Chapter 9

Front-end planning is important

Front-end planning refers to the initial planning and design phase of an impact evaluation. Ad hoc commissioned impact evaluations usually do not have a long planning period, thereby risking a suboptimally planned and executed evaluation process.

As good impact evaluation relies on good data, preferably including baseline data, attention to proper front-end planning should be a priority issue. Ideally, front-end planning of impact evaluations should be closely articulated to the initial design and planning phase of the policy intervention. Indeed, this articulation is most clearly visible in an RCT, in which intervention and impact evaluation are inextricably linked.

9.1. Planning tools

Clear definition of scope (chapters 1 and 2) and sound methodological design (chapters 3–6) cannot be captured in standardized frameworks. Decision trees on assessing data availability (see § 8.2.) and method choice (see appendix 6) are useful, though they provide only partial answers to methodological design choice issues. Pragmatic considerations of time, budget, and data (see § 8.4) but culture and politics also play a role. Two tools that are particularly helpful in the planning phase of an impact evaluation are the *approach paper* and the *evaluation matrix*.

The approach paper outlines what the evaluation is about and how it will be implemented. This document can be widely circulated and gives stakeholders and others a chance to comment and improve upon the intended evaluation design from an early stage. It also helps to generate broad "buy-in" or at worst to define the main grounds of potential disagreement between evaluators and practitioners. In addition, it is wise to use an evaluation matrix when planning and executing the work. This tool ensures that key questions are identified, together with the ways to address them, sources of data, role of theory, etc. This can also play an important role in stakeholder consultation, ensuring that important elements are not omitted.

9.2. Staffing and resources

Resources are important, and spending should be safeguarded up front. The longer the time horizon of a study, the more difficult this is. Resources are also important to realize the much-needed independence of an evaluator and the evaluation team. A template for assessing the independence of evaluation organizations can be downloaded from http://www.ecgnet.org/docs/ecg.doc. This document specifies a number of criteria and questions that can be asked.

Evaluation is not only a *financial resources business* but even more a *people's business.* So is the *planning of an evaluation*. As evaluation projects are usually no longer "*lonely hunter*" *activities*, *staffing* is crucial. So when starting the preparation of the study, a crucial point concerns addressing a number of questions:

- Who are the people who do the evaluation?
- Under which (contractual) conditions are they doing the job?
- What is their expertise?
- Which roles will they be carrying out?

Topics that deserve attention are the following:

- The *mix of disciplines and traditions* that are brought together in the team.
- The *competencies* the team has "in stock." Competencies range from methodological expertise to negotiating with institutional actors and stakeholders, getting involved in "hearing both sides" (those evaluated and the principal) and in the clearance of the report.
- The *structure of the evaluation team.* For the evaluation to be planned and carried out effectively, the roles of the project director, staff, and other evaluators must be made clear to all parties.
- The *responsibilities* of the team members.
- The more an evaluation is linked to a political "hot spot," the more it is necessary that at least one member of the team have a "political nose"—not primarily to deal with administrators and (local) politicians, but to understand when an evaluation project becomes too much of what is known as a *partnerial evaluation* (Pollitt, 1999).
- Also, staff should be active in realizing an *adequate documentation and evaluation trail*.

A range of skills is needed in evaluation work. The quality and eventual utility of the impact evaluation can be greatly enhanced with coordination between team members and policymakers from the outset. It is therefore important to identify team members as early as possible, agree on roles and responsibilities, and establish mechanisms for communication during key points of the evaluation.

9.3. The balance between independence and collaboration between evaluators and stakeholders

One of the questions within the world of impact evaluations is what degree of institutional separation to put between the evaluation providers and the evaluation users. There is much to be gained from the objectivity provided by having the evaluation carried out independently of the institution responsible for the project being evaluated. Pollitt (1999) warned against "partnerial" evaluations, where positions of stakeholders, commissioners, and evaluators blurred too much.[1] However, evaluations often have multiple goals, including building evaluation capacity within government agencies and sensitizing program operators to the realities of their projects once they are carried out in the field. At a minimum, the evaluation users, who can range from government agencies in client countries to bilateral and multilateral donors, international NGOs, and grass roots/civil society organizations, must remain sufficiently involved in the evaluation to ensure that the evaluation process is recognized as legitimate and that the results produced are relevant to their information needs. Otherwise, the evaluation results are less likely to be used to inform policy. The evaluation manager and his or her clients must achieve the right balance between involving the users of evaluations and maintaining the objectivity and legitimacy of the results (Baker, 2000).

9.4. Ethical issues

It is important to take the ethical objections and political sensitivities seriously. There can be ethical concerns with deliberately denying a program to those who need it and providing the program to those who do not; this applies to both experimental and non-experimental methods. For example, with too few resources, randomization may be seen as a fair solution, possibly after conditioning on observables. However, the information available to the evaluator (for conditioning) is typically a partial subset of the information available "on the ground" (including voters/taxpayers). The idea of "intention-to-treat" helps alleviate these concerns; one has a randomized assignment, but anyone is free to not participate. Even then, the "randomized out" group may include people in great need. All these issues must be discussed openly and weighed against the (potentially large) longer-term welfare gains from better information for public decision making (Ravallion, 2008).[2]

9.5. Norms and standards

As noted before, impact evaluations are often designed, implemented, analyzed, disseminated, and used under budget, time, and data constraints while facing diverse and often competing political interests. Given these constraints, the management of a real-world evaluation is much more complicated than textbook descriptions.

Evaluations sometimes fail because the stakeholders were not involved, or the findings were not used because they did not address the stakeholders' priorities. Others fail because of administrative or political difficulties in getting access to the required data, being able to meet with all the individuals and groups that should be interviewed, or being able to ask all the questions that the evaluator feels are necessary. Many other evaluations fail because the sampling frame, often based on existing administrative data, omits important sectors of the target population—often without anyone being aware of this. In other cases the budget was insufficient, or was too unpredictable to permit an adequate evaluation to be conducted. Needless to say, evaluations also fail because of emphasizing stakeholders´ participation too much, leading to partnerial evaluations (Pollitt, 1999), and because of insufficient methodological and theoretical expertise.

Although many of these constraints are presented in the final evaluation report as being completely beyond the control of the evaluator, in fact their effects could very probably have been reduced by more effective management of the evaluation. For example, a more thorough scoping analysis could have revealed many of these problems, and the client(s) could then have been made aware of the likely limitations on the methodological rigor of the findings. The client(s) and evaluator could then strategize to either seek ways to increase the budget or extend the time, or agree to limit the scope of the evaluation and what it promises to deliver. If clients understand that the current design will not hold up under the scrutiny of critics, they can find ways to help address some of the constraints: "We have found that impact evaluations generally provide rudimentary documentation of the data being used. There is evidently a trade-off between decision makers' and bureaucrats' appeal for short and crisp reports and principles for scientific documentation, but we want to emphasise that displaying descriptive statistics improves the transparency of the methodological approach" (Jerve and Villanger, 2008: 34).

For the sake of honest commitment to development, evaluators and evaluation units should ensure that impact evaluations are designed and executed in a manner that limits manipulation of processes or results that lean toward any ideological or political agenda. They should also ensure that there are realistic expectations of what can be achieved by a single evaluation within existing time and resource constraints, and that findings from the evaluation are presented in ways that are accessible to the intended users. This includes finding a balance between simple, clear messages and properly acknowledging the complexities and limitations of the findings.

International evaluation standards (such as the OECD-DAC or the United Nations Evaluation Group Norms and Standards and/or the standards and guidelines developed by national or regional evaluation associations) should be applied where appropriate (Picciotto, 2004).

Greater emphasis on impact evaluation for evidence-based policy making can create greater risk of manipulation aimed at producing desirable results (House, 2008). Impact evaluations require an honest search for the truth and thus place high demands on the integrity of those commissioning and conducting them. For the sake of honest commitment to development, evaluators and evaluation units need to ensure that impact evaluations are designed and executed in a manner that limits manipulation of processes or that produces results favoring any ideological or political agenda.

9.6. Ownership and capacity building

Capacity building at the level of governmental or non-governmental agencies involved should be an explicit purpose in impact evaluation. In cases where sector-wide investment programs are financed by multidonor co-financing schemes, participating donors would make natural partners for a joint evaluation (OECD-DAC, 2000).

Other factors in selecting other donors as partners in a joint evaluation work may also be relevant. Selecting donors with similar development philosophies, cultures, evaluation procedures and techniques, and regional affiliations, and that are geographically close may make working together easier. Another issue may be keeping the total number of donors "manageable." Where more donors are involved, a key group of development partners (including national actors) could assume management responsibilities and the role of others can be more limited. Once appropriate donors that have a likely stake in an evaluation topic are identified, the next step is to contact them and discern whether they are interested in participating. In some cases, an appropriate consortium or group may already exist, where the issue of a joint evaluation can be raised and expressions of interest easily solicited. The DAC Working Party on Aid Evaluation, the United Nations Evaluation Group, and the Evaluation Cooperation Group have a tradition of cooperation, shared vision on evaluation, and longstanding relationships and have fostered numerous joint evaluations.

The interaction among the international development evaluation community, the countries/regions themselves, and the academic evaluation communities should also be stimulated, as it is likely to affect the pace and quality of capacity building in impact evaluation. Capacity building will also strengthen (country and regional) *ownership* of impact evaluation. Providing a space for consultation and agreement on impact evaluation priorities among the different stakeholders of an intervention will also help enhance utilization and ownership.

Key message

Front-end planning is important. It can help manage the study, its reception, and its use. When managing the evaluation, keep a clear eye on items such as costs, staffing, ethical issues, and level of independence—of the evaluator and the team, versus the level of collaboration with stakeholders. Pay attention to country and regional ownership of impact evaluation and capacity building and promote it. Providing a space for consultation and agreement on impact evaluation priorities among the different stakeholders of an intervention will help enhance utilization and ownership