GLOSSARY OF EVALUATION TERMS

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

INTRODUCTION

This "Glossary of Evaluation and Related Terms" was jointly prepared by a committee consisting of evaluation staff of the Office of the Director of U.S. Foreign Assistance of the State Department and USAID to develop a broad consensus about various terms used in monitoring and evaluation of foreign assistance. The primary intended users are the State Department and USAID staff responsible for initiating and managing evaluations of foreign assistance programs and projects. In preparing this glossary, the committee reviewed a wide range of evaluation glossaries published by bilateral and multilateral donor agencies, professional evaluation associations and educational institutions. The committee particularly relied on the glossaries issued by the Evaluation Network of the OECD's Development Assistance Committee, the U.S. Government Accountability Office, the National Science Foundation, the American Evaluation Association and Western Michigan University.
(Core definitions are asterisked.)

Accountability:
Obligation to demonstrate what has been achieved in compliance with agreed rules and standards. This may require a careful, legally defensible, demonstration that work is consistent with the contract.

*Activity:
A specific action or process undertaken over a specific period of time by an organization to convert resources to products or services to achieve results. Related term: Project.

Analysis:
The process of breaking a complex topic or substance into smaller parts in order to examine how it is constructed, works, or interacts to help determine the reason for the results observed.

Anecdotal Evidence:
Non-systematic qualitative information based on stories about real events.

*Appraisal:
An overall assessment of the relevance, feasibility and potential sustainability of an intervention or an activity prior to a decision of funding.

*Assessment:
A synonym for evaluation.

Assumptions:

A proposition that is taken for granted, as if it were true. For project management, assumptions are hypotheses about causal linkages or factors that could affect the progress or success of an intervention.

***Attribution:**

Ascribing a causal link between observed changes and a specific intervention(s) or program, taking into account the effects of other interventions and possible confounding factors.

Audit:

The systematic examination of records and the investigation of other evidence to determine the propriety, compliance, and adequacy of programs, systems, and operations.

***Baseline:**

Information collected before or at the start of a project or program that provides a basis for planning and/or assessing subsequent progress and impact.

Benchmark:

A standard against which results are measured.

***Beneficiaries:**

The individuals, groups, or organizations that benefit from an intervention, project, or program.

***Best Practices:**

Methods, approaches, and tools that have been demonstrated to be effective, useful, and replicable.

Bias:

The extent to which a measurement, sampling, or analytic method systematically underestimates or overestimates the true value of a variable or attribute.

Case control:

A type of study design that is used to identify factors that may contribute to a medical condition by comparing a group of patients who already have the condition with those who do not, and looking back to see if the two groups differ in terms of characteristics or behaviors.

***Case Study:**

A systematic description and analysis of a single project, program, or activity.

*Causality:

The relationship between one event (the cause) and another event (the effect) which is the direct consequence (result) of the first.

Cluster sampling:

A sampling method conducted in two or more stages in which each unit is selected as part of some natural group rather than individually (such as all persons living in a state, city block, or a family).

*Conclusion:

A judgment based on a synthesis of empirical findings and factual statements.

Confounding variable (also confounding factor or confounder):

An extraneous variable in a statistical model that correlates (positively or negatively) with both the dependent and independent variable and can therefore lead to faulty conclusions.

*Construct Validity:

The degree of agreement between a theoretical concept (e.g. peace and security, economic development) and the specific measures (e.g. number of wars, GDP) used as indicators of the phenomenon; that is the extent to which some measure (e.g., number of wars, GDP) adequately reflects the theoretical construct (e.g., peace and security, economic development) to which it is tied.

*Content Validity:

The degree to which a measure or set of measures adequately represents all facets of the phenomena it is meant to describe.

Control Group:

A randomly selected group that does not receive the services, products or activities of the program being evaluated.

Comparison Group:

A non-randomly selected group that does not receive the services, products or activities of the program being evaluated.

*Counterfactual:

A hypothetical statement of what would have happened (or not) had the program not been implemented.

*Cost Benefit Analysis:

An evaluation of the relationship between program costs and outcomes. Can be used to compare different interventions with the same outcomes to determine efficiency.

Data:
Information collected by a researcher. Data gathered during an evaluation are manipulated and analyzed to yield findings that serve as the basis for conclusions and recommendations.

Data Collection Methods:
Techniques used to identify information sources, collect information, and minimize bias during an evaluation.

*Dependent Variable:
Dependent (output, outcome, response) variables, so called because they are "dependent" on the independent variable; the outcome presumably depends on how these input variables are managed or manipulated.

*Effect:
Intended or unintended change due directly or indirectly to an intervention. Related terms: results, outcome.

*Effectiveness:
The extent to which an intervention has attains its major relevant objectives. Related term: efficacy.

*Efficiency:
A measure of how economically resources/inputs (funds, expertise, time etc.) are used to achieve results.

Evaluability:
Extent to which an intervention or project can be evaluated in a reliable and credible fashion.

Evaluability Assessment:
A study conducted to determine a) whether the program is at a stage at which progress towards objectives is likely to be observable; b) whether and how an evaluation would be useful to program managers and/or policy makers; and, c) the feasibility of conducting an evaluation.

*Evaluation:
A systematic and objective assessment of an on-going or completed project, program or policy. Evaluations are undertaken to (a) improve the performance of existing interventions or policies, (b) asses their effects and impacts, and (c) inform decisions about future programming. Evaluations are formal analytical endeavors involving systematic collection and analysis of qualitative and quantitative information.

Evaluation Design:

The methodology selected for collecting and analyzing data in order to reach defendable conclusions about program or project efficiency and effectiveness.

Experimental Design:

A methodology in which research subjects are randomly assigned to either a treatment or control group, data is collected both before and after the intervention, and results for the treatment group are benchmarked against a counterfactual established by results from the control group.

*External Evaluation:

The evaluation of an intervention or program conducted by entities and/or individuals which is not directly related to the implementing organization.

*External Validity:

The degree to which findings, conclusions, and recommendations produced by an evaluation are applicable to other settings and contexts.

*Findings:

Factual statements about a project or program which are based on empirical evidence. Findings include statements and visual representations of the data, but not interpretations, judgments or conclusions about what the findings mean or imply.

Focus Group:

A group of people convened for the purpose of obtaining perceptions or opinions, suggesting ideas, or recommending actions. A focus group is a method of collecting information for the evaluation process that relies on the particular dynamic of group settings.

*Formative Evaluation:

An evaluation conducted during the course of project implementation with the aim of improving performance during the implementation phase. Related term: process evaluation.

*Goal:

The higher-order objective to which a project, program, or policy is intended to contribute.

*Impact:

A results or effect that is caused by or attributable to a project or program. Impact is often used to refer to higher level effects of a program that occur in the medium or long term, and can be intended or unintended and positive or negative.

*Impact Evaluation:
A systematic study of the change that can be attributed to a particular intervention, such as a project, program or policy. Impact evaluations typically involve the collection of baseline data for both an intervention group and a comparison or control group, as well as a second round of data collection after the intervention, some times even years later.

*Independent Evaluation:
An evaluation carried out by entities and persons not directly involved in the design or implementation of a project or program. It is characterized by full access to information and by full autonomy in carrying out investigations and reporting findings.

*Independent Variable:
A variable that may influence or predict to some degree, directly or indirectly, the dependent variable. An independent variable may be able to be manipulated by the researcher (for example, introduction of an intervention in a program) or it may be a factor that cannot be manipulated (for example, the age of beneficiaries).

*Indicator:
Quantitative or qualitative variable that provides reliable means to measure a particular phenomenon or attribute.

*Inputs:
Resources provided for program implementation. Examples are money, staff, time, facilities, equipment, etc.

*Internal Evaluation:
Evaluation conducted by those who are implementing and/or managing the intervention or program. Related term: self-evaluation.

*Internal Validity:
The degree to which conclusions about causal linkages are appropriately supported by the evidence collected.

Intervening Variable:
A variable that occurs in a causal pathway from an independent to a dependent variable. It causes variation in the dependent variable, and itself is caused to vary by the independent variable.

*Intervention:
An action or entity that is introduced into a system to achieve some result. In the program evaluation context, an intervention refers to an activity, project or program that is introduced or changed (amended, expanded, etc).

*Joint Evaluation:
An evaluation in which more than one agency or partner participates. There can be varying levels of collaboration ranging from developing an agreed design and conducting fieldwork independently to pooling resources and undertaking joint research and reporting.

*Lessons learned:
Generalizations based on evaluation findings that abstract from the specific circumstances to broader situations. Frequently, lessons highlight strengths or weaknesses in preparation, design, and implementation that affect performance, outcome and impact.

Level of Significance
The probability that observed differences did not occur by chance.

Logic Model:
A logic model, often a visual representation, provides a road map showing the sequence of related events connecting the need for a planned program with the programs' desired outcomes and results.

*Logical Framework (Logframe):
A management tool used to improve the design and evaluation of interventions that is widely used by development agencies. It is a type of logic model that identifies strategic project elements (inputs, outputs, outcomes, impact) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure. Related term: Results Framework.

Measurement:
A procedure for assigning a number to an observed object or event.

*Meta-evaluation:
A systematic and objective assessment that aggregates findings and recommendations from a series of evaluations.

*Mid-term Evaluation:
Evaluation performed towards the midpoint of program or project implementation.

Mixed Methods:
Use of both quantitative and qualitative methods of data collection in an evaluation.

*Monitoring:
The performance and analysis of routine measurements to detect changes in status. Monitoring is used to inform managers about the progress of an ongoing intervention

or program, and to detect problems that may be able to be addressed through corrective actions.

*Objective:
A statement of the condition or state one expects to achieve.

Operations Research:
A process to identify and solve problems in program implementation and service delivery. It involves identifying a problem and testing a strategy to address the problem. The goal is to increase the efficiency, effectiveness, and quality of services delivered by providers, and the availability, accessibility, and acceptability of services desired by users.

*Outcome:
A results or effect that is caused by or attributable to the project, program or policy. Outcome is often used to refer to more immediate and intended effects. Related terms: result, effect.

*Outcome Evaluation:
This form of evaluation assesses the extent to which a program achieves its outcome-oriented objectives. It focuses on outputs and outcomes (including unintended effects) to judge program effectiveness but may also assess program processes to understand how outcomes are produced.

*Outputs:
The products, goods, and services which result from an intervention.

*Participatory Evaluation:
An evaluation in which managers, implementing staff and beneficiaries work together to choose a research design, collect data, and report findings.

*Performance Indicator:
A particular characteristic or dimension used to measure intended changes. Performance indicators are used to observe progress and to measure actual results compared to expected results.

*Performance Management:
Systematic process of collecting and analyzing performance data to track progress towards planned results to improve resource allocation, implementation, and results.

Performance Measurement:
Ways to objectively measure the degree of success that a program has had in achieving its stated objectives, goals, and planned program activities.

Primary Data:
Information collected directly by the researcher (or assistants), rather than culled from secondary sources (data collected by others). In program evaluation, it refers to the information gathered directly by an evaluator to inform an evaluation.

Process:
The programmed, sequenced set of things actually done to carry out a program or project.

Process Evaluation:
An assessment conducted during the implementation of a program to determine if the program is likely to reach its objectives by assessing whether or not it is reaching its intended beneficiaries (coverage) and providing the intended services using appropriate means (processes).

*Program:
A set of interventions, activities or projects that are typically implemented by several parties over a specified period of time and may cut across sectors, themes and/or geographic areas.

*Program Evaluation:
Evaluation of a set of interventions designed to attain specific global, regional, country, or sector development objectives. A program is a time-bound intervention involving multiple activities that may cut across sectors, themes and/or geographic areas.

*Project:
A discrete activity (or 'development intervention') implemented by a defined set of implementers and designed to achieve specific objectives within specified resources and implementation schedules. A set of projects make up the portfolio of a program. Related term: activity, intervention.

Project Appraisal:
A comprehensive and systematic review of all aspects of the project — technical, financial, economic, social, institutional, environmental — to determine whether an investment should go ahead.

*Project Evaluation:
An evaluation of a discrete activity designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader program.

Qualitative Data:
Observations or information expressed using categories (dichotomous, nominal, ordinal) rather than numerical terms. Examples include sex, survival or death, and first.

Quantitative Data:
Information that can be expressed in numerical terms, counted, or compared on a scale.

*Quasi-experimental Design:
A methodology in which research subjects are assigned to treatment and comparison groups typically through some sort of matching strategy that attempts to minimize the differences between the two groups in order to approximate random assignment.

Random Assignment:
The process of assigning research subjects in such a way that each individual is assigned to either the treatment group or the control group entirely by chance. Thus, each research subject has a fair and equal chance of receiving either the intervention being studied (by being placed in the treatment group), or not (by being placed in the "control" group). Related term: randomization.

Randomized Controlled Trials (RCT):
Research studies that use an experimental design. Related terms: experimental design.

*Rapid Appraisal Methods:
Data collection methods, which fall within the continuum of formal and informal methods of data collection, that can be used to provide decision-related information in development settings relatively quickly. These include key informant interviews, focus group discussions, group interviews, structured observation and informal surveys.

*Recommendations:
Proposals based on findings and conclusions that are aimed at enhancing the effectiveness, quality, or efficiency of an intervention.

*Reliability:
Consistency or dependability of data with reference to the quality of the instruments, procedures and used. Data are reliable when the repeated use of the same instrument generates the same results.

*Result:
The output, outcome or impact intended (or unintended).

*Results Framework:
A management tool, that presents the logic of a project or program in a diagrammatic form. It links higher level objectives to its intermediate and lower level objectives. The diagram (and related description) may also indicate main activities, indicators, and

strategies used to achieve the objectives. The results framework is used by managers to ensure that its overall program is logically sound and considers all the inputs, activities and processes needed to achieve the higher level results.

Risk Analysis:
An analysis or an assessment of factors (called assumptions in the log frame) that affect, or are likely to affect, the successful achievement of an intervention's objectives. It is a systematic process to provide information regarding undesirable consequences based on quantification of the probabilities and/or expert judgment.

*Scope of Work:
A written description of the objectives, tasks, methods, deliverables and schedules for an evaluation.

Sector Program Evaluation:
An evaluation of a cluster of interventions in a sector within one country or across countries, all of which contribute to the achievement of a specific goal.

*Self-Evaluation:
An evaluation by those who are entrusted with the design and implementation of a project or program.

*Stakeholders:
Entities (governments, agencies, companies, organizations, communities, individuals, etc.) that have a direct or indirect interest in a project, program, or policy and any related evaluation.

*Summative Evaluation:
Evaluation of an intervention or program in its later stages or after it has been completed to (a) assess its impact (b) identify the factors that affected its performance (c) assess the sustainability of its results, and (d) draw lessons that may inform other interventions.

Survey:
Systematic collection of information from a defined population through interviews or questionnaires.

Sustainability:
The degree to which services or processes continue once inputs (funding, materials, training, etc.) provided by the original source(s) decreases or discontinues.

Target:
The specified result(s), often expressed by a value of an indicator(s), that a project, program, or policy is intended to achieve.

*Target Group:
The specific individuals, groups, or organizations for whose benefit the intervention is undertaken.

Treatment:
A project, program, or policy that is the subject of an evaluation.

*Thematic Evaluation:
An evaluation that focuses on a specific cross cutting area.

*Validity:
The extent to which data measures what it purports to measure and the degree to which that data provides sufficient evidence for the conclusions made by an evaluation.

Variable:
An attribute or characteristic in an individual, group, or system that can change or be expressed as more than one value or in more than one category.