TIME TO LEARN: AN EVALUATION STRATEGY FOR REVITALIZED FOREIGN ASSISTANCE

PRIVATE SECTOR DEVELOMENT IMPACT ASSESSMENT INITIATIVE

JANUARY 2010

This publication was produced for review by the United States Agency for International Development. It was prepared by DAI.

TIME TO LEARN: AN EVALUATION STRATEGY FOR REVITALIZED FOREIGN ASSISTANCE

PRIVATE SECTOR DEVELOMENT IMPACT ASSESSMENT INITIATIVE

Program Title:	**AMAP Business Development Services Knowledge & Practice Task Order**
Sponsoring USAID Office:	**USAID's Microenterprise Office**
Contract Number:	**GEG-I-800-02-00014-00, Task Order No. 01**
Contractor:	**DAI**
Date of Publication:	**January 2010**
Author:	**Lucy Creevey, Jeanne Downing, Elizabeth Dunn, Zan Northrip, Don Snodgrass, and Amy Cogan Wares**

CONTENTS

TIME TO LEARN

Nearly a dozen high-profile reports and commissions have offered recommendations to the Obama administration on how U.S. foreign assistance should be restructured. Ranging from the organizational to the thematic, these recommendations provide a wealth of sometimes conflicting advice. There is a common thread, however, running through nearly all the recommendations: the need for improved evidence on foreign assistance effectiveness and accountability, with a particular call for "rigorous evaluation."

This paper suggests fulfilling that recommendation by institutionalizing evaluation as a driving force in a system of continuous learning and program improvement at USAID. Simply increasing the number of evaluations conducted will not improve the quality and cost-effectiveness of development assistance. Most important is ensuring that evaluations result in learning that is integrated into decision-making as part of a learning cycle.

Unfortunately, the current discussion over evaluation of foreign assistance is dominated by a series of distracting critiques and methodological debates that obscure evaluation's central role in learning and project improvement. Three of the most common assertions include:

USAID doesn't learn from its mistakes because it no longer evaluates its projects. USAID's evaluation and policy planning functions should be restored immediately. Despite wistful remembrances to the contrary, there never was a golden age of evaluation at USAID. While it is true that USAID once produced more and better evaluations, the link between evaluation and program planning and design was weak. We support recent efforts to reestablish these functions, but they need to be elevated, integrated, and significantly restructured if they are to be useful guides to U.S. foreign assistance strategy. A system for continuous learning with senior political support is required to make evaluation worthwhile.

Evaluation means "impact assessment." Efforts to look at project outputs and outcomes are not worth the money, because they cannot establish whether those outcomes are attributable to the project; yet impact assessment is expensive and time consuming. Evaluation should be seen as a continuum, with the monitoring of outputs on one end, and the attribution of project impacts at the other. There is value in every part of the continuum, and sometimes the most value in the middle.

Randomized controlled trials (RCTs) are the gold standard and the only acceptable method for impact assessment. RCTs are wonderful where they can be done appropriately, but they have limited applicability in the unsterile and unstable "laboratory" where most development projects take place. One high-quality RCT of a microlender in Kenya found significant impact after one year, yet the institution collapsed shortly thereafter. External relevance is as important as internal rigor. Despite these challenges, insightful and credible results can be obtained by evaluating development projects using a variety of methods providing different aspects of rigor.

USING EVALUATION TO LEARN

To succeed, U.S. foreign assistance requires a program of evaluation and learning that responds to leadership's most important questions. Current examples might include:

- How can the private sector best be harnessed to strengthen food security?
- How can local incentives reduce climate change by slowing environmental degradation?
- What interventions are most effective in preventing HIV/AIDS?
- What is the contribution of job growth to political stability?
- In all areas, how can U.S. foreign assistance generate credible results and value for taxpayer dollars?

CREATING RESEARCH PLANS RESPONSIVE TO A PRIORITIZED LEARNING AGENDA

USAID should lead an annual process that generates and prioritizes such questions by their importance to foreign assistance objectives. For each question ultimately selected for in-depth analysis, a Research Leader needs to be appointed and a comprehensive Research Plan established that includes:

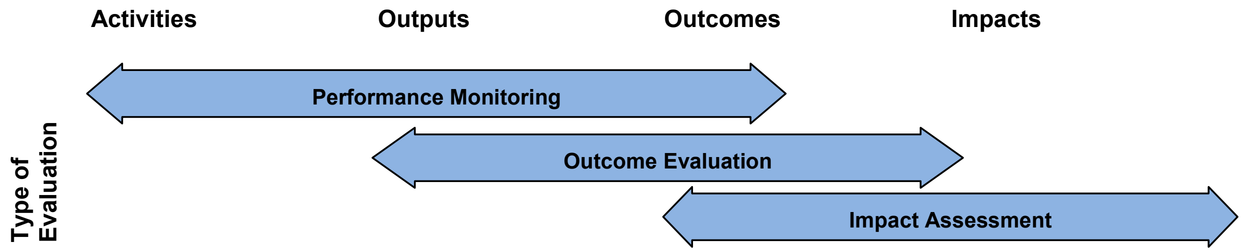
- A selection of projects focused on the question at hand;
- A causal model that articulates the cause-and-effect relationships expected;
- Metrics based on the causal model that clarify what to measure and why;
- Plans to collect monitoring data from each studied project;
- For a smaller number of the studied projects, plans to conduct outcome and impact assessments using methods appropriate to the question being asked and to the causal models and implementation contexts of the specific projects.

The totality of this evidence - drawing on outputs, outcomes, and impacts from a range of relevant projects - should be collected, considered, and synthesized by the Research Leader. Conclusions should be shared and fully leveraged in future foreign-assistance programming, enforced through directives and incentives from the top.

MATCHING QUESTIONS TO EVALUATION TYPES

Evaluation questions should be linked to specific components in a project causal model. Once this linkage is made, the evaluation methodology required to answer the question being asked can be selected. Figure 1 provides a simplified model of the types of evaluation that are appropriate for answering questions at a project's output, outcome, and impact levels.

FIGURE 1. TYING THE CAUSAL MODEL TO EVALUATION QUESTIONS AND TYPES



Performance monitoring should be conducted by all projects to track the timing and execution of their planned activities, and the degree to which activities result in anticipated outputs. Outputs are results that are discrete and can be counted, e.g., the number of farmers reached by a program aimed at improving agricultural productivity and the number of new technologies introduced. While this information cannot attribute results to a USAID-funded project, it is important to project management, because it offers quick feedback and allows for mid-course adjustments. Performance monitoring also feeds the reports and success stories vital to communicating with USAID stakeholders.

Outcome evaluation focuses on the link between project outputs and outcomes. It seeks to determine whether expected behavioral changes at the outcome level have occurred – for example, adoption of new technologies by farmers and improvements in their productivity. Some outcome indicators are simple and impartial enough to be incorporated into a project's internal performance monitoring system. Others depend on independent quantitative and qualitative analysis. Outcome evaluation is important both because it verifies that the project is on the right trajectory toward the desired impacts and because it tests the accuracy of the underlying causal model. Also, outcomes are often developmentally important achievements in their own right, even if the expected impacts cannot be proven.

Impact assessment focuses on the program's higher-level objectives: the outcomes and impacts achieved. It differs from performance monitoring and outcome evaluation in that it attempts to establish that the changes that occurred can be attributed to the activities of the project. Policy makers tend to be most interested in impact questions, but - methodologically - impact questions are the most difficult and expensive to answer. Requiring a control group, impact assessments call for consistency in project implementation between baseline and endline surveys that can be difficult if not impractical and inadvisable to achieve, especially for projects that operate in dynamic and unstable environments. In other words, measuring impacts successfully and credibly requires mixed methods, creativity as well as rigor.

CHOOSING METHODS THAT ARE BOTH CREDIBLE AND ATTAINABLE

If the questions to be answered require an outcome evaluation or impact assessment, the choice of methodology will become a critical determinant of the credibility of the evaluation's findings. If the project is innovative, new or implemented in n unstable environment, as represented by many developing countries, the typical experimental approaches may be challenged. The Degrees of Evidence approach (developed by the Private Sector Development Impact Assessment Initiative from the insights of Madhabi Chatterji and others) provides guidance on constructing outcome evaluations and impact assessments to attain both credibility and relevance given the many complications that arise when evaluating development assistance programs. The Degrees of Evidence approach is based on several principles:

- There is more than one acceptable way to conduct outcome evaluations and impact assessments. No one method is perfect and all evaluations can be criticized, but carefully selecting and combining methods will increase the credibility of evaluation results.
- Rigor is not a binary concept, but a matter of degree and often multidimensional. One characteristic of rigor is the extent to which the weaknesses of the study are acknowledged and explained by the evaluators. The user of the evaluation needs to understand the methodological choices that were made and the trade-offs those choices required.
- Even evaluations that are imperfect can have value. Not all questions can be answered with the highest degree of credibility, but in most cases, information with caveats is better than no information at all.

Above all, the imperative that USAID learn from its investments dictates that the results of its projects be analyzed and the lessons learned applied to future programming. A decision to restrict USAID to RCTs and quasi-experimental methodologies will mean than many of its most promising projects go unevaluated, condemning the agency to ignorance in the name of rigor. A decision to continue the nearly exclusive focus of recent administrations on the evaluation of implementers would be even worse, foregoing the chance to measure and learn from more meaningful project results.

Instead, USAID should seize the opportunity to extract greater value from the full spectrum of evaluation types and methods. Doing so will require strategic decisions, funding, and skills and the support of senior leadership to rebuild USAID as a learning organization that both replicates success and learns from failure.

INSTITUTIONALIZING EVALUATION AND LEARNING AT USAID

This paper has outlined an ambitious scope of work to leverage evaluation as a strategic tool in the revitalization of U.S. foreign assistance. Four imperatives - that USAID's learning agenda prioritize leadership's most pressing questions, that those questions be answered by examining a range of projects from across the Agency, that research be carried out with adequate funding and credible methods, and that conclusions be reached impartially - suggest strongly that this critical function should be centralized. A central Evaluation and Planning Unit should be mandated to:

- Act as an agent for USAID leadership in identifying an evolving set of priority questions that can be answered through Research Plans that feature multiple types of evaluation across a range of projects.
- Select a small number of those projects for investment in outcome evaluation and impact assessment, based on those projects' ability to answer leadership's most pressing questions;
- Fund and oversee the implementation of those evaluations, using credible methods and independent analysts;
- Capture and synthesize the knowledge generated from those research activities;
- Disseminate this knowledge widely in manageable pieces, through training modules, accessible publications, online media, and other channels;
- Support and enforce the application of lessons learned in future USAID decision making, programming, and implementation choices.

Additionally, to ensure that all USAID projects meet the basic requirements for evaluation of any kind, and could therefore be examined in future Research Plans, the Evaluation and Planning Unit should have the responsibility to:

- Ensure that all USAID projects have causal models that articulate their development logic;
- Require and support the development of performance monitoring systems in each USAID project, tracking indicators that are both relevant to their causal models and sufficiently rigorous for meaningful reporting to USAID stakeholders in the Administration, Congress, and the American public;

A New Spring?

In the late 1960s, USAID initiated an evaluation process known as the Spring Reviews. Planned a year or more in advance, these reviews gathered a range of evidence on specific issues critical to USAID programming. In some cases the reviews made important contributions to development activities in USAID and beyond. One of the better known examples is the Spring Review of agricultural credit, which proved to be a milestone on the road to the microfinance revolution. Managing such a mechanism to focus Agency attention on critical evaluation questions should be one of the primary functions of USAID's Evaluation and Planning Unit.

What won't work is a simple return to the past. USAID needs more than annual targets for evaluations completed, and more than a central repository for the reports and evaluations it funds. In order to improve the quality and cost-effectiveness of development assistance, evaluation must be institutionalized in a central office with the mandate to lead USAID in becoming a learning organization in which new knowledge is captured strategically, disseminated quickly, and applied under a framework of continuous feedback and adaptation, as shown in Figure 2. If evaluations are to be more than shelf accessories, USAID will need to synthesize their key lessons in accessible media and trainings that clarify best practices and are immediately linked to new program design. With a well executed learning cycle, USAID can emerge from the Obama administration as a donor that uses evidence-based programming to advance both stakeholder support in the United States and much-needed development results in the field.

FIGURE 2: LEARNING CYCLE