Volume 4: Impact Evaluation and Feed the Future

What are Impact Evaluations?

According to the USAID Evaluation Policy, "impact evaluations measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provides the strongest evidence of a relationship between the intervention under study and the outcome measured." Unlike general evaluations, which can answer many types of questions, impact evaluations are structured around one particular type of question: What is the impact (or causal effect) of a program on an outcome of interest? An impact evaluation looks for the changes in outcome that are directly attributable to the program. For example, did the adoption and diffusion of drought resistant rice seed among farmers in Dodoma District in Tanzania, caused household income to increase?

.

The Purpose of Impact Evaluation

Impact evaluations are aimed at providing feedback to help improve the design of programs and policies. In addition to providing for improved accountability, impact evaluations are a tool for dynamic learning, allowing development stakeholders to improve ongoing programs and ultimately better allocate funds across programs. Impact evaluations are also needed to inform policy makers on a range of decisions, from curtailing inefficient programs, to scaling up interventions that work, to adjusting program benefits, to selecting among various program alternatives.

The Role of Impact Evaluation in Monitoring and Evaluation

In the project cycle, there are clear and important differences, as well as, linkages between 'monitoring' and 'impact evaluation'. Monitoring focuses primarily on tracking inputs and outputs, and evaluation focuses on measuring outcomes and impact. In general, project inputs and outputs, should produce outcomes and impacts.

Impact evaluations fit into the chain of monitoring and evaluation process in several ways:

1. They help to assess the causal link between an intervention and an outcome;

2. Impact evaluations provide evidence for the effectiveness of an intervention, which can be compared with other similar interventions. Through this process, impact evaluations assist in establishing credible cost-effectiveness comparisons;

3. Impact evaluations can serve to build the knowledge base of what works. With an increasing demand for evidence of USAID effectiveness, rigorous evaluations offer a method through which development successes can be highlighted.

To measure the impact of an intervention, a clear, well-designed evaluation strategy is necessary. Incorporating an impact evaluation into a development program requires a well-planned monitoring and evaluation process. Missions are advised to develop a comprehensive impact evaluation plan as part of their FTF M&E plan.

The impact evaluation plan, include establishing the type of question to be answered by the evaluation, constructing a theory of change that outlines how the project is supposed to achieve the intended results, developing a results chain, formulating hypotheses to be tested by the evaluation, and selecting performance indicators. All of these help to determine the evaluation questions. These questions should be identified at the start of the project, and missions are advised to engage a range of stakeholders from policy makers to program managers, to forge a common vision and agree on the questions the impact evaluations should answer. By providing critical feedback about what works, and what does not, impact evaluations can help to solidify a results-based project structure.

When to Conduct an Impact Evaluation

Impact evaluations require a substantial amount of information, time, and resources. Therefore, it is important to select carefully the interventions that will be evaluated. USAID's new Evaluation Policy requires that operating units conduct an impact evaluation of "any activity within a project involving untested hypotheses or demonstrating new approaches that are anticipated to be expanded in scale or scope through US[G] foreign assistance or other funding sources." In order to justify the technical and financial resources to conduct an impact evaluation of an FTF program intervention, the following questions must be answered regarding the intervention:

1. Is the intervention INNOVATIVE? Is it testing a new, promising approach?
2. Is the intervention REPLICABLE? Can it be scaled up or can it be applied to a different setting?
3. Is the intervention STRATEGICALLY RELEVANT? Is it a flagship intervention that requires substantial resources; covers, or could be expanded to cover, a large number of people; or could generate substantial savings.
4. Is the intervention UNTESTED? That is very little known about the effectiveness of the intervention globally or in the specific context in which it is implemented?
5. Is the intervention INFLUENCIAL? Can the results be used to inform key policy decisions?

Impact Evaluation Design under FTF

The impact of the FTF program will be independently and scientifically evaluated by recognized experts in agricultural development and nutrition, using the most rigorous evaluation methods possible. USAID's new approach to evaluation is different due to the rigorous social-science designs and methods that are to be used.

Under FTF, impact evaluations will use target and control groups to compare situations and changes between those groups receiving interventions those that are not. The control group, or counterfactual, should be a group which is as similar as possible in observable and unobservable dimensions to those receiving the intervention in order to define a hypothetical situation of what would occur in the absence of the program. This comparison allows for the establishment of definitive causality--attributing observed changes to the program, while removing confounding factors.

Impact evaluations will be grouped into two categories: experimental design (Category 1) and quasi-experimental design (Category 2). Category 1 evaluations will be those that are designed as Randomized Control Trials (RCTs). This category of evaluations will use an experimental design by constructing credible counterfactual scenarios, with the most credible being the random selection of treatment and control groups. Category 1 evaluation will typically commence at the beginning of a project's implementation, to give evaluators the opportunity to work closely with program implementation staff to design the evaluation, and obtain data throughout the life of the project.

When randomization is not possible, a quasi-experimental design will be applied. In this design, a comparison group that resembles the treatment group will be generated, at least in observed socio-cultural, economic, ecological, and geographic characteristics. Quasi-experimental methods include propensity score matching method, double differences methods, instrumental variable methods, or reflexive comparisons. Depending on the circumstances, and context in which the project is being implemented, Category 2 evaluations will commence near the beginning of a project's implementation, sometimes during implementation, or after a project is completed. Ideally an evaluation plan should be built into a program intervention as the program is

designed, which means that the evaluation should be planned as early as possible. Although it may not be possible to establish a distinct control group, Category 2 evaluations should use the most rigorous methodology possible to assess whether anticipated results were achieved.

Impact Evaluation Methods under FTF

Data collection and analysis under FTF impact evaluations can employ quantitative or qualitative methods or both, although an evaluation design that uses "mixed methods" (both quantitative and qualitative) is usually best in that it offers both rigor and richness of data. Good research that convinces a range of clients on the difficult questions of causality requires a combination of techniques.

Quantitative methods help to understand and establish the basic relationship between two or more variables, often specifically looking for correlation and causation. Under FTF, quantitative data collection will likely center around individual or household surveys that gather data from a representative sample of the intervention's population on those variables or indicators critical to the evaluation questions being tested. Those data collected under the survey would then be analyzed using regression-based statistical methods. Available secondary data sets might also be useful for analysis.

Qualitative methods are also useful to also study relationships between variables and can substantiate quantitative findings by providing richer, more complex data on those relationships. Qualitative methods provide for a deeper understanding of particular phenomena, facilitate a wider range of explanations, and can help identify unintended impacts. Under FTF, the most common or useful data collection methods would likely be interviews (structured to unstructured), focus groups, participant and direct observation. Rigorous analysis qualitative data will require systematic coding of data to be analyzed through methods such as content analysis, analytical induction, or constant comparison methods.

References

Baker, Judy L., Evaluating the impact of development projects on poverty: a handbook for practitioners, Washington DC: The World Bank

Gertler, Paul J. Gertler, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch, 2011, Impact evaluation in practice, Washington DC: The World Bank (http://www.worldbank.org/ieinpractice)

USAID/PPL/LER, 2011, USAID evaluation Policy, Washington DC: USAID