

# **BEYOND SUCCESS STORIES: MONITORING & EVALUATION FOR FOREIGN ASSISTANCE RESULTS**

*EVALUATOR VIEWS OF CURRENT PRACTICE AND RECOMMENDATIONS  
FOR CHANGE*

**MAY 2009**

**This paper was produced independently  
by Richard Blue, Cynthia Clapp-Wincek, and Holly Benner.**

# CONTENTS

- Author Biographies..... i
- EXECUTIVE SUMMARY ..... II**
- SECTION 1: STATE OF PLAY.....1**
  - I. Background and Problem..... 1
  - II. Modernizing Foreign Assistance Management, Monitoring and Evaluation ..... 3
    - A. A Learning Culture for Foreign Assistance..... 3
    - B. Policies and Processes for Monitoring and Evaluation In a Learning Culture ..... 4
  - III. Changing Foreign Assistance Landscape and Actors in M&E..... 6
    - A. The Office of the Director of Foreign Assistance ..... 6
    - B. The U.S. Agency for International Development (USAID)..... 6
    - C. U.S. Department of State..... 8
    - D. PEPFAR ..... 9
    - E. Millennium Challenge Corporation ..... 9
    - F. The Department of Defense: ..... 11
- SECTION 2: USG M&E PRACTICES .....12**
  - I. Monitoring ..... 12
    - A. Issues ..... 12
    - B. Guidance and Getting the Job done..... 14
    - C. Link between Monitoring and Evaluation:..... 14
  - II. Evaluation ..... 14
    - A. Should every activity be evaluated?..... 14
    - B. How Much Should Be Spent on Monitoring & Evaluation?..... 15
    - C. Methodology- How Rigorous?..... 16
    - D. Expertise and training..... 17
    - E. Using Evaluations in Managing Foreign Assistance..... 18
- SECTION 3: MODERNIZING FOREIGN ASSISTANCE MONITORING AND EVALUATION PRACTICE .....20**
  - A. Create a Center for Monitoring and Evaluation..... 20
  - B. Final words of advice..... 22

## **AUTHOR BIOGRAPHIES**

### **Dr. Richard Blue**

Dr. Blue was Associate Assistant Administrator for Evaluation in USAID responsible for development of the USAID Impact Evaluation Program, and for establishing the Center for Development Information and Evaluation. Since 1997 he has conducted over thirty evaluations and assessments for USAID, mainly in the Democracy and Governance sector.

### **Cynthia Clapp-Wincek**

Ms. Clapp-Wincek was the Acting Director of the Office of Program and Budget for the Coordinator for Assistance to Europe and Eurasia in the Department of State and was the Evaluation Officer for the Africa Bureau at USAID. She has specialized in monitoring, evaluation and strategic planning in USAID and the Department of State as well as over 15 years consulting independently.

### **Holly W. Benner**

Ms. Benner is Assistant Director of the Managing Global Insecurity Project at the Brookings Institution. She has worked in conflict assessment, monitoring, evaluation, and strategic planning for the Office of the Coordinator for Reconstruction and Stabilization at the Department of State, the Office of Conflict Management and Mitigation at USAID, and U.S. Embassy Kathmandu.

# EXECUTIVE SUMMARY

The U.S. Government defines defense, diplomacy, and development as the three pillars of U.S. foreign policy. U.S. development assistance has traditionally fallen under the remit of the U.S. Agency for International Development (USAID). However, with increasingly complex and interconnected global challenges, and the creation of specialized offices and agencies across the U.S. government, delivery of U.S. foreign assistance is now a tangled web that spreads across at least 50 different objectives, 20 government agencies and 50 offices.<sup>1</sup>

The U.S. Department of State, the U.S. Department of Defense, and specialized agencies such as the Millennium Challenge Corporation (MCC) have emerged as players in foreign assistance. USAID now administers less than half of U.S. foreign aid. President Obama and Secretary of State Hillary Clinton have repeatedly underscored a commitment to reform and investment in U.S. development assistance, pointing to its central importance to U.S. national security and restoring America's reputation globally.

Critical to a more effective and efficient U.S. foreign assistance strategy is a robust monitoring and evaluation system that assesses impact of foreign assistance programs, helps agency leaders better target foreign aid, and contributes to improved U.S. development investments globally. Yet, current monitoring and evaluation of most U.S. foreign assistance is uneven across agencies, rarely assesses impact, lacks sufficient rigor, and does not produce the necessary analysis to inform strategic decision making. Many programs are not evaluated at all and others only sporadically.<sup>2</sup>

Done well, monitoring and evaluation can serve multiple stakeholders—from host country citizens and institutions to U.S. decision makers in the field and in Washington. But effective utilization of evaluation does not come without effort. No matter how rigorous or relevant, a completed report is just the first step in a process of dissemination, discussion, and debate. To achieve more effective foreign assistance, U.S. foreign assistance leadership must create a 'learning culture' which values knowledge, and accepts the reality that not all programs will be successful or should be sustained.

## Scope and Intent of This Report

This report takes a unique perspective on the present status of monitoring and evaluation efforts as practiced in the principal U.S. agencies that provide foreign assistance—USAID, MCC, and the Department of State—and makes recommendations for a monitoring and evaluation system within a learning culture. Findings derive from literature review, interviews with senior USG officials and primarily interviews with and survey responses from 'external evaluators'—individuals who conduct evaluations of U.S. foreign assistance programs, either as part of consulting firms, non-governmental organizations (NGOs), or as individual consultants. External evaluators were chosen because: 1) the authors are external evaluators themselves with prior USAID and State experience; 2) in recent years, the majority of evaluations completed of USG foreign assistance programs have been contracted out to external evaluation experts; and 3) evaluators are hired to investigate whether foreign assistance efforts worked, or didn't work, and to ask why results were, or were not, achieved. This gives them a unique perspective.

---

<sup>1</sup> Brookings-CSIS Taskforce, "Transforming Foreign Assistance for the 21<sup>st</sup> Century, Executive Recommendations," (June 2008) 4.

<sup>2</sup> See <[foreignassistancereform.wikispaces.com](http://foreignassistancereform.wikispaces.com)> for a Center for Strategic and International Studies compilation of a bibliography on foreign assistance reform.

## Key Findings

**Monitoring:** The role of monitoring is to determine the extent to which the expected outputs, outcomes or impacts of a program or activity were achieved. When done well, monitoring can be invaluable to project implementers and managers to make mid-course corrections to maximize project impact. While monitoring requirements and practice vary across U.S. Agencies and Departments, the following themes emerged;

- The role of monitoring in the USG foreign assistance community has shifted dramatically in the last 15 years. The role of USG staff has shifted to primarily monitoring contractors and grantees. Because this distances USG staff from implementation of programs, it has resulted within the Agencies in the loss of dialogue, debate and learning from monitoring.
- The myriad of foreign assistance objectives has led to onerous reporting requirements that try to cover all bases.
- There is an over reliance on quantitative indicators and outputs, deliverables over which the project implementers have control (number of people trained) rather than qualitative indicators and outcomes, expected changes in attitudes, knowledge, and behaviors that would be the consequence of the outputs.
- There is no standard guidance for monitoring foreign assistance programs – the requirements at MCC are very different from those at DOS and USAID.

**Evaluation:** There is also great diversity in evaluation policies and practices across USG agencies administering foreign assistance. MCC has designed a very robust impact evaluation system for its country compacts but these evaluations have yet to be completed. The Education and Cultural Affairs Bureau at the State Department has well respected evaluation efforts, but there is limited evaluation work in other bureaus and offices. USAID has a long and rich evaluation history but neglect and lack of investment as well as recent foreign assistance reform efforts have stymied those functions. The survey responses reflect the current state of practice.

*The decision to evaluate: when, why and funding:*

- The requirements on the decision to evaluate vary across U.S. Agencies. There is no policy or systematic guidance for what should be evaluated and why. More than three quarters of Survey respondents emphasized the need to make evaluation a requirement and routine part of the foreign assistance programming cycle.
- Evaluators rarely have the benefit of good baseline data for U.S. foreign assistance projects which makes it difficult to conduct rigorous outcome and impact evaluations that can attribute changes to the project's investments.
- While agencies require monitoring and evaluation plans as part of grantee contracts, insufficient funds are set aside for M&E, as partners are pressured to limit spending on “non-programmatic” costs.

*Executing an Evaluation:*

- Scopes of Work for evaluation often reflect a mismatch between evaluation questions that must be answered and methodology, budget and timeframe given for an evaluation.
- Because of limited budget and time, the majority of respondents felt that evaluations were not sufficiently rigorous to provide credible evidence for impact or sustainability.

*Impact and Utilization of Evaluation:*

- Training on M&E is limited across USG agencies. Program planning, monitoring and evaluation are not included in standard training for State Department Foreign Service Officers or senior

managers, a particular challenge when FSOs and Ambassadors become the in-country decision makers on foreign assistance programs.

- Evaluations do not contribute to community-wide knowledge. If “learning” takes place, it is largely confined to the immediate operational unit that commissioned the evaluation rather than contributed to a larger body of knowledge on effective policies and programs.
- Two thirds of external evaluators polled agreed or strongly agreed that USAID cares more about success stories than careful evaluation.
- Bureaucratic incentives do not support rigorous evaluation or use of findings – with the possible exception of MCC which supports evaluation but does not yet have a track record on use of findings.
- Evaluation reports are often too long or technical to be accessible to policymakers and agency leaders with limited time.

## **Conclusions and Recommendations**

The call to modernize U.S. foreign assistance sets the stage for the necessary reform of the structures, policies and processes that govern monitoring and evaluation in the U.S. foreign assistance community. The organizations responsible for implementing our foreign assistance efforts must be able to manage well, learn from experience, and be broadly transparent and accountable for results to the U.S. Congress and the American people. This requires a transformation in values, practices and organizational capacities of USG agencies responsible for implementing foreign assistance programs.

### **Modernizing Foreign Assistance Monitoring and Evaluation: An Independent M&E Center**

A more robust M&E and learning culture for foreign assistance results will not occur without the commitment of USG interagency leadership and authoritative guidance. Whether or not calls to consolidate agencies and offices disbursing foreign assistance are heeded, the most efficient and effective way to accomplish this learning transformation would be to establish an independent Center for Monitoring and Evaluation (CME).

The CME 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.

The Director of CME must have the status, political support and bureaucratic independence to ensure high-level interagency access and to represent the M&E agenda to other senior policy makers and administrators. If a new Foreign Assistance Authorization Act is passed by the U.S. Congress, the CME would benefit from explicit statutory authority outlining its independence, authorities and reporting responsibilities.

CME responsibilities would include designing a streamlined monitoring system, developing a comprehensive annual evaluation agenda and M&E budget request, reporting annually to the Executive Branch and US Congress, conducting interagency and multinational evaluations for best practices and develop programs to facilitate learning. CME should play a strong role in capacity building including developing common M&E standards and requirements across all foreign assistance agencies, developing career incentives to support a learning culture, mentor and support agencies’ M&E offices, strengthening

training and providing leadership in building host country capacity to monitor and evaluate their own development.

Creating an independent center would be the best foundation for stronger monitoring and evaluation and therefore, greater foreign assistance results. Without the authority to set and enforce uniform standards for M&E practices, recent history has shown that implementing agencies will not produce the kind of systematic and useful products that are possible and necessary. Given the demonstrated tendency to spend as little on monitoring and evaluation as possible, these vital functions will continue to be treated as ad hoc, poorly planned and executed functions, if done at all. Strengthening monitoring, evaluation and results without a center is possible, and clearly necessary, but much less likely.

## **Conclusion**

As the Obama Administration looks to increased investments in foreign assistance and reforms to create a more effective and efficient USG foreign assistance structure, monitoring and evaluation should play a central role. A robust M&E Center, that is supported by senior policymakers and managers, would promote evidence-based results for American dollars.

An effective M&E system will make management of foreign assistance “smarter,” and will be a critical component of a successful foreign assistance reform package. Knowledge-driven planning and decision-making maximizes impact and systematic M&E supports a culture of transparency and learning from experience. As overdue efforts move forward to transform U.S. foreign assistance for the 21<sup>st</sup> Century, it is time to revitalize the monitoring and evaluation function, demonstrating USG commitment to policies and programs based on demonstrated results.

# SECTION I: STATE OF PLAY

Central to a more effective U.S. foreign policy and foreign assistance strategy is a robust monitoring and evaluation system that assesses outcomes and impact, helps agency leaders better target foreign aid, and contributes to an overall improvement of U.S. foreign assistance investments around the world. This report covers a number of issues associated with the present status of monitoring and evaluation efforts as practiced in the principal U.S. agencies that provide foreign assistance: USAID, the Department of State and MCC. It takes as its starting point the 2001 study that the authors wrote for USAID on evaluation experience<sup>3</sup> and is supplemented by the authors' recent evaluation work as well as their work at USAID and Department of State. It draws on observations gathered in a survey of external evaluators – those that are hired to investigate whether foreign assistance efforts worked, or didn't work, and to ask why results were, or were not, achieved. This gives them a unique perspective. The work of the Center for Global Development and the HELP Commission both make significant contributions to this discussion as well.

This report begins with the problem, addresses structural issues and the need for smart management, and introduces the primary actors in foreign assistance. We discuss issues with the practices of how M&E are conducted and conclude with our proposal for the creation of a monitoring and evaluation structure and function that would support strengthened U.S. foreign assistance.

## I. BACKGROUND AND PROBLEM

International development, as determined by the U.S. Government (USG), is one of the three principal pillars of U.S. foreign policy, along with defense and diplomacy. The goal of U.S. foreign assistance<sup>4</sup> is to further America's foreign policy interests in expanding market economies and democratic polities while improving the lives of those in the developing world.<sup>5</sup> The 2006 U.S. National Security Strategy states that the “goal of our statecraft is to help create a world of well-governed, democratic states that can meet the needs of their citizens and conduct themselves responsibly in the international system.”<sup>6</sup>

The USG, through the U.S. Department of State (DOS), the U.S. Agency for International Development (USAID), the Millennium Challenge Corporation (MCC) and other U.S. departments and agencies, provides technical and material assistance to promote security, economic growth, good governance, health, education and resource management. In response to emergencies, the USG also provides humanitarian assistance and disaster relief. In recent years, U.S. assistance has focused on fragile, failed and post-conflict states which following 9/11 are seen as increasingly central to U.S. national security interests. A growing body of assistance practice has also developed with the emergence of transnational threats including global terrorism, international criminal activity, climate change and environmental degradation, and infectious disease. Addressing these 21<sup>st</sup> Century challenges will be critical to

---

<sup>3</sup> Cynthia Clapp-Wincek and Richard Blue, Evaluation of USAID's Recent Evaluation Experience, USAID Working Paper No. 320, June 2001.

<sup>4</sup> In this report we focus on foreign assistance managed by civilian USG agencies—the Department of State, the U.S. Agency for International Development, and the Millennium Challenge Corporation. We have excluded PEPFAR programs only because it was a somewhat unique program and outside of our experience. While DOD is an increasingly big player in the foreign assistance landscape, we have focused only on civilian agencies. DOD remains outside the scope of this study. There are brief comments about both PEPFAR and the Military in Part III of Section I.

<sup>5</sup> [www.usaid.gov](http://www.usaid.gov)

<sup>6</sup> 2006 National Security Strategy at <http://www.whitehouse.gov/nsc/nss/section1.html>.

sustainable development and promoting well-governed and prosperous states that contribute to a stable international system.

The Bush Administration’s “Global War on Terrorism” (GWOT) has absorbed much of the increases in foreign assistance in recent years, and has led to the most significant shifts of foreign assistance responsibilities to the Department of Defense (DOD) and other USG agencies beyond USAID and the Department of State. USAID—which has traditionally led U.S. foreign assistance efforts—administers less than half of U.S. foreign aid while the share administered by DOD has grown from 3.5 percent in 1998 to 18 percent in 2006.<sup>7</sup>

With increasingly complex and interconnected challenges, many additional U.S. departments and agencies have also been added to the foreign assistance mix including the U.S. Treasury, the Department of Justice and FBI, the Department of Homeland Security, the Department of Labor, along with the already well established presence of the Department of Agriculture and the Department of Commerce. In the health field, the Center for Disease Control has been especially prominent. The creation of separate U.S. agencies and offices for development outside of USAID such as the MCC and the U.S. President’s Emergency Fund for AIDS Relief (PEPFAR) has also contributed to a growing fragmentation of U.S. development assistance. U.S. foreign aid is a tangled web that spreads across at least 50 different objectives, 20 government agencies and 50 offices.<sup>8</sup>

Together, U.S. foreign assistance investments should contribute to U.S. economic interests by expanding markets for trade and investment, to U.S. political interests by expanding the number of healthy, democratic states that can serve as global partners, and to U.S. security interests by reducing the threat of transnational challenges—from terrorism and conflict to climate change and infectious disease. This collection of benefits, alongside robust diplomacy and defense, are critical to a just, humane and stable international system, as well as to U.S. global leadership in the 21<sup>st</sup> Century.

Effective U.S. foreign assistance rests upon the ability to carefully analyze and carry out assistance programs that address the key problems in a specific location, to monitor the implementation of that assistance, and to rigorously evaluate and learn from the results. Systematic knowledge-driven monitoring and evaluation make management of foreign assistance smarter by:

- Demonstrating that management is committed to results – if management puts priority on learning about results, it shows their commitment to results;
- A culture of transparency – if you hide your mistakes (as many project managers would like to do), others won’t know to avoid them;
- Accepting enough failure to make risk-taking possible – as long as we learn from the mistakes;
- Designing programs with systematic analysis and the most current sectoral learning – from evaluations and related research;
- Learning from both monitoring and evaluation about the progress and problems of a program to make the right mid-course corrections ;
- Learning together with local counterparts so that we are best informed and together we make the best programming decisions;
- Finding the information necessary to make conflict prevention and mitigation more agile and focused; and
- Informing policy through sufficiently rigorous evaluation of outcomes and impact.

---

<sup>7</sup> Oxfam America, “Smart Development: Why U.S. Foreign Aid Demands Major Reform,” February, 2008.

<sup>8</sup> Brookings-CSIS Taskforce, “Transforming Foreign Assistance for the 21<sup>st</sup> Century, Executive Recommendations,” (June 2008).

This kind of knowledge-driven management reinforces the efficiency and careful use of foreign assistance resources, strengthens program effectiveness in order to have the greatest impact on people’s lives, and therefore, expands our ability to meet our foreign policy and security objectives.

## **II. MODERNIZING FOREIGN ASSISTANCE MANAGEMENT, MONITORING AND EVALUATION**

Calls for reform of foreign assistance have come from many sources in academia, think tanks, foundations, the international NGO community, and current and former USG officials, including former USAID Administrators. Reform proposals stress the need for more coherence, transparency, accountability and learning from experience to improve U.S. assistance. Modernizing foreign assistance requires smarter management and smarter management is necessary to strengthen learning, monitoring and evaluation. There is much about the current practice of M&E that could be improved (and will be discussed later in this report) but changes in how we **do** monitoring and evaluation will not make any difference unless the policies, management and incentives of the foreign assistance organizations are changed.

We all share a common goal. We want our foreign assistance programs to produce demonstrable and positive results, serving our best interests and those of governments and citizens of developing countries. Only a ‘smart’ program can do this – to be “smart”, the foreign assistance policies and the organizations responsible for implementing them must build learning into the process.

### **A. A LEARNING CULTURE FOR FOREIGN ASSISTANCE**

A *culture* means a widely accepted set of norms and behaviors that define what ‘ought’ to be done. We have no magic formula for building a learning culture, but experience does suggest a few clues. First, leadership must value and understand the importance of good monitoring and evaluation, as well as the time, financial and human resource requirements necessary. Second, leadership must have a “center” of competence, expertise, and experience which it can use to inculcate values, provide guidance, and insure high quality work from all implementing agencies. Third, good monitoring and evaluation work must be rewarded with recognition, **strong career incentives**, and policy action. And last, negative findings must be as acceptable as positive ones, at least some of the time. A learning culture once created must be sustained by visible indications that it is valued and is being used to make changes if needed.

Norms imply ethical standards, as well as technical ones, by which an organization carries out its monitoring and evaluation responsibilities. There are many such standards to which evaluation researchers subscribe, including protection of human subjects, transparency as to purposes, confidentiality, and voluntary participation. We single out three which are especially relevant in the foreign assistance business.

1. The obvious one...truth is sometimes hard to accept. It serves the larger purposes of learning and accountability, even when it causes problems, limits choices, and threatens immediate stakeholders. Improving support for monitoring and evaluation will mean nothing unless ‘truth’ is encouraged to speak to “power”. Leadership sets the tone: if leadership asks for truth, the system can be created to provide it.
2. Monitoring and evaluation activities should be conducted with USG partners in the developing world. If we believe that a robust monitoring and evaluation system helps our foreign assistance to become more efficient and effective, then it obviously should be a part of programs to build institutional capacity in our host country governments and civil society sectors.

3. Putting a spotlight on evaluating and monitoring impact highlights its importance and our commitment to achieve positive change in the world.

## **B. POLICIES AND PROCESSES FOR MONITORING AND EVALUATION IN A LEARNING CULTURE**

*Foreign Assistance Outcome Monitoring* – For the senior policy makers and managers in the foreign assistance agencies to know what progress is being made toward results, a set of standard outcome indicators reported by agencies would be very useful. Program outcomes are not “what we are buying with our foreign assistance” but whether that has had the desired result. This is also what needs to be reported to the Congress and the American people to show that foreign assistance makes a difference. If policy makers and senior managers want such outcome reporting, they must:

- **Focus** agency strategy/policy for foreign assistance on a limited set of priorities so that it is possible to determine a reasonably limited set of indicators. The proliferated foreign policy goals we have now are what led to the ridiculously onerous reporting system.
- **Concentrate resources** in sufficient amounts on those limited priorities that you have a reasonable plan to affect outcomes. Currently, money gets stretched across so many “priorities” and earmarks that resources are too rarely sufficient to achieve outcomes; the best that can be done is to implement a series of loosely connected activities.

*Management monitoring* – Program managers need enough systematic data to be able to make good management decisions.

- A prerequisite is that the foreign assistance agencies have sufficient staff working on projects and programs that they have time to think through systematic information such as monitoring and evaluation. Staff’s skill levels in collecting, analyzing and interpreting systematic information also must be enhanced. Even though, they would not be collecting information themselves, they need sufficient skills to know the strengths and weaknesses of the data they are using.
- Program managers will need mentoring and advice on how to think about the outcomes of their programs to meet new priorities, what information they would need to manage well, and ways to get that information;
- Monitoring processes should be participatory between agency staffs, implementing partners and in most cases local partners as well (governments and civil society);
- Foreign assistance programs should build local M&E capacity as an important development activity – the people in the countries we work with should be developing the skills to manage their own development using good systematic information
- The requirement for data quality assessments should be dropped. Our priorities are in the wrong place. We are assessing the quality of data without supporting managers to have the time and resources to be thoughtful about what they are managing. This sends the messages of micromanagement and distrust. It would be a much better use of resources to send technical advisors to help to collect and analyze the data and as mentioned above mentor and support staff in planning for and using information.

*Outcome and Impact Evaluation* – Individual evaluations do not “roll up” into some whole that will tell the story of which foreign assistance approaches work and which do not. This requires multi-country evaluations on a type of project or sector that are designed for learning about the outcomes, impacts and best practices. These types of meta-evaluations require advance planning, robust research designs, and commitment of time and resources to do well. They will not be done if they are left to ‘post hoc’ decision making by field project managers. These are the studies that inform policy makers and senior managers.

*Activity/project evaluations* are an opportunity for managers to build on their systematic monitoring information with external input from evaluators who will address the issues of why the changes reported in monitoring are occurring. Managers need participatory, quick and not-too-expensive approaches. The related activities of semiannual program reviews and after actions reviews can serve some of these same evaluative functions. Not every activity needs a full evaluation during implementation, but managers should carve out time for benefiting from the thinking of implementing partners, clients, government and/or civil society about progress and problems.

An independent *Center for Monitoring and Evaluation* should be established to strengthen monitoring and evaluation capacity across all foreign assistance agencies, to carry out comparative evaluation work and to report to agency decision makers and external stakeholders. Some might question the need for an ‘independent’ Center. In the same way that the independence of the GAO and the Inspectors General gives their reports visibility and credibility, the proposed Center’s independence is critical to its authority and credibility. The Center must function as the highest level of accountability and transparency for insuring that all Federal Agencies providing foreign assistance are implementing effective monitoring and evaluation systems and processes. Properly organized, an independent Center will also strengthen the role of Monitoring and Evaluation Offices and systems in the various agencies. Bureaucratic placement of such a Center depends on the structure of the foreign assistance community in the new administration. Wherever it is placed, it should be able to facilitate “truth” speaking to “power” about such comparative assessments across the whole community. If a new Foreign Assistance Authorization act is passed by the US Congress, the CME would benefit from explicit statutory authority outlining its independence, authorities and reporting responsibilities. The proposed Center is discussed further in section 3.

Yet, despite development’s increasingly important role in U.S. foreign policy, current monitoring and evaluation of most of U.S. foreign assistance is unsystematic, lacks sufficient rigor to assess impact, and does not produce the necessary analysis to inform strategic decision making, or even tactical adjustments. Many programs are not evaluated at all and others only sporadically.<sup>9</sup> The growing fragmentation of U.S. foreign assistance makes systematic knowledge-driven monitoring and evaluation and smart management that much more challenging, but also that much more necessary.

<sup>9</sup> See <[foreignassistancereform.wikispaces.com](http://foreignassistancereform.wikispaces.com)> for a Center for Strategic and International Studies compilation of a bibliography on foreign assistance reform.

## **Reporting**

Some consider “evaluation” to be a tool for program managers and not a tool to assess the outcomes and impacts of foreign assistance more broadly. But with somewhat different approaches, systematic learning efforts can serve a variety of stakeholders:

- Activity monitoring and evaluation are most directly relevant to first-line field managers – what are the progress and problems for their specific activity? Activity (or project) evaluations and monitoring (ongoing data collection) are required in every grant, contract or cooperative agreement. This system is designed to capture what is and isn’t working for the specific activity.
- Agency managers, policy makers, and funders in the Congress need more programmatic (or sectoral/big picture) monitoring and evaluation to inform their decisions. This is most notably on funding choices for sectors and subsectors e.g. more for economic growth and specifically more for microfinance **or** for health and specifically HIV/AIDS. These require multi-country program evaluations and impact monitoring with data collection in a manner than compares apples to apples. Experience has clearly demonstrated that individual activity/project evaluations do not “roll up” to allow conclusions on the effectiveness and impact in a sector. Studies of broader scope and different methods are necessary. USAID’s Democracy and Governance work has recently demonstrated this. Experience has also shown that monitoring at the output level leaves questions about the results of foreign assistance unanswerable.

Local counterparts need information at both levels, just as we do.

### III. CHANGING FOREIGN ASSISTANCE LANDSCAPE AND ACTORS IN M&E

As we have discussed, one of the most notable shifts in U.S. foreign assistance has been the proliferation of U.S. actors who disburse international aid. Because of the longer and deeper experience that USAID has had in monitoring and evaluation, the report addresses issues at USAID with the greatest frequency. However, the following agencies and offices all hold a piece of the puzzle.

#### A. THE OFFICE OF THE DIRECTOR OF FOREIGN ASSISTANCE

The *Office of the Director of Foreign Assistance* (referred to as “F” or more recently “DFA”) in the Department of State. The foreign assistance reforms put in place in 2006 by Secretary of State Rice and Ambassador Tobias as Director of Foreign Assistance were an effort to bring some standardization to the U.S. assistance program but only covered programs at the Department of State and USAID. However, the Millennium Challenge Corporation and PEPFAR, two large and notable programs were excluded, and the F process did not engage DOD—limiting attempts at coherence. The new reforms emphasized centralized planning, the development of a set of common indicators for all sectors of foreign assistance activity, and an increased emphasis on monitoring.

The new guidance and requirements determined shortly after F was created strengthened the focus on accountability – that is, reporting on processes and outputs to external audiences including the White House and Congress. Although some effort has been made to include outcome indicators, F’s original list of over 600 standard indicators (it is now in the 400+ range) were primarily output indicators. These are reported in the Annual Performance Plan and Report through a comprehensive and cumbersome database called the Unified Foreign Assistance Coordination and Tracking System (FACTS)<sup>10</sup>.

One shortcoming had been that F had emphasized budgeting and monitoring and had ignored evaluation. In 2008, F created an evaluation unit and defined the F role as coordinating, facilitating and advocating for:

- policy, guidance, and standards for foreign assistance;
- cross-cutting and multi-agency evaluations;
- technical assistance (advice and support) to program staff;
- analysis of evaluation data including gaps and priorities;
- documentation and dissemination.<sup>11</sup>

#### B. THE U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT (USAID)

The *U.S. Agency for International Development (USAID)*, the agency that was the predominant actor in U.S. foreign assistance efforts, and is now subject to coordination by F. Although USAID’s leadership role in M&E has eroded, the Agency does have longstanding guidance for staff on results management in the unfortunately named Automated Directives System (ADS).

After a period where there was a lack of clarity on whether the ADS was still operative (in spite of guidance that said it was), USAID has issued the revised guidance. There is a Performance Unit within the Office of Management Policy and Budget Performance in the Management Bureau (M/MPBP) that is responsible for monitoring and ADS guidance. The effects of layering F monitoring requirements on top of existing ADS requirements will be discussed further in the section on monitoring.

---

<sup>10</sup> The Foreign Assistance Country Tracking System.

<sup>11</sup> Source: Donna Stauffer, Bureau of the Director of U.S. Foreign Assistance, American Evaluation Association Annual Meeting, Panel Presentation 504, November 7, 2008.

USAID was once considered a global leader in evaluation. With the creation of the LogFrame in the 1970s and first the Office of Evaluation and then the Center for Development Information and Evaluation (CDIE) that undertook evaluation studies and provided support to the rest of the Agency, USAID prioritized and directed Agency resources toward evaluation. Regional Bureaus also had evaluation units that supported their field missions and each mission had someone designated as an evaluation officer.

As demonstrated in the Clapp-Wincek/Blue 2001 meta-evaluation of USAID's evaluation practices, USAID's commitment to evaluation declined sharply in the 1990s, a decline driven by the removal of the mandatory requirement to conduct mid-term and final evaluations on ALL projects, and the rise of performance monitoring planning, which were mandatory for all projects.<sup>12</sup> These two policy changes, in combination with the long term decline in the number of USAID professional staff, made serious attention to evaluation almost impossible.

Following the 2001 study, several USAID professionals with evaluation experience began a long campaign to resuscitate USAID's evaluation leadership. An indefinite quantity contract for analytic services managed by USAID/PPC was used to promote improved quality in the preparation of scopes of work, provide more training, and in some cases, to referee disputes between unhappy field missions and evaluation contractors whose reports were judged 'unacceptable' by local mission staff.

In 2006, these advocacy efforts paid off with the issuance of a directive by Administrator Andrew Natsios, stating that Missions and other operating units should conduct a strategic/program level evaluation of a concluding strategy before embarking on a new one. However, Administrator Natsios resigned shortly thereafter, and the "F" process began, resulting in the demise of the Natsios Directive, along with USAID's Policy Bureau and the Center for Development Information and Evaluation.

USAID is in the process of reconstituting an evaluation function through the new Unit in M/MPBP which partners with the Performance Unit that focuses on monitoring mentioned above. The evaluation function has been downgraded bureaucratically – now it is a unit when previously it was an office. Moreover the choice to put the evaluation function in the Management bureau rather than with the policy function calls into question the purposes, audience and types of evaluation work planned. The new unit has an ambitious set of priorities that will belie the bureaucratic location if they are fulfilled:

- Strengthening evaluation capacity;
- Implementing a new program of more rigorous impact evaluations;
- Working with development partners on collaborative evaluations;
- Participating in evaluation organizations and forums; and
- Providing intellectual leadership.<sup>13</sup>

---

<sup>12</sup> Cynthia Clapp-Wincek and Richard Blue, "Evaluation of Recent USAID Evaluation Experience", Working Paper No. 320, Center for Development Information and Evaluation, USAID, June 2001.

<sup>13</sup> Source: Gerald Britan, Chief Performance Division, USAID American Evaluation Association Annual Meeting, Panel Presentation 504, November 7, 2008.

### C. U.S. DEPARTMENT OF STATE<sup>14</sup>

*In the Department of State*, foreign assistance is hardly a core function. Nonetheless, State Department programs are significant constituting nearly \$10 billion in 2008 (not including assistance to Europe and Eurasia which is State oversight of largely USAID implemented programs) In general, results management within the State Department has focused on responding to the U.S. Congress, Office of Management and Budget (OMB) and Government Accountability Office (GAO). Prior to coming under F's authority in 2006, the primary vehicle for State Department monitoring and reporting on foreign assistance was the Mission Performance Plan (now revised to Mission Strategic Plans). In addition, each bureau has its own unique guidance and practices for program management.

The Bureau of Resource Management (RM/SPP) created an evaluation function in 2008. One of its first tasks was to assess capacity throughout the Department to evaluate programs or functions. The assessment shows that some internal capacity exists<sup>15</sup>:

- Public Diplomacy (PD) has staff devoted to evaluation of PD programs particularly the strong evaluation practices in the office Educational and Cultural Affairs;
- EUR/Assistance Coordinator for Europe and Eurasia has some evaluation staff;
- Bureau for Population, Refugees and Migration has a three-person M&E unit which provides training and is strengthening monitoring practices; and
- Bureau for Democracy, Human Rights and Labor (DRL) and Bureau for Near Eastern Affairs (NEA) have hired evaluation consultants to assess some programs.

It is notable that most of these offices/bureaus have programs similar to USAID, including EUR/ACE which is responsible for the assistance programs at both State and USAID for that region of the world. However, five functional bureaus that disperse international assistance have little or no M&E capacity:

- International Security and Nonproliferation;
- International Organizations;
- Oceans and International Environmental and Scientific Affairs;
- Political Military Affairs; and
- International Narcotics and Law Enforcement.

It should be noted that individual State Department offices are doing some very interesting analytical work, such as INL's Criminal Justice Sector Assessment Rating Tool, but they are not being used as part of the monitoring and evaluation functions. With greater institutional emphasis on a learning culture, there are efforts such as these to build on.

RM/SPP is drafting a policy and undertaking efforts to build up the very minimal capacity for programmatic evaluation in the Department of State. They have also convened a rich program evaluation discussion series, although it may be over the heads of much of their target audience at State.

---

<sup>14</sup> It should be noted that the authors use "Department of State" in this report to refer to programs and policies of those offices in the Department that implement foreign assistance (and occasionally foreign affairs) programs. We distinguish the office of the Director of Foreign Assistance by calling it "F" even though it is a part of the Department. In our survey and key informant interviews, many referred to F as the Department of State—creating some lack of clarity regarding whether they mean F or other foreign assistance programs at State, e.g. "*FACTS is a DOS creature foisted on USAID.*"

<sup>15</sup> Reported by Melinda Crowley from the Office of Strategic and Performance Planning, U.S. Department of State at the American Evaluation Association Annual Meeting, Panel Presentation 504, November 7, 2008.

## D. PEPFAR

The President's Emergency Plan for AIDS Relief (PEPFAR) began in 2003 to combat global HIV/AIDS with tuberculosis and malaria added at the time of authorization. A requirement of the act which established PEPFAR was that the Institute of Medicine (IOM) provide Congress with an evaluation of the initiative's progress three years after the passage of the Leadership Act. In response, the IOM Committee for the Evaluation of PEPFAR Implementation examined the initiative from 2005-2007, including visits to most of the fifteen PEPFAR focus countries.<sup>16</sup>

Many things about PEPFAR are controversial, but "overall, the Committee concluded that PEPFAR has made a promising start, but it warns that there is an enduring need for U.S. leadership in the effort to respond to the HIV/AIDS pandemic." Several recommendations made by the committee touch on monitoring and evaluation:

- A transition from its focus on emergency relief to an emphasis on long-term strategic planning and capacity building for a sustainable response.
- Addressing the long term factors that underlie the epidemics in each country including:
  - Emphasizing and enhancing prevention with accumulation of better data to determine the most appropriate interventions needed in the countries.
  - Expanding the knowledge base by emphasizing evidence-based approaches, learning from experience, and adaptation to new developments; as well as conducting operations research and robust program monitoring and evaluation.

PEPFAR exists within the context of the international health community's strong track record in knowledge-driven performance management, including monitoring and evaluation. The authors are not health experts, but it appears to a layman that some of the exhortations for better data above are reflective of PEPFAR policy choices that were not supported by data and perhaps set it apart from that community. This is too complex an issue to be included in this report other than comments by our external colleagues below.

It should be noted that many of the observations in this report do not apply to the same extent to the health sector within the USG foreign assistance community. Their use of monitoring and evaluation and their learning culture are deep and rich.

## E. MILLENNIUM CHALLENGE CORPORATION

The *Millennium Challenge Corporation* was created in 2004 with the aim of "reducing poverty through growth" in "well-governed countries." Developing countries are eligible to apply for funds across a broad set of program areas if they obtain relatively high scores on a set of 17 indicators within the categories of Rule Justly, Invest in People, and Encourage Economic Freedom.

MCC also implements a program for threshold countries—countries that do not qualify for MCC assistance but are close to qualifying and have demonstrated a commitment to meeting the MCC-eligibility requirements in the future. The program provides 'threshold funds' to support reform efforts in areas identified as weaknesses across the MCC's indicators. Threshold programs are implemented by USAID; therefore the discussion about MCC in this paper refers to the "compacts" that the Corporation signs with Governments.

---

<sup>16</sup> Institute of Medicine, [PEPFAR Implementation: Progress and Promise](#), Washington, D.C., March 2007.

As a program newly created five years ago, MCC has been allowed the comparatively narrow objective of “reducing poverty through growth”. Within this context, impact is measure by the “income metric” directly linked to poverty. This one overarching impact objective for all compacts “allows project appraisal across sectors based on this same objective.”<sup>17</sup>

MCC’s Chief Economist describes a “Framework for Results” that is used to ensure Aid Effectiveness, an approach that begins at compact conception and continues through completion:

- Pre-Investment Analyses
  - Constraints Analysis
  - Benefit-Cost Analysis
- Monitoring and Assessment in Implementation
  - Baseline surveys
  - Implementation performance against the Benefit-Cost Analysis expectations
- Rigorous impact evaluations, as appropriate.

MCC plans to invest in the careful collection and analysis of data throughout the life of the compact. Each compact has a monitoring and evaluation plan with quarterly progress reports against that plan.<sup>18</sup> Currently one to two-page updates of implementation status are available on MCC’s website, including key milestones as well as some output data. MCC’s stated expectation for implementation monitoring and reporting goes beyond this: “reports of implementation provide an initial indication of whether the project met its targets in implementation, not just in terms of outputs ...but also in terms of outcomes...This information will help inform any discussion of whether a project should be funded by MCC in a second program (either expanded or sustained).” Current implementation reporting as presented on the website would be insufficient for those kinds of decisions. Presumably there are more detailed reports that are not evident.

MCC has a greater commitment to impact evaluation than any of the other USG agencies. As of February 2009, 18 impact evaluations were under contract (12 funded centrally by MCC and 6 funded by local implementing agencies out of compact funds). These evaluations cover 47% of program activities and 58% of program obligations.<sup>19</sup> At that time, another 11 impact evaluations were being considered by MCC and its local counterparts that , if all were contracted, would raise the total under rigorous and independent evaluation to almost three-fourths (74.\$%) of program obligations.

MCC (or their local counterparts) contracts with teams of professional impact evaluation specialists to “provide independent and rigorous measurement of MCC-funded programs”. MCC seeks to use the most rigorous methods possible to evaluate impact depending on feasibility, cost and potential learning. When the evaluations are completed, MCC has committed to making them publicly available. MCC’s contracts with these external evaluators empower them to publish their results independently.

---

<sup>17</sup> This section is drawn from a presentation by Franck Wiebe, “MCC and Impact Evaluation: Focusing on Results”, at DOS Program Evaluation Discussion Series, January 15, 2009

<sup>18</sup> MCC, “Guidelines for the Development of Monitoring and Evaluation Plans”, November 2006 and Franck W. Wiebe, “Aid Effectiveness: Putting Results at the Forefront; MCC’s New Institutional Approach” Working paper, October 27, 2008

<sup>19</sup> These statistics exclude the Burkina Faso and Namibia compacts, which have been signed but have not yet entered into force and do not have impact evaluations contracted yet. Source: communication with Franck Wiebe.

## **F. THE DEPARTMENT OF DEFENSE:**

With an increase in “nation-building” efforts as part of the Global War on Terror, and ongoing operations in Iraq and Afghanistan, the Department of Defense has emerged as a new player in international development. The Pentagon is recognizing that conventional “kinetic” military operations have limited utility in countering the threats posed to U.S. forces battling insurgencies in failed and conflict-ridden states. Therefore, they are searching for “nonkinetic,” and “hearts-and-minds” approaches, dispersing international aid and investing in infrastructure, schools, and other community projects in conflict zones. Indeed, the Department of Defense Directive 3000 elevated nation building as an objective equal to war fighting and as we mention above, roughly 18% of U.S. foreign assistance is now dispersed through the Department of Defense. Granted that over three quarters of that assistance goes to Iraq and Afghanistan, it still reflects a shift in DOD priorities. U.S. military involvement in development activities is not without controversy.

While DOD has extensive metrics, lessons-learned staff and institutional capacity, and monitoring and evaluation functions for traditional war-fighting efforts, there is no publicly available evidence that the military has a rigorous methodology for assessing the strategic effectiveness of their development activities as yet.<sup>20</sup> There are multiple efforts underway to address this including an M&E stream to Multinational Experiment 6, Austere Challenge 09, a draft handbook for projects funded with Overseas Humanitarian, Disaster and Civic Aid and likely many others. There are also likely to be significant challenges for coordination between evaluation of DOD foreign assistance activities and those of civilian agencies, including:

- The overall assistance goals are quite different -- the military is security oriented and tactical; development is often based on poverty reduction, humanitarian need and focused on longer-term institution building;
- Using the same terms to mean very different things e.g. effect and performance; and
- The staggering differences in scale of resources will exacerbate the difficulties.

Further discussion of the DOD role is beyond the scope of this report. But coordination between civilian and military foreign assistance will be a key challenge of the new administration.

---

<sup>20</sup> Reuben Brigety, Testimony for the Senate Foreign Relations Committee, Aid for the Future, July 31, 2008.

# SECTION 2: USG M&E PRACTICES

This section presents overall patterns in the way USG agencies are doing M&E, primarily USAID and Department of State. For brevity, it paints with a broad brush to highlight important issues. But this overshadows both the professionalism and commitment of staff and the sectors that do a very sound job. Several communities in USAID have demonstrated strong commitment to evaluation and monitoring. Both the health community and the microfinance community have invested in better evaluative approaches and are known for their strong learning. The DG community has invested heavily in recent years in improving monitoring and evaluation practices as well. MCC has excellent plans but no track record. Our data for this discussion draws on the evaluators who responded to our survey, our own experience, and the information provided by expert key informants.

## I. MONITORING

### A. ISSUES

The role of monitoring in the USG foreign assistance community has shifted dramatically in the last 15 years, particularly within USAID. In the 1970s and early 1980s, the principle was that good monitoring served the needs of project management. Now, monitoring is shifting to emphasize external accountability: reporting results to headquarters, to OMB and to the Congress.

In the 1970s, the USAID LogFrame served as the foundation for the Agency's monitoring system. It was used as a critical part of project planning at the outset of a project. It contained a set of project objectives, indicators and measures for tracking progress. Associated monitoring requirements included a baseline for each measure, targets for expected results and reporting on actual results at quarterly time points for the life of the project. Language changed as the LogFrame *input-output-purpose-goal* hierarchy was replaced by the Results Framework system at USAID. In the broader development community, the hierarchy that took over was *input-output-outcome-impact*.

Projects can and should be monitored to determine the extent to which the expected outputs, outcomes or impacts were achieved. When done well, monitoring can be invaluable to project implementers and managers to make mid-course corrections to maximize project impact. This type of monitoring is still needed and indeed encouraged, for example in the revised ADS Guidance at USAID and the performance monitoring systems at MCC. But overall, the role of monitoring in U.S. foreign assistance programs has undergone a shift. The emphasis is no longer just on good activity management but on accountability, reporting to agency headquarters, OMB and the

#### Terminology

It may be helpful to explain some of the terminology that pervades most discussions of foreign assistance programs and monitoring and evaluation. Outputs are those 'deliverables' over which the project implementers have control, such as the number of training programs and the number of teachers trained. Outcomes are generally those *expected changes in attitudes, knowledge/skills, and behaviors that are to be the immediate consequences of the outputs*. Put simply, a foreign assistance project runs training courses for civil servants on public infrastructure projects. Whether on return to their jobs the trained personnel (output) actually behave differently using the new skills and attitudes taught (outcome) is not within the direct control of the project implementers, it can be influenced by implementers and policy decisions by the Embassy or USAID. Impacts, sometimes called 'results', are the consequences of combinations of outcomes e.g. improved public health infrastructure results in healthier people.

Congress. The F requirements, intended to streamline monitoring and reporting, are extraordinarily rigid and labor-intensive. The authors have frequently heard USAID and State Department staff say that they find it difficult to continue regular monitoring practice when overwhelmed by F requirements for program management

As the size of the USAID staff has shrunk, Agency personnel have been forced to increasingly take on the role of contract/grant managers. What had previously been programmatic monitoring has evolved to accountability monitoring of contractor performance. The authors have seen too many cases where harried USAID field staff does not have the time to review or use their monitoring data. For example, during a training of USAID officers on monitoring plans, several of the more experienced USAID officers in attendance were ‘sitting out’ the training exercises. The author was told, “we don’t do these ... ourselves; we contract out the work”. If USAID staff is not actively involved in the process of thinking about what will be monitored and reported, experience has shown that they will not be actively involved in using that kind of systematic information to make decisions.

Within the Department of State, the only department-wide monitoring tool is the externally focused Mission Strategic Plan (MSP). MSPs were a series of goal papers with self selected indicators whose quality varied significantly – most were not tied very clearly to foreign assistance programs. MSPs were seldom consulted other than during the annual drafting and review exercise. In most cases, it was a “box to check” to meet OMB and Congressional requirements and rarely involved extensive discussion and debate within a country team.<sup>21</sup> They are an exercise in external reporting and not in programmatic management.

The initial decision by the F Bureau to include only input-output monitoring in their FACTS system was also a clear indication of the shift towards external reporting and away from programmatic monitoring to increase project impact. In early meetings to develop this set of indicators, the question that was repeatedly asked was, “what are we buying with this money?” F leadership emphasized the need to report to OMB and the Congress on what USAID and State were *doing* with their budgets rather than what they were *achieving* with their budgets.<sup>22</sup> This set of signals to USAID and State Department program mangers was troublesome in two ways:

- With the scarcity of staff time to manage programs, the amount of effort to respond to the new F reporting system was effort taken away from existing responsibilities – too often programmatic monitoring, including such things as gathering baseline data;
- By focusing on outputs, F sent a message, intended or not, that outcomes were not as important as external reporting on whether foreign assistance resources were being expended. This is much closer to the waste fraud and abuse perspective of the Office of the Inspector General (IG) than the program focus of using systematic information to maximize the results of foreign assistance.

Both USAID and State Department send out annual messages requesting “Success Stories”. Over 65% of the external evaluators who answered our questionnaire agreed with the statement “DOS/USAID care more about success stories than careful evaluations of their program”. This was the most answered question in the survey.

What information you collect sends signals about the priorities for USG foreign assistance. One external evaluator said “USAID doesn’t want to know the facts/truth...Branding and success stories are the primary interests.” Combining the message that success stories are more important than systematic

---

<sup>21</sup> Government Performance and Results Act (GPRA) and questions in OMB’s Performance Assessment Rating Tool (PART) designed to be sure agencies were meeting requirements such as GPRA.

<sup>22</sup>Hyman, Gerald F. “Assessing Secretary of State Rice’s Reform of US Foreign Assistance.” Carnegie Papers: Democracy and Rule of Law Program No. 90, CSIS. Feb. 2008.

information with the emphasis on output monitoring undercut the sense that these organizations are committed to results.

By contrast, MCC monitoring and evaluation guidance emphasizes impact and sends a clear message that results are important. MCC defines impact as “reduced poverty through growth” using income as the indicator for poverty. This clear and measurable objective is the basis for a system to compare activities in the context of their progress in achieving that one impact. However, it is early in the life of the MCC compacts to have much understanding of the realities of their monitoring experience.

## **B. GUIDANCE AND GETTING THE JOB DONE**

One explanation for the challenges with monitoring has to do with guidance. There is no standardized monitoring guidance for foreign assistance programs – the requirements at MCC are very different from the requirements at State and USAID. When F was created, it was an opportunity to standardize across at least USAID and the Department of State. When F focused on input-out external reporting, that resulted in continuing the plethora of existing agency practices for management monitoring. Moreover, F’s very strong early messages about streamlining led many at USAID to believe that their ADS system with Results Frameworks and Performance Monitoring Plans (PMPs) had been superseded by the new F system. This confusion is reported by a survey respondent: “Right now they can’t decide whether to use Results Frameworks or FACTS. There is obviously no guidance since every mission I’ve done PMPs for (36 in 1-1/2 years) approaches it differently.” Even with the mandatory requirement in the guidance for baseline measurement of each indicator at all levels, practice is falling seriously short as indicated above.

Monitoring requirements have increased, staff size has decreased and the monitoring function has been delegated to contractors and grantees. This would be workable if agency staff had the time and skill to both oversee the monitoring function and the skill to understand the strengths and weaknesses of the information they are using. This outsourcing has resulted in the loss of dialogue, debate and learning by agency staffs. These issues are discussed further in the next section.

## **C. LINK BETWEEN MONITORING AND EVALUATION:**

Monitoring can inform the project manager when a project is veering off course which may occasion the call for a more in depth management or process evaluation to provide an independent assessment of what has gone wrong, why, and what might be done to fix it. In addition to being used for such management decisions, and external reporting, monitoring data is the necessary foundation for strong evaluation. Evaluation of programmatic results without monitoring data is available as the starting point for an evaluation is much more difficult and the limited in scope. This data is necessary to look at the basic evaluation questions of *why did things happen?* or *was the value of the project as expected when the investment was made?* and *whether the achievements are likely to be sustained?*

# **II. EVALUATION**

## **A. SHOULD EVERY ACTIVITY BE EVALUATED?**

At the Department of State, the de facto decision has been to not evaluate. State Department does have the well respected evaluation work that has been done by the office of Education and Cultural Affairs, but quite limited evaluation work on other programs. Some effort is being made to promote evaluation by RM/SPP and F but these efforts are very new and of limited scope. Management of foreign assistance programs remains outside the purview of the typical foreign service and civil service officers at the Department of State and functional bureaus have implemented ad-hoc approaches to program evaluation. There is a brand new evaluation policy that institutes a department-wide requirement for evaluation.

Through the PART process, OMB has put pressure on State to do evaluations and a number of these are being undertaken.

MCC has designed a very robust impact evaluation system and is working toward conducting evaluations for a large number of activities within their compacts. Decisions are made in consultation with their implementing partners (check language in first section) for ownership of the findings. When impact evaluation is not a priority for host country partners, some impact evaluations are centrally funded in order to have the knowledge about the impact that MCC programs have on income in that country. But few activities have been completed; therefore evaluations haven't been completed so there is little in the way of track record.

USAID's rule for the last decade was that the decision to evaluate should be left to managers in the operational units, when they identified a need. Previously, the mandatory requirement to do mid-term and final evaluations led to the production of many, largely un-used reports. Today, in practice, it is the USAID technical officers<sup>23</sup> who usually call for an evaluation. The newly approved ADS has included a mandatory evaluation requirement but not for each activity.<sup>24</sup>

Operational managers are reluctant to use scarce budget and time to commission an evaluation for the purpose of building the Agency knowledge base about what works, what doesn't, and why. Therefore the sectoral centers in Washington<sup>25</sup> have been providing leadership in evaluation. For example, Global Health, the Democracy and Governance Office and the Micro-enterprise Office have all supported evaluations as well as efforts to strengthen evaluation methods in their sectors. The Office of Private and Voluntary Cooperation has continued to require evaluations in their cooperative agreements. The Europe and Eurasia Bureau has a very strong system called Monitoring Country Progress to assess progress in outcomes for economic growth and democratization.

## **B. HOW MUCH SHOULD BE SPENT ON MONITORING & EVALUATION?**

Many reports and studies about monitoring evaluation budgeting advance the idea that somewhere between three and five percent of an agency's program budget should be used for monitoring and evaluation.<sup>26</sup> Using the five percent figure, the evaluation share of the State/USAID joint \$22 billion budget request for 2009 would be \$1.1 billion<sup>27</sup>. If that amount were split between monitoring and evaluation on a 50/50 basis, State/USAID would manage an evaluation budget of approximately \$525 million dollars. How does this compare to what was spent on evaluation in 2008? Because most M&E budgets are embedded in the contractual or cooperative agreements between USAID and implementing partners and dispersed in sectoral centers in Washington, it is impossible to ascertain how much money is spent on monitoring and evaluation. But, it is highly likely that neither State nor USAID have come close

---

<sup>23</sup> USAID had formally designated officers in charge of each activity called "Cognizant Technical Officers." (CTOs)

<sup>24</sup> ADS Section 203.3.6.1 "**MANDATORY:** AO Teams must conduct at least one evaluation aimed at understanding progress or lack thereof and the types of actions that need to be taken to improve performance during the life of each AO as defined by their respective USAID Mission/Office."

<sup>25</sup> The "Pillar Bureaus"

<sup>26</sup> See for example, "The New York State Public Service Commission's (Commission) Order establishing the EEPS took several important steps to enhance evaluation of programs funded with System Benefits Charge (SBC) funds, including increasing from 2 percent to 5 percent the portion of the program budgets dedicated to evaluation requiring development and implementation of evaluation protocols, and the establishing a statewide Evaluation Advisory Group" .www3.dps.state ny.us/PSCweb/PIOweb.nsf.

<sup>27</sup> The FY 2009 Foreign Operations request for the Department of State and USAID totals \$22.7 billion, a \$598 million increase over the FY 2008 enacted level, or an increase of almost 3%, including the Global HIV/AIDS Initiative and P.L. 480 Title II food aid. **Congressional Budget Justification: Foreign Operations. Fiscal Year 2009**, US Department of State-USAID.

to that number. It is possible that MCC has planned something approaching several per cent of budget, but it will be very interesting to see if they are more successful in sustaining that level of investment than history suggests.

One explanation of why there isn't better planning can be found in the incentives that are operating at the beginning of a program for both USAID and the selected implementing partner. Cost is an important factor in winning bids, and potential partners do their best to keep the budget as low as possible. On the USAID side, the practice is increasingly to require the bidder to include in its proposal the preparation of the Performance Monitoring Plan (mandatory) and the evaluation (if needed). Neither side has a sufficient incentive to invest a lot of up front money in setting up a baseline and structuring a rigorous evaluation design, in the face of the lowest cost being a primary threshold for winning bids.

Poor budget planning leads to ad hoc and under-funded levels of efforts for outside evaluators, who must "make do" in answering questions, with a weak evaluation design usually without benefit of a baseline against which to measure change, or without built-in control groups against which to compare the effects of the assistance project, or to address questions about attributable impact and sustainability.

The absence of leadership and guidance that provide funding, incentives, and standards for operational units to invest in careful evaluation planning produces an ad hoc approach to thinking about field evaluations. The result is too often a poorly conceived scope of work, a "do the best you can in three weeks" evaluation design, and a set of findings and conclusions that rely as heavily on the experience and judgment of the team members as it does on the careful assembly of facts.

This causes the evaluation team to narrow the scope of their effort limiting the import of what they learn. Development and other foreign assistance efforts are complex and highly reactive to their difficult and changing settings. The by-product of slimmed down evaluation efforts are missed opportunities to get the best information on activities within their context.

### **C. METHODOLOGY- HOW RIGOROUS?**

The rigor of an evaluation derives from the evaluation methods which are designed to minimize bias and therefore increase confidence in the findings. Social science research methods call for testing whether a treatment group's improvement is attributable to the treatment by having a 'control' or comparison group. In one author's experience with 31 evaluations, control or comparison groups were established, and had to be 'found' by the evaluation team as part of the evaluation process. This is a very imperfect solution.

#### **Evaluators' Perspective: Scopes of Work**

The Scope of Work (SOW) prepared by the client is an all important document for external evaluators and therefore for the quality of the evaluation. There are several critical elements: first, evaluation questions, second, team qualifications, and third, time and budget available for the evaluation. SOWs may also suggest methodologies to be used, and, rarely, the confidence level the client expects in the evaluation's data, findings and conclusions.

In the authors' experience, too many evaluation scopes of work have in the past been poorly thought out, most notably in the tendency to set out long lists of questions to be answered without regard to the time needed to develop and analyze empirical evidence on which to base an answer. Even when the SOWs have more focused questions, there often remains a gap between questions posed and the resources needed to provide high quality answers. Moreover, it is rare that the resources provided for an evaluation are sufficient to develop and apply more rigorous research methods that would produce valid empirical evidence regarding outcomes, and attributable impact.

One colleague summed it up: "USAID does a good job of posing questions but the corresponding weakness is that it does not provide a broad enough scope of work or level of effort to insure that findings have any empirical validity. The result is that project officers end up using M&E findings to justify their own pre-established opinions (which may in fact have an experiential basis but which never are the result of systematic applied evaluation research)."

Professional evaluators tend to want more rigorous methods to increase the confidence in the findings, but many recognize the real world limitations.<sup>28</sup> Rigor is costly in terms of time and money.

The Center for Global Development has led a strong effort calling for more rigorous foreign assistance evaluation leading to the creation of the International Institute for Impact Evaluation (3IE) in 2007. Taken to the extreme, this call for rigor assumes a set of independent, well trained social scientists setting up randomized control trials or quasi-experimental designs to test the ultimate impact of a foreign assistance investment. This has raised the issue of participation and the ownership and utilization of results. There is a substantial movement among a set of evaluation professionals who strongly advocate for a different approach called participatory evaluations, largely on the grounds that participation enhances utilization<sup>29</sup>. Many argue that scientific rigor and participatory evaluation practices are not contradictory, and indeed there are many other evaluation methods in use throughout the world. This debate continues but there has been recognition that the “appropriate” method should be chosen for each evaluation.

When implementing new and untested approaches, foreign assistance managers would be wise to identify some type of comparison group, collect careful baseline data on both groups and keep a close eye during implementation. Occasionally there will be instances to employ the fully rigorous methods that professional evaluators are capable of providing. But the reality for most foreign assistance is that the great should not be the enemy of the good. Matching the issues to be addressed to the length and scope of the evaluation, more clarity about before and after comparisons, and actively looking for competing explanations to explain the change observed will go a long way. Even approximations of rigorous designs and methodologies will require more resources, better advance planning, and the development of a more sophisticated ‘consumer’ of evaluation reports than now exists in the U.S. foreign assistance community.

#### **D. EXPERTISE AND TRAINING**

Program planning, monitoring and evaluation are not included in the normal training for State Department Foreign Service officers. Department of State does have both a course at the Foreign Service Institute and an online course that deals with monitoring, but neither is required and comparatively few have taken either one of them. This leads to a very different culture in the State and USAID foreign services. As State and the Embassy’s role in foreign assistance decision making increases, this becomes increasing problematic.

Throughout the 1990s, most of USAID’s training investment in M&E focused on the monitoring side, to the detriment of evaluation. By 2000, USAID began a correction, contracting with Management Systems International to conduct a series of intensive Evaluation Certificate courses for USAID evaluation officers and others. This course continues to be presented. Although there is an “M&E” module in most every other USAID training course, what is taught is too thin to prepare USAID staff adequately to manage partners to carry out these functions. In the best scenario, at least some staff would be prepared to carry out monitoring and evaluation themselves so that they had a deep enough understanding to backstop M&E functions.

---

<sup>28</sup> Bamberger, Jim Rugh and Linda Mabry *Real World Evaluation*, Sage 2006 [www.realworldevaluation.org](http://www.realworldevaluation.org) Real world evaluators, which includes most of our respondents, have been struggling with the issue of how to reduce bias and improve validity of their findings while coping with limited resources, short time frames, and poorly conceived scopes of work. This has spawned a literature dating back to the 1970s beginning with Rapid Appraisal methods from the work of Robert Chambers at the Institute for Development Studies, Sussex, England, and the USAID Impact Evaluation Program of the same era.

<sup>29</sup> Michael Quinn Patton, *Utilization Focused Evaluation*, Fourth Edition, Sage 2008 and Others.

USAID has also invested in outside expertise to conduct assessments of evaluation practices, such as the National Academy of Science study for the Democracy and Governance Office, or the Teller study for the Global Health Bureau. These critical studies have stimulated USAID thinking about the need for better evaluation work.

## **E. USING EVALUATIONS IN MANAGING FOREIGN ASSISTANCE**

Most evaluation professionals would agree that evaluations that are not used are a waste of time and money.<sup>30</sup> Are evaluations actually used – even at USAID?

In the author’s experience, the immediate client, usually the USAID officer who commissioned the evaluation, did have a purpose in mind, usually connected with the development of the follow-on activity. But the utilization usually stops there.

At present, there are no empirical studies that examine the use of evaluation findings in USAID or, to our knowledge, other U.S. foreign assistance agencies. From our survey respondents and our own experience, we identify some of the factors which may contribute to evaluation utilization, or lack thereof.

- The perception reported by external evaluators that “*USAID cares more about success stories than careful evaluation*” would explain why there is very little motivation to invest in serious program or project evaluation.
- The linking of U.S. foreign assistance to short term foreign policy objectives works to make outcome or impact evaluation irrelevant. As a colleague put it: “...with USAID increasingly responding to short term foreign policy directives—with impact on the recipient country coming a distant second in priority -- evaluation seems somewhat besides the point.”
- Bureaucratic incentives do not support rigorous evaluation or use of findings. One thoughtful colleague said: “*If the USG feels that evaluation research is necessary to assure a solid empirical foundation for decision-making, then there should be a concurrent commitment to assuring that the evaluation research which is conducted addresses ‘actionable’ issues, that it is methodologically sound, and that it is actually used by the broadest possible range of stakeholders. Currently, I consider the majority of such efforts “ceremonial”—because the primary objective is “to get an evaluation expert”, “to get some numbers for the next annual report”, and/or to get something for a presentation at an upcoming conference or management meeting.*”
- Make evaluation report findings more accessible. Long, detailed reports will not be read by decision makers. “*Write up evaluation findings, conclusions and recommendations in one and two page formats for decision-makers....Crisp, professional briefings are also necessary.*”

Done well, evaluations can serve multiple stakeholders at various levels, from host country citizens and institutions, to U.S. decision makers in the field and in Washington. But different kinds of evaluations need to be designed to address the management issues of activity managers and host country managers than the types of studies that address the policy and strategic choices about assistance mechanisms such as projects or budget support; or whether sufficient capacity has been built to deliver HIV/AIDS drugs; or whether support to media or elections are making a greater contribution in a country.

Perhaps in the long run, the most important consumer of evaluation based knowledge should be policy makers and the US Congress. Because Congress appropriates the funds, and has a history of specifying through “earmarks” or strongly suggested allocations exactly how the funds should be used, in many ways, the Congress has operated as the policy and budget office for the US foreign assistance program. If

---

<sup>30</sup> The strongest advocate for this point is Michael Quinn Patton. See his “Utilization Focused Evaluation” 4<sup>th</sup> Edition, Sage Publications. 2008.

Congress is going to play this role, the members and staff should be thirsty consumers of evaluation based knowledge about what works, what doesn't and why. Unfortunately, the main reporting document required by the Government Performance Results Act does not provide enough in the way of synthesis and analysis of foreign assistance outcomes and impact. While Inspector General reports are regularly sent to the Congress, evaluations are not. Perhaps this is at least partly because the types of evaluation being done are not of the type to address the broader policy issues.

No matter how rigorous or relevant, a completed report is just the first step in a process of dissemination, discussion, and debate. For this cycle of learning to be complete to achieve the purpose of contributing to 'smart' policies and programs, there must be leadership and a system wide commitment to creating a 'learning culture' which values knowledge, and accepts the reality that not all programs will be successful or should be sustained.

# SECTION 3: MODERNIZING FOREIGN ASSISTANCE MONITORING AND EVALUATION PRACTICE

Ideally, foreign assistance reform will have already changed the context in which monitoring and evaluation are conducted. To strengthen monitoring and evaluation practices:

**Design streamlined monitoring system with a better balance between learning and accountability.** Most monitoring data originates in the field – even for DOS centrally managed programs. First, conduct a review of all required monitoring and performance reporting. Then, sufficient expertise and resources should be allocated to meet the requirements. If this is not possible, requirements need to be changed – renegotiated with external stakeholders if necessary.

**Develop a minimal degree of common standards and requirements across agencies that provide foreign assistance,** recognizing that some agencies will need a long transition period to meet those standards. Work with agencies to meet the standards in a way appropriate to each program.

**Introduce results management into training of all staff that will have a role in managing foreign assistance.** This includes State Foreign Service Officers. All agency training should be reviewed to ascertain what is being taught now.

Develop an interagency evaluation agenda that not only collates evaluations being conducted but includes what evaluations should be conducted. In addition to senior agency and department managers, external stakeholders including the Congress and NGOs should be consulted to identify priority evaluation topics.

**Conduct multi-country comparative evaluations that address foreign policy priorities and strategic choices.** Resources and expertise must be allocated. Stakeholders should be included throughout the process, including host country policy and decision makers.

## A. CREATE A CENTER FOR MONITORING AND EVALUATION

In section 1, the authors recommended the creating of an independent Center for Monitoring and Evaluation. The creation of such a Center would signal the importance of results management and M&E. In addition, a Center would provide leadership to carry out the recommended changes.

Evaluation offices and programs would still be needed within agencies to focus on their particular needs, but an independent center would provide policy guidance, enforce standards, implement training programs, provide technical assistance, and set overall budget requirements. It would also conduct policy, sector and program useful meta-evaluations, and operate a timely documentary archive of all agency analytic work for use by all foreign assistance implementing agencies, and other stakeholders. It would also synthesize, analyze and report on an annual basis to a variety of stakeholders, including the Executive and the Congress, on the evidentiary basis for foreign assistance results and impact, including strengths, weaknesses, and lessons learned.

The CME would be responsible for the following functions:

- Develop a comprehensive annual evaluation agenda including an overall M and E budget request for all FA agencies, including the preparation of a separate budget for CME;
- Prepare guidance and training on monitoring and evaluation methods and approaches based on best practices as well as conducting random assessments of FA agency practices;
- Require career incentives and personnel policies to create and support a learning culture throughout each agency;
- Mentor and support agencies in their collection and use of M&E information, program reviews and managers evaluations. In agencies where M&E units are very weak or do not exist, CME staff could support monitoring and evaluation directly;
- Develop programs to facilitate dissemination of evaluation findings and conclusions, and to promote organizational learning and knowledge management by foreign assistance agencies. This would include a user friendly archival system;
- Conduct multi-country evaluations on a type of project or sector of the overall progress in achieving outcomes, looking for best practices and lessons learned to apply to future programming. These meta-evaluations should be focused on foreign policy priorities and must serve the expressed needs of policy and program decision-makers; and
- Report annually to the Executive and the US Congress on foreign assistance agencies' principle findings, conclusions and recommendations as well as monitoring and evaluation activities.

While the specific structure, personnel levels, and allocation of functional responsibilities would be developed once the primary policy decision to establish a CME has been made, two structural principles need to be kept in mind.

- The director of CME **must have** the political status to insure high-level interagency access and to represent the M&E agenda to other senior policy makers and administrators, as well as the technical experience to engage with confidence with M&E professionals in setting M&E policies, procedures and standards; and
- The best international models have an independent authority reporting to, but not under the direct control of the department secretary responsible for foreign assistance. CME priorities would be established in consultation with senior policy makers.

Without some authority setting and enforcing reasonably uniform standards for monitoring and evaluation practices, recent history has shown that implementing agencies will not produce the kind of systematic and useful monitoring and evaluation products that we and our external evaluator colleagues believe is possible and necessary. Given the demonstrated tendency to spend as little on monitoring and evaluation as possible, these vital functions will continue to be treated as ad hoc, poorly planned and executed events, if done at all.

It is important to find a balance between a central monitoring and evaluation office such as the proposed CME and the need for monitoring and evaluation units in all foreign assistance agencies. In a diverse and decentralized system, different agencies will have different needs, work in different task environments, and have differing responsibilities and objectives. The CME would be responsible for bringing coherence to learning across the programmatically diverse but interconnected elements of different FA agencies, as well as for insuring that ALL foreign assistance programs are held to the highest possible standard for monitoring and evaluation. While different FA agencies will have different needs, there should be a common monitoring and evaluation culture and set of institutional practices across all FA agencies.

## **B. FINAL WORDS OF ADVICE**

*Build strong monitoring and evaluation back into foreign assistance demonstrating commitment to results and transparency of management* – not only for the external stakeholders but for the beleaguered staff in at least several of the agencies implementing foreign assistance.

*Stop asking for success stories which sends the message that only success is acceptable.* In foreign assistance, we need to take some risks and learn from them. Foreign assistance policy should include a statement about the acceptable level of failure. A critical examination of clearly established successes is as important a learning experience as failure, but it is time to move beyond the overemphasis on success stories.

*Shift the balance of M&E from partners back into agencies:*

- Missions and other programmatic offices at USAID,
- State Department Bureaus with foreign assistance programs and to an increasing degree Embassy staffs,
- Review the experience with MCC to ascertain the balance between internal and external monitoring and evaluation roles.

Implementing partners of USAID and State will continue to be responsible for collecting most systematic information and should be included in the analysis and review process – but now too much of this is just delegated to partners.

*Balance learning and accountability* by building strong enough foreign assistance monitoring and evaluation systems that the auditors can move back to the space of inspections and audits that focus on waste fraud and abuse with the confidence that programmatic results will be fully addressed.

Current monitoring and evaluation of foreign assistance programs do not serve the learning, decision making and management needs of the primary USG agencies implementing foreign assistance discussed in this paper. We know how to do the job. What is missing is leadership commitment, policy, organization, resources, and positive incentives for learning from what works, what doesn't and why.