promoting burden hour. reductions, improving burden estimates, and enhancing the usefulness of the data collected.

OMB has not yet acted on the advice provided by the EPTF and others. However, recently the Administrative Conference of the United States has decided to undertake its own evaluation of the Paperwork Reduction Act and interviewed the AEA's Evaluation Policy Consultant as part of that study. He did this based on the merits of AEA's comments to OMB. Here, then, we see another example of AEA being invited to the table during consideration of evaluation policies.

Home Visiting Programs. Health care reform legislation contains a provision authorizing a Maternal, Infant, and Early Childhood Home Visiting Program. The law requires that to be eligible for funding, a program must have been evaluated using rigorous, evidence-based methods. HHS published for comment its methodology for determining what evaluation methods would qualify under this standard. The proposed methodology, automatically reserves its "high" level rating to randomized assignment studies and relegates all other methods to "moderate" or low levels. The AEA Board and the Evaluation Policy Task Force worked together in preparing [comments](#) which were signed by AEA's Presidential Rotation--Leslie Cooksy, President, Debra Rog, Immediate Past President, and Jennifer Greene, President-elect.

The AEA comments raise concerns about how the studies upon which the evidence is based are rated and made recommendations to 1) forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings, and to avoid using the label "gold standard" in connection with random assignment designs in the rating methodology, 2) use additional criteria to assess the value of impact evaluations, 3) more specifically identify alternative impact evaluation methods, and 4) emphasize the value of multiple studies and mixed methods.

GAO Auditing Standards. The Government Accountability Office (GAO) requested public input on its sixth revision of the *Generally Accepted Government Auditing Standards*, commonly known as the Yellow Book. Because of the close connection of the Yellow Book to evaluation, AEA decided to respond. After consultation with AEA members and the AEA Board of Directors, AEA's Evaluation Policy Task Force prepared comments which, like the comments mentioned above on the Home Visiting Programs, were signed by AEA's Presidential Rotation.

AEA's comments noted the importance of the Yellow Book, especially to evaluators who conduct performance audits. However, AEA also noted "overly rigid requirements . . . that inappropriately constrain all performance audits within the mold of a normative methodology." The comments also called for greater emphasis on a standard of completeness, namely that "Being complete . . . also means the report states evidence and findings without omission of significant relevant information related to audit objectives."

Defending Evaluator Independence. A New York Times article by Ian Urbina, "[Pressure Limits Efforts to Police Drilling for Gas](#)," (external link) documents political and industry efforts to suppress the findings of Environmental Protection Agency (EPA) research on and recommendations regarding

hazardous waste from a gas-drilling technique known as "hydrofracking." The article says, "More than a quarter-century of efforts by some lawmakers and regulators to force the federal government to police the industry better have been thwarted, as EPA studies have been repeatedly narrowed in scope and important findings have been removed."

AEA leadership, including members of the Board of Directors and the Evaluation Policy Task Force were alarmed enough about this to authorize a letter to the Editor of the Times. The [AEA letter](#) was signed by AEA President Jennifer Greene

III. PUBLIC PRESENCE INITIATIVE (BRANDING)

Summary

We believe that the EPTF public presence initiative has contributed to a growing awareness of the importance of evaluation policy within Federal agencies and within AEA itself.

For example, it is now common place for executive branch and congressional leaders to link evaluation to government functions like

- budgeting (as evidenced in annual budget guidance documents) and
- program administration (such as, health care reform and international development programs and diplomacy).

Its importance within AEA is reflected in

- Board decisions to involve both Board members and the AEA membership at large in emerging evaluation policies whenever possible, and
- Establishment of an Evaluation Policy TIG,
- Establishment of AEA policies that institutionalize processes for responding to policy-shaping opportunities

These results followed concerted efforts of the EPTF to engage AEA members and the broader evaluation community in evaluation policy through:

- Annual EPTF updates at AEA conferences
- Monthly AEA "Policy Watch" newsletter articles
- The Evaluation Policy Discussion List
- Evaluation Policy page on AEA's website, and
- Presentation to evaluation associations such as the:
 - Eastern Evaluation Research Society
 - Washington Evaluation Research Network
 - Department of State Evaluation Conferences
 - Environmental Evaluators Network
 - CDC/AEA Evaluation Summer Schools
 - Inspectors General Evaluation Round Table
- Invitations to comment on emerging policies such as:
 - The AEA Evaluation Roadmap
 - Paperwork Reduction Act
 - GAO Auditing Standards, and
 - Protection of Human Research Subjects.

The ability to influence public evaluation policy cannot depend solely on the EPTF's own activities or providing comments on emerging policies. AEA may, after all, be unaware of policies in the making. influence may be more powerful if its "in the air," if people or organizations interested in an evaluation policy matter have a place to go to find out about it, and if AEA or the EPTF is "invited to the table" when evaluation policies are being considered. Furthermore, it is important that AEA members be aware of what is happening in the policy world, know what positions its leadership is taking, and have the opportunity to speak up on evaluation policy matters. Meeting these objectives is the purpose of the EPTF's "Public Presence Initiative," sometimes known in policy worlds as "branding." The AEA Board of Directors, in its Charge to the EPTF, called for a vigorous public presence initiative. The EPTF's actions in this arena are described here.

Outreach to Public Policy Groups and to Evaluators

From the outset, the EPTF initiated actions to reach out to the broad community of evaluators, whether or not they were members of AEA, Specific activities included presentations mad to the:

- Washington Evaluation Research Network
- Eastern Evaluation Research Society
- Department of State evaluation conferences
- Environmental Evaluators' Network
- CDC/AEA Evaluation Summer School, and
- Council of Inspectors General on Integrity and Efficiency

The Evaluation Policy Consultant met in person with numerous evaluators in Federal agencies and with policy makers in the Executive agencies and congressional committees staff members. This included:

- Prominent evaluators
- Agency and OMB staff involved with implementing federal programs for which evaluation requirements exist or would be highly relevant, and
- Congressional committee staff working on evaluation legislation or on legislation for which evaluation would be highly relevant

A complete accounting of such outreach efforts can be found in the EPTF Evaluation Policy Consultant's [log of outside meetings](#).

Outreach to AEA Evaluators

The AEA Board and the EPTF have been particularly focused on keeping AEA members informed of evaluation policy activities and involving them to the extent possible in developing evaluation policies that would be advocated by AEA. For this reason the EPTF has:

- Provided EPTF update sessions at each annual AEA conference since 2007

- Established and maintained an [Evaluation Policy section](#) (external link) within the AEA website
- Established an Evaluation Policy Discussion List which has hosted 349 messages involving 81 AEA members on current evaluation policy topics
- Published 34 monthly "[Policy Watch](#)" articles in the AEA Newsletter
- Supported the formation of the Evaluation Policy AEA Topical Interest Group (TIG)
- Invited AEA members' advice on
 - the final version of the AEA Evaluation Roadmap
 - the Paperwork Reduction Act
 - GAO Auditing Standards
 - Protection of human research subjects

Formal Procedures for Board and Member Involvement in Evaluation Policy

Based on early experience with involvement of AEA members in emerging evaluation policies such as those mentioned in the previous paragraph, the AEA Board directed the EPTF to prepare practical policies and procedures to involve AEA members and the Board to the greatest extent possible in

- Future revisions to the AEA Roadmap, and
- Development of official AEA positions on evaluation policies emerging through regulations, administrative issuances, and other policy mechanisms.

In response, the EPTF developed proposed policies, [AEA Board and Member Involvement in Public Documents on Evaluation Policy](#), which were discussed by the Board at its Spring, 2011 meeting and approved for incorporation into AEA's standing policy document.

Advocacy Documents

The Board's Charge to the EPTF includes, under the "Public Presence Initiative" the

- "Development of general "talking points" on evaluation and its role that would be likely to have broad consensus from AEA members and could be used in consultative work on evaluation policies (possible examples of talking points might be, for example, that: evaluation is a critically important endeavor; legislation and regulation should have explicit written requirements for evaluation of federally-funded programs; a broad range of evaluation methods are likely to be applicable and legislation and regulation should require that multiple methods and approaches appropriate to the program being assessed should be considered; etc.)." and
- "Development and enhancement of AEA collateral materials (e.g., brochures, Guiding Principles, public statements) that will support the consultative campaign."

In response, the EPTF has developed a small number of such documents. These include:

- Documents prepared for initial contacts with incoming Obama Administration officials
- [Draft Talking Points](#)
- [One pager: Program Evaluation for Management and Results](#)

Later, as EPTF and Board policies and procedures were formalized, the Board approved two foundation documents now used as reference points for public statements made by EPTF members of the EPTF Evaluation Policy Consultant

- [Evaluation Advocacy Statements Approved by the Board](#)
- [Evaluation Roadmap for a More Effective Government](#)

Cultural Change in AEA

Perhaps one of the most abiding and influential impacts of the Evaluation Policy Task Force initiative is the widespread acceptance of the legitimacy and importance of evaluation policy improvement as an objective of AEA as an organization. In the early days of the initiative very few AEA members or leaders were aware of evaluation policy or were accepting of it as an appropriate arena for AEA activities. In fact some were skeptical of success and, even more important, of the appropriateness of AEA becoming involved in evaluation policy, including advocating for legislation, regulations, and general awareness of the importance of evaluation as an integral feature of effective government. Not only has this notion become acceptable, but the AEA Board has required that AEA members and the Board of Directors be involved in this activity to the extent possible.

Furthermore, in the early days of the EPTF initiative, AEA Board members and many other AEA leaders did not understand the demands of tight deadlines in influencing public policy and exploiting policy opportunities. Nor were there any procedures to follow in developing and approving evaluation policy. Today the need for timeliness and the benefits of policy engagement seem to be widely accepted within AEA leadership circles and the AEA membership at large.

Appendices

EPTF PERFORMANCE SUMMARY APPENDIX A: ASSESSMENT AND RECOMMENDATIONS TO THE AEA BOARD OF DIRECTORS – JULY 2009

SUMMARY

Purpose

To present recommendations of the AEA Evaluation Policy Task Force (EPTF) to the AEA Board of Directors about the continuance of evaluation policy influencing initiative.

Background

The EPTF was established on July 1, 2007. *The Charge to the Task Force* establishes a requirement for it to report to the AEA Board with its recommendations about the future of the initiative. (A copy of *The Charge* is attached)

Assessment

During its first two years of operations, the EPTF laid the groundwork for influencing the formation of policies on matters affecting the practice of evaluation by directly engaging Federal policy makers on major legislative and administrative matters and publishing papers describing and promoting important evaluation policies.

Achievements included: advice to the Associate Director of OMB for Administration and Government Performance on how evaluation should be used in connection with OMB's PART program; advice to the incoming Director of OMB on an "Evaluation Roadmap" for incorporating evaluation as an essential aspect of good government; budget language adopted by the Appropriations Conference Committee expressing the need for increased funding for evaluation of U.S. global HIV/AIDS programs; and the introduction of a bill establishing evaluation as a major component of a proposed reform of all U.S. foreign assistance programs.

As a prelude to these initiatives, the EPTF established its presence in the evaluation community and involved them in its policy influencing activities through presentations at AEA 2007, 2008, and 2009 national conferences; establishment of an evaluation policy discussion group; AEA newsletter articles through a monthly "Policy Watch" column; establishment of an EPTF webpage within the AEA Website; and listening to other evaluators in various external forums, including meetings of Federal Evaluators (a Washington, D.C. based confederation of senior evaluators working at Federal Agencies); the Washington Research Evaluation Network (WREN); and the Eastern Evaluation Research Society (EERS).

Recommendations

The EPTF recommends that the AEA Board take the following actions:

- 1. Three Year Continuation.** Continue the EPTF for three more years.
- 2. EPTF Membership.** Increase to 10 the limit on the number of members of the EPTF.

3. Chair. Change the requirement that the Chair be a member of the Executive Committee to a requirement that at least one member of the EPTF be a member of the Presidential rotation.

4. Policy Clarifications

A. Adopt a policy that clarifies AEA's role in evaluation policy-shaping, which might include

- serving as a broker to resolve major evaluation issues
- providing a forum for discourse
- taking public positions on evaluation issues.

B. Authorize the Executive Director and the EPTF to routinely advocate for:

- Broad use of evaluation in public programs, especially those of the Federal Government
- Use of multi-method approaches
- Tailoring evaluation across the life-span of a program
- Adequate funding for evaluation, and
- Use of qualified, experienced evaluators as appropriate.

5. Vetting Processes

Develop processes for vetting documents and statements of the EPTF.

ASSESSMENT AND RECOMMENDATIONS TO THE AEA BOARD OF DIRECTORS – JULY 2009

DETAIL

Purpose

To present recommendations of the AEA Evaluation Policy Task Force (EPTF) to the AEA Board of Directors about the continuance of evaluation policy influencing initiative.

Background

The EPTF was established on July 1, 2007. *The Charge to the Task Force* establishes a requirement for it to report to the AEA Board with its recommendations about the future of the initiative. (The full copy of *The Charge* is attached). The EPTF conducted a self assessment in August 2008 and then updated it in May 2009. This assessment is summarized below and is followed by recommendations to the Board.

I. ASSESSMENT

The EPTF established two work groups to prepare its self assessment, one on its consultative campaign and one on the public presence task. Following are results, prepared by George Grob, EPTF consultant, based on deliberations of the two work groups and on further discussions with the Chair and EPTF Members.

A. What Did the EPTF Accomplish?

Evaluation Policy Advice

The EPTF directly engaged Federal policy makers on legislative and administrative matters of major importance to the practice of evaluation. These engagements involved senior officials with influence over evaluation policies at the highest levels of the Federal government. Examples include:

- **Evaluation Roadmap.** Prepared a paper entitled “Evaluation Roadmap for a More Effective Government” which was sent to the newly appointed Director of OMB on February 3, 2009, under the signatures of three AEA Presidents (current, immediate past, and president-elect). The document was used by senior officials of OMB in their discussions of new initiatives to improve the use of evaluation as an integral part of developing and managing Federal programs. Because of the Roadmap, AEA President Debra Rog has been asked to speak at the “Inspection and Evaluation Training Conference” that is being convened in June by the Council of the Inspectors General on Integrity and Efficiency’s Inspection and Evaluation Committee. Similarly, EPTF members Stephanie Shipman and Patrick Grasso have been asked to make a presentation about the Roadmap to a Department of State evaluation conference in June. The Roadmap was also discussed by Debra Rog and George Grob at the Environmental Evaluators Network’s national conference in June.
- **U.S. Global HIV/AIDS Program.** Collaborated with the Lundy Foundation to provide advice and technical assistance to the Congress on the importance of providing adequate funding for evaluation of the U.S. global HIV/AIDS program.
 - Provided the advice in formal written comments to the House Appropriations Committee. Subsequently, the Appropriations Conference Committee included our recommended language in its bill report.
 - Prepared a detailed list of evaluation requirements, responsibilities, and deadlines in the newly enacted reauthorization of the global HIV/AIDS program for the purpose of tracking evaluation funding actions and provided it to senior officials of USAID and the Office of the U.S. Global AIDS Coordinator as well as numerous stakeholders interested in international HIV/AIDS programs.
- **International Assistance.** Again in collaboration with the Lundy Foundation, prepared, at the request of staff of the House Foreign Affairs Committee, a draft legislative proposal with supporting documents that would make evaluation an essential function in the development and implementation of all U.S. foreign assistance programs. The draft bill, which includes language based on the AEA/Lundy proposal, was introduced by Representative Howard Berman, Chair of the House Foreign Affairs Committee, April 28, 2009.
- **Government Efficiency and Effectiveness.** Provided advice, based on the Evaluation Roadmap, to staff of the Oversight and Government Reform Committee on its bill, “The Government Efficiency, Effectiveness, and Performance Improvement Act.”
- **OMB Evaluation Guidance in the PART Program.** Met with Robert Shea, Associate Director of OMB for Administration and Government Performance. At his request, the EPTF sent a formal paper to him on March 7, 2008, with comments on OMB’s PART Guidance entitled *What Constitutes Strong Evidence of a Program’s Effectiveness*. The paper was considered by OMB staff in their development of a demonstration program of expert advice on impact evaluation methods to be used by Federal Agencies

- **Federal Performance Improvement Council.** Addressed senior Federal Government evaluators of the Evaluation Working Group of the Performance Improvement Council.
- **NIH.** Established ongoing engagement with NIH to facilitate improvement of evaluation of biomedical research programs.

Infrastructure

Established the infrastructure of an evaluation policy function in AEA, including:

- **Talking Points.** Talking points established the overall themes relating to how evaluation can help bring about substantial improvement in government programs; effective evaluation practices that AEA believes should be adopted in public program administration; and the hallmarks of an effective national framework for ensuring that effective evaluation practices are routinely adopted in the development and management of public programs.
- **Ongoing operations.** The EPTF has evolved as a practical mechanism for developing and monitoring overall evaluation policy functions. It is knowledgeable, experienced, balanced, available for regularly scheduled meetings, and able to respond, on occasion, to demands for intensive advice giving and reviewing of documents. A consultant provides a steady resource for professional advice and staff work.
- **Webpage.** A webpage publishing its activities and fundamental documents is included on the AEA website
- **Priorities.** The EPTF identified priorities related to broad government evaluation policies, education, large Federal research programs, international development programs, and national health insurance.

Evaluator Involvement

Involved evaluators in the policy influencing activities, including both AEA members and other evaluators. This included:

- **AEA Members.** The EPTF invited AEA members to participate through AEA 2007, 2008, and 2009 conference presentations, established an evaluation policy discussion group, invited members to serve as on-call experts, invited members to provide examples of successful evaluations, published AEA newsletter articles through a monthly “Policy Watch” columns
- **Other Evaluators.** The EPTF listened to evaluators’ concerns and suggestions through meetings with Federal Evaluators (a Washington, D.C. based confederation of senior evaluators working at Federal Agencies); the Washington Research Evaluation Network (WREN); the Eastern Evaluation Research Society (EERS); OMB senior program performance assessment staff; and one-on-one meetings with senior evaluators in Federal agencies.

B. What Did the EPTF Learn?

Intense, Crisis Atmosphere of Some Policy Interventions

The opportunity to influence evaluation policy often comes in intensive bursts of time that require sporadic change of schedules and single minded dedication to preparing rapid input regarding evaluation policy in the legislative, regulatory, or administrative development processes. The EPTF

had only a week and a half to draft its paper on the PART guidance, and just two months to develop a detailed legislative proposal for the Foreign Assistance Act. Having coherent materials with the right tone and scope, such as the Evaluation Roadmap and the talking points, is crucial to being able to provide advice when requested.

Ease of Access to Policy Officials

It is possible, even relatively easy, to gain access to senior Federal policy officials to explain the importance of evaluation and its role in the development and management of Federal programs.

Favorable Climate for Evaluation Policies in the Underlying Legislation for Specific Programs

Contrary to our expectations, we were surprised to find that in some quarters, particularly in the Congress, there is a growing interest in embedding the use of evaluation and a variety of program assessment processes into the workings of government programs. Policy development staff in some congressional committees and program managers in some Federal agencies would like to know more about the practical methods for doing so. Policy officials and program managers would be more easily persuaded of the value of evaluation if they had easy to read explanations along with examples of how evaluation can help them. The persuasiveness of such materials would be enhanced if they were periodically updated and routinely distributed. Policy officials and program managers appreciate easy to read, balanced, professionally reliable documents. They work in high pressure, rapid turn around environments and are most easily persuaded by information they get at the time that decisions are made.

Continuous Interventions Needed to Achieve Success

Our ability to convince policy makers to incorporate evaluation policies in key program documents will depend on continuous involvement in policy making processes. Legislation, regulations, and administrative guidelines go through many iterations, sometimes reflecting major swings in policy content. To be successful, AEA will need to maintain close surveillance of the activities in program areas whose evaluation policies it hopes to influence, so it does not miss key steps or opportunities to intervene. Sometimes missing such steps could result in losing irretrievable opportunities to get our proposals accepted.

Building General Evaluation Policy on a Foundation of Iterative Successes

With the exception of a few general evaluation policies, such as PART or across-the-board changes in evaluation funding set-aside percentages, policy makers in the Congress and Executive Branches are not interested in general discussions of evaluation policy. They focus on specific programs, usually during reauthorizations or major changes in regulations. As a result, major successes in spreading effective evaluation policies throughout Federal programs may hinge on working on them one at a time and then using the emerging evaluation policies as examples for successive programs. For example, the evaluation policies now embedded in the recently draft foreign assistance bill could be used as models for evaluation policies to be included in other laws.

Networks of Policy Contacts

Networks of mutually respected individuals on both sides of the policy development function (in our case evaluation policy experts on the one hand and programmatic policy staff of the Congress and

Executive Branch on the other) are the oil that makes the policy machinery run. It is therefore important to establish, grow, and maintain such networks. Such connections need to be systematically documented and institutionalized so that AEA's relationships with key outsiders will not depend exclusively on personal relationships of EPTF members or the EPTF consultant or staff.

Dynamism of the Evaluation Function

Proposed evaluation policies must reflect the dynamic nature of the evaluation function itself. To be sure of what is happening in evaluation workplaces and how that may be changing, the EPTF needs to learn more about their on-going experience.

Need for Supportive Materials

We have already reached the stage where people are asking for examples of influential evaluations and examples of legislative or administrative language they can adopt to strengthen evaluation provisions of the laws, regulations, or administrative documents they are working on.

Connection with AEA Board

The EPTF was fortunate to have the support of the current and past AEA presidents. For the future it will be necessary to institutionalize the connection between the evaluation policy function (however organized and wherever located organizationally within AEA) and the AEA Board.

Need for Evaluation Policy Making Machinery

The EPTF's desire to share information with AEA members was thwarted by the lack of a clear and efficient policy decision mechanism. Given the lack of pathways for member consultation, EPTF members were uncomfortable about the propriety and usefulness of publishing unapproved documents that could be rescinded later on or that do not truly represent the positions of the AEA or its members at large. Closely connected with this is the need for the AEA Board and AEA members to validate what the EPTF consultant is saying in his outside contacts.

Need to learn more about evaluation policies of specific Federal agencies and the Congress

The EPTF learned a great deal about the Executive Branch's evaluation policy at its highest levels, e.g., at OMB. It also learned a bit about how evaluation policy is made in some Federal agencies, such as Department of Education, the Department of Health and Human Services, and the State Department, but not as much about the evaluation policies of other agencies. It currently knows very little about how evaluation policy is made within the Congress, where the work is of an ad hoc nature and varies considerable from committee to committee.

II. RECOMMENDATIONS

The EPTF recommends that the AEA Board take the following actions:

1. Three Year Continuation

Continue the EPTF for three more years. Initial success would seem to justify a continuation of this activity. At the same time, a three year extension will provide additional time for the AEA Board to evaluate it.

2. EPTF Membership

Raise the limit on the number of members of the EPTF to 10. This will allow greater diversity in viewpoints and advice and permit the gradual absorption of new members, but still keep it small enough to promote effective internal communication and flexibility.

3. Chair

Change the requirement that the Chair be a member of the Executive Committee to a requirement that at least one member of the EPTF be a member of the Presidential rotation. This will allow the Board greater flexibility in appointing the Chair.

4. Policy Clarifications

A. Adopt a policy that clarifies AEA’s role in evaluation policy-shaping, which might include

- serving as a broker to resolve major evaluation issues
- providing a forum for discourse
- taking public positions on evaluation issues.

B. Authorize the Executive Director and the EPTF to routinely advocate for:

- Broad use of evaluation in public programs, especially those of the Federal Government
- Use of multi-method approaches
- Tailoring evaluation across the life-span of a program
- Adequate funding for evaluation, and
- Use of qualified, experienced evaluators as appropriate.

5. Vetting Processes

Develop processes for vetting documents and statements of the EPTF. In deciding this, the Board should take into account that:

- The EPTF works under intensive time pressures that are sometimes incompatible with the normal procedures used for Board deliberations.
- There may need to be different types or classes of documents that require different levels of vetting. For example, major statements, like the recent Evaluation Roadmap document, should require more vetting than others, like the consultant’s earlier more “internal” memo provided to an OMB staffer with a clear indication that it has not yet been fully vetted within AEA.

The EPTF is ready to assist the Board in developing such policies.

The next section (“The Charge”) is a [excerpt from] the original charge to the EPTF that has been edited to conform to the first three recommendations. Proposed changes are highlighted in yellow.

EPTF CHARGE RECOMMENDED REVISIONS

Proposed changes are highlighted in yellow

....Earlier Sections to remain unchanged and are not included here for purpose of brevity

Appointment, Time Commitment, Structure and Support

The Evaluation Policy Task Force is appointed for an additional ~~two~~ **three**-year period beginning in July, ~~2007~~ **2009**. The Evaluation Policy Task Force is a Task Force of the AEA Executive Committee and, as such, will report through the EC and may use the EC as a sounding board as desired.

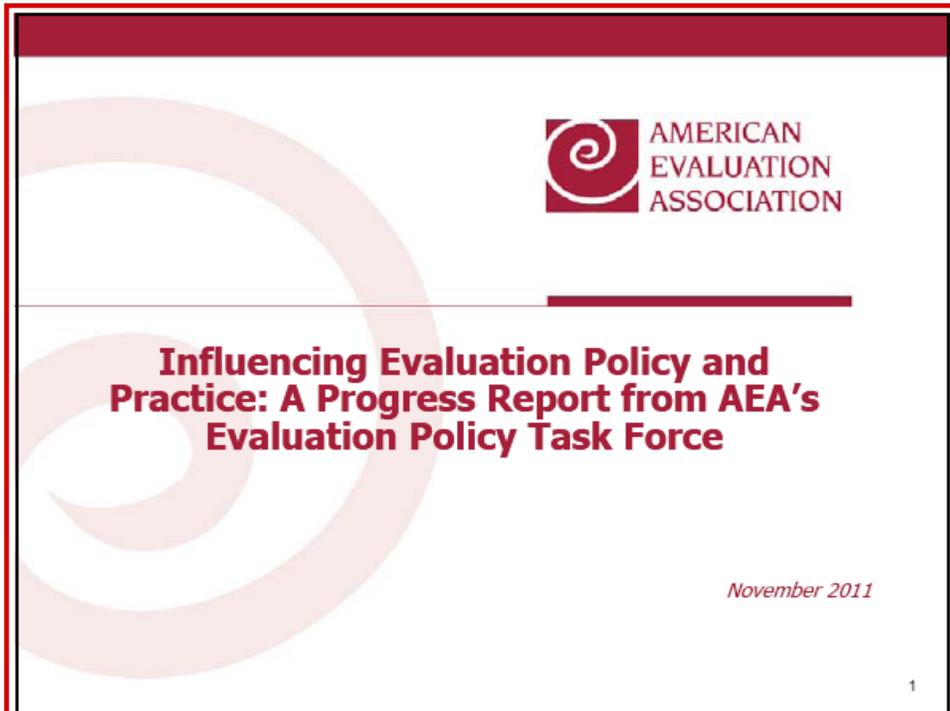
The Task Force will consist of ~~5 members up to 10 members, including~~ the Executive Director, ~~a member of the Presidential rotation~~, and a liaison from the Public Affairs Committee (either the Chair or a PAC member). ~~The Chair of the Task Force will be a member of the Executive Committee.~~ The Evaluation Policy Consultant will be integrally involved in the Task Force and will participate in meetings and discussions as appropriate. The Task Force Chair and AEA Executive Director will identify potential members, subject to approval of the Executive Committee, with the following criteria to guide selection of members:

- Commitment to supporting the mission and goals of the American Evaluation Association.
- Knowledge of and a history of prior involvement with the American Evaluation Association.
- Familiarity with the field of evaluation and capacity to understand and represent the field to others.
- A broad perspective on evaluation and willingness to advocate for the many and diverse views of evaluation.
- Knowledge of the Federal policymaking process.
- Experience with policy development initiatives in the Federal Government.
- Experience with public presentations of evaluation to a variety of audiences.
- Diversity of the Task Force and representativeness of the breadth of members and interests of AEA.

The Task Force will provide a written annual report to the AEA Board for review at its Winter Board meeting and updates as needed through the Task Force Chair at other Board meetings. ~~At the completion of the first year of their work, the Task Force will be expected to report to the Board with their recommendations about the feasibility and desirability of continuation of this effort and the best mechanisms for doing so (e.g., continuation of the Task Force; establishment of a standing committee or assignment of this effort as a subcommittee to an existing committee).~~ The AEA Board will provide ongoing feedback and guidance based on its reviews of these reports.

.....continued

EPTF PERFORMANCE SUMMARY APPENDIX B: 2011 POWERPOINT SUMMARY OF EPTF ACTIVITIES AND ACCOMPLISHMENTS

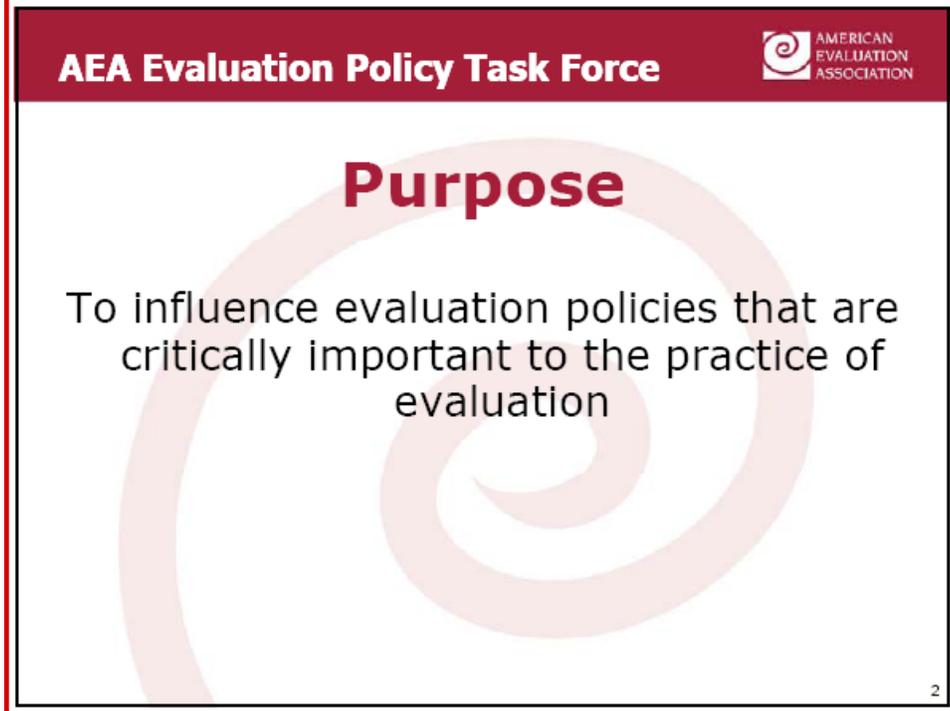


AMERICAN EVALUATION ASSOCIATION

Influencing Evaluation Policy and Practice: A Progress Report from AEA's Evaluation Policy Task Force

November 2011

1



AEA Evaluation Policy Task Force AMERICAN EVALUATION ASSOCIATION

Purpose

To influence evaluation policies that are critically important to the practice of evaluation

2

Members

- **Eleanor Chelimsky**
- **Katherine Dawes**
- **Patrick Grasso (Chair)**
- **Jennifer Greene**
- **Susan Kistler (AEA Executive Director)**
- **Mel Mark**
- **Stephanie Shipman**
- **George Grob (Advisor)**

3

Consultative Campaign

- **Federal Government**
 - **Office of Management and Budget**
 - **The Administration**
 - **Priority Legislation**
 - **Federal Regulations**
 - **Administrative Guidances**

4

Connection with Policy Makers



- ④ Evaluation Roadmap to OMB and Others
- ④ Meetings with Senior Federal Officials (Bush and Obama Administrations)
- ④ Health Care Reform Advice
- ④ Advice to OMB on PART Program
- ④ Presentation to Federal Performance Improvement Council

5

Connections with Policy Makers



- ④ Initial Contact with House and Senate Budget Committees
- ④ Material for US Global AIDS Office on PEPFAR Implementation
- ④ Advice to Congressional Staff Developing Evaluation Legislation
 - ④ Senator Mark Udall
 - ④ Senators Thomas Carper and James M. Inhofe

6

Policy Successes



- ☉ OMB Impact Evaluation Policy
- ☉ Evaluation Policies in the Federal 2011 and 2012 Budget Guidances
- ☉ U.S. Global HIV/AIDS Appropriations Guidance
- ☉ Evaluation Requirement of the Global Health Initiative
- ☉ Requirements for oversight of Health Care Reform
- ☉ GAO Review of International Food Assistance Programs Using Criteria Developed from the Roadmap

7

Policy Successes



- ☉ USAID Evaluation Policy
 - ☉ Integration of evaluation and program planning
 - ☉ Definitions and Distinctions of evaluation types
 - ☉ Requirements for evaluations of major programs and untested interventions
 - ☉ Acknowledgement of need for both quantitative and qualitative methods
 - ☉ 3% set aside of major program funds for evaluation

8

Policy Successes



State Department Evaluation Policy

- Evaluation framework covering all programs, projects, and activities in bureaus and missions
- Evaluation requirements
- Evaluation plans
- Evaluator Independence
- Evaluation Documentation

9

Policy Successes



GPRA Modernization Act

- Carries over existing policies requiring evaluation
- Schedules. Evaluating success in meeting program goals
- Requires identification of skills and competencies for goal setting and evaluation
- Requires Federal website to publish goals and assessments

10

GAO Report on International School Feeding Program

-  Uses AEA Evaluation Roadmap as the standard for evaluation of this program

11

Other Policy Advocacy

-  To OMB on the Paperwork Reduction Act
-  To HHS Health Resources and Services Administration on evaluating evidence of the effectiveness for home visiting programs
-  To GAO on Auditing Standards
-  To the Inspection and Evaluation Committee of the Council of Inspector General on Integrity and Efficiency on evaluation training
-  To the Editor of the New York Times on Evaluator Independence

12

Public Presence Initiative

-  **AEA Member Engagement**
-  **Outside Meetings**
-  **AEA Newsletter Articles**
-  **Media Coverage**

13

Engagement of AEA Members

-  Annual EPTF updates at AEA conferences
-  Monthly AEA “Policy Watch” Newsletter articles
-  Evaluation Policy Discussion List
-  Evaluation Policy TIG
-  Invitation to comment on emerging policy related
 -  The AEA Roadmap
 -  Paperwork Reduction Act
 -  GAO Auditing Standards
 -  Protection of human research subjects

14

Engagement of Other Evaluators



Presentations to evaluation associations

-  EERS
-  Washington Evaluation Research Network,
-  Dept. of State Evaluation Conferences
-  Environmental Evaluators Network
-  CDC Summer School

15

On the Podium



 Inspector Generals' Evaluation Training	Debra Rog
 Department of State Evaluation Conference	Stephanie Shipman Patrick Grasso
 Environmental Evaluators Network Conference	Debra Rog George Grob
 CDC Summer Institute	George Grob

16

Evaluation Policy Infrastructure



- ❏ AEA Evaluation Roadmap for a More Effective Government
- ❏ Advocacy statements for public representation of AEA evaluation policy
- ❏ Website

17

Evaluation Policy Infrastructure



- ❏ Advocacy statements summarizing key evaluation policy approved by AEA Board
- ❏ AEA Evaluation Roadmap for a More Effective Government
- ❏ Requirements and Procedures for involving AEA members and Board of Directors in formulating evaluation policy proposals
- ❏ AEA Membership in Consortium of Social Science Associations (COSSA)

18

Possible Future Priorities



- 🌀 Targets of Opportunity
 - 🌀 Protection of human research subjects
 - 🌀 Clinical and translational research awards
 - 🌀 Other legislative and regulatory policy opportunities as they emerge
- 🌀 New EPTF Members
- 🌀 Formal assessment of the EPTF by the AEA Board
- 🌀 More opportunities for AEA members to influence evaluation policy
- 🌀 Strengthen connection with other evaluation organizations

19

AEA Members' Policy Watch



Staying In Touch

- 🌀 **Email:** evaluationpolicy@eval.org
- 🌀 **Website:** <http://www.eval.org/EPTF.asp>
- 🌀 **Discussion Group:** <http://www.eval.org/EPTF.signup.discussion.asp>

20

EPTF PERFORMANCE SUMMARY APPENDIX C: AN EVALUATION ROADMAP FOR A MORE EFFECTIVE GOVERNMENT, SEPTEMBER 2010 (SEE NEXT PAGE)



An Evaluation Roadmap for a More Effective Government

September 2010



The American Evaluation Association (AEA) is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness. AEA has over 6,500 members representing all 50 states in the United States as well as over 75 foreign countries.

Mission:

To improve evaluation practices and methods

Increase evaluation use

Promote evaluation as a profession and

Support the contribution of evaluation to the generation of theory and knowledge about effective human action.

**An Evaluation Roadmap
for a
More Effective Government**

In keeping with our mission, the American Evaluation Association hereby describes its vision of the role of evaluation in the federal government. We provide a roadmap for improving government through evaluation, outlining steps to strengthen the practice of evaluation throughout the life cycle of programs.

Evaluation is an essential function of government. It can enhance oversight and accountability of federal programs, improve the effectiveness and efficiency of services, assess which programs are working and which are not, and provide critical information needed for making difficult decisions about them.

AN EVALUATION ROADMAP FOR A MORE EFFECTIVE GOVERNMENT

The Challenge

Like all governments, the United States government faces challenges in both foreign and domestic policy arenas. Today, these challenges span subject areas such as national security, foreign aid, energy, the environment, health care, education, and the economy. Program or policy interventions are typically developed in response, in an effort to mitigate, resolve, or better understand the problems involved.

To determine the merit, quality, and usefulness of these interventions, credible information is needed about what the program or policy in question has achieved and at what cost. Such information is crucial if government officials are to ensure that the chosen interventions are working, that taxpayers' money is being spent wisely, and that the government is accountable to the public for the interventions and their results.

Why Program Evaluation Is Essential

Evaluation involves assessing the strengths and weaknesses of programs, policies, and organizations to improve their effectiveness. It provides a useful and important tool to address the need for credible information, well-grounded decision making, and governmental transparency. Within a government context, the legitimacy of evaluation can be seen as deriving from the structure of the government it serves and from the functions it fills.

In the United States, evaluation can serve information needs that arise within any of the three branches of government. For example, evaluation can provide information about a new program's initial outcomes, allowing for better management within the Executive Branch. It can also be used to assess the relative performance of a set of alternative policy options, informing legislative deliberations. The courts may cite evaluation findings as a basis for their judgments. More fundamentally, evaluation can contribute the evidence needed to support the system of checks and balances established by the United States Constitution. For example, evaluation enables congressional oversight and executive accountability, along with the development of new knowledge, innovation, and organizational learning in both branches. This commitment to accountability and transparency, on the one hand, makes evaluation essential to democratic government and, on the other, requires evaluation in a government context to be independent and to resist advocacy for particular positions.

Evaluation provides needed feedback for managing any program. It uses systematic data collection and analysis to address questions about how well government programs and policies are working, whether they are achieving their objectives, and, no less importantly, why they are or are not effective. Evaluation produces evidence that can be used to compare alternative programs, guide program

development and decision making, and reveal effective practices. By its very nature, it supplies the publicly accessible information that is at the heart of transparency and open government.

Since the inception of modern program evaluation, federal agencies have conducted many evaluations and applied their results to make reasoned program decisions. But for the most part, these evaluations have been sporadic, inconsistently applied, and inadequately supported. The units formed to conduct evaluations too often are short lived and under resourced. Training and capacity building for evaluation have been inconsistent across agencies and, in many cases, insufficient to achieve the needed evaluation capacity and sustain it over time.

Yet there is a strong case to be made for a commitment to evaluation as an integral feature of good government, whether the goal is better performance, stronger oversight and accountability, or more data-informed and innovative decision making. The lessons learned in agencies that have applied evaluation constitute a solid knowledge base upon which to build.

The U.S. government would benefit significantly from using program evaluation to

- Address questions about current and emerging problems
- Inform program and policy planning efforts
- Monitor program performance
- Provide timely feedback to decision makers to make changes when needed
- Increase accountability and transparency
- Reduce waste and enhance efficiency
- Improve programs and policies in a systematic manner
- Support major decisions about program reform, expansion, or termination
- Assess whether existing programs are still needed or effective
- Identify program implementation and outcome failures and successes
- Identify innovative solutions that work
- Inform the development of new programs where needed
- Examine the requirements for the transfer of promising programs to new sites
- Share information about effective practices across government programs and agencies

The key is to make program evaluation integral to managing government programs at all stages, from planning and initial development, through start up, ongoing implementation, appropriations, and reauthorization. In short, what is needed is a transformation of the federal management culture to one that incorporates evaluation as an essential management function.

Recommendations

We recommend that each federal agency adopt the following framework to guide the development and implementation of its evaluation programs.

Scope and Coverage

- Conduct evaluations of public programs and policies throughout their life cycles and use evaluation to both improve programs and assess their effectiveness
- Evaluate federal programs and policies in a manner that is appropriate for program stewardship and useful for decision making
- Build into each new program and major policy initiative an appropriate framework to guide evaluations throughout the life of the program or initiative
- For existing programs, assess what is already known and develop evaluation plans to support future decision making

Management

- Assign senior, experienced evaluation officials and managers to administer evaluation centers or coordinate evaluation functions at appropriately high levels of government agencies
- Prepare annual and long-term evaluation plans to guide decision making about programs
- Provide sufficient and stable funding to support professional evaluation activities
- Coordinate and communicate about evaluation efforts across agencies with overlapping or complementary missions
- Develop written evaluation policies across and within federal agencies that can guide evaluation efforts and help ensure their quality
- Ensure that evaluation units and staff receive high-level, public, and consistent support

Quality and Independence

- Develop and adopt quality standards to guide evaluation functions consistent with the American Evaluation Association’s Guiding Principles for Evaluators
- Promote the use and further development of appropriate methods for designing programs and policies, monitoring program performance, improving program operations, and assessing program effectiveness and cost
- Safeguard the independence of evaluation design, conduct, and results
- Preserve and promote objectivity in examining program operations and impact

Transparency

- Consult closely with Congress and non-federal stakeholders in defining program and policy objectives and critical operations and definitions of success
- Disseminate evaluation findings and methods relating to public accountability to policy makers, program managers, and the public
- Create clearinghouses to share information about effective and ineffective program practices

In this Roadmap, we more fully develop these ideas. We describe the general principles that should guide a government-wide effort to strengthen evaluation functions. We propose broad administrative steps to institutionalize evaluation in federal agencies. Finally, we discuss how the Executive Branch and Congress can jointly make the most effective and efficient use of evaluation.

General Principles

The following general principles should guide efforts to integrate evaluation into program management.

Scope. Evaluation should be integral to planning, developing, managing, and implementing government programs at all stages. Evaluation activities should be used to:

- Make sure that program designs are appropriate to achieve program goals
- Identify problems during start-up and correct them before they become entrenched
- Identify and share promising approaches that emerge during program implementation
- Assess the extent to which programs and policies are being well implemented
- To the extent feasible, establish expectations and performance standards at program inception and involve stakeholders in refining them as the programs mature
- Develop appropriate and efficient data collection and reporting systems and information technology support to provide a continuing flow of evaluative information to policy makers and program managers
- Examine the extent to which programs reach their intended beneficiaries
- Periodically examine selected program features to improve their effectiveness and efficiency
- Periodically assess program results and service quality
- Systematically examine whether an apparently successful program can be expanded to another setting before scaling it up

Coverage. In general, federal programs and policies should be subject to evaluation.

Analytic Approaches and Methods. Which analytic approaches and methods to use depends on the questions addressed, the kind of program evaluated, its implementation status, when the evaluation results are needed, what they are needed for, and the intended audience.

No simple answers are available to questions about how well programs work, and no single analytic approach or method can decipher the inherent complexities in the program environment and assess the ultimate value of public programs. Furthermore, definitions of “success” may be contested. A range of analytic methods is needed, and often several methods—including quantitative and qualitative approaches—should be used simultaneously. Some evaluation approaches are particularly helpful in a program’s early developmental stages, whereas others are more suited to ongoing and regularly implemented programs.

The broader policy and decision-making context also can influence which approach is most appropriate. Sometimes information is needed quickly, requiring studies that can use existing data or rapid data collection methods; at other times, more sophisticated long-term studies are required to understand fully the dynamics of program administration and beneficiary behaviors.

Over the years, the evaluation field has developed an extensive array of analytic approaches and methods that can be applied and adapted to a wide variety of programs, depending on the program’s characteristics and implementation stage, how the results will be used, and the kinds of decisions that

will be made. All evaluation methods should be context sensitive, culturally relevant, and methodologically sound. Evaluation approaches and methods include, but are not limited to:

- Case studies
- Surveys
- Quasi-experimental designs
- Randomized field experiments
- Cost-benefit and cost-effectiveness analyses
- Needs assessments
- Early implementation reviews
- Logic models and evaluability assessments

Resources. Evaluation should be supported through stable, continuous funding sources and through special one-time funds for evaluation projects of interest to Executive Branch and congressional policy makers. The stable, continuous evaluation funds should be provided through appropriations or program fund set-asides. These methods can also be combined to support viable evaluation programs. Program managers should authorize and require periodic evaluations of each program throughout its life to provide rich evaluative information to policy makers during annual appropriation and cyclical reauthorization and amendment discussions.

Professional Competence. Evaluations should be performed by professionals with appropriate training and experience for the evaluation activity (such as performing a study, planning an evaluation agenda, reviewing evaluation results, or performing a statistical analysis). Evaluation is an interdisciplinary field that encompasses many areas of expertise. Many evaluators have advanced degrees in, and often work collaboratively with colleagues in allied fields, such as economics, political science, applied social research, sociology, anthropology, psychology, policy analysis, statistics, and operations research. Federal agencies should ensure that the required diversity of disciplines is appropriately represented in internal and independent evaluation teams.

Evaluation Plans. Each federal agency should require its major program components to prepare annual and multiyear evaluation plans and to update these plans annually. The planning should take into account the need for evaluation results to inform program budgeting; reauthorization; agency strategic plans; ongoing program development and management; and responses to critical issues concerning program effectiveness, efficiency, and waste. These plans should include an appropriate mix of short- and long-term studies to produce results of appropriate scope and rigor for short- or long-term policy or management decisions. To the extent practical, the plans should be developed in consultation with program stakeholders.

Evaluation questions can spring up unexpectedly and urgently in response, for example, to a sudden need for information to address a presidential initiative, a management problem, or questions raised by Congress. Therefore, evaluation plans should allow for flexibility in scheduling evaluations.

Dissemination of Evaluation Results. The results of all evaluations related to public accountability should be made available publicly and in a timely manner (except where this is inconsistent with the Freedom of Information Act or Privacy Act). They should be easily accessible and usable through the internet. Similarly, evaluations of promising and effective program practices should be systematically and broadly disseminated to potential users in federal agencies. Evaluation data and methods should also—to the extent feasible and with sufficient privacy protections—be made available to professionals and the public to enable secondary analysis and assure transparency.

Evaluation Policy and Procedures. Each federal agency and its evaluation centers or coordinators (discussed below) should publish policies and procedures and adopt quality standards to guide evaluations within its purview. Such policies and procedures should identify the kinds of evaluations to be performed and the criteria and administrative steps for developing evaluation plans and setting priorities, selecting evaluation approaches and methods to use, consulting subject matter experts, ensuring evaluation product quality, publishing evaluation reports, ensuring independence of the evaluation function, using an appropriate mix of staff and outside consultants and contractors, appropriately focusing evaluation designs and contracts, and promoting the professional development of evaluation staff.

Independence. Although the heads of federal agencies and their component organizations should participate in establishing evaluation agendas, budgets, schedules, and priorities, the independence of evaluators must be maintained with respect to the design, conduct, and results of their evaluation studies.

Institutionalizing Evaluation

Significant progress has been made in establishing evaluation as an integral component of government program management. However, additional steps are needed.

Background

Some federal agencies, such as the Department of Defense, established evaluation offices in the late 1950s and early 1960s. Others, including the Departments of Education and of Health and Human Services, developed their evaluation functions in the 1970s within the then-Department of Health, Education, and Welfare. The authorizing statutes for some of these agencies set aside a fixed percentage of program funds for evaluation. Other departments have added evaluation offices to their organizations although these offices have grown and shrunk over the intervening years.

One relatively stable evaluation organization has been the Government Accountability Office (GAO), previously known as the General Accounting Office. It has remained the largest single government agency producing evaluations at Congress's request.

One of the most enduring evaluation-related functions has been the Government Performance and Results Act (GPRA) of 1993. This law encourages each agency to develop an agency-wide strategy and mission and also requires them to determine whether their programs achieve their goals and objectives. GPRA defines evaluation as assessing the "manner and extent to which" agencies achieve their goals, thus addressing both implementation and results. In practice, government agencies have

implemented GPRA by using performance indicators and measurement to determine whether they have reached a goal and have conducted few evaluation studies that might shed light on how programs reached their goals, why programs do or do not meet their goals and objectives, and how to improve programs. As a result, the GPRA process produces little information to guide programmatic or policy action.

We propose that government agencies, policy makers, and program managers build on the progress already made by embracing evaluation as an integral feature of good government. Agencies should consistently use program evaluation and systematic analysis to improve program design, implementation, and effectiveness and to assess what works, what does not work, and why. This comprehensive vision recognizes that evaluation is more than simply “looking in the rearview mirror” and needs to be used throughout a program’s life as an integral part of managing government programs at all stages.

For this approach to work, the Executive Branch and Congress will need to take action, as described below.

Executive Branch Role

As noted earlier, the infrastructure and practice of program evaluation in federal agencies is somewhat of a mixed story. Some agencies have well-developed and stable evaluation offices; others do not. The same can be said for evaluation funding, scope, policies, planning, and dissemination.

Different federal agencies and programs have different evaluation needs, and the maturity and breadth of their evaluation programs vary. In addition, the evaluation function might be a component of other offices focused on such functions as management, planning, research, and policy development, including legislative or regulatory development. For example, several agencies have offices of planning and evaluation, research and evaluation, or monitoring and evaluation, and some inspectors general have offices of inspections and evaluations. In some agencies, the evaluation function is highly centralized or within a large program area; in other agencies, the evaluation function is scattered in small offices throughout the agency.

No single best practice exists for organizing evaluation offices and functions. All of the arrangements described above have emerged in response to such factors as substantive area, kind of agency, or type of evaluation focus. They may or may not be the most effective models for current circumstances. Whatever model is chosen, the evaluation office must include the functions and possess the attributes described above under general principles.

Based on the general principles discussed in the previous section, we propose that agencies in the Executive Branch establish one of the following organizational frameworks to support evaluation.

Option 1: Evaluation Centers. Agencies could establish one or more evaluation centers to promote evaluation capacity and provide stable organizational frameworks for planning, conducting evaluation, or procuring evaluation advice or studies from outside organizations. Every program in the agency should be assigned to one of the centers for program evaluation.

The heads of these evaluation centers should report directly to the senior executive of their center's organizational component. Each of these centers would:

- Have a stable budget with sufficient funds to plan and carry out an appropriate level of program evaluation over several years
- Issue policies and procedures to guide its evaluation work, including guidance on appropriate methods for conducting formative and summative evaluations, as well as developmental evaluations to improve evaluative capabilities within agencies.
- Strategically plan a body of evaluation work for the agency and each agency component for which it has evaluation responsibility
- Consult with agency program and budget offices and, in concert with the agency's legislative liaison office, with Congress in developing evaluation plans
- Hire professional evaluators or engage consultants or contractors with the diverse skills necessary to plan and execute (or procure) independent evaluation studies
- Publish the results of evaluations related to public accountability of the programs within their jurisdictions
- Share information about effective programs and evaluation methods with other government agencies
- Promote and facilitate the ongoing training and professional development of the center's evaluators

Option 2: Evaluation Coordinators. Agencies that choose to distribute their evaluation offices, associating small evaluation offices with individual programs or small groups of programs, should promote evaluation capacity and performance by appointing one or more senior officials to:

- Advise the agency head or senior officials on matters pertaining to evaluation
- Ensure that each program or program group in the agency has a current annual evaluation plan.
- Promote, facilitate, and coordinate the development of evaluation plans for programmatic issues that cut across agency lines
- Facilitate the preparation of evaluation budgets
- Establish appropriate standards, frameworks, and procedures for evaluation activities in the agency
- Facilitate the development and efficient and effective production of evaluation plans, designs, instruments, and reports by government agency staff or outside evaluators

- Facilitate the dissemination of evaluation reports related to public accountability
- Share with other agency components information about effective programs and evaluation methods
- Promote and facilitate the ongoing training and professional development of evaluators in the agency

Option 3: Combined Approach. Federal agencies may find it advantageous to use Option 1 and Option 2—evaluation centers for large programs, program groupings, and overall evaluation support, and evaluation coordinators for distributed evaluation offices—to ensure the viability of the evaluation function.

Congress’s Role

The GAO, Congressional Research Service, Congressional Budget Office, National Academies, and temporary commissions carry out evaluation and analysis for Congress, usually in fulfillment of their oversight role. Congressional committees or subcommittees also conduct some evaluative studies or investigations.

We do not propose to change these organizational structures. Instead, we offer recommendations to strengthen the connection between evaluations and the laws that Congress passes. This can be done by building evaluation expectations into authorizing legislation and explicitly setting aside adequate resources for evaluation.

Authorizing legislation. Program authorization and periodic reauthorization provide opportunities for Congress to establish frameworks for systematic evaluation of new and continuing programs. Congressional committees can, through authorizing legislation, provide guidance on or stipulate such activities and products as:

- Early implementation reviews to identify start-up problems in such areas as scheduling, contracting, and grant making and to correct them before they become more serious
- Requirements for developing evaluation plans
- Evaluation of promising approaches to share among program implementers
- Development of performance indicators and the means to collect meaningful data on them once the program starts
- Studies reviewing the efficiency of federal program management as well as the fidelity of program implementation to the congressional mandate that instituted the program
- Studies assessing program effects and identifying why programs are or are not effective
- Evaluations of topics of interest to Congress and reports on the results to Congress in support of its oversight and appropriations functions and to inform future reauthorizations

- Establishment, expansion, or amendment of ongoing surveys or other data-collection mechanisms to become permanent sources of reliable data
- Establishment of evaluation centers or evaluation coordinators, as described above
- Funding for evaluation activities

Collaboration Between Executive Branch and Congress

The utility of evaluation results may be maximized if Congress and the Executive Branch jointly specify broad evaluation expectations and concerns in authorizing statutes and appropriations. We recognize that such collaboration will not always be easy or even possible to achieve. Nevertheless, experience suggests that, when possible, a partnership of this kind can help increase the benefits that evaluation provides.

Looking to the Future

The U. S. government faces major challenges in the years to come, as well as significant opportunities to improve lives, protect the planet, and create efficiencies. With more thoughtful and more systematic integration of evaluation into the responsible planning, management and oversight of programs and the application of evaluation results to planning and decision making, the performance of today's programs can be improved. Institutionalizing evaluation can also help achieve a more accountable, open, and democratic system of governance for future generations.

EPTF PERFORMANCE SUMMARY APPENDIX D: EVALUATION ROADMAP CITATIONS AND USE RECORD

AEA published its most recent version of *An Evaluation Roadmap for a More Effective Government* in September 2010. AEA and the Evaluation Policy Task Force have used it in commenting on emerging national policies such as maternal, infant, and children's home-visiting programs, health care reform, and foreign assistance. A broader goal was to use it as AEA's "place at the table" when evaluation policies were being considered or would be appropriate, even if an AEA representative couldn't be there in person. The latter goal is now being realized.

The Roadmap has been cited in congressional testimony, Government Accountability Office (GAO) reports, and other settings. Below are examples we know about.

SEPTEMBER 2011: Washington Evaluators/United States Government Accountability Office - *AEA's Evaluation Roadmap and its Utility for Improving Agency Evaluation [Brown Bag Session]*

"Stephanie Shipman, a member of the AEA Evaluation Policy Task Force will discuss key components of the Roadmap. Martin de Alteriis will discuss ways that the GAO has used the Roadmap to assess agency evaluation efforts. He will illustrate how the Roadmap can be used to improve agency evaluation efforts through three examples." This brown bag luncheon was sponsored by the Washington Evaluators AEA Affiliate while the speakers came from GAO.

[../we.example.eptf.htm](http://we.example.eptf.htm)

MAY 2011: United States Government Accountability Office – *International School Feeding: USDA's Oversight of the McGovern-Dole Food for Education Program Needs Improvement*

"The American Evaluation Association's *An Evaluation Roadmap for a More Effective Government* recommends that agencies develop policies and procedures to guide evaluation and assess the strengths and weaknesses of programs to improve their effectiveness." [from summary] Goes on to discuss USDA's efforts in this area. The *Roadmap* is referenced in the text, summary, and opening page of findings. "Finally, we compared USDA's oversight and internal control practices to our Standards for Internal Control in the Federal Government and the American Evaluation Association's *An Evaluation Roadmap for a More Effective Government*." [report page 2]

<http://www.gao.gov/products/GAO-11-544>

APRIL 2011: The Economist - Economist Intelligence Unit – *Creating Value in the Public Sector: Intelligent Project Selection in the US Federal Government*

"Programme evaluation would profit from an influx of trained and dedicated people. Most of the agency planning offices that carry out the project evaluation and analysis now rely on a very small workforce, which will be stretched even further with the demands from the OMB, the new GPRA and Congress. The American Evaluation Association made the same point in its *Evaluation Roadmap*. The units formed by agencies to conduct evaluations are too often under-resourced. Training and capacity building for evaluation have been inconsistent across agencies 'and, in many cases, insufficient to achieve the needed evaluation capacity and to sustain it over time', the association reports." [page 12]

http://www.businessresearch.eiu.com/sites/businessresearch.eiu.com/files/downloads/Oracle_PublicSector_WEB.pdf

MARCH 2011: United States Government Accountability Office – *Employment and Training Administration: More Actions Needed to Improve Transparency and Accountability of Its Research Program*

"While there is no single or ideal way for government agencies to conduct research, several leading national organizations have developed guidelines that identify key elements that promote a sound research program. These guidelines identify five elements as key: agency resources, professional competence, independence, evaluation policies and procedures, and evaluation plans." [page 6] Cites *Roadmap* in accompanying footnote.

<http://www.gao.gov/new.items/d11285.pdf>

FEBRUARY 2011: Grantmakers for Effective Organizations – *Scaling What Works - Briefing Paper – What Do We Mean by Scale?*

"Choosing from among the extensive set of analytic approaches and methods developed by the evaluation field over many years of practice— including logic models, case studies, surveys, quasiexperimental designs and needs assessments— depends on what is being scaled, its implementation stage, how results will be used and the kinds of decisions an evaluation is meant to facilitate." [page 5] *Roadmap* cited as basis for this quote.

<http://www.geofunders.org/document.aspx?oid=a0660000005uSOAAA2>

JANUARY 2011: United States Government Accountability Office – *Program Evaluation: Experienced Agencies Follow a Similar Model for Prioritizing Research*

"These [Department of Education] plans align well with the American Evaluation Association's (AEA) recommendation, made in a recent policy paper on federal government evaluation, that federal agencies prepare annual and multiyear evaluation plans to guide program decision-making and consult with the Congress and nonfederal stakeholders in defining program and policy objectives, critical operations, and definitions of success." [page 15] Cites *Roadmap* in accompanying footnote.

<http://www.gao.gov/new.items/d11176.pdf>

JANUARY 2011: IBM Center for the Business of Government – Jonathan Breul in a blog post on *Producing an Effective Program Evaluation Agenda*

Highlights GAO's Program Evaluation: Experienced Agencies report, adds as a final aside "In addition, you might check out the American Evaluation Association's *Evaluation Roadmap for a More Effective Government*."

<http://www.businessofgovernment.org/blog/business-government/producing-effective-program-evaluation-agenda>

DECEMBER 2010: NASA – *NASA Informal Education: Final Report*

"The call for improved evaluation practices is strongly conveyed in the American Evaluation Association's (AEA) paper, *An Evaluation Roadmap for a More Effective Government* (2009). In it, the AEA Evaluation Policy Task Force outlines a framework and set of guiding principles for each federal agency to inform the development of its own evaluation program. Moreover, it advocates that evaluation be used to inform a variety of decisions, all of which are particularly relevant to NASA's

Informal Education Program...Although our discussions with NASA OE personnel suggest that their vision of the role that evaluation would serve is in sync with the purposes outlined by the AEA *Roadmap*, it is clear this vision is not being implemented in a way that allows for effective practice." [page 50]

http://www.nasa.gov/pdf/511273main_2010_InformalEvaluation.pdf

SEPTEMBER 2010: Brookings Institution Center on Children and Families and the National Institute for Early Education Research - *Investing in Young Children: New Directions in Federal Preschool and Early Childhood Policy*, Chapter on *Strengthening Home-Visiting Intervention Policy: Expanding Reach, Building Knowledge*

"Defining the evidentiary base necessary for estimating the potential impacts of a given intervention is complex and particularly challenging when the reform involves multiple strategies. Randomized control trials are often the best and most reliable method for determining whether changes observed in program participants over time are due to the intervention rather than to other factors. Maximizing the utility of program evaluation efforts, however, requires more than just randomized clinical trials. As noted by the American Evaluation Association in a February 2009 memo to Peter Orszag, the Director of the Office of Management and Budget: "There are no simple answers to questions about how well programs work and there is no single analytic approach or method that can decipher the complexities that are inherent within the program environment and assess the ultimate value of public programs." [page 85]

http://www.brookings.edu/reports/2010/1013_investing_in_young_children_haskins.aspx

MARCH 2010: United States Government Accountability Office – *Afghanistan Drug Control: Strategy Evolving and Progress Reported, but Interim Performance Targets and Evaluation of Justice Reform Efforts Needed*

"To assess U.S. monitoring and evaluation of counternarcotics programs, we first utilized, as a framework, the Government Performance and Results Act, which outlines good management practices such as establishing strategic, long-term goals and planning and reporting progress toward these goals on an annual basis. We also referenced good management practices outlined within previous GAO reports and considered monitoring and evaluation principles established by the American Evaluation Association." [page 42] Cites *Roadmap* in accompanying footnote.

<http://www.gao.gov/new.items/d10291.pdf>

SEPTEMBER 2009: United States Government Accountability Office – *International Food Assistance: USAID Is Taking Actions to Improve Monitoring and Evaluation of Nonemergency Food Aid, but Weaknesses in Planning Could Impede Efforts*

Roadmap is cited explicitly throughout as a core framing document for the GAO's review including providing principles for evaluation policies and procedures [page 26], for independence [page 27], for scope and coverage [page 28], for dissemination of results and for professional competence [page 30], for resources [page 33], for evaluation plans [page 34].

<http://www.gao.gov/new.items/d09980.pdf>

JUNE 2009: US House of Representatives – *Hearing Before the US House of Representatives Ways and Means Committee Subcommittee on Income Security and Family Support on Proposals to Provide Federal Funding for Early Childhood Home Visitation Programs*

“We considered established principles for monitoring and evaluation—especially the “Roadmap” principles of the American Evaluation Association (AEA), which have been developed to integrate monitoring and evaluation with program management...” [page 4]

<http://www.gpo.gov/fdsys/pkg/CHRG-111hrg52502/pdf/CHRG-111hrg52502.pdf>

2009: Wellcome Trust Project – *Developing Metrics and Measures for Dual-Use Education*

"It is apparent that metrics, measures and evaluation have become increasingly salient in the US where there is evidence of the emergence of a shift towards revising and enhancing the process of evaluation and assessment of government projects. The American Evaluation Association’s (AEA) Evaluation Policy Task Force is worth quoting at length in this regard and in the 2009 report, *An Evaluation Roadmap for a More Effective Government* the AEA state: [extended Roadmap quote on the effectiveness of federal programs]" [page 8]

http://www.brad.ac.uk/bioethics/media/SSIS/Bioethics/docs/Education_metrics_and_measures_2nd_DRAFT.pdf



March 7, 2008

Robert Shea
Associate Director for Administration and Government Performance
Office of Management and Budget
Eisenhower Executive Office Building
Washington, DC

Dear Mr. Shea,

I am pleased to submit, in response to your request on February 26, 2008, the comments of the Evaluation Policy Task Force of the American Evaluation Association (AEA) on the evaluation guidelines represented in the document entitled "What Constitutes Strong Evidence of a Program's Effectiveness?"

We support efforts to implement systematic processes for evaluating federal programs and we see the OMB Program Assessment Rating Tool (PART) as an important step in that direction. With appropriate revisions to the existing evaluation guidelines, the PART could be even more useful, reliable, and sustainable. PART is well constructed to encourage a comprehensive evaluation approach for the diversity of federal programs.

Our primary overall recommendations are:

- **Develop new guidance around evaluation in PART Questions 2.6 and 4.6:** Clarifying and expanding the evaluation guidance for PART Questions 2.6 and 4.5 would improve PART's legacy, sustainability, and validity. To provide practical guidance for OMB examiners and agency officials, such a document would include:
 - a) How evaluation information from other PART questions on program purpose and design, program strategic planning and measurement, and program management is needed and used as a foundation for designing high quality effectiveness evaluation.
 - b) How to identify and select evaluation methods for assessing program effectiveness that are appropriate to the needs and development level of a program.
- **Present Randomized Control Trials (RCTs) in context:** More balanced and considered presentation of the role of RCTs in assessing the effectiveness of federal programs would significantly improve the document.
 - a) There are important alternatives to RCTs for assessing effectiveness that should be included and may have distinct advantages under specific circumstances.

- b) Situating RCTs, the focus of the current document, within a balanced presentation of the spectrum of appropriate, rigorous, evaluation methods will improve the likelihood of selecting appropriate measures and methods to assess and improve program performance.
- c) The utility of RCTs can be significantly enhanced when mixed with additional methods that enable identification of why and how observed effects occur.
- **Draw on a range of expertise:** Calling upon experts who could provide relevant assistance during the development of the evaluation guidance would enhance the knowledge-base of the drafters and improve the rigor and feasibility of the guidance. We would be interested in either working with you directly on revisions of the guidance or in working with you and your staff to identify appropriate evaluation expertise.

Given the short timeframe, we have not been able to include detailed citations and references to the literature. Should such additional be beneficial, please do not hesitate to contact us.

We hope that our comments are helpful to you and your staff at OMB and we look forward to working with you to improve the PART system so that it is both effective and sustainable as a management mechanism in the federal government.

Sincerely,



William M.K. Trochim
President, American Evaluation Association
Chair, AEA Evaluation Policy Task Force



Attachments:

[Comments on "What Constitutes Strong Evidence of a Program's Effectiveness?"](#)

EPTF PERFORMANCE SUMMARY APPENDIX F: COMMENTS ON WHAT CONSTITUTES STRONG EVIDENCE OF A PROGRAM'S EFFECTIVENESS? MARCH 7, 2008

**Comments on
What Constitutes Strong Evidence of a Program's Effectiveness?
Evaluation Policy Task Force
American Evaluation Association
March 7, 2008**

The Evaluation Policy Task Force supports the general idea behind the PART process, and particularly its recognition of the important role that program evaluation can play in improving program performance and accountability. We recognize the hard work that has gone into developing the guidance to date and commend OMB for building on their investment by continually improving the PART program in critically important ways.

The PART process presents an opportunity to promote high-quality program evaluation methodologies designed to assess and improve federal government programs. The paper at issue is useful in that it brings together considerable information on randomized experimental designs or Randomized Clinical Trials (RCTs). However, it has serious limitations that threaten the viability and sustainability of the PART process. We provide these comments in the hope that OMB will consider significant revision and enhancement of this statement, and are confident that addressing these concerns would strengthen the PART process and lead to improved federal program performance.

This paper misses a key opportunity to help federal programs align their most pressing evaluation questions with the most rigorous, practicable evaluation methodology. While the paper notes (p. 1) that "RCTs are not suitable for every program and generally can be employed only under very special circumstances," and thus protects itself from criticisms of advocating only one methodology, it subsequently neglects consideration of how agencies would determine where, when, and why other methods might be used. Instead, the paper focuses almost exclusively on RCTs.

We wish to preface this discussion with a general statement: *there is broad-based consensus in the field of evaluation that RCTs are an important methodological tool when applied under the correct circumstances.* However, if we are to apply RCTs correctly, we must have a rigorous understanding of what those circumstances are and when they are likely to be present.

We offer here some general comments followed by more specific comments organized by paper section.

General Comments

- **The paper needs to be more comprehensive and balanced.** The title of the paper ("What Constitutes Strong Evidence of a Program's Effectiveness?") leads the reader to expect a broad exploration of evidentiary strength and consideration of major relevant methods for achieving that strength. Yet internal validity is the only aspect of evidentiary strength that is seriously addressed and the RCT is the only method seriously examined. Other methods are merely mentioned under Section II, or are used as foils to compare them unfavorably with RCTs.

We believe that the guidance needs to accomplish two objectives that are essential for OMB examiners and agency budget analysts: (1) explain what types of evaluations and methods are, and are not, acceptable and eligible to receive credit from OMB under PART questions 2.6 and 4.5; and (2) provide a framework for how to tell whether an evaluation submitted under the PART is technically sound and addresses the most important questions related to program performance. The document currently does not sufficiently achieve either of these objectives.

- **RCTs are weak with respect to the goal of program improvement.** RCTs have their greatest value in confirming or disconfirming specific hypotheses about whether or not a program is effective. RCTs are not very helpful in explaining *why* things happen and, except in the crudest sense (did the program work or not), do not contribute information for program improvement.

Furthermore, RCTs do not provide information about *how* the program did or did not affect the outcomes in question. There is an important literature in evaluation on the rigorous modeling and study of program theory that was expressly developed to address this deficiency in RCTs. However, the paper does not mention this literature or suggest that studies of why and how the program affected the outcomes are essential in providing evidence of a program's effectiveness.

- **RCTs do not by themselves explicitly address construct validity.** In rigorous program effectiveness evaluation it is essential to assure that the program is implemented with high fidelity and that measures are reliable and valid. In technical terms, one needs to demonstrate the construct validity of the cause (i.e., the program) and the effect (i.e., the outcome measures). This is an empirical endeavor that should be carried out prior to mounting an RCT. The paper does not discuss the issue of construct validity in RCTs.
- **RCTs are weak with respect to generalizability or external validity.** Historically, one of the major criticisms of RCTs is that they are relatively weak with respect to external validity or generalizability. This critically important issue is not acknowledged in the paper, and, on the contrary, the reader is left with the impression that no such weakness exists. For example, under Section II, the paper states, "the difference in outcomes between the groups demonstrates the 'outcome' or impact one would expect for the intervention more generally" (p. 2). This is misleading. RCTs, if well done, assess outcome or impact at one point in time, in a particular setting or settings, at the current economic juncture, with the types of

people who participated in the study, but surely not generally. In the Appendix, the guidance (under Quantity of Evidence Needed) calls for measures (multiple "typical" sites, etc.) that target some aspects of external validity but avoids addressing them as such.

The point here is not just that RCTs are strong with respect to internal validity but relatively weak for external. It is also that issues of external validity must be addressed because (a) using evaluation as an input to policy often requires generalizable information; and (b) the costs involved in dealing seriously with issues of generalizability might well invalidate what is written in Section IV, C, on the unlikelihood of prohibitive costs.

- **Addressing RCTs' validity problems often entails investment in companion program evaluations that have methodological designs other than RCTs.** This idea is well understood in biomedical research (where RCTs originated). In the clinical trials model in medicine, it is assumed that a potential intervention has already been through significant basic research that justifies the plausibility of the clinical intervention. Clinical trials involve four basic phases.
 - Phase I trials are exploratory small sample studies that examine tolerance of the treatment and potential side effects.
 - Phase II trials typically demonstrate whether the program is capable of achieving effects (efficacy) and examines its correlates.
 - Phase III trials are typically controlled effectiveness studies.
 - Phase IV trials typically examine generalizability and the fidelity of transfer of controlled interventions to field settings.
- Randomized designs are usually not used until late in phase II or more likely in phase III studies when effectiveness is the focus. The FDA reports that the vast majority of interventions (approximately 70-75%) that begin clinical trials do not survive to controlled Phase III randomized trials because they do not meet the basic conditions that warrant subsequent efforts. This critically important structure of clinical trials in medicine is not mentioned in the paper, despite the fact that the paper correctly cites this tradition as exemplary with respect to use of RCTs. OMB PART is already structured well to emulate this multi-phased model that situates RCTS as an essential component at the appropriate phase. By attending more carefully to the need for rigorous evaluation in connection with program development, measurement, and implementation as implied in PART Questions 1-3, the OMB would significantly enhance the quality, rigor, and potentially cost efficiency of effectiveness evaluations undertaken in connection with Question 4.
- **The importance of mixed methods.** There is a general consensus regarding the value of mixed methods in contemporary evaluation based on the recognition that all methods, however good, are fallible and have unique biases and weaknesses. The use of multiple methodologies, especially those that are less likely to share related biases, is recognized as one of the strongest mechanisms for identifying and controlling such biases. In addition, evaluation,

unlike traditional basic science, is typically required to address multiple purposes, rather than simply the generation of new knowledge. Evaluations need to provide information about what happened and why, and to suggest ways that programs might be improved. RCTs alone are a relatively limited mechanism for some of these purposes. However, when RCTs are appropriate, they often can also be augmented with other methods, either sequentially or synchronously, that can significantly help broaden both their rigor and utility. For example, high quality RCTs conducted for the purpose of program improvement should, in order to be considered rigorous for this purpose, be required to incorporate other methods such as interviewing, participant observation or theory-driven modeling in order to identify how and why the observed effects occurred.

- **The need to address feasibility and resource issues realistically.** In this guidance it is critically important to explicitly acknowledge feasibility considerations (including, but not limited to, cost and time) and provide agencies the flexibility for choosing to conduct evaluations that best leverage limited evaluation resources.
- **The need to address equity and human subjects concerns realistically.** Virtually no mention is made of the requirement of the RCT that people be denied a potentially beneficial intervention for the sake of the experimental design. While this position may be defensible ethically, and is often so defended in medical clinical trials, it is one of the most common concerns raised regarding RCTs and deserves mention in this context. It is important to note that the calculus for weighing the costs and risks to individuals against the potential gain to society from high-quality evaluation may differ from field to field. For instance, the U.S. Congress recently passed legislation to halt an RCT of the Upward Bound program apparently because they did not believe that denying access to that program for the sake of the controlled test of its effectiveness was a reasonable political or policy trade-off.

Specific Comments

Introduction

The paper begins by stating that “The revised PART guidance this year underscores the need for agencies to think about the most appropriate type of evaluation to demonstrate the effectiveness of their programs” (p. 1). One would expect that the document would review some of the key terms or phrases. For instance, what is an “appropriate” type of evaluation? What is meant by evidence of effectiveness? Perhaps more important, one expects that there would be a presentation of the variety of methods that provide useful information about program effectiveness, with some consideration of the strengths and weaknesses of each. Instead, the paper immediately “points to the randomized controlled trial (RCT) as an example of the best type of evaluation to demonstrate actual program impact” (p. 1). The author then inserts several caveat sentences:

“Yet, RCTs are not suitable for every program and generally can be employed only under very specific circumstances. Therefore, agencies often will need to consider alternative evaluation methodologies. In addition, even where it is not possible to demonstrate impact, use of evaluation to assist in the management of programs is extremely important.” (p. 1)

Several approaches are mentioned in passing (Experimental or Randomized Controlled Trials, Direct Controlled Trials, Quasi-experimental Studies, and Non-Experimental Studies (Direct or Indirect)), to be considered later.

The paper lists several questions that will be covered. But critically important questions that should be included are not raised. For instance, it makes sense that agency staff would want to know about things like:

- When is the right time to evaluate a program in terms of effectiveness? How frequently should the program be evaluated?
- What types of programs should be evaluated with the program evaluation methodology as described in this paper?
- How does this evaluation guidance apply to different types of programs like:
 - Intervention programs
 - R&D programs
 - Service programs
- How should program evaluation be adapted to large complex program systems that involve different types of subprograms or program components?
- How should agency staff make the decision about what type of evaluation is most appropriate for their program?
- How should resource issues be addressed?
- How should an agency develop the capacity to implement or manage evaluations of program effectiveness?

It seems that much more detail on these types of questions would enable OMB examiners and agency staff to make a more informed judgment about how to decide what is appropriate evaluation for purposes of PART and how to help ensure that rigorous high-quality evaluation will be carried out.

I. How is program evaluation addressed in PART?

The Role of Evaluation in Program Development and Implementation

This section points out that PART Questions 2.6 and 4.5 are relevant to evaluation and provides the full text for those questions. While the paper mentions that many of the questions in Section III of PART are relevant to evaluation, it does not consider them further. However, the paper does not recognize at all that evaluation plays a critically important role in Section I: Program Purpose and Design.

Evaluation is essential in addressing such questions. For instance, if a program is not well designed and its implementation does not reflect its intent, the program does not have construct validity, an essential precondition for assessing the program’s effectiveness. One approach to address such

questions is evaluability assessment, a method developed in response to evaluation problems related to program design, and one that has been used in the federal government for quite awhile.

The Role of Program Design and Implementation Evaluation in Impact Evaluation

Section II of PART on Strategic Planning is also integrally dependent on evaluation but receives no mention in this paper. Questions 2.1 – 2.4 relate to the specification of goals and measures, both foundational to conducting an evaluation. It is essential that program goals are well articulated and that measures are of high quality (reliability and validity) and there are well-established evaluation procedures for accomplishing these tasks. Every program effectiveness design, including RCTs, makes the assumption that these steps have been done well prior to undertaking the effectiveness study, and the quality of effectiveness studies depends on this.

Section III of PART on Program Management is at the heart of evaluation and involves the measurement and monitoring of program performance. It makes little sense to undertake a high quality assessment of program effectiveness (Section IV) without assuring first that the program is functioning well. In fact, one might even argue that it would be irresponsible both fiscally and methodologically to initiate an effectiveness study for programs that cannot demonstrate that they have well defined models, goals, measures and procedures for managing implementation. Yet, no mention is made in this paper that a precondition of doing the effectiveness evaluations of Section IV (RCT or otherwise) should be prior evidence of good performance on these foundational areas.

The Role of Internal and External Stakeholders in Evaluation

Question 2.6, which is a focus of this paper states:

2.6: Are independent evaluations of sufficient scope and quality conducted on a regular basis or as needed to support program improvements and evaluate effectiveness and relevance to the problem, interest, or need?

There is a fundamental problem that is raised by the use of the term “independent” in this question. Independent is defined in the guidance as:

To be independent, non-biased parties with no conflict of interest should conduct the evaluation. Evaluations conducted by the program itself should generally not be considered “independent;” (p. 31)

The primary problem is that this does not acknowledge the critically important role of all stakeholders in developing the evaluation. For instance, many of the questions in Sections I-III of the PART relate to the articulation of the model, goals and measures of the program and of its management when being implemented. High-quality evaluation practice and common sense require that the “internal” program staff be an integral part of these evaluation activities. In many organizations, program development and planning are conducted largely internally and involve a number of evaluation activities. Rigorous high-quality assessment of effectiveness depends on their being done well using the best evaluation methodologies. But Question 2.6 mentions only “independent” or external evaluation. While external evaluators may play a role in helping to facilitate the articulation of program models, goals and

measures and in helping set up systems for monitoring implementation, it is inconceivable that these tasks could be done well only through independent external evaluators. Yet these activities are essential for addressing the requirement of Question 2.6 that the evaluation address “support program improvements and evaluate effectiveness and relevance to the problem, interest, or need.” Some guidance for agency staff on these issues would be critically important.

Summary

The paper should describe the range of evaluation methods and approaches that can be used to address questions other than 2.6 and 4.5, should consider their relative strengths and weaknesses in different settings, and should discuss the fact that high quality and rigorous effectiveness evaluation is predicated on doing these tasks well. There are many well-established evaluation methods that would include (but not be limited to): needs assessment, program logic modeling, performance measurement, interviewing, focus groups, survey development and qualitative methods, measurement systems development, the assessment of the quality of measurement (reliability and construct validity), implementation assessment, process evaluation, and evaluability assessment, to name a few. Each of these approaches has a considerable well-developed literature and would be standard fare in even an introductory evaluation course. It is generally suggested in the evaluation field that high quality rigorous evaluation requires that the issues in Sections I-III of the PART guidance be accomplished well and that programs that have not addressed these issues cannot be considered ready for outcome or effectiveness evaluation.

II. What are the most common ways to evaluate program performance?

This section begins with the statement “The most significant aspect of program effectiveness is *impact*—the outcome of the program, which otherwise would not have occurred without the program intervention” (p. 2). In one sense, this is true. All program activities are ultimately directed towards this end. However, one cannot reach impact without assuring first that the pathways to impact have been successfully navigated. Impact is, from this perspective, the end of a long chain of events and is dependent on assuring that none of these links is weak. This chain is well established in the PART structure and would include the work done on defining program purpose and design, program strategic planning and measurement, and program management and implementation. The evaluation of program design and implementation discussed in the previous section are an essential foundation to evaluations focused on impacts and effectiveness. This is foundational work, much of which also involves evaluation (e.g., program modeling, needs assessment, measurement development, implementation and process evaluation, etc.). In fact, it is likely that the vast majority of day-to-day evaluation work for most federal programs will need to be directed to the foundational evaluation efforts that must be in place before an effectiveness evaluation can be justified or warranted. No mention is made of this in the paper. Instead, the paper mentions in passing that “Where it is feasible to measure the impact of the program...” and then moves on to recommending RCTs. But agency staff and OMB examiners have no guidance about how these foundational evaluation requirements need to be addressed in order for high-quality impact assessment to be feasible. For instance, how long does it typically take to achieve the conditions required for effectiveness evaluation?

Making Informed Choices about Design

The paper states unequivocally that “RCTs are generally the highest quality, unbiased evaluation to demonstrate the actual impact of the program” (p. 2). They then continue with

However, these studies are not suitable or feasible for every program, and a variety of evaluation methods may need to be considered because Federal programs vary so dramatically. Other types of evaluations may provide useful information about the impact of a program (but should be scrutinized given the increased possibility of an erroneous conclusion) or can help address *how* or *why* a program is effective (or ineffective) (i.e., meeting performance targets, achieving efficiency, fulfilling stated purpose). (p. 2)

However, they do not present any guidance for how one decides whether an RCT is the most appropriate method and present no detailed accounting of the strengths and weaknesses of the many alternatives. It would seem that some detailed explication of methods and perhaps something like a decision tree would help both OMB examiners and agency staff to make this complex decision.

The paper briefly presents five broad classes of methods. However, the level of detail and the breadth of coverage do not provide the information necessary for guidance of OMB examiners or agency personnel who have not been trained in these methods.

Specific Issues Related to “Classes of Methods”

First, the authors present the RCT. While they mention that there are feasibility challenges with RCTs (“There are many programs for which it would not be possible to conduct an RCT” (p. 2)), they do not consider here any of the other potential problems or biases associated with these designs. However, since the remainder of the paper is essentially devoted to RCTs, we will address these issues in subsequent sections.

The second of these methods – Direct Controlled Trials – seems to be inaccurately labeled. For example, both RCTs and many quasi-experimental designs would be considered “controlled trials” and it is not clear how the term “direct” describes the distinction the paper is trying to make. They seem to be referring to the traditional notion of an “experiment” in physical sciences research, a design which does not utilize random assignment because such a mechanism is not needed when the phenomenon being studied is “well-behaved” and expectations of the non-experimental condition have been well established in previous research. Historically it would probably be most accurate to label these “experimental designs” as distinct from *randomized* experiments where the mechanism of random assignment is required for control of extraneous variables. They present no assessment of the strengths or weaknesses of this class of design, but the history of science suggests that they predate the randomized experiment and have played the foundational role in scientific research. Thus, it is puzzling that the paper does not list them first (since the ordering they choose seems to be from stronger to weaker in internal validity) or discuss the circumstances in which such designs might appropriately be selected for PART evaluation.

The presentation of the third type of method, the quasi-experiment, is incomplete and somewhat confusing. It neglects any mention of the prominent types of quasi-experiments (regression-discontinuity, nonequivalent groups, interrupted times series), several of which are strong alternatives to RCTs when internal validity is the priority and have several distinct advantages over RCTs in certain circumstances. The paper describes quasi-experiments as “comparison group studies”, but this is a term that applies equally well to RCTs.

The fourth class of methods is “non-experimental direct analysis” and encompasses a wide variety of approaches including pre-post and longitudinal studies and “correlation analyses, surveys, questionnaires, participant observation studies, implementation studies, peer reviews, and case studies.” This is a huge area of methodology that the paper immediately dismisses with the statement “These evaluations often lack rigor and may lead to false conclusions if used to measure program effectiveness, and therefore, should be used in limited situations and only when necessary.” We find this dismissal inappropriate because it does not acknowledge that for each of these methods, there is an extensive literature on how to assure that they are accomplished with high quality and rigor. For example, the method of peer review is considered foundational to rigor and quality in scientific research, even in the assessment of the quality of RCTs in biomedical and health research. While peer review has its weaknesses, it remains a central method in the evaluation of scientific research projects. Similarly, the area of longitudinal research is well established both qualitatively and quantitatively and is capable of yielding causal inferences that rival the internal validity of RCTs. Yet the paper dismisses this whole area out of hand.

Finally, the paper presents non-experimental indirect analysis, described primarily as expert review. They almost rule out this approach entirely except when no alternative is available. Again, this fails to recognize: (a) that this approach can, when done well, be carried out with rigor and with the systematic use of empirical data; and (b) is a dominant approach in how the federal government and many other entities evaluate complex problems. For example, the hearings on the Challenger Space Shuttle disaster constituted a systematic expert panel review that included expert testimony and the consideration of data and was successful in identifying the cause of the disaster from an extremely complex set of variables. Many federal entities like the Institute of Medicine, the National Academies of Sciences, the National Research Council and the Government Accountability Office use such methods effectively to assess program effectiveness. Even in the area of clinical medicine that is often cited as exemplary by advocates of RCTs, the expert review panel increasingly constitutes a critically important evaluative mechanism that supplements for weaknesses in RCTs and meta-analysis, especially when a complex evidence base is being assessed. It is reasonable to expect that for many federal programs such approaches, if accomplished well, would provide an excellent and cost-effective mechanism for assessing program effectiveness and it is puzzling why it was simply ruled out in this context.

Value of Broad Consultation, Including for RCTs

The paper correctly advocates that agency staff should “consult with internal or external program evaluation experts, as appropriate, and OMB to identify other suitable evaluation methodologies to demonstrate a program’s impact. Some sources of evaluation expertise may include the peer-reviewed literature for the relevant discipline, scientific organizations such as the National Academy of Sciences,

think tanks, and research organizations” (p. 3). We agree that such consultation is an important factor in building the capacity of federal agencies to accomplish high quality effectiveness evaluation. However, the paper prefaces that statement with the phrase “When it is not possible to use RCTs to evaluate program impact...” suggesting that such consultation would not be warranted for RCTs. We disagree. The construction and implementation of high quality RCTs is a complex endeavor and federal agencies should be encouraged to seek outside professional assistance on such matters. It would be extremely helpful for agency staff if the paper provided clearer and more detailed guidelines on how appropriate consultants might best be identified for evaluation consultation.

Utility of References; the Need for More

The brief set of references that are provided to “assist in the decision of what type of evaluation will provide the most rigorous evidence appropriate and feasible, the PART guidance provides several links to references on program evaluation” are a good start but need to be annotated and augmented with other high quality citations. Many of these sources would be confusing to agency staff without training in evaluation or further orientation regarding their appropriateness and use under different circumstances.

III. What sorts of tests provide strong evidence of a program’s effectiveness?

This section is essentially a consideration of how RCTs provide strong evidence of a program’s effect. We agree that under certain circumstances RCTs indeed provide strong evidence of effectiveness. However, there is consensus in the field that in other circumstances other methods are more feasible or appropriate than RCTs and may also provide strong evidence of effectiveness. In several parts of this section, it is clear that the comments about the advantages are especially delimited to or focused on social programs rather than programs in general. It is important to recognize that the federal government addresses a wide array of programs other than social ones and that the comments in this section should be qualified appropriately for agencies in these other areas.

Conditions Under Which RCTs May Be Appropriate

A key statement in this paper and one that has led to considerable confusion is the following:

Well-designed and implemented RCTs are considered the gold standard for evaluating an intervention’s effectiveness across many diverse fields of human inquiry, such as medicine, welfare and employment, psychology, and education. (p. 4).

We believe that the labeling of the RCT as a “gold standard” is inaccurate in that it does not precisely describe the conditions under which the RCT would even by its proponents be considered the “strongest” design.

Recommended pre-requisites to undertaking an RCT often include:

- the program is well defined and has an articulated program model
- the program has been implemented consistently and with high fidelity

- there are high-quality (e.g., valid and reliable) outcome measures
- the program as implemented is capable of producing change
- there is sufficient statistical power to accomplish the study with high quality
- the participants can be kept unaware of the group (intervention or control) to which they have been assigned
- the random assignment can be implemented and maintained
- drop-out rates do not occur or do not differ by group
- the interest is in confirming whether the implemented program caused the observed outcomes
- there is no well-established empirical baseline or standard against which the program could be compared
- there is the potential for preexisting differences between the groups that would obscure the treatment effect
- there is little interest in whether the program would be effective generally in other settings, with other persons or in other time periods
- ethical and human subject protections have been approved and are in place
- it is morally and socially acceptable to deny the intervention to participants in order to test effectiveness

Put in other terms, it might be accurate to say that RCTs are the “gold standard” only when conditions like the above have been met. When such conditions are not present, RCTs may be susceptible to bias or may be unfeasible. Some of these conditions are described in the last section of the paper (pps. 12-13) and in the Appendix. But their omission here is important to the question of “What sorts of tests provide strong evidence of a program’s effectiveness?” The “gold standard” terminology may misrepresent RCTs when presented without these critically important and limiting qualifiers.

Potential for Inappropriate Use of Evaluation Resources

We especially want to avoid the danger that well-intentioned OMB examiners might say to Agencies: “RCTs are the Gold Standard, therefore the RCT is all we are willing to accept.” The real danger exists that RCTs will be undertaken in circumstances where they are neither appropriate nor feasible. This problem will in the end reflect poorly on the PART process and runs the danger of wasting precious evaluation resources and reducing the possibility that more appropriate and rigorous methods will be used to improve federal programs. Finally, the “gold standard” language inappropriately sets the RCT “above” other methods. In fact, most alternative methods when implemented with quality and rigor could be considered “gold standards” under the appropriate circumstances. There are situations where a quasi-experimental design, a survey instrument or an expert peer review panel might legitimately be described as the “gold standard” for the work at hand. The more important issues for establishing standards of rigor and quality for program improvement are how to direct agency staff to the most appropriate evaluation method for the program and how to assure that the method is implemented with the highest quality and rigor feasible. The “gold standard” language does not enhance the task of improving the rigor of effectiveness evaluation.

Specific Issues

When describing the “unique advantage of random assignment” the paper states that “...assuming the trial is properly carried out (as described in the Appendix) – the resulting difference in outcomes between the intervention and control groups can confidently be attributed to the intervention and not to other factors” (p. 5). This is, of course, a big assumption. The design and implementation of experiments is a complex endeavor that requires considerable technical expertise that most federal agencies do not readily have. It cannot be assumed that without considerable effort and expertise this condition will be met.

The paper states “Properly designed, RCTs are the only method that can eliminate the risk of bias, which can adversely affect the results of the evaluation” (p. 5). This is simply incorrect. There are many types of biases in any evaluation project. RCTs are particularly designed to address one important type: selection bias, or the threat that prior differences between treated and control groups might affect outcomes.

In this section the paper argues that “‘single group pre-post’ study designs often produce erroneous results” (p. 5). While such designs are comparatively weak relative to RCTs with respect to internal validity, they can be strong designs where there is a well-established empirical database of longitudinal data that can act as an appropriate comparison standard. There are whole fields of study, most notably economics and epidemiology, where such approaches are frequently and legitimately used and RCTs would not be appropriate substitutes.

On the top of page 6 the paper discusses a few important design considerations for RCTs. But it ignores a host of other important design issues including (but not limited to): interaction effects; multiple treatment comparisons (e.g., in factorial designs); or controls for variability in data (e.g., ANCOVA or blocking). While these cannot be covered adequately in so introductory a paper, they are important in RCT design and omission of this level of complexity may create for agency staff the false impression that experimental design is more straightforward and feasible than may be the case in many applied contexts.

The paper describes investigations that underscore the limitations of comparison group studies relative to RCTs and state that “these investigations have shown that comparison-group studies in social policy (employment, training, welfare-to-work, education) often produce inaccurate estimates of an intervention’s effects, because of unobservable differences between the intervention and comparison groups that differentially affect their outcomes” (p. 6). But the relatively few such investigations often compare well-designed RCTs with poorly designed alternatives (as is the case in the example shown in the figure on page 7). Furthermore, the paper goes on to state that “Even when statistical techniques have been used to adjust for observed differences between the two groups, problems have been found” (p. 6). But problems have been found in all major evaluation methodologies, including RCTs (consider the literature in clinical medicine that shows how well-designed clinical trials led to results that have subsequently been overturned). The literature on statistical adjustments for selection bias is extensive and is a major focus in a number of fields including economics where the Nobel Prize was awarded for work on such adjustments.

IV. The application of Randomized Controlled Trials: where they are / are not possible

In general we find this discussion both useful and accurate. It helps to provide some critical considerations that are essential to deciding on whether RCTs are feasible, and includes a useful list of examples of RCTs in a wide variety of federal program evaluations. The section on costs of conducting RCTs does a nice job of distinguishing between the relative advantages of large multi-site trials and smaller, more cost-effective ones, and provides some indication of the length of time that RCTs might require. Subsection D in this section and the appendix provide some important caveats and considerations that should be taken into account when deciding upon and implementing RCTs.

EPTF PERFORMANCE SUMMARY APPENDIX G: ORSZAG COVER LETTER, FEBRUARY 3, 2009



February 3, 2009

The Honorable Peter Orszag
Director, Office of Management and Budget
725 17th St. NW
Washington, DC 20503

Dear Mr. Orszag,

On behalf of the American Evaluation Association, we want to thank you for accepting, and wish you the very best in carrying out, your important duties in President Obama's administration, especially in these challenging times.

We are writing to propose for your consideration a major initiative to improve oversight and accountability of Federal programs by systematically embracing program evaluation as an essential function of government. In the attachment we describe how evaluation can be used to improve the effectiveness and efficiency of Federal programs, assess which programs are working and which are not, and provide critical information needed for making difficult decisions about them. We provide a roadmap for improving government through evaluation, outlining steps to strengthen the practice of evaluation throughout the life cycle of programs.

We understand how complex and demanding is the work before you. We hope our suggestions will be useful to you and we stand ready to assist you on matters of program evaluation.

Sincerely,

Handwritten signature of Debra Rog in black ink.

Debra Rog
President

Handwritten signature of William M Trochim in black ink.

William Trochim
Immediate Past President

Handwritten signature of Leslie Cooksy in black ink.

Leslie Cooksy
President Elect

Attachment: An Evaluation Roadmap for a More Effective Government

Copy to:
Dustin Brown, Deputy Assistant Director for Management
Daren Wong, Program Examiner

EPTF PERFORMANCE SUMMARY APPENDIX H: PROGRAM EVALUATION FOR MANAGEMENT AND RESULTS

The Obama administration faces a national debt in the trillions of dollars, annual deficits in the hundreds of billions, and uncertainties about financial institutions and the economy. At the same time, concerns remain about national security, health care, education, energy development, and many other facets of American life.

Program evaluation can make substantial contributions to address these issues. It uses systematic analysis to answer questions about how well a program or policy is working, whether it is achieving its objectives, and why. The new administration can use it to address new questions about current and emerging problems, stop wasteful spending, increase accountability, support major decisions about program reforms, and improve programs we need.

Use program evaluation throughout the program cycle

The key is to make evaluation integral to managing government programs at all stages, from initial development through start up, ongoing implementation, and reauthorization. Evaluation can be adapted to all part of the program and policy process and can help:

- Make sure that program and policy designs make sense and can plausibly achieve the goals that are set out for them
- Identify problems encountered during start up and correct them before they become permanent features of programs
- Identify and share promising approaches that evolve during early implementation
- Establish expectations and performance standards at program startup and monitor and refine them as the program matures
- Develop appropriate and efficient data collection systems that provide a continuing flow of evaluative information to policy makers and program managers
- Periodically examine selected program features to improve their effectiveness and efficiency
- Assess program outcomes and impacts
- Examine whether a successful program can be transferred to another setting before scaling up

Improve program evaluation infrastructure

Significant progress has already been made in measuring and managing program performance; but more can be done to:

- Invest in a body of evaluative work on individual public programs that can be used on an ongoing basis to guide management decisions about each program
- Include evaluation expectations in authorizing statutes and in executive agency initiatives and program plans
- Establish independent evaluation offices at appropriately high levels of government
- Set-aside sufficient funds to support professional evaluation activities
- Take advantage of evaluation as a formative tool for improving programs
- Expand the use of practical, appropriate models for assessing program impact

EPTF PERFORMANCE SUMMARY APPENDIX I: COMMENTS TO SENATE AND HOUSE, JUNE 2009



June 26, 2009

Dear Members of the Senate Finance Committee

On behalf of the American Evaluation Association's Evaluation Policy Task Force (EPTF), I am pleased to submit proposals for your consideration in drafting a bill to reform our nation's health care system. The proposals relate to evaluating and monitoring the reforms once they are enacted and are being implemented.

In summary, we believe that a sound foundation needs to be laid in the health care reform legislation itself for oversight and evaluation of all of its provisions. This is especially true for those provisions related to fundamental goals like insurance coverage, related insurance market reforms, access to care, health care disparities, quality of care, preventive health services, affordability, and cost control. Without such a foundation, information on unexpected successes and shortcomings, promising practices, and lessons of experience will not be available to identify needed adjustments in program administration, regulation, or future legislation.

We offer specific proposals to provide for such evaluation and oversight. They relate to national measures of success, key evaluation questions, responsibilities, annual evaluation plans, annual evaluation result reports, a Health Care Evaluation Advisory Group, and evaluation resources.

We hope these suggestions are helpful. Should you have any questions or wish to discuss this further, please do not hesitate to contact me at 607-255-0887 or EPTF Coordinator George Grob at 540-454-2888.

Sincerely,

A handwritten signature in black ink that reads "William M. Trochim".

William Trochim

Chair, AEA Evaluation Policy Task Force

**Proposals for the Senate Finance Committee on
Evaluation and Oversight
of
Health Care Reforms**

**Prepared by
The American Evaluation Association's Evaluation Policy Task Force
June 2009**



The American Evaluation Association is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness. AEA has over 5500 members representing all 50 states in the US as well as over 75 foreign countries.

Mission:

To improve evaluation practices and methods

Increase evaluation use

Promote evaluation as a profession and

Support the contribution of evaluation to the generation of theory and knowledge about effective human action.

Evaluation Policy Task Force

The goal of the American Evaluation Association’s Evaluation Policy Task Force is to promote evaluation policies that are critically important to the practice of evaluation.

The members of the Task Force are:

- William Trochim, Chair
- Eleanor Chelimsky
- Leslie Cooksy
- Katherine Dawes
- Patrick Grasso
- Susan Kistler
- Mel Mark
- Stephanie Shipman
- George Grob, Consultant

The term “evaluation policy” encompasses a wide range of potential topics that include (but are not limited to): when systematic evaluation gets employed, and on what programs, policies and practices; how evaluators are identified and selected; the relationship of evaluators to what is being evaluated; the timing, planning, budgeting and funding, contracting, implementation, methods and approaches, reporting, use and dissemination of evaluations; and, the relationship of evaluation policies to existing or prospective professional standards.

Proposals for the Senate Finance Committee on Evaluation and Oversight of Health Care Reforms

This paper provides advice regarding the establishment of evaluation and oversight provisions in the forthcoming Senate Finance Committee's draft bill to reform our nation's health care system. The comments are those of the American Evaluation Association's Evaluation Policy Task Force (EPTF). They are based on the principles outlined in "An Evaluation Roadmap for a More Effective Government," published by the EPTF in February 2009 (<http://www.eval.org/aea09.eptf.eval.roadmapF.pdf>).

In summary, we believe that a sound foundation needs to be laid in the health care reform legislation itself for oversight and evaluation of all of its provisions. This is especially true for those provisions related to fundamental goals like insurance coverage, related insurance market reforms, access to care, health care disparities, quality of care, preventive health services, affordability, and cost control. Without such a foundation, information on unexpected successes and shortcomings, promising practices, and lessons of experience will not be available to identify needed adjustments in program administration, regulation, or future legislation.

We offer specific proposals to provide for such evaluation and oversight. They relate to national measures of success, key evaluation questions, responsibilities, annual evaluation plans, annual evaluation result reports, a Health Care Evaluation Advisory Group, and evaluation resources.

Systematic Evaluation and Oversight

Utility and Importance. Based on its long history in the development of major health care legislation, including Medicare, Medicaid, and the Children's Health Insurance Program, the Committee members and staff are well aware of the need for systematic evaluation and oversight of the programs enacted through its efforts. Nevertheless, it is useful to briefly recall why such ongoing assessments are so important and to reflect briefly on how they can best be carried out.

Evaluation and oversight mechanisms will provide benefits almost immediately after legislation is enacted, during the early stages of implementation. That is the time when policy makers and program managers need to identify weaknesses that can be quickly corrected and promising practices that can be shared and used widely.

Within a few years' time, some fundamental questions about the results of the legislation will be inevitable. This will include questions on the achievement of the legislation's main goals: access to health care services for greater segments of the population; affordability of those services to patients; and quality of care for all Americans. More detailed questions will arise about the costs of the reforms, the possible effects of those reforms on health disparities among groups, the effects of preventive care on immediate and long-term costs as well as on health outcomes over time, and potential system overload or increased waiting-times, for example.

While many of these short and long term questions are predictable, other unforeseen issues will also arise regarding on-going implementation and management issues.

If the needed evaluation and oversight mechanisms are in place, such questions will be answerable in terms of valid and reliable data-based comparisons between what exists today (vis-a-vis specific measures of, say, access and affordability), and what will have been accomplished by the legislation over a given period. By answering important empirical questions, evaluation can advance learning about what works, as well as how and why and whether it is transferable, and also about what does not work, thus avoiding waste of resources.

Evaluation will also bring necessary transparency and accountability to efforts at reform in a field fraught with controversy.

Finally, the evaluation language itself, once incorporated into the bill, should, by its advocacy of public accountability, improve both the general perceptions and prospects of the legislation.

Scope and Methods of Evaluation. As implied by the above and as discussed more thoroughly in the EPTF “Evaluation Roadmap” mentioned previously, we take a broad view of evaluation, recognizing the need for a body of evaluative work that covers public programs and policies throughout their life cycle and for using evaluation as a tool to improve programs as well as assess their effectiveness and impact.

Critical evaluation questions change over the life of a program, and the evaluation approaches and methods must be adapted accordingly. A complete set of evaluation approaches and methods would include but not be limited to case studies, surveys, quasi-experimental designs, randomized field experiments, cost-benefit and cost-effectiveness analyses, needs assessments, early implementation reviews, logic models and evaluability assessments. It also encompasses the field of performance management.

Based on experience in the Medicare and Medicaid programs, safeguarding the integrity of the new reforms, protecting them from fraud, waste, and abuse, must be in the forefront of concerns of the Senate Finance Committee.

However, it is not sufficient to limit the oversight and monitoring efforts of the government to such integrity issues. It is just as important to learn, starting from the first days of implementation, whether the reforms are working as intended and ultimately whether they are effective in achieving their goals.

Responsibilities. The Federal agencies that implement and operate the new programs must have a hand, if not the lead, in the evaluation and oversight of the new reforms. Indeed, evaluation and oversight are essential functions of good government. Nevertheless, it would appear beneficial to also involve recognized experts in the evaluation field and traditional accountability and evaluation agencies such as the Institutes of Medicine, the Government Accountability Office, and the Inspectors General of relevant Federal agencies. Of course, the Congress is fundamentally involved through its oversight role, and the relevant committees can contribute significantly by identifying specific evaluation questions for which they will need answers.

Proposals

We recognize that within the very short timeframes during which the Senate Finance Committee must prepare a draft bill, it may be difficult to develop the kind of detail needed for a fully specified evaluation and oversight system. Therefore, in preparing our proposals, we considered what might be feasible to include in the bill under this compressed timeframe. Our tactic was to identify some overarching organizational principles and general steps that, if mandated in the resulting law, could evolve quickly into an adequate foundation for evaluation and oversight of the most important health care reforms. With this in mind, we recommend the following:

Annual National Report Card. Identify key recurring success measures for health care reform, such as percent of uninsured persons, health care cost indices, status of meeting modernization goals, and possibly key health status indicators, that should be reported to the Congress and broadly published annually; and identify who will prepare these reports.

Key Evaluation Questions. Identify the key evaluation questions that need to be posed. These could be organized along the lines of the key concepts underlying the bill, such as insurance coverage, access to care, affordability, quality of care, and cost control. Identify who will prepare evaluation studies to address these questions, when they are to be completed, and how they will be published or made public.

Federal Evaluation Units. Require each Federal agency with responsibilities for health care reforms covered under this reform bill to establish or designate one or more evaluation units to be responsible for evaluation of that agency's health care activities affected by the bill. (By "agency" we mean "department" or other separately identifiable unit of the Federal Government responsible for administering or regulating health care activities.)

Annual Evaluation Plans. Require each such agency to prepare annual evaluation plans related to programs or activities covered in the bill. Require such plans to be published on the agency's website and to be sent to the appropriate congressional oversight committees.

Annual Evaluation Result Reports. Require each such agency to prepare an annual report of evaluation accomplishments, summarizing evaluation results and explaining how the results from the current and previous years' reports were used.

Health Care Evaluation Advisory Group. Establish a Health Care Evaluation Advisory Group with responsibility for (1) preparing a biennial report on the state of the nation's health care system, submitting the report to the relevant congressional committees, and making it publicly available, (2) making recommendations for the consideration of the Congress and the President; (3) providing non-binding advice to Federal evaluation units identified in the above agencies about important health care evaluation issues; and (4) providing technical advice on methods for evaluation of health care issues.

Evaluation Resources. Authorize adequate funds to develop and sustain such evaluation and oversight activities.

We hope this analysis and these proposals are useful. We are available to assist the committee in the further development of these concepts.



June 26, 2009

Dear Members of the
House Committee on Education and Labor
House Committee on Energy and Commerce
House Committee on Ways and Means

On behalf of the American Evaluation Association's Evaluation Policy Task Force (EPTF), I am pleased to submit the attached comments on the House discussion draft of its bill to provide affordable, quality health care for all Americans and reduce the growth in health care spending, and for other purposes. Our comments focus on evaluation and oversight provisions of the bill.

In summary, various provisions in the bill will provide a strong basis for the timely evaluation and ongoing oversight of many aspects of the health care reform proposals. However, we believe that a sound foundation has not been fully laid for monitoring and assessing some of its most fundamental provisions, such as guaranteed coverage, insurance market reforms, essential benefits, affordability, shared responsibility, and cost control measures. Without such a foundation, we are concerned that as the reforms are implemented their effects on the costs of, access to, and outcomes of the health care system will not be evaluated adequately. As a consequence, information on unexpected successes and shortcomings, promising practices, and lessons of experience will not be available to identify needed adjustments in program administration, regulation, or legislation. We offer recommendations to strengthen the evaluation and oversight provisions.

We hope these suggestions are helpful. Should you have any questions or wish to discuss this further, please do not hesitate to contact me at 607-255-0887 or EPTF Coordinator George Grob at 540-454-2888.

Sincerely,

A handwritten signature in black ink that reads "William M. Trochim".

William Trochim

Chair, AEA Evaluation Policy Task Force

**Evaluation and Oversight
of
Health Care Reforms in the House Discussion
Draft Bill**

**Prepared by
The American Evaluation Association's Evaluation Policy Task Force
June, 2009**



The American Evaluation Association is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness. AEA has over 5500 members representing all 50 states in the US as well as over 75 foreign countries.

Mission:

To improve evaluation practices and methods

Increase evaluation use

Promote evaluation as a profession and

Support the contribution of evaluation to the generation of theory and knowledge about effective human action.

Evaluation Policy Task Force

The goal of the American Evaluation Association’s Evaluation Policy Task Force is to promote evaluation policies that are critically important to the practice of evaluation.

The members of the Task Force are:

- William Trochim, Chair
- Eleanor Chelimsky
- Leslie Cooksy
- Katherine Dawes
- Patrick Grasso
- Susan Kistler
- Mel Mark
- Stephanie Shipman
- George Grob, Consultant

The term “evaluation policy” encompasses a wide range of potential topics that include (but are not limited to): when systematic evaluation gets employed, and on what programs, policies and practices; how evaluators are identified and selected; the relationship of evaluators to what is being evaluated; the timing, planning, budgeting and funding, contracting, implementation, methods and approaches, reporting, use and dissemination of evaluations; and, the relationship of evaluation policies to existing or prospective professional standards.

Evaluation and Oversight of Health Care Reforms in the House Discussion Draft Bill

This paper provides analysis and advice regarding the evaluation and oversight provisions of the House draft discussion bill to provide affordable, quality health care for all Americans and reduce the growth in health care spending, and for other purposes. The comments are those of the American Evaluation Association's Evaluation Policy Task Force (EPTF). It is based on the principles outlined in "An Evaluation Roadmap for a More Effective Government," published by the EPTF in February, 2009 (<http://www.eval.org/aea09.eptf.eval.roadmapF.pdf>).

In summary, various provisions in the bill will provide a strong basis for the timely evaluation and ongoing oversight of many aspects of the health care reform proposals. However, we believe that a sound foundation has not been fully laid for monitoring and assessing some of its most fundamental provisions, such as guaranteed coverage, insurance market reforms, essential benefits, affordability, shared responsibility, and cost control measures. Without such a foundation, we are concerned that as the reforms are implemented their effects on the costs of, access to, and outcomes of the health care system will not be evaluated adequately. As a consequence, information on unexpected successes and shortcomings, promising practices, and lessons of experience will not be available to identify needed adjustments in program administration, regulation, or legislation. We offer recommendations to strengthen the evaluation and oversight provisions.

Why Systematic Evaluation and Oversight Are Needed

It is obvious that the authors of this draft legislation are well aware of the need for systematic evaluation and oversight. This is evidenced by the many evaluation provisions that have already been written into the draft bill. Before we analyze these further, though, it is useful to recall why evaluation and oversight are so important.

Evaluation and oversight mechanisms will provide benefits almost immediately after legislation is enacted, during the early stages of implementation. That is the time when policy makers and program managers need to identify weaknesses that can be quickly corrected and promising practices that can be shared and used widely.

Within a few years' time, some fundamental questions about the results of the legislation will be inevitable. This will include questions on the achievement of the bill's three main goals: accessibility of health care services to greater segments of the population; affordability of those services to patients; and quality of care for all Americans. More detailed questions will arise about the costs of the reforms proposed, and about the possible effects of those reforms on health disparities among groups as a result of increased access, and on potential system overload or increased waiting-times, for example. Ultimately, it will be necessary to examine some assumptions of the bill, such as the effects of preventive care on immediate and long-term costs, as well as on health outcomes over time.

While many of these short and long term questions are predictable, other unforeseen issues will also arise regarding on-going implementation and management issues.

If the needed evaluation and oversight mechanisms are in place, such questions will be answerable in terms of valid and reliable data-based comparisons between what exists today (vis-a-vis specific measures of, say, access and affordability), and what will have been accomplished by the legislation over a given period. By answering important empirical questions, evaluation can advance learning about what works, as well as how and why and whether it is transferable, and also about what does not work, thus avoiding waste of resources.

Evaluation will also bring necessary transparency and accountability to efforts at reform in a field fraught with controversy.

Finally, the evaluation language itself, once incorporated into the bill, should, by its advocacy of public accountability, improve both the general perceptions and prospects of the legislation.

As implied by the above discussion and as discussed more thoroughly in the EPTF “Evaluation Roadmap” mentioned above, we take a broad view of evaluation, recognizing the need for a body of evaluative work that covers public programs and policies throughout their life cycle and using evaluation as a tool for improving programs as well as assessing their effectiveness and impact. A complete set of evaluation approaches and methods would include but not be limited to case studies, surveys, quasi-experimental designs, randomized field experiments, cost-benefit and cost-effectiveness analyses, needs assessments, early implementation reviews, logic models and evaluability assessments. It also encompasses the fields of performance management.

Assessment of Evaluation and Oversight Provision of the Draft Bill

Our review of the draft bill reveals numerous evaluation and oversight provisions that will go a long way to answering key questions and enabling both the Executive and Legislative Branches to take corrective actions during the early implementation stages of the reforms and assess their value after the passage of a few years. Examples are existing provisions related to

- health insurance exchange integrity and beneficiary protections
- healthcare quality performance measurement
- comparative effectiveness research
- health workforce development
- special needs programs
- drug rebates
- access to care for Medicare beneficiaries with limited English proficiency
- accountable care organization pilot programs
- “medical home” pilot programs
- nursing home transparency
- home visitation programs

The first three of these—integrity reviews, health care quality measurement, and comparative effectiveness research of medical procedures—are particularly noteworthy because of their scope and

reach. They go to the very heart of the health care reforms—access to and quality of care and control of cost growth.

The care with which the evaluation provisions of these programs were constructed is also commendable. In these sections we see the articulation of important evaluation questions to be addressed, the assignment of the agencies responsible for answering them, the schedules for their completions, the reporting of the results to the Congress, and the general publication of results.

Nevertheless, we are concerned that the foundations for evaluation and oversight of the some of the most basic and overarching aspects of the proposed reforms are missing or ambiguous. This would include evaluation of some parts or all of such fundamental features as

- Effectiveness and quality of care under the public health insurance option
- Guaranteed coverage and other insurance market reforms
- Development and availability of essential benefits
- Caps on out-of-pocket spending
- Increased competition
- Medicaid expansion
- Medicare improvements
- Individual responsibility for obtaining health insurance
- Employer responsibilities
- Assistance for small employers
- Affordability
- Expansion of community health centers
- Prohibition of cost-sharing for preventive services
- Addressing health disparities
- Expansion of community based programs
- Strengthening of state, local, tribal, and territorial public health departments
- Medicare modernization and efficiency
- Prevention of hospital readmissions
- Consumer protections in Medicare Advantage
- Innovative value-based public health insurance option payment methods
- Administrative simplification and reduction of paperwork burdens

A pivotal issue is the scope and reach of the evaluation activities of the proposed Special Inspector General for Health Insurance Exchange (hereafter, Special IG), as described in section 206 (c). We find ambiguity in several respects.

The first ambiguity relates to the focus of the Special IG’s duties and activities as described in section 206(c)(3). On the one hand, they encompass evaluations as well as audits and investigations. On the other hand, this same section seems to narrow the focus of such activities “to protect the integrity of the Health Insurance Exchange, as well as the health and welfare of participants in the exchange.” The phrase “integrity” in the sphere of Inspector General functions usually relates to matters of fraud, waste, or abuse. Even the requirement to “protect . . . the health and welfare of participants in the

exchange” is ambiguous in terms of what kind of protection will be afforded and how generalized this concept of protection is.

Secondly, it is not clear how many of the fundamental health care reforms are reflected in the “Health Insurance Exchange.” For example, would the Special IG be responsible for overseeing the quality of care in the public insurance options, or only the administration of participants’ election of one of those options?

Thirdly, the evaluation activities of the special IG would not reach the quality of care and affordability of qualified health benefit plans not offered through the Health Insurance Exchange or the protection of the health and welfare of individuals not participating in it.

Finally, the success of the Health Insurance Exchange and of the entire package of health insurance reforms will be influenced by many other organizations responsible for and provisions of law covering various aspects of health insurance coverage, quality of care, affordability, and financing. These include the U.S. Departments of Labor, Treasury, and Health and Human Services and many state agencies and officials including insurance commissioners and state insurance regulators. It is not clear how far the Special IG can go in evaluating the activities of these other organizations as they affect the Health Insurance Exchange, or how practical it would be to do so.

In summary, we were unable to determine just how many of the reforms described in the draft bill are covered under the scope of the evaluation activities of the Special IG, and we are uncertain that the competencies, authorities, and resources to be made available to the Special IG would suffice to oversee and assess, with the needed methodological expertise, the effectiveness of improved access to health care and its quality and affordability, if the evaluation functions of the Special IG were to be defined very broadly.

Furthermore, except for the specific evaluation requirements mention in the first paragraph of this section, and acknowledging the important but limited and ambiguous role of the Special IG, we could not ascertain how the evaluation of the most fundamental features of the reforms would be provided. Unlike the specific programs mentioned in the first paragraph, for the most fundamental reforms we found no articulation of important evaluation questions, no schedules, no required reporting, no general publication of results, and no clear identification of who would be responsible for conducting the evaluations.

With regard to responsibilities, it would appear that the evaluation of the most fundamental reforms would benefit from the participation of recognized experts in the evaluation field, as well as traditional accountability and evaluation agencies such as the Government Accountability Office, the Institutes of Medicine, and the evaluation units and Inspectors General of relevant Federal agencies.

Recommendations

We recognize that within the very short timeframes during which the House discussion draft was prepared, it would have been nearly impossible to develop the kind of detail needed for a fully specified evaluation and oversight system. Therefore, in preparing our recommendations, we considered what might be feasible to include in the bill during the forthcoming mark-up. Our tactic

was to identify some overarching organizational principles and general steps that, if mandated in the resulting law, could evolve quickly into an adequate foundation for oversight and evaluation of the most important health care reforms. With this in mind, we recommend the following:

Annual National Report Card. Identify key recurring success measures for health care reform, such as percent of uninsured persons, health care cost indices, status of meeting modernization goals, and possibly key health status indicators that should be reported to the Congress and broadly published annually; and identify who will prepare these reports.

Key Evaluation Questions. Identify key evaluation questions for which the relevant congressional committees will require answers. These could be organized along the lines of the key concepts underlying the bill, such as coverage and choice, affordability, shared responsibility, prevention and wellness, workforce investments, and cost control. Identify who will prepare evaluation studies to address these questions, when they are to be completed, and how they will be published.

Scope of Special IG Evaluation Responsibilities. Clarify the scope of evaluation responsibilities of the Special IG with respect to 1) the particular sections of the bill to which the evaluation functions of that office apply and 2) the kinds of evaluation issues to be addressed. The latter detail could provide illustrative examples of such issues without necessarily limiting the questions the special IG could ask.

Federal Evaluation Units. Require each Federal agency with responsibilities for health care reforms covered under this reform bill to establish one or more evaluation units to be responsible for evaluation of that agency's health care activities affected by the bill. (By "agency" we mean "department" or other separately identifiable unit of the Federal Government responsible for administering or regulating health care activities.)

Annual Evaluation Plans. Require each such agency to prepare annual evaluation plans related to programs or activities covered in the bill. Require such plans to be published on the agency's website and to be sent to the appropriate congressional oversight committees.

Annual Evaluation Result Reports. Require each such agency to prepare an annual report of evaluation accomplishments, summarizing evaluation results and explaining how the results from the current and previous years' reports were used.

Health Care Evaluation Advisory Group. Establish a Health Care Evaluation Advisory Group with responsibility for (1) preparing a biennial report on the state of the nation's health care system, submitting the report to the relevant congressional committees, and making it publicly available, (2) making recommendations for the consideration of the Congress and the President; (2) providing non-binding advice to Federal agencies identified above of important health care evaluation issues; and (4) providing technical advice on methods for evaluation of health care issues.

Evaluation Resources. Authorize adequate funds to carry out such evaluation and oversight activities.

We hope this analysis and these recommendations are useful. We are available to assist the House committees in the further development of these concepts.

EPTF PERFORMANCE SUMMARY APPENDIX J: COMMENTS ON PAPERWORK REDUCTION ACT

American Evaluation Association

16 Sconticut Neck Road, #290 • Fairhaven, MA 02719 • www.eval.org • aea@kistcon.com • (888) 232-2275 • (508) 748-3326

DATE: December 16, 2009

TO: Office of Information and Regulatory Affairs, Office of Management and Budget

SUBJECT: Improving Implementation of the Paperwork Reduction Act

The American Evaluation Association (AEA) is pleased to submit comments on improving the implementation of the Paperwork Reduction Act, in accordance with Federal Register /Vol. 74, No. 206 /Tuesday, October 27, 2009 /Notices 55269.

AEA is a professional association of evaluators devoted to the application and exploration of evaluation in all its forms. AEA has approximately 5800 members representing all 50 states and the District of Columbia as well as over 60 foreign countries.

In summary, AEA recognizes the importance of minimizing paperwork burdens on the American public. However, we believe that the current paperwork review process has unintended negative impacts, particularly in denying Federal managers and policy makers timely access to information that is critically needed to address emerging problems and take advantage of promising approaches. We offer several options to speed up the reviews while promoting burden hour reductions, improving burden estimates, and enhancing the usefulness of the data collected.

We hope our attached comments are helpful. The AEA stands ready to assist OMB in improving the current process and would be happy to work with OMB, perhaps to sponsor sessions at AEA's national conferences or other venues on ways to reduce reporting burdens and enhance the quality of evaluation related surveys and the usefulness of the data collected. If we can be of assistance, or if you need more information on our comments, please do not hesitate to call on us or to contact George Grob, our senior advisor for evaluation policy (GeorgeFGrob@cs.com, 540-454-2888).

Sincerely,



Debra Rog
President



William Trochim
Immediate Past President



Leslie Cooksy
President Elect

Attachment: Comments on the Paperwork Reduction Act

COMMENTS ON THE PAPERWORK REDUCTION ACT

The American Evaluation Association (AEA) is pleased to submit the following comments on improving the implementation of the Paperwork Reduction Act, in accordance with Federal Register /Vol. 74, No. 206 /Tuesday, October 27, 2009 /Notices 55269.

AEA is a professional association of evaluators devoted to the application and exploration of evaluation in all its forms. AEA has approximately 5800 members representing all 50 states and the District of Columbia as well as over 60 foreign countries.

In summary, AEA recognizes the importance of minimizing paperwork burdens on the American public. However, we believe that the current paperwork review process has unintended negative impacts, particularly in denying Federal managers and policy makers timely access to information that is critically needed to address emerging problems and take advantage of promising approaches. We offer several options to speed up the reviews while promoting burden hour reductions, improving burden estimates, and enhancing the usefulness of the data collected.

Importance of the Paperwork Reduction Act to Evaluators

The professional practice of evaluation often makes use of surveys and other data gathering instruments. Many of our members are employees of the Federal government or work as consultants to Federal agencies. These evaluators would, for all practical purposes, be unable to perform evaluations of Federal programs without being able to gather information from the programs' beneficiaries, grantees, and stakeholders. Thus, the effectiveness and efficiency of the Paperwork Reduction Act, with its requirements for comment by the public, by the Federal agencies sponsoring program evaluations, and by the Office of Management and Budget (OMB) via the Office of Information and Regulatory Affairs (OIRA), is of critical importance to their work. Therefore, AEA appreciates OIRA's invitation to submit comments on improving implementation of the Paperwork Reduction Act.

Focus of AEA's Comments—Unintended Adverse Consequences

In requesting public comments on the implementation of the Paperwork Reduction Act, OIRA specifically requested comments on 1) reducing current paperwork burdens, especially on small entities; 2) increasing the practical utility of information collected by the Federal Government; 3) ensuring accurate burden estimates; and 4) preventing unintended adverse consequences.

All of these topics are important to evaluators, but the last one, unintended consequences, is of special concern to them. Many evaluators believe that the review process does have significant unintended adverse consequences in that it prevents important evaluative information from reaching the hands of policy makers and program managers in time to be useful. As a result, many ineffective practices and inefficiencies that could easily be remedied early in the implementation of programs are allowed to continue to the detriment of the programs' beneficiaries and at increased cost to taxpayers.

We believe it is possible to avoid these unintended adverse effects while still fulfilling the purpose of the Paperwork Reduction Act and simultaneously streamlining the review process, reducing

paperwork burdens on the public, improving estimates of burdens, and improving the utility of the data that is collected. However, to lay the groundwork for our proposals to address the four topic areas raised by OIRA, we first wish to discuss what evaluators regard as three fundamental shortcomings of the review process—1) its duration, 2) the limitation to nine of the number of respondents allowed without OMB’s approval, and 3) the orientation of OMB’s *Standards and Guidelines for Statistical Surveys*.

The length of the review process. For all practical purposes it is nearly impossible, without seeking a waiver from OMB in extraordinary circumstances, to get a survey or other information gathering instrument approved in less than 6 months. Often the process takes longer than that, sometimes a year or more. The math is straightforward. Such delays are inherent in a process which starts with an internal agency review followed by two periods of public comment, the first of 60 days, and the second of 30. On top of this, OIRA has up to 60 days to make its decision about the information collection request after it has been notified of the start of the second 30 day comment period. Even if no other administrative action were needed, the mandatory intervals of public comment and OMB approval would consume 4 months. Of course, time is needed in advance of the first public comment period to secure the approval of designated agency officials, then to absorb the comments received during the first public review, then again to absorb comments from the second public review, and finally to react to OIRA’s decisions.

The impact of the delay caused by the paperwork review process is that it makes it practically impossible to get timely evaluative information to policy makers and program managers when it is needed and would be most useful. This is particularly true during the early stages of the implementation of new programs, regulations, and legislative amendments related to such things as eligibility determinations, changes in benefits, and new administrative requirements. The delay can be particularly harmful when program managers become aware of potential waste, inefficiencies, unrealistic program objectives or goals, administrative shortcomings, or other emerging problems that require quick correction. The paperwork review process adds six or more months of delay than would otherwise occur in getting critically needed evaluative information to address emerging program issues. As a result, decision makers must make decisions based on anecdotal and incomplete information, or else delay corrective actions well beyond the period in which they could be most salutary.

Perhaps a more serious unintended consequence of the long duration of the current review process is its chilling effect on evaluation activity. Almost all evaluators and many decision makers familiar with the current paperwork review process will state that they have consciously decided not to propose evaluations of important programmatic questions, such as the likelihood of goal achievement given unexpected difficulties of implementation, because of the prospects of disheartening efforts to obtain approval of their information gathering efforts. We have no exact measurement of the scope of this problem, but we can confidently report that such behavior is commonly expressed in meetings and private conversations about evaluations under consideration. A sentiment often expressed relative to the appropriate design of a needed evaluation is that it makes no sense to plan such a survey because the OMB paperwork review process will prevent getting the needed information in time.

The nine respondent limit. The paperwork review process is not required when nine or fewer respondents are involved. This is seldom a useful exception. The principles underlying the statistical sciences, especially those pertaining to random sampling, generally preclude the use of sample sizes less than 30 for making statistically reliable projections. Even when probabilistic sampling methods are not needed, the limit of nine respondents precludes the use of the most common evaluation methods.

The exception to the paperwork review process involving nine or fewer respondents is useful for pretesting of information collection instruments intended for larger numbers of respondent. Such pretesting is an essential step in the design of surveys. However, even for this limited purpose, the limitation to nine respondents can be counterproductive. Many surveys could benefit from more extensive pretesting which would improve the value of the information collected and reduce confusion and burdens on respondents.

Furthermore, one of the topics about which OIRA has requested comments is the accuracy of the burden estimates that accompany proposals for information requests that it receives. The best way to estimate the burden hours of a survey is through the experience of pretesting. The current nine respondent limit hampers even this aspect of survey development and management.

Orientation of OMB's Standards and Guidelines for Statistical Surveys (2006). OMB requires that surveys submitted for approval under the Paperwork Reduction Act follow its *Standards and Guidelines for Statistical Surveys*. This document is a thoughtful, carefully prepared compilation of 20 standards related to the planning, design, execution, and reporting of statistical data. It certainly is a useful reference for anyone interested in sound statistical practice. However, this document may often be misinterpreted as applying to any collection of information for evaluations or other analytical studies. As discussed in the next section, "Proposals for improving the paperwork review process," modern evaluation tools include the use of such approaches as case studies, focus groups, open ended web based surveys, and probe samples for which many of OMB's standards and guidelines are irrelevant, distracting, or misleading. For example, the standards require "selecting samples using generally accepted statistical methods (e.g., probabilistic methods that can provide estimates of sampling error). Any use of nonprobability sampling methods (e.g., cut-off or model-based samples) must be justified statistically and be able to measure estimation error." They require a "Plan for a nonresponse bias analysis if the expected unit response rate is below 80 percent." They then go on to provide detailed guidance on the formulas to be used in such a nonresponse analysis, including both a "weighted and unweighted response rate," with the "unweighted response rate (RRU)" defined as "the ratio of the number of completed cases (or sufficient partials) (C) to the number of in-scope sample cases (AAPOR, 2004)." A similar definition of "weighted response rate" is given, and both definitions are accompanied by complex mathematical formulas that would be inappropriate to reproduce here.

The above excerpts show that the standards are clearly intended for larger quantitative statistical surveys and, as such, are irrelevant to focus groups, case studies, and many kinds of qualitative analysis. Perhaps OMB did not intend these standards to be used for all evaluation studies involving surveys or other forms of information collections. However, in practice, most surveys are judged on the basis of these standards, whether relevant or not, because of misconceptions surrounding their

purpose. Ironically, these standards sometimes promote the use of studies with a very large number of respondents, thus increasing the paperwork burden on the public.

Proposals for Improving the Paperwork Review Process

AEA recognizes the importance of minimizing paperwork burdens on the American public and acknowledges the statutory duties of officials in Federal agencies and OMB under the Act. We offer the following recommendations for OMB to consider within the framework of the current law or in requesting statutory changes to the Paperwork Reduction Act if necessary.

While there are practical problems connected with the paperwork review process for large evaluations, especially summative evaluations that are needed in the latter stages of a program's life cycle, our recommendations here focus on smaller data gathering processes that are most useful for formative evaluations, especially during the early or mid stages of a program's life cycle.

Decisions about many emerging program problems or opportunities could be better informed with quick evaluations of limited scope and relatively small public paperwork burdens. The field of evaluation has matured considerably over the last ten to twenty years, especially with respect to the quality and usefulness of such rapid study methodologies. This includes the effective use of low burden approaches like short electronic email or web based surveys, structured interviews, on-site inspections, focus groups, case studies, probe samples, and analysis of existing data sets. Many of these methods involve qualitative rather than quantitative data, or both. Using several of these methods simultaneously to triangulate the results can greatly enhance the validity and reliability of results as well as the relevance of findings. Thus, these approaches, now commonplace in the field of evaluation, could be used to address OMB's stated interests in minimizing burdens, improving burden estimates, improving the utility of data collected, and avoiding the unintended adverse consequences of the current paperwork review process--if the approval process could be expedited.

Facilitating the use of such approaches could be accomplished with relatively modest changes to current procedures. Recognizing the need for independent review to ensure the quality, scope, focus of surveys and other information gathering instruments, we are not proposing to eliminate the review of such instruments by designated agency officials. Instead, our proposals are centered on decreasing or eliminating the requirements for public comments and OMB approval in certain cases. In this vein, we offer the following options for consideration by OMB, to be used either alone or in combination.

Exemptions based on higher number of respondents. This proposal is to increase the number of respondents allowed before triggering the requirement for public comments and OMB review. Any increase would be efficacious, but we suggest consideration of increases to 50, 100, or 500 respondents. Needless to say, the higher the number the stronger the analytic power. However, even at the lower end of this field, a sample of up to 50 respondents can yield important insights about such things as emerging program problems and opportunities and significant changes over time. Larger (but still relatively small) samples can also provide stronger platforms for pretesting surveys and estimating burden hours. This could improve the analytic value of the data collected through, and, in some cases, reduce the burden of the ultimately approved instrument.

Burden based limits. This proposal is to allow an exemption from public comments and OMB review based on burden hour limits instead of number of respondents. For example, a burden hour limit of 100 hours could support information requests of 200 respondents based on average length of 30 minutes of burden per response, or 400 respondents based on average length of 15 minutes of burden per response.

Burden hour budgets. This proposal is to delegate annual burden hour budgets to Federal agencies to be used at their discretion, and allow these budgets to be allocated to offices within the agencies that conduct research, evaluation, or other such studies. This could be coupled with burden based limits or respondent thresholds discussed above. This approach would preserve control over burden hours, but without the need for public comments or OMB clearance. Information requests above the burden based or respondent based limits would still need to go through public comment and OMB clearance processes. Agencies could also request increases in their burden hour budget to accommodate more small projects if they can justify them to OMB's satisfaction.

Broader exemptions and limits for voluntary responses. The above policies could even further ameliorate the paperwork burdens on the public if they were limited to surveys for which responses could be purely voluntary. This could alleviate concerns of citizens that failure to respond to a request for information from a Federal government agency might result in adverse consequences for them. Implementation of this option would be accompanied by the prominent display of a label, designed by the responsible agency, in consultation with OMB, that assures the potential respondent of the voluntary nature of the request.

Generic clearances. This proposal is to allow exemptions for information requests in certain situations where information requests are now widely regarded as fairly routine, appropriate and useful. Examples include: *early implementation studies* intended to quickly gather information through interviews and document reviews to determine if there are any obvious problems with grantees meeting schedules or otherwise successfully getting started in implementing the new program; *promising approaches and obstacles studies*, intended to identify common problems among grantees and useful solutions or practices to be shared among them; *accountability reviews* to determine if funds are being spent only for authorized purposes; *routine pre- and post- tests* administered in connection with training programs; and *customer satisfaction surveys* made after a program beneficiary has an interaction with an agency or grantee providing a service or other benefit. The latter category is a good example of methods used by many private sector businesses shortly after providing a service to their customers. Allowing exemptions to the public comment period and OMB review could facilitate improved service in public programs, as now routinely occurs in the private sector.

There are many other examples of more substantive types of evaluations that an agency might use in connection with its particular programs. Examples might be an evaluation of activities of volunteer partners in certain Federal social interventions that are designed to enhance compliance with regulations or to improve communications and stakeholder collaborations.

What we would propose for these generic clearances is that Federal agencies identify categories of evaluations that they would like to do routinely under certain circumstances, such as early implementation reviews for new waves of grants. For such categories, the agency could produce a

guide or standard for conducting the studies. The guide could be subjected to the statutory paperwork reviews process and then be approved for use for a period of time, say three to five years. It could then be updated if necessary and reviewed again under current procedures. These generic categories could be subject to the burden hour budgets or respondent size limits discussed above. This approach would maintain the advantages of the current paperwork review process, but would significantly improve the timeliness and utility of the reviews.

Standards and Guidelines. This proposal is to limit the applicability of OMB's *Standards and Guidelines for Statistical Surveys* to large statistical surveys intended to produce scientifically projectable quantitative results. Clarifying that these standards do not apply to qualitative studies or to methods that are intended to produce quick insights about program operations, timely information to inform management decisions in the light of emerging problems, or to exploit promising opportunities would help reduce paperwork burdens, provide timely and useful results to top level managers and policy makers, speed up the paperwork review process, and reduce frustrations and misunderstandings inherent in the current system.

Conclusion

All of the above policies would promote improvements in the four topic areas for which OIRA has requested comments. They could help minimize burdens, improve the estimating of burdens, improve the utility of data requested, and avoid the unintended consequences attendant upon the current paperwork review process. At the same time, critically needed quality controls would still be in place at the agency level for all information requests, and at the OMB level for the larger information collection instruments associated with major data systems, quantitative statistical studies, and large and complex summative studies.

AEA's Offer of Technical Assistance

The AEA stands ready to assist OMB in achieving the four types of improvements about which it has requested comments. AEA would be happy to work with OMB, perhaps to sponsor a session at AEA's national conference, or perhaps in other venues on ways to reduce reporting burdens and enhance the quality of evaluation related surveys and the usefulness of the data collected.

Again, AEA wishes to reiterate its understanding of the need to minimize public paperwork burdens. We hope the above analysis and recommendations are useful and thank OMB for inviting comments on this important subject.

EPTF PERFORMANCE SUMMARY APPENDIX K: COMMENTS ON HOME VISITATION PROGRAM, AUGUST 2010

American Evaluation Association

16 Sconticut Neck Road, #290 • Fairhaven, MA 02719 • www.eval.org • aea@kistcon.com • (888) 232-2275 • (508) 748-3326

DATE: 8/17/2010

TO: Mary K. Wakefield, Administrator, Health Resources and Services Administration

SUBJECT: Comments on Evidence of Effectiveness of Maternal, Infant, and Early Childhood Home Visiting Programs

The American Evaluation Association (AEA) is pleased to submit comments on proposed criteria for evidence of effectiveness of home visiting program models for pregnant women, expectant fathers, and primary caregivers of children birth through kindergarten entry, in accordance with Federal Register Notice of 7-23-10.

AEA is a professional association of evaluators dedicated to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. AEA has approximately 6000 members representing all 50 states and the District of Columbia as well as over 60 foreign countries.

In summary, AEA recognizes the importance of using “evidence based” models as a basis for distributing funds available under the Affordable Care Act’s Maternal, Infant, and Early Childhood Home Visiting Program. We believe that the proposed criteria and methodology for a systematic review of such models represent a thoughtful starting point for assessing the evidence of their effectiveness. However we do have concerns about the studies upon which the evidence is based are rated. We offer recommendations to 1) forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings, and avoid using the label “gold standard” in connection with random assignment designs in the rating methodology, 2) use additional criteria to assess the value of impact evaluations, 3) more specifically identify alternative impact evaluation methods, and 4) emphasize the value of multiple studies and mixed methods.

We hope our attached comments are helpful. If we can be of assistance, or if you need more information on our comments, please do not hesitate to call on us or to contact George Grob, our senior advisor for evaluation policy (GeorgeFGrob@cs.com, 540-454-2888).

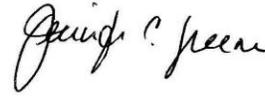
Sincerely,



Leslie Cooksy
President



Debra J Rog
Immediate Past President



Jennifer Greene
President Elect

Attachment: Comments on Evidence of Effectiveness of Maternal, Infant, and Early Childhood Home Visiting Programs

Comments on Evidence of Effectiveness of Maternal, Infant, and Early Childhood Home Visiting Programs

The American Evaluation Association (AEA) is pleased to submit comments on the proposed criteria to be considered in assessing evidence of effectiveness of maternal, infant, and early childhood home visiting program models and on the methodology for a systematic review of such evidence, in accordance with Federal Register Notice of 7-23-10.

AEA is a professional association of evaluators dedicated to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. AEA has approximately 6000 members representing all 50 states and the District of Columbia as well as over 60 foreign countries.

In summary, AEA recognizes the importance of using “evidence based” models as a basis for distributing funds available under the Affordable Care Act’s Maternal, Infant, and Early Childhood Home Visiting Program. We believe that the proposed criteria and methodology for a systematic review of such models represent a thoughtful starting point for assessing the evidence of their effectiveness. Our comments focus on the criteria for judging the effectiveness of the models. We offer recommendations to 1) forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings, and avoid using the label “gold standard” in connection with random assignment designs in the rating methodology, 2) use additional criteria to assess the value of impact evaluations, 3) more specifically identify alternative impact evaluation methods, and 4) emphasize the value of multiple studies and mixed methods.

AEA Evaluation Principles and Practices Regarding Evaluation Methodologies

The AEA has sponsored numerous discussions and presentations on evaluation methods, including impact evaluations, through its annual conventions and in its professional journals, the peer reviewed *American Journal of Evaluation* and *New Directions for Evaluation*. In addition it has prepared “An Evaluation Roadmap for a More Effective Government,” a paper describing its vision of the role of evaluation in the Federal Government. The Roadmap outlines steps to strengthen the practice of evaluation throughout the life cycle of programs. It presents evaluation as an essential function of government that can enhance oversight and accountability of Federal programs, improve the effectiveness and efficiency of services, assess which programs are working and which are not, and provide critical information needed for making difficult decisions about them.

Particularly germane to our discussion here about the relative merits of impact evaluation methods and their application to the home health visitation program is the following excerpt from the Roadmap regarding analytic approaches and methods.

“Which analytic approaches and methods to use depends on the questions addressed, the kind of program evaluated, its implementation status, when the evaluation results are needed, what they are needed for, and the intended audience.

“No simple answers are available to questions about how well programs work, and no single analytic approach or method can decipher the inherent complexities in the program environment and assess the ultimate value of public programs. Furthermore, definitions of ‘success’ may be contested. A range of analytic methods is needed, and often several methods—including quantitative and qualitative approaches—should be used simultaneously. Some evaluation approaches are particularly helpful in a program’s early developmental stages, whereas others are more suited to ongoing and regularly implemented programs.”

These principles thus maintain that rigorous and useful evaluation is firmly anchored in the evaluation questions to be addressed and the developmental stage of the program to be evaluated. Programs with demonstrated efficacy and potential for broad implementation (among other preconditions for assessment) are best suited for rigorous impact evaluation.

It is in the context of these broad principles that we offer the following comments on the Department of Health and Human Services (HHS) proposal for assessing the value of studies that measure the impact of the Maternal, Infant, and Early Childhood Home Visitation Program.

Summary of Proposed Criteria for Assessing the Evidence of Home Visitation Models

As a point of reference for our comments, we include here an excerpt from the Federal Register Notice that summarizes the rating scheme that the Department of Health and Human Services (HHS) proposes to use in deciding which home visitation models will be regarded as effective based on “well-designed, rigorous impact research” and therefore eligible for funding under the home visitation program.

“HHS proposes that an impact study will be considered high, moderate or low quality depending on the study’s capacity to provide unbiased estimates of program impact. Studies that are rated ‘high’ and ‘moderate’ (see Table 1 below) [not included here], therefore, would meet requirements to be considered ‘well-designed, rigorous impact research.’ In brief, the high rating would be reserved for random assignment studies with low attrition of sample members and no reassignment of sample members after the original random assignment. The moderate rating would apply to studies that use a quasi-experimental design and to random assignment studies that, due to flaws in the study design or execution (for example, high sample attrition), do not meet all the criteria for the high rating. To receive the moderate rating, studies would have to demonstrate that at the study’s onset, the intervention and comparison groups were well matched on specified measures (i.e. baseline equivalence), such as a pretest measure of targeted outcomes or race and maternal education. Studies that do not meet all of the criteria for either high or moderate quality would be considered low quality.

“As summarized in Table 1 [not included here], the rating scheme would consider five dimensions: (1) Study design, (2) attrition, (3) baseline equivalence, (4) reassignment, and (5) confounding factors.”

In our comments below we may cite other sections of the Notice, but we believe that the excerpt above contains the key concepts that are germane to our comments, which now follow.

AEA Comments and Recommendations

In summary, we believe that the HHS proposal for assessing the scientific rigor of evaluation designs and the resulting relative impact and value of home visitation models may result in over-rating the quality of some studies, under-rating some high quality evidence from quasi-experimental studies, and possibly mis-valuing the impact of some potentially effective home visitation models. We discuss our concerns and offer recommendations in the sections below.

I. Randomized Designs Are Not Necessarily Superior to Quasi-Experimental Designs in All Circumstances.

No Consensus on “Gold Standard.” As stated in the Notice, HHS premises its preference for random assignment studies on its understanding that “Randomized control designs are often considered the ‘gold standard’ of research design because personal characteristics (before the program begins) do not affect whether someone is assigned to the program or control group.”

Within the professional evaluation community, there is no consensus on the relative weight of evidence accorded random assignment studies compared to other impact studies. In fact, this is a matter of some controversy. There is widespread agreement that randomized experiments can provide the most credible evidence of effectiveness under certain conditions. However, there are also limitations on what such studies can offer, and in other circumstances other methodologies may be better suited for measuring impact.

Recently, this topic has been the subject of three comprehensive reviews completed by the Government Accountability Office (GAO), the Congressional Research Service (CRS), and the Network of Networks for Impact Evaluation (NONIE) comprised of the Organization for Economic Co-operation and Development’s Development Assistance Committee (OECD/DAC) Evaluation Network, the United Nations Evaluation Group (UNEG), the Evaluation Cooperation Group (ECG), and the International Organization for Cooperation in Evaluation (IOCE).

For your convenience, the following citations and internet links are provided for easy reference:

- Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions, GAO (2009), <http://www.gao.gov/new.items/d1030.pdf>
- Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues, CRS (2006), <http://digital.library.unt.edu/ark:/67531/metacrs9145/m1/>
- Impact Evaluations and Development, NONIE Guidance on Impact Evaluation, NONIE (2009), http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf

The last of these three studies was prepared in the context of international development. We recognize that HHS has legitimately excluded from its review impact studies of international development home

visitation programs. Nevertheless, the NONIE report provides a more general perspective on multiple types of impact evaluations that are germane to U.S. domestic programs.

All three of these studies reached similar conclusions about random assignment studies. The following excerpt from the GAO report is typical of the conclusions in all three references.

“In our review of the literature on program evaluation methods, we found general agreement that well-conducted randomized experiments are best suited for assessing intervention effectiveness where multiple causal influences lead to uncertainty about what has caused observed results but, also, that they are often difficult to carry out. Randomized experiments are considered best suited for interventions in which exposure to the intervention can be controlled and the treatment and control groups’ experiences remain separate, intact, and distinct throughout the study. . . . Several other research designs are generally considered good alternatives to randomized experiments, especially when accompanied by specific features that help strengthen conclusions by ruling out plausible alternative explanations” (p. 20).

Federal Practice in Assessing Effectiveness Also Varies. In this regard, it is worth noting that as part of its review, GAO examined the practices of six Federally supported initiatives that identify effective interventions. While all of them consider the value of randomized studies, three of the six, all within HHS, (Evidence-Based Practice Centers at the Agency for Healthcare Research and Quality; Guide to Community Preventive Services at the Centers for Disease Control and Prevention; and National Registry of Evidence-Based Programs and Practices at the Substance Abuse and Mental Health Services Administration) do not require randomized experiments for interventions to receive their highest evidence rating.

Potential Shortcomings of Randomized Experiments. The rating system proposed for the home visitation program does attempt to deal with the limitations of random assignment studies by restricting its highest rating to those randomized studies with low attrition of sample members and no reassignment of sample members after the original random assignment. However, there are additional fundamental shortcomings that can limit the value of random assignment studies in providing evidence of the effectiveness of a program model, such as lack of fidelity to program design in the program’s implementation, inadequate construct validity of outcome measures, limitations regarding the generalizability of the results, and dearth of information about the reasons for high or low impact. Here we highlight the last two of these shortcomings.

Generalizability. The argument for randomized experiments is that, if successfully conducted, they provide strong internal validity, that is, they provide a good test of whether the program as implemented made a difference in the outcome as measured. However, there may be a variety of reasons or conditions which limit the application of the program to settings other than those examined by the study.

For example, it may be that the program was not implemented well or that the outcomes were not well measured. Or, the program being evaluated and hence the entire impact evaluation may be narrowly limited to particular geographic, socioeconomic, and programmatic

conditions that constrain the ability to predict whether the intervention would be successful under other circumstances.

Explanatory Value. In addition to methodological challenges in conducting randomized experiments, such studies provide little explanatory information regarding important causal mechanisms or contextual contingencies related to observed changes in valued outcomes. Random assignment studies may be able to demonstrate the impact of a program model, but are less likely to shed much light on the reasons for success or failure, or the opportunities to improve the intervention models.

With these and other kinds of limitations, when used alone, random assignment evaluations can fail to determine with confidence if the program had a strong impact on policy-relevant outcomes. Of course, problems of program implementation fidelity, valid outcome measurement, and generalizability of results are shared by other study methods as well. Nevertheless, these examples demonstrate the problems with focusing on only one type or aspect of validity when assigning priorities across methods.

All this having been said, we would not want to leave the impression that we underestimate the value of randomized assignment evaluations. Not only are there circumstances where they provide strong evidence of impact, but also the cumulative information they bring can be very helpful in meta-analytic studies. But as with all evaluative efforts applied outside the laboratory and in the real world, they have both strengths and weaknesses. For these reasons, we believe they should be rated on a broader set of criteria than are suggested in the HHS proposal.

Recommendation 1.a. Forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings.

In HHS's proposed rating scheme, a high rating would be reserved for random assignment studies with low attrition of sample members and no reassignment of sample members after the original random assignment. The proposed system risks two kinds of errors: rating as high quality a study with a randomized design that does not meet all the criteria for valid inference; or rating as moderate or low what is actually high quality evidence from studies using quasi-experimental or other designs that have adequately addressed challenges to causal inference.

We therefore recommend that the high rating not be reserved for random assignment studies only, but that impact evaluations using other designs could also be eligible for a high rating if they, individually or in combination with other studies (possibly but not necessarily including random assignment studies) provide rigorous and credible evidence of effectiveness. Conversely, we recommend that those random assignment studies that fail to provide rigorous and credible evidence be rated as moderate or low, depending on a number of factors that we will discuss in this and other sections of our comments.

Recommendation 1.b. Avoid using the label "gold standard" in connection with random assignment evaluation designs in the rating methodology.

Given the lack of consensus in the professional evaluation community on the matter, it would be prudent to avoid using the label ‘gold standard’ which, as the HHS proposal notes, has sometimes been applied to random assignment studies alone. The continuing use of this designation can be harmful in that it may lead program sponsors or public agencies to commission expensive and time consuming randomized studies under conditions where other methods might be more appropriate. Furthermore, it may cause senior government officials and others to regard as unevaluated or unproven the value of effective programs which have been reviewed using well validated and contextually appropriate methodologies.

II. Many Other Factors, Especially Real World Conditions and Fidelity of Implementation, Affect the Quality of Scientific Evaluation Evidence

The value of an impact study depends on more than its design type. It also depends on the details of the design, the manner of its implementation, and the results obtained, among other things. The HHS proposal recognizes these broader considerations by establishing a number of criteria to be used in assessing the relative effectiveness of home visitation models based on the quality of the studies it will review for that purpose. However, we believe that there are a number of additional criteria that might be useful to assess the quality of the studies. We offer the following standards as necessary for high-quality, rigorous impact assessment:

- adequacy of sample size to detect the effects that are expected
- the magnitude of the impact of the home visitation model found by the study
- validity and reliability of outcome measures
- implementation fidelity to program design
- appropriate correction for differences in baseline characteristics and nested effects
- appropriateness of data analyses and reporting
- clarity of description of the comparison group's experience
- the potential for or limitations to scaling up the program model or using it in circumstances that are different from those in the study
- evidence of the potential sustainability of results after the intervention has ended
- information from the study that helps explain the reasons for success or failure of the model being reviewed
- ethical considerations in replicating the model
- the extent to which the study rules out alternative explanations of success or failure of the model being reviewed
- the nature of, and process for selecting, study sample sites and participants
- assessment of the contextual factors of importance in program success or failure

For many of these criteria, there is no guarantee that a random assignment study will provide the kind of information that is most beneficial in deciding which program models are most effective. Other study types may be better suited in those cases.

Recommendation 2. Provide additional criteria to use in assessing the value of impact evaluations.

We believe additional criteria should be considered in rating the value of impact evaluations. Studies of poorly implemented models or narrowly designed studies likely warrant moderate or low ratings, despite having a random assignment design. Similarly, certain quasi-experimental approaches that have incorporated strategies for producing valid results may warrant the highest rating.

III. Other Evaluation Designs Can Support Causal Attribution, Especially When They Can Rule Out Other Potential Causal Factors.

The HHS proposal recognizes that quasi-experimental evaluation methods may produce rigorous evidence of effectiveness of home visitation program models. However, it does not identify or describe what these methods are. They are considered together in a category that accords them a value that is characterized simply as being inferior to randomized assignment studies in all evaluation contexts and for all home visitation programs. In fact, some of these methods are now highly developed and widely used by professional evaluators, many with strong results under appropriate circumstances. Quasi-experimental designs vary considerably in how well they can be expected, on average, to address the selection problem that is generally handled well by a random assignment experiment. However, they may be effective in dealing with selection bias in certain conditions or contexts. Furthermore, selection bias is not the only limitation of impact methodologies. The overall quality of the design depends on the extent to which, by the design and other study features, selection bias and other potential threats have been rendered implausible. All evaluation designs have various limits, including random assignment studies.

Recommendation 3. More specifically identify alternative impact evaluation methods.

We therefore recommend that the commonly accepted evidence based methods be named and described in the rating methodology, along with their relative advantages and the conditions under which they bring value in discerning evidence based impact of program models. The purpose of doing so would be to assist advocates of program models to determine whether their models have a chance of being considered acceptable in the context of submitting proposals for funding under the home visitation program.

It is beyond the scope of these comments to construct the details of such a comparison, although they are described in the three references that we cited earlier. Here we would like to suggest that methodologies such as the following be specifically identified, described, and briefly assessed. Most of them are very different from the short examples in the HHS proposal which refers to quasi-experimental studies as involving self-selection or program selection by level of risk.

- **Purposefully chosen** comparison group designs, where individuals are selected to serve as a control group that resembles the treatment group as much as possible on variables related to the desired outcome

- **A pipeline approach**, which compares outcomes for the treatment group with households or individuals who have not yet experienced the treatment (delayed receipt), and notes the danger of contamination
- **Propensity scoring**, which statistically matches the treatment group with others with the same cluster of characteristics
- **Regression discontinuity design**, which compares outcomes for treatment and comparison groups that are formed by having scores above or below a cut-point on a quantitative eligibility or selection variable
- **Interrupted time-series design**, which compares trends in repeated measures of an outcome for a group before and after an intervention or policy is introduced

As noted earlier, each of these (and other) designs may be effective in certain conditions or contexts, perhaps with auxiliary methods or evidence. What counts here is not the study design per se, but rather the degree to which the design and other study features are successful in ruling out selection bias and other threats.

IV. Knowledge of Program Impact is Enhanced by Considering Multiple Studies and Using Mixed Methods.

Regardless of the perceived value of random assignment studies, relative to other approaches to impact evaluation, there is almost universal agreement within the evaluation profession that knowledge of program impact is enhanced by considering the results of multiple evaluations of programs and of evaluations performed using a combination of methods. In fact, independent replication of program outcomes under different conditions and with different methodologies can provide especially convincing evidence of program effectiveness.

<p>Recommendation 4. Emphasize the value of multiple studies and mixed methods.</p>
--

For this reason, we recommend that in assessing the evidence of the effectiveness of home visitation program models, the HHS review team consider all valid studies pertaining to the models it reviews. This could include in depth comparative case studies, contextually driven normative studies, process tracing, meta analyses, and various forms of syntheses of evidence, if conducted in tandem with or in addition to experimental or quasi-experimental studies. Ideally, conclusions about the relative effectiveness of various program models should be based on the body of work on that model and not just on a single study, if more than one valid study is available.

Conclusion

In summary, AEA recognizes the importance of using “evidence based” models as a basis for distributing funds available under the Affordable Care Act’s Maternal, Infant, and Early Childhood

Home Visiting Program. We believe that the proposed criteria and methodology for a systematic review of such models represent a thoughtful starting point for assessing the evidence of their effectiveness. However we do have concerns about how the studies upon which the evidence is based are rated. We offer recommendations to 1) forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings, and avoid using the label “gold standard” in connection with random assignment designs in the rating methodology, 2) use additional criteria to assess the value of impact evaluations, 3) more specifically identify alternative impact evaluation methods, and 4) emphasize the value of multiple studies and mixed methods.

EPTF PERFORMANCE SUMMARY APPENDIX L: HYDROFRACKING LETTER, MARCH 2011

March 7, 2011

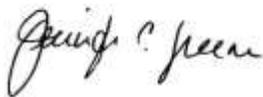
Letters to the Editor
The New York Times
620 Eighth Avenue
New York, NY 10018

Dear Editor,

The problem of political pressure on policy-oriented research and evaluation findings, such as that documented in Ian Urbina's recent article, goes far beyond environmental issues, urgent as these are. In a democratic society, policymakers and citizens alike need carefully collected evidence and even-handed evaluations about government initiatives that render their effects transparent to the public gaze. But, such evaluations cannot be credible or serve democratic interests if they are politically manipulated.

Political pressures on evaluators have been endemic since time immemorial, in both Democratic and Republican administrations. This is why the American Evaluation Association, in its recent policy statement on government evaluation (*An Evaluation Roadmap for a More Effective Government*), has insisted on independence in the design, execution, report-writing, publication, and dissemination phases of their work.

Thank you for bringing this important matter to public attention, and especially for the article's careful reporting and documentation.



Jennifer Greene, Ph.D.
President, American Evaluation Association

EPTF PERFORMANCE SUMMARY APPENDIX M: CONSULTANTS LOG OF OUTSIDE MEETINGS AND CONTACTS

Date	Attendees	Subject and Outcomes
11-3-11	John Tunna, Director of Research and Development, DOT, and MikeCoplen	I met today with John Tunna, Director of Research and Development, and MikeCoplen of the Department of Transportation. Mr. Tunna was interested in promoting evaluation of research programs to buttress budget decisions. He gave examples of studies he had completed quickly about the impact of years of research on railway transportation safety projects, focusing on injury reduction. I encouraged him to pursue this line of work more formerly, in essence to “mass produce” such studies and also to build in an emphasis on potential impact of research grants at the time the grants are made.
10-4-11	NIH CTSA Evaluation Key Function Committee	<p>Patrick Grasso and I made a presentation at the face to face meeting of the Clinical and Translational Science Award (CTSA) Evaluation Key Function Committee. We explained to the 100+ assembled members AEA’s evaluation policy initiative, with emphasis on the EPTF and the Evaluation Roadmap. We then reflected on lessons learned from this work that might be useful to the CTSA Evaluation Committee in fashioning their own more specific Roadmap like policy. Bill Trochim served as chair of our panel. I can report that that interest was strong, with a good exchange of questions and answers after our presentation. Bill told me later that our session was enthusiastically received.</p> <p>After our panel, Bill Trochim led a separate break out session of people more intensely interested in this project, soliciting ideas to refine the policy as it has been developed so far.</p> <p>Bill told me that the next step will come in three weeks or so with a revised version of the CTSA evaluation policy reflecting input received from the Evaluation Committee. He will then ask for EPTF comment once again.</p>
9-24-11	Margaret Johnson Cornell University	I exchanged emails with Margaret Johnson of Cornell University regarding the AEA’s Evaluation Roadmap. Margaret is doing her Ph.D. dissertation on evolution policy and will be featuring the Roadmap in it. I answered questions from her regarding the nature of changes made between the very first edition of the document and the current officially published version. I explained the nature of input from AEA members and the process used by the EPTF in incorporate many but not all of the AEA member inputs, as well as the input directly from the EPTF and the Board and our use of a professional editor’s services.

9-12-11	John Collins and Luke Holland, Senate Legislative Aides	Today I met with John Collins, legislative aide for Senator Carper on the Subcommittee on Federal Financial Management, Government Information, Federal Services and International Security, and Luke Holland, who works for Sen. James M. Inhofe. The two of them are working together on legislation to conduct objective reviews of government programs. Our meeting was very productive. In addition to providing them the Roadmap, I answered their questions about the pros and cons of various options to organize such a function within Federal agencies. They told me that they will be back in touch with an outline of proposals. They seemed very anxious to pursue this.
8-30-11	Clinton Brass, Congressional Research Service	Clinton Brass of the Congressional Research Service called me to alert me to the fact that he and his group are beginning to write a paper on GPRA. He said that he had given my contact information to a couple of his staff who would be calling me soon regarding the paper. He had given them the Roadmap as part of their introduction to this topic.
8/4/11	Laura Rothlisberger, EPA intern	Laura Rothlisberger is an intern working this summer for Katherine Dawes. She prepared a paper on evaluation authorities in EPTA, a framework for understanding and advocating strong evaluation authorities for EPA programs. The paper catalogued various legislative authorities in EPA for funding programs called "evaluation," many of which tend to be more about research than what we call evaluation. I provided technical advice to her to clarify existing evaluation authorities and organizations within HHS and the Department of Education. I also advised her in describing current evaluation authorities not to suggest that they could not be used for program evaluation even if they are currently being used for research.
7/21/11	Clinical and Translational Science Awards (CTSA) Evaluation Key Function Committee-- William Trochim, Chair, and others	I attended a meeting of the NIH Clinical and Translational Science Awards (CTSA) Evaluation Key Function Group, chaired by Bill Trochim. Bill facilitated a brainstorming meeting to identify topics for a white paper describing key evaluation functions and resources in the field of clinical translational science. The paper would be somewhat like the AEA Evaluation Roadmap, but applied for this field. This was the first meeting that will lead to a paper to be discussed at a meeting of the Key Function Group this October 3, with Patrick Grasso and I serving as introductory plenary speakers.
7/7/11	Hillary Daniels, staff to Senator Udall	Following up on our meeting of 3/11, Hillary Daniels sent me for my comments a draft bill that would require the Federal Inspectors General to exercise oversight over Federal agency evaluation organizations and activities. Because of a conflict of interest related to my current position as an intermittent employ of an IG organization, I recused myself from

		providing such comments. I asked Eleanor Chelinski, and she agreed, to provide the comments. I had originally offered to provide comments on behalf of the EPTF. However, Ms. Daniels emphasized that because of the preliminary nature of the concepts in the draft bill, she would prefer that I simply offer my own comments. Subsequently, with my recusal, she is very pleased to be receiving comments from an experienced evaluator like Eleanor.
5/23/11	Kathy Beuler, Peace Corps IG and Co-Chair of the IG Inspections and Evaluation Round Table; Mark Jones, Executive Director of the Council of Inspectors General for Economy and Efficiency (CIGIE); Tom Caufield, Executive Director of the CIGIE Training Institute; and others	Patrick Grasso, Stephanie Shipman, and I met today with members of the OIG Inspection and Evaluation Round Table. We made a presentation promoting greater interaction between AEA and the Round Table members (directors of all the IG inspections and evaluation units). In addition to explaining what AEA could offer in terms of learning, training, and networking opportunities, we also discussed what the IG community could bring to the professional evaluation community. Stephanie also provided information about the Federal Evaluators group and distributed a list of resources available within the Federal Government evaluation community. Based on feedback I received afterwards, this all seemed to go over pretty well with the group. Tom Caufield, the new Executive Director of the CIGIE Training Institute, approached me suggesting we meet to talk about evaluation training.
5/4/11	Stuart Wright, Deputy Inspector General for Evaluation and Inspections, HHS	I met with Stuart Wright by telephone conference to follow up to an earlier meeting that Patrick Grasso and I had with him and Dan Levinson, the HHS Inspector General. Dan is Co-Chair of the Evaluation Committee of the Council of Inspectors General for Integrity and Efficiency. In that earlier meeting, Patrick and I had encouraged stronger links between AEA and the IG evaluation community. In this telephone meeting, Stuart invited Patrick and me to address a meeting of the IGs' Inspections and Evaluation Round Table on May 23 or 24. The Round Table is the working group of IG Directors of Evaluation and Inspections. The meeting will also be attended by the Executive Director of the newly formed IG Academy which provides training for all IG disciplines including audit, investigations, evaluation, and management. Naturally, I agree to making the presentation. Patrick will join me.
5/1/11	Cynthia Clapp Wincek Director of Learning Evaluation and Research, USAID	I called Cynthia Wincek to discuss option for responding to EVALTALK postings critical of my favorable treatment of the USAID evaluation policy. The criticisms centered on

		<p>contradictory language in the policy regarding the preference to be given to random assignment studies for impact evaluation. She encouraged me not to respond, preferring instead to let the issue pass for now. She stated that she was more in favor of the USAID policy to not give any preferred status to any particular evaluation method.</p>
4/13/11	<p>Robert Shea, Grant Thornton Shelley Metzenbaum, Associate Director for Performance and Personnel Management Peter Grace, HUD Jay Hoffman, Dept. of Energy Ron Raborg, SSA</p>	<p>Stephanie Shipman, Katherine Dawes, and I attended a presentation and panel discussion of a report, issued today: <u>"Critical Role at a Critical Time: Survey of the Performance Improvement Officers."</u> The report was prepared jointly by the <u>Partnership for Public Service</u> and Grant Thornton (a leading accounting firm). Robert Shea, former Associate Director of OMB for Administration and Government Performance, who led the Grant Thornton effort in conducting the survey of 23 Performance Improvement Officers (PIO) of the largest federal agencies, served as master of ceremonies for a panel of 3 of the PIOs.</p> <p>The survey found mixed opinions among the PIOs, newly appointed under the recently enacted GPRA Modernization Act. Some say they are trusted members of the leadership team of the agencies and have improved agency performance. However, others said they were largely unknown in their agencies and have little impact. Many of the PIOs are Chief Financial Officers (CFO) or budget directors and believe that their PIO responsibilities are less important than their financial management functions. The report provided a litany of shortcomings, including lack of top-leader support within the agencies and insufficient authority to improve government performance and results. However, the three panelists, none of whom were CFO's of budget directors, were more optimistic about progress and potentialities. The overall tone of the presentations was that progress is being made but at a slow rate.</p> <p>Also on the positive side, many of the survey respondents cited the 2009 American Recovery and Reinvestment Act as a potential model for best practices in the performance measurement arena.</p> <p>Shelley Metzenbaum, Associate OMB Director for Performance and Personnel Management, provided some introductory remarks in which she emphasized the value and</p>

		<p>need for analyzing and using data as well as collecting it.</p> <p>Among the report's recommendations is one that calls for moving beyond compliance reporting to include performance oversight and management. It also calls for developing strategies, measuring outcomes, and evaluating programs and services.</p>
3-2-11	Hillary Daniels, Legislative Assistant to Senator Udall's	<p>At her invitation, I visited today with Hillary Daniels, legislative assistant to Senator Mark Udall. Senator Udall is interested in developing legislation to promote more effective use of evaluation within federal programs. She had previously met with Katherine Dawes who suggested that she contact me.</p> <p>She has developed a preliminary proposal with options to establish a position of Chief Evaluation Officer within each Federal agency and to strengthen the evaluation function within Inspector General offices. I encouraged her to pursue evaluation legislation, answered questions, and offered suggestions for the broad authority that Senator Udall is interested in. I used the Evaluation Roadmap as the framework for further development and refinement of her ideas, gave her the names of prominent people, such as Thomas Chapel of CDC who has been appointed the Chief Evaluation Officer of CDC, and sent her copies of the AEA Roadmap as well as the Obama Administration evaluation policies, the new USAID evaluation policy, and the newly enacted GPRA Modernization Act.</p> <p>Afterwards, she wrote back that she would like to check in again with me as she further develops ten proposal.</p>
2-16-11	Thomas Chapel and Ann O'ConnOr of CDC	<p>Thomas Chapel, CDC's Chief Evaluation Officer, and Ann O'Connor, CDC's Program Planning and Advancement Chief called me to ask for advice regarding an ongoing revamp of CDC's evaluation policy. they have been working on this for a number of months but were inspired by USAID's issuance to put theirs to bed. I will be looking over some of their documents and offering advice from an AEA perspective. Interestingly he told me that they had already used the Roadmap as a reference for adapting their evaluation policies--more proof that the Roadmap is being used.</p>
2-9-11	Vic Dukay, Matt Oresman, Michelle Orza, Patrick Grasso	<p>I met today with the informal Foreign Assistance Evaluation Work Group that we formed two years ago to collaborate on promoting evaluation policy for foreign assistance programs.</p>

		<p>This is the same group that worked with staff of Senate and House foreign relations committees, the House Appropriations Committee, the Office of the U.S. Global AIDS Coordinator to successfully get appropriations language for evaluation of the U.S. HIV/AIDS program and which provided advocacy and technical guidance for evaluation language that was included in House and Senate bills to reauthorize the Foreign Assistance Act. Ruth Levine, now responsible for evaluation at USAID, credited this group and the AEA Evaluation Roadmap for much of the push and technical support for the recently issued USAID Evaluation Policy.</p> <p>Today's meeting was intended to start planning for next steps in evaluation policy for foreign assistance. The group decided on the following four goals: To institutionalize evaluation policy through legislation; to ensure adequate funding of evaluation; to support USAID's implementation of its new policy, and to do this in a way that is not harmful to current evaluation policy and implementation. The last of these four objects takes into account that congressional policies may focus intensely on budget cutting, for which evaluation may be a useful tool, but which itself may be the targets for such cuts.</p> <p>The group members will individually think about next steps and meet again in early March, after the President's budget has been published. and analyzed.</p>
2-2-11	Brian Robinson of the Economic Research Unit	<p>Today I was interviewed by Brian Robinson, a writer with the Economic Research Unit. This is an office that conducts research for "The Economist." He has been asked to write a report on how evaluation is used to help government officials make decisions about what programs are needed, whether and how they should be amended, or whether they should be dropped. He called Susan Kistler seeking an interview because he had read the Evaluation Roadmap, and much of that seemed relevant to his research. I gave him an extensive background briefing not only about evaluation as we know it, but also, based his inquiry, on a looser nod more general meaning of that term as it relates to any of the various kinds of analyses and reports that government policy makers use to make program policy.</p>
2-1-11	Ruth Levine of USAD and Others	<p>Patrick Grasso and I attended a by-invitation presentation on the new USAID Evaluation Policy sponsored by Georgetown University's Walsh School of Foreign Service. Ruth Levine made a presentation on the policy. Those in attendance included some GU professors, development evaluators/executives from such groups as the International Initiative for Impact Evaluation (3IE), Millennium Challenge Corporation (MCC), Innovation Network (which does evaluation for smaller non-profit organizations), the Center</p>

		<p>for Global Development, and USAID. Some highlights: Importantly, Ruth began by citing the Roadmap as one of the important sources for the policy, calling it "hugely influential". While she said there were many other sources, aside from the Roadmap the only one she mentioned was the work by the MCC, although in that case she noted that the approach they pursue (almost always RCTs) works for MCC but would not fit at USAID.</p> <p>She also noted that the evaluation policy not only is intended to revitalize the once-strong USAID evaluation function, but also to fit into a broader USAID initiative on knowledge management, project design, and (with the State Department) revamped performance indicators. The policy is only on evaluation, not monitoring and evaluation (M&E) because of concerns that monitoring has become dominant in USAID over the last decade, and also because evaluation is the purview of the Bureau of Policy, Planning, and Learning, which produced the policy, whereas monitoring is "owned" by most parts of USAID.</p> <p>The major points of the policy she identified were: (1) focusing on both accountability and learning, which usually are complementary, but sometimes need to be weighed; (2) clarifying key terms, such as "impact evaluation" and "performance evaluation"; (3) requiring that the terms of reference for evaluations specify the methods to be used (not now usually included in the scope of work); (4) engaging local stakeholders to help identify evaluation questions and to build local capacity; and (5) increasing transparency through broader disclosure of reports, including a registry of all evaluations under way.</p> <p>In addition, the Bureau is providing two training programs in evaluation, one for program managers and one for evaluators; these are oversubscribed. They also are developing a warehouse for baseline data, and improving the search function for the existing Evaluation Warehouse.</p> <p>Finally, she noted that they have requested a significant budget increase for USAID evaluation, and that OMB had supported this request. However, she also cautioned that the Bureau has only 7 staff, with 2 more positions to be filled (if they can beat an expected hiring freeze).</p> <p>There was a lively conversation following this presentation. Most of the issues raised were old chestnuts in the field (how do you ensure independence? how do you increase evaluation use? etc.). Overall, there was a lot of support for the policy from the group.</p>
1-20-11	Vic Dukay, Lundy Foundation	I talked to Vic Dukay about reconvening our foreign assistance group--Vic, Michele Orza (National Health Policy Forum), Matt Oresman (Patton Boggs), Patrick Grasso, and myself--to discuss next steps for evaluation policy in the field

		of foreign assistance. High on the list is reconnecting with the House and Senate oversight committees to promote evaluation language in any new bills to reauthorize the Foreign Assistance Act. This could lend the weight of law to the kind of policies that USAID has just issued.
1-6-11	Sharon Stout, Department of Education	I talked today with Sharon Stout, Ph.D., Education Research Analyst, U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, about the newly enacted GPRA Modernization Act of 2010. She expressed concerns about the feasibility of its implementation and some dangers to the proper evaluation of programs under the Act. The implementation issues derive from a significant increase in the frequency of reporting progress against program goals and objectives and the scope of reporting. Performance reporting is increased from annual to quarterly, and includes requiring individual agencies to report on the contribution of results by other federal programs and by outside groups. Perhaps more significantly, she raised concerns about requirements to analyze the value of individual contribution of single Federal programs to results that will stem from collaborative efforts among Federal agencies, outside organizations, and the Congress.
12-13-10	Lisa Rajigah, 3Ie	I met with Lisa Rajigah of the Washington staff of 3Ie at her invitation. She gave me a rather thorough background briefing about the goals, structure, and current activities of 3Ie. The organization is undertaking a wide variety of activities to conduct and promote high quality impact evaluations of international development programs. It defines high quality impact evaluations as those built around a credible counterfactual with an evaluation design based on the underlying program theory to learn what works and why, and also at what cost.
12-1-10	Robert Shea, Principal, Grant Thornton, and Shelley Rappaport, also of Grant Thornton	<p>Today I met with Robert Shea, former Associate Director of OMB for Administration and Government Performance, widely known also as the father of PART. This meeting was by invitation, facilitated by his assistant Shelley Rappaport, who was in the audience at the AEA conference session on evaluation in the Obama Administration. She suggested I meet with Mr. Shea, primarily because of his past and continuing interests in government evaluation.</p> <p>Mr. Shea continues to have a strong interest in evaluation, some stemming from his past position at OMB and some stemming from his new responsibilities in his new job. I emphasized to him that that AEA is a non-partisan association, and I saw some benefit of letting him know about AEA activities given the fact that he is still called upon for advice and to comment on matters of government performance. I gave him a copy of the Evaluation Roadmap, and suggested that he read it and that if he is called upon for</p>

		<p>advice on matters related to evaluation he might want to consider ideas described in that document. One of his suggestions was a closer collaboration between AEA and the National Academy of Public Administration, as he sees matters of common interest between the two organizations. I told him I would pass that advice along to the EPTF for its consideration.</p> <p>I thanked him for receiving Bill Trochim and myself, and told him that that led to continuing contacts with OMB staff.</p> <p>As a matter of full disclosure, we also met to talk about possible collaborations on the private enterprise side of things</p>
9-17-10	Audry Yowell, Maternal and Child Health program	I talked with Audry Yowell of the Maternal and Child Health program about AEA comments on the Maternal and Child Health Home Visiting program. She said that they had received 183 comments, some of the quite long. The comments are being reviewed by a committee, and they are taking them all very seriously.
9-15, 16-10	Sharon Stout, Department of Education Evaluation Office	Sharon Stout and I exchanged views on an emerging Senate Budget Committee draft bill, GPRA Modernization Act of 2010. This bill, not yet introduced, is the Senate counterpart to a similar House bill introduced by Representative Cuellar . Both would codify some of the performance improvement functions introduced during the Bush and Obama administrations (e.g. portions of President Bush’s PART and President Obama’s Performance Improvement proposals). We will follow this closely.
8-16-10	Dan Levinson, IG at HHS, Kathy Buller, IG at Peace Crops; Stuart Wright, Dep. IG for Evaluation and Inspections at HHS; Elise Stein, HHS Coordinator for Council of IG’s for Integrity and Efficiency	Patrick Grasso and I met today with Dan Levinson, HHS Inspector General, Kathy Buller, Peace Crops Inspector General; Stuart Wright, Dep. IG for Evaluation and Inspections at HHS; and Elise Stein, HHS Coordinator for the Council of Inspectors General on Integrity and Efficiency http://www.ignet.gov/pande/cigiecharter061909.pdf . The Council is established by the Inspector General Reform Act of 2008. It is a council of all Federal Government Inspectors General. Among its responsibilities is to increase the professionalism and effectiveness of personnel by developing policies, standards, and approaches to aid in the establishment of a well trained and highly skilled workforce in the Offices of Inspectors General. Dan Levinson is the Chair

		<p>and Kathy Buller is the Vice Chair of the Inspection and Evaluation Committee of the CIGIE. Patrick Grasso and I met with them and senior staff of the HHS IG to introduce them to AEA.</p> <p>Our purpose was to promote membership in AEA by the inspectors and evaluators within the entire IG community. Mr. Levinson commented that he intends to join, and Stuart Wright indicated that he will follow up with us to present information about AEA to the Inspection and Evaluation Round Table, a group that represents the practicing evaluators within the IG community and serves as staff support for the CIGIE Evaluation Committee.</p>
5-21-10	Jonathan Breul, Executive Director, IBM Center for the Business of Government	Jonathan Breul and I discussed options and tactics for improving HR 2142, Government Efficiency, Effectiveness, and Performance Improvement Act of 2010, which just passed the House Oversight and Government Reform Committee. The goal is to get some recognition of the role of evaluation in the performance management bill.
5-18-10	Stuart Wright, HHS Deputy Inspector General for Evaluation and Inspections	<p>I spoke with Stuart Wright today about the scope of the HHS OIG's responsibilities under the health care reform bill. My goal was to make sure he understood that the OIG is responsible for the oversight of all matters pertaining to health insurance reforms as they pertain to HHS. I sent him a copy of the legislation and an excerpt of the broad OIG authority. While he was not specifically aware of this provision he assured me that he had already directed his staff to include the major aspects of health reform legislation in preparing the next evaluation plan.</p> <p>On another matter, Stuart suggested to me that it might be worthwhile to meet with Dan Levinson, HHS Inspector General. Mr. Levinson is the Chairman of the Evaluation Committee of the Council of Inspectors General. He has indicated his interest in promoting the professionalism of the evaluation function throughout the evaluation community and may be amenable to some kind of interaction with AEA to bring this about. I indicated that this sounded like a good idea to me and that I would see what I could do to follow up on that idea.</p>
1-5-10	Shelley Metzenbaum, Associate Director of OMB for Performance and Personnel Management; Katherine Stack, OMB Budget Director for Education; Debra Rog,	<p>Summary Debra Rog, Patrick Grasso, and George Grob met with Shelley Metzenbaum and Katherine Stack of OMB to discuss priorities and opportunities for improving the linkage between evaluation and Federal program management. Shelley is the most senior OMB Executive directly responsible for evaluation and performance management, and Katherine Stack is a long time senior executive</p>

	<p>Immediate Past President of AEA; Patrick Grasso, Chair of AEA's Evaluation Policy Task Force</p>	<p>responsible for budget formulation in the field of education. Both expressed a strong interest in collaborating with AEA in promoting a more effective connection between evaluation and improved performance of Federal programs. Shelley put forward a number of ideas and concepts, including the notion of more effectively lining up performance measures with evaluation. She sees performance indicators as providing data for deeper analysis of the factors underlying program performance. She is particularly interested in identifying programs that appear on the surface to be successful, identifying performance standards that seem to be related to specific program goals, and then developing evaluations that can be used to determine if the programs are meeting those goals and whether the underlying theories supporting the goals are valid. She wants to help agencies understand that performance can indeed be measured and linked with evaluation.</p> <p>She is hoping that AEA could promote these ideas and help provide sources of training to make substantial improvements along these lines. She is also looking more generally for AEA to identify effective practices and experts that can advise OMB and Federal agencies on how to make such improvements. She also asked if there was some way to involve the OMB budget examiners in the evaluation training (presumably both on the receiving and giving ends).</p> <p>Katherine Stack spoke about the "Invest In Innovation Program" (also known as the 3 I's), active now in the Department of Education through \$650 million provided by the Recovery Act. It provides a tiered set of funding for programs and associated evaluations that are designed to methodically promote 1) new, innovative approaches to improving education, 2) validation of current initiatives, 3) scaling up current programs with strong evidence of success. She remarked about the need to identify the most effect and appropriate methods for evaluating each type of program.</p> <p>Both Shelley and Katherine also promoted the ideas of encouraging problem solving networks, quality control for evaluations, and peer review of proposed methods.</p> <p>They asked if AEA could provide expert advice and training sources to promote these idea.</p> <p>We responded that AEA could provide a bridge between OMB and the evaluation community, invited Shelley to speak at the 2010 AEA conference in San Antonio, offered her the opportunity to write an article for publication in the AEA Newsletter, and promised to get back to her with ideas about</p>
--	---	---

		<p>how to provide venues to promote better use of evaluation to improve Federal programs.</p> <p>Shelley pressed us to meet again in about three weeks. We promised to meet soon in the future after consulting with the AEA Evaluation Policy Task Force.</p>
11-02-09	Michael Coplen	<p>Mr. Coplen, senior evaluator at Federal Railroad Administration, called to discuss OMB's new policy emphasizing the important of evaluation. He requested assistance in understanding OMB's new voluntary program for funding impact evaluations. I gave him background on the new program and volunteered to review any application he might prepare for funding from OMB. We also discussed evaluation policy more broadly, including his interest in wok on Science of Science policy. We agreed to meet at the AEA conference to share ideas about evaluation in the Federal Government.</p>
11-17-09	House Budget Committee	<p>I attended the House Budget Committees' first hearing on performance management. Testimony was given by the Honorable Jeffrey D. Zients, Deputy Director for Management and Chief Performance Officer, Office of Management and Budget. His testimony is available at The Honorable Jeffrey D. Zients (House Hearing Testimony) . This followed along the same lines as his testimony before the Senate Budget Committee on October 10-28-09 (See below)</p> <p>After the hearing I again spoke to Shelley Metzenbaum, Associate Director of OMB for Performance and Personnel Management. She requested that I identify dates in December to meet with her. I will follow up on this after the Thanksgiving Day break</p> <p>I also chatted briefly with Morna Miller, staff coordinator for the House Budget Committee and promised to follow up with them with more information about evaluation. I also reconnected with Amy Edwards of the Senate Budget Committee, who was also present at the hearing.</p>
10-29-09	Senate Budget Committee	<p>I attended the Senate Budget Committee's first hearing of its new Task Force on Government Performance. Testimony was given by</p> <p>The Honorable Jeffrey D. Zients -- Deputy Director for Management and Chief Performance Officer, Office of Management and Budget</p> <p>Sir Michael Barber -- Partner, McKinsey & Company/ Former Chief Advisor on Delivery to British Prime Minister Tony Blair</p> <p>☐ Dr. Paul L. Posner -- Director, Master's in Public Administration Program, George Mason University</p> <p>Before the hearing I met briefly with Shelley Metzenbaum,</p>

		Associate Director of OMB for Performance and Personnel Management. I will follow up with shortly to inform her of EPTF activities and products
10-23-09	Amy Edwards, Senate Budget Committee	<p>Today I spoke with Amy Edwards, a staff member on the Senate Budget Committee who is working on the newly established Task Force on Government Performance. I told her about the work of the Evaluation Policy Task Force and the Roadmap. I will be sending that to her this afternoon.</p> <p>We talked about the important role of evaluation in effective government. Our discussions circled around performance measurement, impact evaluation, and the larger body of evaluation studies needed throughout the life of a program. She emphasized that the Budget Committee was particularly interested in finding examples of where evaluations have led to savings. I told her about documented cases from my old IG office. While I tried to move the discussion to a broader field, she was clearly interested in the savings aspects of things. I believe that many of the people she will be talking with will also try broader out the discussion include the concept of value, not just savings.</p> <p>I will be attending the first hearing of the Task Force at 10:00 AM in the Budget Committee's hearing room, 608 Dirksen Senate Office Building. Among those testifying will be Jeffrey Zients of OMB and Michael Barber of the United Kingdom. Mr. Barber was an advisor to Tony Blair on his government performance initiatives. I will do my best to meet key players while I am there.</p>
9/14/09	Ruth Gillett, Cleveland/Cuyahagoa Office of Homeless Services	<p>I spoke with Ruth Gillett of the Cleveland/Cuyahago Office of Homeless Services. This is the office that turned down the grant funding request of Cleveland's Homeless Congress on the basis that Federal regulations do not allow for funding of evaluation projects other than to collect data and to assist HUD with its evaluation if HUD requests a grantee to do so. Based on information she provided me, I checked the application requirements for grants and discovered that she had correctly interpreted HUD's guidance. HUD is running a central evaluation, comparing the effectiveness of various approaches to helping the homeless.</p>
9/15/09	Andy Schneider, House Energy and Commerce Committee	<p>I talked by telephone today with Andy Schneider, staff on the House Energy and Commerce Committee. He is one of the staff members working on health care reform. I explained to him our concerns about the shortcomings of the evaluation of the insurance reform provisions of all the current bills. Once I explained it to him he seemed interested in helping us get something into the bills. He asked me to send him a mark-up of HR 3200 with the most important changes that can be made with the least amount of drafting. We focused on the</p>

		inclusion of important evaluation questions that can be specified as part of the mission of the Special IG that would be established to oversee the health insurance “exchange.” I also mentioned to him the efficacy of the Senate HELP Committee’s bill’s provision for an annual report on the nation’s health care system. I suggested that if the language were tweaked to include mention of the adequacy of health insurance coverage as one of the topics to be included in the annual report, then we could push for the House bill to accede to Senate on that matter. We discussed other possible changes as well. He sounded ready to help during the work of the House and Senate conference committees’ actions on the current bills.
7/21/09	Emilia Disanto, Senate Finance Committee	I met with Emilia Disanto of the Senate Finance Committee in her Washington, DC, office to go over with her what I thought were the shortcomings of evaluation provisions in the several health care reform bills currently circulating. She had previously provided our written comments to the staff person on the Senate Finance Committee who is responsible for actually drafting the committee’s health care bill. She was sympathetic with the points I was raising. I asked for her advice on how I could get other hill staffers to pay attention to our concerns. She will make referrals to two other staffers that I can follow up on—one on the House Energy and Commerce Committee and one on the Senate HELP Committee.
8/6/09	Matt Forti of Bridgespan, Susan Kistler, Patrick Grasso	Susan Kistler, Patrick Grasso, and I met with Matt Forti of the Bridgespan Group, a non-profit group specializing in social impact. The firm has been hired by the Council on Foundations to provide advice on how to identify effective social programs funded through foundation grants that are effective and have the potential for greater impact if they are scaled up and expanded. This advice will likely be provided to the Social Innovation Fund established under the Federal Corporation for National Community Service. In response to their questions, we advised them to avoid establishing a set of best methods for rigorously evaluating the impact of social programs, instead promoting the idea that evaluation should be carried out during the life of programs and that many different methods are available to fit the needs of the evaluation issues at play. We referred them to resources on the AEA website, including the AEA Guiding Principles for Evaluators, the Evaluation Roadmap, and links to other practice standards.
8/1/09	Richard Blue, Cynthia Wincek, and Holly	I met with Richard Blue, Cynthia Wincek, and Holly Benner for State Department employees, now consultants, who

	Benner	<p>prepared a proposal for an independent Center for Monitoring and Evaluation (CME) of Foreign Assistance programs, reporting to the Office of the Secretary of State or the Deputy Secretary of State for Management and Resources. According to their proposal, the Center would be placed within the Secretary or Deputy Secretary's Office to ensure M&E efforts become a central feature of foreign assistance decision-making. It would support the development of effective monitoring and evaluation systems and practices in all USG agencies responsible for foreign assistance, while undertaking policy relevant comparative evaluations on major programs that involve multiple USG agencies. The Center would reinforce a 'learning culture' for U.S. foreign assistance and demonstrate that USG leadership values the importance of good monitoring and evaluation, as well as the time, financial, and human resource requirements necessary for improving the effectiveness of U.S. foreign assistance investments. A "center" of competence, expertise, and experience could also inculcate values, provide guidance, and insure high quality work from all implementing agencies.</p> <p>We discussed the Center proposal as well as the current activities of the EPTF in the area of foreign assistance. We agree to keep one another informed of developments in this field and to support one another whenever practical.</p>
7/30/09	Jim Grenier of OPM and Janet Griffith of ICF	<p>I met with Jim Grenier of the U.S. Office of Personnel Management (OPM) and Janet Griffith of ICF International at their request to provide advice on an evaluation initiative currently under development within the OPM. They brought materials outlining their first efforts at developing a broad policy for evaluating all program areas of OPM, including evaluating the effects of these programs government-wide. Their proposals were based on extensive interviews with OPM officials and stakeholder groups. I provided advice to them on defining the scope of their evaluation plans, on key definitions relating to program results, and on standard ways to approach evaluation designs.</p> <p>This initiative, if it coalesces and is adopted, will be a significant improvement in evaluation at this OPM and may provide models that can be adopted in other agencies. We agreed to stay in touch.</p>
7/30/09	Daren Wong, Office of Management and	Daren called with inside information about OMB efforts to promote evaluation. OMB may propose a combined

	Budget	leadership role of OMB and the Council of Economic Advisors. Daren requested that at the upcoming meeting with Jeffery Liebman we emphasize the need to institutionalize evaluation in agencies and build up their evaluation capacity. He suggested we use examples from CDC, ACF, and USAID.
7/23/09	Dustin Brown, Deputy Assistant OMB Director for Management	I talked to Dustin Brown to get advice on next steps for reconnecting with OMB staff on evaluation. He suggested we meet with Jeffrey Liebman, Executive Associate Director and Associate Director for Economic Policy. He said that Mr. Liebman is leading the way at OMB on evaluation, is quite interested in it, and meets regularly with OMB on this subject. He asked me to send him some times when we would be available to meet with Jeffrey Liebman and he will set it up, including inviting other senior staff at OMB interested in this topic. He also thought that inviting Jeffrey Liebman to speak at the AEA conference was a really good idea.
6/12/09	Daren Wong, Office of Management and Budget	<p>I talked to Daren Wong of OMB today about a possible meeting between Bill Trochim and Jeffrey Zientz, the nominee for the position of Deputy Director of OMB for Management and the Chief Performance Officer. OMB staff provided Bill's name to Mr. Zientz in response to a request for names of individuals to talk to before he takes office. We made arrangements for Mr. Zientz to reach Bill should he wish to talk to him.</p> <p>We also discussed the desirability of Bill's responding to the recent blog of Director of OMB Peter Orzag. While generally supportive of evaluation in Federal decision making, the blog also raised potential issues about over reliance on "evidence based," evaluations and "top tier" methodologies. I told him that I would bring up the idea of such a response to Bill and the EPTF.</p>
5/20/09	Daren Wong, Office of Management and Budget	<p>I called Daren Wong of OMB to run by him our idea of inviting Jeffry Zients or some other senior OMB official to speak at the AEA conference in November. He confirmed that we should wait until Mr. Zients is confirmed, but then be ready to move quickly to invite him. He thought such an invitation would be welcomed. I had previously sent an email about this to both Daren and to Dustin Brown, Deputy Assistant Director of OMB for Management.</p> <p>Daren and I caught up on the situation at OMB. He said that they expect intensive work to be done on management policy as soon as Mr. Zients arrives and that the idea of promoting a broader role for evaluation is still very much alive within current OMB staff.</p>
7/21/09	Vic Dukay, Lundy Foundation	I talked to Vic Dukay of the Lundy Foundation to go over with him my concerns about some of the provisions in the early (as yet not introduced) Senate draft bill reauthorizing the Foreign Assistance Act. The version that just passed the

		House is completely to our liking, but the Senate version places undue emphasis on impact evaluation at the expense of evaluation of inputs, outputs, and outcomes. I wanted to see if he agrees with my concerns. He does. We agreed that I would write to our foreign assistance group describing the provisions and our concerns about them and then to seek a meeting with Senate staff before the bill is introduced.
6/17/09	Matt Oresman, Patton and Boggs	<p>Matt Oresman I and compared strategies for influencing the State Department on matters related evaluation. He is going to take a slow, opened ended, non-advocacy approach with regard to foreign assistance. He will start at junior levels within the office of Jack Lew, Deputy Secretary of State. He will not push now on the Berman bill since the State Department is the source of the Administration's objection to the bill. His idea is to simply make contact, express interest in the views of the State Department, and offer to help when the time is right.</p> <p>I told him that we would try to approach senior people at the State Department about our Evaluation Roadmap in the wake of their evaluation conference. Our approach would be to garner support for the Roadmap not just on foreign assistance, but on all State Department activities, including diplomacy and administration.</p> <p>We both agreed to keep each other informed of our contacts and progress.</p>
5/29/09	Yvette Fontenot, Senate Finance Committee staff	Today I talked with Yvette Fontenot, primary staff contact on health reform on the Senate Finance Committee. She called me back after my leaving a message on her phone earlier this week about the building evaluation into the health reform legislation. I promised to send her a written explanation of our proposals and offered to meet with her or her staff if she is interested.
5/26/09	Andrew Dawson and Marci Harris, congressional staff	Today I talked by phone with Andrew Dawson of the House Ways and Means Health Subcommittee and Marci Harris, staff to Rep. Fortney "Pete" Stark, Chair of the Health Subcommittee, about building evaluation into the health care reform legislation. Both were at least politely interested and asked me to send them more information.
5-13-09	Jeremiah Rigsby	Today I met in person with Jeremiah Rigsby of Rep. Henry Cuellar's staff (see introductory information below). I provided him the Roadmap, the evaluation "leave behinds," AEA Guiding Principles, and copies of the Berman Foreign Assistance Strategic Planning bill, the PEPFAR evaluation responsibility table, and President Obama's analytic budget supplement on achieving a high performance government (containing his proposals to improve PART and embrace evaluation.) The purpose was to give him material that shows a growing support for evaluation as an essential function of government and to give him examples of how others are attempting to do this.

		<p>Our discussion focused on the strength of Rep. Cuellar’s draft bill insofar as it uses program assessments and monitoring as ways to improve programs as well as to make budget decisions about them. I described an even broader vision of evaluation as a tool to use in various ways throughout the life of a program.</p> <p>He wants to stay in touch and talked about drawing me in to reauthorizations of homeland security and other programs that are under the purview of the House Oversight and government reform Committee.</p>
5-8-09	Jeremiah Rigsby, staff member to Rep. Henry Cuellar	<p>Today I made contact with Jeremiah Rigsby, on the staff of Rep. Henry Cuellar (D) (Texas), one of the two sponsors of H.R. 2142, the Government Efficiency, Effectiveness, and Performance Improvement Act. I will be meeting with him on Wednesday, May 13, at 4:00 PM. I also tried to reach Adam Miles, on the staff of Rep. Dennis Moor (D) (Kansas), the other sponsor. Adam was not available and will not be till Tuesday, but he may be at my meeting with Jeremiah. The staff are hoping to get one or more hearing’s on this bill and are looking for advice on content. I will share the Roadmap paper with them and single out features of the bill that are compatible with it. I will also highlight other aspects of the Roadmap that they may want to consider and will offer to be available to provide technical assistance as the bill wends its way through Congress, if it does.</p> <p>After the meeting I will formulate advice on whether the EPTF should send written comments to the two sponsors and what the content of such advice might be.</p>
3/16/09	Vic Dukay, Matt Oresman, Michele Orza	<p>The purpose of this meeting was to review the comments received on a draft proposed section for a proposed recodification of the foreign Assistance Act as mentioned in the summary of our meeting of 2/27/09 discussed below. As a result of that meeting, Patrick Grasso, Michele Orza, and I prepared a draft set of legislative specifications and explanations thereof for review by Vic Dukay and Matt Oresman. Vic and Matt provided their comments, as well as those of several of Vic’s colleagues, to us along with draft legislative language to include in the package for the House Foreign Affairs committee. In our meeting today, we discussed issues that emerged from the first round of drafting. These primarily involving the definition of evaluation, the need for an interagency Evaluation Advisory Group to facilitate coordination of evaluation activities by various Federal agencies and to prepare a report to the Congress on the evaluation of Federal foreign assistance programs as a body every three to five years. We agreed to prepare such a revised set of specifications and explanations</p>

		by the end of this week.
3/9/09	Nicole Cathcart, The Performance Institute	<p>Nicole Cathcart of The Performance Institute asked to meet with me (by phone) to discuss how the Institute could be helpful in getting the word out about the Roadmap paper. She thought it would be of considerable interest to their clients. She plans to email all of the Institute’s clients telling them about the Roadmap paper and sending them the link on the EPTF page of the website. I told her that she could invite readers to email me if they have questions. I also took the opportunity to explain to her what the EPTF is doing and to describe other products, such as the comments to OMB on what constitute strong evidence.</p> <p>The Performance Institute offers a number of training conferences during the year with attendance by Federal Government staff and others. In the past, it has worked closely with OMB in promoting and supporting the movement toward greater use of performance management through GPRA and PART. It also is represented on the Government Performance Coalition, a non-profit organization that meets regularly to promote government performance management.</p> <p>Nicole says that the The Performance Institute is beginning to move in the direction of promoting evaluation and auditing. She invited me to make a presentation at an “Innovations in Government Summit” to be held near the end of September. I agreed to be available for that.</p>
2/27/09	Vic Dukay, President of Lundy foundation, Mathew Oresman of Patton Boggs, and Patrick Grasso	<p>The purpose of the meeting was to plan next steps for promoting the funding of evaluation of the PEPFAR program. I presented a table outlining the various evaluation provisions in the new law, identifying who is responsible for each, and when action is required. Vic Dukay will share this table with his colleagues in the international field that promote care of orphans. A large group of them are now in DC planning their own strategies and meeting with Department of State staff.</p> <p>Mathew Oresman reported that an appropriations bill was just enacted with the supporting committee report containing language expressing the expectation that funds for evaluating PEPFAR programs be increased. The language was prepared on the basis of materials sent to the committee by the Lundy Foundation on which the EPTF had collaborated.</p> <p>Mathew also reported that he had been requested by staff of the House Foreign Affairs Committee to provide proposed language about evaluation for a bill the committee is planning to prepare to overhaul the Foreign Assistance Act.</p>

		Patrick Grasso and I agreed, working with Michele Orza, former Director of the Institute of Medicine (IOM) study of the implementation of PEPFAR, to provide draft language explaining and justifying such evaluation requirements.
2/6/09	Daren Wong	A high level group of OMB staff met to brainstorm the idea of promoting a stronger evaluation focus within the Federal Government. Among the attendees was Jeff Liebman, Executive Associate Director and Associate Director for Economic Policy, now also informally serving as in the role of Associate Director of OMB for Administration and Government Performance, previously Robert Shea's position. The attendees all had copies of our Roadmap paper. Daren was commissioned to prepare a paper on options for strengthening the evaluation function.
1/22/09	Daren Wong	Daren reported progress being made in setting up the review panels for an OMB pilot program on impact evaluation. Five panelists have been appointed, including Steward Karenski, Deputy Director, Institute of Education Science at the Dept. of Education; Naomi Goldstein, Director, Office of Planning, Research, and Evaluation at the Administration for Children and Families; Jerry Britten of USAID; Janet Curley at Peace Corps; and Ted Kiniker (agency unknown.) One impact evaluation of manpower programs proposed by the Department of Labor will be reviewed.
1/6/09	Daren Wong	<p>I called Daren Wong of OMB to compare notes about current and planned actions relating to promoting the institutionalization of evaluation in the Federal Government. The breaking news today was President Elect Obama's announcement of Nancy Killefer as Chief Performance Officer and also as Deputy Director of OMB for Management. The combination of duties will give her an extraordinarily strong role in the oversight of federal programs.</p> <p>Daren told me that she was briefed by Jeff Liebman for whom OMB staff had, at his request, prepared a paper on strengthening evaluation in the Federal government. That paper contained many of the principles that we articulate in our pending draft paper on institutionalizing evaluation, an early draft of which was sent to OMB staff.</p> <p>The sense at OMB is that the role of the management office will be expanded and strengthened. It is too soon to speculate about PART, but indications are that some refined version of it, perhaps with a broader and more flexible set of assessment tools, will be used.</p> <p>I shared with Daren our own plans for contacting leaders in both Legislative and Executive Branches with our ideas about evaluation. We agreed to stay in touch.</p>
12/03/08	Daren Wong	This was a follow-up to my having sent him, on 12/1 a draft

		of the EPTF paper on institutionalizing evaluation. He shared with me his own comments on the paper and showed me the paper that he had prepared for Jeff Liebman of the Obama transition team. His paper incorporated key concepts or our paper
12/03/08	Victor Dukay, Matthew Oresman, Michele Orza	<p>The purpose of this meeting was to discuss strategies for promoting increased funding for evaluation of the PEPFAR program. The attendees were Victor Dukay of the Lundy Foundation, Matthew Oresman of Patton-Boggs (doing pro-bono work for the Lundy Foundation), and Michele Orza, former Director of the Institute of Medicine (IOM) study of the implementation of PEPFAR.</p> <p>Now that the PEPFAR program has been reauthorized, funding for evaluation will be decided largely in the Executive Branch, especially the Office of the Global AIDS Coordinator within the State Department. We discussed ways to contact that office after the start of the Obama Administration. We also considered contacting other organizations that under the new law are mandated to perform evaluations of the program, such as IOM, GAO, and selected Inspectors General. They may be able to include in their evaluations an examination of implementation of the requirements in the law that all aspects of the PEPFAR program have evaluation built into them.</p>
11/20/08	Matthew Oresman	Matthew is with Patton-Boggs. He is continuing to work with the Lunday Foundation to promote adequate funding of the newly authorized PEPFAR program. He contacted me to arrange to meet in early December with himself and Victor Dukay, President of the Lunday Foundation, to strategize about how to approach State Department agencies involved with making budget decisions for PEPFAR. I agreed to meet with them and also, at his request, provided him with contact information for Michele Orza, the director for the evaluation report on PEPFAR that was prepared at the request of Congress by the Institute of Medicine.
11/20/08	Andrew Rock	Andrew Rock is a staff member of the Office of the Assistant Secretary for Planning and Evaluation (ASPE) at Department of Health and Human Services (HHS) He is responsible for preparing the Department's annual Performance Improvement Report. This congressionally mandated report summarizes and provides information on how to obtain copies of all evaluation reports prepared under auspices of the 2.4% evaluation set aside fund under the Public Health Act. He confirmed for me that all of the HHS evaluations funded under the set aside are required to be summarized in this report. The Performance Improvement Report also contains summaries of HHS evaluations funded under other authorities within the Department. The goal is to include as many as possible in this annual report. The report is

		available online through the Policy Information Center, a resource of evaluation information published by ASPE for public consumption.
11/17/08	Elise Stein	Elise Stein is a staffer on the Senate finance Committee working on health issues. I called her to get an update of the work of that committee, under the direction of its Chair, Senator Max Baucus, in drafting health reform legislation. At my request she gave me the contact information for the staff members working on this project. I hope to contact them shortly to share ideas about incorporating evaluation into the legislation. It is likely that these same people will be involved in drafting a version of health reform legislation in concert with the Obama administration.
11/6/08	Daren Wong	<p>This conversation was a follow up to my discussion with Dustin Brown of OMB on 9/26/08. See the description of that meeting below. Daren Wong works for Mr. Brown and is the person that he looks to for technical support on matters pertaining to PART and, more generally, evaluation. My telephone conversation today was part of my continuing engagement of OMB staff on matters related to promoting more effective evaluation in the Federal Government. Our efforts now are focused on connecting with the transition team of the Obama Administration.</p> <p>On 10/26 I sent a one page document on strengthening the use of evaluation in the Federal Government, along with our policy handouts, to both Dustin Brown and Daren Wong.</p> <p>In this conversation with Mr. Wong, he brought me up to speed on what OMB staff are considering with respect to proposals for evaluation policy in the Obama Administration.</p> <p>First, OMB staff believes that the Obama Administration will first focus its attention on the most pressing policy priorities such as the economy, the financial system, the Iraq and Afghanistan wars, health care, education, and the like. In the meantime, OMB staff is considering proposing management initiative to institutionalize evaluation, strengthening it throughout the Federal agencies. They are thinking of using the Evaluation Working Group of the Council of Performance Improvement Officers (recently established by Federal Executive Order) to develop the mechanics of such a policy. They may issue a formal request might be issued from OMB for actions to be taken by Federal agencies to strengthen their performance management, impact assessment, and general evaluation capabilities.</p> <p>Daren invited me to offer suggestions with greater specificity than we had previously sent to him on what might be included in such a proposal and what specific administrative</p>

		<p>arrangements should be made to carry it out. I promised him I would prepare such recommendations promptly.</p> <p>We also discussed the question of just how formal his request to me was. We agreed that his verbal request o me was sufficient for me to reply by email to him in his capacity as senior OMB staff working on evaluation issues and from my capacity as consultant to the EPTF. We both agreed that he would consult with Dustin Brown to see if Dustin would like to more formally request our views in an email.</p> <p>In this context we also discussed the possibility that AEA’s views could be sought more routinely in the future. We both agreed, however, that OMB would be better positioned to do so if their requests for outside opinions on such matters were to be made to several rather than just one outside organization. This would enhance the sense of inclusiveness and also prevent OMB from appearing to be tied too closely to the one set of views from AEA. It would also prevent AEA from being seen as “in the pocket” of OMB. We agreed that I would provide him with the names of other professional organizations from which OMB could solicit opinions on matters of management policies related to evaluation.</p> <p>On this latter point, I am not certain that OMB will have the option to single certain organizations out for advice. They normally must make such requests open to all comers through some kind of public notice. However, they might want to make sure that certain organizations receive the invitation.</p> <p>I also brought up the topic of OMB’s current pilot study involving agencies volunteering to have their impact evaluation plans reviewed by a committee involving all major sectors of OMB (budget, management policy, and office of information and regulatory affairs) as well as prominent and respected Federal evaluators. I suggested that OMB make public the fact that such a pilot study was underway, and also make the governing documents of the pilot available to the public. This would they could be referenced, something that would make it clear that OMB is very open to considering a wide variety of impact assessment methods. Daren said that he would consult with Dustin Brown and other s at OMB about making information about the pilot study public.</p>
10/17/08	Jonathan Breul	Jonathan Breul is Executive Director, IBM Center for The Business of Government, and Board Member of the Council for Excellence in Government. He is a former high level executive of OMB, serving both Republican and Democratic administrations. I called to discuss what he knew about Presidential candidates’ transition teams and the interests of

		<p>policy makers in the field of government management. (Jonathan was recently interviewed by Government Executive magazine about Senator McCain’s views on government reform and generally stays in touch with government management experts of both parties.) He believes that interest in the topic of government reform is picking up in both the political parties, but not in an organized way. It is difficult to know who to talk to at this time. He believes that connecting with senior OMB career staff is still the best tactic as no leading “management guru” has yet emerged in either party. He believes that the transition teams will not be the real centers of policy making. That will be the incoming political appointees who will move very quickly to press the agenda of the new President. Initially, the incoming administration will have to focus on very large issues such as the U.S. and world financial and banking systems and broad economic policies, Then they move on programmatic issues in such fields as health, education, etc. H is recommendation is to field a very short but hard hitting paper at the highest levels of the career executive corps, who will be the first communicators with the incoming appointees.</p>
10/9/2008	Michele Orza	<p>Michele Orza, now at the George Washington University National Health Policy Forum, was previously the Director of the IOM study evaluating the implementation of the PEPFAR program. I followed up with her today to get an update on PEPFAR, which was recently reauthorized with a five year authorization of \$48 billion, up considerably from the original authorization of \$15 billion. She provided assistance to me in locating key documents about the reauthorization and appropriation. I also received advice from her on the construction of a table to compare the recommendations that were made in the IOM report to the changes that were made to the program in the reauthorization bill. I intend to use her evaluation report as an example of an evaluation that had direct impact on program legislation and to use the PEPFAR authorization bill as a model for how to build evaluation into a program in ways that mirror the EPTF recommendations in its talking points.</p>
9/26/2008	Dustin Brown	<p>Dustin Brown is the Deputy Assistant Director of OMB for Management. I spoke to him about getting ready for the transition teams of the next administration, particularly for any such team that focuses on management of the Federal Government. It is likely that Dustin, the most senior career executive at OMB for management, will be the point of contact for such a transition time after the Presidential election. I informed him of AEA’s interest in seeing evaluation institutionalized in Federal programs. He was very receptive to the idea of an exchange of ideas on this subject and asked me to send him materials describing AEA’s</p>

		views. We also agreed to meet soon, along with other relevant officials, to discuss next steps.
9/23/2008	Daren Wong	Daren Wong is the principal staff person for routine contacts at the Office of Management and Budget on the Program Assessment Rating Tool (PART) and is Co-Chair of the Evaluation Working Group of the Performance Improvement Council. More generally, he has strong interests in advancing evaluation as an essential function of government. He reports to Dustin Brown, the most senior career civil servants in the chain of command at OMB dealing with management issues. I discussed with Daren the desirability of our approaching Dustin Brown now with our ideas of building evaluation into the management of Federal programs. I did this because I believe that Dustin is likely to be one of the initial points of contact at OMB for any presidential transition team dealing with general management issues. Hence, it might be beneficial if we could exchange ideas with him about tactics for making evaluation an important part of any management initiative of a new administration. His opinion was that such contact by us would be welcome.
9/12/2008	David Intracasso	David Intracasso is a healthcare analyst in the Marwood Group, a business and congressional advisory services company. David previously worked in the Office of Planning and Evaluation and the Administration for HealthCare Research and Quality in the U.S. Department of Health and Human Services and more recently for Senator Steny Hoyer (D, Md). He initiated the call. He is interested in promoting the use of evaluation in government programs, especially for health care reform. He wanted to know what the EPTF had been doing along these lines. I briefed him on our dealings with OMB and more generally about our efforts to have a positive impact on government evaluation policy. In turn I asked about his own work in this area. He outlined a general strategy to engage the Subcommittee on Government Management of the House Oversight and Government Reform Committee on the topic of GPRA and evaluation and possibly the House Ways and Means Health Subcommittee with respect to health care reform. We discussed other possibilities such as contacting the Senate Finance Committee with respect to health care reform and the Subcommittee on Oversight of Government Management, Federal Workforce, and the District of Columbia of the Senate Homeland Security and Governmental Affairs Committee on general evaluation policy. We agreed stay in touch.
8/26/08	Julie King	Julie King is a director with the executive search division of The Bridgespan Group, which has been retained to find candidates for Deputy Director, Impact Planning and Improvement, U.S. Initiatives, for the Bill & Melinda Gates Foundation.

		<p>The newly-formed Impact Planning and Improvement unit is an internal consultancy that sits at the intersection of the Foundation’s three primary areas of operation: U.S. Initiatives, Global Health, and Global Development providing support and strategic planning for measurement of short-term results and long-term impact, and ongoing learning and adaptation.</p> <p>Julie King had written to me for help in identifying candidates for the above position. I provided several names for her consideration. However, I also took advantage of this discussion to delve more deeply into the Gates Foundation’s policies regarding impact evaluation. I did this because I was intrigued by the following excerpt from her email to me:</p> <p>“The winning candidate will be capable of engaging all stakeholders, understanding the breadth of life cycle of a project, tracking progress and potential for success beyond randomized control trials (emphasis added), and be able to articulate strategy. He/she should have 15+ years in impact and evaluation and a broad knowledge of performance measurement tools.”</p> <p>It turns out that the Gates people, including Bill Gates, wanted to use more than randomized trials after their experience with that approach by the Center for Global Development to whom the Gates Foundation had given a fairly large grant for evaluation.</p> <p>I discussed with her at length the ideas of the EPTF regarding evaluation, especially impact evaluation, and sent her a copy of the EPTF advice to Michael Shea of OMB on the PART Guidance. I believe we may have an opportunity here connect directly with the Gates Foundation’s Impact Planning and Improvement, and perhaps through them promote a broader understanding and practice of evaluation in the field of international development.</p>
8/19/08	William Trochim Daren Wong Erin O’Keefe Katherine Dawes	<p>This telephone conference was a follow up to the meeting of 8/4/08 described below. Daren Wong, Katherine Dawes, and Erin O’Keefe briefed Bill Trochim and myself about the progress and next steps of the impact evaluation pilot program of the Performance Improvement Council’s Evaluating Working Group. The program described in the summary of the 8/4/08 meeting described below has been refined and will be presented to the Performance Improvement Council on 8/20/08. In brief, the final draft, to be used at that session, represents a balanced approach to</p>

		<p>impact evaluations performed by federal agencies. If the pilot is successfully carried out, it may lead to a substantive change in OMB's PART guidance along the lines recommended by the paper it sent to OMB. Daren Wong thanked Bill Trochim and the EPTF for its advice.</p> <p>Katherine Dawes is Co-Chair of the Evaluation Working Group. She has been asked to moderate the meetings of the Advisory Panel that will review the Federal Agency Impact Evaluation proposals.</p>
8/4/08	Daren Wong and Katherine Dawes	<p>I met today with Daren Wong of OMB and Katherine Dawes of EPA. Also present at the meeting was Erin O'Keefe, assistant to Daren Wong.</p> <p>Daren Wong is the principal staff person for routine contacts at the Office of Management and Budget on the Program Assessment Rating Tool (PART) and is Co-Chair of the Evaluation Working Group of the Performance Improvement Council. He represents OMB on the Working Group. Katherine Dawes is the Director of the Evaluation Support Division in the US Environmental Protection Agency. She is the other Co-Chair of the Evaluation Working Group, representing its Executive Agency members.</p> <p>The purpose of the meeting was for Daren and Katherine to brief me, as a representative of the EPTF, on OMB's planned pilot test of selected Federal agency impact evaluations. They wish to make another, more formal briefing to Bill Trochim and myself when the design of the pilot is completed but before it is implemented.</p> <p>In brief, the Evaluation Working Group hopes to sponsor a pilot test of impact evaluations proposed by a handful of Federal Agencies. These impact evaluations will be reviewed by a panel of experts, who will advise on the appropriateness of the overall approaches and the methods for the evaluations.</p> <p>The purpose of the pilot is to</p> <ul style="list-style-type: none"> • Promote strong evaluation • Promote capacity building for evaluation • Improve evaluation guidance issued by OMB for PART • Improve the programs themselves.

		<p>OMB plans to present the results of the pilot to the transition team of the next Administration.</p> <p>Following are some of the key administrative features of the pilot test:</p> <ul style="list-style-type: none">• Federal agencies will be invited to nominate their program impact evaluations for inclusion in the pilot test.• The evaluations that will be chosen to be included in the pilot will be limited to programs that directly help people (such as health, education, employment, and public assistance. Certain programs, such as research or science, technology, engineering and mathematics education, will not be included in the pilot because of the significant difference in their scope and purpose because they have not yet been subjects of impact evaluation as those related to direct services or general education.• A review panel of evaluation experts will be assigned to review the proposed evaluations. The experts will consist of experienced senior evaluators primarily from Federal agencies.• Each agency sponsoring an impact evaluation that is selected for the pilot test will nominate 5 reviewers from a panel of experts assembled for this purpose by Evaluation Working Group.• The Evaluation Working Group will then select 3 of the 5 experts nominated by the agency.• Each agency will submit written materials describing present its proposed evaluations to their expert review panel, and then meet with the panel to obtain advice.• The panel meeting will be attended by other interested parties, who may ask questions and raise concerns during the meeting. This will include, but not be limited to, representatives of three components of OMB—budget examiners, PART staff, and the Office of Information and Regulatory Affairs (OIRA). The latter office is responsible for review and approval of information collection proposals, such as the surveys that will be used in the impact evaluations. Thus, these OMB offices can each hear the comments and concerns that may be raised by the others. It is believed that their presence at the meeting will also encourage federal agencies to nominate impact evaluations for the pilot. These offices ordinarily provide their comments, assessments, and decisions regarding such evaluations sequentially and without the benefit of hearing one another’s concerns. Although their presence at the meeting will not substitute for their separate formal reviews, it is hoped that the pilot may
--	--	--

		<p>will promote a more consistent OMB response to evaluation proposals.</p> <ul style="list-style-type: none"> • After the meeting is concluded, the expert panel will prepare it advice. The advice will be non-binding. <p>To support the pilot, OMB</p> <ul style="list-style-type: none"> • Is using GAO’s evaluation guidance as the overall framework for its review. This has been translated into four questions related to <ul style="list-style-type: none"> --design adequacy and appropriateness --feasibility in terms and resources and schedules --data collection burdens <p>For example, the appropriateness of the design will deal with</p> <ul style="list-style-type: none"> --the purpose of the study --how the results will be used and disseminated, and the general relevance of the study. <ul style="list-style-type: none"> • Is developing a catalogue of evaluation designs with critiques of various methodologies.
7/17/08	Mathew Oresman	<p>I talked today with Mathew Oresman of Patton Boggs. He is handling congressional relations for the Lundy Foundation on the PEPFAR legislation. Yesterday the Senate passed its version of the PEPFAR reauthorization bill. Since the House had previously passed its bill, this sets the stage for a Senate/House conference to iron out differences and send a bill to the floor of each Chamber. The conference committee proceedings may give us an opportunity to persuade staff of both House and Senate of the need for the evaluation language we are requesting and the melding of the best evaluation sections of the House and Senate bills. However, there is a strong possibility that the House will simply cede to the Senate, removing the conference committee as a vehicle for change. We continue to watch this situation closely.</p>
6/11/08	Connie Chang	<p>Connie Chang, Director of Federal Services for OCEAN TOMO, was one of the organizers of the WREN conference, discussed below. She wanted to express her strong support for what the EPTF is doing for evaluation but also to emphasize the need to institutionalize effective evaluation practices through written works and websites. She suggested that formal texts from the AEA describing how the various types of evaluations could be most appropriately and effectively be used in different settings would be valuable. She had developed such a “tool kit” of evaluation methods for the Advanced Technology Program of the Commerce Department. She suggested that the AEA fund such a project, but on a broader scale.</p>
6/6/08	Michele Orza	<p>I met with Michele Orza (see notes below on my previous introductory discussion with her on 6/4/08), bringing her</p>

		copies of our proposed edits to the PEPFAR authorization bills. She agreed with the proposed changes and agreed to discuss them with her contacts in the Senate authorization committee. Unlike the House bill, which has passed the full House, the Senate bill has not yet been reported out of the authorization committee. Hence if the committee agrees to our edits, then they can be negotiated for inclusion in the floor bill during the House-Senate conference committee after passage of the Senate bill. I discussed this strategy with Matt Oresman of Patton Boggs (see notes about our discussion on 4/7/08) and shared copies with Dr. Victor Dukay of the Lundy foundation, who has agreed to our approach.
6/6/08	WREN	I attended the Spring 2008 Interactive workshop of the Washington Research Evaluation Network (WREN) at the George Washington University Marvin Center. Bill Trochim, Susan Kistler, Stephanie Shipman, and Katherine Dawes were also there. Bill made two presentations—one on the EPTF and one on evaluating R&D programs as part of a broader panel on the current state of prospective evaluation methodologies. Stephanie made a presentation as part of the opening plenary session. She discussed how Federal R&D agencies responds to PART and GPRA, and GAO's advice to OMB and the Congress. I chaired a break-out session with a small number of WREN members who had questions about the EPTF initiative.
6/4/08	Michele Orza	This was an introductory conference call with Michele Orza, now at the George Washington University National Health Policy Forum, but previously the Director of the IOM study evaluating the implementation of the PEPFAR program. I filled her in on our efforts to revise the evaluation language in the draft PEPFAR authorization and appropriations bills. She was in agreement with our objectives and generally with the nature of our proposed language. I promised to send her our proposed edits for her review.
4/7/08	Mathew Oresman of Patton Boggs	Discussed more details about legislative strategy for PEPFAR, as well what is known about the payoff for performing impact evaluations. The legislative strategy now is to combine the best provisions of House and senate versions of the reauthorization bill and to seek language in the Appropriations Conference Committee report describing the committees expectations for substantial and appropriate allocations of funds for evaluation (rather than putting a dollar amount in the appropriations language itself.)
4/2/08	Dr. Madeleine Wallace, Acting Chief, Evaluation and Systemic Assessments Branch	Discussed evaluation set-asides in NIH. According to Dr. Wallace, in FY 2006 NIH contributed \$646 million, as its share of the 2.4% evaluation set-aside for HHS programs. However, NIH was allotted only \$15 million. The rest of their share was allotted to the Agency for Health Care Research and Quality, the Office of the Assistant Secretary for Planning

		and Evaluation, the Centers for Disease Control and Prevention, the Substance Abuse and Mental Health Services Administration, the Health Resources and Services Administration, the Administration for Families, and the Office of the Secretary.
4/2/08	Nanci B. Csutsi, Director of Evaluation Colorado Trust	Dr. Victor Dukay of the Lundy Foundation had stated that he had heard that the Colorado Trust set aside 10% of its program funds for evaluation. I discussed this with Nanci Csutsi, the Director of Evaluation for the Trust. She said that the trust is a strong supporter of evaluation funds and uses 2% to 30% of its program funds for evaluation, probably averaging 20%. However, the funds are not set aside within the grants themselves, but rather are pooled in her office to conduct evaluations of national scope on important issues connected with each family of grants that they fund. She says that some foundations are increasingly dedicating money to evaluation and that she would send me a list of people in such foundations after we obtain the results of the survey being conducted by the Grantmakers for Effective Organization.
3/31/08	Gerald Britain of USAID	Over a long career at USAID, Mr. Britain has been involved in policies relating to evaluation of USAID programs. Donna Stauffer (see interview notes below) referred me to him. He confirmed what she had said, but pointed me to the written policy in Section 203, Assessing and Learning, of the USAID Automated Directive Systems. It suggests, but does not require, using 3-10 percent of program funds for performance management (a GPRA-like approach that establishes performance indicators for USAID strategic objectives) and that at least one assessment (defined to be equivalent to evaluation) should be performed during the life of such a strategic objective. The assessment could range from gathering stakeholder views to performing sophisticated impact studies.
3/31/08	Donna Stauffer, Office of Director of Foreign Assistance, USAID	She explained the historical background of evaluation of USAID programs, stating that while programs were encouraged to do evaluate programs, there has been no mandate to do so, and no evaluation set -aside.
3/25/08	Laura Leviton of Robert Wood Johnson Foundation	Discussed evaluation set-asides. Robert Wood Johnson Foundation has none. She expressed her own opinion that a simple set-aside would not be a good idea since the circumstances surrounding international programs are so varied, not only in terms of what evaluation methodologies would be appropriate for each, but also the economics of performing evaluations in developing countries. She agreed, though, to send a request for such information to the membership of Grantmakers for Effective Organizations . I subsequently sent her an email on this topic, which she forwarded to that organization. We received 17 responses. Some of these simply expressed interest in or emphasized

		<p>the importance of the question. Others, though, provided information about how they do in fact support this policy. The level of support varied considerably, and many expressed the notion that there can be no simple rate or formula for deciding how much to spend on evaluation</p>
3/13/08	<p>Alan Ginsberg, Director, Office of Policy and Program Studies, Department of Education, Department of Education Jaye Noell (same office)</p>	<p>Discussed evaluation activities within the Department of Education and the current atmosphere relating to requirements for using experimental methods. The House authorization committee prescribes evaluation agendas (known as “national assessments”) for programs in reauthorization legislation. These are studies that the committee asks to be completed before hearings on the next reauthorization.</p> <p>Requirements for using RCT’s are now leading to the use of other evaluation approaches. This is because the RCT movement gave rise to calls for more evidence about program impact; RCT’s cannot always deliver such evidence, due to cost and methodological limitations; hence there is a demand for other evaluation methods to be used, including more traditional service and outcome studies.</p> <p>The Institute for Education Science (IES) is looking for new approaches to use in areas that are lacking strong evaluations. IES is now using consensus panels to provide expert judgment to assess the evidence used in evaluation studies.</p> <p>The Office of Policy and Program Studies is now making extensive use of student test scores (which are being accumulated from the No Child Left Behind tests) as a data source to test program effectiveness.</p> <p>This office is also translating evaluation results into practical guidance for school systems that want to improve their programs.</p> <p>Alan suggested a need for evaluation standards related to the methodological practice of evaluation, such as considering the need for logic models and examination of unintended consequences.</p>
2/27/08	<p>Debra Duran and Madeleine Wallace of NIH William Trochim</p>	<p>Exchanged information and concepts about EPTF and NIH evaluation policies; introduced to NIH the idea of a collaborative policy initiative involving large research programs at NIH. Agreed to meet again to discuss more specifics about such an initiative.</p>
2/26/08	<p>Robert Shea of OMB William Trochim</p>	<p>A “getting acquainted” meeting with exchange of information, ideas, and advice about AEA and OMB PART. Mr. Shea invited EPTF to submit comments on OMB PART guidance, to advise the Evaluation committee of the Program Improvement Council, and to meet with other key OMB staff.</p>
2/25/08	<p>Henry Aaron, Ph.D. of the Brookings Institute</p>	<p>Discussed possible approaches to building evaluation into major health care reform initiatives of the next Administration. Mr. Aaron suggested several ways to connect</p>

		with incoming administration officials, including several promising references for future discussions.
2/25/08	Andy Rock of the Office of Assistant Secretary for Planning and Evaluation (ASPE) at HHS	Discussed ASPE's project to assess the influence of evaluations produced at HHS. Mr. Rock provided detailed information about the approach to the project and its schedule. He agreed to stay in touch keep AEA up to date on this project.
2/7/08	Federal Evaluators in Washington DC. Federal evaluators from CRS, DOE, DOJ, ED, EPA, GAO, HHS, HUD, State, Treasury, and VA were in attendance. AEA was represented by William Trochim, Susan Kistler, Stephanie Shipman, Katherine Dawes, Damon Thompson, and George Grob	This was primarily a listening session for EPTF to hear about the concerns of Federal evaluators. See attached minutes. Parties agreed to stay in touch. Further meetings with evaluators working in research setting may be scheduled.
1/24/08	Heather Foster, legislative assistant to Rep. Diana DeGette of Colorado, Deputy Whip and Vice Chair of the Energy and Commerce Committee	Purpose was to introduce the EPTF interest in evaluation policy and in building evaluation into legislation, particularly health reform legislation should it be introduced in the future.
1/24/08	James Scanlon, Deputy Assistant Secretary for Planning and Evaluation (ASPE) (Science and Data Policy), HHS	This was a listening session to hear perspectives and needs of evaluators from a senior official responsible for health research policy. Agreed to stay in touch. He provided names of other senior evaluators to talk to.
1/24/08	Barbara Broman, Deputy to the Deputy Assistant Secretary for Planning and Evaluation (ASPE) (Human Services Policy), HHS	This was a listening session to hear perspectives and needs of evaluators from a senior official responsible for human development policy. She suggested that AEA could effectively help senior career evaluators by sharing "best practices" on evaluation through seminars and other communications.
12/10/07	Laura Fuller, senior staff member of the Senate HELP	Discussed the importance of evaluation from the perspective of a senior committee staff members and sought advice on the logistics of accessing key congressional staff members.

	Committee	She said congressional staff are quite interested in evaluation but don't know how to incorporate it into legislation. She suggested that AEA hold short training briefings for congressional staff on this subject. She also provided helpful advice on keeping up with congressional actions and resources for locating congressional contacts.
12/7/07	Emilia Dissanto, senior staff member of Senate Finance Committee	She felt that right now there is not a lot of interest in the committees on the use of evaluation, but that previous efforts such as GPRA do demonstrate an institutional interest at least. She thought the idea of making evaluation an inherent part of major initiatives was an attractive one and should apply to general policies such as personnel, conflict of interest, contractor reforms, and competitive sourcing. She also provided tips on how to access key members of congressional staffs.
12/3/07	Jonathan Breul Executive Director, IBM Center for The Business of Government, and Board Member of the Council for Excellence in Government (Jonathan is a recipient of the AEA Myrdal Government Award)	Discussed a broad ranging set of issue surrounding the establishment of Federal evaluation policy. We covered his concerns about PART (similar to our own) and his interest in a stronger connection between GPRA and evaluation. He suggested contacting senior careerists at OMB and GAO as a way of influencing evaluation policy, since many of them will survive the transition to a new Administration.

EPTF PERFORMANCE SUMMARY APPENDIX N: AEA NEWSLETTER POLICY WATCH COLUMNS

01/09 – Policy Watch

As newly elected President Barak Obama makes his way to Washington, AEA's Evaluation Policy Task Force (EPTF) and I are watching closely for signals about his approaches to management and policy making. We believe that evaluation should be a central focus of such efforts and that evaluation should be recognized as an essential function of government.

During the last year, we established working connections with staff at the Office of Management and Budget (OMB). In a previous newsletter, 2008 AEA President Bill Trochim reported how, at OMB's request, we provided advice on how evaluation should be integrated into the government's Program Assessment Rating Tool (PART). We shared ideas about what to consider when deciding which methods to use in conducting program impact analyses and encouraged them to draw on broader expertise in the evaluation community to develop future guidance on evaluation for the PART program. For those of you who missed it, the EPTF response to OMB re: *What Constitutes Strong Evidence of a Program's Effectiveness* may be found on the EPTF website at <http://www.eval.org/EPTF.asp>.

We are now starting to reach out to others, working to connect with key incoming staff in the executive and legislative branches of the United States government. We have prepared a set of evaluation policy handouts describing evaluation and evaluators and what they can contribute to good government. These handouts explain the role that evaluation can play in

- Improving knowledge and understanding of how programs work
- Strengthening public accountability
- Assessing program effectiveness and efficiency, and
- Identifying opportunities and pathways to achieving objectives, outcomes, and efficiencies.

The handouts emphasize that evaluation is needed throughout the life cycle of programs from initial development, through implementation, and during re-authorization. You can read these papers as well on the EPTF website.

The EPTF and I will be contacting key staff who are active in the transition process to share these handouts and encourage further dialogue. Meanwhile, we are preparing materials that focus more on how to integrate evaluation with the development and management of federal programs. If you are interested in commenting on future draft documents and providing input on our plans as they evolve, please join the EPTF discussion group by signing up at <http://www.eval.org/EPTF.signup.discussion.asp>.

02/09 – An Evaluation Roadmap for a More Effective Government

On February 3, AEA sent to Peter Orszag, Director, Office of Management and Budget, our paper, "*An Evaluation Roadmap for a More Effective Government*." It was prepared by AEA's Evaluation Policy Task Force (EPTF). I am happy to share it with you, along with the cover letter signed by AEA President Debra Rog, President Elect Leslie Cooksy and William Trochim who is both Immediate Past President and EPTF Chair. The paper, along with the cover letter, is available on the EPTF webpage.

The paper was sent to OMB at a time when President Obama's key staff responsible for management policy are coming to grips with how transparency, accountability, and oversight of Federal programs will be institutionalized. Our paper is in their hands as they meet and deliberate on how best to do this.

The paper articulates principles and suggests concrete actions that can be taken to make evaluation an essential and abiding ingredient of effective governance. As stated in the cover letter, it describes "how evaluation can be used to improve the effectiveness and efficiency of Federal programs, assess which programs are working and which are not, and provide critical information needed for making difficult decisions about them."

It explains key principles related to the scope and coverage of evaluation, analytic approaches and methods, resources, professional competence, planning, dissemination of evaluation results, evaluation policies and procedures, and independence.

If implemented, the paper would move evaluation beyond the current practices of performance management and impact analysis that have characterized former and current accountability programs such as GPRA and PART. It describes evaluation as a tool to be utilized earlier in the life of programs and as an integral part of managing government programs at all stages, from initial development through start up, ongoing implementation, appropriations, and reauthorization.

Finally, it suggests a variety of ways to organize the evaluation function within Federal agencies and how to connect it to the decisions made in both the Executive and Legislative Branches of the Federal Government.

We plan to make the paper widely available to policy makers throughout the Federal Government as well as to other interested parties. I encourage any of you who are in a position to do so to share it with people who can be influential in implementing it.

I want to thank members of the AEA evaluation policy discussion group who offered helpful comments on an earlier draft of this paper. Again, if you are interested in commenting on future draft documents and providing input on our plans as they evolve, please join the discussion group by signing up at <http://www.eval.org/EPTF.signup.discussion.asp>.

[Go to the EPTF webpage](#)

03/09 – A Look at President Obama's 2010 Budget

On February 26, United States President Barack Obama released his 2010 budget. He described it as a "once in a generation...look at where the country has been" and "charting a new path."

We looked at the budget to see if we could discern what the President thinks about evaluation. The results, compiled by Susan Kistler, our Executive Director, can be accessed through the Evaluation Policy Task Force's page of AEA's website, or more directly at <http://www.eval.org/aea09.obama.budget.evaluation.pdf>.

The budget proposes to make better use of evaluation in education programs, increasing funding for "rigorous evaluation" as a first step to doubling the Department of Education's support for education research. It links evaluation to decision making, proposing to invest in what works and to eliminate funding for education programs that "have either never been seriously evaluated or have received

weak evaluations." The phrase, "rigorous evaluations" is the most oft repeated phrase among the statements made about evaluation of education programs.

More broadly, the Administration will "fundamentally reconfigure the Program Assessment Rating Tool." They want to "open up the insular performance measurement process to the public, the Congress and outside experts" and "eliminate ideological performance goals and replace them with goals Americans care about and that are based on congressional intent and feedback from the people served by Government programs. Programs will not be measured in isolation, but assessed in the context of other programs that are serving the same population or meeting the same goals."

The overall tone is one that plans to take evaluation seriously, link it to decision making, open it up to the public, and use it in context. To measure program performance and ascertain which programs work, the Administration will be leaning on evaluations that are professionally done and wide open to public view.

The bottom line seems to be that the President is looking for performance from evaluators as well as programs.

[Go to the EPTF webpage](#)

04/09 - Looking for Policy in All the Right Places

With the coming of the new Administration and Congress, major new pieces of legislation are now being introduced and more are coming. In this context, we can propose integrating evaluation into Federal programs as an essential feature of good government.

AEA members can help by letting us know where the opportunities are. We work with multiple parties to scan for new legislation, but we also know that one of AEA's most valuable resources is the knowledge base of its members. We need to contact legislative staff at the time they are just beginning the bill drafting process. If you know of federal legislation at such a point, please email me directly at evaluationpolicy@eval.org. While we will not be able to pursue every opportunity, the more we know about where those opportunities lie, the more we can make thoughtful and strategic decisions about next steps and can work with potential partners in pursuit of policy influence.

Some of you may hear about legislation in the making from your colleagues. You can also learn more about pending congressional action from the Library of Congress's website (named after Thomas Jefferson) at <http://thomas.loc.gov/>. This website posts current information about all bills, starting when they are formally introduced by Members of Congress. For those who are somewhat new to the legislative process, the following brief summary may help you understand where to look for legislatively-based evaluation policy.

Authorization laws establish Federal programs, stipulating what must or will or may be done and how much money may be spent on them. Thus, major authorization bills are our first targets of opportunity for making evaluation an integral part of the programs themselves. However, the amount of money actually available for a program is usually limited to the amount appropriated for it in Appropriations legislation. Thus, if we want evaluation to be part of program administration, we have to make sure the Appropriations legislation provides funding for it.

Funding for evaluation is often only implicitly part of funds appropriated for general management. That leaves the evaluation funding decisions up to the Executive Branch officials. However,

congressional Authorization and Appropriations committees can provide additional guidance in what is known as "report language," expressions of congressional intent in the reports issued by the committees at the time bills are sent to the floor for voting or to the President for signature. While non-binding, such guidance carries significant weight in the minds of Executive Branch officials responsible for program implementation.

A good example of these processes is the recently enacted "Tom Lantos and Henry J. Hyde United States Global Leadership Against HIV/AIDS, Tuberculosis, and Malaria Reauthorization Act of 2008." We worked with the Lundy Foundation to request the Authorizations or Appropriations committees to stipulate explicit funding levels for evaluation of this program. They were reluctant to do so. However, based on our written advice, the House and Senate Appropriations Conference Committee did include the following sentence in the Appropriations report: "USAID and OGAC [Office of the Global AIDS Coordinator] are expected to increase funding for operations research, impact evaluation research, and program monitoring to ensure that interventions and approaches to service delivery are evidence-based and continuously improved over time." We have good reason to believe that this guidance will be taken seriously by USAID and OGAC. Through the Lundy Foundation, we have provided materials to reinforce the importance of the evaluation provisions and to clarify timelines, responsibilities, and evaluation actions.

To learn more about the work of AEA's Evaluation Policy Task Force, or to join the conversation via the EPTF discussion list, visit the EPTF website.

[Go to the EPTF webpage](#)

05/09 – Starting Out on the Right Foot in Foreign Assistance

On April 28, Representative Howard Berman, Chair of the House Committee on Foreign Affairs, introduced legislation requiring a comprehensive strategy for United States efforts to reduce global poverty and promote broad-based economic growth in developing countries. It is called the Initiating Foreign Assistance Reform Act (H.R. 2139), and was co-sponsored by Representative Mark Kirk. This is the first phase of what Representative Berman hopes will be a complete restructuring of the Foreign Assistance Act.

A major portion of this bill is devoted to the establishment of an evaluation program to support the development and execution of foreign assistance programs. It is based on draft legislation and supporting materials that were prepared by a small working group, including members of the AEA Evaluation Policy Task Force (EPTF) and the Lundy Foundation, and reflects the principles embodied in the EPTF's "Evaluation Roadmap for a More Effective Government."

This legislative language is probably the most complete evaluation policy ever included in a Federal draft bill. This is just the beginning of what will be a long and complicated legislative process, and there are no guarantees that this bill will become law. However, evaluation will be an integral part of the discussions that surround a new assessment of this nation's foreign assistance activities.

Some of the key evaluation provisions, which are found in section 3 of the bill, are:

- A requirement for the President to develop and implement a rigorous system to monitor and evaluate the effectiveness and efficiency of United States foreign assistance
- The establishment of measurable performance goals
- Criteria for selection of programs to be subject to various evaluation methodologies

- Establishment of an evaluation organization unit in each Federal agency involved in foreign assistance activities
- Requirements to apply the lessons learned and results from evaluation activities in the planning and implementation of foreign assistance programs
- Requirements to publish all evaluation plans and reports
- Requirements for annual evaluation plans
- Consultations among Federal agencies, governments of host countries, international and indigenous nongovernmental organizations, and other relevant stakeholders
- Capacity building for evaluation in Federal agencies and for recipient countries
- Annual budgeting for evaluation
- Establishment of a Foreign Assistance Advisory Council with biennial reports of its activities to the President and the Congress
- Annual reports from the President to the Congress on the use of evaluation
- Definitions of key evaluation terms
- A 5% set aside of foreign assistance funds to pay for evaluations

While it is unlikely that the evaluation language in this bill will be modified any time soon, we may have opportunities to comment in the future. To facilitate the availability of advice from AEA members on this bill, **AEA will be hosting three webinars**. Each is the same and will incorporate a brief description of the bill's provisions as well as next steps in the legislative process, with extended time for questions and discussions. We are holding them in different time zones so as to encourage participation from our colleagues around the world.

If you would like to participate, please send your preferred time to Ashley in the AEA office at office@eval.org. She will send information so that you may view the webinar online and/or listen via VOIP or dial-in.

Time 1: Friday, June 5, 7:00 AM EST

Time 2: Wednesday, June 10: 3:00 PM EST

Time 3: Thursday, June 11: 8:30 PM EST

[Go to the foreign affairs website for the announcement and a link to the bill itself](#)

06/09 – The Value of Environmental Evaluation

On June 8 and 9, AEA President Debra Rog, AEA Evaluation Policy Task Force (EPTF) Members Katherine Dawes and Stephanie Shipman, and I were privileged to attend the Environmental Evaluators Network Forum. This was the fourth annual gathering of environmental evaluators, co-sponsored by the United States Environmental Protection Agency (EPA) and the National Fish and Wildlife Foundation, and hosted by the George Washington University Trachtenberg School of Public Policy and Public Administration.

The agenda of the conference read like a standard reference guide for bringing evaluation to bear on the many challenges confronting our planet. Major presentations and sharing events focused on evaluation polices, including: connections between evaluators, managers, and policy makers; application of evaluation throughout the life cycle of programs; use of evaluation in economic recovery and global warming initiatives; setting environmental goals and standards; the uptake and use of evaluation results by program decision makers; understanding of the scope and complexity of the environment itself; evaluation capacity building; and idea sharing through evaluator networks. Debra

Rog's featured presentation blended her Presidential theme of evaluation in context and the EPTF-developed *Evaluation Roadmap for a More Effective Government*.

Those of us who were there were quite impressed with the scale and scope of innovations and applications of evaluation in the world of environmental policy making and program administration. Those interested in evaluation policy and, more broadly, how evaluation can be used in policy making and program management, can learn much from what these environmental evaluators are doing. Fortunately, anyone wanting to learn more can consult the Environmental Evaluators Network's web site and find agendas and products from this and the previous three forums. Evaluators can also join their LinkedIn group at <http://linkedin.com/> (just look for the "Environmental Evaluators Network" group).

For more information go to the Environmental Evaluators Networking Forum Website or contact Katherine Dawes, EPA's Director of the Evaluation Support Division, dawes.katherine@epa.gov.

Those who want to learn even more about environmental evaluation may also want to read the just-released edition of *New Directions for Evaluation (NDE)*, Volume 2009 Issue 122, *Environmental Program and Policy Evaluation: "Addressing Methodological Challenges,"* Issue Edited by Matthew Birnbaum, Per Mickwitz. AEA members have free online access to this and other volumes of NDE through the AEA website.

[Go to the Environmental Evaluators Networking Forum Website](#)

07/09 – Sizing Up Health Care Reform Proposals

As I write this article, the United States Congress is working intensely on legislation to reform our nation's health care system. The House Ways and Means Committee passed its bill on July 17, 2009. Two other House Committees - Education and Labor, and Energy and Commerce - continue to work on theirs. The Senate Health, Education, Labor, and Pensions (HELP) Committee passed its bill on July 15, while the Senate Finance Committee continues to work on its bill.

Needless to say, the Evaluation Policy Task Force (EPTF) is concerned that health care reforms are systematically evaluated from the date of their enactment through the many years in which they will be in effect. The best way to make this happen is to have evaluation requirements and resources built into the new legislation. For this reason, the EPTF has sent written comments to House and Senate Committees on the emerging legislation. Copies of EPTF comments to the three House Committees and to the Senate Finance Committee can be found online on the EPTF website.

One way to think about the proposed legislation and the EPTF comments is to realize that the various bills cover two broad categories of proposals: (1) insurance reforms, such as guaranteed availability, renewability of coverage, elimination of discrimination based on health status, coverage of preventive health care, assistance for low income families, cost control, and financing; and (2) public health provisions such as those relating to quality of care, health care workforce enhancement, community health centers, women's health, immunization, and school based clinics.

What we found in reviewing the current bills is that evaluation is fully and explicitly built in to the legislation relating to the public health proposals, but to a much lesser extent in the insurance reforms. In fact, many of the evaluation provisions in the public health sections could serve as excellent models for legislation in many public laws, not just health care legislation. The health insurance sections provide oversight and evaluation through the Government Accountability Office, existing Inspectors

General, or a Special Inspector General. The public health provisions reflect the growing reliance on evaluation in that field, and the insurance provisions reflect longstanding concerns about "oversight" of Medicare and Medicaid programs, including fraud, waste, and abuse.

The EPTF developed an overall evaluation framework for *all* the provisions of the bills, including

- **An annual national report card** on our nation's health care system
- **Key evaluation questions** that need to be posed by the Congress
- **Evaluation units** designated, or if necessary established, within agencies responsible for the reforms
- **Annual evaluation plans** related to programs or activities covered in the bill
- **Annual reports** summarizing evaluation results and explaining how they were used
- **A Health Care Evaluation Advisory Group** with responsibility for preparing a biennial report on the state of the nation's health care system, making recommendations for the consideration of the Congress and the President; and providing non-binding advice to Federal evaluation units

The EPTF will continue to monitor the development of health care reforms and offer advice when appropriate as the legislation evolves.

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

08/09 – Science and Technology Priorities

Evaluation will be an essential ingredient in meeting the President's science and technology priorities. That is the message that comes through loud and clear in an August 4 [memorandum](#) for the heads of Executive Departments and Agencies issued jointly by Peter Orszag, Director of the Office of Management and Budget (OMB) and John Holdren, Director, Office of Science and Technology.

After describing what it will take to address the practical challenges of developing research and technology policies and programs that address the economy, energy dependence, global warming, green jobs, health and health care costs, and national security, the memorandum immediately puts evaluation at center stage. The opening words of the program guidance section are:

"In their budget submissions, agencies should describe the expected outcomes from their research in relation to these four practical challenges and cross-cutting areas, providing quantitative metrics where possible, and describe how they plan to evaluate the success of various techniques to increase support for high-risk research.

"Budget submissions should also describe how agencies are strengthening their capacity to rigorously evaluate their programs to determine what has been demonstrated to work and what has not."

The guidance sees evaluation as a tool to enable "... agencies to eliminate or reduce funding for less-effective, lower-quality, or lower-priority programs ..." as well as "... target investments toward high-performing programs."

This language is heartening in that it is generally consistent with the principles described in *An Evaluation Roadmap for a More Effective Government* that three AEA Presidents sent to Peter Orszag in February. We are continuing to stay in touch with OMB staff on how evaluation can be made an

essential element of governance. Hopefully, future budget guidance documents will also make the kind of strong connection between evaluation and policymaking as this one does.

Meanwhile, evaluators will have to deliver on their end, working with program managers to:

- Choose appropriate metrics and evaluation methods
- Apply those methods and metrics to both current and evolving policies and programs
- Assess the results of science and technology policies that are implemented in more than one place within a single lab, agency, or firm
- Measure success not only at the program level, but also in terms of the influence of science and technology on the four practical challenges described at the beginning of the memorandum
- Report successes, failures, and unexpected consequences of research and new technologies
- Interpret evaluation results in terms of their ramifications for programs.

Hopefully, evaluators will rise to meet the challenges posed in this directive.

[Go to the Evaluation Policy Task Force website to learn more](#)

09/09 - Next Steps on Foreign Assistance

In our [May Policy Watch article](#) we alerted you to some promising developments in the field of foreign assistance, a new bill in the House of Representatives with enlightened evaluation policies, the "Initiating Foreign Assistance Reform Act" ([H.R. 2139](#)). While we were hopeful that this bill or something like it would be enacted, we were not getting our hopes too high. A stronger sign that something might eventually emerge from the Congress would be similar evaluation provisions in bills that reauthorize foreign assistance programs. We can now report progress on that front.

Reauthorization bills are now pending in both the House of Representatives and the Senate. Both contain strong, but very different, evaluation sections.

The House bill, "Foreign Relations Authorization Act, Fiscal Years 2010 and 2011" ([H.R. 2410](#)) passed the House on June 10 and has been remanded for consideration by the Senate. It is peppered with provisions requiring evaluation of specific programs authorized by the act, such as: new, secure U.S. public diplomacy centers in foreign countries, libraries and resource centers, employee compensation, assistance to Iraq refugees, and aid to Pakistan. Perhaps more importantly it contains a new, sweeping evaluation authority in section 1114 for evaluation of all foreign assistance programs with almost all the provisions that are in the "Initiating Foreign Assistance Reform Act." The only provision that is missing is the establishment of a United States foreign assistance evaluation advisory council.

The Senate bill, "Foreign Assistance Revitalization and Accountability Act of 2009" ([S. 1524](#)), was introduced on July 28. It fills the gap in the House bill by establishing two new entities to coordinate evaluation of foreign assistance programs. In section 5, it would establish an Office for Learning, Evaluation, and Analysis in Development that will develop, design, coordinate, guide, and conduct the complete range of evaluation activities relating primarily to the United States Agency for International Development (USAID). In section 6, it establishes a Council on Research and Evaluation of Foreign Assistance to conduct evaluations of the impact of foreign assistance programs carried out by any Federal agency or international and multilateral assistance programs receiving financial assistance from the United States. However, it contains none of the overarching evaluation provisions included in the House bill.

If the Senate passes its reauthorization bill, the odds are high that the joint House/Senate committee that will be established to combine the two bills will include significant evaluation provisions. Hopefully, such a resolution would be to combine the best provisions of both bills. We are watching this situation closely and working with the Lundy Foundation and other independent evaluators to advise the Congress on how to improve evaluation of foreign assistance programs.

[Go to the EPTF website to learn more](#)

10/09 – OMB Policy on Impact Evaluation

On October 7, United States Federal Office of Management and Budget (OMB) Director Peter Orszag issued an [OMB memorandum](#) to Heads of Executive Departments and Agencies establishing a broad policy to place increased emphasis on impact evaluations. It focuses initially on "social, educational, economic, and similar programs whose expenditures are aimed at improving life outcomes (such as improving health or increasing productivity) for individuals." Procurement, construction, taxation, national defense, and drug and clinical medical evaluations are beyond the initial scope of the initiative.

Within this field of impact evaluation the policy picks up on many of the ideas that AEA's Evaluation Policy Task Force (EPTF) has advocated for in its [Evaluation Roadmap for a More Effective Government](#) and the [EPTF memorandum](#) last year to OMB on the Program Assessment Rating Tool (PART) program. For example, consistent with the Roadmap, it emphasizes that "evaluations can help policymakers and agency managers strengthen the design and operations of programs," and it seeks to improve the institutional capacity of Federal agencies to perform evaluations, asking for an assessment of agency capacity to conduct rigorous, independent evaluations and to attract and retain talented researchers in an office with standing within the agency. To assist in building this capacity, it reconstitutes the Inter-agency Evaluation Working Group under the Performance Improvement Council.

Consistent with last year's EPTF comments to OMB on PART, it avoids reference to a "gold standard" method. Instead it establishes that a key goal of the Working Group will be to "help agencies determine the most rigorous study designs appropriate for different programs given their size, stage of development, and other factors."

The policy formalizes and provides funding for the voluntary program started last year for agencies to submit information about, and requests for funding for, high priority evaluation activities.

The policy is strong on public disclosure of evaluations, directing OMB and agencies to expand information about program evaluations that they make public. This is being done not only to promote transparency of evaluation results but also to allow experts inside and outside the government to engage early in the development of program evaluations. In particular, OMB welcomes input on the best strategies for achieving wide consultation in the development of evaluation designs. Clearly, OMB is making an open invitation for evaluators to weigh in on important, concrete evaluation issues.

Interestingly, it also expresses openness to the idea of evaluations of potential program policies, not just existing programs. More specifically, it establishes a goal of making researchers, policymakers, and the public aware of evaluations that study alternative approaches for achieving outcomes to determine which strategies are most effective.

Needless to say, the Evaluation Policy Task Force will be monitoring the implementation of this new policy. If you become aware of significant developments in this area, please share them with the EPTF discussion group or directly with me, George Grob, consultant to the Task Force, at evaluationpolicy@eval.org.

[Go to the EPTF website to join the discussion group and learn more about its work](#)

11/09 – Paperwork Reduction Act

The United States Federal Office of Management and Budget (OMB) is seeking comments on implementation of the Paperwork Reduction Act. AEA is planning to send comments and is seeking your input by Wednesday, December 2. Please read on to learn more and make your voice heard.

On October 27, OMB published a [request for comments](#) on the requirements of the [Federal Paperwork Reduction Act](#). Under the Act, all surveys and other data collection instruments issued by or on behalf of any Federal agency to more than 9 responders must be approved in advance by OMB. OMB is seeking comments on reducing current paperwork burdens, especially on small entities; increasing the practical utility of information collected by the Federal Government; ensuring accurate burden estimates; and preventing unintended adverse consequences.

The review process is complex. It involves an initial review by the originating Federal agency, two rounds of public comments (first for 60 days, and then again for 30 days), and up to 60 days of review by OMB. The entire process usually takes at least six months and often substantially longer than that. While addressing legitimate concerns about paperwork burdens on the public, it may delay the availability of critically important evaluation results.

All AEA members may respond directly to OMB's request for comments on their own. In addition, AEA plans to send a short statement to OMB on this important issue. To inform this statement, we would like member input. We are particularly interested in hearing from AEA members who have had experience in seeking approval of evaluation surveys or other data collection instruments under the Paperwork Reduction Act. The schedule for making our reply is very tight. **We need to hear from you by 5:00 PM EST on Wednesday, December 2, via our online questionnaire.**

AEA Paperwork Reduction Act Questionnaire Link: <http://bit.ly/prasurvey>

Of course, this should not preclude your submitting your own comments directly to OMB and I encourage you to read the [request for comments](#) to learn more regarding options for direct input.

Please feel free to share this information with your colleagues and invite them to respond as well.

We are grateful for your assistance.

[Go to the EPTF website to join the discussion group and learn more about its work](#)

12/09 – AEA Comments on the Paperwork Reduction Act

Last month, we informed you that the United States Federal Office of Management and Budget (OMB) published a [request for comments](#) on the requirements of the [Federal Paperwork Reduction Act](#). This

is the law that requires solicitation of public comments and prior approval by OMB of surveys and other data collection instruments for all research and evaluations sponsored by the Federal Government. Knowing how important surveys and this process are for the many AEA members that are involved in Federal projects, we invited input to comments that AEA would send to OMB. Thirty-three members responded. Based on that input, as well as discussions with other interested members and the Evaluation Policy Task Force, we sent formal comments to OMB signed by the AEA Presidential Rotation - President Debra Rog, Immediate Past Present William Trochim, and President-elect Leslie Cooksy.

In summary, we told OMB that AEA recognizes the importance of minimizing paperwork burdens on the American public. However, we believe that the current paperwork review process has unintended negative impacts, particularly in denying Federal managers and policy makers timely access to information that is critically needed to address emerging problems and take advantage of promising approaches. We offered several options to speed up the reviews while promoting burden hour reductions, improving burden estimates, and enhancing the usefulness of the data collected. We also advised OMB on the interpretation of its own [Standards for Statistical Surveys](#), which OMB analysts use in connection with their reviews of proposed surveys.

We are grateful to OMB for inviting public comments on the survey review process and hope our comments are useful to them. We recognize that most of the requirements that our members and others find so burdensome have not originated from OMB, but rather are required by the Paperwork Reduction Act itself. Therefore, we offered our comments to OMB for their consideration either in proposing changes to the law or in administratively adopting changes that are compatible with the current statute. We also offered assistance to OMB by perhaps sponsoring sessions at our national conferences to help promote an understanding of the Act, OMB's procedures, ways to minimize paperwork burdens on the public, estimate burdens accurately, and improve the quality and usefulness of the data that we collect.

I wish to thank all those who provided us with useful advice, either by responding to our online survey or providing comments in person or by phone.

[Go to the EPTF webpage to view AEA's Comments on the Paperwork Reduction Act](#)

1/10 - Pillars of Evaluation Policy

One of the difficult challenges any professional organization faces in trying to influence legislative, regulatory, and administrative policies is to articulate a clear and consistent set of cardinal goals or values that are of the utmost importance to the professional practice of its members and the field in which they work. On the surface, that might seem easy to do. In reality it is difficult to come up with a list of "policy pillars" which capture the essence of a profession, can be easily remembered, and can be used as points of reference from which to derive appropriate responses to myriads of elaborate proposals advanced by legislators and other government policy makers.

AEA's Board of Directors, with input from the Evaluation Policy Task Force (EPTF), took on this task at its November meeting during the AEA national conference. Following are the common pillars they adopted. The Board has authorized the EPTF and me to promote them in meetings and communications with outside groups involved in formulation of evaluation policies.

AEA Pillars of Evaluation Policy

- Broad use of evaluation in public programs, especially those of the Federal Government
- Using methods appropriate to the evaluation questions
- Adapting the size and scope of evaluations to be appropriate to the program's context and needs
- Adequate funding for evaluation
- Use of qualified, experienced evaluators as appropriate
- Evaluator independence
- Transparency of results

If you would like to discuss the pillars or explore other issues of evaluation policy, please consider joining the EPTF discussion list at <http://www.eval.org/EPTF.signup.discussion.asp>.

2/10 – Evaluation Policy in the President's Budget

When President Obama took office a little more than a year ago, we noticed his interest in program evaluation. Now we have fresh evidence of it in his 2011 budget. AEA's Executive Director, Susan Kistler, has prepared an analysis of that budget, which we are sharing with you [here](#). Evaluation is woven throughout the document. Here are the highlights.

Evaluation of specific Federal programs. This includes proposals to increase the number of effective teachers and principals; boost development of clean energy on tribal lands; improve services for seniors and people with disabilities; and promote wellness initiatives in the Federal workplace. Especially noteworthy is a request for \$500 million to expand the Investing in Innovation Fund, to evaluate and expand proven models for achieving student success.

Terminations, reductions, and savings. The budget cites evaluations to support budget reductions in several programs, such as certain children and families services' job demonstration programs, rural health care services programs, and watershed and flood prevention programs.

Sweeping evaluation policy. Far and away, the most impressive evaluation message is a set of overarching policies found in an annex to the budget called *Analytical Perspectives*. The parts that are relevant to evaluation are the first two sections (labeled 7 and 8) in a chapter called "Performance and Management."

Section 7, Delivering High-Performance Government, provides a new look at performance measurement and substitutes a new performance management system that builds on but also replaces the PART system used by the previous Administration.

Section 8, Program Evaluation, provides detailed steps and explanations to implement the policies enunciated by OMB Director Peter Orzag in his October 7 memorandum and discussed in our [October column](#). The introductory paragraph puts evaluation in the context of performance management. "Performance measurement is a critical tool managers use to improve performance, but often cannot conclusively answer questions about how outcomes would differ in the absence of a program or if a program had been administered in a different way. That is where program evaluations play a critical role."

The second paragraph is equally telling. "A central pillar of good government is a culture where answering . . . questions [about how well programs work] is a fundamental part of program design and where agencies have the capacity to use evidence to invest more in what works and less in what does not."

The budget allocates approximately \$100 million to 17 agencies to conduct new evaluations with strong study designs that address important, actionable questions or to strengthen agency capacity to support such strong evaluations.

We cannot do justice to these new policies in this short article. I urge evaluators to read these two sections themselves. For the evaluators involved in the work of the 17 agencies named in section 8, my take on this is that the ball is now in your court. Actually, I think it is in all of our courts.

[Go to the EPTF web page](#)

3/10 – New Teams Forming Up: Your Help Needed

We need your help. We've been keeping a close eye out for Presidential appointees and other senior officials with significant roles in the field of evaluation. A few names have popped up, but typically these appointments don't draw much attention until well into the first year or so of new Presidential administrations. Most of their early attention goes to filling top posts such as Cabinet members and heads of Federal departments and agencies. They then have a hand in selecting other officials, including those in the field of evaluation. Still more will follow in the third tier of appointments, influenced by those in the first and second tiers. Nevertheless, these are very senior and important appointees who will likely have significant influence not only on evaluation but on public policy more generally.

As you know, our AEA Presidents wrote to Peter Orzag, Director of the Office of Management and Budget (OMB) within days of his arrival, sending him the *Evaluation Roadmap*. And, we have been in touch with other senior officials including Shelley Metzenbaum, Associate Director of OMB for Performance and Personnel Management. It will be equally important for us to contact senior evaluation officials in all Federal agencies as they officially arrive for duty.

That's where you come in. I am writing to invite those of you in touch with Federal agencies to let us know when you become aware of such appointments. It would be useful to have contact information, a note regarding their role in terms of evaluation, and - if easily available - any background information from official announcements or common public sources. We are not asking for detailed background

checks or for opinions about the potential ramifications of such appointments. Nor are we interested in rumors that someone is being considered for appointment. We do want to know about those who are officially appointed and have arrived for duty.

Once we become aware of such appointments, we will follow up to welcome them aboard and send them information about AEA. You can email me with that information at EvaluationPolicy@eval.org. Thank you for your help.

[Go to AEA's Evaluation Policy Task Force page](#)

4/10 - An Invitation to Comment on the Evaluation Roadmap

Earlier this week AEA Evaluation Policy Task Force (EPTF) Chair Patrick Grasso sent a message to AEA members inviting all of you to comment on the *Evaluation Roadmap*. I encourage you to take advantage of this opportunity if you have not already done so.

The *Evaluation Roadmap* is the name now commonly used for a document entitled "*An Evaluation Roadmap for a More Effective Government*," which three AEA Presidents (Leslie Cooksy, Debra Rog, and William Trochim) sent to Peter Orszag, Director, Office of Management and Budget, in February 2009, shortly after he took office. The paper explains the importance of evaluation as an essential ingredient of good government, and it describes concrete steps to make this happen.

This document is important. It has been widely circulated among government policy makers and has been recognized as a useful reference source and practical guide for promoting effective use of evaluation in government. However, it can be made even better with your input.

The EPTF prepared the original version of the *Roadmap* to reach incoming officials of United States President Obama's administration as they were formulating their management agenda. We are now interested in framing the *Roadmap* as a foundational document of AEA to guide evaluation policy development work well into the future. The EPTF has redrafted the *Roadmap* to eliminate references to the Obama and Bush Administrations in order to eliminate any partisan interpretation of it and to emphasize the importance of evaluation in the context of abiding, fundamental principles of governance.

The AEA Board, its Presidents (the three mentioned above plus President-elect Jennifer Greene), and the EPTF invite your comments. On for before Friday, May 21, please login at the link below to provide your feedback.

Your username:

Your password:

[Go to the Roadmap Vetting Page](#)

Thank you for your advice.

[Go to the EPTF page](#)

5/10 - Promising Developments in Health Care Reform

Health care reform legislation is now yesterday's news, but the reforms themselves will be felt for years to come. A natural question for us is, "How will the reforms be evaluated?" We followed the development of the legislation quite closely and provided copies of the *Evaluation Roadmap for a More Effective Government* to key staff members of the congressional committees working on the legislation. We encouraged them to consider evaluation provisions in the bills leading up to the enactment of the *Patient Protection and Affordable Care Act*.

The health insurance reforms were originally lacking appropriate evaluation and oversight mechanisms. To be sure, there were (and still are) important requirements for evaluation of administrative simplification of health insurance transactions and for a survey of enrollee satisfaction with health plans offered through state health insurance exchanges. The Government Accountability Office is also required to review all aspects of the administration of the health insurance exchanges. However, overall, evaluation requirements of insurance reforms were sparse. We pointed this out to the relevant congressional staff. The final bill now includes oversight of the insurance reforms through the Office of the Inspector General of the Department of Health and Human Services. The Inspector General's oversight of all aspects of this title provides a framework for the conduct of independent evaluations as well as audits and investigations.

While the national debate raged over health insurance reform, less public attention was paid to the numerous public health and health systems reforms that were contained in the bill. Only one sixth of the law, Title I, *Quality, Affordable Health Care for All Americans*, deals directly with health insurance reforms. Title II, *Role of Public Programs*, adds health insurance improvements through public programs, but also includes demonstrations of innovative health care delivery and financing mechanisms, primarily Medicaid. Fully two-thirds of the remaining pages of this 2400 page bill, titles III through X, address public health programs and health care systems reforms.

Fortunately, these public health and health systems reforms were drafted with evaluation in mind. Starting with title II, requirements for the kinds of evaluation that the Roadmap calls for are prolific, including evaluations of new demonstrations authorized under Medicaid, health care indicators, quality improvements, community health, nursing home reforms, improvements in coverage of prescription drugs, health care workforce enhancements, prevention of chronic diseases, improved access to medical therapies, and preventive health. To get a feel for this, I suggest you pull up the pdf version of the bill at the link provided above and do a search with the partial word "evaluat."

We cannot take credit for this. The culture of evaluation within the public health community is largely responsible for the successful embedding of evaluation throughout the bill. Indeed, it is encouraging to see this development, a reflection that evaluation may be coming of age.

I will share more about this in future editions of the AEA newsletter and on the evaluation policy discussion list.

[Go to the Evaluation Policy Signup page to join the discussion list](#)

6/10 - Rounding Out Health Care Reform Evaluation Policy

Last month, I wrote to you about evaluation of the new health reform law. As discussed then, the Evaluation Policy Task Force's most immediate concern was ensuring that evaluation would be used to assess the timely and effective implementation of the complex insurance reforms. Much could go wrong, not all of it turning on fidelity of implementation. There was, and still is, legitimate concern

about waste and fraud. The institutional responsibilities of the Inspector General -audit, investigation, and evaluation - provide one set of strong safeguards against such challenges.

I was also initially concerned about the evaluation of the many separate bits and pieces of the public health and systems improvements that are scattered throughout the legislation. Reforms related to health services, health professions, preventive health, health care quality, health information technology, and refinements to Medicare and Medicaid programs also need to be effectively implemented and their impact evaluated. Fortunately, this potential problem was addressed as a result of the apparent acculturation of evaluation within the public health community and the corresponding congressional authorization committees. They did a good job of building evaluation into the very fiber of these programs.

A much larger concern was how to use evaluation to help ensure the success of the whole enterprise. Would the legislation result in a healthcare system that is accessible and affordable? Would it provide quality healthcare to those who need it? Will we as a nation be healthier five, ten, fifteen or more years out?

Fortunately, the new law does, in fact, provide one tool for addressing these overarching questions. It is section 5605, Key National Indicators, of the enacted health reform legislation, the *Patient Protection and Affordable Health Care Act*.

This section establishes a congressionally appointed Commission on Key National Indicators to oversee the development of such a system and authorizes the National Academy of Sciences to determine how best to establish it. The Academy will convene a multi-sector, multi-disciplinary process to define major scientific and technical issues associated with developing, maintaining, and evolving the indicator system. The system will be subject to annual reports by both the Commission and the Academy, and to financial audits and programmatic reviews by the Government Accountability Office.

Of course, a national indicator system is not an evaluation. However, it will hopefully provide a means to help all Americans and policy makers see where things stand and promote the commissioning of evaluations to follow up on both problems and promising results revealed by the indicators.

As most insiders remarked upon passage of health care reform legislation, now the real work begins. Perhaps national indicators will help policy makers and evaluators stay on top of our evolving health care system.

[Go to AEA's Evaluation Policy Task Force page](#)

7/10 – Evaluation and the Global Health Initiative

In past articles, we have focused on evaluation policies established through budget decisions and legislative changes. However, there are other important policy making mechanisms, including program management initiatives. Such initiatives are often associated with budget proposals, but go far beyond the dollars. They can be powerful forces that result in enduring changes in attitudes, expectations, goals, strategies, organizations, and administrative procedures. In essence, they can establish new ways of doing business. One good example is the President's Global Health Initiative (GHI).

On May 5, 2009, U.S. President Barack Obama announced his proposal to spend \$63 billion over six

years to significantly reduce deaths from AIDS, malaria and tuberculosis. However, he also noted the need to "improve health systems around the world, focus our efforts on child and maternal health, and ensure that best practices drive the funding for these programs." The emphasis on improving health systems and adoption of "best practices" has subsequently been refined and promoted within all U.S. international health programs, including those funded and operated by the Office of the U.S. Global AIDS Coordinator, the Administration for International Development (USAID), and the Department of Health and Human Services (HHS). The strategy and implementation details are described in *Implementation of the Global Health Initiative, Consultation Document*.

The following goals guide the initiative:

- Implement a woman- and girl-centered approach
- Increase impact through strategic coordination and integration
- Strengthen and leverage key multilateral organizations, global health partnerships and private sector engagement
- Encourage country ownership and invest in country-led plans
- Build sustainability through health systems strengthening
- Improve metrics, monitoring and evaluation
- Promote research and innovation

Evaluation is a fundamental aspect of the initiative. References to it are pervasive. I invite those who are interested to word search the string "evaluat" in the pdf to see just how embedded evaluation is. One indicator of the importance accorded to evaluation is an unusual 10 percent evaluation set aside of GHI funds that will be provided to up to 20 "GHI Plus" countries. These countries will be selected from among those that provide significant opportunities for impact, evaluation, and partnership with governments. The 10 percent set aside will be used to design and implement an intensive monitoring and evaluation effort, and to broadly disseminate findings.

From an evaluation policy standpoint, a new dynamic and instrumentality is at work in the Global Health Initiative. Evaluation policy will evolve on the ground. What people actually do will determine how ingrained evaluation will become as a *sine qua non* of international public health programs. If evaluation is truly useful, it will be in demand. If not, it won't.

I am curious to know what AEA members think about all this and what evaluators can do to make evaluation truly useful in the GHI initiative. The Policy Discussion List is a good way to share ideas.

[Go to the Evaluation Policy Discussion List and Join the Conversation](#)

8/10 - Weighing the Evidence on Home Visiting Programs

The recently enacted health care reform legislation contains a provision authorizing a Maternal, Infant, and Early Childhood Home Visiting Program. The law requires that to be eligible for funding, a program must meet a number of stringent conditions. Among them is that it has been evaluated using rigorous, evidence based methods. It further stipulates that the Department of Health and Human Services (HHS) review the evaluations that have already been performed on various home visiting models in order to determine which ones have been proven successful using rigorous evaluation methods and that the methodology for conducting that review be transparent and subject to public comment. True to the law, HHS published for comment its methodology for conducting that review.

The proposed methodology, among other things, automatically reserves its "high" level rating to randomized assignment studies and relegates all other methods to "moderate" or low levels. You can [read the announcement](#) requesting public comments online.

The AEA Board and the Evaluation Policy Task Force worked together in preparing [comments](#) which were signed by our "three presidents" - Leslie Cooksy, President, Debra Rog, Immediate Past President, and Jennifer Greene, President-elect.

In summary, AEA's comments recognize the importance of using "evidence based" models as a basis for distributing funds available under the home visiting program and that the proposed criteria and methodology for a systematic review of such models represents a thoughtful starting point for assessing the evidence of their effectiveness. However, the AEA comments raise concerns about how the studies upon which the evidence is based are rated. They include recommendations to 1) forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings, and avoid using the label "gold standard" in connection with random assignment designs in the rating methodology, 2) use additional criteria to assess the value of impact evaluations, 3) more specifically identify alternative impact evaluation methods, and 4) emphasize the value of multiple studies and mixed methods.

The relative value of randomized control, quasi-experimental, and other study types has been a matter of intense interest among evaluation professionals. The AEA Board, including its three presidents, and the Evaluation Policy Task Force recognize that there is much more to be said about such evaluation methodologies than is included in its comments on the HHS home visiting program. However, it was important to offer balanced advice in response to a formal invitation for public comments from a Federal Government agency on a topic of such relevance to the evaluation community.

For anyone interested in this topic, I wish to emphasize the importance of reading the HHS notice and the AEA comments in their entirety. For those interested in further discussion of these issues, I would like once again to invite you to join the evaluation policy discussion group.

[Go to the Evaluation Policy Discussion List and Join the Conversation](#)

9/10 – AEA Member Input to Evaluation Policy

I want to thank AEA members for their participation in the development of evaluation policy. While the AEA Board of Directors and the Evaluation Policy Task Force (EPTF) were anxious to get input from AEA members on emerging evaluation policy issues, initially it was not clear how to go about it. The most fundamental barrier has always been the short deadlines under which national policy making machinery functions. To an outsider, it looks like laws, regulations, and budgets take years to enact. That is true, but the individual steps of those processes are carried out in bursts of rapid fire consultations. Opportunities for public input on a particular policy are often limited to just a few weeks or even less. The EPTF has sometimes had to review proposed policies and draft advice in a matter of days.

But this last year has been a season of opening doors for all of us. The Board, the EPTF, and AEA members, have learned how to engage one another in the policy processes. All have been involved in redrafting of the *Evaluation Roadmap*, advising the Office of Management and Budget on ways to improve implementation of the Paperwork Reduction Act, and, more recently, commenting on the

Department of Health and Human Services' proposal for rating the effectiveness of delivery mechanisms for the Maternal, Infant, and Early Childhood Home Visiting program.

As I write this article, two opportunities for AEA member input are pending:

- Please consider providing input by October 1 to AEA's comments on the General Accountability Office's Audit Guide. [Use this link to provide input](#)
- Please vote by October 7 on whether to approve the Evaluation Roadmap as a public statement on behalf of AEA. [Use this link to vote](#)

It is also noteworthy that in several cases (foreign assistance legislation, the Paperwork Reduction Act, and the GAO Audit Guide), it was AEA members who alerted us to the opportunity for AEA to provide advice. I encourage you to consider [joining the EPTF discussion list](#) to share your ideas regarding possibly policy-influencing opportunities or to email me directly at evaluationpolicy@eval.org.

It has been my pleasure to work with so many of you. But more than that, your participation has been very important to AEA and to the evaluation profession. The AEA Board and the EPTF hope you will continue to stay engaged on the policy front, and to be eyes and ears, alerting us to evaluation policies as they arise. Thank you so much.

10/10 – Evaluation Roadmap: It's Official!

Max Ehrmann, in his poem *Desiderata*, advised his son to "Enjoy your achievements as well as your plans." That is good advice for all of us now with the approval by AEA members of *An Evaluation Roadmap for a More Effective Government*.

Two years ago this month, the Evaluation Policy Task Force (EPTF) began an intensive campaign to connect with the incoming administration of President Obama about the importance of evaluation. The *Roadmap* emerged from the initial papers and interactions with Office of Management and Budget (OMB) staff.

This document is worth celebrating for several reasons. First and foremost is its content - a digested but still detailed statement about the need for evaluation as "the DNA" of federal programs, as outgoing OMB Director Peter Orszag called it, and a set of principles to guide adoption of effective evaluation practices within the federal government. The *Roadmap's* themes can be seen in the evaluation policies of the Obama administration, which we have described in several "Policy Watch" articles.

Equally important is the process whereby this document was crafted. AEA has long struggled to produce policy statements that are both timely and reflective of its memberships' views. With the *Roadmap* we found a way forward. This version reflects comments provided by interested members through an open call for comments and was not finalized until its approval this month by an open membership vote. The process that emerged during the *Roadmap* development has been adapted to the formulation of comments on the Paperwork Reduction Act; the Maternal, Infant, and Early

Childhood Home Visiting program; and the Government Auditing Standards. We are now working to develop final policy for the Board to review at its January meeting.

So what's next? First and foremost, we urge all AEA members to read the *Roadmap* and to distribute it broadly through any of their own professional networks that are tied to the federal government. The more who know about it, the better.

Beyond that, there is much more evaluation policy to develop, publish, and implement. The current *Roadmap* should be regarded as "version 1." It is a living document that can be updated periodically to reflect both the needs of government policy makers and the maturation of the evaluation field itself. There may also be needs and opportunities to develop official positions on other policy matters. For example, one member has suggested developing a statement on the practical aspects of promoting evaluator independence within government agencies.

The EPTF encourages AEA members to identify policy areas of critical concern to evaluators. The evaluation policy discussion list provides a mechanism to exchange ideas on such matters and I encourage you to subscribe via the EPTF homepage.

[Go to the EPTF Homepage to view the Roadmap or Join the Discussion List](#)

11/10 – AEA Comments on GAO Auditing Standards

As many of you may know, the Government Accountability Office (GAO) is seeking public input on its sixth revision of the *Generally Accepted Government Auditing Standards*, commonly known as the Yellow Book. Because of the close connection of the Yellow Book to evaluation, AEA decided to respond. After consultation with AEA members and the AEA Board of Directors, AEA's Evaluation Policy Task Force prepared comments which were sent to GAO on November 22 under the joint signature of AEA's President, President-Elect, and Past President.

The Yellow Book has evolved in recent years, gradually encompassing studies called performance audits that overlap with evaluations. For example, the Yellow Book explicitly refers to AEA's *Guiding Principles for Evaluators*, and *The Program Evaluation Standards* put forward by the Joint Committee on Standards for Educational Evaluation.

AEA's comments noted the importance of the Yellow Book, especially to evaluators who conduct performance audits. However, AEA also noted "overly rigid requirements . . . that inappropriately constrain all performance audits within the mold of a normative methodology." The comments also call for greater emphasis of a standard of completeness, namely that "Being complete . . . also means the report states evidence and findings without omission of significant relevant information related to audit objectives."

Based on the limited response from AEA members to the invitation to comment (only one member submitted input) and from a panel discussion of the Yellow Book at the recent AEA conference, it

appears that very few AEA members are familiar with the Yellow Book. Perhaps their being called "Auditing Standards" accounts for this. Some of us (including myself) have been more intimately involved with them, and as noted in the request for comments, two of our outgoing AEA Board members - Mike Hendricks and Rakesh Mohan - serve on GAO's Advisory Board for the Yellow Book (although not as representatives of AEA). Even though the *Auditing Standards* are not comprehensive in their treatment of evaluation related topics, they do contain useful advice. Particularly noteworthy are their treatments of various types of performance audits, evidence, conflict of interest, and threats to independence.

Hopefully GAO will react favorably to AEA's comments, but whether they do so or not, evaluators can learn some useful practices from auditors. The parts of the [draft Yellow Book](#) that may be of special interest to evaluators are:

- Chapter 1 - Ethical Principles
- Chapter 2 - Section on Performance Audits (pages 12-14)
- Chapter 3 - General Standards
- Chapter 6 - Field Work Standards for Performance Audits (paragraphs 6.15, 6.37-6.75)
- Chapter 7 - Reporting Standards for Performance Audits
- Appendix I - Additional Information (pages 166-167 and 180-187)

For those interested, here is wishing you all happy reading.

[Go to the AEA Comments on the GAO Auditing Standards](#)

12/10 – Ringing in the New Biennium

The New Year is fast upon us, and it is time for reflection and resolutions. In the federal policy world, we also count two years at a time. For example, now is the 111th Congress, with the 112th only weeks away. Similarly, it is common to think of presidential terms in two year chunks. We are now ending the first half of the Obama Administration and about to start the second. So for our policy work, we need to reflect on the last biennium, and strategize about the coming one.

Looking back is pleasant. AEA began by engaging senior officials of the New Administration. An early version of the Evaluation Policy Task Force's (EPTF) document [An Evaluation Roadmap for a More Effective Government](#) was delivered to senior officials at the Office of Management and Budget (OMB) upon their arrival. I can't say that they took action because they read it, but many of its principles are embedded in the Administration's policy emphasizing impact evaluation and in its budget guidance for the 2011 and 2012 fiscal years.

We also made inroads in the Congress. We see our handiwork in the oversight of health insurance reforms in the new health care legislation, in emphasis on evaluation funding in a foreign assistance appropriations committee report, and in the House reauthorization bill for foreign assistance.

More generally, embracing evaluation as an essential feature of good government is emphasized throughout the health care reform bill and in the President's global health initiative.

AEA also made good progress in managing evaluation policy. It has always been difficult for professional organizations to respond rapidly to opportunities to comment on emerging policy issues while still engaging its membership and governing bodies. However, AEA has done so through its membership vote of approval of the *Roadmap* and in its public comments to OMB on the Paperwork Reduction Act; to the Department of Health and Human Services on the Maternal, Infant, and Early Childhood Home Visiting Programs; and to the Government Accountability Office on its Auditing Standards.

Now, it is time to plan the next two years. Resources are limited. However, it is better to start with a long list and narrow it down than vice-versa. So the EPTF would like to invite your input. Some of you have already started without our call, and that is great. I am referring to a recent string of conversations on the evaluation policy discussion list responding to a question by one AEA member, Margaret Johnson of Cornell University, "What would you say are the top few issues in evaluation policy right now?" An excellent question, Margaret, and one to which we would like to have a lot of answers. What better way to ring in the new biennium than to lay down some challenges? So I encourage discussion list members to put in their two cents, and for others to consider joining the EPTF Discussion List to share your ideas regarding policy-influencing opportunities.

[Go to the EPTF Website and Join the EPTF Discussion List](#)

1/11 - Whatever Happened to GPRA and PART?

On January 4, the President signed HR 2142, the [GPRA Modernization Act of 2010](#). [A summary](#), prepared by the Congressional Research Service (CRS), is also available. The bill amends the Government Performance and Results Act (GPRA) and other statutory provisions related to performance reports, and incorporates some broad principles underlying the Program Assessment Rating Tool (PART) of President Bush's administration and many of President Obama's policies related to a highly performing government. The act provides a three tiered approach to performance management that includes four-year strategic plans, annual performance plans, and high priority goals.

Of special interest to evaluators are the following provisions:

1. Evaluation policies carried over from the original GPRA legislation, including a definition of "program evaluation" as "an assessment, through objective measurement and systematic analysis, of the manner and extent to which Federal programs achieve intended objectives;" and requirements to describe program evaluations used in establishing or revising general goals and objectives in agencies' strategic plans and provide a schedule for future program evaluations, evaluate agency performance plans against performance goals, and include in annual performance reports a summary of relevant program evaluation findings.

2. New evaluation requirements: A requirement for the "Director of the Office of Personnel Management, in consultation with the Performance Improvement Council, . . . [to] identify the key skills and competencies needed by Federal Government personnel for developing goals, evaluating programs, and analyzing and using performance information . . ."

3. New roles for the Office of Management and Budget (OMB): The responsibility to assess program performance and to inform the agency, the Congress, and the Government Accountability Office of unmet goals. The head of the agency may need to prepare plans to correct performance deficiencies.

4. Transparency: The establishment of a Federal website to publish performance goals and assessments.

There is much good news here in the bill's retaining (and thus emphasizing) evaluation as a central aspect of performance management. The new requirement to identify key skills and competencies for evaluating programs will also have a strong and enduring impact on Federal evaluation functions, depending on how it is implemented. A key concern of federal staff may well be in the magnitude of the administrative tasks and the feasibility of carrying out all the requirements of the law. There is also some uncertainty about how active OMB will be in its independent assessment of performance.

This is necessarily a very brief summary of a law that will profoundly impact the management and assessment of Federal programs for years to come. We will discuss this more in the future. Meanwhile, as always, it will be helpful for us to hear from you about your comments and concerns.

[Go to the EPTF Website and Join the EPTF Discussion List](#)

2/11 – Evaluation in the President's 2012 Budget

On January 14, President Obama released his 2012 budget proposals. What does the Obama budget say about evaluation? *Winning the Future: President Obama's Budget* repeatedly asserts the importance of evaluation in helping to determine what works and promotes evidence-based decision making. The budget narrative calls for "appropriately rigorous" evaluations employing "strong study designs" that "address important actionable questions." It proposes to fund a selection of strong evaluations in the Departments of Education, Health and Human Services, and Treasury; and at the Millennium Challenge Corporation, the United States Agency for International Development, the National Aeronautics and Space Administration, and the Office of Personnel Management. Other efforts to reinvigorate evaluation activities across the Federal Government include an evaluation-focused inter-agency working group to promote information sharing.

Want to learn more? Try these resources:

- The [Analytical Perspectives](#) chapter offers a section titled [Performance and Management](#) that gives an overview of the Administration's FY2012 Evaluation Initiative. It also contains a truly fascinating picture of where OMB stands, not only on evaluation, but on social indicators, cost-benefit analysis, etc.

- The [Terminations, Reductions and Savings](#) chapter recommends programs for either termination or education. You will see the significance that program evaluation plays in the budget narrative. The administration often cites "lack of evidence" that a program is working as a reason for recommending a program be cut or redirecting funding to other initiatives with a stronger focus on evidence-based practices.
- The [White House budget website](#) contains the full-text budget as well as tools for exploring the details.
- [AEA's 2-page compilation](#) of evaluation-related references in the federal budget.
- On February 17, each Federal agency was to post to their website a more detailed explanation of the President's budget proposal for that agency. These are known as the Agency's "Congressional Justifications" where you can find more about an agency's funding for program evaluation studies and evaluation capacity.

As always, it will be helpful for us to hear from you about your comments and concerns.

[Go to the EPTF Website and Join the EPTF Discussion List](#)

3/11 - Defending Evaluator's Independence

One of the most valuable assets and service offerings of evaluators is our independence. It is not just something we have, but also something we give. Our independence makes our evaluations more valuable to our clients and to the stakeholders of the programs we review. This is also true for our kindred professionals - social researchers, policy analysts, program analysts, and the like. Threats to our independence do loom from time to time, for some more commonly or seriously than others.

An excellent analysis of the potentially serious adverse results of such threats is found in a March 3 New York Times article by Ian Urbina, "[Pressure Limits Efforts to Police Drilling for Gas.](#)" which documents political and industry efforts to suppress the findings of Environmental Protection Agency (EPA) research on and recommendations regarding hazardous waste from a gas-drilling technique known as "hydrofracking." Unfortunately, as the article says, "More than a quarter-century of efforts by some lawmakers and regulators to force the federal government to police the industry better have been thwarted, as EPA studies have been repeatedly narrowed in scope and important findings have been removed."

Your AEA leadership, including members of the Board of Directors and the Evaluation Policy Task Force were alarmed enough about this to write to the Editor of the Times. Here is the text of the letter signed by AEA President Jennifer Greene on March 7.

"Dear Editor,

The problem of political pressure on policy-oriented research and evaluation findings, such as that documented in Ian Urbina's recent article, goes far beyond environmental issues, urgent as these are. In a democratic society, policymakers and citizens alike need carefully collected evidence and even-handed evaluations about government initiatives that render their effects transparent to the public

gaze. But, such evaluations cannot be credible or serve democratic interests if they are politically manipulated.

Political pressures on evaluators have been endemic since time immemorial, in both Democratic and Republican administrations. This is why the American Evaluation Association, in its recent policy statement on government evaluation (*An Evaluation Roadmap for a More Effective Government*), has insisted on independence in the design, execution, report-writing, publication, and dissemination phases of their work.

Thank you for bringing this important matter to public attention, and especially for the article's careful reporting and documentation.

Jennifer Greene, Ph.D.
President, American Evaluation Association"

While the letter was not published, AEA is on record in its defense of independence for evaluators and the related professionals from whose ranks most of us have sprung and with which we remain associated. We wanted you to know. As always, it will be helpful for us to hear from you about your comments and concerns.

[Go to the EPTF Website and Join the EPTF Discussion List](#)

4/11 – Evaluation of International Development

Earlier this year, the U.S. Agency for International Development (USAID) announced a sweeping new evaluation policy [<http://www.usaid.gov/evaluation>] that greatly expands and improves the conduct and use of evaluation as an integral part of USAID's planning, programming, and implementation. Of particular note are:

- integration of evaluation and program planning
- definitions and distinctions of various types of evaluations that together cover the life cycle of programs. These include *impact evaluations* that measure the change in a development outcome that is attributable to a defined intervention; *performance evaluations* that focus on descriptive and normative questions about what programs have achieved (either at an intermediate point in execution or at the conclusion of an implementation period); and *performance monitoring* through performance indicators that reveals whether desired results are occurring and whether implementation is on track.
- requirements for at least one performance evaluation for each major program and untested and innovative intervention; and that major interventions be subject to impact evaluations whenever feasible
- acknowledgement of the need for both qualitative and quantitative methods

- a clear statement that no method is superior to others but that methods must be chosen that are appropriate for the program to be evaluated and the evaluation to be performed, and
- a 3% set aside of major program office funds for conducting evaluations.

This policy is a good example of the influence of AEA's Evaluation Policy Task Force (EPTF). In many ways, the USAID mirrors AEA's *Evaluation Roadmap for A More Effective Government*. This is no accident. We provided a copy of the *Roadmap* to senior evaluators at USAID and were told that the USAID policy team consulted the *Roadmap* in developing its policy. Ruth Levine, USAID's Deputy Assistant Administrator, Bureau of Policy, Planning and Learning, who chaired the internal group that prepared the policy, invited AEA's Executive Director Susan Kistler to send an AEA representative to a reception and discussion of the USAID policy hosted by Georgetown University's Mortara Center for International Studies. I was privileged to represent AEA. She also invited EPTF Chair Patrick Grasso to attend. In her opening remarks, Ruth cited AEA's *Roadmap* as a significant resource for her group's efforts. In turn, what USAID has done here could well serve as a model for what other Federal agencies might do.

This is a necessarily brief summary of a policy that could profoundly impact our country's international development efforts for years to come. I suggest you start with [Ruth's blog](#) and then read [the policy](#) itself. We will discuss more about evaluation of international policies and programs in future columns. Meanwhile, as always, it will be helpful for us to hear from you about your comments and concerns.

[Go to the EPTF Website and Join the EPTF Discussion List.](#)

5/11 - Safeguarding Evaluator Independence

In March, we discussed the topic of threats to evaluator independence. This was spurred by a *New York Times* article about political pressures to suppress policy research on the dangers of gas-drilling techniques known as hydrofracking. AEA President Jennifer Greene signed a letter to the *Times* editor emphasizing the need to shield evaluators from such interference.

This leads to the more general question of how to safeguard evaluator independence through evaluation policies. Fortunately, Frederick M. Kaiser and Clinton T. Brass of the Congressional Research Service (CRS) developed an excellent treatise on this subject, *[Independent Evaluators of Federal Programs: Approaches, Devices, and Examples](#)*.

There is no simple answer to the question of just how independent an evaluator should be. For example, reasonable people may legitimately disagree about how broadly an evaluation report's distribution should remain entirely within the discretion of the evaluation commissioner. Similar disagreements can arise with respect to an evaluator's right to access certain data sets, or whether evaluators should include recommendations in their reports. There are also disagreements regarding circumstantial factors that may affect independence such as whether evaluators are employees of the agency whose programs are being evaluated or work for consulting firms who are paid by such an agency.

The CRS report points out that one of the best ways to protect evaluator independence is through formal evaluation policies. This has been done through legislation for the Government

Accountability Office and inspectors general. But it might also be accomplished through administrative procedures. The CRS report describes possible attributes of such independence, such as criteria for selection of the chief evaluator, recruitment of evaluators, tenure, funding, supervision, purpose of the office, scope and type of evaluations authorized, standards and procedures, reporting schedules, report availability and dissemination, and obligations of the evaluated program agencies to respond to evaluators' findings and recommendations.

The report ends by concluding that: "independent evaluators follow no single path or set of directions. Instead, they reveal numerous ways and directives for possible approaches to assess federal programs; provide relevant information and data to the executive, legislature, stakeholders, and the general public; enhance oversight of affected programs; and aid in the development of new legislation or executive directives."

AEA's [Evaluation Roadmap for a More Effective Government](#) provides advice on evaluator independence in several places, including the following:

"Independence. Although the heads of federal agencies and their component organizations should participate in establishing evaluation agendas, budgets, schedules, and priorities, the independence of evaluators must be maintained with respect to the design, conduct, and results of their evaluation studies." (p. 8)

Evaluators are often asked or need to talk about their independence with clients, and they may be called upon by agency officials to help develop policy or procedures about it. They may find the AEA Roadmap and the CRS report a handy resource for insight on this matter.

[Go to the EPTF Website and Join the EPTF Discussion List.](#)

6/11 – Rounding Out U.S. International Evaluation Policy

In April, we described a major breakthrough in U.S. evaluation policy for international development. The U.S. Agency for International Development (USAID) had announced a new [evaluation policy](#) to expand and improve the conduct and use of evaluation as an integral part of USAID's foreign assistance programs.

On May 12, another shoe dropped. The State Department issued an even broader, [overarching policy](#) covering all of its programs, including diplomacy as well as foreign assistance. Its importance is highlighted in the opening paragraph:

"The policy supports the Department's goal of connecting evaluation, an essential function of effective performance management, to its investments in diplomacy and development to ensure they align with the agency's overarching strategic goals and objectives."

Important features of this policy include:

- requiring that all major programs be evaluated at least once;
- requiring evaluation plans to be prepared annually by each bureau that includes: a list of projects and programs and the strategic goal(s) that the projects and programs support; a status report on current evaluation efforts and resources and on recently completed

evaluations; a plan for conducting new evaluations; and a discussion of proposed use and dissemination;

- recognizing the need for and requiring arrangements for the funding of evaluation. (The policy notes that "the cost of an evaluation will vary by program, and no set amount is prescribed, although industry averages suggest that 3-5% of the program cost is a reasonable baseline.");
- linking evaluations to strategic and program planning;
- making the Chief Performance Officer responsible for reviewing Bureau evaluation plans;
- recognizing the importance of both internal and independent evaluations;
- recognizing the importance of evaluation for new program requests;
- maintaining an archive of completed evaluations; and
- encouraging posting of evaluation reports and results on the Department's intranet.

I describe this new policy as "another" rather than "the other" shoe dropping since U.S. international programs are not limited to the State department (including USAID). Significant international policy and programs are formulated and carried out by other departments and agencies, including but not limited to the Department of Agriculture, Health and Human Services, Defense, and Labor. U.S. international evaluation policy will not be fully developed until these other organizations make evaluation an essential feature of their programs as well. This is already beginning to happen in the President's Emergency Program for AIDS Relief (PEPFAR) and in the Centers for Disease Control and Prevention's programs for malaria and tuberculosis. A more recent development is a growing interest in the use of evaluation in international food programs of both USAID and the Department of Agriculture. Evaluation, including the relevance of AEA's *Evaluation Roadmap*, is being considered for discussion at an upcoming International Food Assistance Development Conference in Kansas City, Missouri also this month.

AEA has been on the forefront in advancing evaluation policy in U.S. international programs, and we are pleased to see ideas such as those expressed in AEA's *Evaluation Roadmap* as well as by our colleagues in the various Federal departments and agencies take root.

[Go to AEA's Evaluation Policy Task Force website page](#)

7/11 - The Roadmap Gets Legs

AEA published its most recent version of *An Evaluation Roadmap for a More Effective Government* in September 2010. AEA and the Evaluation Policy Task Force have used it in commenting on emerging national policies such as maternal, infant, and children's home-visiting programs, health care reform, and foreign assistance. A broader goal was to use it as AEA's "place at the table" when policies were being considered, even if an AEA representative couldn't actually be there in person.

The latter goal is now being realized. The *Roadmap* has been cited in congressional testimony, Government Accountability Office (GAO) reports, and other settings. AEA's Executive Director Susan Kistler recently posted on the AEA website the results of her search for [citations to the Roadmap](#) in public documents. Here is a brief summary.

Criteria for Effective Research and Evaluation Programs

Four GAO reports cited the *Roadmap* as the criteria against which the adequacy of Federal agencies' research and evaluation programs could be judged: the McGovern-Dole Food for Education program, USAID's international feeding program, Afghanistan drug control programs, and the Department of Labor's Employment and Training program.

Establishing and Prioritizing Research and Evaluation Plans

GAO cited the *Roadmap* in a report evaluating the methods used by Federal departments to plan their research programs. An IBM Center for the Business of Government blog, commenting on this GAO, further emphasized the value of the *Roadmap* as a useful reference on this topic.

The Need for Federal Program Evaluation and A Framework for Institutionalizing It

NASA, in its report on informal education, cited the *Roadmap* as establishing the need for evaluation and providing a framework for developing an effective evaluation program. Similarly, the Welcome Trust Project quoted the *Roadmap* at length regarding systemic and systematic evaluation of Federal programs. The Economist's Intelligence Unit cited the *Roadmap* in making its point about the need to fully embrace and fund evaluation of Federal programs.

Avoiding the Over-Emphasis on a Single Method for Evaluation

The Brookings Institution Center on Children and the National Institute for Early Education Research, in a report on early childhood policy, used language in the *Roadmap* to emphasize the need for more than randomized clinical trials to evaluate program interventions. Grantmakers for Effective Organizations, in its briefing paper, *Scaling What Works*, cited the *Roadmap* as a basis for emphasizing the context of program decision making in choosing the most appropriate evaluation approaches and methods. Similarly, Deborah Daro, in her testimony before the House Ways and Means Subcommittee on Income Security and Family Support, quoted the *Roadmap* in emphasizing the need to integrate monitoring and evaluation with program management and also to avoid reliance on a single evaluation approach.

Now that the *Roadmap* is catching on, those of you with connections to Federal Government agencies might want to share it with your contacts. Maybe it will go viral. If you are aware of any citations, please send them to info@eval.org.

[Go to AEA's Evaluation Policy Task Force website page](#)

8/11 - Do You Want to Help?

I sometimes get offers from AEA members to help in promoting evaluation policies or questions about what they can do to get policy makers interested in the results of their evaluations. I thought it might be useful to offer some suggestions along these lines.

First, let's consider evaluation policy. This has to do with rules, whether formal or informal, that an organization establishes for conducting or using evaluation. Evaluation policies include such things as authorizations, requirements, funding, methods, planning, publishing, and quality assurance for evaluation. They may be promulgated through laws, regulations, administrative procedures, budgets, organizations, and standards. They may be established at the Federal level by the Congress or executive agencies. Similar rules can be promulgated by State and local governments, by foundations, or any organization that wishes to make evaluation part of the way they do business.

AEA has established the Evaluation Policy Task Force (EPTF) to promote efficacious evaluation policies, with a particular focus on Federal policies. It is an advisory body and has no authority to speak on behalf of AEA except when specifically authorized to do so by AEA's President and Board of Directors.

AEA's Board is particularly anxious to encourage AEA members' input to the formulation of evaluation policies. At its most recent meeting (in June) it approved a policy to facilitate such involvement under a variety of circumstances, including long term efforts to produce carefully vetted position papers, such as AEA's [Evaluation Roadmap for a More Effective Government](#), and shorter term policy opportunities such as comments on proposed government regulations. You can go to [AEA's website](#) to get a sense of the kinds of public evaluation positions that AEA takes and on which the Board wants AEA members' input.

One way for AEA members to participate in promoting effective evaluation policies is to alert the EPTF about policy influencing opportunities. This is what happened, for example, with regard to AEA's advice to the Office of Management and Budget about Paperwork Reduction Act requirements that affect evaluators' abilities to conduct surveys. Another AEA member alerted the EPTF about an opportunity to provide technical assistance to congressional staff on funding evaluation of the President's Emergency Relief Program for AIDS. This in turn led to opportunities to influence the development of evaluation policies of USAID.

These are just two examples of AEA members' contributions to evaluation policy. Please do not be shy about alerting the EPTF to policy influencing opportunities. You can reach the Evaluation Policy Task Force via email at evaluationpolicy@eval.org, or by joining the [evaluation policy discussion list](#).

AEA members can also influence evaluation policy by providing policy makers (such as congressional and Executive Branch staff) copies of AEA's formal evaluation policy positions found on the [AEA website](#), including the [Evaluation Roadmap](#).

In a future column I will discuss what AEA members can do to get their evaluation studies in the hands of policy makers.

[Go to AEA's Evaluation Policy Task Force website page](#)

9/11 – Protection of Human Research Subjects – Advice Needed

In last month's column, I mentioned how useful and important it is that AEA members notify us about emerging evaluation policy issues. Well, it happened again! And this one is particularly important. An AEA member sent an email alerting me that the Department of Health and Human Services (HHS), in concert with the Office of Science and Technology Policy, is beginning a major overhaul of government-wide procedures (known as the "Common Rule") pertaining to the protection of human research subjects. HHS issued an advance notice describing current policies, outlining reforms under consideration, and inviting public comment. On September 6, I sent the AEA membership [an email](#) about this unique opportunity to comment.

This policy overhaul is quite important for several reasons. First and foremost, it is about protecting vulnerable populations - ensuring informed consent by subjects of biomedical and social and behavioral research, guaranteeing their privacy, and mostly ensuring that they will not be harmed in scientific experiments.

At the same time, it recognizes that the current procedures may impede the conduct of valuable research. The advance notice describes this fundamental clash of values. It "seeks comment on how to better protect human subjects who are involved in research, while facilitating valuable research and reducing burden, delay, and ambiguity for investigators."

Because of the extraordinary complexity of the human research subject procedures and of the issues involved, it would be inappropriate to outline or summarize them here. In my email to all of you, I highlighted policies related to exemption of most forms of social and behavioral research, simplification of procedures for review of multi-site experiments, and coverage of evaluation. Here I want to single out evaluation. HHS is explicitly "seeking comment on whether and, if so, how, the Common Rule should be changed to clarify whether oversight of quality improvement, program evaluation studies, or public health activities are covered." See question 24 in the proposed rules.

After considering input received as a result of the notice, HHS plans to redraft the Common Rule. When it does, it will again seek public comment. At that time AEA would like to be in a position to submit formal comments. To do that, we need to get a sense of your concerns now. And so does HHS. For all these reasons, the Evaluation Policy Task Force urges you to take this opportunity to weigh in on this important topic.

The first step is to re-read the email sent earlier this month. It includes links to the HHS notice and convenient explanatory materials. It also explains how to make comments and share them with us by the October 26 deadline.

Thank you in advance for your comments. Please don't hesitate to contact me at evaluationpolicy@eval.org if you need additional information about this important policy.

[Go to AEA's Evaluation Policy Task Force website page](#)

EPTF PERFORMANCE SUMMARY APPENDIX O: RECOMMENDED PROCESSES FOR AEA BOARD AND MEMBER INVOLVEMENT IN PUBLIC DOCUMENTS ON EVALUATION POLICY

The following proposed policies were discussed and the key concepts within them approved for inclusion in the Association's policy document by the Board at its Spring, 2011 meeting

I. Procedure for Involving AEA Board and Members in Development of Public Documents Focusing on Evaluation Policy

In order to increase AEA Board and member participation in, and ownership of, the EPTF's work, a process is needed for vetting new policy documents. However, one process cannot accommodate the wide range of circumstances that may arise; experience to date indicates that there are wide variations in the amount of time available for AEA to influence evaluation policy, and the scope of work involved also ranges widely. Therefore, and building on previous decisions, the following guidelines are proposed:

A. Full-fledged process—for new policy of scope similar to *Roadmap*

Invite all AEA members to make suggestions (e.g., through the AEA Website)

- EPTF prepares policy document
- Board reviews policy document, provides comments, and has option to veto progress
- If approved for further action, EPTF revises based on Board comments
- AEA Members invited to make comments on revised EPTF version
- EPTF revises based on member comments
- Board approves to send to members for vote, or requests further revisions
- EPTF revises further if needed (cycle through steps 7 and 8 until approved)
- AEA members vote, up or down
- Presidential Rotation signs if appropriate

B. Moderate process—for comments on policies, regulations, or proposals; with time allowed for each step dependent on allowable comment period

- EPTF prepares policy document, using as a starting point existing vetted documents, such as the Roadmap and previous statements
- Board invited to comment during the EPTF drafting process, and has option to veto progress
- AEA members invited to make comments on EPTF draft
- EPTF revises based on member comments
- Board approves statement, or requests further revisions
- EPTF revises further if needed (cycle through steps 5 and 6 until approved)
- Presidential Rotation signs

C. Expedited process—short-term turnaround, such as letters to editors

- EPTF prepares policy document based on existing vetted documents, such as the Roadmap and previous statements
- President approves statement, or requests further revisions
- EPTF revises further if needed (cycle through steps 2 and 3 until approved)
- President informs Board
- President signs, or the full Presidential rotation signs if desirable and feasible

II. Procedure for Updating the Roadmap

The following outline is provided as a basis for updating the Roadmap. It is based on a two-pronged approach depending on the scope, significance, and urgency of the proposed changes.

A. Major update every four years or more frequently if necessary

- Invite all AEA members, including Board, to make suggestions
- EPTF prepares new version
- Board approves to send to members for vote, or requests further revisions
- EPTF revises further if needed (cycle through steps 3 and 4 until approved)
- AEA members vote, up or down

B. Periodic, Limited Revisions

- EPTF prepares limited revisions based on sporadic AEA member suggestions or emerging circumstances
- Board Members have one week to react to revisions
- Serious concerns or potentially elaborate change negotiations abort the change
- Without Board objection or with easily negotiated changes, AEA publishes revised version with explanation of changes to AEA members

EPTF PERFORMANCE SUMMARY APPENDIX P: DRAFT TALKING POINTS

Drafted by George Grob (Consultant), 11-20-07

Based on Task Force members' comments at meeting of 11-20-07

This document provides draft talking points that will provide a common “sheet of music” for Evaluation Policy Task Force members and the consultant to use in discussing the Task Force’s (and ultimately AEA’s) policy development goals and objectives with AEA members and outside groups (e.g., in the consultative campaign and public presence initiative).

As a way to get started, we present here some tentative talking points in three categories: How evaluation can help; principles of practice; and possible public position on the role of evaluation in public program development and management. It must be emphasized that the talking points will evolve and be revisited from time to time by the Evaluation Policy Task Force. All three parts (or other sections as yet to be determined) do not need to be finished at one time, and some of the points will continue to be discussed within AEA through a variety of venues and mechanisms.

A. How can evaluation help bring about substantial improvement in government programs?

Based on past experience,

- 1) Evaluation can help programs
 - a) achieve objectives, outcomes, and efficiencies
 - b) improve knowledge and understanding
 - c) improve public accountability and documentation of accomplishments
 - d) provide tools for improving the efficiency and effectiveness of both government and private sector (e.g. foundations) programs
- 2) Evaluation can be helpful at all stages of program development, management, and closeout. It is a systematic way of asking questions about program performance and using appropriate analytic methods to:
 - a) help clarify and articulate stakeholders’ expectations for program and help identify and focus on stakeholders’ most important questions about program performance
 - b) identify and summarize relevant research to assess the likely success of the design of a new program (or provision)
 - c) provide helpful feedback to managers on problems experienced in starting up new programs or provisions
 - d) help incorporate tailored, practical accountability provisions in regulations, grants, and agreements
 - e) provide for a staged set of performance measures and studies to provide ongoing feedback on the quality and efficiency of program
 - f) to answer questions that unfold during program implementation about:
 - i) the appropriateness and relevance of the program design to address the initiating problem
 - ii) the causes of any difficulties observed
 - iii) its success in achieving outputs, and short, medium, and long-term outcomes
 - iv) whether it is avoiding negative side-effects
 - v) whether it is addressing concerns of program advocates and critics
 - g) support the sustainability of effective programs by providing data and insights program operations and results

- h) assess the transferability of successful programs to other locations, organizations, and service environments
 - i) obtain lessons for improving the design and execution of the next generation of programs, and
 - j) learn whether a mature program remains relevant and effective after significant demographic, cultural, or technological changes
- 3) Evaluation uses a variety of approaches in accomplishing its work:
- a) Building an appropriate collection of evaluation questions, requirements, administrative structures, and resources into laws and regulations when major programs are being proposed and developed for the first time
 - b) Engaging multiple groups of stakeholders in articulating their expectations regarding the program and what it might achieve. Identifying the evaluation information that different stakeholders expect or look for.
 - c) Doing the same (but at a more specific level of detail) when program regulations and grant announcements are issued
 - d) Assessing the progress and fidelity of implementation and identifying emerging effective practices and implementation problems in the early stages of program implementation
 - e) Developing models of how the program is intended to operate and the outputs/outcomes that are expected (e.g., logic modeling and program theory)
 - f) Establishing performance indicators, output and outcome measures, supporting data to gather information about the success or weaknesses of programs throughout their life cycle and in subsequent reauthorizations and enhancements
 - g) Selecting the most appropriate approaches (including considerations of available resources, schedules, and methodologies) for assessing the effectiveness and impact of programs, and to the extent possible, the likely causes of successes and failures.
- 4) Evaluation uses a wide variety of methodologies to appropriately respond to program and policy issues, including (but not limited to):
- a) logic models and program theories
 - b) needs assessments
 - c) early implementation reviews
 - d) sampling methodology
 - e) compliance reviews
 - f) performance reviews
 - g) case studies
 - h) quasi-experimental design
 - i) randomized field experiments
 - j) special focus studies addressing emerging issues
 - k) performance measurement systems
 - l) meta analysis
 - m) client and participant satisfaction surveys
 - i) These methods are not interchangeable but are used separately or in combination based on a number of factors, especially the kind of policy question to be answered and whether a cause/effect relationship is to be demonstrated.
- 5) Evaluation is an interdisciplinary professional practice, encompassing but much broader than narrow functions or objectives such as performance management and monitoring. . Many evaluators often work collaboratively with colleagues in allied professions such as:
- a) Economics
 - b) Psychology
 - c) Public Administration (Research and Evaluation)
 - d) Performance Audits
 - e) Policy Analysis

- f) Educational Research
 - g) Social Science Research
 - h) Statistics and Operations Research Analysis
- 6) Evaluators can help formulate criteria by which programs can be identified
 - 7) Evaluators can effectively advise on tradeoffs in the use of evaluation resources for particular programs or families of programs
 - 8) Evaluators can help executive agency leaders and lawmakers formulate evaluation goals and strategies to maximize the value of information that can be obtained within the short deadlines and limited resources generally available for policy related evaluation studies,
 - 9) Evaluators can also help them understand the advantages and limitations that result from any combination of evaluation methodologies.
 - 10) Evaluators can advise legislators and executives on ways to improve existing broad evaluation policies such as GPRA and PART
 - 11) Evaluators recognize:
 - a) The advantages and limitations of the various methodologies
 - b) The advantages of using mixed and multiple methods.
 - c) Opportunities for using available data sets in conducting evaluations and the advantages of improving their scope and reliability for this purpose.
 - 12) Evaluators can provide many examples of strong work that attest to the above strategies and functions and that have been successfully used to inform public policy (provide examples)
 - 13) Evaluators use professional standards such as the AEA Guiding Principles and other appropriate professional evaluation standards
 - 14) The field of evaluation is interested in helping public managers and policy officials understand the strengths and limitations of different evaluation approaches and methodologies.
 - 15) The field of evaluation generally and AEA in particular publish broad principles and functional frameworks on the effective performance and use of evaluation and its incorporation into public program laws, regulations, and administrative procedures.
 - 16) The field of evaluation seeks to establish, using the perspectives and views of both evaluators and potential evaluation users (lawmakers and executive managers),
 - a) What they see as the advantages and disadvantages of evaluation
 - b) Which approaches for incorporating evaluation into program development and management process they believe are most useful – and in what contexts/circumstances, and
 - c) Which approaches are problematic
 - 17) The field of evaluation is concerned about questionable or inappropriate requirements in laws, regulations, and administrative functions regarding evaluation practices
 - 18) AEA can advise on challenging, emerging, evaluation policy issues.
 - 19) The AEA is able to facilitate the availability of senior, experienced teams of evaluators to testify at public hearings of program development and accountability

[CAUTION: The AEA needs to be clear about the distinction between its facilitative role in connecting policy makers and experienced evaluators and the fact that this does not constitute formal AEA endorsement of the positions such advisors might take without AEA review or approval (which would be inappropriate except for individuals who serve as witnesses on behalf of AEA).]

B. What are the effective evaluation practices that AEA believes should be adopted in public program administration?

- 1) Adopt AEA Guiding Principles and other appropriate professional standards and cite their adherence to them in their reports
- 2) Consult with all stakeholders in the design of evaluations
- 3) Use multiple methods whenever appropriate
- 4) Match the evaluation methodology to the stage of program development or evolution
- 5) Prepare as full a set of evaluations as possible over the life of a program to ensure that evaluation insights are available at every stage of program development and implementation and that the full impact of the program can be ascertained from the body of work and not just one or two studies.
- 6) Build evaluation components into the program itself, so that output and outcome information begins to flow from program operations as soon as possible and continues to do so (with appropriate adjustments) throughout the life of the program.
- 7) Involve evaluators in the initial design of programs through such approaches as program logic models and broader analysis of environmental systems so that their independent professional advice can promote effective design and program evaluability from the very early stages of a program.
- 8) Promote the formation of evaluation teams with rich and appropriate mixes of disciplines and capabilities to follow the emergence, implementation, and effectiveness of programs.
- 9) Stress the importance of relevant education and experience in evaluation, while recognizing that evaluation is a complex multi-disciplinary endeavor.
- 10) Seek collaborations with other evaluators and professionals from allied fields to develop families of evaluation reports and information for programs, covering the entire spectrum of evaluative information including accountability, efficiency, and effectiveness.
- 11) While seeking advice from all sides, retain control of the evaluation design, performance, and reporting, as evaluative independence is necessary for evaluative credibility and success.

C. What are the hallmarks of an effective national framework for ensuring that effective evaluation practices are routinely adopted in the development and management of public programs?

- 1) Public program evaluators, both within and outside the government, should abide by AEA Guiding Principles and other appropriate professional standards in conducting their work, and should cite these standards in the reports that they issue.
- 2) To the extent practical, mid level goals and objectives, neither too general or specific, of public laws and regulations should be stated up front; or, if this is impractical during the bill drafting or regulating writing stage, the law or regulation should specify requirements or establish a process for the development of such goals and objectives.
- 3) Measures of key program processes and outcomes should be developed while the program is being developed and should be put into place when program implementation begins; however, they should be modified as appropriate from time to time to reflect what has been learned about the program, especially during the early stages of implementation.
- 4) The mechanisms for gathering the data about the measures of program effects should be provided for in the authorizing legislation or mandating regulations.
- 5) Sufficient resources and requirements should be embedded in the authorizing legislation and regulations to provide for the ongoing evaluation of the program through its various stages of implementation. The overall approach should be to authorize and require the periodic evaluation of the program throughout its life so that a rich source of evaluative information is available from a family of studies that surround the history of the program and will be known to policy makers during the cyclical reauthorizations and amendments that are typical of public programs today.

- 6) At a broader level, sufficient resources should be made available and organizations with appropriate authorities and independence should be established to plan and produce, or to oversee the production of, evaluation studies common to families of programs or agencies at various levels with a Federal Department.
- 7) At an even broader level, Departmental or government-wide level, institutions such as the General Accountability Office, Inspectors General, and top tiered evaluation offices reporting directly to the Secretaries of major Federal Departments, should be supported with the resources, organizational independence, competencies, and authorities necessary for the effective evaluation and oversight of public programs.
- 8) In addition, there should be a robust private practice of non-government independent evaluators and evaluation groups with a broad range of viewpoints and capabilities that can provide input and feedback about the performance of public programs.
- 9) Evaluators both within the Federal Departments and agencies and from the private sector should produce a wide range of studies, recognizing the advantages and limitations of various methodological approaches, in order to provide public officials with timely and useful evaluative information and to provide longer-term data that can be used cumulatively to enhance our learning about what works.
- 10) Data collection for programs of national scope should be developed and operated in concert with state and local governments and other major stakeholders in order to promote the efficiency and effectiveness inherent in unified systems.
- 11) University programs that train evaluators should take account of all of the relevant methods available and concern themselves as well with the various program and policy contexts in which evaluation results are used.
- 12) Evaluators should have access to, and they should avail themselves of, continuing professional development and training opportunities.

EPTF PERFORMANCE SUMMARY APPENDIX Q: PROGRAM EVALUATION FOR MANAGEMENT AND RESULTS

The Obama administration faces a national debt in the trillions of dollars, annual deficits in the hundreds of billions, and uncertainties about financial institutions and the economy. At the same time, concerns remain about national security, health care, education, energy development, and many other facets of American life.

Program evaluation can make substantial contributions to address these issues. It uses systematic analysis to answer questions about how well a program or policy is working, whether it is achieving its objectives, and why. The new administration can use it to address new questions about current and emerging problems, stop wasteful spending, increase accountability, support major decisions about program reforms, and improve programs we need.

Use program evaluation throughout the program cycle

The key is to make evaluation integral to managing government programs at all stages, from initial development through start up, ongoing implementation, and reauthorization. Evaluation can be adapted to all part of the program and policy process and can help:

- Make sure that program and policy designs make sense and can plausibly achieve the goals that are set out for them
- Identify problems encountered during start up and correct them before they become permanent features of programs
- Identify and share promising approaches that evolve during early implementation
- Establish expectations and performance standards at program startup and monitor and refine them as the program matures
- Develop appropriate and efficient data collection systems that provide a continuing flow of evaluative information to policy makers and program managers
- Periodically examine selected program features to improve their effectiveness and efficiency
- Assess program outcomes and impacts
- Examine whether a successful program can be transferred to another setting before scaling up

Improve program evaluation infrastructure

Significant progress has already been made in measuring and managing program performance; but more can be done to:

- Invest in a body of evaluative work on individual public programs that can be used on an ongoing basis to guide management decisions about each program
- Include evaluation expectations in authorizing statutes and in executive agency initiatives and program plans
- Establish independent evaluation offices at appropriately high levels of government
- Set-aside sufficient funds to support professional evaluation activities
- Take advantage of evaluation as a formative tool for improving programs
- Expand the use of practical, appropriate models for assessing program impact

EPTF PERFORMANCE SUMMARY APPENDIX R: STATEMENTS APPROVED BY AEA BOARD, JULY 2010

2.1.1.1.1

The consultant and guiding Task Force may advocate, in relation to the U.S. federal sector, on behalf of the Association for the following items in these or analogous terms:

- A broader and more routine application of evaluation requirements to public programs generally;
- The assurance of adequate and consistent funding for evaluation over time and across subject areas;
- The flexible targeting of the size and scope of evaluations to the context, specific policy questions, and individual stage of development within the lifespan of a particular program;
- A clearer understanding of the need for evaluation not only to determine whether a program is promising, but also to determine the criteria for scaling up or transferring a promising program to new and different locations or contexts;
- More consistent use of qualified, experienced evaluators for both the performance of evaluations and the planning and application of evaluation policy;
- The more frequent provision, by evaluators and agencies, of evaluation designs in which the major rationale for the choice of methods lies in their appropriateness for answering the policy questions posed;
- Stronger support for the independence of the evaluation process as a whole, and of the evaluators conducting it; and
- A greater commitment to the transparency and use of the evaluation results.

EPTF BUDGET EXPENDITURES

The American Evaluation Association's Fiscal Year (FY) runs from July 1 to June 30. The EPTF began its work in 2007 and hired George Grob as a consultant in September of that year. The expenditures each year represent primarily the contract with Mr. Grob plus 10% for payment to AEA's Association Management Company to cover staff support and the participation of the Executive Director on this initiative. The jump in 2009-2010 is related to development of the Roadmap for which a professional editor was brought on board to finalize the version of the document to be distributed at that time.

FY	Total
2007-2008	\$35,897.31
2008-2009	\$42,993.29
2009-2010	\$50,719.28
2010-2011	\$47,022.95
Total	\$176,632.83

EVALUATION POLICY ADVISOR'S QUARTERLY ACTIVITY REPORTS

Compiled, November 2011

Table of Contents

Quarterly Reports	Page
September - November 2007	198
December 2007 - February 2008	200
March - May 2008	203
June - August 2008	206
September - November 2008	209
December 2008 - February 2009	211
March - May 2009	214
July - December 2009	216
January - June 2010	221
July - December 2010	224
January - March 2011	228
April - June 2011	231
July - September 2011	234
October - December 2011	NA

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of First Quarter Activities (September –November, 2007)

The First Quarter of activities of the AEA Evaluation Policy Task focused on the broad policies and administrative procedures that will provide the foundation for its future activities and accomplishments. This was accomplished through

- discussions among the Task force Members at 6 formal meetings
- the drafting of several key documents that express the underlying themes, goals, scope, focus, and administrative processes for carrying out its work
- the creation and initial implementation of key administrative tools
- the commissioning of an initial list of policy opportunities

Meetings

I helped to prepare for and follow-up on 6 meetings of the Task Force

- September 12
- September 26
- October 10
- October 24
- November 7 (at AEA national conference in Baltimore)
- November 20

Key Documents

I drafted and then refined several key documents that facilitated Task Force Members' discussions and decisions about the Task force's scope and direction, and administrative mechanisms for achieving its goals. These included

- Preliminary Concept Paper on the Consultative and Public Presence Campaigns
- Talking Points
- Invitation for Nominations of Successful, Useful Evaluations
- Invitation for Nominations of On-Call Experts
- Welcome Message for the EPTF ListServe Discussion Group
- PowerPoint Slides for the Task Force's Session at the AEA National Conference in Baltimore
- PowerPoint Slides for the Public Issues Forum at the AEA National Conference in Baltimore

Administrative Tools

Though primarily the work of AEA's Executive Director, I participated in the development of and am now the day to day administrator of three administrative tools that provide participation in the Task Force's work by AEA members. These are:

- Nomination of Successful, Useful Evaluations
- Nominations of On-Call Experts
- Evaluation Policy ListServe Discussion Group

So far we have received nominations of 8 successful evaluations and 5 on-call experts. The ListServe discussion Group has 17 enrolled members

Initial Policy Opportunities

The Task force has provided me instructions, and I am now making initial preparations, to reach out to

- OMB regarding the President's recent Executive Order on Improving Government Program Performance. I have drafted an email for the Chair's signature, offering assistance to the Associate Director for OMB for Administration and Government Performance and requesting a meeting to facilitate mutual support in improving government performance through evaluation policy
- Presidential Candidates
- Congressional staff, and
- Senior Executive Branch staff and senior evaluators interested in and well positioned to affect evaluation policy

Operational Plan

- I completed a draft plan for the next three quarters of the Task force's first year of operations

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Second Quarter Activities (December 2007 – February 2008)

During the second quarter, the AEA Evaluation Policy Task Force focused on the selection and development of five potential policy initiatives. The policy initiatives and their current status are as follows:

- Large Federal Research Initiatives: preliminary development plan completed; initial meeting held with NIH officials
- National Health Insurance Evaluation: preliminary development plan completed; initial meeting held with senior Brookings policy expert
- Broad Government Evaluation Policies: preliminary development plan completed; initial meeting held with Robert Shea, Associate Director of OMB for Administration and Government Performance
- Education Evaluation Policy: preliminary development plan completed.
- Development Assistance: (preliminary development plan received 3/2/08, after the close of the second quarter)

EPTF Meetings

I helped to prepare for and follow-up on 4 meetings of the Task Force

- December 12
- January 8
- January 30
- February 13

Documents

I drafted and then refined several key documents that facilitated Task Force Members' discussions and decisions about the Task force's scope of work, direction, and administrative mechanisms for achieving its goals. These included

- Guidance for Preliminary Development Plans for Policy Initiatives
- Preliminary Development Plan for National Health Insurance
- Preliminary Development Paper for Government-wide Evaluation Policies
- Draft Agenda for Meeting with Robert Shea of OMB
- Draft Handout Materials

Administrative Tools

I continue to administer the EPTF Discussion Group. I kicked off the discussions of this group in January. Membership of the group has increased to 48.

I also tend to incoming nominations of successful evaluations and on-call experts. These have largely dropped off now. However, there were additional nominations in both categories since my first quarter report. We have now registered 17 suggested successful evaluations and 11 on-call experts.

Outside Meetings

I arranged for and attended several meetings in pursuit of EPTF goals, as follows

Date	Attendees	Subject and Outcomes
2/27/08	Debra Duran and Madeleine Wallace of NIH William Trochim George Grob	Exchanged information about EPTF and NIH evaluation policies; introduced to NIH the idea of a collaborative policy initiative involving large research programs at NIH. Agreed to meet again to discuss more specifics about such an initiative.
2/26/08	Robert Shea of OMB William Trochim George Grob	A “getting acquainted” meeting with exchange of information, ideas, and advice about AEA and OMB PART. Mr. Shea invited the EPTF to submit comments on OMB PART Guidance, to advise the Evaluation Committee of the Program Improvement Council, and to meet with other key OMB staff.
2/25/08	Henry Aaron, Ph.D., of the Brookings Institute George Grob	Discussed possible approaches to building evaluation into major health care reform initiatives of the next Administration. Mr. Aaron suggested several ways to connect with incoming Administration officials, including several promising references for future discussions.
2/25/08	Andy Rock of the Office of Assistant Secretary for Planning and Evaluation (ASPE) at HHS George Grob	Discussed ASPE’s project to assess the influence of evaluations produced at HHS. Mr. Rock provided detailed information about the approach to the project and its schedule. He agreed to stay in touch and keep AEA up to date on this project.
2/7/08	Federal Evaluators in Washington, DC. Federal evaluators from CRS, DOE, DOJ, ED, EPA, GAO, HHS, HUD, State, Treasury, and VA were in attendance. AEA was represented by William Trochim Susan Kistler Stephanie Shipman Katherine Dawes Damon Thompson, and George Grob	This was a listening session for EPTF to hear about the concerns of Federal evaluators. See attached minutes. Parties agreed to stay in touch. Further meeting with evaluators working in research setting may be scheduled.
1/24/08	Heather Foster, legislative assistant to Rep. Diana DeGette of Colorado, Deputy Whip and Vice Chair of the Energy and Commerce Committee George Grob	Purpose was to introduce the EPTF interest in evaluation policy and in building evaluation into legislation, particularly health reform legislation should it be introduced in the future. Explored ways to do this and agreed to stay in touch.

1/24/08	James Scanlon, Deputy Assistant Secretary for Planning and Evaluation (ASPE) (Science and Data Policy), HHS George Grob	This was a listening session to hear perspectives and needs of evaluators from a senior official responsible for health research policy. Agreed to stay in touch. He suggested that AEA could effectively help senior career evaluators by sharing “best practices” on evaluation through seminars and other communications. He provided names of other senior evaluators to talk to.
1/24/08	Barbara Broman, Deputy to the Deputy Assistant Secretary for Planning and Evaluation (ASPE) (Human Services Policy), HHS George Grob	This was a listening session to hear perspectives and needs of evaluators from a senior official responsible for human development policy. Agreed to stay in touch.
12/10/07	Laura Fuller, Senior staff member of the Senate HELP Committee George Grob	Discussed the importance of evaluation from the perspective of a senior committee staff member and sought advice on the logistics of accessing key congressional staff. She said congressional staff are quite interested in evaluation but don’t know how to incorporate it into legislation. She suggested that AEA hold short training briefings for congressional staff on this subject. She also provided helpful advice on keeping up with congressional actions and resources for locating congressional contacts.
12/7/07	Emilia Dissanto, senior staff member of Senate Finance Committee George Grob	She felt that right now there is not a lot of interest in the committees on the use of evaluation, but that previous efforts such as GPRA do demonstrate an institutional interest at least. She thought the idea of making evaluation an inherent part of major initiatives was an attractive one and should apply to general policies such as personnel, conflict of interest, contractor reforms, and competitive sourcing. She also provided tips on how to access key members of congressional staffs.
12/3/07	Jonathan Breul Executive Director, IBM Center for The Business of Government, and Board Member of the Council for Excellence in Government (Jonathan is a recipient of the AEA Myrdal Government Award) George Grob	Discussed a broad ranging set of issue surrounding the establishment of Federal evaluation policy. We covered his concerns about PART (similar to our own) and his interest in a stronger connection between GPRA and evaluation. He suggested contacting senior careerists at OMB and GAO as a way of influencing evaluation policy, since many of them will survive the transition to a new Administration.

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force Summary of Third Quarter Activities (March 2008 – May 2008)

During the three month period of March through May 2008, the AEA Evaluation Policy Task Force (EPTF) continued meeting with senior evaluators and policy makers who use evaluation in their work to understand their concerns about the role of evaluation in public policy making. The EPTF also initiated actions to promote effective Federal evaluation policy in selected areas. Significant breakthroughs occurred in the areas of general government evaluation policies, NIH research programs, and international AIDS relief.

Outreach to Evaluators

I continued meeting one-on-one with evaluators and others in positions to influence evaluation policy. I have attached a summary of my meetings during this period.

Promoting Evaluation Policies

The EPTF had previously identified the following programs as timely opportunities for promoting more effective evaluation policies.

- Broad Federal Government policies, especially those reflected in the Government Performance and Results Act (GPRA); the Program Assessment Rating Tool (PART); and the use of randomized control trials to measure the impact of public programs
- Major Federal research programs
- Health care reform
- International development programs, and
- Science, technology, engineering, and mathematics (STEM) education

For all but one of these areas (international development), the EPTF had previously prepared preliminary development plans to map out the general approach that would be taken. The preliminary plan for International Development was completed in March. During this quarter, I helped edit the preliminary plan for international development. I also met (sometimes together with Dr. Trochim and sometimes on my own) with key government officials in positions to influence evaluation policy in the first four areas listed above. These meetings are summarized in the attachment mentioned above and the status of the EPTF's work in each of these areas is summarized in the following sections.

Now under development is the attached set of draft summary documents that can be used as handouts in outside discussions about evaluation policies. Based on current plans, the number of such discussions is likely to increase in the near future, as discussed in the next section.

Broad Federal Government Policies. As a result of a meeting of Dr. Trochim and myself with Robert Shea, Associate Director for OMB for Administration and Government Performance, on

February 26, Mr. Shea invited AEA to provide comments on OMB guidance related to impact evaluation. I helped the EPTF prepare the attached detailed comments. They were favorably received by Mr. Shea, who distributed them to key OMB and Federal agency officials involved in revising OMB's guidance for PART. In addition, Mr. Shea invited Dr. Trochim to address the initial meeting of the Evaluation Committee of the Performance Improvement Council, which he did on April 14. He was assisted by Susan Kistler, who joined him at his presentation in Washington, D.C. The Council, established by Presidential Executive Order, consists of the collective membership of executive level appointees responsible for overseeing and making recommendations to improve the performance of Federal programs. Katherine Dawes of the EPTF serves as co-chair of the Evaluation Committee. I have continued to remain in contact with OMB staff assigned to the Evaluation Committee and am now in the process of gathering information and formulating advice on how to select appropriate methods to measure the impact of Federal programs.

Major Federal Research Programs. After Dr. Trochim and I met with the two key leaders and staff of the Division of Evaluation and Systematic Assessments (DESA) in the Office of the Director at NIH (Deb Duran, Chief of the Systemic Assessments Branch and Dr. Madeleine Wallace, Acting Chief of the Evaluation Branch of NIH) on February 27, I followed up with Dr. Wallace to discuss policies with regard to evaluation set aside funding of both the Department of HHS and NIH.

Health Care Reform. The EPTF decided to wait until both national political parties have selected their nominees for President and then to approach the nominees' health policy staff with the idea of building evaluation into any health reform legislation the candidates plan to advance if they are elected. In the meantime, the I identified the individuals within the campaign offices of the current candidates who are likely to play significant roles in health policy development.

The President's Emergency Program for AIDS Relief (PEPFAR). The EPTF is working with the Lundy Foundation to refine language in the reauthorization and appropriations bills of PEPFAR to require appropriate evaluation activities to be carried out under this program. The Lundy Foundation approached the AEA asking for help in obtaining funding for such evaluations. The EPTF is assisting them in influencing the Congress to emphasize the need for evaluation of this program, clarify what kinds of evaluation would be appropriate for it, and provide funds to support evaluation activities. I met with Dr. Vincent Dukay of the Lunday Foundation and with Matthew Oresman of Patton Boggs, who has been assisting the Lundy group in developing language for the congressional committees involved in the reauthorization of and appropriations for PEPFAR program. Based on these discussions, I helped the EPTF formulate the attached proposed amendments to the appropriations language prepared by the Lunday foundation and Patton Boggs. As a result of my continuing discussions with them, they revised their language along the lines of the EPTF recommendations. The revised language is attached.

Federal International Development Programs. An interest in systematic evaluation of international development programs is gaining strength within both government and non-government agencies. This is reflected in the attached preliminary development plan for international development which I helped to edit. An outcome of the EPTF's work on PEPFAR

will be the development of model language that can be used in the reauthorization of other Federal international development programs to promote ongoing evaluation of them.

Science, technology, engineering, and mathematics (STEM) education. The EPTF had previously prepared an initial development plan for promoting the use of appropriate evaluation methods in STEM education programs and will turn its attention to carrying out this plan as resources and opportunities allow.

Informing AEA Members

Activities of the EPTF and emerging evaluation issues of interest to it are now being featured in the AEA Newsletter, thanks to the initiative of Executive Director Susan Kistler and Damon Thompson, Communications Officer. I provided editorial comments when requested to do so.

Meetings

The Task Force has met approximately monthly throughout this period and the President, Executive Director, and Consultant, in various configurations, have met regularly in between EPTF meetings in order to prepare for presentations.

Key Documents

During this quarter, I helped the Task Force complete key documents related to its initiatives to influence evaluation policy. Among them are the following attachments mentioned above:

- Comments on OMB guidance related to impact evaluation
- Initial development plan for International Development
- Comments on proposed appropriations language for the PERPFAR program
- Draft policy handouts

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Fourth Quarter Activities (June 2008 – August 2008)

During the three month period of June through August 2008, the AEA Evaluation Policy Task Force (EPTF) continued its outreach to working evaluators, followed up on significant initiatives started in the previous quarter to promote effective evaluation policies in selected areas of the Federal Government sector, and undertook a self-assessment of its first year of operations. We continue to make progress in the areas of general government evaluation policies, NIH research programs, and international AIDS relief through the PEPFAR program. In addition, we expanded our public presence through the publication of three evaluation policy articles in the AEA Newsletter.

Outreach to Evaluators

I attended the Spring 2008 Interactive Workshop of the Washington Research Evaluation Network (WREN) at the George Washington University Marvin Center. Bill Trochim, Susan Kistler, Stephanie Shipman, and Katherine Dawes were also there. Bill made two presentations—one on the EPTF and one on evaluating R&D programs as part of a broader panel on the current state of prospective evaluation methodologies. Stephanie made a presentation as part of the opening plenary session. She discussed how Federal R&D agencies responds to PART and GPRA, and GAO's advice to OMB and the Congress. I chaired a break-out session with a small number of WREN members who had questions about the EPTF initiative.

Promoting Evaluation Policies

I provided intensive support for two of the EPTF's policy initiatives this quarter—broad Federal government policies and PEPFAR.

Broad Federal Government Policies. We followed up on the contacts made in the previous quarter with OMB staff and on the advice that the EPTF sent to OMB on its PART Guidance. In particular, I initiated another telephone conference between Bill Trochim and Robert Shea, Associate Director for OMB for Administration and Government Performance, to follow up on the interview questions we had sent him and also to renew and strengthen our relationship OMB. In addition, I actively maintained contact with Daren Wong of OMB to keep communications open and offer advice on policy activities that were in the making. As a result, OMB invited Bill Trochim and myself to meet and comment on a proposed pilot program to revise OMB's approach to impact evaluation. I met twice with Daren Wong and Katherine Dawes, who were co-chairing a Working Group on Evaluation under the auspices of the government-wide Performance Improvement Council. The first meeting occurred at Daren's office. The second, which included Bill Trochim, was by teleconference. Based on these meetings, OMB made substantive changes to its proposed pilot program. At Daren's request, our comments were made verbally rather than in writing. I have included a summary of those two meetings in the attached list and description of my outside

meetings. I have also attached a copy of the resulting revised protocol which reflects the EPTF's advice on impact evaluation.

I continue to stay in touch with OMB staff on these and other related matters.

The President's Emergency Program for AIDS Relief (PEPFAR). We followed up on the work we had started in the previous quarter to influence evaluation language and budget policy with respect to the reauthorization of the PEPFAR program. After further consultations with Vincent Dukay of the Lunday Foundation and Matthew Oresman of Patton Boggs, they agreed to further changes in their proposed appropriations language that further reflected the advice that we previously sent them. I have attached their revised language for their letter to the House/Senate Budget Conference Committee.

I also made contact with Michele Orza, now at the George Washington University National Health Policy Forum, but previously the Director of the IOM study that evaluated the implementation of the PEPFAR program. I filled her in on our efforts to revise the evaluation language in the draft PEPFAR authorization and appropriations bills. She agreed with our objectives and generally with the nature of our proposed language. I sent her our proposed edits for her review, and she agreed to present them to her contacts in the Senate authorization committee. I have attached copies of the edits we proposed to both the House and Senate bills.

The reauthorization of the PEPFAR program was made without our proposed edits. However, the contacts we made and the discussions we held had the effect of improving our mutual understanding of the role of evaluation in this program and provide a strong basis for seeking improvements in other international development programs. I am also continuing to stay in touch with the Lunday foundation, discussing potential strategies to ensure that evaluation is appropriately funded in the PEPFAR program through the decisions that will be made by the USAID and its international collaborators.

Informing AEA Members

During this quarter, the AEA Newsletter featured three articles on our evaluation policy work: the EPTF advice to OMB on PART Guidance; the interview with Robert Shea; and an explanation of the importance of our work in the PEPFAR program and more generally in the legislative arena. I have attached copies of my drafts of two of the articles (on the Shea interview and the PEPFAR program), and a copy of a rough draft that Bill Trochim asked me to prepare as a starting point for the article that he himself wrote on the EPTF advice on OMB's PART Guidance.

EPTF Self Assessment

I facilitated a self assessment by the EPTF at the end of its first year of operations, to be used in making decisions about its future. I helped Bill Trochim formulate four questions which served as a basis of the assessment. I have attached a copy of my email message to him on this matter. In addition, I prepared several documents which were used by two EPTF working groups, one on the consultative campaign, and one on public presence. This included a detailed accounting of what the EPTF had done in both areas, and two drafts of answers to the four questions posed to the group by

Bill Trochim. I have attached copies of the detailed accounting of what was done and the latest draft of the answers to the four questions.

EPTF Meetings

The EPTF met on June 11, the Working Group on Public Presence met on July 30, and the Working Group on the Consultative Campaign met on August 5 .

Key Documents

In addition to the documents mentioned in the sections above, I finalized the draft of the policy handouts (“leave behinds”) for the AEA Executive Committee and the AEA Board. I have attached a copy of the latest version of this document.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during this quarter, all of them related to the activities described above.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachments

Email from Daren Wong with key documents for OMB’s impact evaluation pilot
Draft language for a letter on PEPFAR to the House/Senate Budget Conference Committee
Proposed Edits to Senate PEPFAR Legislation
Proposed Edits to House PEPFAR Legislation
Draft AEA Newsletter Article on Robert Shea Interview
Draft AEA Newsletter Article on PEPFAR Legislation
Draft Article for Bill Trochim on OMB PART Guidance
EPTF Self Assessment—Email to Bill Trochim RE Four Questions
EPTF Self Assessment—Actions to Date for Consultative Campaign
EPTF Self Assessment—Actions to Date for Public Presence Initiative
EPTF Self Assessment—Draft of Answers to Four Questions
Draft Policy Handouts
Outside meetings (June through August 2008)

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities for September - November 2008

During the three month period of September through November 2008, the EPTF focused its attention on the AEA national conference in Denver, the future of the evaluation policy initiative, tactics for influencing Federal evaluation policy in the incoming OBAMA administration and the new Congress, and on the evaluation of the PEPFAR program.

AEA National Conference

Evaluation policy was the theme of AEA's national conference in Denver, Colorado, during the week of November 3. In addition to Bill Trochim's presidential address on this subject, numerous panels and other presentations were offered on this topic. The EPTF hosted a panel session to provide an update to AEA members on its work and to once again listen to the concerns of AEA members. I co-chaired this meeting with Bill, one which was characterized by active involvement of the audience and a healthy exchange of ideas on next steps. I also made my own presentation, a "demonstration" or instruction on the relationship between evaluation and policy. I also assisted Leslie Cooks in a more intimate meeting with the winners of the student travel awards. In addition, I joined Susan Kistler in a meeting with Miguel Fontes and other representative of Brazil to discuss ways in which AEA might assist Brazil evaluators in promoting the integration of evaluation into its national programs.

Future of the Evaluation Policy Initiative

With the completion of the first year of the evaluation policy initiative, I met with Susan Kistler and President Bill Trochim in discussions to prepare the way for its future. These meetings considered the ramifications of the new administrative structure of AEA, involving an important distinction between overall AEA policy, to be established by the Board of Directors, and the implementation of its policies under the Executive Director. We agreed to discuss this in January with the EPTF, and I was asked to prepare a plan of action for the coming year. Among the ideas discussed were the need for criteria for deciding which evaluation policies to pursue and clarifications of which matters which would be of interest to the AEA Board of Directors.

Evaluation Policy in the Obama Administration and the New Congress

During this time the EPTF took actions to influence Federal evaluation policy in the incoming Obama Administration and the new Congress. With input from the EPTF, I drafted a one page summary of the major concepts underlying the need for and ways to ensure a stronger role for evaluation in the development and management of Federal programs. After providing this to OMB staff, we were invited to provide a more detailed version of the paper to help OMB staff in briefing the Obama transition team on institutionalizing evaluation in Federal programs. Again with advice from EPTF, I prepared such a paper and got it to OMB staff in time for them to incorporate many of

our ideas into a paper they were preparing for one member of the Obama transition team that expressed a strong interest in evaluation. I followed this up with a meeting at OMB to further exchange ideas and tactics for promoting the idea of building evaluation into federal programs.

Outreach to Evaluators

While unintended, I distributed the draft paper to the Evaluation discussion Group. We received helpful comments from several members as well as expressions of their gratitude for being invited into the process.

PEPFAR

The President's Emergency Plan for AIDS Relief (PEPFAR) program was reauthorized this year. The focus on its evaluation has now shifted to the actions that will be taken by the Executive Branch and several other organizations, such as the Institute of Medicine, GAO, and the relevant Inspectors General to conduct evaluation studies mandated in the new law. I followed up with the Lundy Foundation and others to be gin the process of strategizing about how to best influence these future events, especially how to ensure that adequate resources are budgeted for evaluation of the program.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during this quarter, all of them related to the activities described above.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, offering ideas and suggestions for consideration by the members, and handling other administrative matters.

Attachments

In addition to the summary of my outside meetings, I have attached the following documents related to the topics discussed above

Power Point Slides Used at the AEA Conference
--EPTF Update and Listening Session
--My Demonstration Session on Evaluation and Policy
On Page Paper on Evaluation Policy
Paper for Institutionalizing Evaluation in Federal Programs

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Second Quarter 2009 Activities

(December 2008 – February 2009)

During the three month period of December 2008 through February 2009, the AEA Evaluation Policy Task Force (EPTF) undertook significant work to reach policy makers about the value of evaluation as an intrinsic feature of government. This took the form of a paper entitled “An Evaluation Roadmap for a More Effective Government,” which AEA sent to the Director of OMB. In addition, new work was started on promoting increased funding for evaluation of the PEPFAR program and preparing advice for the House Foreign Assistance committee on how to incorporate evaluation into a recodification of the Foreign Assistance Act.

Promoting Evaluation Policies

This quarter was quite significant in terms of the EPTF’s connection to and influence on evaluation policy. In November I took the initiative to facilitate the development by the EPTF of a one page paper entitled “Program Management for Evaluation and Results.” I contacted a senior level OMB official, sending him this paper as well as the “Policy Handouts” that the EPTF had approved earlier. This exchange revealed an openness of OMB staff to considering a wider role for evaluation in the routine functioning of the Federal Government. Subsequently, I responded to an urgent request from another OMB staff member who needed a short paper outlining the EPTF’s views on how evaluation could be more effectively used by the incoming Obama administration. The paper was needed to support OMB’s own efforts to prepare within a very short time period its own paper on this subject for a member of the Obama transition team who had requested ideas from and a meeting with OMB staff on this subject. OMB staff used our draft paper to prepare materials to brief that transition team member. Based on this, I facilitated the EPT’s development of a longer paper entitled “An Evaluation Roadmap for a More Effective Government.” Shortly after President Obama’s taking office, that same transition team member took a high level position at OMB. He called for a staff meeting to discuss the role of evaluation in management of the Federal government. Again in response to an urgent call for help, I was able to deliver an advance copy of the “Roadmap” paper to OMB. That document was in the hands of an expanded cadre of senior OMB officials who met to develop a more vigorous government-wide evaluation policy. The document was formally transmitted to the Director of OMB on February 3, 2009, signed by the current, immediate past, and newly elected AEA Presidents. I have begun the process of more widely distributing that paper.

During the last week of this quarter I reconnected with the Vic Dukay Lundy Foundation and Matt Oresman of Patton Boggs to pick up again on promoting increased funding for evaluation of the PEPFAR program. I prepared a spread sheet listing the evaluation activities authorized under the new PEPFAR law, along with the due date and who is responsible for each. This document has been sent to USAID and the Office of the Global AIDS Coordinator (who is responsible for the implementation of PEPFAR) and 30 representatives of large non-government organizations who

were meeting in Washington D.C. to jointly plan their efforts to promote funding of care of disadvantage orphans.

Patrick Grasso joined me at subsequent meeting with Vic Dukay and Matt Oresman of Patton Boggs. There we learned that the Appropriations Conference Committee adopted language that we had recommended promoting increased funding for evaluation of PEPFAR programs.

At the same time, Patrick Grasso and I were asked to be on a small team to prepare proposed legislation requested by the House Foreign Affairs Committee to include in a recodification of the Foreign Assistance Act that the committee is planning on completing during this Congress.

Outreach to Evaluators

Quite by accident I engendered a stronger interest in evaluation policy among members of the Evaluation Policy Discussion Group that we had established but was not active. In the course of circulating a draft of the “Roadmap” paper to EPTF members, I accidentally sent it to the discussion Group as well, asking for their comments. This led to an enthusiastic exchange of ideas within the group and comments that helped to improve the paper.

At the request of the Executive Director, I have begun to prepare a monthly column for the AEA Newsletter entitled “Policy Watch. So far, two of these columns have been published, one introducing the Policy Handouts” mentioned above and another on the “Roadmap” paper. Both have been well received, and another is in preparation on the Obama budget as I write this report. I provided same day messages to the Discussion Group on the publication of these newsletter articles. Membership in the Discussion Group has increased to 88.

EPTF Meeting

The EPTF met on January 12 to discuss the Roadmap paper.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during this quarter, all of them related to the activities described above.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachments

Program Evaluation for Management and Results
An Evaluation Roadmap for a More Effective Government

Letter to Director of OMB forwarding “Roadmap” Paper
Evaluation Provisions of PEPFAR Statute—Responsibility Table
Evaluation Provisions of PEPFAR Statute—Short Version
Evaluation Provisions of PEPFAR Statute—Full Version
Policy Watch--February, Evaluation Roadmap
Policy Watch—December, Policy Handouts
Outside Meetings and Contacts 12-01-08 to 2-01-09

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Third Quarter 2009 Activities (March – May 2009)

During the three month period of March through May, 2009, the AEA Evaluation Policy Task Force (EPTF) produced a sweeping legislative proposal for evaluating Federal foreign assistance programs, sent the “Evaluation Roadmap” to several congressional committees and to the White House Office of Health Reform, prepared an assessment of the EPTF’s activities and recommendations to the AEA Board of Directors for a three year continuation of it, and promoted a livelier discussion of evaluation policy through the Evaluation Policy Discussion List

Evaluation of U.S. Foreign Assistance Programs

The EPTF, through Member Patrick Grasso and myself, joined forces with AEA member Vic Dukay, Mathew Oresman of Patton-Boggs, and Michele Orza, former director of the Institute of Medicine’s evaluation of PEPFAR program, to prepare draft legislation to require evaluation of all foreign assistance programs authorized by the Foreign Assistance Act. This was done at the request of the House Foreign Affairs Committee in connection with a proposed recodification of the Act. Our group prepared a proposal embodying the principles of the AEA EPTF’s “Evaluation Roadmap for a More Effective Government.” We provided the committee a written description of the proposal, draft legislative language, and draft committee report language. Copies of these documents are attached. The committee largely adopted our proposal. In advance of the proposed recodification bill, which will take many months to prepare, Committee Chair, Rep Howard Berman, included the evaluation framework in his proposed bill, “Initiating Foreign Assistance Reform Act (H.R. 2139), which was co-sponsored by Rep. Mark Kirk. A copy is attached.

Assessment of and Recommendations for Continuation of the EPTF

The EPTF was established on July 1, 2007. *The Charge to the Task Force* establishes a requirement for it to report to the AEA Board with its recommendations about the future of the initiative. During this quarter the EPTF prepared a decision paper for the AEA Board describing its accomplishments and recommending a three year extension. The paper also recommends increasing the limit on the number of EPTF members to 10; changing the requirement that the Chair be a member of the Executive Committee to a requirement that at least one member of the EPTF be a member of the Presidential rotation; adopting a policy that clarifies AEA’s role in evaluation policy-shaping; authorizing the Executive Director and the EPTF to routinely advocate for fundamental principles of evaluation practice; and developing a process for vetting documents and statements of the EPTF. A copy of the paper is attached

Evaluation Roadmap Follow-up

I sent the Evaluation Roadmap to the Senate Finance and Ways and Means Committees and to the White House Office of Health Reform in connection with health care reform and to the House

Oversight and Government Reform Committee in connection its draft legislation on performance management improvements.

Newsletter Articles

I prepared the attached articles for the “Policy Watch” Column of the AEA Newsletter

- March: “Obama 2010 Budget”
- April: “Looking for Policy in All the Right Places”
- May: “Starting Out on the Right Foot in Foreign Assistance”

Outreach to Evaluators

I revitalized the Evaluation Policy Discussion Group through posting 6 messages on

- Foreign assistance bill of Rep. Berman, Chair of the House Foreign Affairs Committee
- The President’s Budget agenda statement regarding a high performing government
- New legislation on evaluation and performance management developed by the House Oversight and Government Reform Committee
- Nomination of Jeffrey Zients to be Deputy Director of OMB for Management and the Chief Performance Officer
- Evaluation provisions in President Obama’s budget
- The AEA “Evaluation Roadmap for a More Effective Government”

The Discussion Group now has 106 members

EPTF Meetings

The EPTF met on March 12 and May 12 to complete its assessment of the first two years of activities of the EPTF and to develop recommendations to the AEA Board of Directors for a three year continuation of it.

Other Activities

I have actively maintained contact with OMB staff, and prepared for presentations on evaluation policy and evaluator thinking skills for the AEA/CDC Summer Institute.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during this quarter, all of them related to the activities described above.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachments

Draft Legislative Language for Foreign Assistance Evaluation

Draft Legislative Proposal (Explanation) for Foreign Assistance Evaluation

Draft Committee Report Language for Foreign Assistance Evaluation

Draft “Initiating Foreign Assistance Strategy Bill,” H.R. 2139

Draft Assessment of the EPTF and Recommendations to the AEA Board

March Policy Watch Newsletter Column—Obama 2010 Budget

April Policy Watch Newsletter Column—Looking for Policy in All the Right Places

May Policy Watch Newsletter Column—Starting Out on the Right Foot on Foreign Assistance

Outside meetings and contacts 3-1-09 to 5-30-09

Combined Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (July 1 – December 31, 2009)

Significant Headway in Federal Evaluation Policy

During the six month period of June through November, 2009, results of the AEA Evaluation Policy Task Force (EPTF)'s efforts to influence evaluation policy yielded results on several fronts in the Federal Government.

Chief among them was the issuance of a new Office of Management and Budget (OMB) policy on impact evaluation. The new policy reflects many of the principles of the Evaluation Roadmap that three Presidents of AEA sent to OMB Director Peter Orszag in February shortly after he took office. The same three Presidents sent him a letter of support on this new policy. AEA support for the policy was featured in a news article in GovExec.com, which extensively quoted EPTF consultant George Grob on this matter.

OMB also issued priorities for science and technology, also reflecting Evaluation Roadmap principles

The Government Accountability Office (GAO) issued a report on evaluation activities within USAID's international food assistance program. In that report, GAO used the principles embodied in the Roadmap to develop its criteria for conducting its study.

In another GAO report on alternative ways to conduct impact evaluation, GAO cited the EPTF's memorandum to OMB on its PART Guidance to buttress its findings on the relative values of various impact evaluation methods.

Other significant advances were the inclusion of evaluation and accountability provisions suggested by EPTF consultant George Grob in both the House and Senate health care reform bills and significant evaluation provisions in pending House and Senate bills reauthorizing Federal foreign assistance programs.

Other Evaluation Policy Developments in the Making

Both the Senate and House Budget Committees held hearings on government performance and accountability, both of them featuring testimony by OMB Director Peter Orszag. His testimonies made several references to OMB's strong support for evaluation.

OMB has published a request for advice from the public on how to more effectively implement the Paperwork Reduction Act, including ways to reduce paperwork burdens on the public and to better estimate the burdens OMB also wants advice on how to avoid unintended effects of its survey

review process, especially with respect to surveys whose response is entirely voluntary. The EPTF prepared comments that were sent to OMB on behalf of AEA on December 16.

EPTF and the AEA Board of Directors

EPTF consultant George Grob presented a progress report to the AEA Board of Directors at their July meeting and raised the issue of how to effectively vet its public statements with the Board before issuing them. The Board appointed Patrick Grasso as EPTF Chair, thanking Bill Trochim for his two years of service in that role. It approved a two year extension of the EPTF while requesting an evaluation of it by the end of the two year period. It also requested the EPTF to develop for its consideration proposed mechanisms to more effectively vet public statements of the EPTF

At its November meeting, based on recommendations of the EPTF, the Board approved the following policy regarding public statements issued by the EPTF and its paid consultant. “The paid consultant and guiding Task Force may advocate, in relation to the U.S. federal sector, on behalf of the association for the following items in these or analogous terms:

- Broad use of evaluation in public programs, especially those of the Federal Government
- Using methods appropriate to the evaluation questions
- Adapting the size and scope of evaluations to be appropriate to the program’s context and needs (e.g., adapting an evaluation based on the specific stage of development within the life-span of a program)
- Adequate funding for evaluation
- Use of qualified, experienced evaluators as appropriate
- Evaluator independence
- Transparency of results”

The Board requested further proposals for improving the process for vetting public statements of the EPTF for discussion at its February, 2010, meeting and more details on the proposed evaluation at its June meeting.

Outreach to Evaluators

Part of the process that the EPTF has initiated to prepare comments to OMB on the Paperwork Reduction Act (discussed on the previous page) is to survey AEA members, requesting their input to inform AEA’s formal response.

The Evaluation Roadmap and other evaluation concepts promoted by the EPTF were featured in the following presentations made during this period:

- Three webinars on the Evaluation Roadmap for AEA members
- Inspector Generals’ Evaluation Training—Debra Rog
- Department of State Evaluation Conference—Stephanie Shipman and Patrick Grasso
- Environmental Evaluators Network conference—George Grob
- AEA/CDC Summer School—George Grob
- AEA National Conference
 - EPTF Update—Bill Trochim, Patrick Grasso, and George Grob (with Leslie Cooksy, Katherine Dawes, Mel Mark, and Stephanie Shipman, in attendance)

- Open Discussion of New OMB Impact Evaluation Policy—Debra Rog, Patrick Grasso, and George Grob
- Introduction of Evaluation and Policy—George Grob
- Evaluation in the Obama Administration—presidential strand panel, with Stephanie Shipman and George Grob among the presenters
- AEA Newsletter “Policy Watch” Columns
 - Value of Environmental Evaluation
 - Sizing Up Health Care Reform
 - Science and Technology Priorities
 - Next Steps in Foreign Assistance
 - OMB Policy on Impact Evaluation
 - Paperwork Reduction Act

The Evaluation Policy Discussion List has grown to 126 members, and during the last six months messages were posted on 26 topics, mostly related to news items discussed above.

EPTF Meetings

The EPTF met on Jul7 23, September 9, and October 8.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during the last two quarters, all of them related to the activities described above.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

Outside Meetings and Contacts for June-November 2009

Other Documents Available Upon Request

AEA Comments to OMB on the Paperwork Reduction Act

GovExec Article on OMB Policy 10-07-09

EPTF Update for 2009 AEA Conference

Evaluation and Policy for AEA 2009 Conference

Evaluation in the Obama Administration 11-10-09

AEA Letter to OMB Director Peter Orszag Supporting Increased Emphasis on Evaluation 10-16-09

OMB memo on Increased Emphasis on Evaluation 10-07-09

To Andy Schneider RE Oversight and Evaluation of Health Insurance Reforms 10-06-09
To Yvette Fontenot RE Evaluating Health Care Reform 9-14-09
To Yvette Fontenot with comments on health care reform 6-2-09
More comments to Lundy-AEA team on Senate Bill 8-13-09
Comments to Lundy-AEA team Senate Bill 7-23-09
EPTF comments on Proposed Health Care Reform Legislation 7-1-09
EPTF Update for Board of Directors 7-11-09
Assessment and Recommendations to the Board 6-5-09
Webinar on Foreign Assistance Act 6-3-09

Combined Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (January 1 – June 30, 2010)

Significant Headway in Federal Evaluation Policy

The President's Budget. President Obama issued his budget proposals for FY 2011 in February. A prominent feature of this budget is the Performance and Management portion [<http://www.whitehouse.gov/omb/budget/fy2011/assets/management.pdf>] of the Analytic Perspectives, a budget volume that outlines important cross cutting features of the entire Federal budget. The Performance and Management section gives major emphasis to Delivering High Performance Government (Section 7) and Program Evaluation (Section 8). The high performance government section continues, but also refines and prioritizes, longstanding emphasis on the use of performance measures and management processes to get the most of government programs. The evaluation section is new, and it reflects a major emphasis on the role of program evaluation in the planning and management of government programs. This is the first time that a President's budget has given such strong emphasis to evaluation, and the policy it expresses is very much in line with principles of the Evaluation Roadmap that was sent to the incoming OMB Director, Peter Orszag in February 2009. This budget proposal is highlighted in the February 2010 "Policy Watch" AEA Newsletter column

The Budget also makes references to the use of evaluation in connection with past and upcoming budget decisions for various programs. These are more fully described in an analysis of the budget prepared by AEA Executive Director Susan Kistler.

The endurance of the new evaluation related budget policies is reflected in the most recent budget policy issued by the Office of Management and Budget for the 2012 Budget [http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-19.pdf]. It re-emphasizes the importance of evaluation and the essential connection between evaluation and budgeting. It also expands the Obama Administration's commitment to evaluation with the addition of another \$100 million for the development of evaluation capacity in Federal agencies, above and beyond approved agency budgets.

Health Care Reform. National health care reform legislation was enacted on December 24, 2009, shortly after our last quarterly activity report was issued. The new legislation contains several evaluation requirements connected to the authorization and planned implementation of public health and health care systems programs. Of special importance is the inclusion in the bill of a requirement for oversight of health insurance reform initiatives—the leading purpose of the reform bill—by the Inspector General (IG) of the Department of Health and Human Services (HHS). This will bring to bear on the insurance reforms not only the IG's audit and investigative units, but also its evaluation office. The consultant to the EPTF had been calling the need for such a provision to the attention of key staff members of the five congressional committees responsible for the development of the legislation.

A fuller description of these provisions can be found in the May and June “Policy Watch” columns of the AEA Newsletter

Meeting with Associate Director of OMB

On January 5, EPTF Chair Patrick Grasso, EPTF consultant George Grob, and Immediate AEA Past Present Debra Rog met with Shelley Metzenbaum, Associate Director of the Office of Management and Budget for Performance and Personnel Management, and Katherine Stack, OMB Budget Director for Education. The purpose of the meeting was to improve the linkage between evaluation and Federal program management. Both OMB officials expressed a strong interest in collaborating with AEA in promoting a more effective connection between evaluation and improved performance of Federal programs. They put forward a number of ideas, such as more effectively lining of up performance measures with evaluation and involving OMB budget examiners in evaluation training programs. AEA has extended an opportunity for Shelley or some other senior OMB official to interact with AEA members at AEA’s 2010 national conference.

EPTF and the AEA Board of Directors

At its April meeting, the AEA Board of Directors requested the EPTF to provide its advice on three topics:

- 1) Rewording of the seven advocacy principles it had previously approved (see last combined quarterly report for July through December 2009) along with the addition of a new principle emphasizing the appropriate care needed in generalizing the results of evaluations to circumstances outside those of the originally evaluated program
- 2) Revision of the Evaluation Roadmap to eliminate any possible partisan biases implied by the current version and appropriate changes to reflect comments from AEA members, and
- 3) A plan to assess the activities of the EPTF at the end of its current two year authorization period (July 2011)

The EPTF has prepared the requested documents for submission to the AEA Board for consideration at its upcoming meeting in July.

Outreach to Evaluators

More than 50 AEA members submitted comments on the Roadmap in response to an open invitation extended to all AEA members in May.

Evaluation concepts promoted by the EPTF were featured in the following presentations, publications, and EPTF presence during this period:

- Environmental Evaluators Network conference—presentation by Eleanor Chelimski; Susan Kistler, Patrick Grasso present
- AEA/CDC Summer School—George Grob
- Department of State Evaluation Conference—Stephanie Shipman in attendance
- AEA Newsletter “Policy Watch” Columns
 - Rounding Out Health Care Reform Evaluation

- Promising Developments in Health Care Reform
- Request for Comments on Roadmap
- New Teams Forming Up
- Evaluation Policy in the President's Budget
- Pillars of Evaluation Policy

The Evaluation Policy Discussion List has grown to 138 members, and during the last six months 26 messages were posted on 5 topics--evaluation in President Obama's 2011 budget, 2012 budget guidance, the Roadmap, congressional requests for evaluations, and the Cuellar bill.

EPTF Meetings

The EPTF met on January 15, March 18, May 18, and June 11

Outside Meetings

I have attached a list and description of the outside meetings that I attended during the last two quarters.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

Outside Meetings and Contacts for June-November 2009

Other Documents Available Upon Request

What Does the 2011 President's Budget say about Evaluation? by Susan Kistler
 Summary of Meeting at OMB 1-5-10
 Advocacy Statements for Presentation to the AEA Board 6-11-10
 Proposed Revisions to Evaluation Roadmap 6-11-10
 Disposition of AEA Members' Comments on the Evaluation Roadmap 6-11-10
 EPTF Proposed Assessment for Board Consideration 6-16-10
 Copies of "Policy Watch" columns in the AEA Newsletter

Combined Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (July 1 -- December 31, 2010)

The six month period starting on July 1 was characterized by the fulfillment of the Evaluation Policy Task Force's and the AEA Board of Directors' long held goal to develop practical ways to both timely develop AEA positions on emerging policy issues, and to the extent possible, to do so in a way that involves AEA members and Board members in the process.. It was also a period of expanded member involvement in evaluation policy through the formation of an Evaluation Policy TIG.

Publication of Formal AEA Policy Positions.

- **Final Approval of the Evaluation Roadmap.** During May, AEA members were invited to comment on a draft version of "An Evaluation Roadmap for a More Effective Government." More than 50 members responded. Based on their comments, the EPTF revised the Roadmap and presented it to the Board of Directors, who then requested an up or down vote on it by AEA members. The members were given that opportunity in late September and early October and, based on the approval of those voting, the final version was issued in October.
- **Auditing Standards.** The Government Accountability Office (GAO) made an open solicitation for public comments on an updated version of its Auditing Standards (commonly known as the "Yellow Book"). In September, the EPTF invited AEA members to provide comments for possible inclusion in what would become official AEA comments. On a parallel track, the EPTF also invited AEA Board members to offer their comments. One AEA member replied, and the substantive aspects of his comments were included in AEA's official comments, signed by AEA's Immediate Past, Current, and Elected Presidents on November 22.

Overall, AEA's comments reflected an appreciation for important principles and practical advice to evaluators that were contained in the draft auditing standards. However, it also expressed concerns about an unnecessary definition of performance audits and overly rigid requirements regarding the use of criteria that inappropriately constrain all performance audits within the mold of a normative methodology. They also advised that that there would be little support in the world of program analysis and evaluation for using a purely normative methodology to determine program or policy effectiveness, since this approach alone cannot rule out factors other than the program which might plausibly be responsible for the audit findings.

- **Home Visiting Programs.** On August 17, the AEA provided comments to the Administrator of the Health Resources and Services Administration within the federal Department of Health and Human Services on principles for evaluating evidence of the

effectiveness of maternal, infant, and early childhood home visiting programs. The AEA comments made recommendations to: forego assigning an automatic high rating for random assignment designs and automatically relegating all other evaluation designs to moderate or low ratings; avoid using the label "gold standard" in connection with random assignment designs ; use additional criteria to assess the value of impact evaluations; more specifically identify alternative impact evaluation methods that can be used; and emphasize the value of multiple studies and mixed methods.

The narrow time slot allowed by HHS for public comments on the proposal precluded solicitation of input from the AEA membership at large, but the AEA Board of Directors was given an opportunity to review and comment on the draft comments prepared by the EPTF.

Health Care Reform. As noted in my last semi-annual report, national health care reform legislation was enacted on December 24, 2009. Of special importance is the inclusion in the bill of a requirement for oversight of health insurance reform initiatives by the Inspector General of the Department of Health and Human Services (HHS). During this period the HHS Office of Inspector General has begun hiring 35 new evaluators, in addition to numerous auditors and investigators, to help implement this responsibility.

Advocacy Statements. During its July meeting, the AEA Board of Directors approved, with modifications, a final set of evaluation advocacy statements that may be used by EPTF members and its consultant in explaining broad evaluation policies in public settings. The latest version includes a new plank proposed by the EPTF regarding the choice of evaluation methods based on their appropriateness for answering the policy questions posed. The Board, in approving the entire set of eight planks, also edited them in two places in order to give them broader applicability (outside the federal government).

Meeting with the Officials of the Federal Inspector General Community

Patrick Grasso and I met with the Chair and Vice Chair of the Inspection and Evaluation Committee of the Council of Inspectors General on Integrity and Efficiency. The purpose of the meeting was to better inform these key officials of the benefits of closer ties between the Inspector General community and the AEA. We provided them with AEA's Guiding Principles for Evaluators, lists of relevant TIGS, information about accessing the AEA website, and excerpts from the recent scan of AEA members which demonstrated the significant number of evaluation practitioners in AEA as well their fields of interest, many of which are relevant to the inspector general community. We highlighted the many resources available to AEA members. We encouraged them to promote AEA membership for IG evaluators and their attendance at the annual AEA conferences, membership in relevant TIGS, volunteering for leadership positions within AEA, and publication of their work in the American Journal of Evaluation and New Directions in Evaluation. We are waiting to hear back from them about opportunities to address a broader group of IG evaluators. Meanwhile, Dan Levinson, Chair of the Evaluation Committee, has joined the AEA.

AEA National Conference

During its annual conference in November, The EPTF sponsored several sessions. Chief among these was its annual update session, in which it provided an overview of all of its accomplishments during the three years of its existence and outlined major issues it is facing in looking to the future. In addition, it hosted a session on evaluation in the Obama Administration, chaired by Jennifer

Greene. During this session I reviewed all the major policy statements and budget proposals to date, while EPTF members Katherine Dawes and Stephanie Shipman provided their insights based on their connections with OMB evaluation policy staff and their appointment as members of the Evaluation Policy Advisory Group of the Federal Council of Performance Officer. Patrick Grasso served as discussant. In addition, I gave a presentation on evaluation and policy, aimed at helping evaluators understand how to effectively influence public policy. The EPTF also sponsored the first meeting of the newly formed Evaluation Policy TIG which was attended by several EPTF members, and it sent observers to session on the GAO Audit Standards.

Outreach to Evaluators

AEA members Patrick Grasso and Bill Trochim and I promoted the formation of an Evaluation Policy TIG, coaching interested members on how to organize themselves and develop proposals for presentation AEA's national conference. Other EPTF members also provided advice and were in attendance at the TIG's first meeting at the national conference.

For the AEA Newsletter I authored six monthly "Policy Watch" columns:

- Ringing in the New Biennium--Top Evaluation Policy Topics (forthcoming)
- AEA Comments on GAO Auditing Standards
- Approval of the Roadmap
- AEA Member Input to Evaluation Policy
- Weighing the Evidence on Home Visiting Programs, and
- Evaluation in the President's Budget

The Evaluation Policy Discussion List has grown to 167 members, and during the last six months 34 messages were posted on 6 topics--top evaluation policy issues, evaluation in international development, the President's message to senior executives, evidence of effectiveness of home visiting program, OMB's new performance management website, and the Senate Budget Committee's field hearing on the "Stat" performance improvement system.

Significantly, the first of these topics--top evaluation policy issues--was the result of spontaneous interest by AEA members themselves who started this string of conversations.

EPTF Meetings

The EPTF met on September 22 to discuss comments on the GAO Yellow Book and on November 11 while at the AEA national conference to discuss broad strategies for the coming year.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during the last two quarters. These include meeting with Lisa Rajigah of the Washington staff of 3Ie; Robert Shea, former Associate Director of OMB for Administration and Government Performance; Sharon Stout of the Department of Education's Evaluation Office; and Dan Levinson and Kathy Butler, Co-Chairs of the Evaluation Committee of the Council of Inspectors General for Integrity and Efficiency.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

- Outside Meetings and Contacts for July - December 2010

Other Documents Available Upon Request

- Power Point presentations made at the AEA national conference on behalf of the EPTF
 - Update on EPTF Activities
 - Evaluation in the Obama Administration
 - An Introduction to Evaluation and Policy
- AEA comments on
 - GAO Auditing Standards
 - Maternal, Infant and Early Childhood Home Visiting Programs
- Final version of "An Evaluation Roadmap for a More Effective Government" as approved by AEA members
- Copies of "Policy Watch" Columns in the AEA Newsletter
- Evaluation Advocacy Statements

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (January 1 -- March 31, 2011)

The three month period starting on January 1 was significant in encompassing the issuance of significant evaluation policies and early signs of more in the making.

Evaluation Policies of the Federal Government

USAID. The United States Administration for International Development (USAID) published a sweeping [evaluation policy](#) covering all of its foreign development programs. It greatly expands and improves the conduct and use of evaluation as an integral part of USAID's planning, programming, and implementation. Policy features of particular note are:

- integration of evaluation and program planning
- definitions and distinctions of various types of evaluations
- requirements for performance evaluation for each major program and untested and innovative intervention and impact evaluation of interventions whenever feasible
- acknowledgement of the need for both qualitative and quantitative methods
- a clear statement that no method is superior to others but that methods must be chosen that are appropriate for the program to be evaluated and the evaluation to be performed, and
- a 3% set aside of major program office funds for conducting evaluations.

The official most responsible for preparing this policy has publically acknowledged the value of AEA's Evaluation Roadmap as a significant influence on the policy.

GPRA Modernization Act of 2010. On January 4, the President signed HR 2142, the [GPRA Modernization Act of 2010](#). The bill amends the Government Performance and Results Act (GPRA) and other statutory provisions related to performance reports, and also incorporates some broad principles underlying the Program Assessment Rating Tool (PART) of President Bush's administration and many of President Obama's policies related to a highly performing government.

Of special interest to evaluators are:

- Evaluation policies carried over from the original GPRA legislation, including requirements to describe program evaluations used in establishing or revising program goals and objectives, to provide a schedule for future program evaluations, to evaluate agency performance plans against performance goals, and to summarize relevant program evaluation findings
- A new evaluation requirement for the " Director of the Office of Personnel Management, in consultation with the Performance Improvement Council, . . . [to] identify the key skills and competencies needed by Federal Government personnel for developing goals, evaluating programs, and analyzing and using performance information . . . ", and
- The establishment of a Federal website to publish performance goals and assessments.

While we cannot prove a direct AEA influence on this legislation, on two occasions I did brief key congressional staff who were preparing this bill on the importance of evaluation for Federal programs and shared with them copies of the Evaluation Roadmap and other materials.

Centers for Disease Control and Prevention (CDC). The Chief Evaluation Officer of CDC asked me for input on key documents he was preparing as a foundation for a major updating of the agency's evaluation policy. I provided comments as requested and have been invited to remain in contact with him as the policy is refined.

Potential Congressional Initiative. At her invitation, I visited with Hillary Daniels, legislative assistant to Senator Mark Udall, who is interested in developing legislation to promote more effective use of evaluation within federal programs. She had previously met with Katherine Dawes who suggested that she contact me. Two options under consideration are to establish a position of Chief Evaluation Officer within each Federal agency and to strengthen the evaluation function within Inspector General offices.

Evaluator Independence

A March 3 New York Times article by Ian Urbina, "Pressure Limits Efforts to Police Drilling for Gas," documented political and industry efforts to suppress the findings of Environmental Protection Agency (EPA) research on and recommendations regarding hazardous waste from a gas-drilling technique known as "hydrofracking." In response to this article, AEA President Jennifer Greene signed a letter to the Times editor on the importance of protecting the independence of evaluators. She pointed out that "evaluations cannot be credible or serve democratic interests if they are politically manipulated" and noted that AEA has "has insisted on independence in the design, execution, report-writing, publication, and dissemination phases of their work." While the letter was not published, it did put AEA on record on this important issue.

AEA National Conference

I prepared two proposals for the 2011 national conference, one on the EPTF update, and the other on my expert lecture on evaluation and policy. However, I have also been advising the Evaluation Policy TIG and others on conference ideas some of which have been formally submitted. This includes a proposal on state evaluation policy and another on the relationship of government policies and quality of care.

Outreach to Evaluators

For the AEA Newsletter I authored three monthly "Policy Watch" columns:

- In Defense of Evaluators' Independence
- Evaluation in the President's 2012 Budget
- Whatever Happened to GPRA and PART?

The Evaluation Policy Discussion List has grown to 178 members, and during the last three months 16 messages were posted on 6 topics--developmental disabilities, evaluator independence, the President's budget, top evaluation policy issues, international evaluation, GPRA Modernization Act, and federal agencies' evaluation agendas.

EPTF Meeting

The EPTF met on January 13 to review priorities for the coming year, including programmatic areas potentially ripe for evaluation policy, procedures for updating the Roadmap, involvement of AEA members and the Board of Directors in evaluation policy decision making, plans for an evaluation of the EPTF, and next steps on the "Building Bridges" initiative.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during the last quarter. These include meetings with Hillary Daniels, Legislative Assistant to Senator Udall; Thomas Chapel and Ann O'Connor of CDC; Vic Dukay, Matt Oresman, Michelle Orza, and Patrick Grasso (our foreign policy "brain trust"); Brian Robinson of the Economic Research Unit; Ruth Levine of USAID; and Sharon Stout of the Department of Education's Evaluation Office.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

- Outside Meetings and Contacts for January - March 2011

Other Document Available Upon Request

- March 7 email to Editor of the New York times with embedded letter from AEA President Jennifer Greene on evaluator independence
- Copies of "Policy Watch" Columns in the AEA Newsletter

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (April 1--June 30, 2011)

The three month period starting on April 1 saw continuing favorable development in evaluation policy and the resolution of significant AEA evaluation policy procedures.

Recent Evaluation Policies of the Federal Government

State Department. On May 12 the State Department issued its [evaluation policy](#). It is broader than the USAID policy that we discussed last quarter's report in that it provides an evaluation framework that applies to all programs, projects, and activities carried out by its various bureaus and missions, including not only international assistance but also its diplomacy activities. Some of its key provisions are:

- Evaluation Requirements. Evaluations are encouraged for programs and projects at all funding levels.
- Evaluation Plans. Each bureau of the Department will develop an annual evaluation plan.
- Evaluation Independence. Evaluations are categorized as either *internal* or *independent*. Internal evaluations are performed at the direction of bureau/office management which plays an important role in performance management and program oversight. Independent evaluations are performed by outside contractors, free from interference from the bureau/office or operating unit that commissions them.
- Evaluation Documentation. Evaluations are to be collected into a central repository. Bureaus and posts are encouraged to post copies of their evaluation reports and results on OpenNet (Department Intranet) or ClassNet website homepages to facilitate learning within the agency and other interested U.S government agencies.

GAO Report on International School Feeding Program. In May, GAO issued a report on the [McGovern-Dole Food for Education Program](#). It uses the AEA Roadmap as a reference for its recommendation to improve evaluation of this international school feeding nutrition program. It basically puts the AEA Roadmap on the same category as GAO's Standards for Internal Control in the Federal Government as a standard for government agencies to follow. The Department of Agriculture agreed to the GAO recommendations regarding evaluation.

Emerging Federal Evaluation Policies

Evaluation of Translational Research. The EPTF approved a new project to provide advice and technical assistance to the NIH-funded Clinical and Translational Science Award Evaluation Key Function Group. NIH is seeking to make evaluation an essential feature of its initiative to integrate the worlds of biomedical research and medical practice, possibly leading to NIH establishing a new National Center for Advancing Translation Science. EPTF Chair Patrick Grasso and policy consultant George Grob will address the group's face-to face meeting in October. Meanwhile, the

EPTF will be invited to comment on the draft papers under development to spell out the role of evaluation in this area.

Potential Congressional Initiative. In last quarter's report I described my initial contacts with Hillary Daniels, legislative assistant to Senator Mark Udall, who is interested in developing legislation to promote more effective use of evaluation within federal programs. Since then, on behalf of the Senator, she has prepared a draft bill that would strengthen Federal evaluation by assigning oversight of agency evaluation programs to the Inspectors General of each Agency. She has requested comments on the draft bill which will be provided by Eleanor Chelimsky on behalf of the EPTF.

AEA Evaluation Policy Development Procedures

The AEA Board of Directors approved two evaluation policy procedures recommended to it by the EPTF

- Involvement of AEA members and the Board of Directors in the process for developing AEA evaluation policies, and
- Updating AEA's Evaluation Roadmap

The common theme of these two procedures is to engage to the extent possible both AEA members and the Board in the development of evaluation policies, while, at the same time, recognizing the need for timely action to influence government evaluation policies when deadlines are short.

Outreach to Evaluators

For the AEA Newsletter I authored three monthly "Policy Watch" columns:

- Evaluation of International Development
- More on Evaluator Independence
- Rounding Out U.S. International Evaluation Policies

During the last three months 6 messages were posted to the evaluation policy discussion list on international development evaluation and the roles of Federal Performance Improvement Officers.

EPTF Meeting

The EPTF met on

- April 4 to discuss processes for updating the Roadmap; the "Building Bridges" initiative; process for engaging Board and AEA members in evaluation policy decisions; and distribution of Roadmap to key congressional and Executive Branch leaders.
- June 10 to discuss Bill Trochim's request for EPTF input on the emerging evaluation policy for clinical translational science awards.

Outside Meetings

I have attached a list and description of the outside meetings that I attended during the last quarter. These include a presentation to the Inspectors General Inspections and Evaluation Round Table;

Cynthia Clapp Wineck, new Director of Learning, Evaluation, and Research at USAID; and the Partnership for Public Service at a panel on the role of Performance Improvement Officers.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

- Outside Meetings and Contacts for April-June 2011

Other Document Available Upon Request

- Copies of "Policy Watch" Columns in the AEA Newsletter

Quarterly Report of George F. Grob

Consultant to the AEA Evaluation Policy Task Force

Summary of Activities (July 1—September 30, 2011)

The three month period starting on July 1 saw the emergence of potentially new evaluation polices on which the EPTF has been preparing or providing advice.

NIH Clinical and Translational Science Awards (CTSA) Evaluation Policy

The Evaluation Key Function Group of the NIH Clinical and Translational Science Awards (CTSA) initiative had requested input and advice from the EPTF on developing an evaluation policy for this activity. The CTSA initiative aims to provide more efficient and effective transitions of biomedical research to practical adoption in national and community level health care delivery systems. The Evaluation Key Function Group will key off AEA's Roadmap to provide more detailed evaluation policy for the emerging CTSA program.

Based on guidance from the EPTF, I worked with former AEA President William Trochim who is chairing the Key Function Group to specify the nature and procedures for EPTF involvement in this project. Subsequently I met with the group to provide initial input and then coordinated input from the EPTF and reviewed two subsequent redrafts of the policy. EPTF Chair Patrick Grasso and I will be making a presentation to the CTSA program officials and group leaders on October 3. That group will provide input and guidance for the evaluation policy. The EPTF will be invited to review and provide advice on the resulting document. The document will credit the Roadmap and the EPTF for setting the stage for and providing advice on the policy.

Potential Federal Legislation to Strengthen Federal Evaluation

At the request of a legislative aide to Senator Mark Udall, I had previously met with her and provided her the AEA Evaluation Roadmap and other materials and advice to assist the Senator in formulating legislation to strengthen the practice of evaluation in the Federal Government. More recently, she provided a draft bill that would accomplish this goal by authorizing and requiring the Inspectors General of Federal agencies to monitor and assess the evaluation programs of their respective agencies.

Due to my current consultancy with the Inspector General of the Federal Housing Financing Agency, I was restricted by a conflict of interest from providing further advice on Federal OIG matters. However Eleanor Chelinski agreed to pick up on this and provided such advice on behalf of the EPTF.

Along the same lines, I met with John Collins, legislative aide for Senator Carper on the Subcommittee on Federal Financial Management, Government Information, Federal Services and International Security, and Luke Holland, who works for Sen. James M. Inhofe, on possible new legislation to ensure periodic review of Federal programs. I provided copies of the AEA Evaluation

Roadmap as well as advice on several options for structuring such legislation. They will be sending me an outline or draft bill for additional comments.

Protection Human Research Subjects

An AEA member alerted me to the fact that the Secretary of Health and Human Services (HHS), in concert with the Office of Science and Technology Policy (OSTP), has issued an advance notice of its intention to update the Federal protections for human research subjects. In response, I prepared and email that was sent to all AEA members inviting their direct input to HHS on this emerging policy. I highlighted for their special attention issues which could be of special importance to evaluators. These included: exemptions from reviews by institutional review boards common methods of social and behavior research and possibly for evaluation as a whole, establishment of uniform mandatory protections of individually identifiable data, and elimination of redundant reviews of multi-site research projects.

Outreach to Evaluators

For the AEA Newsletter I authored three monthly “Policy Watch” columns:

- Human Research Protections
- Do You Want to Help (How AEA Members Can Promote Effective Evaluation Policies)
- The Roadmap Gets Legs

During the last three months 4 messages were posted to the evaluation policy discussion list on: evidence based policy, South African Evaluation Policy, the role of Chief Financial Officers in performance-related activities, and human research subject protections.

EPTF Meeting

The EPTF met on

- September 26 on strategic planning, building bridges, recruiting new EPTF members, human research subjects protections, involving AEA members in evaluation policy, translation research evaluation policy
- July 13 on feedback from AEA’s Board of Directors meeting, translational research evaluation policy, and policy for recruiting new EPTF members.

Outside Meetings

I have attached a list and description of the outside meetings during the last quarter. These include meetings with: two separate congressional staff offices, the Evaluation Key Function Group of the NIH Clinical and Translational Science Awards initiative, the Congressional Research Service, an EPA intern working on evaluation policy at EPA, and a Ph.D candidate preparing her thesis on evaluation policy.

Role of the Contractor

I have continued to provide ongoing support to the EPTF in drafting and editing documents such as those discussed above, proposing meeting agendas, meeting with evaluators and outside organizations, preparing draft correspondence and discussion outlines for the Chair, offering ideas and suggestions for consideration by the members, and handling routine correspondence and other administrative matters.

Attachment

- Outside Meetings and Contacts for July-September 2011

Other Documents Available Upon Request

- Copies of "Policy Watch" Columns in the AEA Newsletter
- Email to AEA members promoting their involvement in Federal policy on human research protections