



Impact Evaluation

A Discussion Paper for AusAID Practitioners

AusAID supports a broad definition of impact evaluation that is focused on systematic assessment of cause and effect. Impact evaluations are best conducted where there is clear need and intent to use the evidence for decision-making. Questions regarding value for money of the evaluation approach may also be considered. A range of approaches can be used for impact evaluation; the choice of which approach to use will be based on the purpose, questions and context of the evaluation.

Introduction

There are diverse views about what impact evaluations are and how they should be conducted. These views are affected by different theories about how development change occurs, and debate about what constitutes 'rigorous' standards of evidence. AusAID, along with others in the field of impact evaluation, is exploring what these debates mean for its evaluation practice.

This discussion paper is an initiative of AusAID's Impact Evaluation Working Group¹ and arises from two concerns:

- Most aid evaluations focus mainly on whether aid outputs and objectives have been achieved (for example, schools built and teachers trained). It is more difficult to assess whether aid has actually made a difference in people's lives (children in school longer and learning better). Impact evaluation attempts to fill this gap. As there is increasing interest in AusAID for commissioning impact evaluation, there is a need for a consistent and quality approach to impact evaluation within the agency. The Impact Evaluation Working Group identified common confusion in AusAID about what impact means and what constitutes a quality approach. The Working Group also had concerns that, without a good understanding of impact evaluation, there was a risk it could be used in situations when other approaches might be more suitable.
- Some organisations and evaluators limit the definition of impact evaluation to one methodological approach (for example, Randomised Control Trials)—however, others have shown that no single evaluation approach will be suitable for all development contexts. It is not always easy for AusAID staff to identify and understand good alternatives, which may limit the value that AusAID can obtain from impact evaluation.

¹ An Impact Evaluation Working Group was formed in 2011 by a sub-group of the Performance and Quality Network. It was led by the Office of Development Effectiveness (ODE) and the Program Effectiveness and Performance Division, and included representatives from the Research Section and program performance & quality areas. This discussion paper was authored by ODE with input from the working group members. This paper was also technically reviewed by Patricia Rogers, Howard White, and AusAID's Independent Evaluation Committee.

This paper aims to support appropriate and effective use of impact evaluations in AusAID by providing AusAID staff with a benchmark for navigating the diverse views about impact evaluation. It provides staff who commission impact evaluations with a definition, principles and minimum standards. It does not provide guidance—staff commissioning impact evaluations should seek assistance from their Performance & Quality area on how to put these principles into practice. ODE is also a potential source of assistance and advice.

What is impact and impact evaluation?

Impacts are positive or negative changes produced by a development intervention—directly or indirectly, intended or unintended—in the context of its environment, as it interacts with the multiple factors affecting development change.²

Impact occurs at multiple levels and timeframes—there can be short-term, intermediate and long-term changes resulting from an intervention. How and when impact occurs will differ depending on the type of intervention and the context.

An impact evaluation is a systematic and empirical investigation of the impacts produced by an intervention—specifically, it seeks to establish whether an intervention has made a difference in the lives of people. It aims to answer questions about what works or does not work, how, for whom, and why.

To provide these answers, impact evaluation **links cause and effect**: it assesses the direct and indirect **causal contribution of the intervention** to change in people's lives.

Therefore, an impact evaluation:

- evaluates the positive and negative, primary and secondary long-term impacts that result from a development intervention
- assesses the direct and indirect contribution of these interventions to such impacts on final beneficiaries, especially for the poor, whether intended or unintended
- explains how policy interventions contribute to an impact so that lessons can be learned.³

Impact evaluations programs should be built into the design of an intervention and be conducted throughout its life, or commence during or after the life of an intervention. The timing and duration of an impact evaluation depends on the methodological approach used, and the agreement of stakeholders. Some impact evaluations will take multiple years to generate findings—although it is possible to conduct shorter impact evaluations.

Impact evaluation complements other types of reviews and evaluation that mainly focus on the nature, process and intended results of the intervention.⁴ Other types of evaluation may include assessment of the intervention's outcomes and unintended consequences, but generally do not aim to establish the causal contribution of the intervention to development outcomes.

² This is based on the full definition of impact in the OECD DAC criteria for evaluating development assistance: 'The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results and must also include the positive and negative impact of external factors, such as changes in terms of trade and financial conditions'. See <http://www.oecd.org/dataoecd/42/6/49756382.pdf>

³ Elliot Stern et. al. (2012) [Broadening the range of designs and methods for impact evaluations: report of a study commissioned by the Department for International Development](#), p. 12.

⁴ See 'Evaluations and Reviews' guidance note for more information.

When should we do impact evaluations?

Impact evaluation is not required, or suitable, for every program. There must be both a clear need for the evaluation and intent by program staff and decision-makers to use the findings. Value for money of the evaluation approach may also be a consideration.

Clear need

The decision to undertake an impact evaluation should be based on its purpose and intended uses, and consideration of its merits relative to other monitoring and evaluation activities that are needed to inform decisions.

Other factors that might affect whether you do an impact evaluation are the significance of the intervention in terms of aid investment, its importance to strategic objectives or the degree of developmental risk.

Situations in which impact evaluation can be valuable include:

- to establish the value of innovative and effective programs, including pilots
- to test the casual logic and assumptions of an intervention that has long been used, but has little and/or contested evidence on how it contributes to development outcomes
- to test the validity of a successful intervention in a different context
- to prove the worth of an intervention to policy-makers and decision-makers.

Value for money

Impact evaluations are generally more expensive, and have higher transaction costs on partners, than other types of evaluation. Therefore it is imperative to design and manage the evaluation to ensure that the chosen option represents value for money over other evaluation approaches and is targeted towards strategic aid investments.

Intent to use

A major contributor to value for money is ensuring that impact evaluations are **demand-driven** and responsive to key questions of decision-makers. All impact evaluations should be **user-focused**, that is, designed and managed to focus on meeting the intended uses of primary intended users. The evaluation process needs to build ownership and interest among key stakeholders, and promote the use of the results. At a minimum:

- All key stakeholders should be consulted before the evaluation is commissioned, and engaged regularly as the evaluation is underway.
- Results of impact evaluations should be appropriately communicated to its intended users—for example, by producing a policy brief that summarises the key findings in a format accessible and relevant to decision-makers.⁵

Impact evaluation approaches and methods

AusAID recognises that **a range of methodological approaches can be used for impact evaluation**. The need for an impact evaluations starts with a need for evidence (as detailed above). The purpose and

⁵ For an example of a good policy brief, see 3ie's Evidence Matters briefs at http://www.3ieimpact.org/evidence_matters.html

questions of the evaluation, the complexity of the intervention and the context of the intervention, will determine what types of methods are used. Impact evaluations in AusAID should not be methods-driven.

Mixing approaches and methods

It is preferable to combine more than one approach to assess causation, in order to compensate for the weaknesses inherent in any single impact evaluation approach. The key is to focus on what data is needed to answer the evaluation questions and that is available or can be generated. Then use an appropriate range of approaches and methods to collect that data and build a quality evidence base.⁶

Like all evaluations, impact evaluations need to ask relevant questions, and have valid measures, indicators, quality data, and appropriate criteria and standards. In addition, impact evaluations require appropriate approaches and methods to establish cause and effect relationships.

Common strategies include:

- analysis of the factual situation (whether what has happened in the intervention and the broader environment is consistent with the theory that the intervention has contributed to observed changes)
- counter-factual analysis (hypothetically, what might have happened without the intervention)
- identifying and testing plausible alternative explanations for changes.⁷

Particular approaches apply one or more of these strategies, for example: realist evaluation, process tracing (factual analysis), experimental and quasi-experimental approaches (counter-factual analysis), process tracing, general elimination method (alternative explanations).

See Annex A for an example of the range of impact evaluations approaches and the ways causality is established.

Using experimental, quasi-experimental and statistical approaches

Experimental, quasi-experimental and statistical approaches focus on testing what might have happened without the intervention (see Box 1). These approaches are best suited to stable situations and programs with relatively simple causal logic. Given the scale of these evaluations, it is important that there is strong demand for the evidence and involvement from partners.

These approaches have value in certain situations but are not applicable for many of AusAID's programs.

BOX 1: HOW DO EXPERIMENTAL, QUASI-EXPERIMENTAL AND STATISTICAL APPROACHES MEASURE IMPACT?

Experimental designs create a counter-factual through a control group, using random assignment, including various types of randomised controlled trials. Quasi-experimental designs create a counter-factual through a comparison group, including regression discontinuity design, propensity score matching or matched comparisons. Time series designs and econometric models create a counter-factual by developing predictions of what would have happened in the absence of the intervention.

An advantage of these approaches is that, using random sampling, and quantitative measurement, they aim to produce an unbiased estimate of the net effect. This can be communicated clearly and simply to

⁶ See the six aspects to consider in choosing methods: Patricia J. Rogers (2012), p.5.

⁷ Patricia J. Rogers (2012), [Introduction to Impact Evaluation](#), Impact Evaluation Notes No. 1 March 2012, InterAction and the Rockefeller Foundation.

decision makers—for example, “treating primary school children with deworming drugs reduced absenteeism by 25 per cent”.⁸ In addition, high quality studies used mixed methods and theory-based evaluation to develop better explanations of why or why not interventions are successful.

Experimental, quasi-experimental and statistical approaches require a large sample size to be statistically valid, and therefore usually focus on large-scale service delivery programs. Their focus on the average effect across a particular group makes them less useful in situations where contextual factors (either within the project or in other projects where the findings might be used) are likely to have a major effect.

It is recommended that programs wishing to apply experimental approaches partner with experienced evaluation organisations who place an emphasis on linking the results into policy, such as 3ie, the Poverty Action Lab (J-PAL), or the Strategic Impact Evaluation Fund.

Using other impact evaluation approaches

There are a range of other approaches to impact evaluation that are applicable to a broad range of development situations. Examples of specific approaches include realist evaluation, contribution analysis, general elimination method, comparative case studies and process tracing.⁹

These approaches to impact evaluation are more likely to be suitable for small-scale interventions, programs in complex or fragile situations, or cases where an intervention is one of multiple causal factors leading to one or more outcomes.

The main advantage of these approaches to impact evaluation is that they can provide in-depth explanations of ‘how’, ‘why’ and ‘for whom’ an intervention has contributed to development change. By testing a theory or establishing the ‘mechanisms’ for change, they are also more likely to provide findings that can be applied in different contexts. These approaches can also be more flexible in their timing.

The main weakness of these approaches is that they lack of a single measure of impact, which may be less convincing to decision-makers unless effectively communicated. There is increasing interest internationally in exploring the rigorous use of impact evaluation methods.¹⁰ AusAID is supporting the ODI Methods Lab in testing and refining a broad range of methodological approaches for application of impact evaluation in international development.

Minimum standards

In line with international good practice,¹¹ minimum standards for AusAID’s impact evaluations should be:

- **Mixed approaches and methods**, incorporating both a mix of evaluation approaches and collection and integrated analysis of both quantitative and qualitative data.
- **Theory-based**, to identify and test the causal pathways.

⁸ Eric Roetman (2011) [A can of worms? Implications of rigorous impact evaluations for development agencies](#), International Initiative for Impact Evaluation Working Paper 11, p. 5.

⁹ For more information, see Elliot Stern et. al. (2012); Howard White and Daniel Phillips (2012) [Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework](#), International Initiative for Impact Evaluation Working Paper 15.

¹⁰ See, for example, Howard White and Daniel Phillips (2012).

¹¹ For example, see [American Evaluation Association Response to U.S. Department of Education Notice of proposed priority “Scientifically Based Evaluation Methods”](#); [European Evaluation Society Statement: The importance of a methodologically diverse approach to impact evaluation](#); [Impact Evaluation for Development \(IE4D\) Principles for Action](#); [3ie principles for impact evaluation](#).

- **Analysis of sub-populations**, including differential impacts on poor people, men and women, boys and girls.
- **Systematic collection and analysis of quality data.**
- **Appropriately resourced**, including providing teams with the time needed to apply the chosen methodology systematically and rigorously.

Quality standards

Impact evaluation in international development is informed by standards for social science research—although more progress has been made in reaching consensus on how these apply to impact evaluation using experimental methods.

Regardless of the approach and methods used, however, quality assurance should focus on the choice of appropriate methods for the evaluation purpose and context, appropriate criteria for judging the impact of the intervention, how rigorously the methods have been applied (including adequate measurement), whether legitimate conclusions have been drawn from the data and the utility of the evaluation.¹²

The quality of an impact evaluation will also be judged on whether it has both credible findings on the evaluation subject in its context (internal validity) and credible conclusions that apply to other contexts (external validity).

Getting added value from impact evaluation

Build a body of evidence to inform decision-making

At an institutional level, AusAID should think beyond conducting individual evaluations to how multiple evaluations can build together to inform decision-making. Knowing what works in one program is not enough; it is necessary to compare interventions and alternative approaches, as well as understanding how similar interventions work in different contexts. That is the evidence that is most useful for informing policy and program decisions.

Ways of building a body of evidence, in collaboration with partners, include:

- synthesis of evaluations that have already been done
- multiple evaluations in one sector (preferably across multiple donor and partner programs, not just AusAID)
- analysis during the design of a new program to test if what has worked in one context (as proven by impact evaluation) is replicable.

Support your partners to conduct impact evaluation

AusAID recognises that impact evaluation is not just about the needs of donors agencies; internationally, the majority of impact evaluations are focused on developing country social programs for the poor. Consider when it makes more sense to focus on supporting partner governments in impact evaluation, including:

- supporting partners to conduct impact evaluations of national development programs
- aiming to meet partners' needs for evidence through impact evaluations of AusAID-funded programs.

¹² See Elliot Stern et. al., Chapter 6.

Integrate impact evaluation into program design, monitoring and implementation

It is important to plan when and how impact evaluation will be conducted during an intervention's design. In most cases, a good impact evaluation will start at the design of an intervention, and continue throughout implementation.

In a good design, ongoing program monitoring can also facilitate impact evaluation, by collecting baseline data, periodic data and other evidence about the intervention and its context.¹³

Coordinating impact evaluations in AusAID

Planning and management of impact evaluation is decentralised to each country or sectoral program. It would be preferable for AusAID to move towards a more coordinated approach to impact evaluation in the Australian aid program. As a first step, it is important that all impact evaluations are included in a broader plan of evaluation for a country or sectoral program that is maintained by the Performance and Quality managers for those programs.

Feedback

ODE welcomes comments and observations on this paper at ODE@ausaid.gov.au

Links to impact evaluation resources

Elliot Stern et.al. (2012) [*Broadening the Range of Designs and Methods for Impact Evaluation*](#), Department for International Development Working Paper 38.

Patricia J. Rogers (2012), [*Introduction to Impact Evaluation*](#), Impact Evaluation Notes No. 1 March 2012, InterAction and the Rockefeller Foundation.

Howard White and Daniel Phillips (2012) [*Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework*](#), International Initiative for Impact Evaluation Working Paper 15.

Paul J. Gertler et.al. (2011) [*Impact Evaluation in Practice*](#), The World Bank.

Network of Networks on Impact Evaluation (2009) [*Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*](#)

3ie Working Papers (on impact evaluation methods and practice), available at:

<http://www.3ieimpact.org/en/evaluation/working-papers/> and resources:

<http://www.3ieimpact.org/en/evaluation/resources/>

IE4D Group (2011), [*Impact Evaluation for Development: Principles for Action*](#).

[European Evaluation Society Statement: The importance of a methodologically diverse approach to impact evaluation](#)

[American Evaluation Association Response to U.S. Department of Education Notice of proposed priority. Federal register RIN 1890-ZA00, November 4, 2003 "Scientifically Based Evaluation Methods"](#)

¹³ See Burt Perrin (2012), [*Linking Monitoring and Evaluation to Impact Evaluation*](#), Impact Evaluation Notes No. 2 April 2012, InterAction and the Rockefeller Foundation.

Annex A: How different impact evaluation methods establish causality

Table 3.3: Design Approaches, Variants and Causal Inference		
Design Approaches	Specific Variants	Basis for Causal Inference
Experimental	RCTs Quasi Experiments, Natural Experiments	Counterfactuals; the co-presence of cause and effects
Statistical	Statistical Modelling Longitudinal Studies Econometrics	Correlation between cause and effect or between variables, influence of (usually) isolatable multiple causes on a single effect Control for ‘confounders’
Theory-based	<i>Causal process designs:</i> Theory of Change, Process tracing, Contribution Analysis, impact pathways <i>Causal mechanism designs:</i> Realist evaluation, Congruence analysis	Identification/confirmation of causal processes or ‘chains’ Supporting factors and mechanisms at work in context
‘Case-based’ approaches	<i>Interpretative:</i> Naturalistic, Grounded theory, Ethnography <i>Structured:</i> Configurations, QCA, Within-Case- Analysis, Simulations and network analysis	Comparison across and within cases of combinations of causal factors Analytic generalisation based on theory
Participatory	<i>Normative designs:</i> Participatory or democratic evaluation, Empowerment evaluation <i>Agency designs:</i> Learning by doing, Policy dialogue, Collaborative Action Research	Validation by participants that their actions and experienced effects are ‘caused’ by programme Adoption, customisation and commitment to a goal
Synthesis studies	Meta analysis, Narrative synthesis, Realist based synthesis	Accumulation and aggregation within a number of perspectives (statistical, theory based, ethnographic etc.)

Source: Elliot Stern et.al. (2012) *Broadening the Range of Designs and Methods for Impact Evaluation*, Department for International Development Working Paper 38, p. 24.