EVALUATION STANDARDS

Planning and Performance Management Unit
Office of the Director of U.S. Foreign Assistance
Final Version: March 25, 2009

Introduction

These "Evaluation Standards" were jointly developed by the State Department's Director of Foreign Assistance and USAID for their staff responsible for initiating and managing evaluations of foreign assistance programs and projects. They are primarily designed to inform evaluation managers about the idealized norms that should guide the planning and conduct of evaluations.

To develop these standards, DFA reviewed the documents on professional ethnics and evaluation standards issues by professional associations, bilateral and multilateral donor agencies and inter-governmental organizations. It has particularly benefited from evaluation standards identified by the Evaluation Network of the OECD's Development Assistance Committee and the American Evaluation Association. DFA acknowledges its intellectual debt to them.

Purpose and Definitions

This paper briefly identifies a set of evaluation standards that should be followed in planning and conducting evaluations of foreign assistance policies, projects and programs. These standards reflect both the principles and practices of evaluation and as well as the distinctive requirements of foreign assistance programs. Although designed primarily for use by U.S. foreign assistance agencies, they are also relevant to partnering institutions, private contractors and host countries that conduct their own evaluations of development programs.

Evaluations are defined here as formal assessments based on systematic collection and analysis of data from on-going or completed projects, programs or policies. They are primarily undertaken to improve the performance of existing interventions, assess their effects and impacts, and/or inform decisions about future programming. They should not be confused with audits undertaken to determine the propriety, compliance, and adequacy of programs and operations

Evaluation standards refer to the idealized norms for planning and conducting evaluations. They are more abstract and succinct than "evaluation guidelines" which provide more comprehensive guidance. Indeed, evaluation standards are the foundations on which evaluation guidelines are developed.

Underlying Principles:

There is a broad consensus among the foreign assistance and evaluation communities about three general principles, which underlie evaluation standards. These are as follows:

1. Usefulness

The information, ideas and recommendations generated by evaluations should serve the needs of foreign assistance agencies/offices/operating units commissioning them. Evaluations should help in improving the effectiveness, relevance and efficacy of foreign assistance. More specifically, they should serve one or more of the following objectives:

> (a) Improving ongoing projects, programs and policies by examining their implementation and the factors that affect their performance and effects as a basis for actionable recommendations,
>
> (b) Improving future foreign assistance policies, programs and projects by codifying experience and lessons learned (what works or does not work), and
>
> (c) Enhancing accountability of foreign assistance to major stakeholders – the White House, OMB, the Congress, taxpayers and host countries by systematically examining if existing policies, programs and projects are achieving what they were intended to achieve.

2. Methodological and Analytical Rigor

Methodological and analytical rigor is necessary to generate empirically grounded findings and recommendations that are credible to different stakeholders of foreign assistance. The norms of scientific inquiry, which contribute to accurate, valid and reliable information, should be followed during the data gathering and analysis processes. Both qualitative and quantitative methods can be rigorous and both are usually required to answer evaluation questions about foreign assistance.

3. Independence

Evaluations should be free from any interference from the agency/office or operating unit commissioning them. Evaluators should be able to gather and analyze data and information freely and fairly following the cannons of scientific inquiry. The independence of evaluations can be ensured in the following ways:

(a) Evaluations are commissioned by a management unit which is separate from and superior to the one which manages the program.

(b) The responsibility of conducting evaluations is given to an outside research or evaluation organization, which is not accountable to the managers of the program.

(c) The operating unit assures that evaluation findings and recommendations reflect empirical data and analysis and resists the temptation to inappropriately influence evaluation outcomes.

Evaluation Standards

1. Feasibility Standards

The feasibility standards are intended to ensure that an evaluation will be realistic, politically viable and cost effective.

1.1 Research Viability: Research viability requires that (a) evaluation goals and questions are plausible, and (b) relevant data and information can be gathered within the given time and resources. Before commissioning an evaluation, the research viability of evaluation enterprise should be carefully examined.

1.2 Cooperation of Stakeholders: The success of foreign assistance evaluations partly depends upon the cooperation of concerned stakeholders, particularly the host country organizations and partnering institutions. Therefore the consent and cooperation of stakeholders should be secured prior to commissioning evaluations. If they have genuine doubts and concerns, these should be addressed at the outset.

1.3 Cost: Effectiveness: The expected benefits from evaluation should be more or at least of equal value to the resources expended on it. While it is generally not feasible to quantify the potential benefits and costs of an evaluation, the commissioning agency/office/operating unit should examine the cost effectiveness of an evaluation and the way in which its findings will be utilized.

2. Evaluation Design Standards

These standards are intended to ensure that an evaluation will serve the information needs of intended users.

2.1 Evaluation Questions: Evaluation questions should be specific, precise and reflect the information needs of decision-makers. Moreover, it should be possible to answer evaluation questions on the basis of empirical data. As far as possible, the number of questions should be kept limited. A large number of questions can result in an unfocussed evaluation, in addition to adding to the cost of an evaluation.

2.2 Data Collection Methods: The choice of data collection methods should largely depend on the nature of the evaluation questions. Different methods - sample surveys, content analysis, statistical analysis of available data, key informant interviews, focus group discussions, participant observation, content analysis of documents and case studies - have distinct features that make them either more or less appropriate for answering a particular type of question credibly. Most foreign assistance evaluations require methodological triangulation of different methods to enhance the accuracy, reliability and credibility of the findings and recommendations.

2.3 Qualifications of Evaluators: While evaluation skills are critical, because of the complex nature of foreign assistance programming, political and cultural sensibilities involved and the vagaries of data collection in foreign environments, factors in addition to evaluation expertise should be taken into consideration in the selection of evaluators and evaluation teams. These should include language proficiency, subject-matter knowledge, in-country work experience, inter-personal communication skills, gender mix and gender analysis skills, and familiarity with the policies and programs of foreign assistance.

2.4 Host Nationals' Participation: When an evaluation requires data collection and analysis in a foreign country or countries, efforts should be made to secure the participation of host country experts. Their participation provides an "insider perspective," facilitates evaluation capacity-building, and increases the likelihood of the acceptance of evaluation recommendations.

2.5 Resources and Time Frame: Adequate resources and time should be provided to an evaluation team taking into consideration the scope of the evaluation and the nature of data collection methods.

3. Data Quality Standards

The data quality standards are intended to ensure that an evaluation will answer the evaluation questions on the basis of accurate data and information.

3.1 Validity: The collected data and information should measure what they purport to measure and should provide sufficient evidence for evaluation findings and conclusions. This requires that evaluators conform to the norms of social science research and pay particular attention to minimizing biases generated by inappropriate research instruments, flawed interviewing, and unrepresentative sampling.

3.2 Reliability: Consistency or dependability of data with reference to the quality of the instruments and procedures used is necessary for the credibility of evaluation findings and conclusions. The information gathering procedures should ensure that the repeated use of the same research instruments will generate comparable findings and conclusions.

3.3 Timeliness: The gathered data and information should be timely to influence decision-making. When recent data are not available or cannot be gathered and the old data and information have to be used, efforts should be made to upgrade them and/or supplement them with additional information.

3.4 Strengths and Limitations of Data: Because of the vagaries of data collection in foreign countries with limited resources and within limited time, the sources of data and information and their strengths and limitations should be carefully identified. This enables managers of foreign assistance programs to make informed judgments about evaluation findings and recommendations.

4. Evaluation Reporting Standards

Reporting standards are designed to ensure that evaluation findings and recommendations will be presented in a way as to maximize their use by decision-makers.

4.1 Clarity: An evaluation report should be clear, focused and comprehensive. It should clearly describe the project, program or the policy evaluated, evaluation questions and context, data collection methods, the generated data, findings, conclusions and recommendations in a simple, coherent fashion.

4.2 Findings and Conclusions: Findings should logically follow from the gathered data and conclusions from the findings. Because conclusions involve interpretation of collected data, they should be explicitly justified. If and when necessary, the evaluator should state his/her assumptions, judgments and value premises so that managers can better understand and assess them.

4.3 Recommendations: The recommendations should be specific, practical and actionable. Evaluators should take into consideration the economic and political context of the project, program or policy evaluated, the strengths and weaknesses of the implementing organization, available resources, and the feasibility of change and innovation while framing recommendations.

5. Propriety Standards

The propriety standards are intended to ensure that an evaluation will be conducted ethically.

5.1 Rights of Human Subjects: When human subjects are involved, evaluations should be designed and conducted to respect and protect their rights and welfare.

5.2 Privacy and Confidentiality of Information: The privacy and confidentiality of information should be maintained. If sensitive information is involved, efforts should be made to ensure that the identity of informants is not disclosed.

5.3 Conflict of Interest: Any conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.