

METHODS LAB

August 2015

EVALUABILITY ASSESSMENT FOR IMPACT EVALUATION

GUIDANCE, CHECKLISTS AND
DECISION SUPPORT

Greet Peersman, Irene Guijt & Tiina Pasanen

The Methods Lab is an action-learning collaboration between the Overseas Development Institute (ODI), BetterEvaluation (BE) and the Australian Department of Foreign Affairs and Trade (DFAT). The Methods Lab seeks to develop, test, and institutionalise flexible approaches to impact evaluations. It focuses on interventions which are harder to evaluate because of their diversity and complexity or where traditional impact evaluation approaches may not be feasible or appropriate, with the broader aim of identifying lessons with wider application potential.

Readers are encouraged to reproduce Methods Lab material for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI, the Australian Department of Foreign Affairs and Trade (DFAT) and BetterEvaluation.

© Overseas Development Institute 2015. This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 3.0).

How to cite this guidance note:

Peersman, G., Guijt, I., and Pasanen, T. (2015)
'Evaluability Assessment for Impact Evaluation'.
A Methods Lab publication. London: Overseas Development Institute.



BetterEvaluation

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
Tel +44 (0) 20 7922 0300
Fax +44 (0) 20 7922 0399
info@odi.org.uk
www.odis.org

BetterEvaluation
E-mail: bettereval@gmail.com
www.betterevaluation.org



Focus

This guidance note focuses on the utility of, and guidance for, evaluability assessment before undertaking an impact evaluation.

Intended users

The primary audience for this guidance note is evaluators conducting an evaluability assessment for impact evaluation. The secondary audience is people commissioning or managing an evaluability assessment for impact evaluation, as well as funders of an impact evaluation.

How to use it

Sections one and two provide an overview of evaluability assessment and how it can be used for impact evaluation.

Section three provides guidance for planning to undertake an evaluability assessment for impact evaluation. This is informative for all intended users of the guidance note.

Section four includes checklists and decision support for evaluability assessments. The *checklist* is geared to those conducting the evaluability assessment and can be adapted to suit a particular context or purpose. The *decision support* provides those conducting an evaluability assessment with evidence-based recommendations for impact evaluation funders and commissioners, about whether, when and how to proceed with the evaluation.

Sections five and six provide guidance on what to do after the assessment is concluded, and offer lessons learned from evaluability assessments in practice.

Acknowledgements

This guidance note is a Methods Lab publication, written by Greet Peersman (BetterEvaluation), Irene Guijt (Overseas Development Institute) and Tiina Pasanen (Overseas Development Institute). Peer reviewers include: Anne Buffardi (Overseas Development Institute), Rick Davies (Independent Consultant), Simon Hearn (Overseas Development Institute), Bron McDonald (Independent Consultant), and Patricia Rogers (BetterEvaluation).

The authors would like to acknowledge the work of Rick Davies and Elizabeth Dunn, whose papers this guidance note draws from:

- Rick Davies (2013), 'Planning evaluability assessments - A synthesis of the literature with recommendations', published by the Department For International Development (DFID) (available at: <http://bit.ly/1eFbd4u>)
- Elizabeth Dunn (2008), 'Planning for cost effective evaluation with evaluability assessment', published by USAID (available at: http://pdf.usaid.gov/pdf_docs/PNADN200.pdf)

In addition, the guidance note draws on the practical experiences of evaluability assessments conducted by the Methods Lab in Afghanistan, Nepal, and the Pacific. The authors would like to acknowledge the efforts of the participants, funders and evaluators who led assessments (Anne Buffardi, Overseas Development Institute; Maren Duvendack, University of East Anglia; and Bron McDonald, Independent Consultant). It also draws on the experiences of evaluability assessments supported by ITAD (UK) and the Department of Planning, Monitoring and Evaluation, in South Africa.

Acronyms

BE - BetterEvaluation

CBO – Community Based Organisation

DFAT - Australian Department of Foreign Affairs and Trade

DFID – UK Department for International Development

KEQ – Key Evaluation Questions

M&E – Monitoring and Evaluation

NGO – Non-Governmental Organisation

OECD-DAC – Organisation for Economic Co-operation and Development – Development Assistance Committee

ODI – Overseas Development Institute

ToC – Theory of Change

USAID – United States Agency for International Development

Contents

| | |
|---|----|
| 1. Using evaluability assessments | 4 |
| What is evaluability and why assess it? | 4 |
| How can an evaluability assessment be used? | 4 |
| What should an evaluability assessment cover? | 5 |
| 2. How can an evaluability assessment support impact evaluation? | 6 |
| What is particular about impact evaluation? | 6 |
| How can an evaluability assessment support decision-making? | 6 |
| 3. Getting started with an evaluability assessment for impact evaluation | 9 |
| Using an internal or external team | 9 |
| Clarifying the purpose and developing a checklist | 9 |
| Determining appropriate data collection methods | 9 |
| Being transparent about judgements | 10 |
| Being realistic about required resources | 11 |
| 4. Evaluability assessment checklist & decision support for impact evaluation | 12 |
| 5. What happens after the assessment? | 17 |
| 6. Lessons learned from practice | 18 |
| 7. Additional resources | 24 |

1. Using evaluability assessments

What is evaluability and why assess it?

The OECD-DAC defines evaluability as ‘the extent to which an activity or project can be evaluated in a reliable and credible fashion’. It implies the use of a robust evaluation design to reduce the risk of irrelevant or invalid findings.

Provided there is intent to evaluate an intervention¹, assessing its evaluability can usually be done for a small cost of the total evaluation budget. This is particularly relevant when done in relation to an impact assessment, and can prevent wasting valuable time and resources on a premature or inappropriately designed evaluation. As Dunn (2008) states: ‘This investment will more than pay for itself by leading to impact assessment(s) that are valid, efficient, and useful’. Hence, an evaluability assessment is an additional tool in the evaluation management toolbox. However, it provides good value for money *only* if its findings are able to influence if, when and how subsequent evaluation gets done - which is a critical assumption.

Assessing evaluability is different from doing an evaluation. An evaluation aims to judge the merits of a particular intervention, whereas an evaluability assessment occurs *before* an evaluation, with the intent to recommend whether or when the evaluation should take place and how to maximise its value. In organisations with good intervention planning, stakeholder management and a sound monitoring and evaluation (M&E) approach, there may be no need for an evaluability assessment.

For a more in-depth discussion of uses of evaluability assessment and detailed case studies, we refer readers to a recent book by Trevisan and Walser (2015)² entitled *Evaluability assessment - Improving evaluation quality and use*. Trevisan and Walser revisit the history of the theory and practice of evaluability assessments since the 1970s. The authors argue that evaluability assessment practice now supports a much wider variety of programmes, stakeholder needs and purposes than originally conceptualised and used, including: programme development or modification; fostering an understanding of programme culture and context; developing interest in evaluation; building evaluation capacity; and serving as a precursor to a variety of evaluation approaches.

How can an evaluability assessment be used?

Evaluability can be assessed at different stages of the intervention cycle. The timing of an evaluability assessment affects how it is used, as demonstrated below (see also Table 1).

- **Early on in the intervention:** Problems of evaluability are often due to weak intervention logic and a lack of the right kind of, or good quality, data (Davies, 2013). At the intervention design stage, an evaluability assessment can be used to make explicit assumptions about the intervention, identify evidence gaps, assist the design process, and inform the proposed M&E system. Given that an intervention and context is likely to change over time, if an evaluability assessment is done early on in the intervention, it may have to be repeated at a later stage. This is to ensure that the most up-to-date information is available (especially if the assessment aims to inform an evaluation at the end of an intervention’s implementation).
- **At the end of the intervention:** Assessing evaluability at this stage will give the most up-to-date information on the degree of ‘difficulty’ for carrying out an evaluation. However, the scope for corrective action is inevitably more limited. At this point, an assessment is particularly useful to pinpoint the most significant challenges that must be addressed in order to make an intended evaluation worthwhile. This may delay the evaluation and/or put restrictions on how the evaluation takes place and its findings are used. Sometimes, it is decided not to proceed with the evaluation. However, given the range of available evaluation methods, one can usually find a feasible way forward, though it may require re-negotiating how the findings are to be used.
- **During intervention implementation:** Undertaking an evaluability assessment at any other stage in the intervention cycle (i.e., in between design and just prior to the end of the intervention) provides an opportunity to accommodate emerging knowledge or new organisational imperatives, to modify the intervention logic and adjust intervention delivery and data collection, where needed.

1 An ‘intervention’ may be a service, project, programme, portfolio of activities or projects, legislation, policy, strategy, partnership etc.

2 Trevisan M. and Walser T. (2015). Evaluability assessment. Improving evaluation quality and use. Thousand Oaks: Sage Publications, Inc.

What should an evaluability assessment cover?

An evaluability assessment should address three focus areas³:

1. Adequacy of intervention design for what it is trying to achieve.
2. Conduciveness of the institutional context to support an appropriate evaluation.
3. Availability and quality of information to be used in the evaluation.

How much attention you devote to each focus area will depend on the timing of the assessment and is closely linked to the assessment purpose (see Table 1). For example, it makes sense to focus more heavily on intervention design early on, when there is more scope for influencing it. Whereas, focusing on the conduciveness to evaluate becomes more important immediately prior to doing an evaluation.

Specific criteria should be defined to operationalise the three focus areas. These can be compiled in a ‘checklist’ to encourage comprehensive coverage of all relevant issues, or at least, make transparent which issues were not covered (see section 4).

Table 1: Overview of the utility of an evaluability assessment

| Purpose | Focus | Timing | Outputs |
|--|---|--|---|
| What is the main reason for doing the evaluability assessment? | Where should most attention be focused? | At what stage should the evaluability assessment be done? | What outputs can be expected? |
| To improve intervention design | Adequacy of the intervention design | Early on in the intervention (ideally, prior to approval of the intervention design) | Proposed refinements to intervention logic |
| To inform the design of the M&E system | Availability and quality of information | Early on in the intervention (ideally, prior to approval of the M&E system design) | Proposed improvement of the M&E system |
| To decide if a planned evaluation should take place now or later | Availability and quality of information Conduciveness of the institutional context | Just prior to evaluation | Level of difficulty in conducting evaluation now Options for evaluation timing |
| To inform the design of a planned evaluation | Availability and quality of information Conduciveness of the institutional context | Just prior to evaluation | Options for evaluation questions, methods, resources and expertise |

³ Based on the guidance and practices of key international agencies (see Davies 2013).

2. How can an evaluability assessment support impact evaluation?

What is particular about impact evaluation?

'Impact refers to positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.'⁴ This OECD-DAC definition of impact makes clear that an impact evaluation must establish the cause of observed changes. Therefore, an impact evaluation not only measures or describes changes that have occurred, but also seeks to understand the role of a particular intervention in producing these changes. In other words, impact evaluation seeks causal attribution.

Causal attribution is defined by the OECD-DAC as: 'Ascription of a causal link between observed (or expected to be observed) changes and a specific intervention.'⁵ This does not require that changes are produced solely or wholly by the intervention under investigation,⁶ but takes into consideration other causes and contextual factors.

How can an evaluability assessment support decision-making?

Given the specific characteristics of impact evaluation, there are three broad questions⁷ an evaluability assessment should address *before* undertaking the impact evaluation:

1. Is it plausible to expect impact?
2. Would an impact evaluation be useful and used?
3. Is it feasible to assess or measure impact?

These questions respectively address each of the three focal areas discussed in **Section 1**: adequacy of intervention design; conduciveness of the organisational context; and availability and quality of information.

⁴ Organisation for Economic Co-operation and Development – Development Assistance Committee, *Glossary of Key Terms in Evaluation and Results Based Management*, OECD-DAC, Paris, 2010. See <http://www.oecd.org/development/peer-reviews/2754804.pdf>

⁵ Ibid.

⁶ United Nations Evaluation Group, *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*, Guidance Document, UNEG, New York, 2013. See http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=1434

⁷ This section is based on Dunn, 2008, Planning for cost effective evaluation with evaluability assessment. Impact assessment primer series. Publication #6. Washington DC: USAID. http://pdf.usaid.gov/pdf_docs/PNADN200.pdf

1. Plausibility: is it plausible to expect impact?

This question is about the adequacy of the intervention logic.

This includes its clarity, relevance, coherence, and feasibility.

Why is it important to assess plausibility for impact evaluation?

Intervention activities should *reasonably* be expected to lead to the intended outcomes⁸ and impact (i.e. there is a *plausible* relationship). Verifying the logic of this on the basis of the intervention design may reveal the need to modify the intervention and/or revisit the expectations about anticipated outcomes and/or impacts.

What decisions can be informed by assessment of plausibility before proceeding with the impact evaluation?

Assessing plausibility can help to inform what elements of the intervention can be evaluated at different points in time. Depending on the extent to which gaps in intervention logic can be addressed, investing in impact evaluation may not be warranted. Instead, it may be better to focus on re-thinking and re-directing the intervention to increase its potential for effectiveness, and assessing its implementation.

2. Utility: would an impact evaluation be useful and used?

This question is about stakeholders' needs and expectations around an evaluation (also referred to as 'utility of an evaluation') and whether the impact evaluation can be designed to meet these.

Why is it important to assess utility for impact evaluation?

An impact evaluation should only be undertaken when its intended use and users can be clearly identified, and when it is likely to produce useful findings. The availability of resources and the timing of decisions about the intervention may affect the utility of the evaluation. These decisions could include continuing implementation, modifying or replicating implementation, or discontinuing implementation.

In addition, different stakeholders often have different information needs, and a single evaluation may be designed to satisfy multiple needs. However, to manage expectations, it is crucial to clarify who needs what information, when, and for what use(s).

Assessing stakeholder expectations about what 'evidence' is seen as credible is equally important. For example, some stakeholders will value participatory evaluation approaches, while others may place higher credibility on specific evaluation designs. Stakeholders' needs and expectations will affect the timing of the evaluation, the type of data to be collected, the way in which data are obtained and analysed, and the strategies and channels used to present and share the findings with intended users.

If key evaluation questions have already been identified at some stage during the intervention cycle, it is necessary to assess their appropriateness for an impact evaluation. This includes the value added of the information delivered to different stakeholders. These questions may also be based on underlying assumptions, which are essential to identify and make explicit. This may include, for example, whose information needs are being prioritised, expectations about a specific evaluation design or other issues which may affect the utility (and often also the feasibility) of the planned evaluation.

What decisions can be informed by assessing utility before proceeding with the impact evaluation?

It is a waste of time and resources to conduct an impact evaluation of which the findings are not likely to be used. The findings of an impact evaluation may not be used if there is: insufficient demand; limited scope for using findings in decision-making; lack of consensus on key evaluation questions; or lack of potential for credible evidence due to inadequate resources or poor timing of the evaluation.

While use of evaluation findings cannot be guaranteed, it is possible to assess whether there is sufficient demand and whether stakeholders have realistic expectations. In the absence of either, it might be better not to proceed with the evaluation, or at least seek to make these limitations explicit and address them before proceeding.

⁸ Outcomes are defined as the intermediate-term results that are intended to lead to the desired impact.

3. Feasibility: is it feasible to measure impact?

This question asks about the availability and quality of data and the systems for making data available. It also covers the resources and logistics that may affect the feasibility of an evaluation, or of specific evaluation designs.

Why is it important to assess feasibility for impact evaluation?

The feasibility of an impact evaluation is influenced by several factors. The following factors need to be understood in order to make a decision whether to proceed now or later:

Feasibility of measuring what is worth measuring: Not all data are easy to collect, and data that can be more easily collected may not be particularly relevant or appropriate to understand causal pathways. Similarly, the timing of the evaluation is crucial in determining what is worth measuring. An evaluation may be undertaken too late to inform important decisions. However, when undertaken too early, it may lead to inaccuracies. For example understated impact, when there has not been sufficient time for impact to emerge, or overstated impact, when longer lasting impact is yet to be determined.

Availability of good quality data: Ideally, data already exists that can be used in the impact evaluation. This includes data that is routinely collected and reported as part of the performance monitoring system, and data from process evaluations which assesses intervention implementation. Additional data collection can then focus on addressing important data gaps, and would often include gathering data to rule out alternative explanations for the observed changes in order to determine causal attribution. For data to be useful for impact evaluation, it needs to be reliable and valid, as well as readily available.

Characteristics of the intervention: The characteristics of the intervention (such as how it is rolled out over time, its location, levels of client intake and the programme reach) may affect evaluation design options. Design options include the sampling approach and sample size, baseline data needs, options for a control comparison group (where appropriate) or other strategies to investigate causal attribution.

The budget size: For impact evaluation, the budget plays an important role in influencing which designs are possible. For example, the required resources could increase substantially if data is gathered at the beginning and end of an intervention, or both participants and non-participants are included.

What decisions can be informed by assessing feasibility before proceeding with the impact evaluation?

If there is insufficient capacity, resources or time to collect and collate good quality data, this can lead to biased or incomplete data that may lead to misinformed decision making about the intervention. In this situation, it is not worthwhile proceeding. However, an evaluability assessment will usually be able to identify several feasible design options. Presenting these to all relevant stakeholders, including implementing staff, can help verify that all relevant issues (including logistics) have been considered, and that there are no major objections to any of the proposed options.

3. Getting started with an evaluability assessment for impact evaluation

Using an internal or external team

Evaluability assessments can be carried out by staff from the implementing organisation, or can be contracted out to external evaluators. Evaluation expertise is essential to address methodological issues about data and potential evaluation design. Subject matter and context expertise is particularly useful in assessing ‘plausibility’ and ‘utility’ issues. Hence, for ‘internal’ assessments, it is crucial for programme managers and other programme staff to work together with the internal evaluator or evaluation team. When the assessment is contracted out, the team should ideally include both evaluation and subject matter expertise.

For external teams, most of the parameters for the assessment will already be defined in the Terms of Reference developed by the commissioner of the study. Internal teams may have more flexibility to shape the assessment.

Clarifying the purpose and developing a checklist

An important first step is to clarify the purpose of the assessment, including the specific expectations of funders, managers and/or other relevant stakeholders who will be using the assessment results.

As stated in Section 2, an evaluability assessment should address three broad questions: (1) is it plausible to expect impact? (2) would an impact evaluation be useful and used? (3) is it feasible to expect impact? It is advisable to compile a checklist of criteria for each focus area, to encourage comprehensive coverage of all relevant issues or, at least, be clear about which issues are not covered. Section 4 provides an extensive checklist. However, this is not intended to be mandatory; the specific situation in which an evaluability assessment for impact evaluation takes place will determine specific emphasis and require tailoring of the checklist. One checklist will not fit every situation.

Determining appropriate data collection methods

In order to determine the plausibility, utility and feasibility of an impact evaluation, an evaluability assessment typically includes document review and key informant interviews. It may also include site visits, observations and workshops.

The following are recommended data collection methods to address each of the key questions in an evaluability assessment for impact evaluation:

1. Plausibility: is it plausible to expect impact?

Ideally, an explicit intervention logic or theory of change has been documented as part of the intervention design. The evaluability assessment includes a verification of its comprehensiveness and plausibility, through review of programme documents and in-depth interviews with programme managers and staff, funders, beneficiaries and other relevant stakeholders. Triangulation of information and on-site observations may be needed, as there may be divergent perspectives, or the reality on the ground may not necessarily reflect the ‘theory’. When there is no explicit theory of change or it is incomplete, it is often possible to articulate or further clarify this through the same methods. For complex interventions, it is important to articulate the various components, how they interact (such as to affect change at the individual, group, organisational and/or policy level), and how they are expected to achieve the desired change.

2. Utility: would an impact evaluation be useful and used?

Different stakeholders’ needs and expectations can often be identified from a document review. However, this will most likely need to be complemented by in-depth interviews and/or discussions with managers and implementers, funders, beneficiaries and other relevant stakeholders. It may also be useful to bring stakeholders together to discuss and prioritise specific needs and uses for the evaluation. Securing agreement on what is useful and credible often requires skilful facilitated discussions with the relevant stakeholders. It may be useful to consider a workshop during which relevant stakeholders agree on the main purpose(s) of the planned impact evaluation.

3. Feasibility: is it feasible to measure impact?

The likelihood of available existing data sources can be assessed through a review of the intervention's monitoring, evaluation and research plans; the data collection tools and reporting templates used for performance monitoring; any available documentation or formal progress reports; completed programme reviews, process or other evaluations; and research studies.

In addition, a site visit may be required to assess the extent to which raw data is indeed available, and to carry out spot checks or more in-depth assessments of data quality. It may involve interviewing key informants to better understand the intervention's characteristics that may help or hinder different impact evaluation approaches, as well as pertinent logistical and ethical issues, risks, and information about other interventions operating in the area.

Where evaluability assessment questions ask about specific information, you should first draw on a review of available documents. Information can then be added and verified through interviews and/or discussions with key informants. **Box 1** provides an overview of useful documents to review. However, in most situations, the process is more iterative than linear; document review will lead to gap filling with stakeholders, but consultation with stakeholders will also lead to further relevant documents (Davies, 2013).

Being transparent about judgements

Some evaluability assessments include a rating for each of the items on the checklist. This is useful to show how robust existing data are, or the extent to which the specific uses of the impact evaluation are clearly defined. Ideally, these ratings should be accompanied by explicit weightings, as not all items assessed have equal importance for decision-making about the way forward. Also, a total, aggregated, score may be useful to those commissioning the evaluation as an index of difficulty (and, by extension, the level of effort/resources needed) for conducting the planned impact evaluation. Even if scoring and/or weighting are not used, sufficient evidence should be provided to support the judgements that have been made (see also Section 5).

Being realistic about required resources

The resources available for the assessment, including budget and time, will affect what can be done. It is important to be transparent about any limitations as they may affect the results of the assessment and how they can be used.

The duration of an evaluability assessment can range considerably. As a general guideline, Davies (2013) suggests a minimum of five days for a desk-based review (off-site) and about two weeks for on-site consultations. Among 29 examples of evaluability assessments identified in a literature review conducted by Davies (2013), the duration of the assessments ranged from two days to four months, with one week being the most common (not all were impact evaluation-focused). The quickest were typically desk-based, reviewing readily available documents. However, most evaluability assessments will require interviews or discussions with key informants in addition to site visits. The duration also depends on the type of intervention to be assessed and the number and expertise of the evaluators conducting the assessment. Portfolio and multiple project assessment take longer than smaller scale interventions.

Negotiating an appropriate number of days and people, with the assessment commissioners is important. However, often a ceiling budget has already been decided upon. When this is the case, it is essential to be clear about what can and cannot be done.

Table 2 outlines the Method Lab's experience with conducting evaluability assessments for impact evaluation. It describes the type of intervention, the purpose, main methods used and the number of people/days.

Box 1: Useful types of information to review in an evaluability assessment for impact evaluation

- Intervention proposal and budget
- Intervention design document
- Performance monitoring data collection tools, reporting templates and actual data
- Progress reports
- Evaluation reports (such as a mid-term review or a process evaluation)
- Research papers
- Broader literature on relevant themes

Table 2: Level of effort in Methods Lab evaluability assessment

| Type of intervention | Purpose | Main methods used | Number of people & days used |
|---|--|--|------------------------------|
| Portfolio of 13 projects working in the same broad sector in 9 countries | <ul style="list-style-type: none"> Investigating the feasibility of an impact evaluation Providing potential design options for impact evaluation | <ul style="list-style-type: none"> Review of 60 documents 13 interviews 1 stakeholder workshop Follow-up telephone calls | 2 people 53 days |
| Multi-component, multi-department project implemented across the whole country | <ul style="list-style-type: none"> Informing an impact-oriented M&E system Providing potential design options for impact evaluation | <ul style="list-style-type: none"> Document review 2 country visits including several meetings 1 stakeholder workshop | 3 people 64 days |
| Multi-component, multi-site programme implemented by 5 consortia | <ul style="list-style-type: none"> Investigating the feasibility of an impact evaluation Informing a programme-wide impact-oriented M&E system Providing potential design options for impact evaluation | <ul style="list-style-type: none"> Review of 16 documents 20 interviews 2 stakeholder workshops (1 on site; 1 virtual) | 2 people 53 days |

4. Evaluability assessment checklist & decision support for impact evaluation

Using the checklist

The following checklist is specifically geared to those conducting an evaluability assessment. The questions are intended to help operationalise each of the three focus areas for evaluability assessment before an impact evaluation: (1) plausibility (2) utility (3) feasibility. Using a checklist encourages comprehensive coverage of all relevant issues or, at least, makes transparent which issues are and are not covered. The checklist provided here is extensive, but is not mandatory, and the specific situation in which an evaluability assessment for impact evaluation takes place will determine the main emphasis and tailoring of the checklist.

Using the decision support

The decision support helps those conducting an evaluability assessment to provide evidence-based recommendations to impact evaluation funders and commissioners. These recommendations are on whether, when and how to proceed with the intended impact evaluation.

1. Is it plausible to expect impact?

Do stakeholders share an understanding of how the intervention operates? Are there logical links between activities and intended impact?

Checklist on adequacy of intervention design

-
- | | |
|--------------------------------------|---|
| A. Investment decision | <ol style="list-style-type: none">1. Is there a clear statement of the problem that the intervention aims to address?2. Is there a clear rationale for undertaking the intervention activities? Is there a clear rationale for activities implementers are not supposed to undertake?3. Is information available on how the intervention fits with existing strategies (e.g., from government or aid agencies)?4. Is it likely that goals could be achieved, given the planned activities and intervention lifespan? Is there evidence from elsewhere to support or negate this?5. Does the intervention design address funder expectations (activities and results)? |
| B. Beneficiaries | <ol style="list-style-type: none">1. Are the goals and intended impacts clearly defined and argued as relevant to the needs of the target group(s)? How were needs of the target group(s) identified? To what extent were intended beneficiaries consulted?2. Are intended beneficiaries (including sub-groups) clearly identified and correctly targeted? Is potential mis-targeting taken into consideration?3. Is there information on who is excluded or who might experience negative results? |
| C. Theory of Change (ToC) | <ol style="list-style-type: none">1. Is there a theory of change (or can it be derived)? Are inputs, activities, reach, outputs, outcomes, impacts (specified or expected to be emergent) and proposed causal linkages between them clearly indicated?2. What is the extent of agreement among key stakeholders on the ToC?3. Are there multiple interactions between different intervention components? Are they defined in terms of their relative contribution to the intended results?4. Are valid ways of assessing inputs, outputs, outcomes and impacts proposed?5. Is it possible to identify which linkages in the ToC are least understood or will be most critical to the intervention success, and thus may inform the focus of the impact evaluation?6. Are assumptions about the roles of actors not directly involved in the intervention explicit (the enablers or constrainters)? Are there feasible plans to assess these? |
-

Decision support

Proceed with impact evaluation

No major barriers exist

- The intervention activities can reasonably be expected to lead to the intended outcomes and impacts (i.e., there are plausible causal links).

Proceed but address critical issues first

Impact evaluation is assumed feasible in the near future

- There are some gaps or inconsistencies in investment decision, intervention logic and/or beneficiary analysis, but it is reasonable to expect that those can be addressed or redefined. Intervention logic can be modified, at a defined point in time, in terms of what actually happened or is happening, by making clearer linkages and/or revisiting the expectations regarding anticipated outcomes and/or impacts.
- Outcomes and impacts need to be described in detail, and descriptions of what is considered 'success' may be useful. However, evaluation is still possible where outcomes or impacts have not been clearly identified in the design e.g., evaluation can address emergent outcomes/impacts, but this needs to be clearly addressed in the theory of change.
- Not having clearly defined goals does not rule out evaluation e.g., it may be possible and appropriate to use a 'goal-free evaluation'⁹ approach.

Do not proceed with impact evaluation

Critical barriers cannot be addressed easily or in a timely manner

- If important gaps about intervention design cannot be addressed easily or in a timely manner, it may be more useful to focus efforts on re-thinking and re-directing the intervention to increase its potential for effectiveness and assessing its implementation.
- There is a very low plausibility of observing intended impact within the timeframe being studied.

⁹ Goal-free evaluation examines the worth of a programme by investigating what it is doing in terms of activities and outcomes, rather than what it is trying to do based on programme goals. It is, generally, focused on the extent to which and the ways in which a programme meets the needs of its target clients; thus, it relies heavily on robust needs assessment in order to judge the programme.

2. Would an impact evaluation be useful and used?

Are there specific needs that the impact assessment will satisfy, and can it be designed to meet needs and expectations?

Checklist on stakeholders, demand and evaluation purpose

- | | |
|---|---|
| A. Stakeholders | <ol style="list-style-type: none">1. Can you identify who initiated the impact evaluation process and why?2. Who are the key stakeholders in the impact evaluation? What are their roles and responsibilities?3. Are key stakeholders motivated to get involved in the evaluation process, or is there any hesitation or active resistance?4. Do key stakeholders understand the need to invest time and resources to conduct a useful and credible impact evaluation?5. Have previous evaluation experiences influenced any stakeholders' likely interest and/or participation in the process? |
| B. Purpose and use | <ol style="list-style-type: none">1. What is the main purpose of the impact evaluation? Are there significant differences in perspectives or needs between stakeholders? If so, are they likely to have a negative influence on the evaluation process? If negative, are they reconcilable? Can the impact evaluation accommodate different information needs or is prioritising needed?2. What opportunities exist for an impact evaluation to influence decision-making? Who will it influence and for what decisions?3. How do different stakeholders anticipate the results will be used (i.e., what are the primary users or key audiences)? Do their views differ? Is the desired use realistic?4. Do stakeholders expect to be informed of the evaluation findings at the end or will they be involved in different stages of the evaluation process? If so, are levels of (non) involvement clear and acceptable to the different stakeholder groups?5. Are there mechanisms or strategies in place to strengthen the use of evaluation findings (e.g., networks, targeted events or other platforms to discuss and share findings)? If not, is there a plan to develop these?6. Has there been discussion on how potential negative findings will be managed? |
| C. Key evaluation questions (KEQs) & evaluation design interests | <ol style="list-style-type: none">1. Do stakeholders diverge on the intervention's objectives or do they place different emphasis/importance? If so, are these differences likely to affect how impact will be evaluated?2. Have KEQs been identified and are they suitable for an impact evaluation? If yes, is there consensus on KEQs? If not, is there scope to identify/modify and agree on a limited number of KEQs?3. Are KEQs realistic given intervention design, likely data and resources available for impact evaluation? If not, can they be prioritised (using the ToC) to identify the more contested, critical or innovative presumed causal linkages?4. Do stakeholders have specific or differing expectations or interests in particular impact evaluation designs? To what extent could these designs work with: the identified KEQs; likely data availability; intervention implementation and evaluation timing; and the available evaluation resources? |

Decision support

Proceed with impact evaluation

No major barriers exist

- There is a clear purpose and demand for impact evaluation. Key evaluation questions have been agreed and a strategy is in place/planned for discussing, sharing and using findings including how to deal with potential negative findings.
-

Proceed but address critical issues first

Impact evaluation is assumed feasible in the near future

- There are different views and expectations, but it is reasonable to expect that they can be met by the impact evaluation or be resolved where needed.
 - The number of evaluation questions exceeds the budget and timeframe for the evaluation, but a prioritisation process by stakeholders can be undertaken and can be expected to make the impact evaluation feasible.
-

Do not proceed with impact evaluation

Critical barriers cannot be addressed easily or in a timely manner

- Demand for impact evaluation is lacking (motivation, interest, need, use) and the current institutional environment does not seem conducive to the negotiations and/or adaptations needed to make an impact evaluation viable and worthwhile.
-

3. Is it feasible to measure impact?

Is it possible to measure the intended impact, given on the ground realities and evaluation resources available?

Checklist on data availability and quality

- | | |
|--|---|
| A. Documentation | 1. Are relevant documents available and accessible? Can authors (individuals or organisations) be contacted if more information or clarification is needed? |
| B. Previous evaluations and reviews | 1. Have previous reviews or evaluations been carried out (e.g., mid-term review or process evaluation)? If so, are they relevant to the impact evaluation and how can they be used (e.g., as a baseline)? 2. What are their strengths and weaknesses? 3. Is the raw data available? Is the sampling clear and robust? Are the data collection instruments available? |
| C. Existing data sources to answer the key evaluation questions and their component parts | 1. What are existing data sources and are these accessible (e.g., intervention monitoring data, relevant national statistics, data from previous process, outcome or impact evaluations done by others on same/similar interventions)? 2. If relevant data sources exist: <ul style="list-style-type: none">• Are the responsible individuals/organisations available and prepared to support the impact evaluation (by assessing relevance and quality of existing data and ability/willingness to share the data)?• Is data available on actual beneficiaries? Is there a record of who was involved in what intervention activities, when and to what extent?• Are indicators, where relevant, standardised and aligned with key components in the theory of change? Are other types of relevant information available/being collected?• Are data available/being collected for all required qualitative and quantitative indicators? How frequently are the indicators and other data collected and analysed, by whom and how?• Is it possible to assess the quality of the data? Is the raw data available, as well as the sampling process and data collection instruments?• Are data adequately disaggregated to enable assessment of success in line with intervention goals and intended beneficiaries (e.g., disaggregation by sex, gender, age, ethnicity)?• If baseline data are not available, are there feasible plans for collecting good quality data? |
| D. Causal attribution | 1. What are possible strategies and limitations for determining causal attribution? What possibilities are there for using experimental, quasi-experimental or non-experimental designs? 2. Are other interventions taking place in the same geographical area that could make it more complicated to determine the contribution of the intervention to outcomes/impacts observed? 3. Can alternative explanations for observed outcomes or impacts be assessed? |
| E. Capacity of M&E systems | 1. Do existing/planned M&E systems have capacity to deliver relevant and good quality data (raw or analysed) to be used in the impact evaluation? 2. Are roles and responsibilities for data collection and analysis defined and implemented accordingly? 3. Is the M&E budget adequate to establish and maintain a functional M&E system? 4. Where data are not yet available, do existing staff and systems have the capacity to do so in the future? Have there been any capacity assessment exercises? If yes, has capacity strengthening been implemented where needed? |

Implementation considerations

This section is applicable in all cases, but requires urgent consideration if/when the impact evaluation is to take place imminently.

-
- | | |
|-------------------------------------|--|
| A. Timing and resources | <ol style="list-style-type: none">1. Has the intervention been implemented in a manner (i.e., duration, intensity, scale) that impact can be expected and useful lessons extracted?2. Are there any major external events (e.g., elections, seasonal changes) during the planned impact evaluation that may affect its feasibility?3. Are the necessary resources available for the impact evaluation (e.g., adequate funding, adequate time in-country and within the intervention's activity schedule)?4. Are evaluators with the necessary skills available? Have potential conflicts of interest been made explicit and appropriately dealt with, where needed?5. Are there intervention expansion plans that might interfere with the impact evaluation or which might be taken advantage of during an impact evaluation? |
| B. Coordination requirements | <ol style="list-style-type: none">1. Which beneficiaries, donors, government departments, NGOs, CBOs, and/or research institutions need to be involved in the impact evaluation?2. What mechanisms for coordination are possible? Have they been discussed and agreed? Are roles and responsibilities clear? |
| C. Ethical issues | <ol style="list-style-type: none">1. What ethical issues exist? Are they known or knowable? Can they be addressed adequately and in a timely manner?2. What constraints will ethical issues impose?3. What ethical guidelines are in place? How will arising ethical issues be dealt with and by whom? |
| D. Risks | <ol style="list-style-type: none">1. Are major risks to the evaluation identified and discussed (e.g., physical security risks, weather constraints)? Are staff and key stakeholders likely to be present, or absent (on leave or secondment)? |
-

Decision support

| | |
|--|---|
| Proceed with impact evaluation | No major barriers exist <ul style="list-style-type: none">The impact evaluation is feasible both from a design perspective (with viable options available to ensure useful and credible results) as well as from a logistical perspective. |
| Proceed but address critical issues first | Impact evaluation is assumed feasible in the near future <ul style="list-style-type: none">There are some data quality, financial, logistical and/or capacity issues, but they can be resolved by strengthening existing M&E systems or providing additional resources for impact evaluation. |
| Do not proceed with impact evaluation | Critical barriers cannot be addressed easily or in a timely manner <ul style="list-style-type: none">Capacity, financial resources, timing and/or a severe risk do not allow for the evaluation to take place in the near future. Substantial investments need to be made to correct any gaps, and adverse circumstances need to change considerably before any evaluation can take place. |

5. What happens after the assessment?

The result of an evaluability assessment will rarely give a green light in all areas assessed. Hence, the decision to proceed or not with the impact evaluation is, ideally, negotiated with the key users of the assessment results.

To ensure transparency and facilitate discussions with key stakeholders, the evaluability assessment report should provide information on the methods used for the assessment and sufficient details about the evidence obtained in order to substantiate the recommendations.

All of the issues specified as relevant on the checklist should be addressed. Issues that are not addressed, should be identified and a justification provided. Structuring the assessment report according to these detailed issues may be overwhelming for users, who are mostly interested in the key findings and recommendations. Therefore, the appendices can be used to provide full detail where needed (this is also important for transparency).

As with any study, confidentiality of the information provided by key informants needs to be carefully considered and addressed. In addition, some of the issues uncovered in an evaluability assessment are sensitive. Thus, the evaluators need to consider how to report issues and/or what information to make publicly available.

Once a draft report is produced, the evaluators can use this as the basis for a detailed discussion with relevant stakeholders for verification, clarification and finalisation of the information. The report findings and recommendations can then be used to develop the Terms of Reference for the actual impact evaluation if a go-ahead is decided, or an action plan to address the identified challenges before proceeding.

6. Lessons learned from practice

This guidance note is informed by practice. Here are some key lessons learned from using evaluability assessment for impact evaluation as part of the Methods Lab:

Purpose

-  It is crucial to have clarity about the purpose of the evaluability assessment and expectations about results. Lack of clarity at the outset may result in the end product not addressing the needs. It is therefore important to ask: why the assessment is being done; what kind of recommendations will need to be provided; and what will not be possible or necessary to do?
-  Even though there may be a shared interest between key stakeholders (such as funders, programme managers, implementers) for an impact evaluation, the underlying motivations are not necessarily the same. Getting more clarity about the purpose of an impact evaluation during an evaluability assessment is beneficial to its scope and design, but it may require increased time or other resources for the assessment.

Information needs and decision-making processes

-  The key to a successful evaluability assessment for impact evaluation is to be clear on who wants what information, when and why. It should not be assumed that those whom the evaluators can access easily are the key decision-makers or users of the intended impact evaluation findings. Staff who implement the intervention may be most accessible, but may have different perspectives from other stakeholder groups, and may not have the authority to make and enforce decisions based on the evaluation findings.
-  It is critical that the key evaluation questions for the impact evaluation are determined and unpacked. Ideally, those conducting the evaluability assessment will have an opportunity to assess the validity of these questions, especially from a user/use perspective. Different stakeholders have different information needs, and different agendas may be accommodated in one impact evaluation. However, beware too much compromise: the result may be that none of the stakeholders' needs are fully satisfied, or even partially addressed.
-  While it is important to assess the availability and quality of existing information, it is important to not let this be the sole basis for deciding the key evaluation questions for the impact evaluation. There may also be a tendency to prioritise results that are more likely to be observed and/or easier to measure (such as agricultural yields) than results that are more difficult to assess (such as organisational or social change). It is important to be explicit about the criteria and processes for decision-making around prioritising information.

Timing of the evaluability assessment

- ✓ When the assessment takes place at the beginning of the intervention, not all staff may be in place yet. Key informants who are available to interview therefore may not be those who will be involved in intervention implementation later on. This timing may have implications for the demand of the impact evaluation and prioritisation of the information needs. On the other hand, discussing impact evaluation at the start of an intervention can help managers and implementers to be aware of and address challenges in a timely manner. These challenges may include: adjusting implementation where needed; documenting contextual variables and unintended results; and documenting potential contamination in the case of a counterfactual design. Turnover of key decision-makers and implementing staff during the course of an intervention may affect the relevance of key evaluation questions and evaluation uses identified at an earlier stage.

Who should carry out the evaluability assessment

- ✓ While addressing evaluability can be integrated in the organisational processes, there is value to evaluability assessments that are carried out by independent third parties and not managers or those (likely to be) commissioned to carry out the subsequent evaluation. Particularly, separate contracts should be considered for the evaluability assessment and the (potential subsequent) impact evaluation. This can help to avoid potential conflict of interest.
- ✓ Having more than one person conducting the assessment may be beneficial, particularly when there is a large or complicated intervention and/or many stakeholders. Multiple assessors may also provide different perspectives and may pick up on different issues. However, the roles and responsibilities of different team members need to be clearly articulated and agreed to avoid confusion, or even potential tension.

How to do the evaluability assessment

- ✓ Most of the evaluability assessment questions cannot be answered by a document review alone, thus interviews, discussions, meetings with key stakeholders are essential. Ideally, the assessment should involve an 'on-the-ground' visit to where the intervention is being/has been implemented. These requirements need to be considered in the resource allocation.
- ✓ The theory of change is not always explicit, but it may be possible to construct it through reviewing programme documents and consulting stakeholders. The extent to which this activity is included in the evaluability assessment should be agreed upon from the outset, as it can be particularly time-consuming. In addition, there are often different understandings of what a theory of change constitutes and thus, this needs to be clarified. It may be helpful to conduct a workshop involving all relevant stakeholders.
- ✓ A clear understanding of various stakeholders' roles, responsibilities and/or influences in the intervention is needed if this is not already explicit in the theory of change, and could include:
- Funders' expectations (specifically, those with ultimate decision-making power)
 - The intervention team and its service delivery partners
 - The users of the immediate outputs
 - The beneficiaries, as well as those who might not be included or might be impacted negatively by the intervention

Results and recommendations

-
-  Based on the evaluability assessment, the assessment team may examine the feasibility of alternative evaluation designs and provide information about their relative strengths and limitations, taking into account the intervention design, availability and quality of information and intended impact evaluation uses. However, the assessment team should not prescribe a particular evaluation design.
 -  Answering the checklist questions is usually not about a simple 'yes' or 'no', but more likely involves addressing 'to what extent'. Often answers fall into the 'adequate' category, and final decisions need to be made about what information needs are prioritised over others and which of the relationships and assumptions in the theory of change will be the focus of the impact evaluation. One of the biggest challenges of an evaluability assessment that aims to propose alternative evaluation designs is how to make sure that recommendations about the multiple trade-offs are transparent and understood by all those who will be involved in the subsequent evaluation or will be using the evaluation findings.
 -  In situations where there are many information needs, an explicit prioritising session with stakeholders may be necessary. If this cannot take place face-to-face, then a virtual session, albeit more difficult to organise, may still be beneficial. This should be facilitated by an independent person and should use clear criteria and set engagement rules.
-

Resources

-
-  Ensure adequate time for the assessment; making judgements based on short visits or brief conversations is likely to lead to misguided findings.
 -  Many of the questions in the checklist tool do not only require an in-depth understanding of the intervention but also of the context and conditions within which the intervention operates. This requires sufficient time for the assessment. This upfront investment during the evaluability assessment is likely to reduce the cost of the subsequent impact evaluation, as it will be able to draw on the information gathered during the assessment.
-

7. Additional resources

- BetterEvaluation (2015) 'Evaluability assessment.' BetterEvaluation (http://bettterevaluation.org/themes/evaluability_assessment).
- Davies R. (2013) 'Planning evaluability assessments. A synthesis of the literature with recommendations.' Working Paper 40. London, Department for International Development. http://r4d.dfid.gov.uk/pdf/outputs/mis_InfoComm/61141-DFIDWorkingPaper40-finalOct13.pdf
- Davies R. and Payne L. (2015) 'Evaluability Assessments: reflections on a review of the literature.' Evaluation (21): 216–231. <http://evi.sagepub.com/content/21/2/216.short>
- Dunn E. (2008). 'Planning for cost effective evaluation with evaluability assessment. Impact assessment primer series.' Publication #6. Washington DC: United States Agency for International Development. http://pdf.usaid.gov/pdf_docs/PNADN200.pdf
- UN Women Evaluation Unit (2009). *Guidance note on carrying out an evaluability assessment*. Evaluation Guidance Note Series, Nr 4. New York: UN Women. <http://erc.undp.org/unwomen/resources/guidance/Guidance%20Note%20-%20Carrying%20out%20an%20Evaluability%20Assessment.pdf>
- Trevisan M. and Walser T. (2015). *Evaluability assessment. Improving evaluation quality and use*. Thousand Oaks: Sage Publications, Inc. <http://www.sagepub.com/books/Book240728>



ODI is the UK's leading independent think tank on international development and humanitarian issues.

Readers are encouraged to reproduce Methods Lab material for their own publications, as long as they are not being sold commercially.

As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI, the Australian Department of Foreign Affairs and Trade (DFAT) and BetterEvaluation.

© Overseas Development Institute 2015. This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 3.0).

Peersman, G., Guijt, I., and Pasanen, T. (2015) 'Evalubility Assessment for Impact Evaluation'. A Methods Lab publication. London: Overseas Development Institute.

ISSN: 2052-7209
ISBN: 978-0-9941522-1-3

All ODI Reports are available from www.odi.org

Overseas Development Institute

203 Blackfriars Road
London SE1 8NJ
Tel +44 (0)20 7922 0300
Fax +44 (0)20 7922 0399

www.odi.org