

# What are Impact Evaluations?

In collaboration with ETHZ, NADEL, Center for Development and Cooperation



The goal of impact evaluations is to **improve aid effectiveness**. An impact evaluation assesses the effects that **can be attributed** to a particular development project or programme.

© ETH NADEL

## What is an impact evaluation?

An impact evaluation analyses the (positive or negative, intended or unintended) impact of a project, programme, or policy on the target population, and quantifies how large that impact is. Impact evaluations establish the **causal effect** of a project, programme or policy on one or several outcome(s). It involves building of counterfactual. For example, an impact evaluation might assess the impact of a development project or programme that aims to improve child health through the construction of public water pumps. For this purpose, it analyses how much lower the incidence of diarrhea is compared to what would have happened without the installation of the water pump. Impact evaluations thus show whether measurable changes in people's lives **can be attributed** to a particular development project or programme. Impact evaluations are not about quantifying the ef-

fect of an project or programme on big overarching development goals – such as climate change reduction or poverty alleviation – that are only indirectly related to a programme. When speaking about impact evaluation, impact is synonymous with the direct causal effect of a project, programme, or policy on an outcome of interest (e.g. reduction in diarrhea).

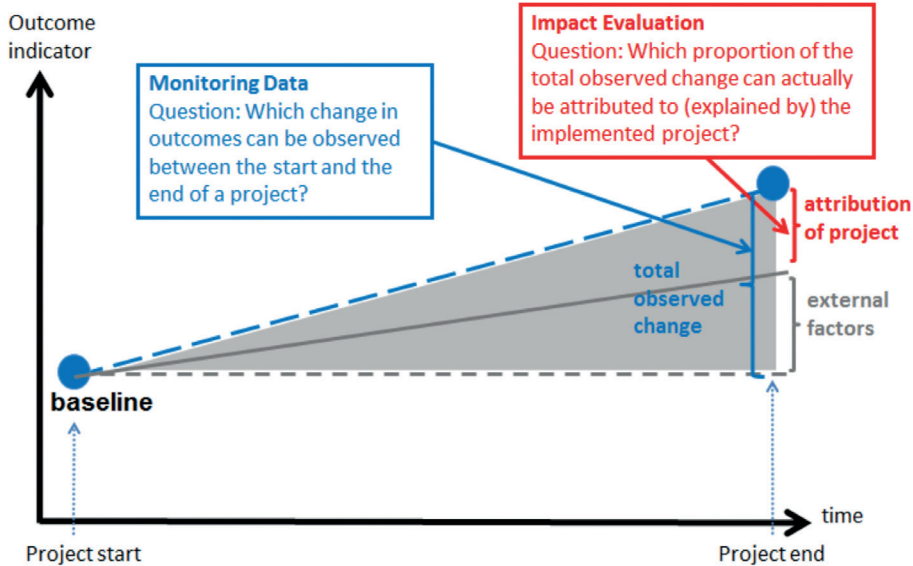
## What is the difference between an impact evaluation and a project evaluation?

While most project or programme evaluations focus on whether implementation of operations or provision of service delivery was successful, impact evaluations focus on whether a development project or programme had a (causal) effect on the lives of people (and why or why not). In above example an evaluation would look at whether the project succeeded in

constructing water pumps and how many people use them, while an impact evaluation aims to assess whether the construction of water pumps led to improved health. Many project and programme evaluations use monitoring data to assess, for example, how many people are using a water pump. Some may even collect monitoring data on whether or not diarrhea has decreased over time. However, such **monitoring data** and before-after comparisons are often not sufficient to establish the causal effect between a development project or programme and its related outcome, because it does not take into consideration external factors.

## Why conduct an impact evaluation?

Impact evaluations contribute to **evidence-based policy making**. Impact evaluations constitute a shift in focus



**Monitoring Data** aims to analyse to what extent there is an observed change of outcomes over time.

**Impact Evaluation** aims to assess the proportion of observed change which can be attributed to the evaluated project or programme.

**from activities to results.** Impact evaluations can help organizations to decide whether **to scale up** projects with proven positive impacts or **to stop** projects or programmes with no impact. Impact evaluations can help **to improve** the design of development project, programmes or policies. Impact evaluations can also be used **to compare** different projects or programmes with regard to their effectiveness. Impact evaluations are thus primarily a tool for **improving** development projects or programmes and **learning**.

### What is the challenge of an impact evaluation?

Ideally one would estimate the effectiveness of a development project or programme by comparing the outcomes for individuals who were part of the development project or programme to **the outcomes for the same individuals if they had not participated in the project or programme**, at the same point in time. This hypothetical outcome is called the **counterfactual**. In the real world, the counterfactual is rarely observable. In the example above where we want to know how much lower the incidence of diarrhea is due to a water pump (in comparison to no water pump) we face the following challenge: while it is possible to measure the incidence of diarrhea in the group of people who received the water pump (treatment group), it is not possible to simultaneously measure the incidence of diarrhea in a state of the world where the same group of people had never received the pump at the same point in time

(counterfactual). Hence a key challenge of impact evaluations is to mimic the counterfactual by finding a group of individuals that is as similar as possible to the individuals being part of a programme, but did not participate in the programme (comparison or control group). The measured outcomes for the comparison group are a proxy for “how participants would have fared without the programme.” **The quality an impact evaluation critically depends on the comparison group.**

### How to define a comparison group?

There are different methods to define a comparison group. Frequently used methods in development policy are the following: Difference in Difference, Statistical Matching, Regression Discontinuity, Randomised Trials (see references for more details), but there are many more. If applied appropriately, these methods can perform well in defining a valid comparison group and thus ensure that the measured impact is not **biased**, i.e. driven by **systematic errors**.

### What do we mean by biased results or systematic errors?

Bias can occur if observed **differences between the individuals who were part of a programme and the comparison group are wrongly attributed to the development project or programme** while these differences can actually be explained by other factors. For

example, a simple comparison of differences in diarrhea for the same group of people **before and after** the installation of water pumps does not always produce reliable impact estimates. Many other factors that are unrelated to the programme itself (e.g. hygiene education, increased income, other improved sanitary infrastructure) could have influenced a change in health outcomes over the same time period. Similarly, a simple comparison of differences in diarrhea between villages **with and without** a water pump might also be problematic. Bias could occur because villages that received a pump might be different (e.g. closer to a main road) from villages which did not receive a pump. Such systematic errors lead to biased (i.e. wrong) impact estimates.

### What about the sample size?

Impact evaluations require a **large enough** sample size to ensure that **we are able to measure an effect of a project, programme or policy**. The number of participants and non-participants should also be large enough to conclude that what we observe is not just a coincidence but in fact the impact of the programme. In the water pump example, we might compare five households that used a water pump with five households that used water from a pond. However, those ten households might not represent pump and pond users in general. Differences in diarrhea between the two groups might not be representative because, by chance, we might have picked sick households from the pump user group and

more healthy households from the pond user group. The **necessary sample size** depends on the **expected impact** of the project or programme we are looking at. The smaller the expected impact of a development project or programme, the larger the sample size required to detect an impact. In addition, sample size also depends on the general **variance of the outcome**: the larger the variance of the outcome indicator, the larger the required sample size. For example, incomes across households usually vary more than the quantity of water consumed. Hence, to analyse the impact of a water pump on households' income would require a larger sample size than to analyse the impact of a water pump on the quantity of water consumed.

### When to conduct an impact evaluation?

To construct a good comparison group, an impact evaluation should be **planned in the conceptual phase, before a project, programme, or policy is implemented** – otherwise, the opportunity to obtain good data and construct a good comparison group may be forfeited. At the same time, the quality of an impact evaluation depends on a clear understanding of the programme. In general, then, an impact evaluation should be planned after operational issues have been addressed, but before the programme has been rolled out.

Impact evaluations demand significant time and resources. Hence, **not all projects, programmes, and policies of an**

**organization** should be evaluated with regard to their impact. Impact evaluations should be only conducted when there is a great **learning potential**. A positive answer to at least one of the following questions is an indication that an project or programme can benefit from an impact evaluation<sup>1</sup>:

- Is there potential to scale-up the project or programme?
- Is the project or programme strategically relevant for the organization for replication or learning purposes?
- Is this an innovative project or programme that is in need of evidence on whether it works?
- Can the results be used to inform national or global policy making?

It is also not possible to analyse the impact of all types of development projects and programmes. In other words, **not all the projects and programmes of an organization** can be evaluated with regard to their impact. Here, three aspects are important:

- Are we reaching enough people with the project or programme so that the sample size of the impact evaluation is sufficiently large?
- Is it possible to establish a good comparison group?
- Are the outcomes (of interest) measurable within a reasonable timeframe?

<sup>1</sup> See also the World Bank website on impact evaluations. <http://www.worldbank.org/en/research/dime>

### Recommended further reading:

#### AFD: Evaluations d'impacts: Quelques concepts et définitions

<http://www.afd.fr/home/recherche/evaluation-capitalisation/Evaluation-impacts/concept-definition>

#### International Initiative for Impact Evaluation:

Summary of existing impact evaluations

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

#### J-PAL: Introduction to Randomized Impact Evaluations

<https://www.povertyactionlab.org/research-resources/introduction-evaluations>

#### World Bank: Impact Evaluation in Practice, Second Edition

<https://openknowledge.worldbank.org/handle/10986/25030>



© SDC