Manager's guide to evaluation

This is a guide for people who are managing an evaluation.

The guide can be used for managing an evaluation that is conducted by an external evaluator or evaluation team, an internal team, or by a combination of these. It can be used for different types of evaluations and for evaluations of different types of interventions, including projects, programs, policies and clusters of projects. It can also be used for evaluation of research.

The guide aims to support decision-making throughout the process of an evaluation, from planning its purpose and scope, designing it, conducting it, reporting findings and supporting use of its findings. In many organizations, this process will draw on the expertise of several individuals. Additional help may need to be obtained for one or more steps in the process.

The information is organized in 9 steps. In some cases, the order in which the steps are addressed in the evaluation process might be slightly different, or earlier steps might need to be revisited in response to changing circumstances and needs.

Roles and responsibilities

Determine how decisions will be made on: the focus of the evaluation; choosing the evaluator / evaluation team; approving the evaluation design; approving evaluation report(s) and who can access the report(s) and data.

There are many decisions to be made in an evaluation. It is important to be clear about who will be involved in making these decisions, what their role will be and how the decisions will be made.

Always check if existing processes and structures can be used; if these are not appropriate or adequate, then new ones may need to be established for the purposes of the specific evaluation. Also consider any pre-established agreements such as, for example, partnership agreements.

Control of the evaluation process may be centralised in a dedicated manager or committee or it may be shared by a working group involving representatives from many different stakeholders. It is important to describe clearly each actor / entity's role to avoid confusion, duplication of effort or things falling through the cracks.

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation management plan
- Decision making matrix
- List of responsibilities of the evaluation manager
- List of responsibilities of the evaluator
- Evaluation partnership agreement

Identify who will be involved in decisions and what their roles will be

It is important to be clear about who will be involved in the various decisions involved in an evaluation, and what their roles will be.

Who might be involved in making decisions?

Consider the possible involvement of:

- The manager of the intervention
- An evaluation steering committee
- A technical advisory group or some individual technical advisors
- A community consultation committee or key informants from the community

What will be their role in decision making?

The role of each individual or group in relation to specific decisions can be specified as follows:

- Consulted - Those whose opinions are sought; they are engaged in two-way communication.
- Recommends - Those who are responsible for putting forward a suitable answer to the decision.
- Approves - Those who are authorised to approve an answer to the decision.
- Informed - Those who are informed after the decision has been made; they are engaged in one-way communication.

How will decisions be made?

Decisions may be made in different ways; one or more of the following processes may be used:

- Majority decision making - Decisions are made on the basis of the support of the majority of the decision makers; in contentious decisions, it is important to be clear about who is able to vote including whether proxy votes are allowed.
- Consensus decision making - Decision making processes that aim to find decisions which everyone can accept; in practical terms, that can mean giving all decision makers the right of veto.
- Hierarchical decision making - Decisions are made on the basis of formal positions of authority.

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Decision making matrix - It can be useful to summarise decision makers and types of decisions in a matrix which can be referred to when developing and implementing evaluation processes. It is important to ensure that -where relevant decision making structures and processes already exist- these are incorporated.

An example of a decision making matrix

	Technical advisory committee	Evaluation steering group	Program manager	Senior management
Focus of evaluation	Consulted	Recommends	Approves	Informed
Selection criteria for evaluator/evaluation team	Consulted	Recommends	Approves	Informed
Choosing evaluator/evaluation team	Consulted	Approves	(included in steering group)	Informed
Evaluation design	Consulted	Approves	(included in steering group)	Informed
Evaluation report	Consulted	Approves	(included in steering group)	Informed
Release of report and data	Consulted	Consulted	Recommends	Approves

Example

An Evaluation Steering Group was used for ongoing evaluation of a large multi-donor initiative:

The Think Tank Initiative (TTI) is a major program funded jointly by the William and Flora Hewlett Foundation, the Bill & Melinda Gates Foundation, the International Development Research Centre (IDRC), the UK Department for International Development (DFID) and the Norwegian Agency for Development Cooperation (Norad).

TTI's operations are implemented by the International Development Research Centre (IDRC), while strategic decisions are taken by an Executive Committee composed of one senior member from each donor organization and the TTI Program Manager.

TTI Phase II will be independently evaluated throughout the 5 years. The evaluation is intended to provide timely and actionable feedback to allow for the adaptive management, as well as rigorously documented and validated learning about the program. This balances both accountability and learning purposes.

The evaluation is being commissioned and managed by IDRC, and the evaluators will have a single point of contact at IDRC with whom to interact on all evaluation-related matters. IDRC is managing the evaluation in collaboration with an Evaluation Steering Group composed of delegates from each of the five TTI funding agencies. Steering Group representatives coordinate evaluation inputs from their respective organizations.

The Evaluation Steering Group is responsible for:

- reviewing and approving work plans;
- participating in annual evaluation planning meetings;
- commenting on and approving evaluation deliverables; and,
- keeping the TTI Executive Committee appraised of development and progress taking place in the TTI Evaluation.

The TTI Executive Committee is responsible for approval of all interim and final reports.

Resources

Decision making structures

- Advisory group

  An advisory group can be established to provide advice on an individual evaluation, a series of evaluations, or the evaluation function within an organization.

- Steering group

  Evaluation management often involves a steering group, which makes the decisions about the evaluation. It is important to distinguish between a steering group (which makes decisions) and an advisory group (which provides advice).

Decision making processes

- Consensus decision making

  Consensus decision is a decision-making method that involves reaching agreement between all members of a group with regards to a certain issue.

Specify responsibilities of the evaluation manager and the evaluator(s)

Evaluation managers are often, but not always, the project / program manager, the head of programming in the organization, or the manager or other staff from the evaluation unit (or the dedicated M&E officer in the organization).

The evaluation manager is responsible for:

- ensuring that the evaluation runs according to plan (see Step 6) and meets the milestones or deliverables on time
- problem-solving where needed (or direct issues to the relevant individual / entity to address)
- ensuring evaluators have access to all relevant project / program documents and stakeholders involved in the project / program and/or the evaluation (see Step 7)

In a large evaluation, the evaluation manager may be assisted by one or more other staff members who will be assigned specific responsibilities in the management process.

Basic management skills also apply to managing an evaluation in terms of personnel management and logistics. Where evaluation management differs is in the content: the evaluation manager needs to have a basic understanding of evaluation methods and processes to assist the evaluation team with making the best possible choices.

It is crucial to the success of the evaluation that the evaluation manager:

- is formally identified;
- has a clear understanding of the scope of her/his authority (i.e., knows what she/he can decide herself/himself related to the running of the evaluation and what she/he needs to get clearance for and from whom);
- can access relevant information and staff (of the organization commissioning the evaluation and of the intervention implementers).

In addition, a clear understanding of the roles of all those involved in the evaluation is essential including the evaluator(s).

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

List of responsibilities of the **evaluation manager** which may include:

- Educate the external evaluator(s) about: the program's objectives, operations and intended beneficiaries; the expectations about the evaluation and any relevant organizational background.
- Provide input and/or collate feedback on the evaluation plan.
- Specify the reporting requirements in terms of progress in the implementation of the evaluation (including reporting of important challenges and their resolution or which potential issues need to be raised for decision making elsewhere)
- Specify what is expected to be included in the formal evaluation report(s).
- Keep the evaluator(s) appraised of any changes in the program's operations or evaluation context.
- Provide regular updates on the evaluation process to all staff.
- Monitor the implementation of the evaluation including completion of milestones/deliverables.
- Facilitate program staff involvement in the evaluation, where relevant and agreed.
- Serve as the trouble-shooter, resolving problems or locating help to resolve them.

Adapted from: The Program Manager’s Guide to Evaluation. Second Edition. Washington DC: Office of Planning, Research and Evaluation. Administration for Children and Families, US Department of Health and Human Services, 2010; p.28

List of responsibilities of the **evaluator** which may include:

- Develop an evaluation plan, in conjunction with the evaluation manager and program staff.
- Provide monthly or quarterly progress reports on the implementation of the evaluation (written or in person).
- Attend program staff meetings, evaluation advisory board or coordinating committee meetings.
- Train data collectors such as on:
    - Participant/case selection for sampling purposes
    - Using data collection instruments
    - Data quality assurance
- Ensure adherence to ethical standards adherence (e.g., confidentiality of data) during all phases of the evaluation.
- Implement of oversee implementation of data collection such as:
    - Interviewing program staff, program participants
    - Conducting focus groups
    - Observing service delivery activities
    - Reviewing participant case records
    - Developing data management procedures and tools (e.g., database)

    - Coding, entering, and cleaning data
    - Analyzing data
- Write interim (quarterly, biannual, yearly) evaluation reports and the final evaluation report.
- Present findings to program staff and others in the organization
- Present findings at meetings and conferences.

Adapted from: The Program Manager's Guide to Evaluation. Second Edition. Washington DC: Office of Planning, Research and Evaluation. Administration for Children and Families, US Department of Health and Human Services, 2010; p.27

Address particular evaluation management issues relating to joint projects, including donor partnerships

Evaluations may involve interventions that are collaborative ventures (such as co-funded or jointly implemented projects).

Different types of partnerships and their implications for evaluation management:

There can be collaborations at various stages and levels of an intervention – such as among implementing organizations, within a network, between implementing agencies and donors, or among donors.



See more detailed information:

Types of partnerships and their implications for evaluation management processes
DOC
41.5 KB

Evaluation of donor partnerships:

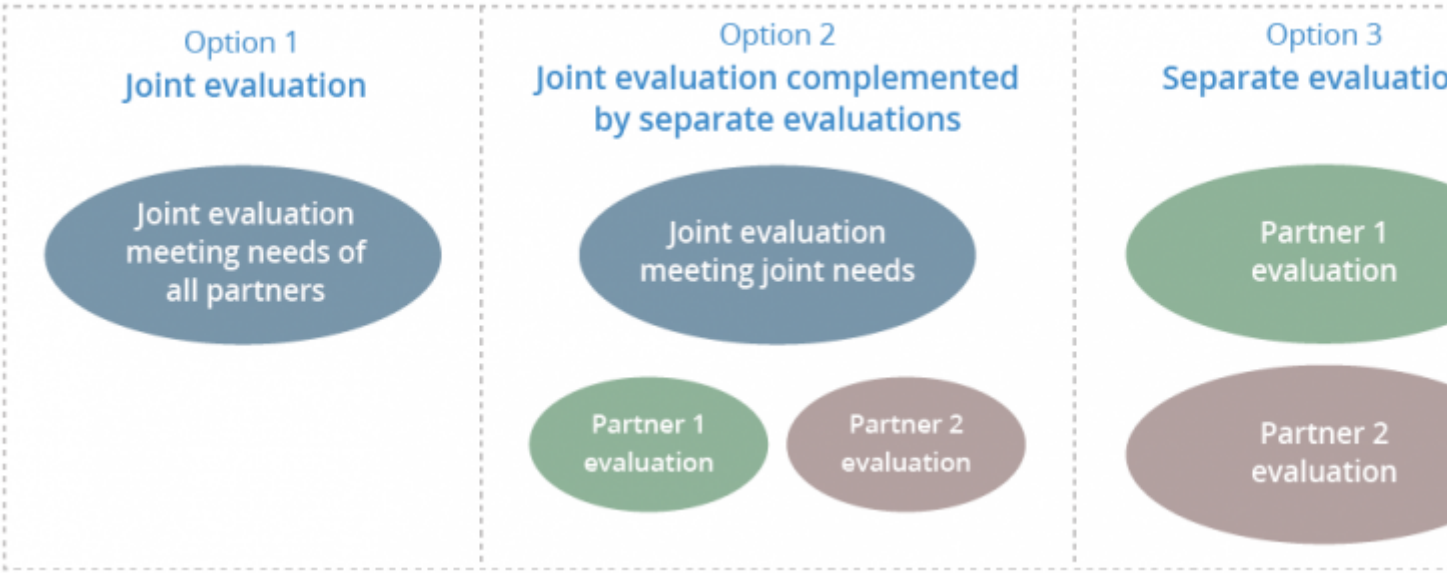
The evaluation of donor partnerships requires careful management and governance. All donors enter a project with their own political context and their own set of expectations. Deciding on key strategic questions -such as what constitutes success and what results donors need to demonstrate- can take time.

Most donors have specific expectations for monitoring, evaluation and reporting related to supported projects or programs. The evaluation manager should be cognizant of these and ensure that the latest information is available as donor's expectations may change over time.

They may include requirements about the use a 'third party' evaluator (i.e., someone who is not affiliated, in any way, with any of the organizations involved in project or program implementation), specific evaluation designs, involvement in the evaluation process, approval or sign-off procedures, data or report sharing and the like.

Where the requirements from different donors are not aligned, or where the needs of a specific evaluation would be better served by a different or more flexible approach, the evaluation manager is responsible for negotiating an agreed way forward with the different donors.

Different options for joint evaluations:



Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation partnership agreement

IDRC-specific information

Donor partnerships, as collaborative ventures that are managed and implemented by IDRC, are increasingly a core part of IDRC's business model. The increasing emphasis on jointly funded projects and other types of partnerships presents particular issues for evaluation management processes. Agreements about evaluation are made during partnership development. Partnership agreements generally lay out when evaluation will take place, who will be responsible for managing it, and what the overall use will be. More detailed discussions take place when it is time to develop the specific terms of reference for the evaluation.

Scope evaluation

It is important to take the time to consider carefully what the evaluation needs to do before thinking through possible evaluation designs.

Ensure all those who need to be consulted during this process are adequately involved.

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Description of the programming that is going to be evaluated (the evaluand)
- Theory of change and/or logic model
- List of primary intended users and their uses for the evaluation
- List of agreed key evaluation questions
- Evaluation timeline
- Evaluator qualities
- Evaluation budget

These products inform the development of a formal Terms of Reference (ToR) or Request for Proposal (RFP) (see Step 3).

IDRC-specific information

IDRC staff and partners may wish to peruse previous evaluations, particularly in similar topic areas, to understand how scope can be delineated and defined in various circumstances. IDRC maintains, in its online open-access digital library, a repository of evaluations conducted throughout IDRC's history.

IDRC staff can also access a repository of Evaluation Terms of Reference compiled by the Policy and Evaluation Division.

Clarify what will be evaluated

An evaluation can focus on a project, a number of projects, a program, a policy, a strategy, an organization, a network.

It is helpful to produce a succinct statement about:

(a) *what is to be evaluated* – which may include information on:

- The rationale: the issue being addressed, what intervention is being done, who is intended to benefit from it, and what the intended results are
- The scale of the intervention, budget and resources allocated and stage of implementation
- The roles of partner organizations and other stakeholders involved in implementation
- The implications of contextual factors – geographic, social, political, economic and institutional circumstances which create opportunities or challenges
- Significant changes in the intervention that have occurred over time – because of changes in contextual factors or lessons learned

(b) *what is considered to be outside the boundaries of the evaluation* – For example, some activities or some longer-term impacts.

Checking this initial description with different stakeholders can be a helpful way of starting to identify where there are disagreements or gaps in what is known about the intervention and/or the boundaries of the evaluation.

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Description of the evaluand

Describe the theory of change

This section explains how and why you might use a theory of change when commissioning and managing an evaluation.

It explains options for how it will be developed or revised, how it will be represented, and how it will be used.

You might be actively involved in these processes or oversee them. In either case it is important to be aware that there are choices to be made and that informed choices will produce more useful theory of change and better evaluation.

A theory of change explains how the activities undertaken by an intervention (such as a project, program or policy) contribute to a chain of results that lead to the intended or observed impacts. Other labels that your colleagues, partners and evaluators might use include: *results chain, logic model, program theory, outcome mapping, impact pathway and investment logic*.

A theory of change is often developed during the planning stage but can also be useful for monitoring and evaluation. A good theory of change can help to: develop better Key Evaluation Questions, identify key indicators for monitoring, identify gaps in available data, prioritize additional data collection, and provide a structure for data analysis and reporting.

Your intervention might already have a theory of change that was developed in the planning stage. You are likely to benefit from reviewing and revising the theory of change as part of commissioning an evaluation in the following circumstances:

- there is disagreement about how valid or comprehensive the current theory of change is
- there are gaps or errors in the current theory of change
- there is little evidence to support the current theory of change (either from the program or from other research and evaluation)
- your understanding of how the project or program works has developed further since the original theory of change was developed,
- the context has changed in significant ways
- the current version is adequate for planning purposes but needs more detail for an evaluation

As a manager, you might be directly involved in developing (or revising) and using the theory of change, or you might oversee the process which internal staff and/or an external evaluator conduct. Whatever your level of direct involvement, you will want to ensure the quality of the process and the product. A key part of this is ensuring there are informed choices made about the processes used to develop (or revise) the theory of change and how to represent it. These choices should take into account how the theory of change is intended to be used and any particular features of the intervention. The following sections discuss these in more detail.

The rest of this section provides guidance in terms of the following key issues:

1. Planning how the theory of change will be used for monitoring and evaluation

2. What the theory of change should cover
3. The process for developing or revising the theory of change
4. The scope of the theory of change
5. Explicit and appropriate change theories and action theories
6. The representation of the theory of change

1. Planning how the theory of change will be used for monitoring and evaluation

Depending on the timing, a theory of change can be used to anticipate what will happen, and establish data collection processes to track changes going forward, or used to make sense of what has happened and the data that have already been collected.

A theory of change can inform the development of a monitoring and evaluation.

Existing data (where available from the intervention and/or previous research and evaluation) can be mapped onto the theory of change then used to identify priority areas for collecting additional data. These might include:

- Assumptions on which the theory of change is based.
- Contextual factors that might be important to gather data on and use to investigate patterns in results – for example, does the intervention work particularly well at certain sites or for certain groups of people
- Indicators of the quality and quantity of inputs and activities to support effective management
- Early indicators of progress or lack of progress in achieving results. This can be particularly important when the intended impacts are longer-term and information about intermediate outcomes is needed to inform decisions
- Links where the causal chain seems to break – where achieving a particular intermediate outcome does not seem to lead to the subsequent outcome
- Causal links which are not well established
- Identifying outliers – "bright spots" that might inform learning and serious problems that need to be addressed immediately

A theory of change can provide a framework for a "performance story" – a coherent narrative about how the intervention makes particular contributions. This can be useful for communicating about the intervention to potential partners, participants and policymakers, and for also providing a consistent point of reference for those involved in implementing and managing it.

2. What the theory of change should cover

A theory of change is not just a list of activities with arrows to intended outcomes. It needs to explain how these changes are understood to come about and the role the intervention will play in this – and the role of other factors, including other interventions.

It therefore needs to include both:

While the core of the theory of change focuses on the links between activities and impacts, it is more useful if it does not only cover these. Check if the following elements are in place and, if not, if it is possible to add them either in the main diagram and narrative or in supplementary documents:

- **Change theory:** this identifies one or more causal mechanisms by which change comes about for individuals, groups and/or communities. (more guidance is provided below on this)

- **Action theory:** this explains how interventions are constructed to activate their theory of change in terms of the activities that will be undertaken and what level of success will be needed for each result to produce the final intended impact (more guidance is provided below on this)
- how **other projects and programs** contribute to producing impacts
    - those who are explicitly collaborating (these are referred to as 'boundary partners' in outcome mapping forms of theory of change)
    - others who have positive or negative influence
- how **the particular contexts in which the intervention is implemented** affect activities and results
- **potential unintended results**, both positive and negative,
- **assumptions** on which the theory of change is based – these are in addition to the cause-effect relationships shown in the logic model and often involve assumptions about the context
- how participants **become engaged** in a project, program or policy,
- how **results are expected to be sustained** after a project, program or policy ends or participants' engagement ends.

A negative theory of change can also be developed to identify possible negative unintended outcomes in order to set in place risk mitigation strategies to avoid them, and data collection that will detect if they have occurred.

3. Process for developing or revising the theory of change

A sound theory of change draws on a range of evidence – previous similar projects and programs, previous research and evaluation, the mental models of stakeholders (including planners, managers and staff, partner organizations, and intended beneficiaries), and observation of the program and patterns in outcomes and impacts..It is important to ensure that the process is adequately inclusive of relevant perspectives, values and evidence. If the theory of change has only used a group meeting to build it, it is likely that some more systematic analysis and review of relevant research and evaluation will improve its quality

If you are developing a new theory of change, or reviewing an existing one, check if these different processes have been included and, if not, if it is possible to add them:

- A **situation analysis** – an assessment of the needs and problems the intervention is intended to address and of the resources and opportunities that might be drawn on to do this?
- Download some questions that can be used to structure this situation analysis.
    - Questions to ask in a situation analysis to develop a theory of change
      DOC
      31.5 KB
- A **review of existing documentation** which explains why an intervention was developed
- **Relevant research, evaluations and other evidence from** similar projects, programs or policies
- **Talking with stakeholders** about how they understand the intervention works or is intended to work – what are the intended outcomes and how do they think they might be achieved (their mental models of the intervention).
- Download some questions that can be used in individual or group interviews with key informants, including those who designed a project, program or policy, those who are currently working to deliver or manage it, and those who are involved in it in other ways.
    - Sample interview questions to articulate the implicit theory of change of a project.doc
      DOC
      33 KB

Check that the process of reviewing or developing the theory of change **involves the right people in the right ways** in the process of developing or reviewing the theory of change. In some cases it will be possible and desirable to involve a range of people in the whole process of gathering information and developing the theory of change; in other cases it will be better to delegate or hire someone to develop a draft and then engage the wider group in reviewing and revising the draft.

- Download this matrix of different people and groups and roles that can be used to support discussion of this issue and document the decisions made.
    - Who should do what in developing a theory of change
      DOC
      40 KB

If you are reviewing and revising an existing theory of change, talk with your staff, colleagues and partners and check previous documentation to review it in terms of these issues:

- What **evidence** was the basis for its development? What additional evidence should be used in the review?
- Whose **mental models** formed the basis of it? To what extent and in what ways were the perspectives and mental models of intended beneficiaries and partner organizations included?
- Were there **different views** about it – in terms of what the intended outcomes and impacts were and/or how these might be brought about?
- Has there been **more recent research and evaluation**, or similar projects and programs, that could inform the theory of change?

If there are gaps in the evidence that has been used to develop the theory of change, or indications that it has changed since being developed, draw on these different sources of evidence to revise it.

A theory of change has most benefit if it provides a common reference point for those working together. This means it needs to be accessible and referred to during discussions and decisions about the project or program. But sometimes it is ignored or forgotten after the initial planning stage, especially if new people come into the program or project and are not aware of what has been done.

Talk with your staff, colleagues and partners to find out:

- Is the current theory of change known, understood and currently used? What can be learned from this?
- If the theory of change isn't being used, is this because of perceived inadequacies?
- If the theory of change is being used, what has been learned about it in use?

4.Identify the outcomes and impacts to be included and who will be involved in producing these

It is important to be clear about the intended impacts of projects, programs and policies. Sometimes there will be different views among partner organisations about these.

The intended impacts might be for:

- Participating individuals – for example, increased skills or knowledge or changes in behaviour
- Other individuals affected by participants – for example, students taught by teachers whose skills have been improved by the project or program
- Organizations
- Communities

- Networks and systems of organizations and services

In some programs and projects there is clarity and agreement about the intended impacts. In other cases there is disagreement (for example, when different partner organizations have different agendas for involvement) or uncertainty (for example, in a capacity development project where the specific changes that will arise are not tightly specified in advance).

Talk with your staff, colleagues and partners and check previous documentation to find out:

- Is there agreement about the intended impacts or do different partners, organizations or individuals have different views, or is there uncertainty?
- If there are different intended impacts, is there tension between them or are they synergistic?

It is also important to be clear about how these intended impacts are expected to be produced – and who will be involved in doing this. In some cases, your project or program might be directly involved – for example, providing direct services. But in many cases, you will be working with other organizations either at the same time or in sequence to bring about the intended changes.

For example, you might work with participants to increase their knowledge and skills, and then they work directly with intended beneficiaries, or you support them to produce research outputs and then organizations are intended to use this research to inform and improve policy and planning.

Talk with your staff, colleagues and partners and check previous documentation to find out:

- Who is expected to be involved in bringing about changes?
- Should they also be involved in developing or reviewing the theory of change?

5. Explicit and appropriate change theories and action theories

For example, behavior changes (such as reduced drink driving or increased uptake of science research) can come about through one or more change theories:

- changing social norms
- changing incentives (higher risk of sanctions or increased rewards)
- capacity development
- increasing opportunities and/or removing barriers.

For each change theory, there are different possible action theories about what activities might be implemented to trigger the change theory. For example, changing incentives in terms of increasing rewards might involve:

- providing an individual monetary bonus for all who comply
- creating a lottery for all who comply with one or more winners drawn randomly
- providing public recognition and praise for high performers

Being explicit about change theories and action theories makes it easier to identify what are appropriate local adaptations of a program and what constitutes good quality implementation. It is likely that there will be different change theories and action theories at different stages of the project or program and at different sites.

Try to ensure that the theory of change has explicit change theories and action theories. Talk with your staff, colleagues and partners, check previous documentation and review relevant research and evaluation to find out:

- What are the change theories underpinning expected changes for individuals, organizations, and communities? How plausible do these seem?
- How well does the theory of change make explicit the change theories underpinning it?
- Are there different change theories at different stages of the project or program?
- Are there different change theories for different people? (For example, motivation for some people who already have capacity, and capacity-building for people who already have motivation)
- Would the theory of change be improved if additional change theories were added in key points?

The project or program activities are intended to contribute to the change process. How they do this can be understood as an action theory – a theory that if the project or program does particular things, these activities will trigger the type of change identified in the change theory.

Talk with your staff, colleagues and partners, check previous documentation and review relevant research and evaluation to find out:

- What are the action theories underpinning the different change theories for individuals, organizations, and communities? How plausible do these seem?
- How well does the theory of change make explicit the action theories underpinning it?
- Are there different action theories for different people? (For example, motivation for some people might be triggered by providing a tangible incentive of public recognition for their work; for others a financial reward might be needed to be seen as motivating)
- Would the theory of change be improved if additional action theories were added in key points?
- Download example change theories and action theories that could produce different types of outcomes and impacts at different stages of a program.
    - Some example change theories and action theories
      DOC
      71.5 KB

6. The representation of the theory of change

A theory of change is often represented in a diagram with an accompanying narrative. There are different types of diagrams that can be used. Diagrams should clearly show the direction of change and are most commonly drawn to be read from left to right, top to bottom, or bottom to top.

Sometimes it is useful to have several different versions – such as an overview diagram for general use with more detailed diagrams of particular components or for particular purposes. For complicated theories of change, it can be helpful to have different diagrams with varying levels of detail. An accompanying narrative can complement the diagram and be more accessible for some people.

There are many different options for representing a theory of change and it is important to choose a format which will communicate clearly. Four main options include:

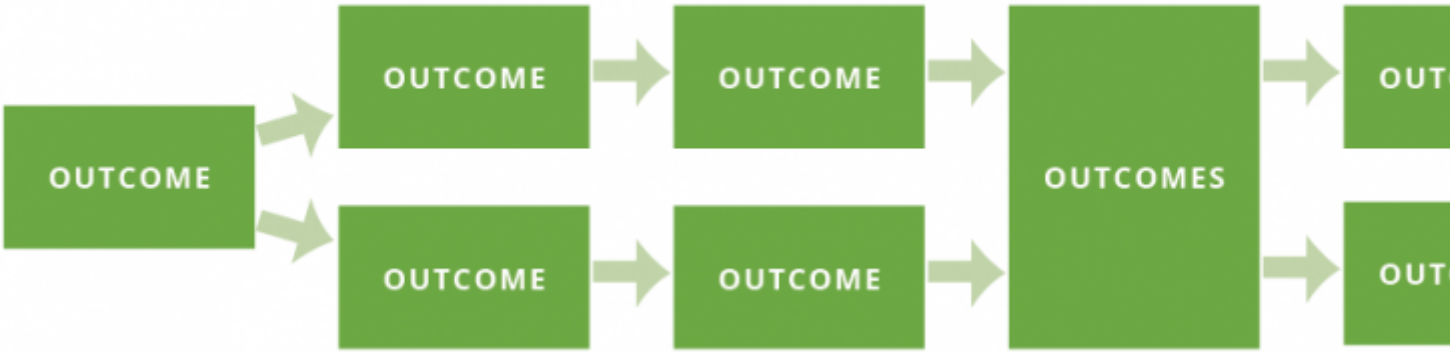
- a simple, linear **results chain** – This has a series of boxes often in the form of inputs, activities, outputs, outcomes and impacts. It is most appropriate for fairly simple interventions, where activities are undertaken at the start and then the consequences flow through in a linear fashion.

Results chain

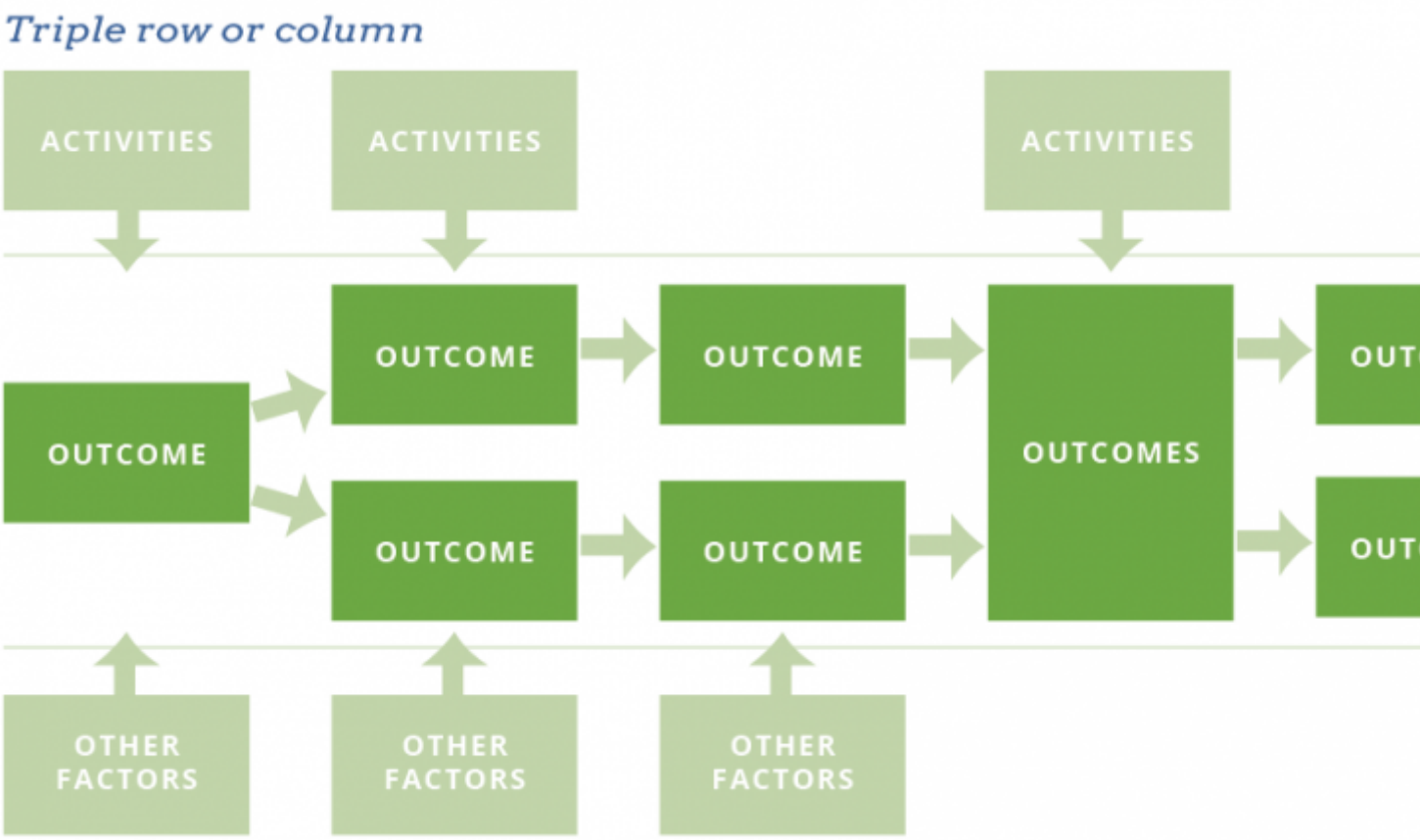


- an **outcomes hierarchy** - This shows the sequence of results, from short-term to long-term. It is appropriate when the causal chain is complicated, with multiple strands. It focuses attention on the causal sequence and provides information about activities in a separate narrative or table

Outcomes hierarchy



- a **triple column/row** . This shows the causal pathway in terms of intermediate outcomes, activities which directly produce these, and the influence of other factors and programs. It can be particularly useful for showing activities that occur along the causal pathway, and for showing clearly the contributions of other partners and contextual factors



- **a set of principles**. This is particularly appropriate for adaptive, emergent projects and programs, in terms of principles. For example, the following principles have been identified for strengthening research capacity in low and middle income countries (Add source):

Principles for strengthening research capacity in low and middle income countries

1. Network, collaborate, communicate and share experiences
2. Understand the local context and accurately evaluate existing research capacity
3. Ensure local ownership and secure active support
4. Build in monitoring, evaluation and learning from the start
5. Establish robust research governance and support structures, and promote effective leadership
6. Embed strong support, supervision and mentorship structures
7. Think long-term, be flexible and plan for continuity

Check the quality of the diagram in terms of its coherence, logic and clarity and revise it as needed:

- Does the diagram provide a **clear overall message** about how the project, program and policy contributes to the end results? If not, can the diagram be redrawn to emphasise the overall narrative? For example:
  - If there are three main parallel elements, create a symmetrical diagram which conveys this message clearly.
  - If there is a lot of detail, provide a summary version that can then be expanded or further explored.
  - Avoid too much detail about the impacts and how they will be measured.
- Can the diagram be read as a **coherent story** about sequence and consequence? In particular:
  - check that every arrow is meaningful (one thing leads to or helps to bring about another thing) and that the wording in each box is appropriate.
  - indicate the direction of expected change (increased or decreased)

Talk with your staff, colleagues and partners, check previous documentation and review other theories of change to find out:

- What would be the best way to represent the theory of change?
- Would it be helpful to have different versions for different users and/or different levels of detail?

Identify who are the primary intended users of the evaluation and what will they use it for

In most cases, the evaluation will have multiple uses. By clarifying and making explicit the intended use(s) of the evaluation for each user, it is easier to have transparent and informed discussions and decisions about the priorities for the evaluation, to focus its attention, and to ensure that all methodological and procedural decisions are made with attention being paid to their likely effect on the utilization of the evaluation.

The primary **intended users** are *not* all those who have a stake in the evaluation, nor are they the general audience. They are the specific people, in a specific position, in a specific organization who will use the evaluation findings and who have the capacity to effect change. From start to end, the evaluation process should be designed and carried out around the needs of the primary intended users. They have the responsibility to do things differently (e.g., make decisions, change strategies, take action, change policies, etc.), because of their engagement in the evaluation process and/or with the evaluation findings.

Determining the **intended use(s)** of an evaluation typically involves a negotiation between the evaluator(s) and the primary intended user(s). By involving all primary intended users in this negotiation, the various perspectives are better represented and consensus can be reached about the priority use(s).

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- List of primary intended users and their uses for the evaluation

IDRC-specific information

See: Identifying Intended Use(s) and User(s) of an Evaluation – this IDRC guideline highlights the importance of identifying the primary intended user(s) and the intended use(s) of an evaluation and outlines a variety of methods that can be used to achieve this in the initial planning stage.

Resources

- Identify primary intended users
  Clarify who will actually use the evaluation—not in vague, general terms (e.g. "decision makers") but in terms of specific identifiable people (e.g. the manager and staff of the programme; the steering committee; funders deciding whether to fund this programme or similar programmes in the future).
- Decide purposes

Clarify the intended uses of this evaluation—is it to support improvement, for accountability, for knowledge building? Is there a specific timeframe required (for example, to inform a specific decision or funding allocations)? If there are multiple purposes, decide how you will balance these.

Develop agreed key evaluation questions

Evaluation, by definition, must answer truly evaluative questions: it must ask not only 'What were the results?' (a descriptive question) but also '*How good* were the results?' (an evaluative question).

Depending on the type of evaluation, causal questions also need to be addressed (to what extent were the results due to the intervention?).

An evaluation should be focused around answering a small number of high-level key evaluation questions (KEQs) which are about performance overall. Each of these key evaluation questions (KEQs) should be further unpacked by asking more detailed questions about performance on specific dimensions of merit (related to evaluative criteria such as relevance, equity, effectiveness, sustainability). The KEQs also need to reflect the intended uses of the evaluation.

Good KEQs are:

- Limited in number: 7 ± 2 questions is a good number in general. This allows for coverage of different aspects of the intervention, but is a small enough number of questions to not get overwhelmed.
- Open questions (not yes/no answers).
- Are specific enough to help focus the evaluation, but broad enough to be broken down further into more detailed questions to guide data collection.

Work with primary intended users of the evaluation to develop an agreed list of key evaluation questions.

Being clear about the intended use of the evaluation and the type of evaluation needed, can help with developing appropriate Key Evaluation Questions.

The following typology can be used to classify the type of evaluation and typical questions:

- Needs analysis
    - What is needed?
    - What are unmet needs?
- Intervention design
    - What is the best way to design the intervention?
- Monitoring
    - How is it going? (regular reporting of metrics)
- Process evaluation
    - Is the intervention being implemented according to plan (periodic investigations)?
    - What has been done in an innovative program?
- Outcome / impact evaluation
    - What results have been produced?
    - What has (and has not) worked for whom in what circumstances?
- Economic evaluation
    - Has the intervention been cost-effective (compared to alternatives)?
    - What has been the ratio of costs to benefits?

[Source: Adapted from Owen J with Rogers P (1999). Program Evaluation: Forms and Approaches. Sydney: Allen & Unwin/London: Sage UK.]

These evaluation types are cumulative: outcome / impact evaluation needs data from process evaluation, and economic evaluation requires data from outcome impact evaluation.

The level of existing knowledge will also be important in developing appropriate evaluation questions:

- When we know what works & why, it is sensible to...
  - …ask if processes are being followed (describe activities compared to an agreed standard)
  - …demonstrate value of what is being done (describe outcomes compared to agreed statement of goals and/or needs)
- When we don't know if it works, it is sensible to...
  - …look at process outcomes / impacts (test theory)
- When we don't know which is the best way, it is sensible to...
  - …document process & context & compare performance (outcomes / impacts, efficiency)
- When we don't know what could work, it is sensible to...
  - …use action research/learning & share results (ask a series of questions about early indications of success or failure)

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- List of agreed key evaluation questions

Examples

- Download an example of key evaluation questions from an evaluation of the African Institute for Mathematical Sciences (AIMS)
  - Example of Key Evaluation Questions (AIMS)
    DOC
    36.5 KB

Resources

- Specify the key evaluation questions
  Articulate a small number of broad evaluation questions that the evaluation will be designed to answer. These are different to the specific questions you might ask in an interview or a questionnaire.
- Good Evaluation Questions: A Checklist to Help Focus Your Evaluation

  Guidance developed by the National Asthma Control Program, US Centers for Disease Control and Prevention.

- What are the key evaluation questions? (Archive link)

  Guidance developed by the Office for Learning and Teaching of the Australian Government.

Decide the timing of the evaluation

Monitoring (the routine tracking and reporting of priority information about an intervention) and **evaluation** (a discrete study to produce an evaluative judgement about merit, worth or significance of an intervention) are distinct but highly inter-related activities.

Generally, monitoring and evaluation findings are used at different times, with different regularity, different resource needs and for different purposes. Both monitoring and evaluation are needed for effective program management and decision making. It is insufficient to conduct monitoring without any kind of evaluative reflection, and, given the episodic nature of most evaluation studies (with notable exceptions such as developmental evaluation), they are, by themselves, inadequate to support adaptive management of an ongoing intervention. Hence, it makes sense to plan for and implement M&E activities in a manner that draws on their respective strengths.

As part of a plan that integrates M and E, decide *when* an evaluation should begin and end.

Once the decision to evaluate has been made, deciding the timing is largely determined by what decisions the evaluation is intended to inform and when the evaluation findings will be needed to be able to do so.

Many organizations refer to mid-term and end-of-term (or final) evaluations. These terms should **not** be literally interpreted as ‘mid-way’ and ‘at the end’ of the intervention implementation period.

A mid-term evaluation often needs to be undertaken very early on (well before the mid-point of a project) –especially with new interventions where it is important to investigate and ensure the quality of implementation.

An end-of-term evaluation might need to be undertaken well before the end if it is intended to inform a decision about whether or not to continue the funding or scale up an intervention. Or, it might need to be undertaken some time after an intervention ends in order to follow up longer-term impacts and the sustainability of results achieved during implementation.

Managers should think through the use of the evaluation findings and decide when it is most appropriate to conduct the evaluation. Mid- and end-of-term/final evaluations can be usefully defined as:

- ***mid-term evaluation*** –primarily intended to inform improvement of implementation. The aim is to maximize the potential for achieving the intended results at the end of the intervention and identifying lessons learned about implementation to inform future interventions. These evaluations can identify (early signs of) unintended, positive and negative, results.
- ***end-of-term or final evaluation*** –primarily focus on project or program results and how and why they were achieved (or not) to inform decisions such as whether to continue the intervention, to improve it, to scale it up or replicate it elsewhere. They can also be used to identify lessons learned to guide implementation and improve results in future interventions.

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation timeline

Resources

- What is the timeline for the evaluation activities? (Archive link)

  Developed by the Office for Learning and Teaching of the Australian Government.

Decide whether the evaluation will be done by an external team, an internal team or a hybrid of both

Expertise, impartiality, cost, and time are key issues in deciding who will conduct the evaluation.

The section below lists specific trade-offs in the decision-making about engaging internal or external evaluators:

- Perspective
  - Internal evaluator(s): May be more familiar with the community, issues and constraints, data sources, and resources associated with the project/program (i.e., insider perspective).
  - External evaluator(s): May bring fresh perspective, insight, broader experience, and recent state-of-the-art knowledge (i.e., outsider perspective).
- Knowledge and skills
  - Internal evaluator(s): Are familiar with the substance and context of research for development programming.
  - External evaluator(s): May possess knowledge and skills that internal evaluators are lacking but it may be difficult to find evaluators who understand the specifics of research for development programming.
- Buy-in
  - Internal evaluator(s): May be more familiar with the project/ program staff and may be perceived as less threatening, and, thus, may have greater buy-in and staff involvement in the evaluation.
  - External evaluator(s): May be perceived intrusive or a threat to the project/program (i.e., perceived as an adversary) and, thus, may have more difficulty obtaining relevant information.
- Stake in the evaluation
  - Internal evaluator(s): May not be seen as an honest broker but may be perceived as having an agenda / stake in the evaluation.
  - External evaluator(s): Can serve more easily as an arbitrator or facilitator between stakeholders as perceived as neutral.
- Credibility
  - Internal evaluator(s): May be perceived as biased when perceived as ‘too close’ to the subject matter which may result in lesser credibility of the evaluation hindering its use.
  - External evaluator(s): May provide a view of the project/program that is considered more objective, not part of the organization’s power structure, and thus, give the findings more credibility and potential for use.
- Resources
  - Internal evaluator(s): May use considerable staff time which is always in limited supply, especially when their time is not solely dedicated to the evaluation.
  - External evaluator(s): May be more costly and still involve substantial management time from the commissioning organization’s staff.

- Follow-up / Use of evaluation findings
    - Internal evaluator(s): More opportunity and authority to follow up on recommendations of the evaluation.
    - External evaluator(s): Contracts often end with the delivery of the final product, typically the final evaluation report which limits or prohibits follow-up. As outsiders, do not have authority to require appropriate follow-up or action.

It is advisable to engage an external evaluator / evaluation team when:

- The scope and/or complexity of the evaluation demand expertise that is not internally available;
- A program or project is politically sensitive and impartiality is a key concern; or,
- Internal staff resources are scarce and timeframes are particularly pressing (i.e., there is little flexibility in terms of evaluation timing).

External evaluators may be an individual, a research institute or a consulting firm.

Planning and executing an evaluation is, in any case, a team effort. A critical decision to be made at this stage is who will lead the evaluation. Consider the following options for assembling the evaluation team:

- External evaluator(s) –one of them serving as the team leader– supported by program staff
- Internal evaluator(s) –one of them serving as the team leader– supported by program staff
- An internal evaluator –serving as the team leader– supported by other internal evaluators and program staff but also external evaluator(s)

Even if an external evaluator is hired to conduct the evaluation, the program manager and other staff must be involved in the evaluation process. Staff are not only primary users of the evaluation but also participants in data collection (such as providing access to records, educating the evaluator about the project/program or being interviewed as a key informant) and/or other evaluation-related tasks. Be realistic about the amount of time needed for this involvement so staff schedules do not get over-burdened.

Although hiring an external evaluator may seem costly, it may, ultimately, be less expensive than channeling considerable staff time into the evaluation. A careful analysis of staff time costs compared to external consultant costs is needed before making a decision.

For partnership evaluations (i.e., co-funded by more than one organization), double-check whether any of the co-funders have a requirement for using a ‘third party’ evaluator (i.e., someone who is not affiliated, in any way, with any of the organizations involved) (see also Step 1).

IDRC-specific information

IDRC uses internal (i.e., in-house staff) or external evaluators or a mix of both. See: Selecting an Evaluation Consultant or Team. Evaluation Guideline, February 2012.

IDRC external reviews between 2010 and 2015 included both a self-assessment and external review component.

Resources

- Decide who will conduct the evaluation
  Clarify who will actually undertake the evaluation. This might include people who are involved in what is being evaluated (such as implementers, clients and community members), an internal or

external evaluator, or some combination of these.

Determine the evaluator qualities

Different types of evaluation (e.g., impact assessment, action-oriented evaluation) will have different demands in terms of technical and other expertise and the degree of ‘distance’ between the evaluator and the subject.

The section below lists some essential evaluator characteristics or qualities matched to the main purpose (or use) of the evaluation - accountability, learning or innovation:

Accountability

When the main purpose of the evaluation is accountability, the emphasis is on determining the worth or merit of a project/program. Some essential evaluator qualities are:

- Should possess qualitative and quantitative expertise and experience.
- Independence and credibility is of central importance.

Learning

When the main purpose of the evaluation is learning, the emphasis is on facilitating project / program improvements. Some essential evaluator qualities are:

- Must be reflective, familiar and comfortable with concepts of adult education and organizational learning, and willing and able to take the role of facilitator.
- Should possess qualitative and quantitative expertise and experience.

Innovation

When the main purpose of the evaluation is innovation, the emphasis is on facilitating the design of new projects/programs based on what works. Some essential evaluator qualities are:

- Should be a strong leader, but also a team player.
- Should possess good analytical skills.

General qualities

Generally, the following basic qualities should be considered:

At the individual level:

- Quantitative and/or qualitative research skills
    - Evaluation experience and expertise
    - Sensitivity to the project’s principles (e.g., empowerment, participatory action, capacity-building)
    - Ability to effectively communicate to the targeted users and audiences
    - Independence

Across the team:

- Quantitative and qualitative research skills
- Multidisciplinary skills (e.g., economic, demographic, environmental, sociological)
- Thematic and contextual knowledge and experience
- Research for development expertise and experience
- Gender and cultural balance
- Language skills
- Familiarity with the organization and its partners

Regardless of the specific evaluator qualities needed to support a quality implementation of the evaluation, there are also important general characteristics: flexibility, ability to problem solve, and credibility.

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- List of evaluator qualities

IDRC-specific information

Evaluators of IDRC programming typically need to have these types of knowledge and skills:

A broad knowledge of the relevant fields of research, policy and practice including issues, trends, institutions, networks and policy communities in the low-and middle-income countries of sub-Saharan Africa, Asia, and Latin America and/or the Caribbean.

Specific knowledge and experience in evaluating the results of development research and the contexts in which research programming takes place, for instance:

- the influence of research and other evidence on changing policy and practice
- capacity development in various aspects of conducting and using development research
- the quality of development research –going beyond traditional academic measures such as bibliometrics
- equity-focused research programming –being sensitive to different dimensions of marginalization
- research programming in conflict-affected settings
- scaling up innovations tested through research to products, processes, businesses
- results of research through a value chain
- research communications, knowledge translation

Identify what resources are available for the evaluation and what will be needed

It is important to develop an estimate of the resources that are available for evaluation and what will be required to do the evaluation well.

The resources needed for an evaluation include:

- Existing data
- Funding to engage an external evaluator or evaluation team or pay for specific tasks to be undertaken and for materials and travel
- Time, expertise and willingness to be involved of staff, partners, technical experts and the wider community, whether as part of the evaluation team, the evaluation governance processes and/or key informants and data sources

When considering what data are already available, look carefully at the quality of existing data and what format it is in.

Also, clarify the skills and availability of any people who will need to be involved in the evaluation.

There are five main ways of developing an estimate of the budget for an external evaluation:

1. **Calculating a percentage of the program or project budget – sometimes 5%-10%.** This is a very crude rule of thumb. Large programs with simple evaluation might need a lot less; small programs with large evaluations – for example, detailed testing and documentation of an innovation – will need much more. It is also better to target evaluation resources across programs where they will be most useful.
2. **Developing an estimate of days needed and then multiplying by the average daily rate of an external evaluator.** This is only useful for simple evaluations, especially those using a small team and a standardised methodology such as a few days of document review, a brief field visit for interviews and then a short period for report write up.
3. **Using the average budget for evaluations of a similar type and scope.** This can be a useful starting point for budget allocation providing that the amounts have been shown to be adequate in the past – otherwise this will perpetuate the problems of underestimates.
4. **Developing a draft design and then costing it, including collection and analysis of primary data.** This can be done as a separate project before the actual evaluation is contracted.
5. **Consider the following options if ongoing evaluation input is needed such as for a Developmental Evaluation**: retainer fee contracts; stepwise funding; or, speculate and allow for contingencies.

- More information about these options can be found in this interview with Michael Quinn Patton:
- Budgeting for Developmental Evaluation
  PDF
  98.49 KB

Allow time to secure resources (for example, including them in an annual or project budget, or seeking someone with particular expertise). If the resources required for the evaluation are more than the resources available, additional resources will need to be found and/or strategies used to reduce the resources required, such as reducing the scope of the evaluation.

It may also be useful to consider ballpark figures for similar types of evaluations. For example, a paper from the Coalition for Evidence-Based Policy suggests it may be feasible to run a randomized controlled trial for an impact evaluation which largely draws on existing data for $50,000 to $300,000.

Possible products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Estimated cost of an evaluation (ideally including in-kind, cash, internal and external costs)

- Justification for expenditure on evaluation
- Statement of available evaluation resources (including budget)

Method

- Strategies to reduce costs

  Reducing costs is something to consider if evaluation costs outweigh the predicted benefits or available resources.

Resources

- **A checklist for developing and evaluating evaluation budgets**
  PDF
  42.64 KB

  This checklist, developed by Jerry Horn as part of the Western Michigan University Evaluation Checklists project, provides useful prompts about items that might need to be included and how they might be estimated.

- **Evaluation budget guide**
  PDF
  100.9 KB

  Blank evaluation budget from CARE Uganda listing many items that might be included.

- Developing an evaluation budget estimate

  Evaluation budget templates in Excel from the National Institutes of Health with different versions for different situations in terms of labour costing (with or without fringe benefits) and travel.

Terms of reference

Develop the formal document that outlines the requirements for the evaluation.

The Terms of Reference (ToR) or Request for Proposal (RFP) are an explicit statement of the resources, roles and responsibilities of the evaluators and the evaluation commissioner or manager including:

- *why* and *for whom* the evaluation is being done;
- *what* it intends to accomplish;
- *how* it will be accomplished;
- *who* will be involved in the evaluation; and,
- *when* milestones need to be reached including when the evaluation needs to be completed.

GeneraTOR - Terms of reference generator
The GeneraTOR is an interactive tool that helps people to write a Terms of Reference (ToR) or Request for Proposal (RFP) document. The GeneraTOR guides users to write different sections of a ToR and generates an editable document that can be downloaded and shared with others.

Step 2 already helped to prepare a lot of the information, so writing the ToR is mostly a matter of bringing the necessary information together in this document.

For an internal evaluation - carried out by staff of the organization - the ToR is often called the 'evaluation brief' or the 'evaluation agreement'. Most of the elements of a formal ToR or TFP for external evaluators are relevant to be included in an internal evaluation brief/agreement.

In addition to the specifics about the project or program and its context, the evaluation –the purpose, scope, key evaluation questions and evaluation methodology (or how they should be developed) – the ToR / RFP should also include reporting requirements, milestones or deliverables, time frames, and relevant contractual requirements.

While the ToR or RFP of any evaluation process will need to be tailored to the particulars of that study and follow the requirements of the organization, there are elements which all ToRs / RFPs should include:

1. Background
2. Purpose/objectives/rationale for the evaluation
3. Intended user(s) and use(s) of the evaluation
4. Key evaluation questions
5. The principles and approach that will guide the evaluation
6. Methodology
7. Roles and responsibilities of different actors
8. Reporting requirements (see immediately below)
9. Timeline and milestones
10. Any specific requirements

Some organizations will include an indicative or ceiling budget.

The reporting requirements for the evaluation (referred to above as section 8 in a ToR) may include:

- The desired format(s) (such oral, written, video, etc.)
- Dissemination materials (such as summary, briefs, presentation materials, newsletter, etc.)
- Intended audience(s)
- Content areas
- Desired length of the report
- Whether the report should include specific recommendations
- Whether the data sets should be returned such as completed questionnaires, surveys, interview notes and tapes, etc.)
- Mode of delivery
- Any specific restrictions or needed permissions to publish information from or based on the evaluation

What it involves

The ToR / RFP should be drafted *before* the evaluation starts. Consultants may be involved in further refining the evaluation design but the evaluation commissioner has to ensure that the ToR specifies what the evaluation needs to accomplish and what is expected from the evaluators. It is important, as part of the decision making processes, to agree on who needs to provide direct input, review and sign-off the ToR before it is released.

The following steps prepared information needed for the ToR:

- Decide whether the evaluation will be conducted internally, externally, or by a mixed team of evaluators (see Step 2 for more detailed guidance). The ToR should include all the information external evaluators need to decide if they will bid for the work.

- Determine the selection criteria for the external evaluator(s) (see Step 2 for more detailed guidance). Among the range of evaluator qualities, the ToR should be clear which qualities are 'desired' and which are 'essential'. These will serve as criteria to support a transparent process for selecting the most suitable consultant(s). Both the criteria and the process for selection of the external evaluator(s) should be specified in the ToR.

- Clarify whether the evaluation design will be developed as part of the ToR, as the first stage of the evaluation, or as a separate project (see Step 5). There are two basic scenarios for addressing the evaluation methodology in the ToR:
    - (1) The ToR specifies a particular evaluation methodology that needs to be used and the evaluator competencies can be specified accordingly; *or*
    - (2) The ToR states that the first task of the external evaluator(s) is to design the evaluation and select an appropriate methodology, in which case it is helpful, at least, to articulate some values and/or principles that need to be upheld in the evaluation process.
- Draft the ToR with the input from relevant staff
- Obtain feedback from key stakeholders

- Obtain sign-off from senior management, as per the organization's specific requirements

The ToR / RFP:

- becomes a free-standing, public document
- can be used as the basis for developing the contract with the external evaluator(s)
- can be used to support discussions with primary evaluation users about their information needs

Product

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation brief or evaluation agreement (for internal evaluations)
- Terms of Reference (ToR) / Request for Proposal (RFP) (for evaluations that include external evaluators)

Use the BetterEvaluation GeneraTOR Terms of Reference tool to write your TOR

You can use a template in a wordprocessing software to write the ToR for your evaluation or you can use the **GeneraTOR** below which will guide you through writing the different sections of a ToR / RFP. It will generate a word document with your saved information which can then be further refined and/or reviewed by others.

To access the **GeneraTOR**, login to your BetterEvaluation account and hit the "GeneraTOR" tab on your profile page.

Not yet a BetterEvaluation member? Click join and follow the instructions to create an account.

IDRC-specific information

Further guidance can be found in Writing Terms of Reference (ToRs) for an Evaluation. Evaluation Guideline Nr 5, March 2004.

Note that it is a requirement to alert evaluator(s) that the quality of the evaluation report they produce will be judged by IDRC's Evaluation Unit on four internationally recognized standards: utility, feasibility, accuracy, and propriety. A copy of "Quality Assessment of IDRC Evaluation Reports"should be given to the evaluator(s) to ensure they understand how the quality of the evaluation report will be assessed (see also Step 8).

IDRC staff can access many examples of actual ToRs used in IDRC-supported evaluations in the repository of Evaluation Terms of Reference which was compiled by the Policy and Evaluation Division.

IDRC staff can also request the Evaluation Unit to review and comment on an evaluation ToR under development.

Resources

- Terms of reference

  A Terms of Reference (ToR) document provides an important overview of what is expected in an evaluation.

- How to perform evaluations – Model ToR

  This guide produced by the Global Affairs Canada (formerly Canadian International Development Agency, CIDA) was developed to support staff to write clear and simple Terms of Reference (ToR) for evaluations.

- UNEG quality checklist for evaluation terms of reference and inception reports

  This guide from the UNEG is a checklist designed to support UNEG members in the design and implementation of evaluations.

GeneraTOR - Terms of reference generator

Welcome to the GeneraTOR

The GeneraTOR will guide you in writing the different sections of a Terms of Reference (ToR) / Request for Proposal (RFP). It will generate a Word document containing your saved information, which others can further refine and review.

You'll need to be a member to access the GeneraTOR tool. Please login or join to become a member .

How to use the GeneraTOR

The GeneraTOR organises the terms of reference into 16 sections. These sections are listed on the left side of the GeneraTor in the section menu.

Click on a section - the project information section is a good place to start. Follow the instructions in the section and write the relevant information in the text field. If the title of a text field has a red star at the end, it means you must write something in this text field.

When you have finished writing a section, click the 'section complete' box at the bottom of the section. A tick will appear by this section in the section menu on the left side of the GeneraTOR. When you return to your saved ToR later, it will be clear which sections you have finished and which still need further work.

Work your way through the sections in the section menu. Complete only the sections that are relevant to your terms of reference.

Save your work

Click 'Save Draft' at the bottom of the page to save your work. The GeneraTOR will periodically save your work automatically, but we recommend that you also manually save your work frequently. Your saved ToRs can be found on your member profile page.

Download your ToR

Your ToR can be found on your member profile page under the heading 'My Terms of References'. Click a download option from the final column to download your ToR as a word document. Once the word document has been downloaded, you can open and edit it as you would any other word document.

Work on a saved ToR

Your ToR can be found on your member profile page under the heading 'My Terms of References'. Click 'edit' from the edit column to open your ToR and keep working on it.

Automatic deletion

If your ToR is not opened for for 365 days, it will be deleted from the BetterEvaluation system.

Engage team

Evaluations can be conducted by in-house staff (internal) or a third party (external) or a mix of both.

Advertising the Terms of Reference (ToR) / Request for Proposal (RFP) is the first step in engaging external evaluators but other important issues need to be addressed to ensure an effective and transparent selection process and to orient the selected evaluator(s).

Once the decision to use external evaluator(s) has been made, sufficient time should be allowed to engage them. Good evaluation consultants are typically busy, so provide enough lead time for them to fit potential new work into their schedules.

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Call for Expression of Interest (IoE)
- Evaluation proposal
- Evaluator(s) selection criteria for shortlisting
- Evaluator(s) final selection questions
- Reference check questions
- Consultant contract
- Briefing book or project documentation center

IDRC-specific information

IDRC staff should consult internal guidance from Procurement about identifying and selecting external consultants. Decisions start with whether to use a competitive process or not. This page http://intranet.idrc.ca/en/ev-94737-201-1-DO_TOPIC.html houses decision-making flow charts to determine what kind of process you should follow, as well as detailed guidance and forms. Involve Procurement early on in making these decisions and ask about the specific forms you will need to fill out if you will be using a competitive process. A lot of what you have already developed for the Terms of Reference will transfer smoothly into the required RFP form. You can contact Procurement colleagues via: purchase_achats@idrc.ca .

- Selecting an Evaluation Consultant or Team. Evaluation Guideline, February 2012.

Advertise the evaluation Terms of Reference (ToR) / Request for Proposal (RFP)

There are different approaches to advertising the evaluation ToR / RFP. These are typically based on the organization's specific procurement procedures and often linked to the cost of the work.

Sources for prospective bidders include: other agencies that have used external evaluators, local universities, research institutes, consulting firms, professional evaluation associations.

For an **open call** (i.e., one that seeks competitive bids), advertising can be done in local newspapers, through online professional networks or listserves, agency newsletters etc. Build in as much time as possible between the posting and the deadline for submissions to maximize the response.

If there is sufficient time, a two-step process may be considered. Firstly, an **Expression of Interest (EoI)** is solicited which often only includes requests to submit a letter of intent and CVs. Subsequently, a **full proposal** is requested from those selected on the basis of the EoI. The first step helps to narrow down the pool of prospective candidates for the second step to those most likely to match the requirements and, thus, reduces the amount of work in reviewing full proposals. However, many organizations use EoIs in different ways and for different reasons.

For a **closed call**, advertising is targeted at a limited number of consultants or organizations / institutions pre-selected in a certain manner, often for their particular expertise needed for the evaluation but also as a means to simplify the bidding process when a need to contract out services arises.

Prospective bidders are encouraged to ask questions about the ToR / RFP to clarify requirements. There are usually specific rules about how this process is managed including time frames for questions and

whether and how answers are shared.

Products

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Call for Expression of Interest (EoI)
- Terms of Reference (ToR) or Request for Proposal (RFP)

IDRC-specific information

A publicly announced request for proposals (RFP) for evaluation contracts are hosted on a central Government of Canada website. In addition, announcements are posted on listserves and websites such as:

- Monitoring and Evaluation News
- Peregrine discussion groups
- sectoral listserves (e.g., agricultural research listserve, peace research networks); the outcome mapping learning community; regional evaluation listserves (e.g., African Evaluation Association, the Community of Evaluators South Asia, ReLAC)

Commissioners may also want to talk to other (similar) organizations to get recommendations for evaluators to contact or talk to their grantees for suggestions.

Select an evaluator / evaluation team

A transparent selection process uses explicit criteria and involves more than one person (often a panel or committee) to discuss and agree on the selection.

The selection criteria and process as well as who will sign-off should, ideally, be discussed and agreed early on in the evaluation process (see Step 1 for further information).

Shortlisting potential candidates

Applications need to be reviewed and a shortlist of potential candidates determined. Typically, applications are ranked in terms of the degree to which they match the selection criteria. Ranking is facilitated by carefully considering the relative importance of the selection criteria for the work at hand. In the first instance, use the ‘essential’ and ‘desired’ qualities but further distinctions may also be needed (such as specific content expertise or previous work in a particular context or with certain stakeholders) (see list of evaluator qualities).

Final selection of the evaluator(s)

The shortlisted candidates are then further queried through a follow up interview (face-to-face or by phone/internet) or other communication means. Decide who will be on the selection panel and what questions will be asked. For example, the following questions may be helpful:

- Knowledge and experience – Does the bidder have the market, industry and/or professional experience to meet your agency's needs?
- Team composition – Are the team members listed in the proposal the people who will actually do the work? Is the team known to your agency?
- Business values and policies – Does the bidder reflect your agency's business values and expectations of quality?
- Understanding the job – How well does the bidder understand the size and scope of the job? Are they up to the challenge?
- Innovative and creative thinking – Will the bidder be able to respond to emerging issues throughout the evaluation?

[The above list of questions were adapted from: Choosing the right consultant. In: Policy Makers Toolkit. Steps in Managing and Evaluation Project. Premier & Cabinet, NSW Government, Australia.]

The selection of a team leader is particularly crucial to the success of the evaluation. A good team leader is results-oriented and focused on management objectives. She/he must have demonstrated ability to manage and synthesize the input and participation of the various team members as well as the range of stakeholders involved. Good communication skills –both verbal and written– and good facilitation and negotiation skills are essential.

Determining potential conflict of interest or safeguarding an appropriate level of 'independence'

Another important issue in the final selection is to determine if the individual or their organization has any potential conflict of interest which may compromise the credibility of the evaluation. Criteria for excluding certain individuals/organizations from undertaking the evaluation could include:

- not having been involved in the design, implementation or oversight of the activities;
- not having received funding from the organization / program in the recent past (such as at least not in the last two years)'
- not having a stake in what the evaluation uncovers.

Conducting reference checks

As the final step before hiring an evaluator, look at their past work, check references, and make sure you feel comfortable that they are credible, competent, and capable of leading a useful evaluation process. Some key issues that can be discussed in the reference check are:

- Timeliness – Did they meet the milestones and deliver the product on time?
- Responsiveness – Were they flexible enough to deal with unexpected challenges or delays? Did they respond to the project / program team in an appropriate manner?
- Relevance – Did they follow the terms of reference?
- Professionalism – What was their work style, communication ability, or degree of cultural sensitivity?
- Evaluator good practice – Did they engage openly with the project / program team? Did they proactively explain their decisions based on good practice in evaluation?
- Evidence-based conclusions – Were the report's conclusions evidence-based or mainly conjecture?

[The above list of questions was adapted from: Willard A. Managing and implementing an evaluation. Guidelines and Tools for Evaluation Managers. USA: American Red Cross, Catholic Relief Services, USAID, 2008.]

Resources

- Checklist for selecting an evaluator (pdf download)

  W.K. Kellogg Foundation Evaluation Handbook (page 63) , 2004.

- Choosing the right consultant

  Premier & Cabinet, NSW Government, Australia.

Prepare the contract

Your organization probably has standard requirements and templates for consulting contracts.

Generally speaking, the contract needs to clarify:

- Who will perform the evaluation tasks, the level of contact and expectations about communication between the evaluator(s) and the funder and/or the project/program (such as formal meetings, written progress updates etc.), specific milestones and/or deliverables, and time frames;
- Agreed total cost (with or without specific costing categories such as professional fees, travel, materials) and payment schedules;
- Who owns the evaluation information and to whom it can be released (including the data, the reports and any other types of publications including in professional journals); and,
- Any legal issues such amendments to the contract, resolving disputes or conditions for terminating the contract.

The ToR can be used as the basis for a contract with the external evaluator(s)

Any negotiations between the consultant and the commissioning organization are best left to the contracting specialist and the project manager. It is useful if the evaluation manager knows if there were particularly contentious issues, but she/he should not be involved in negotiations as this may complicate her/his relationship with the consultant during the evaluation process.

Products

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Consultant contract

IDRC-specific information

The evaluation commissioner should work closely with GAD or the Procurement Department to issue the consultant contract.

Orient the evaluator / evaluation team

The evaluator(s) need a clear understanding of what needs to be evaluated.

They need to obtain as much background information on the project / program as possible, its implementers, the intended beneficiaries, and its specific context.

One of the responsibilities of the evaluation manager is ensure the evaluator(s) have access to all relevant project / program documentation and available data sources (e.g., project monitoring indicators). If the documentation and data sources are well-organized, the evaluator(s) can start their work more quickly rather than having to spend time identifying and compiling materials. If these materials are not easily accessible, the evaluation manager can work with the project / program manager to streamline the documentation process and engage relevant staff in bringing the documentation together.

The ToR / RFP, typically, provides the reporting requirements (see Step 3) for the evaluation but it is useful to discuss these in more detail during the orientation of the evaluator(s) to ensure they are fully understood.

Products

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Project / program documentation: core documentation briefing book, bibliography and/or documentation center including hard and/or electronic copies

Evaluation design

An evaluation design describes how data will be collected and analysed to answer the Key Evaluation Questions.

There are different pathways for you as manager depending on who will develop the evaluation design. In most cases your evaluator will develop the evaluation design. In some cases you will – if you have evaluation expertise and/or the evaluation design has already been developed (for example, in an evaluation that is intended to match an earlier evaluation).

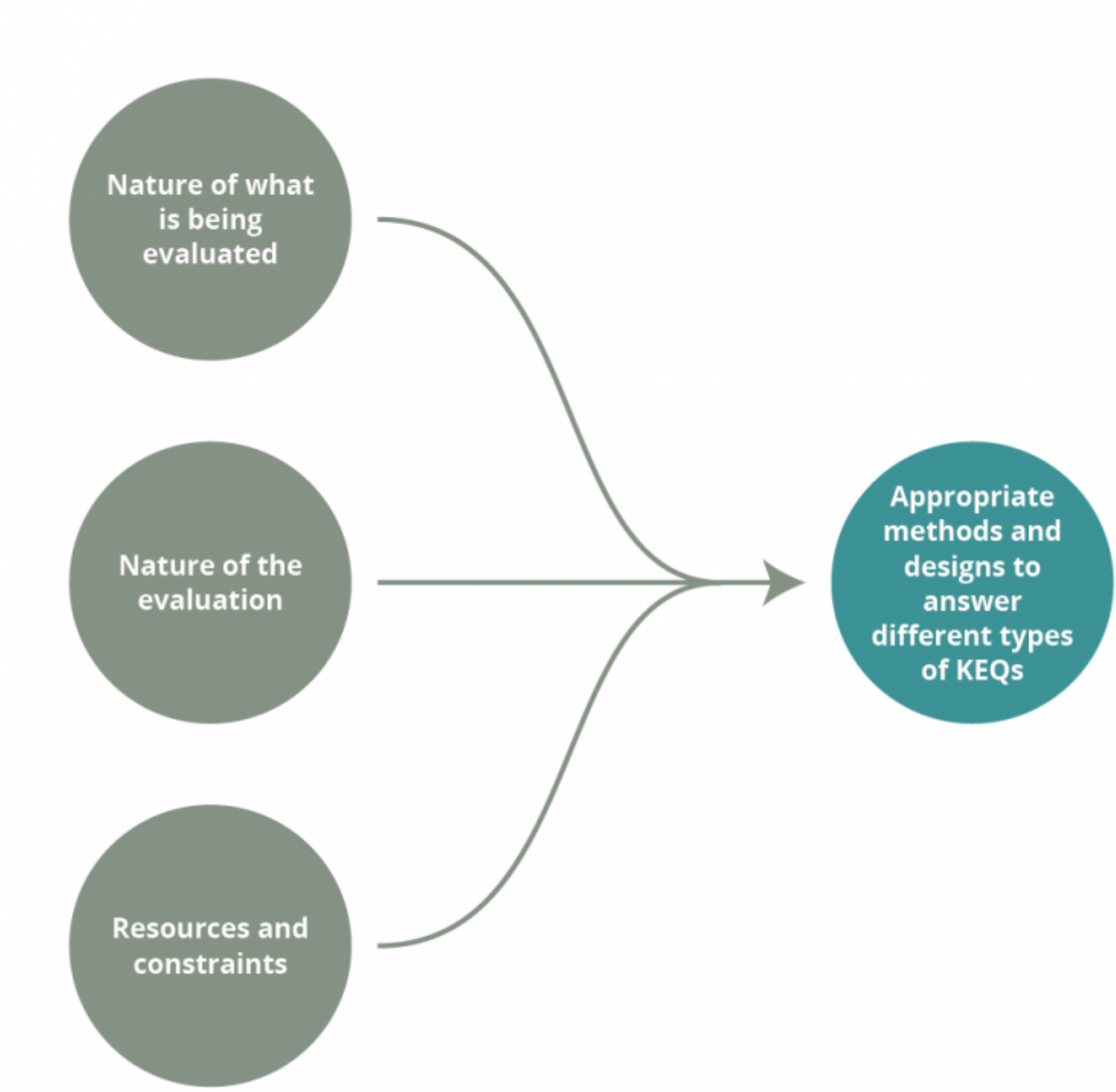
Take into account the following important factors when developing an evaluation design

1. **The nature of what is being evaluated**
   In particular whether there are complicated or complex aspects that need to be addressed in the evaluation, and other particular challenges such as delays before impacts are evident or barriers to collecting accurate data.

Manager's guide to evaluation :
Consider important elements of what is being evaluated

2. **The nature of the evaluation**
   In particular the types of Key Evaluation Questions that are being asked, when the answers are needed.
   Manager's guide to evaluation :
   Consider important aspects of the evaluation

3. **Available resources and constraints**
   Resources including money, existing data, expertise, technical equipment. Constraints including requirements to use certain common indicators, limits to availability of key informants or barriers to accessing existing data.
   Manager's guide to evaluation :
   Consider the implications of the resources available and specific constraints



If an EVALUATOR will develop the evaluation design

- Engage a competent evaluation expert - internal, external or a combination. (See 'Select an evaluator / evaluation team' for advice).
- Work with the expert(s) to ensure they understand important factors that should be taken into account in the evaluation design (see section above)
- The design should provide details of how data will be collected analysed. It is often useful to do this in the form of an Evaluation Matrix which shows how each Key Evaluation Question will be answered.

If YOU (as manager) will develop the evaluation design

- Understand important factors that should be taken into account in the evaluation design (see section above)
- Develop an evaluation design that addresses these important factors.
- Summarise the design in the form of an Evaluation Matrix which shows how each Key Evaluation Question will be answered.

Subsequently, arrange for a **technical review of the evaluation design** and arrange for a **review of the design by the evaluation management structure** (e.g., steering committee). Ideally this will include representation from primary intended users.

Arranging technical review of the evaluation design

Before finalizing the design, it can be helpful to have a technical review of it by one or more independent evaluators. It might be necessary to involve more than one reviewer in order to provide expert advice on the specific methods proposed, including specific indicators and measures to be used. Ensure that the reviewer is experienced in using a range of methods and designs, and well briefed on the context, to ensure they can provide situation specific advice.

Arranging review of the design by the evaluation management structure

In addition to being considered technically sound by experts, it is essential for the evaluation design to be seen as credible by those who are expected to use it.

Get formal organisational review and endorsement of the design by an evaluation steering committee (see ' Identify who will be involved in decisions and what their roles will be' for possible structures, processes and terms of reference for a steering committee)

Where possible do data rehearsal of possible findings with primary intended. This is a powerful strategy for checking the appropriateness of the design by presenting mock-ups of tables, graphs and quotes that the design might produce. It is best to produce at least 2 different versions – one that would show the program working well and one that would show it not working. Ideally the primary intended users of the evaluation will review these and either confirm the suitability of the design or request amendments to make the potential findings more relevant and credible. (For more information see Patton, MQ (2011) Essentials of Utilization-Focused Evaluation. pp. 309-321).

Consider important elements of what is being evaluated

What is being evaluated makes a difference to how it should be evaluated. It is helpful to identify particular aspects of what is being evaluated and check that these have been addressed in the evaluation design.

1. Check the stage of development of the project or program

Firstly, check the implications of the stage of development of the project or program that is being evaluated. Is it still being planned? Is it part–way through implementation? Or is it near the end – or has it in fact already ended?

Stage of development	Consequence	Possible implication for the evaluation design
Not yet started	Can set up data collection from the beginning of implementation	*Possible to gather baseline data as a point of comparison and also to establish comparison groups or control groups from the beginning*
		Opportunity to build some data collection into administrative systems to reduce costs and increase coverage
	Period of data collection will be long	*Need to develop robust data collection systems including quality control and storage*
Part way through implementation	Cannot get baseline data unless this has already been set up	*Will need to construct retrospective baseline data to estimate changes that have occurred*
	Might be able to identify “bright spots” where there seems to be more success and those with less success	*Scope to do purposeful sampling and learn from particular successes and also cases which have failed to make much progress*
Almost completed	Cannot get baseline data unless this has already been set up	*Will need to construct retrospective baseline data to estimate changes that have occurred*
	Depending on timeframes, some outcomes and impacts might already be evident	*Opportunity to gather evidence of outcomes and impacts*
Completed	Cannot get baseline data unless this has already been set up	*Will need to construct retrospective baseline data to estimate changes that have occurred*
	Depending on timeframes, some outcomes and impacts might already be evident	*Opportunity to gather evidence of outcomes and impacts*
	Cannot directly observe implementation	*Will need to depend on existing data or retrospective recollections about implementation.*

2. Is it complex or complicated?

Secondly, consider whether there are important aspects that are either complicated (with many components) or complex (emergent) that should be addressed in the evaluation design.

(i) Focus

Does everyone share the same objectives?

Homogeneity of objectives	*Implications*
Everyone shares a single set of objectives	*Impacts to be included can be readily identified from the beginning.*
There are different objectives valued by different stakeholders (competing objectives, different objectives at different levels)	*Need to identify and gather evidence about multiple possible changes* *Need an agreed way to weight or synthesise results across different domains to produce a judgement of overall performance.*
The stated objectives are changing (often in response to changing needs or opportunities)	*Need nimble impact evaluation systems that can gather adequate evidence of emergent intermediate outcomes or impacts*

(ii) Management

Who has responsibility for management and decision making?

Who is responsible	*Implications*
Single organisation	*Primary intended users and uses easy to identify and address in the development of Key Evaluation Questions and endorsement of the design*
Multiple organisations (which can be identified) with specific, formalized responsibilities	*Likely to need to negotiate access to data and ways to link and co-ordinate data* *Might need to negotiate parameters of a joint impact evaluation, including negotiating scope and focus.*

Who is responsible	*Implications*
Changing list of organizations working together in flexible ways	*Need nimble impact evaluation systems that can gather evidence about the contributions of emergent actors and respond to the different ways they value intended and unintended impacts*

(iii) Consistency

How much variability is there in how activities are implemented?

Level of variability	*Implications*
Standardized – one-size-fits-all program	*Quality of implementation should be investigated in terms of compliance with 'best practice'.*
Adapted – variations of a programme planned in advance and matched to pre-identified contextual factors.	*Quality of implementation should be investigated in terms of compliance with the practices prescribed for that type of situation.*
Adaptive – evolving and personalised program that responds to specific and changing needs.	*Quality of implementation should be investigated in terms of how responsive and adaptive service delivery was.*

(iv) Necessity

How many different options are there for solving the problem or producing the intended impacts? To what extent is this exact initiative needed to solve the problem?

Number of possible interventions	*Implications*
There is only one way to achieve the intended impacts.	*Counterfactual reasoning appropriate.*
The intervention is one of several ways of achieving the impacts, and the options can be identified.	*Counterfactual reasoning not appropriate as it does not accept a causal relationship between the intervention and the impacts unless they would not have occurred in the absence of the intervention.*

Number of possible interventions	*Implications*
Possibly one of several ways of achieving the intended impacts (uncertain).	*Counterfactual reasoning not appropriate as it does not accept a causal relationship between the intervention and the impacts unless they would not have occurred in the absence of the intervention.*

(v) Sufficiency

To what extent will the problem be solved by the intervention alone?

Generalisability of the intervention	*Implications*
The intervention is enough to produce the intended impacts. Works the same for everyone.	*Counterfactual reasoning appropriate* *Reasonable to ask 'Does it work?'*
Works only in specific contexts which can be identified (eg implementation environments, participant characteristics, support from other interventions).	*Impact evaluation question needs to be 'For whom, in what circumstances and how does it work?'* *Counterfactual reasoning only appropriate if the causal package of supportive context and other activities can be identified and included.*
Works only in specific contexts which are not understood and/or not stable.	*Impact evaluation question needs to be 'For whom, in what circumstances and how does it work?'* *Counterfactual reasoning not appropriate as the causal package of supportive context and other activities is changing and/or poorly understood and cannot be adequately identified.*

Change trajectory

How are the impact variables expected to change over time? For example, straight line of increase, or J curve? To what extent are the relationships between variables understandable and predictable?

Relationship between variables	*Implications*

Simple relationship (cause and effect). Predictable.	*Measurement of change can be done at a convenient time and confidently extrapolated*
Complicated relationship that needs expertise to understand and predict.	*Timing of the measurement of changes should be undertaken when it will be most meaningful – expert advice will be needed.*
Emergent factors and multiple causes, sudden changes (tipping points) that are unpredictable. Can only be understood in retrospect.	*Changes will need to be measured at multiple times as the change trajectory cannot be predicted.*

Unintended impacts

To what extent are unintended impacts predictable?

Predictability of unintended impacts	***Implications***
Easily predictable and therefore can be readily included in the data collection plans	*Need to draw on previous research and common sense to identify potential unintended impacts and gather data about them*
Need expertise to predict and address.	*Need advice from experts about potential unintended impacts and how these might be identified.*
Unpredictable - only identified and addressed when they occur.	*Need to include a wide net of data collection that will catch evidence of unexpected and unanticipated unintended impacts.*

Source: Resource Hub for Evaluating C4D 2016 - adapted from Funnell and Rogers (2011), pp.90-91, Rogers 2016.

3. Identify issues to be addressed

Are any of the following issues present? They will need to be addressed in the design.

Issue	**Possible implications for the evaluation design**

Long time until impacts will be evident	Might need to gather data about intermediate outcomes (that will be evident during the timeframe of the evaluation) and use other research and evaluation evidence to predict the likely achievement of impacts
Difficulty observing implementation activities (eg conflict affected or remote areas)	Might need to gather data through remote sensing, key informants, big data or crowdsourcing
Difficulty observing results (outcomes, impacts) (eg sensitive issues, private behaviour)	Might need to gather data through key informant interviews, or unobtrusive measures (for example looking at patterns of wear from foot traffic) or techniques for gathering sensitive data (for example polling booth)

Consider important aspects of the evaluation

Evaluations are designed to answer the Key Evaluation Questions. Different types of questions need different methods and designs to answer them.

In evaluations there are four main types of questions:

Descriptive questions ask about what has happened or how things are – for example:

- What were the resources used by the program directly and indirectly?
- What activities occurred?
- What changes were observed in conditions or in the participants?

Causal questions ask about what has contributed to changes that have been observed – for example:

- What produced the outcomes and impacts?
- What was the contribution of the program to producing the changes that were observed?
- What other factors or programs contributed to the observed changes?

Evaluative questions ask about whether an intervention can be considered a success, an improvement or the best option and require a combination of explicit values as well as evidence – for example:

- In what ways and for whom was the program successful?
- Did the program provide Value for Money, taking into account all the costs incurred (not only the direct funding) and any negative outcomes.

Action questions ask about what should be done to respond to evaluation findings – for example:

- What changes should be made to address problems that have been identified?
- What should be retained or added to reinforce existing strengths?
- Should the program be refunded?

Key Evaluation Questions often contain more than one type of questions – for example to answer the KEQ “How effective has the program been?” requires answering:

Descriptive questions – What changes have occurred?

Causal questions – What contribution did the intervention make to these changes?

Evaluative questions – How valuable were the changes in terms of the stated goals – taking into account the types of changes, the level of change and the distribution of changes.

Check the adequacy of the design by disaggregating each KEQ into the different types of questions and then checking them against the following points.

(i) Checking the adequacy of the design for descriptive questions

The design should make it clear how descriptive questions will be answered. These descriptive questions might relate to:

- Inputs – materials, staff
- Processes – implementation, research projects
- Outputs – eg research publications
- Outcomes – eg changes in policy on the basis of research
- Impacts – eg improvements in agricultural production

It can be helpful to set this out in a table that shows how data will be collected and analysed to answer these descriptive questions.

Descriptive question	Existing data that can be used	Additional data collection/retrieval	Sampling Analysis
What has been the level of resources used for the program?			
Who has participated in the program?			
What changes have occurred in terms of [specific behaviour]?			

The narrative should explain the choices made, addressing:

- Making maximum use of existing data – including a review of the quality and relevance of this
- Appropriate sampling – whether of people, sites, organisations or time periods – what type of sampling has been chosen and why this is appropriate for the type of generalization that will be undertaken.
- Appropriate data collection methods – why these methods have been chosen
- Appropriate data analysis methods – why these methods have been chosen

(ii) Checking the adequacy of the design in terms of evaluative questions

Many evaluations do not make explicit how evaluative questions will be answered – what the criteria will be (the domains of performance), what the standard will be (the level of performance that will be considered adequate or good), how different criteria will be weighted. A review of the design could check each of these in turn:

- Are there clear criteria for this evaluative question?
- Are there clear standards for judging the quality of performance on each criterion?
- Is there clarity about how to synthesize evidence across criteria? For example, is it better to have some improvement for everyone or big improvements for a few?
- Are the criteria , standards and approach to synthesis appropriate? What has been their source? Is further review of these needed? Who should be involved?

Ideally an evaluation design will be explicit about these, including the source of these criteria and standards. They might be set out in a table such as the following.

Table 1: Example table setting out the evaluative criteria, standards, synthesis process and sources

Evaluative aspect	Process for developing agreed standards, criteria and synthesis	Criteria	Standards	Synthesis/Weighting
Adequacy of resources for the program	Using national standards for the provision of services	Number of [services] per 100,000 people	[x] per 100,000 people	Average across all regions, weighted for population
	National Service Standards	Financial accessibility	All people able to access services regardless of ability to pay	
Quality of services provided		Cleanliness	Food handling surfaces free from contamination	
	Community consultation	Cultural appropriateness	People from all ethnic backgrounds feel welcome in the service	

(iii) Checking the adequacy of the design in terms of causal questions

Many evaluations do not make clear how causal questions will be answered. There are many designs and methods that can be used, but they involve one or more of these strategies:

(a) Compare results to an estimate of what would have happened if the program had not occurred (this is known as a counterfactual).

This might involve creating a control group (where people or sites are randomly assigned to either participate or not) or a comparison group (where those who participate are compared to others who are matched in various ways). Techniques include:

- *randomised controlled trials (RCTS)* – a control group is compared to one or more treatment groups
- *matched comparisons* - participants are each matched with a non-participant on variables that are thought to be relevant. It can be difficult to adequately match on all relevant criteria
- *propensity score matching* – creates a comparison group based on an analysis of the factors that influenced people's propensity to participate in the program
- *regression discontinuity* - compares the outcomes of individuals just below the cut-off point with those just above the cut-off point.

(b) Check for consistency of the evidence with the theory of how the intervention would contribute to the observed results

This can involve checking that intermediate outcomes have been achieved, using process tracing to check each causal link in the theory of change, identifying and following up anomalies that don't fit the pattern, and asking participants to describe how the changes came about.. Techniques include:

- *contribution analysis* – sets out the theory of change that is understood to produce the observed outcomes and impacts and then searches iteratively for evidence that will either support or challenge it.
- *key informant attribution* – asks participants and other informed people about what they believe caused the impacts and gathers information about the details of the causal processes
- *qualitative comparative analysis* - compares different cases to identify the different combinations of factors that produce certain outcomes
- *process tracing* - a case-based approach to causal inference which focuses on the use of clues within a case (causal-process observations, CPOs) to adjudicate between alternative possible explanations. It involves checking each step in the causal chain to see if the evidence supports, fails to support or rules out the theory that the program or project produced the observed impacts
- *qualitative impact assessment protocol* – combines key informant attribution, process tracing and contribution analysis, using interviews undertaken in a way to reduce biased narratives

(c) Identify and rule out alternative explanations

This can involve a process to identify possible alternative explanations (perhaps involving interviews with program sceptics and critics and drawing on previous research and evaluation, as well as interviews with participants) and then searching for evidence that can rule them out.

While technical expertise is needed to choose the appropriate option for answering causal questions, as manager you should be able to check there is an explicit approach being used, and seek technical review of its appropriateness.

Causal relationship (between one variable and another – one step in the causal chain)	What strategies and methods/designs are being used for causal inference
eg Participation in program and improved health and wellbeing	Counterfactual – matched comparison groups of participants and non-participants
eg Increased skills and changed behavior	Consistency of evidence and ruling out alternatives – process tracing and key informant attribution

(iv) Check that the design and process answers the action components of KEQs

Answers to action questions are often made in the form of recommendations. These don't necessarily flow straight from the findings. They often need an additional step of identifying possible actions and selecting the most appropriate, given the particular values and the availability of resources.

As manager you should check there is an explicit process for developing and reviewing recommendations, with appropriate levels of input from key stakeholders.

Consider the implications of the resources available and specific constraints

Identify the resources that can be used for the evaluation, and any particular constraints for them.

The following potential resources could be used for the evaluation:

- Funding to engage external individuals or organizations to design and/or conduct the evaluation or review the design and the final report
- Staff time to either conduct the evaluation or to manage an external contractor
- Time and goodwill of other stakeholders who will be involved in the evaluation – such as partner organizations, community members.
- Existing data

Identify any particular constraints for the evaluation such as:

- Short time before findings are needed to inform decisions
- Poor reputation of evaluation due to previous experiences
- Difficulties in engaging particular groups or in working collaboratively
- Missing baseline data
- Difficulties in observing or getting data about implementation or results – for example, when it is being implemented in remote locations, or in fragile, conflict-affected areas.
- Disagreement about what success looks like – for example:
    - Disagreement about the overall goals – for example, is an early childhood program primarily about improving workforce participation of parents or about early learning of children?
    - Disagreement about the criteria that should be used – for example, is good research technically very accurate or produced in time to inform an important decision? Is the goal to improve the average health and wellbeing in a community or to ensure everyone is above the minimum requirement?
    - Diagreement about the standards that should be used - for example, is a 10% increase in published research a good result?

Do an estimate of the costs to collect and analyse the data, as well as the project management and reporting time needed. If available resources are not adequate for the design, adjust the design and/or resources.

When reducing costs it is essential to consider the implications and how to manage these risks. Some possible options for reducing costs are shown below, along with some possible implications and ideas for managing the risks.

- Reduce the number of Key Evaluation Questions
    - Possible implications: Evaluation might no longer meet the needs of the primary intended users
    - How to manage these risks: Carefully prioritise the KEQs. Review whether the evaluation is still worth doing
- Reduce sample sizes
    - Possible implications: Reduced accuracy of estimates
    - How to manage these risks: Check these will still be sufficiently credible and useful through data rehearsal using interval estimates
- Make more use of existing data
    - Possible implications: Might mean that insufficiently accurate or relevant data are used. The cost savings might be minimal if they are not readily accessible.
    - How to manage these risks: This is only appropriate when the relevance, quality and accessibility of the existing data is adequate – need to check this is the case before committing to use them
- Embed data collection in program implementation
    - Possible implications: Might lead to a reduction in data quality
    - How to manage these risks: Ensure staff are trained and motivated to collect data properly and have sufficient time and equipment to do so
- Use fewer waves of data collection, including possibly retrospectively created baselines
    - Possible implications: Will increase the risk of inaccurate data
    - How to manage these risks: Check that retrospective baselines will be sufficiently accurate and that less frequent information on progress will be sufficient to inform decisions

Evaluation work plan

Oversee or engage in the development of a work plan that sets out the specific activities to implement the evaluation.

This should include identification of the reporting requirements and a dissemination plan.

The work plan can refer to existing documentation, so there is no need to start from scratch or duplicate effort. For large, multi-stakeholder evaluations, it may be helpful to spell out the roles and responsibilities of different actors and make the work plan a more formal document.

What should the evaluation work plan contain?

The evaluation work plan is a written document that specifies the evaluation design and details the practices and procedures to use in conducting the evaluation including:

- The evaluation framework –the project / program objectives and theory of change or logic model, the evaluation questions, the evaluation context (such as aspects of the organization, staff, participants etc. which may affect the evaluation) and time frame;
- The evaluation design, methods and processes –types of information needed, data collection and analysis methods, compliance with ethical standards (including informed consent, confidentiality, etc.), quality assurance processes (such as training of data collectors, data cleaning etc.), compliance with reporting requirements, dissemination of evaluation findings and supporting use (in case of external evaluators, the latter tasks may be / may not be included in their responsibilities);
- Who will do which tasks and when in the evaluation implementation;
- Contingencies –how challenges in the evaluation implementation will be identified and addressed.

Estimates of time needed for pre-evaluation as well as evaluation activities should be realistic, but need to incorporate some flexibility (or at least some openness to re-negotiation) to be able to deal with unanticipated challenges in the implementation of the evaluation.

How should the evaluation work plan be developed?

If the evaluation is conducted by internal evaluators, it is their responsibility to develop the plan –with or without the help of an outside consultant– and in consultation with the program manager and/or other program staff.

If the evaluation is conducted by external evaluators, the development of the evaluation plan is usually the first deliverable in their contract.

The evaluation plan should be developed well in advance of the start of the evaluation (at least 2-3 months) to allow for: review by relevant stakeholders; making necessary changes; obtaining ethical approval, where required (such as through an Institutional Review Board, IRB); and, pilot testing data collection instruments, as needed.

Review of the evaluation plan may include determining whether it:

- is consistent with the available evaluation resources and agreed evaluation objectives;
- focuses on the most important types of information to know (i.e., need to know rather than nice to know);
- does not place undue burden on project / program staff or participants;
- is ethical and culturally appropriate.

Reviewers may include: project / program staff, internal or external evaluation experts, advisory board members, the organization's administrators, project / program participants and relevant community members.

What does evaluation logistics planning entail?

If logistics are poorly thought out or under-funded, even the best evaluation team will not be successful. Hence, an evaluation logistics plan should be detailed and well-coordinated to support the different evaluation activities.

Critical logistical planning may include: negotiation of site visit dates, objectives and on the ground requirements (such as staff availability, access to documents/data etc.); necessary notification of officials or community leaders to ensure access and cooperation/collaboration; requirements for transportation, lodging, food, office space and other facilities; requirements for translators or other specific services (such as security).

The commissioning organization may take on the logistics for the evaluation fully or partially, regardless of whether internal or external evaluators are used. In either case, a designated evaluation logistician should be identified to ensure everything runs smoothly and to problem-solve where needed.

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation work plan
- Evaluation logistics plan

IDRC-specific information

Here are some examples of work plans for IDRC-commissioned evaluations.

- Inception Report for the phase II evaluation of the Think Tank Initiative
- Inception Report for the evaluation of the Canadian International Food Security Research Fund
- Work plan and evaluation matrix for the ACACIA program external review, 2010

Implementation

Oversee or engage in specific evaluation activities to collect and analyse data as set out in the evaluation work plan (including revising it as needed) and keeping the evaluation management group(s) informed and involved.

The evaluation plan should specify the sequence of evaluation activities and clearly spell out the roles and responsibilities of the evaluator or the different members of the evaluation team and all others involved in the evaluation process.

The evaluation manager will oversee the implementation of the evaluation plan and help problem-solve if any issues come up. She/he now plays a key role in communicating with the evaluation team, the evaluation

steering committee and technical advisory group (if relevant) but also with the key stakeholders, such as the project / program manager, the community or others.

One of the first tasks is to agree a communication schedule, covering how often and how (such as verbal, e-mail updates, routine written reports etc.) the evaluation team leader (and/or other team members) check in with the evaluation manager on progress with the assigned tasks. Similarly, she/he is also responsible for keeping stakeholders informed of implementation progress through regular communications with the existing governance mechanisms for the evaluation and collecting input / feedback such as on the evaluation design, the evaluation work plan etc. (see Step 1).

It is particularly important to agree on how unforeseen issues that affect the implementation of the evaluation and may jeopardize particular milestones and/or deliverables will be raised, escalated and resolved. Ideally, this has already been agreed early on in the evaluation process (see Step 1).

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Progress reports
- Documented implementation challenges and how these were resolved
- Revised evaluation work plan

Reporting

The evaluation reports should include relevant and comprehensive information structured in a manner that facilitates its use but also provide transparency in terms of the methods used and the evidence obtained to substantiate the conclusions and recommendations.

Evaluation, by definition, answers *evaluative* questions, that is, questions about 'quality' (how good something is) and 'value' (taking into account the specific situation such as the resources used to produce the results and the needs it was supposed to address). Evaluative reasoning is required to synthesize dimensions of quality and value to formulate defensible (i.e., well reasoned and well evidenced) answers to the evaluative questions.

The structure of an evaluation report can do a great deal to encourage the succinct reporting of direct answers to evaluative questions, backed up by enough detail about the evaluative reasoning and methodology to allow the reader to follow the logic and clearly see the evidence base.

The following recommendations will help to set clear expectations for evaluation reports that are strong on evaluative reasoning:

1. The executive summary must contain direct and explicitly evaluative answers to the key evaluation questions (KEQs) used to guide the whole evaluation.
2. Explicitly evaluative language must be used when presenting findings (rather than value-neutral language that merely describes findings). Examples should be provided.
3. Use of clear and simple data visualization to present easy-to-understand 'snapshots' of how the intervention has performed on the various dimensions of merit.
4. Structuring of the findings section using KEQs as subheadings (rather than types and sources of evidence, as is frequently done).

5. There must be clarity and transparency about the evaluative reasoning used, with the explanations clearly understandable to both non-evaluators and readers without deep content expertise in the subject matter. These explanations should be broad and brief in the main body of the report, with more detail available in annexes.
6. If evaluative rubrics are relatively small in size, these should be included in the main body of the report. If they are large, a brief summary of at least one or two should be included in the main body of the report, with all rubrics included in full in an annex.

A hallmark of great evaluative reasoning is how succinctly and clearly key points can be conveyed without glossing over important details.

[Source: Davidson J. Evaluative reasoning. UNICEF Methodological Brief 4 on Impact Evaluation. Florence: UNICEF]

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation report
- Products tailored to different audiences: Evaluation summary, Policy Brief, Newsletter, Conference presentation etc.

IDRC-specific information

The IDRC evaluation manager is responsible for:

- Identifying what report(s) will be needed and the agreed format. This should be done early in the evaluation process.
- Providing feedback on the draft evaluation reports to ensure it is in line with the IDRC Guideline. Only reports that have been approved on the basis of a quality assessment are accepted as final deliverable and released for use.

IDRC staff members, partners, interns, or consultants doing evaluation work for IDRC should use the guideline *Formatting Evaluation Reports at IDRC* *to structure the main evaluation report.*

IDRC Evaluation Report Template:

1. **Cover Page**

- Title
- Evaluator(s) name and organizational affiliation
- Date
- Name of the IDRC team, branch, unit, or person commissioning the evaluation
- IDRC Project or Research Support Project numbers of all the projects covered in the assessment (if applicable)

2. **Executive Summary**

A brief 1-2 page description of the main findings, methodological approach, and recommendations or conclusions of the evaluation.

3. **Body of the Evaluation Report**

- Background of the study:

This should detail the intended user(s) and use(s) of the evaluation process and/or product; what led to the evaluation (e.g. need, purpose, etc.); the specific evaluation questions or issues addressed; the values and principles guiding the evaluation process; and, any capacity building intentions.

- Description of the methodology employed:

This should include an analysis of the strengths and weaknesses of the research design, tools and methods used, the process followed, data sources, and people interviewed. It should describe how the project/program stakeholders and the intended user(s) of the evaluation participated in the process. It should also comment on the validity of the evidence and any ethical considerations.

- Evaluation Findings:

This section should be formulated according to the evaluation plan and the terms of reference (TOR) of the evaluation study.

4. **Annexes**

- List of Acronyms.
- List of people interviewed –with full coordinates if appropriate and not in breach of confidentiality.
- Bibliography of all documents reviewed.
- TOR for the evaluation and/or evaluator.
- Biography of the evaluator(s). This should include the name, sex, organizational affiliation, and contact information for the evaluator(s).

The IDRC guideline on data visualization (PDF, 774KB) provides useful tips for making data easier to understand and use.

The quality of the evaluation report is judged by IDRC's Evaluation Unit on four internationally recognized standards: utility, feasibility, accuracy, and propriety. A copy of "Quality Assessment of IDRC Evaluation Reports" should be given to the evaluator(s) to ensure they understand how the quality of the evaluation report will be assessed.

Resources

- Identify reporting requirements
  Identify the primary intended stakeholders and determine their reporting needs, including their decision-making timelines. Develop a communication plan.
- Develop reporting media
  Produce appropriate written, visual, and/or verbal products that communicate the findings.
- Visualise data
  Decide how to visualise the data to bring clarity during analysis and/or to communicate findings.

Distribution and learnings

Make evaluation reports available and engage with primary intended users to make the results accessible.

Archive the evaluation report and data (if appropriate). Use effective strategies to support use of the findings.

Products

The following items are potential outputs from this step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation report
- Products tailored to different audiences: Evaluation Summary, Policy Brief, newsletter, conference presentation, poster, video etc.
- Archive of evaluation reports and/or evaluation data
- Evaluation use documentation

Make evaluation reports available and engage with primary intended users to make the results accessible

Actual use of the evaluation findings often depends on how well the report meets the information needs of the primary intended users.

Hence, it is important to discuss the content, sharing, and use of the evaluation report and other dissemination products during the initial planning of the evaluation, even though it may be one of the last evaluation tasks.

Findings can also be communicated to those who do not have a direct 'stake' in the evaluation. For example, lessons learned from the evaluation can be helpful to other evaluators or program staff working in the same field; or, it may be worthwhile sharing some of the findings through articles or stories to attract wider attention to the organization's work, or to spread news about a particular situation.

Sharing evaluation findings through a report is only one way to distribute findings. Depending on the audience targeted and the available budget, other options may include:

- Presenting findings at staff fora and subject matter conferences
- Developing a short video version of the findings
- Sharing stories, pictures and drawings from the evaluation (depending on the methods used to collect the data)
- Creating posters or infographics for display
- Producing a series of short memos

Products

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation report
- Products tailored to different audiences: Evaluation Summary, Policy Brief, newsletter, conference presentation, poster, video etc.

Resources

- Develop reporting media
  Produce appropriate written, visual, and/or verbal products that communicate the findings.

Archive the evaluation report and data (if appropriate)

There are benefits of an archive of evaluation data and/or reports in the way they can be used.

Such as:

- to provide examples of well-conducted and well-reported on evaluations
- to identify common weaknesses in evaluations which may indicate the need for specific guidance or training
- to synthesize data across different evaluations (to answer questions such AS “Do these types of interventions work?” or “For whom, in what ways and under what circumstances do they work?”)
- to share data (if appropriate) with other stakeholders such as to contribute to a larger data set on a particular theme or area of shared interest for further analysis

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Archive of evaluation reports and/or evaluation data

IDRC-specific information

IDRC maintains, in its online open-access digital library, a repository of evaluations conducted throughout IDRC’s history. The repository currently holds close to 1000 evaluations.

Support the use of evaluation findings

Following up on the organisation’s response to evaluation findings is an essential part of supporting use.

This is often a management responsibility but it is recommended to build time into the evaluation plan and budget for the evaluator(s) to provide support beyond just delivering the evaluation report.

There are a range of options that can be used to provide support for the use of evaluation findings. For example:

- *Annual reviews*: reviewing major evaluation findings and conclusions based on evaluation studies completed during the preceding year (see, for example, How to conduct an annual review?)
- *Data use calendar*: guides the collection of data and reporting requirements, as well as ensuring that analysis and evaluation data is actively used (see, for example, a data use calendar that includes the

key recommendations and issues for a range of stakeholders, and provides clear next steps and time frames for follow up actions to take place).
- *Management response*: a written response to the recommendations made in the evaluation report. The response might agree, partially agree or disagree with a recommendation and should provide an explanation for any partial acceptance or rejection of a recommendation. For recommendations that have been accepted, or partially accepted, key follow-up actions should be identified, with a time frame specified and the responsible unit named. It is important to identify an individual to coordinate the management response and an agreed deadline by which comments must be provided. The management should respond timely, for example, within two months of receiving the final evaluation report (see, for example, guidance on preparing management responses)
- *Recommendations tracking*: keeping a transparent record of the responses to and action from recommendations (see, for example, a recommendation tracking matrix and how it can be used).

Product

The following item is a potential output from this sub-step. Where possible, it might be useful to research other deliverables that have also been shown to be effective.

- Evaluation use documentation

IDRC-specific information

IDRC programs can include a range or ways in which the use of evaluations is supported. Below are some examples for disseminating evaluation findings and engaging users and influencing change. Keep in mind supporting use is, typically, an additional step after comments have been invited on a draft evaluation report or other output. When people comment on a draft report, they are providing additional data, sharpening analysis, and, generally focusing on critiquing the consultants' report.

Examples of dissemination of evaluation findings:

- When programs inform grantees that a program evaluation is underway and solicit their participation in data collection, the program also lets the grantees know when and how they will be able to access the final evaluation report. See, for example, an email sent out alerting grantees to an external review report.

- Sharing the evaluation findings with users can be done in many different formats: as a summary document, full report, a website, a news piece on a program's website, and/or a blog.

Examples of engagement of users and influencing change:

- Users may write the recommendations: some evaluations limit the work of external consultants to collecting data and providing the evaluation findings and conclusions. The users of the evaluation themselves decide what recommendations they draw from the evaluation.

- Some programs have invited their evaluators to present the evaluation in a team meeting in which the team discusses the recommendations and decides priority actions based on the evaluation findings.

- Program and/or management responses: some evaluations have formally documented the reflections of the users of the evaluation. For example, the Pan Asia Networking team wrote a response to both the process and findings of the evaluation of the program's networking approach

and included it as an annex in the publicly-available report (see: https://idl-bnc-idrc.dspacedirect.org/handle/10625/39270, p.47 onwards). Management responses are filed among each set of documents from the 2015 external reviews (see, for example: https://idl-bnc-idrc.dspacedirect.org/handle/10625/54112).

- Evaluations are often used repeatedly over time. Some IDRC staff wanted to situate the new Climate Change Program as part of trajectory of programming on climate change adaptation that had started more than ten years earlier. So, they started their annual meeting by revisiting their evaluations and their own recollections about major achievements and areas for improvement from three previous climate change programs. They used these as inputs for the new strategy development.

Resources

- Support use
  Plan processes to support primary intended users to make decisions and take action on the basis of the findings.
- UNESCO guidelines for follow-up to evaluation findings

  This four-page paper provides an overview to the United Nations Educational, Scientific and Cultural Organization (UNESCO) procedures for evaluation follow up and a template for managers to detail their action plans in response to evaluation findings.