Rainbow Framework

There are many different methods and processes that can be used in monitoring and evaluation (M&E). The Rainbow Framework organises these methods and processes in terms of the tasks that are often undertaken in M&E. The range of tasks are organised into seven colour-coded clusters that aim to make it easy for you to find what you need: Manage, Define, Frame, Describe, Understand Causes, Synthesise, and Report & Support Use.

The Rainbow Framework can help you plan an M&E activity by prompting you to think about each of these tasks in turn, and select a combination of methods and processes that cover all tasks involved. You might also choose an approach, which is a pre-packaged combination of methods.

Explore the Rainbow Framework

The Rainbow Framework shows the different methods or processes that can be used for each task in an evaluation. Most of these tasks are needed in any M&E activity, but some (marked with an asterisk) might not be needed, depending on the purpose of the M&E activity or the specific type of evaluation you aim to conduct.

Manage an evaluation or evaluation system

Managing an evaluation involves agreeing on how decisions will be made for each cluster of the evaluation (from framing an evaluation to reporting and supporting use) and ensuring they are implemented well.

As you work through the process of planning and implementing the evaluation, you may need to revisit and revise the choices you have made.

Understand and engage stakeholders

Stakeholders are people with a stake in the evaluation, including primary intended users and others.

Understanding and taking into account the priorities and concerns of different stakeholders informs evaluation planning, communication strategies during and after the evaluation and supports the utilisation of evaluation findings.

The primary intended users – people who will be making decisions on the basis of the evaluation findings - are a key group of stakeholders. (Identifying primary intended users is its own important task).

Other stakeholders include people who will be affected by decisions made during or after the evaluation (program staff, program participants and beneficiaries) and secondary users of the evaluation findings. Evaluation findings are often of interest to policy makers and advocates for or against a particular course of action.

Different stakeholders can be engaged for different purposes and at different phases of evaluation planning and implementation. It may not be feasible or appropriate to engage all potential stakeholders.

Involving stakeholders during evaluation planning and implementation can add value by:

- providing perspectives on what will be considered a credible, high quality and useful evaluation
- contributing to the program logic and framing of key evaluation questions
- facilitating quality data collection
- helping to make sense of the data that has been collected
- increasing the utilization of the evaluation's findings by building knowledge about and support for the evaluation.

Engaging stakeholders is also important for managing risks especially when evaluating a contentious program or policy in which key stakeholders are known to have opposing views. It is important to understand different perspectives on what will be considered credible evidence of outcomes and impacts.

Methods

Understand stakeholders

- Community scoping

  Community profiles are good for developing a more in-depth understanding of a community of interest.

- Existing documents

  Reviewing documents produced as part of the implementation of the evaluand can provide useful background information and be beneficial in understanding the alignment between planned and actual implementation.

- Stakeholder mapping and analysis

  Stakeholders are individuals or organizations that will be affected in some significant way by the outcome of the evaluation process or that are affected by the performance of the intervention, or both.

Engage stakeholders

- Community fairs

  A community fair is an event organised within the local community with the aim of providing information about a project and raising awareness of relevant issues.

- Fishbowl technique

  The fish bowl activity is used to manage group discussion.

- Formal meeting processes

Studies have demonstrated that attendance at meetings and conferences, planning discussions within the project related to use of the program evaluation, and participation in data collection foster feelings of evaluation involvement among stakeholders (T

- Informal meeting processes

  Informal meetings can simply be a conversation between an evaluator and a key stakeholder that is not conducted in a formal way.

- Launch workshop

  A launch workshop is a meeting of key stakeholders to both assess and build readiness for evaluation.

Establish decision making processes

A variety of groups may be established within the governance structure in order to advise on the evaluation.

Evaluation decisions are often made by a steering committee, with representatives from different stakeholder groups. An expert or technical reference group or an advisor with specific expertise might provide targeted advice. A diverse range of stakeholders with different perspectives might also be consulted about the scope of the evaluation or on specific issues such as the accuracy of the program logic or the interpretation of findings.

Control may be centralized in a specific manager or committee or it may be shared by a working party involving representatives from many different stakeholders.

It is important to be clear about the roles and responsibilities of steering committees and other stakeholders. They might have the following roles:

- Advise – review material and make suggestions to others who make the decisions
- Recommend – review material and suggestions and make recommendations to others who make the decisions
- Decide – have final control over decisions in the evaluation

Methods

Types of 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.

- Citizen juries

Citizen juries are a method to engage citizens from the wider community in decision-making processes.

- 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).

Ways of exploring issues

- Formal meeting processes

Studies have demonstrated that attendance at meetings and conferences, planning discussions within the project related to use of the program evaluation, and participation in data collection foster feelings of evaluation involvement among stakeholders (T

- Informal meeting processes

Informal meetings can simply be a conversation between an evaluator and a key stakeholder that is not conducted in a formal way.

- Round robin

The "round robin" method is a technique for generating and developing ideas in a group brainstorming setting.

- Six thinking hats

The Six Thinking Hats method encourages participants to cycle through six different ways of thinking, using the metaphor of wearing different conceptual "hats".

Ways of making decisions

- 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.

- Hierarchical decision making

Hierarchical decision-making involves making decisions based on formal positions of authority.

- Majority decision making

Majority decision-making involves making decisions based on the support of the majority of the decision-makers.

Approaches

- Participatory evaluation

  Participatory evaluation is an approach that involves the stakeholders of a programme or policy in the evaluation process.

Decide who will conduct the evaluation

Evaluations can be conducted by a range of different actors including: external contractors; internal staff; those involved in delivering services; by peers; by the community; and by a combined group.

Therefore it is important to make decisions about who is best to conduct the evaluation.

Consider the relative importance of different types of expertise. Relevant expertise may include skills and knowledge in evaluation, in the specific domain (eg education) or program (e.g. delivering health services), or the local culture and context.

Consider the balance of distance and involvement that will be most suitable and that will support use of the evaluation findings. An external, unaligned evaluator may be viewed as more (or less) credible by different stakeholders. Involving staff and communities may be important for supporting cultural change, knowledge building and supporting the utilization of the evaluation findings.

Different management tasks arise depending on who is involved in which evaluative activities. For example, when using an external evaluator you will need to develop a process for selecting and managing them. If internal staff and/or intended beneficiaries are involved there may be a need to ensure processes are well documented and that relevant training in specific evaluation options is conducted to ensure that quality and ethical standards are maintained.

Decisions about who will conduct an evaluation, or components of an evaluation, will also be informed by timelines, resources, and the purpose of the evaluation.

Methods

- Community

  The community, particularly intended beneficiaries of an intervention, can undertake an evaluation or contribute to a combined team.

- Expert review

  Expert review involves an identified expert providing a review of draft documents at specified stages of a process and/or planned processes.

- External consultant

  An external consultant is someone external to the organization who is contracted to conduct the evaluation.

- Hybrid - internal and external staff

  A hybrid evaluation involves both internal and external staff working together.

- Internal staff

  Conducting an evaluation using staff from the implementing agency rather than hiring external consultants.

- Joint evaluation design

  A collaboration is involved in designing the evaluation, which might involve an implementing agency, an evaluation team and/or a community working together.

- Learning alliances

  Learning alliances involve a structured partnership between two or more organisations with the aim of working together to build and share knowledge around topics of mutual interest.

- Peer review

  Conducting an evaluation using individuals/organizations who are working on similar projects.

Approaches

- Horizontal evaluation

  Horizontal evaluation is an approach that combines self-assessment by local participants and external review by peers.

- Positive deviance

  Positive deviance (PD), a behavioural and social change approach, involves learning from those who find unique and successful solutions to problems despite facing the same challenges, constraints and resource deprivation as others.

- Participatory evaluation

  Participatory evaluation is an approach that involves the stakeholders of a programme or policy in the evaluation process.

Resources

Guides

- NSW Government evaluation toolkit

  This web-based toolkit has been developed to help program managers in New South Wales (Australia) government agencies manage evaluations (including those undertaken by internal or external evaluators, or by a combination of both).

- Key considerations for managing evaluations

  This guide from Pact South Africa is aimed at providing an overview of the key considerations that need to be assessed before and during the evaluation process.

- The program manager's guide to evaluation

  This comprehensive guide from the US Administration for Children and Families provides a step-by-step outline of the evaluation process from purpose to reporting.

Blog post

- Is independence always a good thing?

  This blog post from Howard White (May 1, 2014) argues that the benefits of an independent evaluation team can sometimes be overstated. He presents three arguments to support this contention: Institutional independence does not necessarily safeguard against biases toward positive evaluation; independence comes at a cost; and what agency evaluation departments do is only a small part of the evaluation story.

Determine what constitutes high quality evaluation

For any evaluation, there needs to be clarity about what will be considered a quality and ethical evaluation.

Different criteria can be used to determine what constitutes a good quality evaluation, including ethical practice. The options listed below are different criteria that can be used to define what constitutes high-quality evaluation. They are sometimes labelled as evaluation standards or norms.

These can be operationalised through processes and tools. You can read about various ways of doing this on the page Review Evaluation Quality.

Methods

Criteria relating to products

- Accessibility

  Accessibility of evaluation products includes consideration of the format and access options for reports, including plain language, inclusive print design, material in multiple languages, and material in alternative formats (such as online, audio, or braille).

- Accuracy

  Accuracy refers to the correctness of the evidence and conclusions in an evaluation. It may have an implication of precision.

- Credibility

Credibility refers to the trustworthiness of the evaluation findings, achieved through high-quality evaluation processes, especially rigour, integrity, competence, inclusion of diverse perspectives, and stakeholder engagement.

- Transferability

Transferability involves presenting findings in a way that they can be applied in other contexts or settings, considering the local culture and context to enhance the utility and reach of evaluation insights.

Criteria relating to processes

- Bias reduction

Bias reduction involves identifying possible sources of bias and taking steps to reduce it. This is one way of improving the validity of an evaluation.

Types of bias include,

- Common good and equity

Consideration of common good and equity involves an evaluation going beyond using only the values of evaluation stakeholders to develop an evaluative framework to also consider common good and equity more broadly.

- Competence

Competence refers to ensuring that the evaluation team has or can draw on the skills, knowledge and experience needed to undertake the evaluation.

Competence is one of the guiding principles of the American Evaluation Association:

- Cultural competency

Cultural competency involves ensuring that evaluators have the skills, knowledge, and experience necessary to work respectfully and safely in cultural contexts different from their own.

- Ethical practice

Ethical practice in evaluation can be understood in terms of designing and conducting an evaluation to minimise any potential for harm and to maximise the value of the evaluation.

- Evaluation accountability

Evaluation accountability relates to processes in place to ensure the evaluation is carried out transparently and to a high-quality standard.

- Feasibility

Feasibility refers to ensuring that an evaluation can be realistically and effectively implemented, considering factors such as practicality, resource use, and responsiveness to the programme's context, including factors such as culture and politics.

- Human rights and gender equality

  Human rights and gender equality refer to the extent to which an evaluation adequately addresses human rights and gender in its design, conduct, and reporting.

- Impartiality

  Impartiality in evaluation refers to conducting an evaluation without bias or favouritism, treating all aspects and stakeholders fairly.

  Key aspects of impartiality in evaluation can include:

- Inclusion of diverse perspectives

  Inclusion of diverse perspectives requires attention to ensure that marginalised people and communities are adequately engaged in the evaluation.

- Independence

  Independence can include organisational independence, where an evaluator or evaluation team can independently set a work plan and finalise reports without undue interference, and behavioural independence, where evaluators can conduct and report evaluati

- Integrity

  Integrity refers to ensuring honesty, transparency, and adherence to ethical behaviour by all those involved in the evaluation process.

  Integrity is one of the guiding principles of the American Evaluation Association:

- Professionalism

  Professionalism within evaluation is largely understood in terms of high levels of competence and ethical practice.

- Propriety

  Propriety refers to ensuring that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in it and those affected by its results.

- Respect for people

  Respect for people during an evaluation requires those engaged in an evaluation to respect the security, dignity, and self-worth of respondents, program participants, clients, and other evaluation stakeholders.

- Rigour

  Rigour involves using systematic, transparent processes to produce valid findings and conclusions. There are significant differences in what this is understood to mean in evaluation.

- Strengthening national evaluation capacities

  Strengthening national evaluation capacities refers to the ways in which an evaluation can have broader value beyond a single evaluation report by increasing national capacities.

- Systematic inquiry

  Systematic inquiry involves thorough, methodical, contextually relevant and empirical inquiry into evaluation questions.

  Systematic inquiry is one of the guiding principles of the American Evaluation Association:

  Systematic Inquiry

- Transparency

  Transparency refers to the evaluation processes and conclusions being able to be scrutinised.

- Utility

  Utility standards are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.

- Validity

  Validity refers to the extent to which evaluation findings are correct.

  In the Evaluation Standards produced by the Joint Committee on Educational Evaluation, standards relating to this were labelled as Accuracy Standards:

Resource

- What counts as good evidence?

  This paper, written by Sandra Nutley, Alison Powell and Huw Davies for the Alliance for Useful Evidence, discusses the risks of using a hierarchy of evidence and suggests an alternative in which more complex matrix approaches for identifying evidence qu

Determine and secure resources

The purpose and scope of the evaluation needs to be considered when determining the budget.

The amount of resources available may influence the level of an evaluation's rigor or the certainty of its findings. The importance of the program, existing knowledge about the program from previous evaluations and the decisions to which the evaluation will contribute are important factors to consider.

A program that has been thoroughly tested in a context similar to the current implementation setting may require fewer resources to satisfy information needs. A higher proportion of funds may be warranted for:

- Evaluations that will contribute to important decisions, such as whether to roll out a program on a large scale
- Evaluations that require highly defensible findings or will come under scientific scrutiny
- Programs that have not been evaluated before

Very often the available resources (time, money and expertise) will restrict the scope of the evaluation (the number of questions, size of the sample, data collection and analysis options) or influence the choice of evaluation designs. Some organizations have a policy of setting aside a certain percentage of the total program budget for evaluation. Organizations often use a "rule of thumb" to specify considerations in making a budget estimate. Common budget estimates range between 5 – 20% of program costs.

When commissioning an evaluation it is wise to start the budgeting process by consulting with the budget, procurement and/or human resource offices within the organization in order to verify and understand budget process, rules, and stipulations. Engage project staff, stakeholders, and M&E staff or professionals to ensure that the budget is comprehensive and accurate.

The process of developing an evaluation budget may be an excellent opportunity to encourage stakeholders to agree on the value of the evaluation and the amount and type of resources necessary to support it. Sometimes after intended users are engaged and the evaluation purpose and questions decided there is scope to add additional resources in order to undertake the type of evaluation that is required.

Budgets are just as critical for planning an internal evaluation as an external one. Although an internal evaluation draws primarily from resources within the organization, getting agreement on available resources will ensure the evaluation runs much more smoothly. For example, staff may be more flexible than consultants, but developing an accurate calculation of staff time costs early in the process helps to enlist their commitment.

Methods

Determine resources needed

- Evaluation budget matrix

  An evaluation budget matrix specifies various items that need to be costed as individual line items.

- Evaluation costing

  Evaluation expenses are highly situational and there are no magic formulas for calculating costs.

- Resources stocktake

  The resources available for evaluation include people's time and expertise, equipment and funding.

Secure resources needed

- Designated staff time for evaluation

  This strategy for securing sufficient resources for conducting evaluation involves allocating a specified amount of staff time (hours or days per week) to work on evaluation.

- Grant funding for evaluation

  You may also consider approaching a foundation or other donor agency for the funds to undertake an evaluation.

- Work with universities to staff the evaluation

This strategy requires management leadership and uses the rule of thumb approach to estimate the percentage of project funds to spend on evaluation which could be done more accurately by developing an initial evaluation budget.

- Institutionalized budget allocation

  This strategy requires management leadership and uses the rule of thumb approach to estimate the percentage of project funds to spend on evaluation.

  This could be done more accurately by developing an initial evaluation budget.

- Leveraged partnerships to carry out the evaluation

  As many projects are undertaken by a consortium of organisations working together, sometimes it is worthwhile to consider approaching your implementing partners to pool resources and carry out the evaluation jointly.

- Strategies to reduce costs

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

Document management processes and agreements

It is important to document decisions about the management of evaluative activities, including any processes for monitoring compliance with ethical and quality standards during the evaluation.

These documents will also ensure that different stakeholders, whether funders, partner organisations, communities or expert advisors are clear about what is being done, how and when, and their responsibilities and accountabilities for the evaluation.

Different organisations have different forms of documents and different labels for the document that describes what is to be done - the purpose, Key Evaluation Questions and timeline.

Sometimes this document is referred to as Terms of Reference (ToR), Scope of Work (SOW), Statement of Work (SOW), Request for Proposal (RFP), Request for Quotation (RFQ), Invitation To Tender (ITT) or the evaluation brief.

This document can be used for any type of evaluation (internal, external, self-evaluation) but they are particularly useful as part of the process of engaging an external evaluator.

Other types of documents might be developed to formalise the relationships between different organisations working together on the evaluation. These could include a Memorandum of Understanding or a Contractual Agreement.

Methods

Document what is needed in an evaluation

- Expression of interest

    An expression of interest (EoI) is a way for an organisation to publish its intention to appoint an evaluation team to conduct an evaluation of a specific project or program.

- Request for proposal (RFP)

    A Request for Proposal (RFP) is a formal request for evaluators to prepare a response to a planned evaluation and are generally used to select the final evaluator for the evaluation.

- Scope of work

    A Scope of Work (SOW) is a plan for conducting an evaluation which outlines the work that is to be performed by the evaluation team.

- Terms of reference

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

Document how different organisations will work together

- Contractual agreement

    A formal contract is needed to engage an external evaluator and a written agreement covering similar issues can also be used to document agreements about an internal evaluator.

- Memorandum of understanding

    A Memorandum of Understanding (MOU) outlines an agreement between two collaborating bodies in order to identify the working relationships and guidelines that exist between them.

Develop planning documents for the evaluation or M&E system

An evaluation plan (for a particular evaluation) usually specifies: what will be evaluated; the purpose and criteria for the evaluation; the key evaluation questions; and how data will be collected, analyzed, synthesized and reported. It may include a program theory/logic model.

An evaluation framework (sometimes called a Monitoring and Evaluation framework, or more recently a Monitoring, Evaluation and Learning framework) provides an overall framework for evaluations across different programs or different evaluations of a single program (e.g. process evaluation; impact evaluation). This can include guidance on data sources and management processes, as well as an overall program theory/logic model.

However sometimes the term 'evaluation framework' is used to refer to a plan for a single evaluation or to an organisational policy.

Methods

- Aide memoire

  An aide-memoire generally refers to a document that is produced to summarise key findings and important recommendations of an evaluation.

- Evaluation framework

  An evaluation framework (sometimes called a Monitoring and Evaluation framework, or more recently a Monitoring, Evaluation and Learning framework) provides an overall framework for evaluations across different programs or different evaluations of a sing

- Evaluation plan

  An evaluation plan sets out the proposed details of an evaluation - what will be evaluated, how and when.

- Evaluation work plan

  An evaluation work plan involves the development of clear timeframes, deliverables and milestones.

- Inception report

  An inception report of an evaluation is prepared by an evaluator after an initial review of relevant documentation.

Review evaluation quality

Evaluating the quality of an evaluation can be done before it begins (reviewing the plan) or during or after the evaluation (reviewing the evaluation products or processes). This is sometimes called a quality review or meta-evaluation.

Some organisations require formal review of evaluations at specific stages. This is often focused on the evaluation design or plan, the inception report (which might include revising the evaluation design), and the evaluation report or reports. Knowing that specific outputs, such as an evaluation plan, will be subject to external scrutiny can also improve its quality.

Reviewing the evaluation plan and inception report can potentially improve the quality of the evaluation, as it is still possible to revise the design and implementation plans.

Reviewing the evaluation report can lead to improvements in how messages are communicated but there is often limited ability to address any deficiencies in the evaluation. It can however ensure that the key messages from the evaluation are clear and consistent with the findings. A formal review of an evaluation report can be particularly important where its findings are likely to be contentious.

Reviewing the evaluation will also help to identify how key messages may be interpreted, if there are any concerns about the methodology that need to be discussed, and possible ways that the findings will be

used. Being mindful of how the evaluation findings could be received helps in presenting the findings in a way that is likely to support use.

Involving the primary intended users and other key stakeholders in a review of the evaluation also supports the use of the evaluation findings by building the 'personal factor' – the involvement of people who care about the evaluation and how the findings will be used.

The options listed below are different processes and tools for evaluating evaluations. The criteria for evaluating evaluations are shown on the page Determine what constitutes high-quality evaluation.

Methods

- Ethical guidelines

  Ethical guidelines are designed to guide ethical behaviour and decision-making throughout evaluation practice.

- Evaluation standards

  Evaluation standards identify how the quality of an evaluation will be judged. They can be used when planning an evaluation as well as for meta-evaluation (evaluating the evaluation).

- Expert review for meta-evaluation

  An expert review involves experts reviewing the evaluation, drawing in part on their expertise and experience of the particular type of program or project.

- Group critical reflection

  This method involves facilitating group stakeholder feedback sessions on evaluation findings.

- Individual critical reflection

  This method involves facilitating independent feedback from particular individual stakeholders.

- Institutional review board

  Institutional Review Boards (IRBs) are committees that are set up by organizations to review the technical and ethical dimensions of a research or evaluation project.

- Peer review for meta-evaluation

  Reviewing the evaluation by using peers from within or outside of the organisation.

- Validation workshop

  A validation workshop is a meeting that brings together evaluators and key stakeholders to review an evaluation's findings.

Strengthen evaluation capacity

An important aspect of monitoring and evaluation (M&E) 'systems' is strengthening the M&E capacity of individuals, organisations, communities and networks.

While there are other terms used for this, we suggest using the term 'evaluation capacity strengthening' to emphasise the value of recognising, reinforcing and building on existing capacity.

Understanding capacity

M&E capacity is not just about developing competencies for doing monitoring and evaluation. It also includes competencies in effectively designing, managing, implementing and using monitoring and evaluation. It includes strengthening a culture of valuing evidence, valuing questioning, and valuing evaluative thinking. This can include the capacity of evaluators, as well as the capacity of evaluation and programme managers, internal staff, and community members.

When we think about evaluation capacity, it's more than an individual or organisation's ability to undertake technical tasks; it also includes a range of areas such as interpersonal communication and group facilitation, as well as the ability to frame evaluations, make sense of them, support their appropriate use.

Kinds of capacity

When we talk about strengthening evaluation capacity, we refer to building three types of capital:

- **Human capital** — knowledge and skills and the ability to apply them in contextually appropriate ways

- **Social capital** — supportive networks of trust and reciprocity to support work

- **Organisational capital** — including infrastructure and organisational culture

Change theories for capacity strengthening

It can also be useful to consider three broad change theories (drawing on Mitchie et al. 2011 meta-theory of behaviour change):

- increasing motivation
- increasing capacity
- increasing opportunity –including an enabling environment for M&E

Evaluation capacity strengthening is not just about training

One-off training is a common approach to evaluation capacity strengthening but it may not be the most appropriate way to address a capacity strengthening need.

Individuals, groups and organisations should think about different types of capacity strengthening activities and support and consider how these can be integrated to best address their specific needs.

We invite you to explore the full range of methods and processes available to you. Let us know if you have any further suggestions.

Methods

Increasing skills and knowledge

A range of methods related to various strategies to increase skills and knowledge - among evaluators, others doing evaluation, and people who oversee monitoring and evaluation systems (for example, program managers).

Competency assessment

- Self-assessment

  Self-assessment is an individual reflection on one's skills, knowledge and attitudes related to evaluation competencies.

- Peer-assessment

  Peer assessment can provide additional benefits beyond self-assessment – in particular, the opportunity for peer learning through the review process.

Knowledge, skills, attitudes (KSA) development and ongoing development

- Coaching

  Coaching can involve supporting an individual during training or development in order for them to reach a specific personal or professional goal, or providing expert and practical help to improve and apply specific skills and knowledge.

- Dialogues

  Dialogues refer to a range of learning conversations that go beyond knowledge transfer to include knowledge articulation and translation.

- Expert advice

  Expert advice provides advice in response to specific queries.

  It might include a process to clarify and reframe the question that is being asked.

- Fellowship

  A fellowship is an extended position that provides paid employment and support for people who have completed formal coursework in evaluation.

- Internship

  An internship is a paid or unpaid entry-level position that provides work experience and some professional development.

- Mentoring

Mentoring is a process where people are able to share their professional and personal experiences in order to support their development and growth in all spheres of life.

- Learning circle

A Learning Circle allows a group of individuals to meet and explore an issue and learn from each other in the process.

- Peer learning

Peer learning refers to a practitioner-to-practitioner approach in which the transfer of tacit knowledge is particularly important (Andrews and Manning 2016).

- Reflective practice

Reflective practice involves an individual reflecting on their work allowing them to learn from their own experiences and insights and engage in a practice of continual learning.

- Self-paced learning

Viewing learning materials, such as previously recorded webinars, at your own pace.

- Supervised practice in teams

Supervision of practice is an approach often used in social work where it is expected that all practitioners will engage in regular discussions of and reflections on their practice; it is not an approach only intended to support novices.

- Professional development courses

Professional development courses can be a useful way to develop people's knowledge and skills in conducting and/or managing an evaluation.

Building and sharing knowledge

- Community of practice

A community of practice allows a group of people with a common interest or concern to share and learn through a series of interactions, thus reflecting the social nature of human learning.

- Conferences

Attendance at professional conferences to understand how other evaluators frame and discuss their findings is a key component of building evaluation capacity.

- Evaluation library

In many organisations, a print or digital collection of books, manuals and other documents has been gathered to form an evaluation library that can be jointly accessed.

Decisions to be taken include:

- Evaluation journals

Evaluation journals play an important role in documenting, developing, and sharing theory and practice. They are an important component in strengthening evaluation capacity.

- Learning partnerships

  Learning partnerships involve structured processes over several years to support learning between a defined number of organisations working on similar programs, usually facilitated by a third party organisation.

- R&D projects

  Evaluation associations can leverage their membership to engage in knowledge construction through research and development.

Other strategies

Reference points for professional practice

These reference points can be used to guide activities aimed at increasing capacity – for example, when developing a training course or a peer learning program – or activities aimed at increasing motivation – for example, supporting a shared professional identity to motivate individuals.

- Ethical guidelines

  Ethical guidelines are designed to guide ethical behaviour and decision-making throughout evaluation practice.

- Competency frameworks

  Competencies are the skills, knowledge, attributes and behaviours needed to fulfil particular roles.

- Distinct occupational category

  A distinct occupational category or role title recognised at a national or organisational level.

- Expectation of ongoing competency development

  An expectation that members of an association or organisation will engage in ongoing competency development.

- Organisational M&E policy

  Organisational monitoring and evaluation policies are the set of rules or principles that an organisation uses to guide its decisions and actions with respect to monitoring and evaluation across programs and departments.

- Evaluation standards

  Evaluation standards identify how the quality of an evaluation will be judged. They can be used when planning an evaluation as well as for meta-evaluation (evaluating the evaluation).

Engagement with professional associations

Professional associations play an active role in supporting capacity development – for example, by offering workshops and encouraging the development of supportive professional relationships. They can also contribute to motivation by providing inspirational exemplars of practice and practitioners.

- Evaluation societies and associations

  Evaluation societies and associations play a significant role in strengthening national M&E systems.

- Other professional associations

  Associations from different but related sectors and fields can be good places to find useful events and training, network connections, and ideas.

Public recognition of good practice

- Awards

  An award is a formal recognition by peers of outstanding individuals or practice.

  Some awards are made for cumulative good practice, and others are for exemplars of good practice, such as awards for the best evaluation.

- Fellows

  Fellow is a category of membership of an association or society, often awarded to an individual based on their contributions to evaluation.

Increasing opportunity for professional practice

A range of methods for building a better informed and motivated demand side of evaluation and a more conducive enabling environment. Some relate to educating the public and evaluation managers and users about evaluation and evaluators, and others relate to engaging in wider organisational and public processes with implications for evaluation practice.

Educating the public, evaluation managers and users

- Public information about evaluation

  An important part of evaluation capacity strengthening is providing a clear definition or explanation of evaluation in online and printed materials.

- Public information about professional practice

As part of its public advocacy role, a professional association can provide potential clients with information about engaging with evaluators effectively.

Strengthening the enabling environment for good evaluation practice

- Engagement in relevant organisational processes

  For evaluation to be truly useful it needs to be embedded in organisational processes.

  Particularly relevant issues include strategic changes to how government and non-government organisations plan, manage and implement.

- Engagement in relevant public processes

  For evaluation to be truly useful it needs to engage in public discussions about relevant issues.

Review of practice

Some methods which relate to the task 'Evaluate evaluation' can be used as part of evaluation capacity strengthening, as they can both improve a specific product and also develop internal skills and knowledge.

- Expert review

  Expert review involves an identified expert providing a review of draft documents at specified stages of a process and/or planned processes.

- Peer review

  Conducting an evaluation using individuals/organizations who are working on similar projects.

Develop a design for the evaluation

An evaluation design sets out how data will be collected and analysed in terms of the methods used and the research design.

Evaluation designs should suit the particular evaluation in terms of the nature of the evaluation, the nature of what is being evaluated and the availability of resources:

- **The nature of the evaluation:** In particular, answering the key evaluation questions that have been identified, with methods that will answer different types of questions – descriptive, causal, and evaluative.
- **The nature of what is being evaluated:** Especially in terms of complicated or complex aspects that need to be addressed.
- **The availability of resources:** Especially time, money and existing data.

Methods

There are different ways to go about the task of designing an evaluation.

When the evaluation design is developed

- Upfront evaluation design

  An upfront evaluation design is done before or near the beginning of the evaluation and then implemented as designed or as revised at the end of the inception period.

- Iterative evaluation design

  An iterative evaluation design involves setting out an initial overall evaluation design or process at the beginning of the evaluation.

Who develops the evaluation design

- Commissioner-led evaluation design

  The organisation commissioning an evaluation develops an evaluation design as part of setting out the terms of reference for the evaluation.

- Evaluator-led evaluation design

  An evaluation team develops an evaluation design in response to an evaluation brief which sets out the purposes of the evaluation.

- Community-led evaluation design

  A community develops an evaluation design, sometimes with facilitation or technical support.

- Joint evaluation design

  A collaboration is involved in designing the evaluation, which might involve an implementing agency, an evaluation team and/or a community working together.

Combining evaluation methods and approaches

- Single evaluation approach design

  The evaluation design is based on selecting a single existing evaluation model or approach and using it for an evaluation.

- Bricolage evaluation design

A bricolage evaluation design flexibly combines and adapts various data collection and analysis methods, approaches, and conceptual and value frameworks to suit the specific context of the evaluation.

Define what is to be evaluated

This cluster of evaluation tasks develops an initial description of the program and how it is understood to work.

This can be used to:

- engage stakeholders in the task "understand and engage stakeholders" from the 'Manage' cluster of tasks
- guide choices about what data to collect in the 'Describe' cluster of tasks
- inform testing of causal links when planning how to 'Understand Causes'

Develop initial description

It is helpful to develop an initial description of the project, program or policy as part of beginning an evaluation.

Checking this with different stakeholders can be a helpful way of beginning to identify where there are disagreements or gaps in what is known about it.

An overview of what's being evaluated can include information on:

- The rationale: the issue being addressed, what is being done, who is intended to benefit
- 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 can create opportunities or challenges
- Significant changes that have occurred over time - because of changes in contextual factors or lessons learnt

Methods

- Existing documents

  Reviewing documents produced as part of the implementation of the evaluand can provide useful background information and be beneficial in understanding the alignment between planned and actual implementation.

- Existing project description

  Existing project descriptions about what is being evaluated can sometimes be accessed and used by evaluators.

- Peak experience description

  This method provides a succinct and coherent description of a program, project or policy when it is operating at its best.

- Thumbnail description

  A 'thumbnail' is a brief description (short like a thumb nail).

Approaches

- Appreciative inquiry

  Appreciative Inquiry is an approach to organisational change which focuses on strengths rather than on weaknesses - quite different to many approaches to evaluation which focus on deficits and problems.

Develop theory of change / programme theory

A programme theory or theory of change (TOC) explains how an intervention (a project, a programme, a policy, a strategy) is understood to contribute to a chain of results that produce the intended or actual impacts.

It can include positive impacts (which are beneficial) and negative impacts (which are detrimental). It can also show the other factors which contribute to producing impacts, such as context and other projects and programmes.

Different types of diagrams can be used to represent a theory of change. These are often referred to as logic models, as they show the overall logic of how the intervention is understood to work.

Why is a theory of change used?

A theory of change can be used to provide a conceptual framework for monitoring, for evaluation or for an integrated monitoring and evaluation framework.

A theory of change can be a very useful way of bringing together existing evidence about a programme, and clarifying where there is agreement and disagreement about how the programme is understood to work, and where there are gaps in the evidence.

It can be used for a single evaluation, for planning cluster evaluations of different projects funded under a single program, or to bring together evidence from multiple evaluations and research.

When are theories of change developed?

A theory of change is often developed during the planning stage of a new intervention. It can also be developed during implementation and even after a programme has finished. ? When an evaluation is being planned, it is useful to review the programme theory and revise or elaborate it if necessary.

How are theories of change developed?

A theory of change can be developed by programme staff, by an external evaluator, by programme designers, or collaboratively with the community.

How are theories of change represented?

The diagrams used to represent a theory of change (usually referred to as logic models) can be drawn in different ways.

Sometimes they are shown as a series of boxes (inputs->processes->outputs->outcomes->impacts), sometimes they are shown in a table, sometimes they are shown as a series of results, with activities occurring alongside them rather than just at the start. These different types are shown as different methods on this page (below).

Methods

Processes for developing a theory of change

- Articulating mental models

  Articulating mental models involves talking individually or in groups with key informants (including program planners, service implementors and clients) about how they understand an intervention works.

- Backcasting

  Backcasting is a method that involves envisaging alternative futures.

- Existing documents

  Reviewing documents produced as part of the implementation of the evaluand can provide useful background information and be beneficial in understanding the alignment between planned and actual implementation.

- Five Whys

  The Five Whys is an easy question asking option that examines the cause-and-effect relationships that underly problems.

- Generic change theories

  Generic change theories can be applied across different sectors - for example, motivation, deterrence, capacity development.

  This page provides links to some resources that outline these change theories.

- Group model building

  Group model building involves building a logic model in a group, often using sticky notes.

- Previous research and evaluation

Using the findings from evaluation and research studies that were previously conducted on the same or closely related areas.

- SWOT analysis

The SWOT analysis is a strategic planning tool that encourages group or individual reflection on and assessment of the Strengths, Weaknesses, Opportunities and Threats of a particular strategy and how to best implement it.

Ways of representing a theory of change in a logic model:

- Tiny tool results chain

Tiny tool results chain maps both positive and negative possible impacts from an intervention.

- Logframe

Logframes are a systematic, visual approach to designing, executing and assessing projects which encourages users to consider the relationships between available resources, planned activities, and desired changes or results.

- Outcomes hierarchy

An outcomes hierarchy shows all the outcomes (from short-term to longer-term) required to bring about the ultimate goal of an intervention.

Unlike results chains, it does not show the activities linked to these outcomes.

- Realist matrix

A realist matrix focuses on the causal mechanisms at work in a programme or project. It specifies what exactly in the programme creates the outcomes, and under what conditions.

- Results chain

"Results chain or pipeline logic models represent a program theory as a linear process with inputs and activities at the front and long-term outcomes at the end.

- Triple column

A triple column/row theory of change diagram shows the causal pathway in terms of intermediate outcomes, activities that directly produce these, and the influence of other factors and programs.

Approaches

A number of approaches include recommendations about how to develop a logic model as part of undertaking an evaluation:

- Collaborative outcomes reporting

Collaborative outcomes reporting (COR) is a participatory approach to impact evaluation based around a performance story that presents evidence of how a program has contributed to outcomes and impacts, that is then reviewed by both technical experts and

- Outcome Mapping

Outcome Mapping is an approach that helps unpack an initiative's theory of change and provides a framework to collect data on the immediate, basic changes that lead to longer, more transformative change.

- Realist evaluation

Realist evaluation aims to identify the underlying generative causal mechanisms that explain how outcomes were caused and how context influences these.

Resources

- Learning for sustainability: Theory of change

Annotated list of resources about developing and using a theory of change.

- Purposeful program theory: Effective use of theories of change and logic models

This book, by Sue Funnell and Patricia Rogers, discusses ways of developing, representing and using programme theory and theories of change in different ways to suit the particular situation.

- Theory of change

This guide, written by Patricia Rogers for UNICEF, looks at the use of theory of change in an impact evaluation.

- Theory of change software

Describes different options for using software to help create a logic model.

- Theory of Change: Good practice
  PDF
  120.32 KB

This paper sets out some suggestions about what might be considered good practice, adequate practice and inadequate practice in developing, representing and using a theory of change.

Manager's guide to evaluation

- Describe the theory of change

This section of the Manager's guide to evaluation explains how and why you might use a theory of change when commissioning and managing an evaluation.

Identify potential unintended results

Many evaluations and logic models only focus on intended outcomes and impacts - but positive or negative unintended results can be important too.

Use these methods before a program is implemented to identify possible unintended outcomes and impacts, especially negative impacts (that make things worse not better) that should also be investigated and tracked.

Make sure your data collection remains open to unintended results that you have not anticipated by including some open-ended questions in interviews and questionnaires, and by encouraging reporting of unexpected results.

Once you have identified possible unintended consequences use options from the 'DESCRIBE' component to gather information about them if and when they occur. Make sure your data collection remains open to the unintended and unanticipated by including some open-ended questions in interviews and questionnaires, and by encouraging reporting of unexpected results.

Methods

- Journals and logs

  Journals and logs are forms of record-keeping tools that can be used to capture information about activities, results, conditions, or personal perspectives on how change occurred over a period of time.

- Key informant interviews

  Key informant interviews involve interviewing people who have particularly informed perspectives on an aspect of the program being evaluated.

- Negative programme theory

  Most programme theories, logic models and theories of change show how an intervention is expected to contribute to positive impacts; Negative programme theory, a technique developed by Carol Weiss, shows how it might produce negative impacts.

- Risk assessment

  Conducting a risk assessment involves identifying potential negative impacts, their likelihood of occurring and how they might be avoided.

- Six thinking hats

  The Six Thinking Hats method encourages participants to cycle through six different ways of thinking, using the metaphor of wearing different conceptual "hats".

- Unusual events reporting

The reporting of unusual events or incidents is important both for the sake of transparency and to improve policies and procedures.

Understand the situation

A situation analysis examines the current situation and the factors contributing to it. This might include identification and analysis of needs, resources, strengths, weaknesses, opportunities, threats, and/or power analysis.

A situation analysis can be done as part of planning an intervention before implementation, or during implementation to inform ongoing planning, or as part of planning or conducting an evaluation.

Methods

Methods for undertaking a situation analysis include:

- Asset mapping

  Asset mapping is a process of identifying existing assets within a community, organisation or network. It complements the "deficit focus" of needs analysis.

- Demographic mapping

  Demographic mapping is a way of using GIS (global information system) mapping technology to show data on population characteristics by region or geographic area.

- GIS mapping

  GIS mapping will typically display one data variable or indicator, often using colour coding to indicate the density, frequency, or percentage in a given region, allowing quick comparison between regions.

- Interactive mapping

  Interactive mapping involves using maps that allow zooming in and out, panning around, identifying specific features, querying underlying data such as by topic or a specific indicator (e.g., socioeconomic status), generating reports and other means of u

- Needs analysis

  A needs analysis identifies the current needs of an individual, organisation, or community.

  Four different types of need were identified by a classic paper by Bradshaw in 1972:

- Power analysis

  A power analysis identifies the main types of power in a system of interest.

- Social mapping

Social mapping, or 'wellbeing ranking', is used to identify households using pre-determined indicators based on socio-economic factors.

- Stakeholder mapping and analysis

  Stakeholders are individuals or organizations that will be affected in some significant way by the outcome of the evaluation process or that are affected by the performance of the intervention, or both.

- SWOT analysis

  The SWOT analysis is a strategic planning tool that encourages group or individual reflection on and assessment of the Strengths, Weaknesses, Opportunities and Threats of a particular strategy and how to best implement it.

Frame the boundaries for an evaluation

Framing an evaluation involves being clear about the boundaries of the evaluation.

Why is the evaluation being done? What are the broad evaluation questions it is trying to answer? What are the values that will be used to make judgments about whether it is good or bad, better or worse than alternatives, or getting better or worse?

Identify primary intended users

It is important to identify the people who are intended to actually use the evaluation, and to engage them in the evaluation in some way if possible.

This increases the likelihood that the evaluation will be done in ways that will be appropriate and that will actually be used.

Your primary intended users are not all those who have a stake in the evaluation, nor are they a 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 (for example, change policies and procedures, improve management strategies). Who they are will depend on your evaluation.

Research into how evaluation findings are used shows the importance of the 'personal factor'. The personal factor, a specific person or group of people who care about the evaluation findings, is the single most important predictor of evaluation finding use:

'The personal factor is the presence of an identifiable individual or group of people who personally care about the evaluation and the findings it generates. Where such a person or group was present, evaluations were used; where the personal factor was absent, there was a correspondingly marked absence of evaluation impact.'

The tasks of identifying primary intended users and deciding the purposes of an evaluation are interconnected. You might begin by identifying the intended users, who will then decide the purpose of the evaluation. Or the purpose of an evaluation may have already been prescribed,which helps you to

identify intended the users.

Resources

- Identifying the intended user(s) and use(s) of an evaluation

  This guideline from the International Development Research Centre (IDRC) 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

- Utilization-Focused Evaluation: 4th edition

  Useful for practitioners and students alike this book is both theoretical and practical. Features include follow-up exercises at the end of each chapter and a utilization-focused evaluation checklist.

- Utilisation-focused evaluation (U-FE) checklist

  Composed by Michael Quinn Patton in 2002 and updated in 2013, this is a comprehensive checklist for undertaking a utilisation-focused evaluation.

Decide purposes

It is important that key stakeholders agree on the main purpose or purposes of evaluation, and be aware of any possible conflicts between purposes.

The purposes of an evaluation will inform (and be informed by) the evaluation timelines, resources, stakeholders involved and choice of evaluation options for describing implementation, context and impact.

It is not enough to state that an evaluation will be used for accountability or for learning.

Evaluations for accountability need to be clear about who will be held accountable to whom for what and through what means. They need to be clear about whether accountability will be upwards (to funders and policymakers), downwards (to intended beneficiaries and communities) or horizontal (to colleagues and partners).

Evaluations for learning need to be clear about who will be learning about what and through what means. Will it be supporting ongoing learning for incremental improvements by service deliverers or learning about 'what works' or 'what works for whom in what circumstances' to inform future policy and investment?

It may be possible to address several purposes in a single evaluation design but often there needs to be a choice about where resources will be primarily focused.

Methods

Using findings

- Contribute to broader evidence base

  Inform future policy and practice by others outside the organisation.

- Inform decision making aimed at improvement (formative)

  Changing or confirming policies and practices.

- Inform decision making aimed at selection, continuation or termination (summative)

  Identifying best value for money.

- Lobby and advocate

  Justify expenditure and demonstrate achievements.

Using process

- Build trust and legitimacy across stakeholders

  Develop better understandings of each other and demonstrate that expectations are being met.

- Ensure accountability

  Holding someone to account to someone for something.

- Inclusion of diverse perspectives

  Inclusion of diverse perspectives requires attention to ensure that marginalised people and communities are adequately engaged in the evaluation.

Resources

- Exploding the myth of incompatibility between accountability and learning

  This chapter from Capacity Development in Practice examines the conflict in the field of Monitoring and Evaluation (M&E) between the need for 'accountability' and the desire to ensure 'learning'.

- Purposes of assessment - Keystone guide

  This webpage from Keystone Accountability outlines the six major reasons that social organizations monitor, assess and report their performance and results.

  The reasons identified include:

- Seeking surprise: Rethinking monitoring for collective learning in rural resource management

  This PhD thesis from Irene Guijt draws on her extensive knowledge and experience in the field of rural resource management in Brazil.

- Utilization-Focused Evaluation: 4th edition

Useful for practitioners and students alike this book is both theoretical and practical. Features include follow-up exercises at the end of each chapter and a utilization-focused evaluation checklist.

Specify the key evaluation questions

Key Evaluation Questions (KEQs) are the high-level questions that an evaluation is designed to answer - not specific questions that are asked in an interview or a questionnaire.

Having an agreed set of Key Evaluation Questions (KEQs) makes it easier to decide what data to collect, how to analyze it, and how to report it.

KEQs usually need to be developed and agreed on at the beginning of evaluation planning - however sometimes KEQs are already prescribed by an evaluation system or a previously developed evaluation framework.

Try not to have too many Key Evaluation Questions - a maximum of 5-7 main questions will be sufficient. It might also be useful to have some more specific questions under the KEQs.

Key Evaluation Questions should be developed by considering the type of evaluation being done, its intended users, its intended uses (purposes), and the evaluative criteria being used. In particular, it can be helpful to imagine scenarios where the answers to the KEQs being used - to check the KEQs are likely to be relevant and useful and that they cover the range of issues that the evaluation is intended to address. (This process can also help to review the types of data that might be feasible and credible to use to answer the KEQs).

The following information has been taken from the New South Wales Government, Department of Premier and Cabinet Evaluation Toolkit, which BetterEvaluation helped to develop.

Key evaluation questions for the three main types of evaluation

Process evaluation

- How is the program being implemented?
- How appropriate are the processes compared with quality standards?
- Is the program being implemented correctly?
- Are participants being reached as intended?
- How satisfied are program clients? For which clients?
- What has been done in an innovative way?

Outcome evaluation (or impact evaluation)

- How well did the program work?
- Did the program produce or contribute to the intended outcomes in the short, medium and long term?
- For whom, in what ways and in what circumstances? What unintended outcomes (positive and negative) were produced?
- To what extent can changes be attributed to the program?
- What were the particular features of the program and context that made a difference?

- What was the influence of other factors?

Economic evaluation (cost-effectiveness analysis and cost-benefit analysis)

  - What has been the ratio of costs to benefits?
  - What is the most cost-effective option?
  - Has the intervention been cost-effective (compared to alternatives)?
  - Is the program the best use of resources?

Appropriateness, effectiveness and efficiency

Three broad categories of key evaluation questions are often used to assess whether the program is appropriate, effective and efficient .

Organising key evaluation questions under these categories, allows an assessment of the degree to which a particular program in particular circumstances is appropriate, effective and efficient. Suitable questions under these categories will vary with the different types of evaluation (process, outcome or economic).

Appropriateness

  - To what extent does the program address an identified need?
  - How well does the program align with government and agency priorities?
  - Does the program represent a legitimate role for government?

Effectiveness

  - To what extent is the program achieving the intended outcomes, in the short, medium and long term?
  - To what extent is the program producing worthwhile results (outputs, outcomes) and/or meeting each of its objectives?

Efficiency

  - Do the outcomes of the program represent value for money?
  - To what extent is the relationship between inputs and outputs timely, cost-effective and to expected standards?

Example

The Evaluation of the Stronger Families and Communities Strategy used clear Key Evaluation Questions to ensure a coherent evaluation despite the scale and diversity of what was being evaluated – an evaluation over 3 years, covering more than 600 different projects funded through 5 different funding initiatives, and producing 7 issues papers and 11 case study reports (including studies of particular funding initiatives) as well as ongoing progress reports and a final report.

The Key Evaluation Questions were developed through an extensive consultative process to develop the evaluation framework, which was done before advertising the contract to conduct the actual evaluation.

  1. How is the Strategy contributing to family and community strength in the short-term, medium-term, and longer-term?
  2. To what extent has the Strategy produced unintended outcomes (positive and negative)?

3. What were the costs and benefits of the Strategy relative to similar national and international interventions? (Given data limitations, this was revised to ask the question in 'broad, qualitative terms'
4. What were the particular features of the Strategy that made a difference?
5. What is helping or hindering the initiatives to achieve their objectives? What explains why some initiatives work? In particular, does the interaction between different initiatives contribute to achieving better outcomes?
6. How does the Strategy contribute to the achievement of outcomes in conjunction with other initiatives, programs or services in the area?
7. What else is helping or hindering the Strategy to achieve its objectives and outcomes? What works best for whom, why and when?
8. How can the Strategy achieve better outcomes?

CIRCLE (2008) <u>Stronger Families and Communities Strategy 2000-2004: Final Report</u>. Melbourne: RMIT University.

The KEQs were used to structure progress reports and the final report, providing a clear framework for bringing together diverse evidence and an emerging narrative about the findings.

The Managers' Guide

Coming at this from a manager or commissioner's perspective? <u>Step 2: Scope the evaluation</u> in our <u>Managers' Guide</u> has some specific information geared towards making decisions about what the evaluation needs to do, including how to <u>develop agreed key evaluation questions</u>.

Resources

- <u>Practical guide for engaging stakeholders in developing evaluation questions</u>

  This guide from the Robert Wood Johnson Foundation was designed to support evaluators engage their stakeholders in the evaluation process.

- <u>Looking back, moving forward: Sida evaluation manual</u>

  This manual from the Swedish International Development Cooperation Agency (SIDA) is aimed at supporting staff in conducting evaluations of development interventions.

- <u>Evaluation questions</u>

  This site provides a step-by-step guide on how to identify appropriate questions for an evaluation.

- <u>Stakeholders' interest in potential evaluation questions</u>

  This worksheet from Chapter 5 of the National Science Foundation's User-Friendly Handbook for Mixed Method Evaluations provides a template for developing evaluation questions which engage stakeholders interest in the process.

- <u>Prioritize and eliminate questions</u>

  This worksheet from Chapter 5 of the National Science Foundation's User-Friendly Handbook for Mixed Method Evaluations provides a template which allows the organisation and selection of possible evaluation questions.

- CDC: Checklist to help focus your evaluation

  This checklist, created by the Centers for Disease Control and Prevention (CDC), helps you to assess potential evaluation questions in terms of their relevance, feasibility, fit with the values, nature and theory of change of the program, and the level

- Evaluation questions checklist for program evaluation

  Created by Lori Wingate and Daniala Schroeter, the purpose of this checklist is to aid in developing effective and appropriate evaluation questions and in assessing the quality of existing questions.

- Evaluation question examples: Evaluation at country level, regional level, sector or thematic global evaluation

  This document contains example questions, many of which are drawn from country, regional, sector or thematic global evaluations undertaken by the Evaluation Unit.

Determine what 'success' looks like

Evaluation is essentially about values, asking questions such as : What is good, better, best? Have things improved or got worse? How can they be improved?

Therefore, it is important for evaluations to be systematic and transparent in the values that are used to decide criteria and standards.

Criteria

Criteria refer to the aspects of an intervention that are important to consider when deciding whether or not, and in what ways, it has been a success or a failure, or when producing an overall judgement of performance. There are different types of criteria:

Positive outcomes and impacts: for example, should childcare be judged in terms of its success in supporting early childhood development or in supporting parents to engage in education or work? If it is both, how should they be weighted?

Negative outcomes and impacts: for example, an infrastructure development might produce negative unintended effects (e.g. soil erosion caused by a new road) as well as positive intended effects)

Distribution of costs and benefits: for example, is it important for everyone to receive some benefit or the same benefit or for the intervention to be targeted so that the most disadvantaged receive more benefit?

Resources and timing: for example, is there a need for results to be achieved within a certain timeframe?

Processes: for example, use of recyclable materials; providing access to groups with restricted mobility

Standards

Standards refer to the levels of performance required for each of the criteria. For example, if a project aims to reduce maternal mortality, what level of performance is needed for it to be considered successful?

Any reduction? A reduction of at least xx%? A reduction of at least xx in absolute terms? A reduction to a rate of x.x that matches other similar regions, or matches official targets?

Criteria and standards need to be agreed on in order to identify the data that need to be gathered for an evaluation.

In addition, these data need to be combined to form an overall judgement of success or failure, or to rank alternatives against each other. For example, if a road project achieves its economic objectives but produces environmental damage, should it be considered a success overall? How much damage, and at whose cost, would be enough to outweigh the positive impacts? These issues are addressed under the task Synthesise data from a single evaluation.

Methods

Formal statements of values

Some options are used to identify possible criteria and standards that could be used in an evaluation, drawing on formal and informal sources, and some options are used to negotiate which should be used and how they should be weighed.

- Standards, evaluative criteria and benchmarks

  Standards, evaluative criteria, or benchmarks refer to the criteria by which an evaluand will be judged during an evaluation.

- Stated goals and objectives

  Evaluations can use the program's stated objectives and goals to assess program success or failure.

Articulate and document tacit values

- Hierarchical card sorting

  Hierarchical card sorting (HCS) is a participatory card sorting method designed to provide insight into how people categorise and rank different phenomena.

- Open space

  Open Space Technology (OST) is a group facilitation approach for small and large gatherings in which a central purpose, issue, or task is addressed, but which begins with a purposeful lack of any formal initial agenda.

- Photovoice

  Photovoice is a participatory photography method that seeks to empower marginalised people to share their experiences through digital storytelling.

- Rich pictures

A rich picture is a way to explore, acknowledge and define a situation and express it through diagrams to create a preliminary mental model and can help to open discussion and come to a broad, shared understanding of a situation.

- Stories of change

  Stories of change show what is valued through the use of specific narratives of events.

  Structured with a beginning, middle and end, they focus on the change that has taken place due to the program.

- Values clarification interviews

  Values Clarification Interviews involve interviewing key informants and intended beneficiaries to identify what they value.

- Values clarification public opinion questionnaires

  Seeking feedback from large numbers of people about their priorities through the use of questionnaires.

Negotiate between different values

- Concept mapping

  A concept map shows how different ideas relate to each other - sometimes this is called a mind map or a cluster map.

- Delphi study

  The Delphi technique is a quantitative option to generate group consensus through an iterative process of answering questions.

- Dotmocracy

  Dotmocracy is an established facilitation method for collecting and recognizing levels of agreement on written statements among a large number of people.

- Open space

  Open Space Technology (OST) is a group facilitation approach for small and large gatherings in which a central purpose, issue, or task is addressed, but which begins with a purposeful lack of any formal initial agenda.

- Public consultations

  Public consultations are usually conducted through public meetings to provide an opportunity for the community to raise issues of concern and respond to options.

Approaches

- Critical systems heuristics: The idea and practice of boundary critique

  This chapter provides a detailed introduction to critical systems heuristics and the use of its central tool, boundary critique.

- Participatory evaluation

  Participatory evaluation is an approach that involves the stakeholders of a programme or policy in the evaluation process.

Describe activities, outcomes, impacts and context

This cluster of evaluation tasks involves collecting or retrieving data and analyzing it to answer evaluation questions about what has happened - activities, outcomes and impacts - and also important contextual information.

Sample

Sampling is the process of selecting units (e.g., people, organizations, time periods) from a population of interest, studying these in greater detail and then drawing conclusions about the larger population to study them in greater detail.

Methods

Consider why you want to study your population of interest and what you want to do with the information that you have gathered, before you choose your method.

There are three clusters of sampling options: Probability; Purposive (or Purposeful); and Convenience.

Probability

Probability sampling methods use random or quasi-random methods to select the sample, and then use statistical generalization to draw inferences about that population. To minimize bias, these methods have specific rules on selection of the sampling frame, size of the sample, and managing variation within the sample. The methods include:

- Multi-stage: cluster sampling in which larger clusters are further subdivided into smaller, more targeted groupings for the purposes of surveying.
- Sequential: selecting every nth case from a list (e.g. every 10th client)
- Simple random: drawing a sample from the population completely at random.
- Stratified random: splitting the population into strata (sections or segments) in order to ensure distinct categories are adequately represented before selecting a random sample from each.

Purposive (or Purposeful)

Purposive sampling methods study information-rich cases from a given population to make analytical inferences about the population. Units are selected based on one or more predetermined characteristics

and the sample size can be as small as one (n=1). To minimize bias, this cluster of methods encourages transparency in case selection, triangulation, and seeking out of disconfirming evidence. The methods are:

- Confirming and disconfirming: cases that match existing patterns (to explore them) and those that don't match (to test them).
- Criterion: cases that meet a particular condition
- Critical case: a case of particular importance, or that can make a strong point
- Homogenous: cases that are very similar to each other.
- Intensity: selecting cases which exhibit a particular phenomenon intensely.
- Maximum variation: contains cases that are as different from each other as possible.
- Outlier: analysing cases that are unusual or special in some way, such as outstanding successes or notable failures.
- Snowball: asking initial informants to identify additional informants, creating a snowball effect as the sample gets bigger and bigger
- Theory-based: selecting cases according to the extent to which they represent a particular theoretical construct.
- Typical case: developing a profile of what is agreed as average, or normal.

Convenience

Convenience sampling is a cluster of methods that use samples which are readily available and which may not allow credible inference about the population. Convenience methods are:

- Convenience: based on the ease or "convenience" of gaining access to a sample. simply in which data is gathered from people who are readily available.
- Volunteer: sampling by simply asking for volunteers

Resources

Probability

- Probability sample

  This entry from the Encyclopedia of Survey Research Methods provides a detailed overview of probability sampling and the different kinds of designs that can be used for gathering data for this method.

- Sampling for surveys: Simple random sampling

  These instructional videos introduce the topic of sampling for surveys and provide a guide and examples of how to apply simple random sampling.

- Sampling for surveys: Sample size calculations

  This instructional video explains how to calculate a sample size for a survey.

- Sampling for surveys: Stratified random sampling

  These instructional videos provide a guide and examples of how to apply stratified random sampling.

- Sampling for surveys: Clustered random sampling

These instructional videos provide a guide and examples of how to apply clustered random sampling.

Purposive

- Qualitative research & evaluation methods: integrating theory and practice

  The fourth edition of Michael Quinn Patton's Qualitative Research & Evaluation Methods Integrating Theory and Practice, published by Sage Publications, analyses and provides clear guidance and advice for using a range of different qualitative method

- Purposive sampling

  This entry from the Encyclopedia of Survey Research Methods provides a detailed overview of purposive sampling and how it can be used in evaluation. (Academic subscription needed to access).

Use measures, indicators or metrics

As part of an evaluation, it is often important to either develop or use existing indicators or measures of implementation and/or results.

Using an existing indicator or measure can have the advantage of producing robust data which can be compared to other studies, as long as it is appropriate.

Considerable work has been done to develop measures and indicators that can be used for the outcomes of development projects.

The terms "measure", "metric" and indicator" are often used interchangeably and their definitions vary across different documents and organisations. Hence, it is always useful to check what these terms mean in specific contexts.

Terms that are commonly associated with measurements include:

- A **target** is the value of an indicator expected to be achieved at a specified point in time. Often a **benchmark** is used to mean the same thing.
- An **index** is a set of related indicators which intend to provide a means for meaningful and systematic comparisons of performance across programmes that are similar in content and/or have the same goals and objectives.
- A **standard** is a set of related indicators, benchmarks or indices which provide socially meaningful information regarding performance.

Resources

Advocacy

- Outcomes and performance indicators: Advocacy program

This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of advocacy programs.

Education and Training

- Outcomes and performance indicators: Youth tutoring program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of youth tutoring programs.

- Outcomes and performance indicators: Youth mentoring program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of youth mentoring programs.

- Outcomes and performance indicators: Employment training/workforce development program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of Employment Training/Workforce Development Programs.

Governance

- Worldwide Governance Indicators

  Reports aggregate and individual governance indicators for over 200 countries and territories over the period 1996–2020. (World Bank)

- The Ibrahim Index

  Provides a framework and tools that were developed in order to assess the delivery of public goods and services in Africa. (Mo Ibrahim Foundation)

- Data.Gov Open data

  Allows users to interactively access and compare data for governance issues from around the world.

- IADB: Numbers for development

  Explore a snapshot of key development indicators for a country related to its macroeconomic profile, global integration, and social outlook. (IADB)

Health

- Outcomes and performance indicators: Health risk reduction program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of Health Risk Reduction Programs.

Inequality

- Policy impacts on inequality: Simple inequality measures

  This module from the Food and Agriculture Organization of the United Nations (FAO) demonstrates a range of ways to measure inequality by using the statistical concepts of location, shape and variability.

Poverty

- Handbook on poverty and inequality

  This book form the World Bank provides a range of tools which allow the user to measure, describe, monitor, evaluate, and analyze poverty.

- Multidimensional Poverty Index

  Aims to capture the multiple aspects that constitute poverty. (Oxford Poverty & Human Development Initiative)

Welfare

- Outcomes and performance indicators: Transitional housing program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of Transitional Housing Programs.

- Outcomes and performance indicators: Prisoner re-entry program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of Prisoner Re-entry Programs.

- Outcomes and performance indicators: Emergency shelter program

  This set of outcome indicators, developed by the Urban Institute, is aimed at supporting the development, monitoring and evaluation of Emergency Shelter Programs.

Wellbeing

- California county scorecard of children's well-being

  This website from Children Now provides an interactive display of statistics reporting on the wellbeing of children in California.

  It uses 38 indicators to rank child well-being in each of California's 58 Counties.

World peace

- Global peace index

  The Global Peace Index, an initiative of Institute for Economics and Peace (IEP), provides a ranking for each nation in regards to their peacefulness.

Method

- Existing documents

  Reviewing documents produced as part of the implementation of the evaluand can provide useful background information and be beneficial in understanding the alignment between planned and actual implementation.

Collect and/ or retrieve data

This task focuses on ways to collect and/or retrieve data about activities, results, context and other factors.

It is important to consider the type of information you want to gather from your participants and the ways you will analyse that information, before you choose your method. You should also consider triangulating your methods in order to ensure multiple data sources and perspectives.

Methods

The data collection tasks have been organised into five clusters based on the source of the data.

1. Information from individuals
2. Information from groups
3. Observation
4. Physical measurements
5. Reviewing existing records and data

Before choosing methods and collecting data it is essential to consider your key evaluation questions (KEQs) and the type of information you require to address these questions. You also need to consider the context of the evaluation and ensure the methods you choose are suitable and fit for purpose.

1. Information from individuals

- Deliberative opinion polls

  The purpose of Deliberative Opinion Polls (DOPs) is to measure informed opinion on a particular issue.

- Journals and logs

  Journals and logs are forms of record-keeping tools that can be used to capture information about activities, results, conditions, or personal perspectives on how change occurred over a period of time.

- Goal Attainment Scales

  Goal Attainment Scaling (GAS) is a method that can be used as a means of measuring outcome data from different contexts set out on a 5 point scale of -2 to +2.

- Hierarchical card sorting

  Hierarchical card sorting (HCS) is a participatory card sorting method designed to provide insight into how people categorise and rank different phenomena.

- Interviews

  Interviews are conversations between an investigator (interviewer) and a respondent ('interviewees', 'informants' or 'sources') in which questions are asked in order to obtain information.

- Convergent interviewing

  A convergent interview is a type of interview intended to explore issues widely through a combination of unstructured interviews and a maximum diversity sample.

- In-depth interviews

  An in-depth interview is a type of interview with an individual that aims to collect detailed information beyond initial and surface-level answers.

- Key informant interviews

  Key informant interviews involve interviewing people who have particularly informed perspectives on an aspect of the program being evaluated.

- Keypad technology

  Keypads are used in group meetings to gauge audience response to presentations and provide valuable feedback in large group settings.

- Mobile data collection

  Mobile Data Collection (MDC) is the use of mobile phones, tablets or personal digital assistants (PDAs) for programming or data collection.

- Photovoice

  Photovoice is a participatory photography method that seeks to empower marginalised people to share their experiences through digital storytelling.

- Photolanguage

Photolanguage is a projective technique to elicit rich verbal data where participants choose an existing photograph as a metaphor and then discuss it.

- Polling booth

Polling booth is a data collection methodology used to obtain sensitive information from participants.

- Postcards

Postcards can be used to collect information quickly, and they can also be used to provide a short report on evaluation findings (or an update on progress).

- Projective techniques

Projective techniques, originally developed for use in psychology, can be used in an evaluation to provide a prompt for interviews.

- Questionnaires

A questionnaire is a specific set of written questions which aims to extract specific information from the chosen respondents.

- Email questionnaires

Email Questionnaires are surveys or questionnaires that are distributed online via email.

- Face-to-face questionnaires

Face-to-face questionnaires are conducted by an interviewer asking questions of a respondent in person.

- Internet questionnaire

An internet questionnaire allows the collection of data through an electronic set of questions that are posted on the web.

- Mobile questionnaires

Questionnaires and surveys can be conducted through mobile phones which are able to connect to the internet.

- Mail questionnaire

Questionnaires can be mailed out to a sample of the population, enabling the researcher to connect with a wide range of people.

- Telephone questionnaires

Respondents can be surveyed using questionnaires delivered by telephone.

- Seasonal calendars

Seasonal calendars are useful for evaluation as they can help analyse time-related cyclical changes in data.

- Sketch mapping

  Sketch mapping is useful for creating a visual representation ('map') of a geographically based or defined issue drawn from the interpretation of a group or different groups of stakeholders.

- Stories of change

  Stories of change show what is valued through the use of specific narratives of events.

  Structured with a beginning, middle and end, they focus on the change that has taken place due to the program.

- Personal stories

  Personal stories provide qualitative data about how people experience their lives and can be used to make sense of the past and to understand possible futures.

2. Information from groups

- After action review

  The after action review (AAR) is a simple method for facilitating an assessment of organisational performance by bringing together a team to discuss a task, event, activity or project in an open and honest fashion.

- Brainstorming

  Brainstorming involves focussing on a problem and then encouraging participants to come up with as many solutions as possible.

- Card visualization

  Card visualization is a participatory method for capturing data that uses paper cards to allow groups to brainstorm and share their ideas.

- Concept mapping

  A concept map shows how different ideas relate to each other - sometimes this is called a mind map or a cluster map.

- Delphi study

  The Delphi technique is a quantitative option to generate group consensus through an iterative process of answering questions.

- Dotmocracy

  Dotmocracy is an established facilitation method for collecting and recognizing levels of agreement on written statements among a large number of people.

- Fishbowl technique

  The fish bowl activity is used to manage group discussion.

- Future search conference

  A future search conference is a meeting that spans more than one day with the objective that participants identify a shared vision of the future towards which to aim.

- Interviews

  Interviews are conversations between an investigator (interviewer) and a respondent ('interviewees', 'informants' or 'sources') in which questions are asked in order to obtain information.

- Focus groups

  A focus group is a type of group interview designed to explore peoples attitudes.

- Mural

  A mural, a large drawing on the wall, can be used to collect data from a group of people about the current situation, their experiences using a service, or their perspectives on the outcomes from a project.

- ORID

  ORID is a specific facilitation framework that enables a focused conversation with a group of people in order to reach some point of agreement or clarify differences.

- Q-methodology

  Q-methodology (also known as Q-sort) is the systematic study of participant viewpoints.

- Social mapping

  Social mapping, or 'wellbeing ranking', is used to identify households using pre-determined indicators based on socio-economic factors.

- SWOT analysis

  The SWOT analysis is a strategic planning tool that encourages group or individual reflection on and assessment of the Strengths, Weaknesses, Opportunities and Threats of a particular strategy and how to best implement it.

- World cafe

  The world café is a methodology for hosting group dialogue which emphasizes the power of simple conversation in considering relevant questions and themes.

- Writeshop

  A Writeshop is a writing workshop involving a concentrated process of drafting, presenting, reviewing and revising documentation of practice.

3. Observation

Gathering information by observing people, places and/ or processes either directly or through still or moving images (photography or video). This cluster of methods involves watching and documenting the incidence of objects and/ or the behaviour of people.

These methods do not involve gathering data directly from individuals or groups, but rather about observing individuals, groups and things. Evaluators of an education project may observe the physical attributes of a school, the accessibility of the site, the availability of latrines, library, and playground. The evaluator may observe the numbers of boys and girls in a classroom, the teaching techniques used and the types of resources that children use.

- Field trips

  Field trips are organised trips where participants visit physical sites.

- Non-participant observation

  Non-participant Observation involves observing participants without actively participating.

- Participant observation

  Participant observation is used to identify the attitudes and operation of a community by a researcher living within its environs.

- Photography/Video recording for data collection

  This option uses a series of still photographs or videos taken over a period of time to discern changes taking place in the environment or activities of a community.

- Transect

  Transect walks are a method for gathering spatial data on an area by observing people, surroundings and resources while walking around an area or community.

4. Physical measurements

Measuring physical changes based on agreed indicators and measurement procedures. Examples include birth weight, nutrition levels, rain levels, and soil fertility.

- Biophysical measurement

  Biophysical measurement measures physical changes that take place over a period of time related to a specific indicator and using an accepted measurement procedure.

- Geographical

  Capturing geographic information about persons or objects of interest such as the locations of high prevalence of a disease or the location of service delivery points.

5. Existing documents and data

Often information required for an evaluation has already been collected for other purposes. Ministries, government agencies, NGOs, and other organizations often produce valuable reports that you can use to supplement your own data collection. The document review process provides a systematic procedure for identifying, analyzing, and deriving useful information from existing documents such as project documents, information on related projects, government records and publicly available statistics. Document review can assist in triangulating findings collected through other evaluation methods, for example interview and observations. Document review can also reduce duplication.

An evaluator may review existing documents for the following reasons: to gather background information, to determine if implementation of the program reflects the program plan, when you need information to help you develop other data collection tools for evaluation and when you need data to answer what and how many evaluation questions commonly collected by other agencies.

- Big data

  Big data refers to data that are so large and complex that traditional methods of collection and analysis are not possible.

- Journals and logs

  Journals and logs are forms of record-keeping tools that can be used to capture information about activities, results, conditions, or personal perspectives on how change occurred over a period of time.

- Official statistics

  Statistics published by government agencies or other public bodies such as international organizations are often useful in evaluations.

- Previous evaluations and research

  Using the findings from evaluation and research studies conducted on the same or closely related areas is an important first step for evaluation planning.

- Existing documents

  Reviewing documents produced as part of the implementation of the evaluand can provide useful background information and be beneficial in understanding the alignment between planned and actual implementation.

- Reputational monitoring dashboard

  A 'reputation monitoring dashboard' allows users to monitor and quickly appraise reputational trends at a glance and from a variety of different sources.

Manage data

Good data management includes developing effective processes for consistently collecting and recording data, storing data securely, backing up data, cleaning data, and modifying data so it can be transferred between different types of software for analysis.

Good data management is inextricably linked to data quality assurance –the processes and procedures that are used to ensure data quality. Using data of unknown or low quality may result in making the wrong decisions about policies and programmes. Data quality assurance (DQA) should be built into each step in the data cycle ? data collection, aggregation and reporting, analysis and use, and dissemination and feedback.

Even when data have been collected using well-defined procedures and standardised tools, they need to be checked for any inaccurate or missing data. This "data cleaning" involves finding and dealing with any errors that occur during writing, reading, storage, transmission, or processing of computerised data.

Ensuring data quality also extends to presenting the data appropriately in the evaluation report so that the findings are clear and conclusions can be substantiated. Often, this involves making the data accessible so that they can be verified by others and/or used for additional purposes such as for synthesising results across different evaluations.

Commonly referred to aspects of data quality are:

- *Validity*: The degree to which the data measure what they are intended to measure.
- *Reliability*: Data are collected consistently; definitions and methodologies are the same when doing repeated measurements over time.
- *Completeness*: Data are complete (i.e., no missing data or data elements).
- *Precision*: Data have sufficient detail.
- *Integrity*: Data are protected from deliberate bias or manipulation for political or personal reasons
- *Availability*: Data are accessible so they can be validated and used for other purposes.
- *Timeliness*: Data are up-to-date current and available on time.

Methods

- Consistent data collection and recording

  An important aspect of data quality is to ensure data is collected consistently across different sites and different data collectors.

- Data backup

  Data backup refers to onsite and offsite, automatic and manual processes to guard against the risk of data being lost or corrupted.

- Data cleaning

  Data cleaning involves the detection and removal (or correction) of errors and inconsistencies in a data set or database due to data corruption or inaccurate entry.

- Effective data transfer

  Effective data transfer involves processes to move data between systems, including between software packages, to avoid the need to rekey data.

- Secure data storage

  Processes to protect electronic and hard copy data in all forms, including questionnaires, interview tapes and electronic files from being accessed without authority or damaged.

- Archive data for future use

Putting systems in place to store de-identified data so that they can be accessed for verification purposes or for further analysis and research in the future, researchers can extend the range of the data collection efforts and encourage future innovati

Resources

- Data management

  Supports the design of quality data management systems. (Food and Agriculture Organization, Fisheries and Aquaculture Department)

- Data quality tools and mechanisms (archive link)

  Guides to three tools that can be used to assess the quality of data and reporting systems. (The Global Fund)

- Data Quality

  This online course from the Global Health Learning Centre is designed to help learners understand what data quality is, why it is important, and what programs can do to improve it.

Combine qualitative and quantitative data

Using a combination of qualitative and quantitative data can improve an evaluation by ensuring that the limitations of one type of data are balanced by the strengths of another.

This will ensure that understanding is improved by integrating different ways of knowing. Most evaluations will collect both quantitative data (numbers) and qualitative data (text, images), however it is important to plan in advance how these will be combined.

Methods

When data are gathered

- Parallel data gathering

  Qualitative and quantitative data are gathered at the same time.

  For example, a closed-ended questionnaire to many service users is done at the same time as semi-structured observations of the service center.

- Sequential data gathering

  Sequencing is one way of combining qualitative and quantitative data by alternating between them.

When data are combined

- Component design

  Component design is an approach to mixed methods evaluation that conducts qualitative components of the evaluation separately to quantitative components and then combines the data at the time of report writing.

- Integrated design

  Integrated Design is an approach to mixed options evaluation where qualitative and quantitative data are integrated into an overall design.

Purpose of combining data

- Enriching

  'Enriching' is achieved by using qualitative work to identify issues or obtain information on variables not obtained by quantitative surveys.

- Examining

  'Examining' refers to generating hypotheses from qualitative work to be tested through the quantitative approach.

- Explaining

  'Explaining' involves using qualitative work to understand unanticipated results from quantitative data.

  In principle, this mechanism may operate in either direction – from qualitative to quantitative approaches or vice versa.

- Triangulation

  Triangulation facilitates validation of data through cross verification from more than two sources.

Resources

Guides

- Introduction to mixed methods in impact evaluation

  This guide, written by Michael Bamberger for InterAction outlines the elements of a mixed methods approach with particular reference to how it can be used in an impact evaluation.

- Conducting mixed-method evaluations

This technical note from the US Agency for International Development (USAID) provides an overview of using a mixed-methods approach for evaluation.

Analyse data

Analysing data to summarise it and look for patterns is an important part of every evaluation.

The methods for doing this have been grouped into two categories - quantitative data (number) and qualitative data (text, images).

Methods

Numeric analysis

Analysing numeric data such as cost, frequency, and physical characteristics.

- Correlation

  Correlation is ?a statistical measure ranging from +1.0 to -1.0, represented by 'r', that indicates how strongly two or more variables are related and whether that relationship is positive or negative.

- Crosstabulations

  Crosstabulation (or crosstab) is a basic part of survey research in which researchers can get an indication of the frequency of two variables (e.g. gender or income, and frequency of school attendance) occurring at the same time.

- Data mining

  Data mining is the systematic process of discovering patterns in data sets through the use of computer algorithms.

- Exploratory techniques

  Taking a 'first look' at a dataset by summarising its main characteristics, often by using visual methods.

- Frequency tables

  A frequency table provides collected data values arranged in ascending order of magnitude, along with their corresponding frequencies.

- Measures of central tendency

  Measures of Central Tendency provide a summary measure that attempts to describe a whole set of data with a single value that represents the middle or centre of its distribution.

- Measures of dispersion

Measures of dispersion provide information about how much variation there is in the data, including the range, inter-quartile range and the standard deviation.

- Multivariate descriptive

  Multivariate descriptive statistics involves analysing relationships between more than two variables.

- Non-parametric inferential statistics

  Inferential statistics suggest statements or make predictions about a population based on a sample from that population. Non-parametric tests relate to data that are flexible and do not follow a normal distribution.

- Parametric inferential statistics

  Parametric inferential tests are carried out on data that follow certain parameters.

- Summary statistics

  Summary statistics provide a quick summary of data and are particularly useful for comparing one project to another, or before and after.

- Time series analysis

  A time series is a collection of observations of well-defined data items obtained through repeated measurements over time.

Textual analysis

Analysing words, either spoken or written, including questionnaire responses, interviews, and documents.

- Causal mapping

  ?Causal mapping helps make sense of the causal claims (about "what causes what") that people make in interviews, conversations, and documents.

- Content analysis

  Content analysis is a research method in the social sciences used to reduce large amounts of unstructured textual content into manageable data relevant to the (evaluation) research questions.

- Thematic coding

  Thematic coding is a form of qualitative analysis that involves recording or identifying passages of text or images that are linked by a common theme or idea allowing you to index the text into categories and therefore establish a "framework of thematic

- Framework Matrices

  A framework matrix is a way of summarizing and analyzing qualitative data in a table of rows and columns.

- Timelines and time-ordered matrices

  Timelines and time-ordered matrices are useful ways of displaying and analysing time-related data.

Resources

Websites

- WISE: Web Interface for Statistics Education

  WISE's website organises a large amount of statistics resources available on the web into one central place.

Tools

- For an overview of specialist tools for qualitative data analysis, see the CAQDAS site at the University of Surrey which compares ten packages including Atlas.Ti, HyperResearch and NVivo.

Visualise data

Data visualisation is the process of representing data graphically in order to identify trends and patterns that would otherwise be unclear or difficult to discern.

Data visualisation serves two purposes: to bring clarity during analysis and to communicate.

The choice of what type of graph or visualisation to use depends greatly on the nature of the variables you have, such as relational, comparative, time-based, etc. Here we have adopted and modified the categorization system used by ManyEyes (archived link, IBM closed this service in 2015).

That said, sometimes graphing data with an inappropriate visualisation can lead to insights during analysis that would have remained hidden. Experimentation with visualisations during analysis is okay, but when communicating a visualisation, use the graph types listed under the proper methods below. Incorrect visualisation leads to confusion, errors, and abandonment among viewers.

The methods listed here can support both purposes of analysis and communication. You may want to graph data during analysis to see, for example, spikes in website traffic related to your social media campaigns. Visualisation, in this instance, eases data analysis. When communicating that data, however, the visualisation may need to be simplified and key areas may need emphasis in order to call the attention of readers and stakeholders. See the discussion under Report and Support Use for more information about how you may want to repackage a data visualisation for communication purposes.

Each main method below contains several visualisation possibilities. Click on each to see examples and read advice on using and choosing that visualisation method.

This graphic by Andrew Abela from Extreme Presentations provides a good representation of different types charts that can be used to visualise data.

Variable Width Column Chart
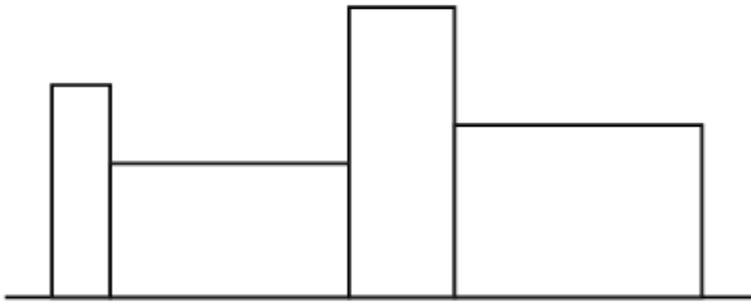
Table or
Embed...

Two Variables
per Item

M...
Cat...
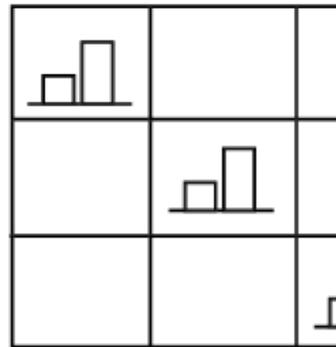
Among Items

Methods

See relationships among data points

- Scatterplot

  A Scatterplot is used to display the relationship between two quantitative variables plotted along two axes. A series of dots represent the position of observations from the data set.

- Matrix chart

  A matrix chart shows relationships between two or more variables in a data set in grid format.

- Network diagram

  A network diagram uses a set of nodes and connecting lines to display of how people (or other elements) in a network are connected.

  It is usually a product of social network analysis.

Compare a set of values

- Bar chart

  A bar chart plots the number of times a particular value or category occurs in a data set, with the length of the bar representing the number of observations with that score or in that category.

- Block histogram

  A histogram is a graphical way of presenting a frequency distribution of quantitative data organised into a number equally spaced intervals or bins (e.g. 1-10, 11-20…).

- Bubble chart

  Commonly used on maps, and x/y-axis plots, or no plot at all, bubble charts communicate the raw count, frequency, or proportion of some variable where the size of the bubble reflects the quantity.

- Bullet graph

  Bullet graphs encode a single variable as a bar.

- Deviation bar graph

  Deviation bar graphs are simply two bar charts aligned, where one of the charts runs right to left rather than left to right.

- Dot plot

Dot plots encode single data points with circles, often on a line.

- Small multiples

Small multiples are an array of graphs on the same scale that are grouped together in a row or grid and are often used to simplify a data display.

Changes over time

- Line graph

A line graph is commonly used to display change over time as a series of data points connected by straight line segments on two axes.

- Slopegraph

A slopegraph is a lot like a line graph, in that it plots change between points however, a slopegraph plots the change between only two points, without any kind of regard for the points in between.

- Split axis bar graph

While many graph types geared toward comparisons ask the viewer to subtract the difference between the heights of two bars or the space between two points on a line, a deviation bar graph simply graphs the difference.

- Stacked graph

Stacked graphs depict items stacked one on top (column) of the other or side-by-side (bar), differentiated by coloured bars or strips.

See the parts of a whole

- Icon array

An icon array is a display in which one shape is repeated a specific number of times (usually 10, 100 or 1,000) and then some of the shapes are altered in some way (usually by colour) to represent a proportion.

- Pie chart

A pie chart is a divided circle, in which each slice of the pie represents a part of the whole.

The categories that each slice represents are mutually exclusive and exhaustive. Data with negative values cannot be displayed as a pie chart.

- Treemap

A treemap displays hierarchical relationships through a set of rectangles, sized proportionately to each data point, clustered together into one large rectangle.

Analyse a text

- Phrase net

  Phrasenets are useful for exploring how words are linked in a text and, like word clouds and word trees, can be informative for early data analysis.

- Word cloud

  Word clouds or tag clouds are graphical representations of word frequency that give greater prominence to words that appear more frequently in a source text.

- Word tree

  Word trees use a visual branching structure to show how a pre-selected word(s) is connected to other words.

See the world

- Demographic mapping

  Demographic mapping is a way of using GIS (global information system) mapping technology to show data on population characteristics by region or geographic area.

- Geo-tagging

  Geo-tagging is the process of adding geographic information about digital content, within "metadata" tags - including latitude and longitude coordinates, place names and/or other positional data.

- GIS mapping

  GIS mapping will typically display one data variable or indicator, often using colour coding to indicate the density, frequency, or percentage in a given region, allowing quick comparison between regions.

- Interactive mapping

  Interactive mapping involves using maps that allow zooming in and out, panning around, identifying specific features, querying underlying data such as by topic or a specific indicator (e.g., socioeconomic status), generating reports and other means of u

Understand causes of outcomes and impacts

Most evaluations require ways of addressing questions about cause and effect – not only documenting what has changed but understanding why.

Impact evaluation, which focuses on understanding the long-term results from interventions (projects, programs, policies, networks and organisations), always includes attention to understanding causes.

Understanding causes can also be important in other types of evaluations. For example in a process evaluation, there often needs to be some explanation of why implementation is good or bad in order to be able to suggest ways it might be improved or sustained.

In recent years there has been considerable development of methods for understanding causes in evaluations, and also considerable discussion and disagreement about which options are suitable in which situations.

When choosing between these different options, consider the different types of causal inference that might be involved:

- One cause producing one effect – it is necessary and sufficient to produce the effect

- Two or more causes combining to produce an effect (for example, two programs or a program when combined with other factors such as particular participant characteristics) – one of the causes alone is necessary but not sufficient

- Two or more causes being alternative ways of producing an effect – either of them are sufficient and neither is necessary

Different labels might be used for these different types of causal relationship - 'causal attribution' implying a single cause, 'causal contribution' implying a package of causal factors, and 'causal inference' being used to refer to all of these.

It is also important to consider the different types of questions that might be asked about cause and effect:

- Did the intervention make a difference?

- For whom, in what situations, and in what ways did the intervention make a difference?

- How much of a difference did the intervention make?

- To what extent can a specific impact be attributed to the intervention?

- How did the intervention make a difference?

You can explore the three broad strategies for causal inference shown below.

Check the results are consistent with causal contribution

One of the tasks involved in understanding causes is to check whether the observed results are consistent with a cause-effect relationship between the intervention and the observed impacts.

Some of the methods for this task involve an analysis of existing data and some involve additional data collection. It is often appropriate to use several methods in a single evaluation. Most impact evaluations should include some methods that address this task.

Methods

Gathering additional data

- Key informant attribution

  A method for testing causal reasoning by asking key informants.

- Modus operandi

  Interventions create distinctive/characteristic patterns of effects.

  Scriven describes the modus operandi as a set of footprints:

- Process tracing

  Process tracing is a case-based and theory-driven method for causal inference that applies specific types of tests to assess the strength of evidence for concluding that an intervention has contributed to changes that have been observed or measured.

Analysis

- Check dose-response patterns

  Evaluators can examine the link between dose and response as part of determining whether the program caused the outcome.

- Check intermediate outcomes

  Intermediate outcomes are identified in a logical model before the final impact.

- Check results match a statistical model

  Program staff may develop a statistical model as part of the project theory design.

  Statistical models can be useful tools to predict elements of the program:
  Cost Time Comparison between groups

- Check results match expert predictions

  Expert predictions can be a useful part of developing the program theory.

  Program staff can draw expert predictions from the literature or by engaging a group of experts.

- Check timing of outcomes

  The program theory may predict the timing of outcomes for the evaluator to check against these dates with the dates of actual changes and outcomes.

  This is another way of checking the results support causal attribution.

- Comparative case studies

  Comparative case studies can be useful to check variation in program implementation.

- Realist analysis of testable hypotheses

  Realist analysis of testable hypotheses tests the program theory by developing a nuanced understanding of 'what works for whom in what circumstances and in what respects, and how?'.

- Multiple lines and levels of evidence

  Multiple lines and levels of evidence (MLLE) is a systematic approach to causal inference that involves bringing together different types of evidence (lines of evidence) and considering the strength of the evidence in terms of different indicators of a

Approaches

- These approaches combine some of the above options together with ruling out possible alternative explanations.

- Contribution analysis

  Contribution analysis is an evaluation approach that provides a systematic way of understanding an intervention's contribution to observed outcomes or impacts.

- Collaborative outcomes reporting

  Collaborative outcomes reporting (COR) is a participatory approach to impact evaluation based around a performance story that presents evidence of how a program has contributed to outcomes and impacts, that is then reviewed by both technical experts and

- RAPID outcomes assessment

  RAPID outcome assessment (ROA) is a method to assess and map the contribution of a project's actions on a particular change in policy or the policy environment.

Compare results to the counterfactual

One of the three tasks involved in understanding causes is to compare the observed results to those you would expect if the intervention had not been implemented - this is known as the 'counterfactual'.

Many discussions of impact evaluation argue that it is essential to include a counterfactual. Some people however argue that in turbulent, complex situations, it can be impossible to develop an accurate estimate of what would have happened in the absence of an intervention, since this absence would have affected the situation in ways that cannot be predicted. In situations of rapid and unpredictable change, when it might not be possible to construct a credible counterfactual it might be possible to build a strong, empirical case that an intervention produced certain impacts, but not to be sure about what would have happened if the intervention had not been implemented.

For example, it might be possible to show that the development of community infrastructure for raising fish for consumption and sale was directly due to a local project, without being able to confidently state that this would not have happened in the absence of the project (perhaps through an alternative project being implemented by another organization).

For a discussion about counterfactual approaches to causal inference, see The Stanford Encyclopedia of Philosophy entry.

Methods

There are three clusters of methods for this task:

Experimental methods (or research designs)

Develop a counterfactual using a control group. Randomly assign participants to either receive the intervention or to be in a control group.

- Control group

  A control group is an untreated research sample against which all other groups or samples in the research is compared.

Quasi-experimental methods (or research designs)

Develop a counterfactual using a comparison group which has not been created by randomization.

- Difference-in-difference

  Difference-in-difference involves comparing the before-and-after difference for the group receiving the intervention (where they have not been randomly assigned) to the before-after difference for those who did not.

- Instrumental variables

  This method is used to estimate the causal effect of variables on an intervention.

- Judgemental matching

Judgemental matching involves creating a comparison group by finding a match for each person or site in the treatment group based on researcher judgements about what variables are important.

- Matched Comparisons

  When using Matched Comparisons, participants (individuals, organizations or communities) are each matched with a non-participant on variables that are thought to be relevant which can be difficult to adequately match on all relevant criteria.

- Propensity scores

  Propensity score matching (PSM) is a quasi-experimental method used to estimate the difference in outcomes between beneficiaries and non-beneficiaries that is attributable to a particular program.

- Regression discontinuity

  Regression Discontinuity Design (RDD) is a quasi-experimental evaluation option that measures the impact of an intervention, or treatment, by applying a treatment assignment mechanism based on a continuous eligibility index which is a variable with a co

- Sequential allocation

  Sequential allocation involves creating a treatment group and a comparison group by using a sequence to choose participants (e.g. every 3rd person on the list).

- Statistically created counterfactual

  A statistical model, such as regression analysis, is used to develop an estimate of what would have happened in the absence of an intervention.

Non-experimental methods

Develop a hypothetical prediction of what would have happened in the absence of the intervention.

- Key informant

  Asking experts of programmes or in the community to predict what would have happened in the absence of the intervention.

- Logically constructed counterfactual

  In some cases it is not possible to construct a counterfactual by creating a control group or a comparison group, but by constructing one logically.

Approaches

- Randomised controlled trial

Randomised controlled trials (RCTs), or randomised impact evaluations, are a type of impact evaluation that uses randomised access to social programmes as a means of limiting bias and generating an internally valid impact estimate.

Investigate possible alternative explanations

All impact evaluations should include some attention to identifying and (if possible) ruling out alternative explanations for the impacts that have been observed.

Methods

- Force field analysis

  A force field analysis is used to support the decision making process by providing a detailed overview of the variety of forces that may be acting on an organisational change issue.

- General Elimination Methodology

  General Elimination Methodology has two stages:

- Key informant

  Asking experts of programmes or in the community to predict what would have happened in the absence of the intervention.

- Multiple lines and levels of evidence

  Multiple lines and levels of evidence (MLLE) is a systematic approach to causal inference that involves bringing together different types of evidence (lines of evidence) and considering the strength of the evidence in terms of different indicators of a

- Process tracing

  Process tracing is a case-based and theory-driven method for causal inference that applies specific types of tests to assess the strength of evidence for concluding that an intervention has contributed to changes that have been observed or measured.

- RAPID outcomes assessment

  RAPID outcome assessment (ROA) is a method to assess and map the contribution of a project's actions on a particular change in policy or the policy environment.

- Ruling out technical explanations

  Ruling out technical explanations involves identifying and investigating possible ways that the results might reflect technical limitations rather than actual causal relationships.

- Searching for disconfirming evidence/following up exceptions

Treating data that doesn't fit the expected pattern not as outliers but as potential clues to other causal factors and then seeking to explain them.

- Statistically controlling for extraneous variables

Statistically controlling for extraneous variables is an option for removing the influence of a variable on the study of program results.

Approaches

- These approaches combine ruling out possible alternative explanations with options to check the results support causal attribution.

- Contribution analysis

Contribution analysis is an evaluation approach that provides a systematic way of understanding an intervention's contribution to observed outcomes or impacts.

- Collaborative outcomes reporting

Collaborative outcomes reporting (COR) is a participatory approach to impact evaluation based around a performance story that presents evidence of how a program has contributed to outcomes and impacts, that is then reviewed by both technical experts and

Synthesise data from one or more evaluations

Bringing together data into an overall conclusion and judgement is important for individual evaluations and also when summarising evidence from multiple evaluations.

Synthesise data from a single evaluation

To develop evaluative judgments, the evaluator draws data from the evaluation and systematically synthesises and values the data.

There are a range of methods that can be used for synthesis and valuing.

Methods

Processes

- Consensus conference

A consensus conference is a formal public meeting, which gives the general public the chance to contribute to and be involved in the assessment of an issue or proposal.

- Expert panel

Expert panels are used when specialized input and opinion is required for an evaluation.

Techniques

- Cost-benefit analysis

This method compares the total costs of a programme/project with its benefits, using a common metric (most commonly monetary units), which enables you to calculate the net cost or benefit associated with the programme.

- Cost effectiveness analysis

Cost-effectiveness analysis (CEA) compares the relative costs of the outcomes of two or more courses of action and is considered an alternative to cost-benefit analysis (CBA).

- Cost utility analysis

Cost utility analysis (CUA) develops an overall measure of utility or value based on the preferences of individuals.

CUA is useful for evaluating, and comparing, programs that aim to reach the same goal in non-monetary terms.

- Lessons learnt

Lessons learnt can take the form of describing what should or should not be done, or describing the outcome of different processes.

- Multi-criteria analysis

A multi-criteria analysis (MCA) is a form of appraisal that measures variables such as material costs, time savings and project sustainability as well as the social and environmental impacts in addition to monetary impacts.

- Numeric weighting

Numeric weighting involves developing numeric scales in order to rate performance against each evaluation criterion and then adding them up for a total score.

- Qualitative weight and sum

Using qualitative ratings (such as symbols) to identify performance in terms of essential, important and unimportant criteria.

"In QWS:

- Rubrics

A rubric is a framework that sets out criteria and standards for different levels of performance and describes what performance would look like at each level.

- Value for money

  Value for money is a term used in different ways, including as a synonym for cost-effectiveness, and as systematic approach to considering these issues throughout planning and implementation, not only in evaluation.

Approaches

- Social return on investment

  Social Return on Investment (SROI) is a systematic way of incorporating social, environmental, economic and other values into decision-making processes.

Synthesise data across evaluations

These methods answer questions about a type of intervention rather than about a single case – questions such as "Do these types of interventions work?" or "For whom, in what ways and under what circumstances do they work?"

The task involves locating the evidence (often involving bibliographic searches of databases, with particular emphasis on finding unpublished studies), assessing its quality and relevance in order to decide whether or not to include it, extracting the relevant information, and synthesizing it. Different options use different strategies and have different definitions of what constitutes credible evidence.

Methods

- Best evidence synthesis

  Best evidence synthesis is a synthesis that, like a realist synthesis, draws on a wide range of evidence (including single case studies) and explores the impact of context.

- Lessons learnt

  Lessons learnt can take the form of describing what should or should not be done, or describing the outcome of different processes.

- Meta-analysis

  Meta-analysis is a statistical method for combining numeric evidence from experimental (and sometimes quasi-experimental studies) to produce a weighted average effect size.

- Meta-ethnography

Meta-ethnography is a method for combining data from qualitative evaluation and research, especially ethnographic data, by translating concepts and metaphors across studies.

- Rapid evidence assessment

  Rapid Evidence Assessment is a process that uses a combination of key informant interviews and targeted literature searches to produce a report in a few days or a few weeks.

- Realist synthesis

  A realist synthesis is the synthesis of a wide range of evidence that seeks to identify underlying causal mechanisms and explore how they work under what conditions, answering the question "what works for whom under what circumstances?" rather than "wha

- Systematic review

  A systematic review is an approach to synthesising evidence from multiple studies. Systematic reviews use methodical approaches and criteria to identify relevant studies for inclusion, assess their quality, extract data and synthesise evidence.

- Textual narrative synthesis

  Dividing the studies into relatively homogenous groups, reporting study characteristics within each group, and articulating broader similarities and differences among the groups

- Vote counting

  Vote counting is a simple but limited method for synthesizing evidence from multiple evaluations and involves comparing the number of positive studies (studies showing benefit) with the number of negative studies (studies showing harm).

Resources

Websites

- Campbell Collaboration
- Evidence for Policy and Practice Information Centre (EPPI-Centre)

Extrapolate findings

An evaluation usually involves some level of generalising of the findings to other times, places or groups of people.

For many evaluations, this simply involves generalising from data about the current situation or the recent past to the future.

For example, an evaluation might report that a practice or program has been working well (finding), therefore it is likely to work well in the future (generalisation), and therefore we should continue to do it

(recommendation). In this case, it is important to understand whether or not future times are likely to be similar to the time period of the evaluation. If the program had been successful because of support from another organisation, and this support was not going to continue, then it would not be correct to assume that the program would continue to succeed in the future.

For some evaluations, there are other types of generalising needed. Impact evaluations which aim to learn from the evaluation of a pilot to make recommendations about scaling up must be clear about the situations and people to whom results can be generalised.

There are often two levels of generalisation. For example, an evaluation of a new nutrition program in Ghana collected data from a random sample of villages. This allowed statistical generalisation to the larger population of villages in Ghana. In addition, because there was international interest in the nutrition program, many organisations, including governments in other countries, were interested to learn from the evaluation for possible implementation elsewhere.

Methods

- Analytical generalisation

  Analytical generalisation involves making projections about the likely transferability of findings from an evaluation, based on a theoretical analysis of the factors producing outcomes and the effect of context.

- Statistical generalisation

  Statistical generalisation involves statistically calculating the likely parameters of a population using data from a random sample of that population.

Approaches

- Horizontal evaluation

  Horizontal evaluation is an approach that combines self-assessment by local participants and external review by peers.

- Positive deviance

  Positive deviance (PD), a behavioural and social change approach, involves learning from those who find unique and successful solutions to problems despite facing the same challenges, constraints and resource deprivation as others.

- Realist evaluation

  Realist evaluation aims to identify the underlying generative causal mechanisms that explain how outcomes were caused and how context influences these.

Resources

Blog post

- Will that successful intervention over there get results over here?

  This blog post and its associated replies, written by Jed Friedman for the World Bank, describes a process of using analytic methods to overcome some of the assumptions that must be made when extrapolating results from evaluations to other settings.

Report & support use of findings

From the first step of the evaluation process, even though it may be one of the last evaluation tasks, explicitly discuss the content, sharing, and use of reports during the initial planning of the evaluation and return to the discussion thereafter. Most importantly, identify who your primary intended users are. Use of the evaluation often depends on how well the report meets the needs and learning gaps of the primary intended users.

Besides the primary intended users (identified as part of framing the evaluation), your findings can be communicated to others for different reasons. For example, lessons learned from the evaluation can be helpful to other evaluators or project staff working in the same field; or it may be worthwhile remolding some of the findings into articles or stories to attract wider attention to an organisations' work, or to spread news about a particular situation.

You will share the findings of the evaluation with the primary intended users and also other evaluation stakeholders.

Don't limit yourself to thinking of sharing evaluation findings through a report. Although a final evaluation report is important it is not the only way to distribute findings. Depending on your audience and budget, it may be important to consider different ways of delivering evaluation findings:

- Presenting findings at staff forums and subject matter conferences
- Developing a short video version of findings
- Sharing findings on the organisation intra-net
- Sharing stories, pictures and drawings from the evaluation (depending on what options you have used to gather data)
- Creating large posters or infographics of findings for display
- Producing a series of short memos

Identify reporting requirements

Before you begin to gather and analyze your data, consider how you can ensure your collection efforts will meet the reporting needs of your primary intended users.

From the very beginning, reporting is an integral part of evaluation which allows you to:

- communicate what you do;
- monitor and track progress;
- demonstrate impact;
- document lessons learned;
- and be accountable and transparent to donors, partners and benefiting communities.

"Evaluation reports may be the only lasting record of a programme or project, including the results achieved and the lessons that were learned from its implementation" (Oxfam Evaluation Guidelines p.11).

Different groups of primary intended users will have varying needs for the evaluation report. When your evaluation plan was developed at the beginning of the process, you should have determined the different groups of primary intended users and begun to ask questions about how the report could be most useful. This information should then be reviewed periodically. Once the reporting deadline nears ensure there is clarity on each of the stakeholder groups' reporting requirements (what needs to be reported and when).

Some questions that may arise include:

- What do you need to include in different kinds of reports?
- At what point do you need to get feedback on your findings - and from whom?
- Will your findings be presented in draft form?
- Are you willing to share draft findings?
- Will you have any influence over the way the findings are re-presented?

Reporting timelines often present a major constraint on the evaluation plan. In particular, the need to report findings in time to inform funding decisions for the next phase of a program often means that reports are needed before impacts can be observed. In these situations, it will be necessary to report on interim outcomes, and to present any research evidence that shows how these are important predictors or pre-requisites to the final impacts. (See the tasks Develop Program Theory/Logic Model and Collect and/or Retrieve Data for more information on this).

Work with the intended users to determine key points in their own reporting and project cycle. For example, the evaluation may be a necessary part of their legislative requirement for an annual review. If that is the case, you need to know their time and internal pressures. Alternatively, they may be presenting at a major conference and want an update from the evaluation team.

With the primary intended users, their learning needs, and their timelines in mind, develop a communication plan to guide the evaluation reporting process. A communication plan can be as simple as a table that organizes this information. Use the communication plan to align data collection activities with reporting needs and to prioritize the time spent on reporting. (Consider the full range of reporting mediums before finalizing the plan. Not everyone will want a full technical report. For ideas on how to make your report more creative, go to the Develop Reporting Media task page.)

Methods

- Communication plan

  A communication plan outlines the strategies that will be used to communicate the results of your evaluation.

- Reporting needs analysis

  Conducting a needs analysis with your client to determine their reporting requirements.

Resources

Guides

- Designing and conducting health systems research projects Volume 2: Data analyses and report writing

  This guide provides 13 modules designed to demonstrate aspects of data analysis and report writing.

- Evaluation strategies for communicating and reporting: guide

  This book from Torres, Preskill and Piontek has been designed to support evaluators to incorporate creative techniques in the design, conduct, communication and reporting of evaluation findings.

Develop reporting media

You may develop a number of reports, in different formats, for different sets of stakeholders.

Work with your primary users and stakeholders to determine when and in what form they want to receive evaluation reports. Also determine who you will involve in viewing draft and interim reports.

Points to consider in choosing the format are:

- How does the audience prefer to receive information – text, graphics, numbers, written, visual or a mixture of all of these?
- What is the preferred length (or duration if an audio/visual presentation)?
- What access does the audience have to information technology (this may inform whether you use web-based formats)?
- What is the purpose of the report and how does this inform the choice of format? Purposes may include:
  - keeping stakeholders engaged during an evaluation
  - providing feedback to and maintaining the commitment of people collecting data during implementation
  - flagging emerging findings and implications for ongoing program development and for the evaluation
  - presenting interim recommendations
  - seeking feedback on draft reports to assist in identifying causal factors
  - informing planning, funding or policy decisions
  - broader dissemination of findings to support use

Methods

Traditionally, written reports have been the main form of media used for evaluation reports. However, we now know that the full technical report is not enough to meet the learning needs of our audiences. The presentation of your report should help your reader quickly and easily understand your key points.

Written

Increasing report readability makes it more likely that readers will be able to learn from the report.

Reporting in the order of importance allows readers to easily access those things which they are most interested in. These are generally the findings and recommendations which, therefore, should appear early in the report. Less relevant details, such as the evaluation background and methodology, belong in an appendix or can even posted online for reference.

- Aide memoire

  An aide-memoire generally refers to a document that is produced to summarise key findings and important recommendations of an evaluation.

- Executive summaries

  The executive summary of an evaluation report is a shortened version of the full report – usually one to four pages – that highlights findings and recommendations and is placed at the front of the report.

- Final reports

  Evaluation reports can be read by many different audiences, ranging from individuals in government departments, donor and partner staff, development professionals working with similar projects or programmes, students and community groups.

- Interim reports

  Interim (or progress) reports present the interim, preliminary, or initial evaluation findings.

- Memos and email

  Memos and emails can be used to help maintain ongoing communication among evaluation stakeholders through brief and specific messages about a particular issue.

- News media communications

  News media communications can include communication channels such as newspapers, magazines, and radio, as well as digital formats such as online news sites, podcasts, social media, and blogs.

- Newsletters, bulletins, blogs, briefs and brochures

  Short communication formats—such as bulletins, briefs, newsletters, blogs and brochures—can be used to highlight particular findings or angles on the evaluation.

- Postcards

  Postcards can be used to collect information quickly, and they can also be used to provide a short report on evaluation findings (or an update on progress).

- Website communications

  These days, having a website is common practice for development organizations working beyond the community level.

This has opened the possibilities of disseminating information such as that coming from evaluations.

Presentation events

Presentation audiences are likely to be most interested in only a portion of the full evaluation report, such as the key findings or a lesson learned about evaluation methods. Thus, it is wise to focus the presentation on only that portion, while making the fuller report available to anyone interested.

- Conferences

  Attendance at professional conferences to understand how other evaluators frame and discuss their findings is a key component of building evaluation capacity.

- Validation workshop

  A validation workshop is a meeting that brings together evaluators and key stakeholders to review an evaluation's findings.

- Teleconference

  Teleconferences can be used to facilitate the discussion of evaluation findings via telephone.

- Verbal Briefings

  Verbal briefings are a way of providing specific information to an audience of interested participants allowing for a structured question and answer format based on that information.

- Webconference

  Webconferencing is a conference hosted on the internet that can allow people who live in different parts of the world to get together.

Presentation materials

- Displays and exhibits

  Through the use of pictures, video or audio representations, maps or models, displays and exhibits can be used to draw attention to certain issues and assist in community engagement.

- Flip charts

  Flip charts are large sheets of paper, usually positioned on a tripod, to be used with thick and differently coloured marking pens.

- Posters

  A good poster communicates your message clearly, quickly and succinctly.

- Powerpoint

Structuring presentations with a series of powerpoint slides is now the most common way of presenting information to groups.

- Video

When produced well, videos provide an excellent means to convey messages coming out of an evaluation.

Creative and/or interactive

Presenting your report in a creative or interactive manner may be the most relevant means to get your information across if the context allows for it. You may consider working with an artist, a graphic recorder or designer to produce creative or interactive displays.

- Cartoons

Cartoon images can be used by evaluators to an understanding of program impact, scenes of program implementation, main findings or issues.

- Data dashboard

Stephen Few defines a dashboard as: "A data dashboard is a visual display of the most important information needed to achieve one or more objectives, with the data consolidated and arranged on a single screen so the information can be monitored at a gla

- Infographics

An infographic (short for 'information graphic') represents data visually so that the information is able to be quickly and easily understood.

- Photographic reporting

Adding photographs to an evaluation report can make it more appealing to readers and also make the key messages more memorable.

- Poetry

When preparing an evaluation report, one way of communicating vividly the experience of participants, or the situation in which the program has been implemented, is to present some of the findings in the form of a poem.

- Reporting in pictures

"A picture is worth a thousand words." Pictures or images provide another way of presenting information, and increasing understanding of your results.

- Theatre

There are several different ways of using theatre to communicate evaluation findings and engage intended users in responding to them.

Graphic design

Simple graphic design principles applied to your reporting documents will ensure readability and maximize learning. You can use design elements and visual depictions of your data to assist the reader.

- Arrangement

  Arranging text and graphics on a page or slide can be a challenge for those not familiar with graphic design. Some basic principles can be easily implemented and boost readability and engagement.

- Colour

  Blocks of background colour can help group similar items or separate reporting elements like sidebars.

  Text intended for narrative reading should be set in black or dark grey on a white or very light background.

- Images

  Written reports and presentations should always include images. Beyond just charts and graphs, photographs or drawings increase the relevancy of the material to the audience and make the report more engaging.

- Text

  Generally speaking, serif fonts support readability in long, narrative-style documents produced on paper.

  Sans serif fonts are easier to read in electronic reporting media.

Visualise data

Also refer to the task visualising data to find methods.

Resources

Guides

- A short primer on innovative evaluation reporting

  This book by Kylie Hutchinson presents a number of innovative ways of reporting, including different methods for presentations, narrative summaries, presenting findings visually and making use of digital outputs.

- Visual language for designers: Guide

"Within every picture is a hidden language that conveys a message, whether it is intended or not. This language is based on the ways people perceive and process visual information.

- Evaluation report layout checklist

This checklist from Stephanie Evergreen distills the best practices in graphic design and has been particularly created for use on evaluation reports.

Ensure accessibility

Plan the reporting products to make sure they are accessible, including addressing issues such as limited time, low literacy, and disabilities.

Methods

General accessibility

- One-Three-Twenty Five (1:3:25) Principle

The 1:3:25 Principle is an evaluation report format with a one page outline of the main messages, a three page executive summary, and 25 pages that present the evaluation findings and methodology.

- Plain language

Plain English is a clear and concise writing style that ensures accessibility to the information for all stakeholders.

- Chartjunk elimination

Often the default settings in graphing programs include too much extraneous graphic detail that can confuse readers and cause them to stop engaging with the report.

- Descriptive chart titles

Descriptive subtitles in a chart can highlight the key takeaway points for the reader.

This is particularly important when graphs must stand alone, without the assistance of the evaluation to help interpret them.

- Emphasis techniques

A key to creating effective and accessible reporting documents is using effective techniques to emphasise important information.

- Headings as summary statements

Headings are extremely important to the readability of reports and documents.

Specific accessibility barriers

- Colour blindness

  People who are affected by colour blindness are unable to distinguish between different hues of certain colours.

- Visual accessibility

  There are a number of ways that documents can be made more accessible to people who are blind or have low vision.

Develop recommendations

Evaluations often make recommendations about how a program can be improved, how the risk of program failure can be reduced or whether a program should continue.

However, not all evaluations include recommendations. It is important to clarify whether recommendations are expected when developing the evaluation brief, terms of reference or scope of work.

If recommendations are developed on the basis of the evaluation findings, processes which involve stakeholders in developing and/or reviewing them will contribute to the use of the evaluation findings. The individual or group who has control of the evaluation – a manager or evaluation steering committee – should be consulted when developing recommendations as their support will probably be very important in order to ensure that the evaluation findings are disseminated and used.

Methods

- Chat rooms

  This method involves setting up an online space where evaluation findings can be discussed.

- Electronic democracy

  Electronic democracy uses new and emergent forms of media to engage community members in seeking to influence the decision-making process by allowing them to apply pressure to those in power over a diverse range of issues.

- External review

  This method involves facilitating a review of the evaluation by an external expert or anonymous reviewer.

- Group critical reflection

  This method involves facilitating group stakeholder feedback sessions on evaluation findings.

- Individual critical reflection

This method involves facilitating independent feedback from particular individual stakeholders.

- Participatory recommendation screening

  This option allows users to test recommendations with key stakeholders.

- World cafe

  The world café is a methodology for hosting group dialogue which emphasizes the power of simple conversation in considering relevant questions and themes.

Support use

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

However, this is often a management responsibility rather than an evaluators. You can work with managers to provide a list of options for follow-up as part of the final report. Indeed, time should be built into the evaluation budget to account for support beyond report delivery.

There are a range of methods that can be used:

Methods

- Annual review

  Annual reviews of major evaluation findings and conclusions, based on evaluation studies completed during the preceding year, can be a useful way to support use.

- Conference co-presentations

  Conference co-presentations take place when evaluators and evaluation commissioners or users jointly present findings or discussions about processes from an evaluation.

- Data use calendar

  A data use calendar is produced to guide the collection of data and reporting requirements, as well as ensuring that analysis and evaluation data is actively used.

- Policy briefing

  Policy briefs are designed to outline findings and recommendations in an accessible manner for specific target audiences.

- Recommendations tracking

  Tracking recommendations involves keeping a transparent record of the responses to and action from recommendations.

- Social learning

Social learning is an approach to learning that focuses on how people learn through social interactions, such as modelling, making connections, sharing experiences and resources, collaboration and self-organization.

Resources

Guides

- Evaluation policy of UNDP

  This evaluation policy from the UNDP has been developed to ensure there is a common basis for evaluations taking place within the organisation.

- 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.

Blogs

- 52 weeks of BetterEvaluation: Week 23: Tips for delivering negative results